Analyzing Genetic Variations and Susceptibility Mechanisms in Lung Cancer

Sharon Li*

Department of Chinese Medicine, Tongji University School of Medicine, Shanghai, China

Corresponding author: Sharon Li, Department of Chinese Medicine, Tongji University School of Medicine, Shanghai, China, E-mail: lei.song@gmail.com

Description

The majority of these genetic variations do not code for proteins and have unknown functions, it is difficult to decipher the molecular mechanisms behind these relationships. Here, we used Massively Parallel Reporter Assays (MPRAs) to evaluate risk-associated variations' allelic transcriptional activity concurrently. At 88% of Genome Wide Association Studies (GWAS) sites, our MPRA method found one or more variations (median 11 variants) that significantly affected transcriptional activity. Potential regulators of the functional variations were identified *via* transcription factor analysis, including those that are anticipated to bind numerous possibly functional variants across the GWAS sites and those that exhibit cell-type-specific expression. Candidate susceptibility genes, including those influencing the proliferation of lung cancer cells, were found by connecting functional variations to target genes using four complimentary techniques. Interference with the top functioning version of CRISPR. verified relationships between different genes, such as *RTEL1*, *SOX18*, and *ARFRP1*. Our results assist clarify the molecular basis of heterogeneity and polygenicity underpinning lung cancer susceptibility and offer a thorough functional investigation of lung cancer GWAS sites.

Genetic basis of lung cancer

One of the most prevalent malignancies that affects a wide range of people and is the leading cause of cancer-related deaths globally is lung cancer. Even though tobacco use is the main risk factor for lung cancer, only a small percentage of smokers get the disease, and up to 25% of lung cancers occur in people who have never smoked. This suggests that there are natural individual differences in lung cancer susceptibility. With an estimated 15%-18% heritability, lung cancer in particular has a significant genetic component, making genetics an essential tool for comprehending the molecular mechanisms behind this fatal illness. Numerous genetic loci linked to the risk of lung cancer in a variety of populations have been found by genome-wide association studies, or GWASs. They conducted a meta-analysis of non-small-cell lung cancer (NSCLC including LUAD and LUSC) in

Chinese populations and discovered six more loci. Significantly, some of the loci identified in these research were more important in LUAD than in LUSC and vice versa, indicating that these two main histological forms of lung cancer have different genetic causes. It was found loci that were not significant in European datasets that primarily included smokers. These GWASs brought to light the heterogeneity of lung cancer according to histology, ancestry and smoking status, as well as its polygenic nature. However, because the majority of risk-associated genetic variants are non-protein coding and their functional implications are mainly unknown, it is difficult to decipher the molecular basis of lung cancer from these GWAS findings. There are currently few comprehensive GWAS follow-up studies on individual lung cancer loci, and it is challenging to link the genetic relationship with biological targets in order to establish.

Identifying lung cancer variants

Target genes from each locus have been identified in prior lung cancer GWAS follow-up studies, although functional variations have not always been identified. To date, the primary method for connecting lung-cancer-associated variations to their target genes has been bulk lung-tissue-based Expression Quantitative Trait Loci (eQTL). For variant-gene linkage investigations, various techniques are now accessible. For example, GWAS variations may be mapped to enhancers specific to lung cells that physically interact with gene promoters using chromatin interaction experiments. The single-cell level. However, the dominant functioning of individual variants is not taken into account by these gene identification techniques alone, which restricts our ability to comprehend which CCVs regulate potential causative genes and how. Furthermore, these variant-gene linkage techniques must be combined and applied to all accessible lung cancer GWAS datasets at the same time. It may be possible to improve target gene prediction and provide a high-level overview of lung cancer susceptibility across many populations by combining these datasets and adding the dominant functional information.