

Description of Additional Supplementary Files

Supplementary Data 1. MPRA Sequences.

Supplementary Data 2. MPRA results of functional regulatory variants (frVars).

Supplementary Data 3. Features included in the LungENN model.

Supplementary Data 4. Stratification analyses between potential causal variants and histological subtypes on risk of NSCLC.

Supplementary Data 5. Linkage disequilibrium with potential causal variants in different blocks at the 5p15.33 locus.

Supplementary Data 6. Associations between causal variants at 5p15.33 and leukocyte telomere length in our WGS dataset.

Supplementary Data 7. Variants in block 1 at 5p15.33 were associated with risk of various tumors.

Supplementary Data 8. Interactions between rs459961 and smoking status on risk of NSCLC.

Supplementary Data 9. Associations between rs10892229 and the expression of MPZL3 across different lung cell types.

Supplementary Data 10. The calculation of different polygenic risk scores.

Supplementary Data 11. Primers in MPRA experiments.

Supplementary Data 12. All guides RNA in the CRISPR experiments.

Supplementary Data 13. qPCR primers for target genes in the CRISPR experiments.