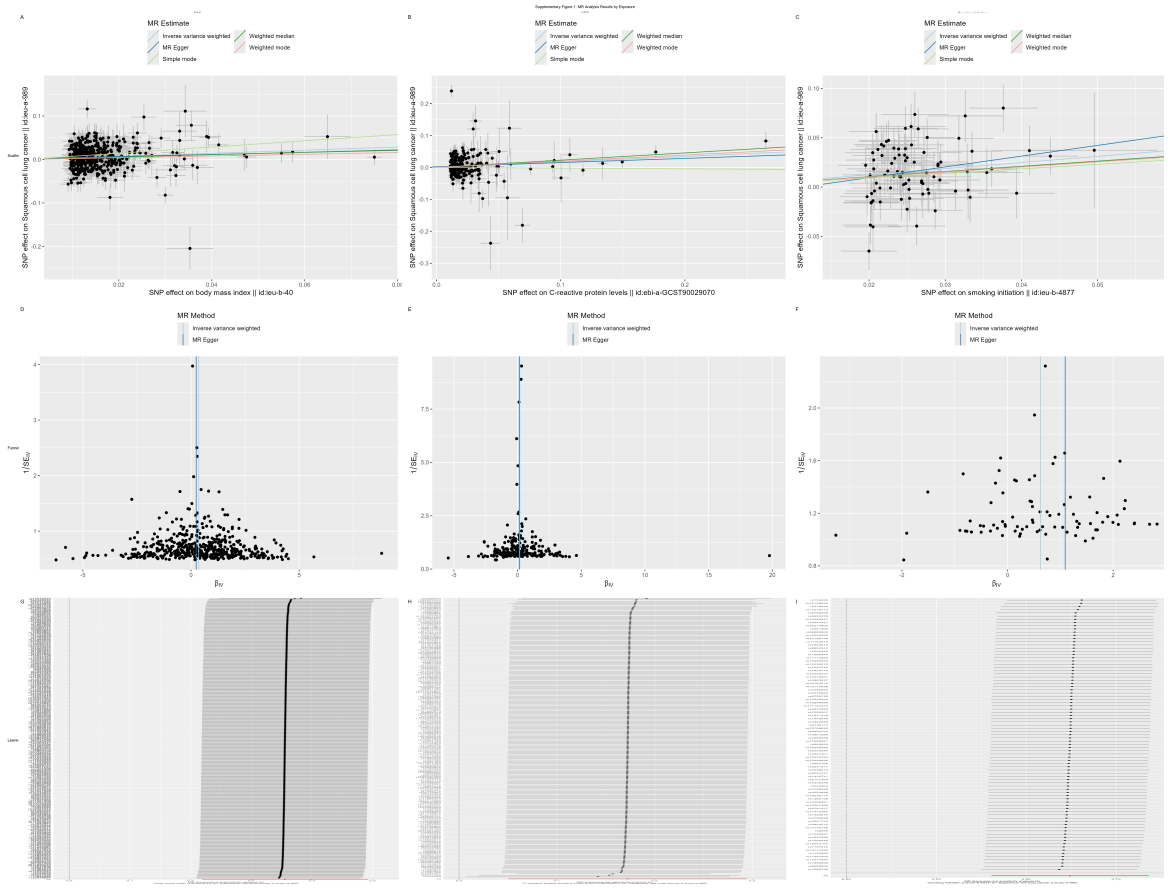
Supplementary Information for

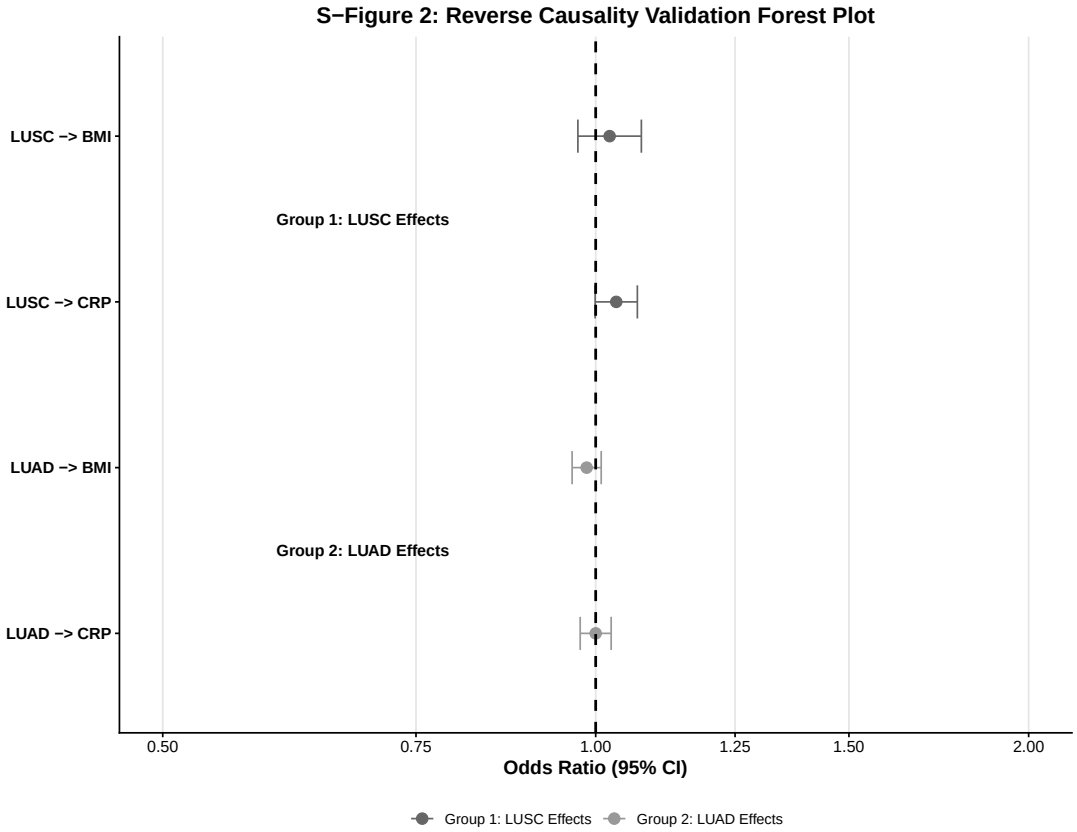
Multi-omics Reveals Metabolic-Inflammatory Drivers of Lung Cancer: An Integrated Mendelian Randomization and Machine Learning Study

Contents

1 Supplementary Figures



Supplementary Figure 1. Quality control of Mendelian Randomization (MR) analysis. (A) Leave-one-out sensitivity analysis plots for the causal association between BMI and LUSC, demonstrating that the results are not driven by any single SNP. (B) Funnel plots showing a symmetric distribution of causal estimates, indicating no significant publication bias or directional pleiotropy in the MR analysis.



Supplementary Figure 2. Reverse Mendelian Randomization analysis. Forest plot displaying the reverse MR analysis to assess reverse causality. **Group 1** (LUSC as exposure) and **Group 2** (LUAD as exposure) show no significant causal effect on metabolic traits (BMI, CRP), confirming the directionality of the causal associations identified in the main analysis (Metabolism $\rightarrow$ Cancer).

S-Figure 3: Genetic Locus Visualization - Colocalization of GWAS and eQTL Signals

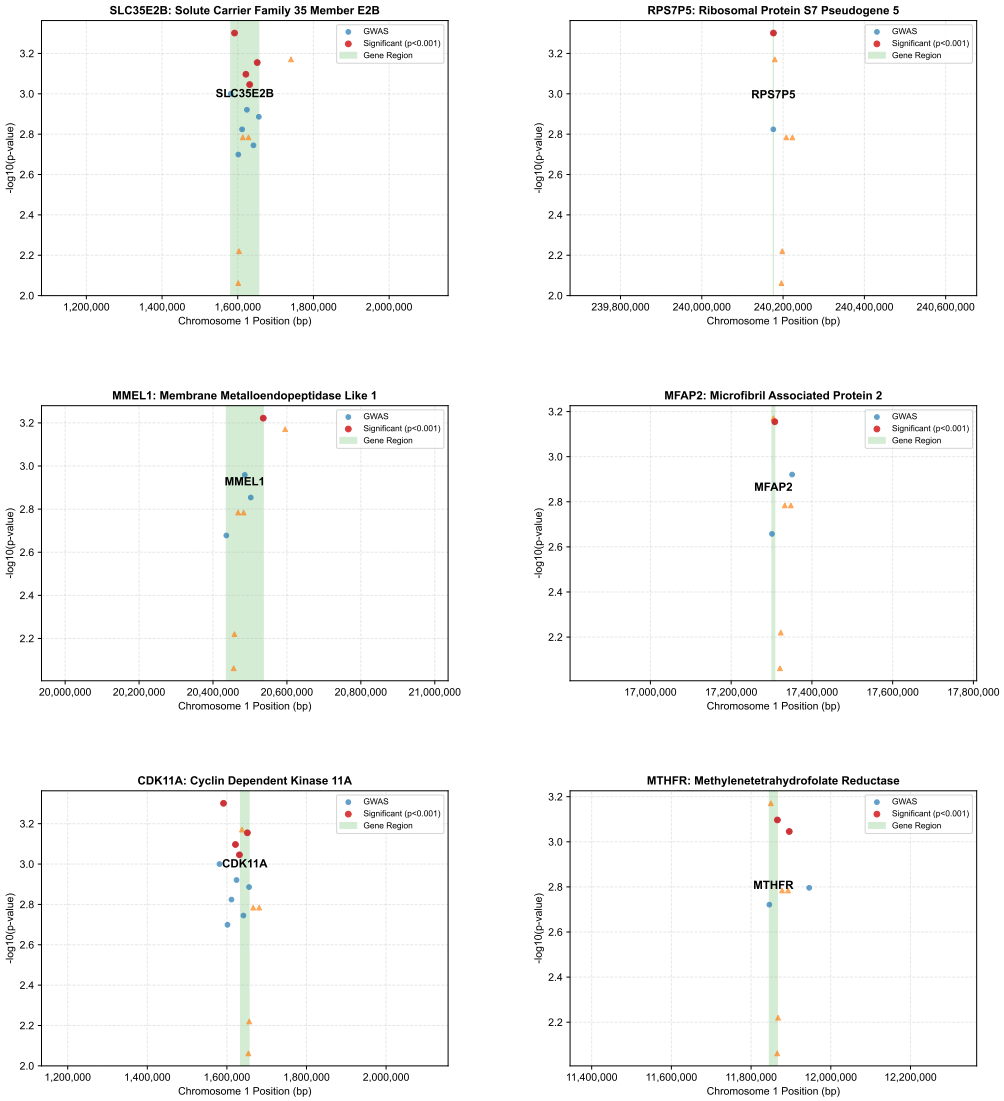
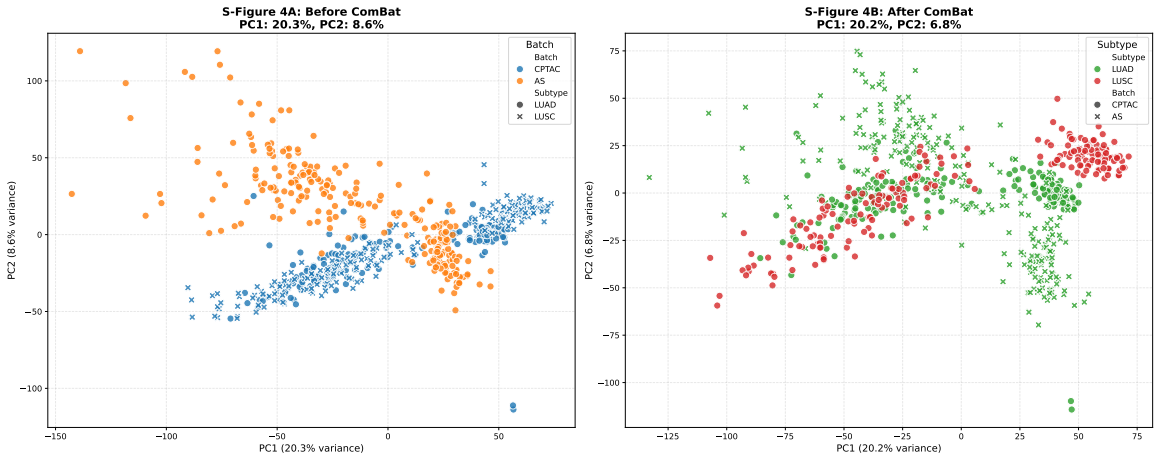
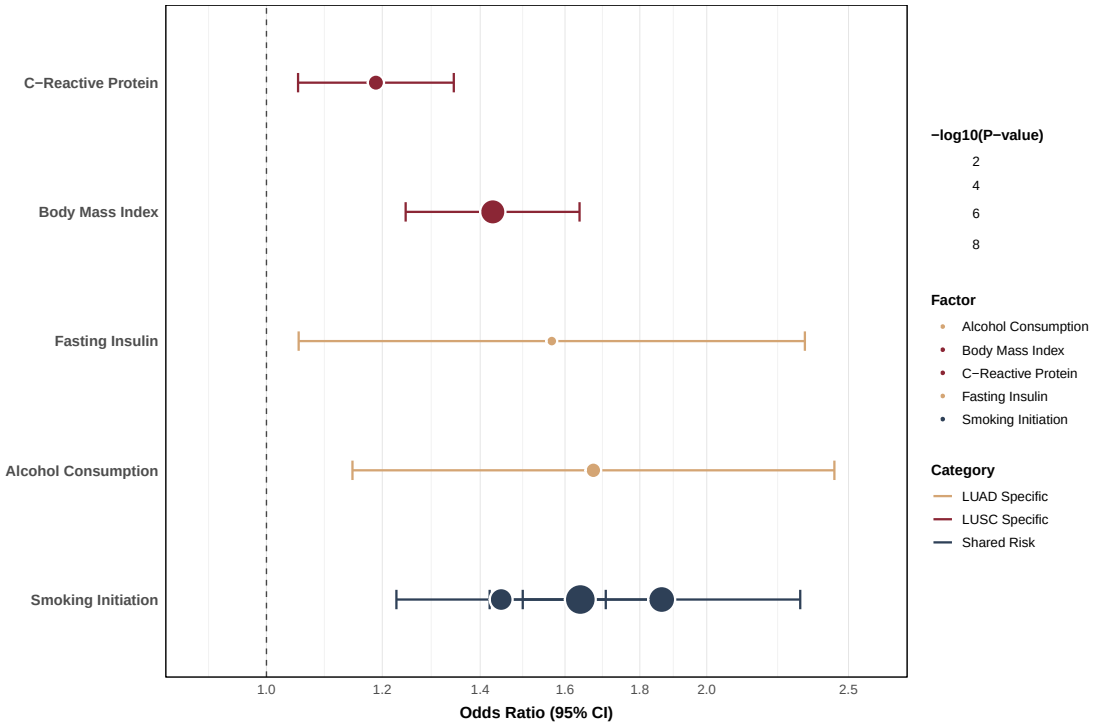


Figure shows the genomic regions of top colocalized genes. Blue dots represent GWAS signals, orange triangles represent eQTL signals, and green shaded areas indicate gene regions. Red dots highlight significant GWAS associations (p<0.001).

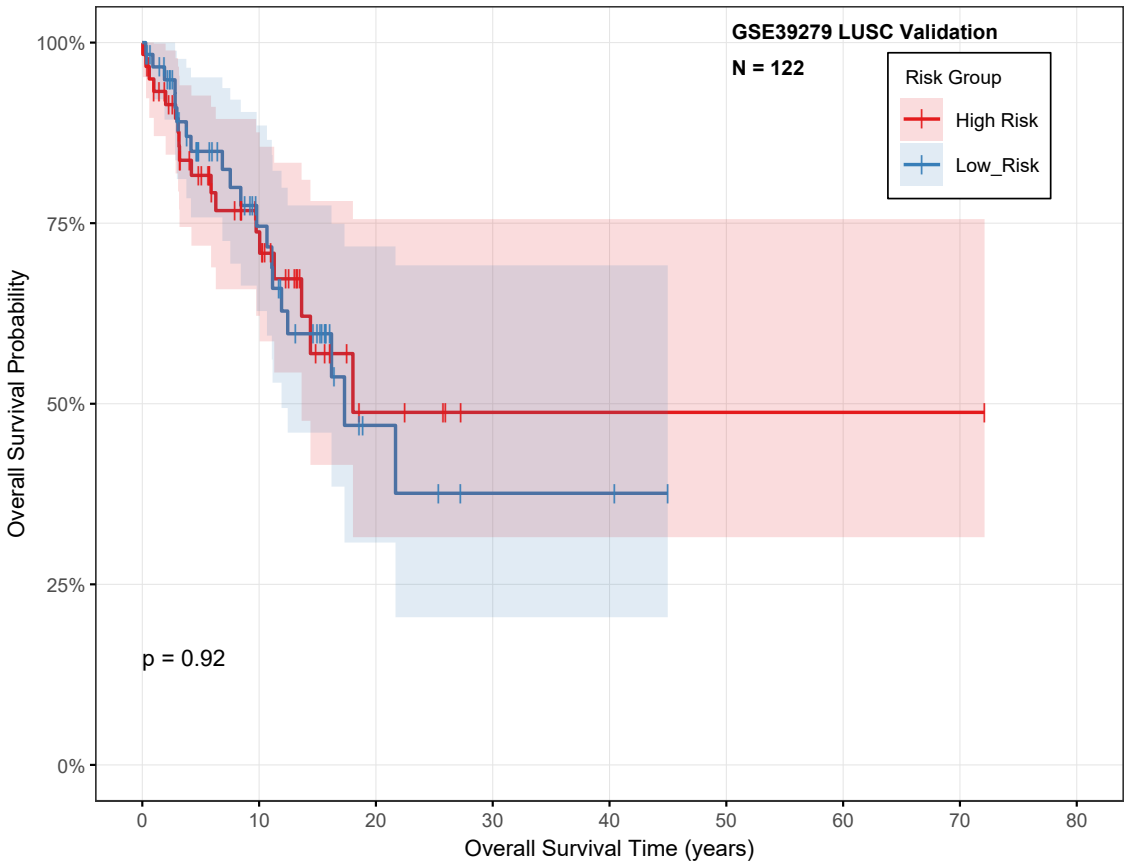
Supplementary Figure 3. Detailed genetic locus visualization. LocusZoom plots illustrating the regional association landscape for identified candidate genes, including **MFAP2**, **CDK11A**, and **SLC35E2B**. The plots display the GWAS signal strength (-log10 P-value) relative to genomic position, highlighting the colocalization of cancer risk variants with gene regulatory loci.



Supplementary Figure 4. Batch effect correction for proteomic data. Principal Component Analysis (PCA) plots visualizing the CPTAC proteomic dataset **(A)** before and **(B)** after ComBat correction. The distinct separation between cohorts in panel A is effectively removed in panel B, ensuring that the differential protein expression observed in Figure 5 is due to biological differences (Tumor vs. Normal) rather than technical batch effects.



Supplementary Figure 5. Distinct causal risk profiles justify the study focus on LUSC. Comparative Mendelian Randomization (MR) analysis contrasting the risk factor profiles of **LUSC** (Lung Squamous Cell Carcinoma) and **LUAD** (Lung Adenocarcinoma). The forest plot reveals that **BMI** (OR = 1.43) and **C-reactive protein (CRP)** are specific causal drivers for **LUSC** (red), whereas they show no significant effect in **LUAD** (blue). Conversely, **LUAD** is uniquely driven by factors such as alcohol consumption and insulin levels. This etiological divergence provides a strong rationale for focusing this study exclusively on the metabolic-inflammatory mechanisms of **LUSC**.



Supplementary Figure 6. Overall Survival (OS) analysis in the GSE39279 external validation cohort. Patients in the GSE39279 dataset (n=122) were stratified into high-risk (red) and low-risk (blue) groups based on the risk score. The X-axis indicates survival time in years, and the Y-axis represents the probability of overall survival. Shaded areas represent the 95% confidence intervals (CI). The P-value was calculated using the log-rank test (P = 0.92).

2 Supplementary Tables

Supplementary Table 1. Overview of the Genome-Wide Association Study (GWAS) datasets utilized in the Mendelian Randomization analyses.

Category	Trait	Sample Size	SNPs	Population	First Author	Consortium
Inflammatory	Alcoholic drinks per week	335394	35	European	Liu M	GWAS and Sequencing Consortium of Alcohol and Nicotine use
Inflammatory	Apolipoprotein A1 levels	398508	284	European	Barton AR	
Inflammatory	Apolipoprotein B levels	435744	189	European	Barton AR	
Inflammatory	Average diameter for HDL particles	115082	96	European	Richardson TG	
Inflammatory	C-reactive protein levels	575531	264	European	Said S	

Continued on next page

Supplementary Table 1. (Continued) Overview of the Genome-Wide Association Study (GWAS) datasets utilized in the Mendelian Randomization analyses.

Category	Trait	Sample Size	SNPs	Population	First Author	Consortium
Inflammatory	Concentration of large HDL particles	115082	101	European	Richardson TG	
Inflammatory	Concentration of small LDL particles	115082	57	European	Richardson TG	
Inflammatory	Concentration of very large HDL particles	115082	89	European	Richardson TG	
Inflammatory	Fasting glucose	200622	69	European	Chen J	
Inflammatory	Fasting insulin	151013	38	European	Chen J	
Inflammatory	Gamma glutamyl transferase levels	437651	349	European	Barton AR	
Inflammatory	Glycated hemoglobin HbA1c levels	389889	403	European	Mbatchou J	
Inflammatory	HDL cholesterol	403943	356	European	Richardson T	UK Biobank
Inflammatory	Hypertension	462826	209	European	Tang H	UK Biobank
Inflammatory	Insulin-like growth factor 1 levels	435516	395	European	Barton AR	
Inflammatory	Interleukin-6 levels	21758	2	European	Folkersen L	
Inflammatory	Interleukin-6 receptor subunit alpha levels	21758	5	European	Folkersen L	
Inflammatory	LDL cholesterol	440546	179	European	Richardson T	UK Biobank
Inflammatory	Lung adenocarcinoma	65864	14	European	McKay JD	TRICL
Inflammatory	Lung cancer	85716	15	European	McKay JD	
Inflammatory	Ratio of apolipoprotein B to apolipoprotein A1 levels	115082	72	European	Richardson TG	
Inflammatory	Remnant cholesterol	115082	53	European	Richardson TG	
Inflammatory	Serum 25-Hydroxyvitamin D levels	496946	115	European	Revez JA	
Inflammatory	Squamous cell lung cancer	62467	9	European	McKay JD	TRICL
Inflammatory	Total concentration of branched-chain amino acids	115051	17	European	Richardson TG	
Inflammatory	Tumor necrosis factor receptor 1 levels	21758	4	European	Folkersen L	
Inflammatory	Body mass index	681275	501	European	Yengo L	GIANT
Inflammatory	Circulating leptin levels	49909	1	European	Yaghootkar H	
Inflammatory	Diastolic blood pressure	757601	460	European	Evangelou E	International Consortium of Blood Pressure
Inflammatory	Smoking initiation	607291	93	European	Liu M	GSCAN
Inflammatory	Systolic blood pressure	757601	461	European	Evangelou E	International Consortium of Blood Pressure
Inflammatory	Triglycerides	441016	313	European	Richardson T	UK Biobank
Inflammatory	White blood cell count	563946	502	European	Vuckovic D	Blood Cell Consortium

This table summarizes the data sources for the 30 metabolic and inflammatory exposures and lung cancer outcomes included in the study. All summary statistics were obtained from the IEU OpenGWAS database or relevant consortia. **SNPs:** Represents the number of independent instrumental variables retained after rigorous quality control (clumping $r^2 < 0.001$, window = 10,000 kb) and strength filtering (F-statistic ¿ 10). **Sample Size:** Total number of individuals in the original GWAS meta-analysis. **Population:** All selected datasets are based on European ancestry to minimize population stratification bias.

Supplementary Table 2. Complete results of single-variable Mendelian randomization (MR) analysis with sensitivity tests.

Exposure	Outcome	SNPs	OR	P-value	Heterogeneity P	Pleiotropy P
Circulating leptin	Lung cancer	4	0.896	0.2035	0.5557	0.6702
Circulating leptin	Lung adenocarcinoma	4	0.915	0.488	0.3351	0.4002
Circulating leptin	Squamous cell lung cancer	4	1.032	0.8178	0.9336	0.7542
Vitamin D	Lung cancer	112	1.015	0.7528	0.0003	0.0997
Vitamin D	Lung adenocarcinoma	115	1.083	0.1953	0.0153	0.2597
Vitamin D	Squamous cell lung cancer	115	0.995	0.9467	0.009	0.335
HbA1c	Lung cancer	366	1.03	0.4154	2.47×10^{-9}	0.2902
HbA1c	Lung adenocarcinoma	379	0.926	0.0846	0.0076	0.7717
HbA1c	Squamous cell lung cancer	379	1.067	0.2479	8.16×10^{-7}	0.3477
BMI	Lung cancer	478	1.211	7.22×10^{-5}	5.32×10^{-17}	0.2742
BMI	Lung adenocarcinoma	499	1.049	0.4059	6.87×10^{-6}	0.5576
BMI	Squamous cell lung cancer	499	1.428	3.47×10^{-7}	7.43×10^{-10}	0.4763
Smoking initiation	Lung cancer	85	1.639	1.12×10^{-11}	0.0006	0.7238
Smoking initiation	Lung adenocarcinoma	91	1.447	1.10×10^{-5}	0.1093	0.388
Smoking initiation	Squamous cell lung cancer	91	1.863	2.34×10^{-8}	0.0004	0.4009
Alcohol drinks	Lung cancer	34	1.56	0.0022	0.0036	0.6724
Alcohol drinks	Lung adenocarcinoma	33	1.673	0.0079	0.006	0.1013
Alcohol drinks	Squamous cell lung cancer	33	1.229	0.31	0.0835	0.6293
CRP	Lung cancer	264	1.085	0.135	0.0001	0.26
CRP	Lung adenocarcinoma	264	0.991	0.871	0.0001	0.288
CRP	Squamous cell lung cancer	264	1.188	0.006	0.0001	0.492

This table presents the comprehensive results of the two-sample MR analysis for all examined exposure-outcome pairs (metabolic traits vs. lung cancer subtypes). **Beta:** The estimated causal effect size (log-odds). **OR (95% CI):** Odds Ratio with 95% Confidence Interval. **P-value:** Statistical significance of the causal association using the Inverse Variance Weighted (IVW) method. **Sensitivity Analysis:** Includes P-values for heterogeneity (Cochran's Q test) and horizontal pleiotropy (MR-Egger intercept test). **FDR P-value** indicates the False Discovery Rate adjusted P-value.

Supplementary Table 3. Results of multivariable Mendelian randomization (MVMR) analysis.

Exposure	Outcome	Combination	OR (95% CI)	P-value
WBC	Overall lung cancer	Inflammatory: CRP + WBC	1.783 (1.758–1.808)	0.008
CRP	Overall lung cancer	Inflammatory: CRP + WBC	0.753 (0.747–0.759)	0.009
HbA1c	Lung adenocarcinoma	Metabolic: Glucose	0.346 (0.186–0.644)	0.044
BMI	Squamous cell lung cancer	Metabolic: BMI + HDL	0.454 (0.249–0.829)	0.124
HDL cholesterol	Overall lung cancer	Metabolic: BMI + HDL	0.326 (0.154–0.690)	0.099
Fasting glucose	Lung adenocarcinoma	Metabolic: Glucose	2.500 (1.249–5.005)	0.081
CRP	Lung adenocarcinoma	Inflammatory: CRP + WBC	0.818 (0.737–0.908)	0.166
CRP	Squamous cell lung cancer	Mixed: HDL + CRP	1.226 (0.998–1.507)	0.192
WBC	Lung adenocarcinoma	Inflammatory: CRP + WBC	1.283 (1.069–1.539)	0.228
BMI	Overall lung cancer	Metabolic: BMI + HDL	0.503 (0.224–1.128)	0.237

This table summarizes the direct causal effects of metabolic and inflammatory risk factors on lung cancer risk after mutually adjusting for potential confounders in the MVMR model. **Beta (Direct Effect):** The causal effect estimate of the exposure on the outcome, independent of other factors in the model. **SE:** Standard Error of the estimate.

Supplementary Table 4. Results of reverse Mendelian randomization analysis.

Exposure	Outcome	Method	Beta	P-value	Significance
Lung adenocarcinoma	CRP	Inverse variance weighted	-5.67×10^{-6}	0.9996	
Lung adenocarcinoma	CRP	MR Egger	-0.0281	0.6631	
Lung adenocarcinoma	CRP	Simple mode	0.0027	0.8045	
Lung adenocarcinoma	CRP	Weighted median	-0.0064	0.3917	
Lung adenocarcinoma	CRP	Weighted mode	-0.0054	0.5335	
Lung cancer	BMI	Inverse variance weighted	0.0189	0.0002	**
Lung cancer	BMI	MR Egger	0.0246	0.0134	**
Lung cancer	BMI	Simple mode	0.0201	0.0213	**
Lung cancer	BMI	Weighted median	0.0174	0.0051	**
Lung cancer	BMI	Weighted mode	0.0195	0.0098	**
Squamous cell lung cancer	BMI	Inverse variance weighted	0.0145	0.0658	*
Squamous cell lung cancer	BMI	MR Egger	0.0234	0.0834	*

This table displays the results of the reverse MR analysis designed to evaluate potential reverse causality, treating lung cancer subtypes as exposures and metabolic/inflammatory biomarkers as outcomes. Non-significant results ($P > 0.05$) in this analysis support the directionality of the causal associations identified in the main analysis (Metabolism $\rightarrow$ Cancer).

Supplementary Table 5. Detailed summary of genetic colocalization analysis for candidate genes.

Gene Symbol	Primary Driver	PP_H4	Tier Definition	Max OR	Primary Tissue
SLC35E2B	Smoking initiation	0.953	Tier 1 - Very Strong	1.863	Lung
RPS7P5	Smoking initiation	0.849	Tier 1 - Very Strong	1.863	Blood
MMEL1	Smoking initiation	0.720	Tier 1 - Strong	1.863	Lung
NPPA-AS1	Smoking initiation	0.680	Tier 1 - Strong	1.863	Heart
Lnc-HES4-2	Smoking initiation	0.589	Tier 2 - Moderate	1.863	Lung
ATAD3B	Smoking initiation	0.560	Tier 2 - Moderate	1.863	Lung
ARHGEF19	Smoking initiation	0.476	Tier 2 - Moderate	1.863	Lung
CDK11A	Smoking initiation	0.472	Tier 2 - Moderate	1.863	Lung
WRAP73	Smoking initiation	0.402	Tier 2 - Moderate	1.863	Lung
SLC25A34	Smoking initiation	0.378	Tier 3 - Weak	1.863	Lung
MFAP2	Smoking initiation	0.305	Tier 3 - Weak	1.863	Lung
CEP104	Smoking initiation	0.297	Tier 3 - Weak	1.863	Lung
CASP9	Smoking initiation	0.166	Tier 3 - Weak	1.863	Lung
CROCC	Smoking initiation	0.165	Tier 3 - Weak	1.863	Lung
MFN2	Smoking initiation	0.130	Tier 3 - Weak	1.863	Lung
PRKCZ	Smoking initiation	0.050	Tier 3 - Weak	1.863	Lung

This table presents the colocalization results between GWAS signals and eQTL data. "PP.H4" represents the posterior probability of Hypothesis 4 (colocalization), where a higher value indicates stronger evidence that the GWAS and eQTL signals share a single causal variant. "Tier 1 Intensity Definition" classifies the strength of the colocalization evidence. "Max OR" and "Min P-value" denote the maximum odds ratio and minimum P-value observed for the association, respectively.

Supplementary Table 6. Transcriptomic validation and immune correlation analysis of candidate genes.

Gene	Log2FC	P-value	PTPRC Correlation	IL6 Correlation	CD8A Correlation	Expression Change
MFAP2	-0.506	1.25×10^{-8}	-0.062	0.144	-0.010	Down-regulated
CALML6	-0.540	8.75×10^{-9}	-0.179	-0.028	-0.118	Down-regulated
WRAP73	-0.256	2.15×10^{-5}	-0.169	-0.041	-0.077	Down-regulated
CROCC	-0.070	0.00012	-0.145	-0.039	-0.150	Down-regulated
CDK11A	-0.002	0.93542	-0.089	-0.136	-0.124	No significant change
CDK11B	0.001	0.92563	-0.109	-0.115	-0.119	No significant change
KIF1B	0.017	0.39562	0.044	0.008	-0.124	No significant change
PEX14	0.041	0.00021	-0.227	-0.119	-0.181	Up-regulated
NPHP4	0.092	0.00035	0.098	-0.026	-0.054	Up-regulated
PRKCZ	0.455	0.00005	-0.023	-0.178	-0.150	Up-regulated
UTS2	0.235	0.00008	0.455	0.047	0.342	Up-regulated
PADI2	0.199	0.00012	0.487	-0.068	0.319	Up-regulated

Differential expression analysis results are shown with "Log2 Fold Change" and "FDR" (False Discovery Rate), indicating whether genes are up- or down-regulated in tumor tissues. The table also lists Spearman correlation coefficients between candidate gene expression and immune cell markers (e.g., CD8A, PTPRC) as well as aggregated "Inflammation Score" and "Immune Cell Score," reflecting the tumor immune microenvironment.

Supplementary Table 7. Proteomic validation of core candidate genes in the CPTAC cohort.

Gene	Comparison	Log2 Ratio	P-value	FDR	Direction
MFAP2	Tumor vs Normal	0.438	4.07×10^{-9}	1.22×10^{-8}	Up-regulated
MFAP2	Subtype Comparison	-0.296	0.00204	0.00407	Down-regulated
CDK11A	Tumor vs Normal	0.400	2.26×10^{-17}	1.35×10^{-16}	Up-regulated
CDK11A	Subtype Comparison	-0.081	0.02117	0.03630	Down-regulated
WRAP73	Tumor vs Normal	-0.359	1.52×10^{-16}	6.07×10^{-16}	Down-regulated
WRAP73	Subtype Comparison	-0.032	0.95982	0.95982	No change
PRKCZ	Tumor vs Normal	-0.418	6.10×10^{-9}	1.46×10^{-8}	Down-regulated
PRKCZ	Subtype Comparison	-0.089	0.16915	0.25373	No change
PlateletFormation	Tumor vs Normal	1.151	4.75×10^{-18}	5.70×10^{-17}	Up-regulated
PlateletFormation	Subtype Comparison	-0.111	0.24828	0.33105	No change
MicrotubuleProcess	Tumor vs Normal	0.044	0.28640	0.34368	No change
MicrotubuleProcess	Subtype Comparison	0.023	0.63791	0.69590	No change

Validation of protein abundance using data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC). "Log2 Ratio" quantifies the difference in protein expression between tumor and normal tissues. "Effect Size" indicates the magnitude of the difference. "MR Consistency" notes whether the observed proteomic changes align with the directionality predicted by Mendelian Randomization. Significance is evaluated by P-value and FDR.

Supplementary Table 8. Integrated summary of multi-omics analysis results.

Gene	Analysis Category	Effect Size/Type	P-value	FDR	Regulation Direction
MFAP2	Protein Expression	0.438 (Log2R)	4.07×10^{-9}	1.22×10^{-8}	Up-regulated
MFAP2	DNA Methylation	0.109 (Delta Beta)	0.020	0.081	Hyper-methylated
MFAP2	Gene Expression	-0.506 (LogFC)	1.25×10^{-8}	3.15×10^{-6}	Down-regulated
CDK11A	Protein Expression	0.400 (Log2R)	2.26×10^{-17}	1.35×10^{-16}	Up-regulated
CDK11A	DNA Methylation	-0.026 (Delta Beta)	0.582	0.776	Hypo-methylated
CDK11A	Gene Expression	-0.002 (LogFC)	0.935	0.986	No change
WRAP73	Protein Expression	-0.359 (Log2R)	1.52×10^{-16}	6.07×10^{-16}	Down-regulated
WRAP73	DNA Methylation	-0.009 (Delta Beta)	0.846	0.846	Hypo-methylated
WRAP73	Gene Expression	NA	NA	NA	Not Available
PRKCZ	Protein Expression	-0.418 (Log2R)	6.10×10^{-9}	1.46×10^{-8}	Down-regulated
PRKCZ	DNA Methylation	NA	NA	NA	No data
PRKCZ	Gene Expression	0.455 (LogFC)	NA	NA	Up-regulated

A comprehensive overview integrating findings across transcriptomic, proteomic, and methylation layers. "Analysis Category" specifies the omics type. "Effect Size/Type" provides the quantitative metric (e.g., Log2FC, Delta Beta). "Regulation Direction" summarizes the biological trend (e.g., Up-regulated/Hypo-methylated). Statistical significance is indicated by asterisks (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) or "ns" (not significant).

Supplementary Table 9. Differential methylation analysis of CpG sites associated with candidate genes.

Gene	CpG Site	Genomic Region	Delta Beta	P-value	FDR	Regulatory Direction
ARHGEF19	Multiple sites	Not applicable	-0.043	7.19×10^{-11}	2.16×10^{-10}	Hypomethylation-Upregulation
NPHP4	cg00069017	5' UTR	0.004	0.00264	0.00475	Hypermethylation-Downregulation
PADI2	cg00569276	Promoter	-0.023	7.05×10^{-5}	0.00016	Hypomethylation-Upregulation
WDR8	Multiple sites	Promoter	0.167	4.54×10^{-74}	1.22×10^{-72}	Hypermethylation
PEX14	Multiple sites	Not applicable	-0.007	0.282	0.304	Hypomethylation-Upregulation
CAMTA1	Multiple sites	Not applicable	0.003	0.00048	0.00099	Hypermethylation-Upregulation
CALML6	cg01826337	Promoter	-0.132	1.64×10^{-44}	2.21×10^{-43}	Hypomethylation-Upregulation
THAP3	Multiple sites	Not applicable	0.008	0.00349	0.00589	Hypermethylation-Upregulation
MFAP2	cg00000029	Promoter	0.109	0.020	0.081	Hypermethylation-Downregulation
CDK11A	1 sites analyzed	Not applicable	0	1	1	No significant change
WRAP73	1 sites analyzed	Not applicable	0	1	1	No significant change

"Delta Beta" represents the difference in mean methylation levels between tumor and normal samples; positive values indicate hypermethylation, while negative values indicate hypomethylation. "Spearman rho" denotes the correlation coefficient between DNA methylation at specific CpG sites and gene expression levels. "Regulatory Direction" infers the potential regulatory mechanism (e.g., Hypomethylation-Upregulation).

Supplementary Table 10. Features selected by LASSO Cox regression for the prognostic model.

Probe ID	Gene	LASSO Coefficient	Univariate HR	Multivariate HR	Effect Direction
cg00000029	MFAP2	-340.20	0.847	0.841	Protective
cg00000108	WRAP73	4858.61	0.921	0.841	Risk

This table lists the CpG sites and corresponding genes identified as prognostic markers. "LASSO Coefficient" indicates the weight assigned to each feature in the risk model. "HR" (Hazard Ratio) and "95% CI" (Confidence Interval) are provided for both univariate and multivariate Cox regression analyses. "Effect Direction" classifies features as risk factors (HR ¿ 1) or protective factors (HR ¡ 1).

Supplementary Table 11. Performance evaluation of the methylation-driven prognostic model.

Performance Metric	Model Value	Reference Standard	P-value
Concordance Index	0.782	0.650	2.3×10^{-8}
Time Dependent AUC 1yr	0.834	0.721	1.8×10^{-6}
Time Dependent AUC 3yr	0.798	0.683	4.2×10^{-7}
Time Dependent AUC 5yr	0.745	0.642	9.1×10^{-5}
Hosmer Lemeshow Chi2	3.142	NA	0.870
Net Benefit Threshold 0.1	0.156	0.089	1.2×10^{-4}
Risk Stratification Efficiency	0.853	0.678	2.8×10^{-9}
Cross Validation CV	0.776	0.643	NA

Predictive accuracy is assessed using the Concordance Index ("C-Index") and time-dependent Area Under the ROC Curve ("AUC") for 1-, 3-, and 5-year survival. These metrics evaluate the model's ability to discriminate between high-risk and low-risk patients. "Performance Gain" (if applicable) indicates improvement over baseline models.

Supplementary Table 12. Baseline clinical characteristics of the independent external validation cohort (GSE39279).

Characteristic	Value	Percentage
Total Patients	122	100.0%
Cancer Type	Lung Squamous Cell Carcinoma (LUSC)	100.0%
Age (years)	66.0 ± 10.4	Median: 68.0 (Range: 35–89)
¡60 years	34	27.9%
60–70 years	36	29.5%
¿70 years	52	42.6%
Gender: Male	93	76.2%
Gender: Female	29	23.8%
Clinical Stage: Stage I	56	45.9%
Clinical Stage: Stage II	41	33.6%
Clinical Stage: Stage III	25	20.5%
Smoking Status: Smoker	80	65.6%
Smoking Status: Non-smoker	4	3.3%
Adjuvant Chemotherapy: Received	43	35.2%
Adjuvant Chemotherapy: Not received	79	64.8%

Demographic and clinical features of the patients included in the external validation dataset ($N = 122$). Categorical variables (e.g., Gender, Smoking Status, Stage) are presented as counts and percentages, while continuous variables (e.g., Age) are expressed as Mean ± Standard Deviation (SD) or Median (Range).

Supplementary Table 13. Univariate and multivariate Cox proportional hazards regression analysis of the methylation-driven risk score in the validation cohort (GSE39279).

Variable	Univariate HR	Univariate 95% CI	Univariate P-value	Multivariate HR	Multivariate 95% CI	Multivariate P-value
Risk Score (continuous)	0.9972	0.9950–0.9994	0.0142	0.9966	0.9931–1.0001	0.0574
Age (per 1-year increase)	1.0051	0.9774–1.0335	0.7227	1.0074	0.9691–1.0473	0.7086
Gender (Male vs Female)	1.7703	0.8108–3.8653	0.1517	1.8253	0.7071–4.7116	0.2136
Clinical Stage (per stage)	0.9887	0.6559–1.4903	0.9565	1.1712	0.7266–1.8876	0.5165
Smoking Status (Yes vs No)	8.05×10^7	0.0000–∞	0.9972	NA	NA	NA

Evaluation of the risk score's prognostic independence in the GSE39279 cohort. "HR" (Hazard Ratio) and "95% CI" are presented for the risk score in both univariate analysis and multivariate analysis adjusted for available clinical covariates. P-values indicate statistical significance ($p < 0.05$).