Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Precision Epigenetic Approach to Non-Invasive Lung Cancer Screening Using Gene- Specific cfDNA Methylation

Santosh Kumar
KING GEORGES MEDICAL UNIVERSITY
Seema Singh
seemasingh4853@kgmcindia.edu

KING GEORGES MEDICAL UNIVERSITY
Surya kant
KING GEORGES MEDICAL UNIVERSITY
Ajay kumar Verma
KING GEORGES MEDICAL UNIVERSITY
M K Ahmad
KING GEORGES MEDICAL UNIVERSITY
Dileep Verma
KING GEORGES MEDICAL UNIVERSITY

Article

Keywords: Lung cancer screening, Circulating cell-free DNA, DNA methylation, Liquid biopsy, Gene-specific methylation signatures, Early detection

Posted Date: March 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9160875/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Precision Epigenetic Approach to Non-Invasive Lung Cancer Screening Using Gene-Specific cfDNA Methylation

Authors
Santosh Kumar^1, Seema Singh^1*, Surya Kant^1, Ajay Kumar Verma^2, M. K. Ahmad^3, Dileep Verma^4

Affiliations
^1 Department of Respiratory Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
^2 Department of Respiratory Medicine, Dr. Ram Manohar Lohia Institute of Medical Sciences (Dr. RMLIMS), Lucknow 226010, Uttar Pradesh, India
^3 Department of Biochemistry, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
^4 Department of Physiology, King George's Medical University, Lucknow 226003, Uttar Pradesh, India

Author details
Santosh Kumar, MD
Professor, Department of Respiratory Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
Email: skumarchest@gmail.com

Seema Singh, PhD
Department of Respiratory Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
Email: seemasingh4853@kgmcindia.edu
ORCID: 0000-0002-6049-3287

Surya Kant, MD
Professor and Head, Department of Respiratory Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
Email: skantpulmed@gmail.com
ORCID: 0000-0001-7520-5404

Ajay Kumar Verma, MD
Professor and Head, Department of Respiratory Medicine, Dr. Ram Manohar Lohia Institute of Medical Sciences (Dr. RMLIMS), Lucknow 226010, Uttar Pradesh, India
Email: ajaykumarverma@kgmcindia.edu
ORCID: 0000-0002-2973-1793

M. K. Ahmad, PhD
Professor, Department of Biochemistry, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
Email: mohdkaleemahmad@kgmcindia.edu

Dileep Verma, MD
Professor, Department of Physiology, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
Email: dileepverma200131@rediffmail.com

Corresponding author
Seema Singh, PhD
Department of Respiratory Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India
Email: seemasingh4853@kgmcindia.edu
ORCID: 0000-0002-6049-3287

Abstract:

Purpose: Lung cancer remains the leading cause of cancer-related death worldwide, underscoring the need for reliable biomarkers for early detection. This study evaluated the promoter methylation and expression of MAX, MTURN, HLA-B, and CAV1 in tumor and tumor-free tissues and validated their methylation status in plasma cfDNA from patients with NSCLC and SCLC. It further assessed the potential of this multi-gene methylation panel as a non-invasive surrogate biomarker for early lung cancer detection and improved therapeutic outcomes.

Methods: In this study, we profiled promoter methylation in LC tissues and evaluated matched plasma-derived circulating cell-free DNA (cfDNA) as a "tissue-mirror" biomarker. mRNA expression of MAX, MTURN, HLA-B and CAV1 was quantified in tumor tissue (TT) from non-small cell (NSCLC) and small cell (SCLC) lung cancers versus tumor-free tissue (TF), and promoter methylation in tissue gDNA was assessed by methylation-specific PCR (MSP). The same loci were interrogated in plasma cfDNA from NSCLC and SCLC cases relative to healthy controls.

Results: In TT, methylation frequencies for MAX, MTURN and HLA-B were significantly elevated 68.45%, 56.94%, and 75.97%—compared with TF (20.42%, 15.28%, and 29.57%;

$p < 0.05$ to $p < 0.001$). Correspondingly, TT showed down-regulation of gene expression ($p < 0.01$) with strong inverse associations to hypermethylation; CAV1 was the exception, demonstrating up-regulation with hypomethylation in NSCLC TT. Plasma cfDNA recapitulated tissue patterns, exhibiting significant hypermethylation in LC versus controls ($p \leq 0.05$ to $p \leq 0.001$) and yielding positive predictive values of 59.33–77.47% and negative predictive values of 52.47–67.33%.

Conclusion: Collectively, these findings indicate that promoter methylation of this four-gene panel in plasma can aid LC diagnosis and functions as a minimally invasive, tissue-reflective biomarker.

Keywords: Lung cancer screening; Circulating cell-free DNA; DNA methylation; Liquid biopsy; Gene-specific methylation signatures; Early detection

1. **Introduction:** Lung cancer (LC) is the most prevalent malignancy worldwide and continues to be the foremost cause of cancer-related mortality in both men and women. Globally, its mortality rate approaches 18%, surpassing that of liver cancer (9.4%), gastric cancer (8.3%), and female breast cancer (7.7%). Globally, lung cancer remains the leading cause of cancer-related death, with 1,817,469 deaths in 2022, exceeding liver cancer (758,725 deaths), breast cancer (666,103 deaths), and stomach cancer (660,175 deaths). (Ferlay et al. 2024) Detecting LC at an early stage greatly improves patient survival and enhances therapeutic efficacy. In this context, molecular biomarkers have emerged as promising tools for early diagnosis as well as for monitoring treatment response.

   Recent advances underscore the pivotal role of epigenetic modifications in cancer pathogenesis. These modifications represent heritable alterations in gene regulation that occur without changes to the underlying DNA sequence (Belinsky 2004, 2005). Among the most stable of these changes are aberrant methylation events—either hypermethylation or hypomethylation—of CpG islands within promoter regions, catalyzed by DNA methyltransferases (Belinsky 2004, 2005). Numerous studies have demonstrated that DNA methylation profiles vary not only between different cancer types but also across specific gene groups (Toyooka et al. 2001; Constâncio et al. 2020).

Importantly, DNA methylation biomarkers have been detected not only in tumor tissues but also in easily accessible body fluids such as sputum, serum, and plasma, even during the earliest stages of disease (Shi et al. 2020; Weiss et al. 2017; Hu et al. 2023). Circulating cell-free DNA (cfDNA)—released into the bloodstream through apoptosis, necrosis, or passive shedding from both malignant and healthy cells—has drawn considerable attention (Moss et al. 2018; Constâncio et al. 2020). A significant portion of cfDNA consists of circulating tumor DNA (ctDNA), and its fraction is influenced by tumor burden, location, and vascularity (Jamshidi et al. 2022). This variability has positioned cfDNA as a powerful substrate for liquid-biopsy approaches. Epigenetic interrogation of cfDNA, particularly DNA methylation analysis, therefore holds substantial promise for non-invasive detection of LC.

In both non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), aberrant promoter methylation of key tumor suppressor genes and oncogenes has been consistently reported (Belinsky 2004, 2005; Toyooka et al. 2001). However, most existing studies are constrained by their focus on single-gene analyses or by limited cohort sizes. To overcome these challenges, multi-gene methylation panels—initially identified in tumor tissue-derived genomic DNA (gDNA) and subsequently validated in plasma-derived cfDNA—are increasingly being developed, offering a more reliable and clinically applicable strategy for advancing methylation-based biomarkers in lung cancer (Weiss et al. 2017; Wei et al. 2021; Hu et al. 2023).

MAX is a central transcription factor that interacts with MYC, MAD, and MXD family proteins, thereby maintaining a delicate balance between cell proliferation and programmed cell death. Within the MYC/MAX/MAD regulatory axis, MAX primarily functions as a tumor suppressor. In lung cancer, MAX dysregulation or inactivation has been associated with tumor progression and adverse biological behavior (Romero et al. 2014; Liu et al. 2019). MTURN, traditionally studied in the context of neuronal differentiation and development, is only recently gaining attention in oncology. Emerging data suggest its inclusion in biomarker panels relevant to lung cancer, particularly alongside MAX and HLA-B (Liu et al. 2019; Ying et al. 2022). HLA-B encodes a major histocompatibility complex (MHC) class I glycoprotein that presents antigenic peptides to cytotoxic T cells, thereby orchestrating immune surveillance. Epigenetic silencing of HLA class I genes through promoter hypermethylation has been associated with reduced expression and immune

evasion in cancer (Nie et al. 2001; Ye et al. 2010; Puttick et al. 2024). CAV1 has also been implicated in methylation-based lung cancer biology and shows subtype-dependent behavior in NSCLC and SCLC (Sunaga et al. 2004; Brodie et al. 2014; Guo et al. 2016). Large-scale prospective studies and recent validations have further demonstrated that DNA methylation panels comprising multiple genes can markedly improve both sensitivity and diagnostic accuracy in lung cancer (Weiss et al. 2017; Wei et al. 2021; Hu et al. 2023). However, cumulative analysis of panels including MAX, MTURN, HLA-B, and CAV1 remains limited. In the present study, we evaluated gene expression patterns in tumor tissue (TT) relative to tumor-free tissue (TF) and correlated these with promoter methylation in tissue-derived genomic DNA (gDNA). Furthermore, methylation profiling of these genes was performed in plasma-derived cfDNA from patients with both NSCLC and SCLC, as well as from healthy controls, to assess their potential as plasma-based surrogate biomarkers that reflect tumor methylation status.

2. Methodology:

Study Designs and Sample Collection: **Ethics statement**This study was approved by the Institutional Ethics Committee of King George's Medical University, Lucknow, India (Ref. No. 110th ECM IIA/P2). Written informed consent was obtained from all participants prior to inclusion in the study. All methods were carried out in accordance with relevant guidelines and regulations, and in accordance with the Declaration of Helsinki.

Written informed consent was voluntarily obtained from all participants (both cases and controls) prior to sample collection, in accordance with the defined inclusion and exclusion criteria. Individuals with HIV, HBV, HCV, autoimmune disorders, pulmonary tuberculosis, diabetes, hypertension, pregnancy, genetic disorders, or any other malignancies were excluded. Only newly diagnosed lung cancer patients who had not undergone chemotherapy were considered eligible.Baseline demographic and clinical variables including age, sex, ethnicity, tumor histology and date of initial diagnosis were documented. Additional details such as smoking status, pathological stage (TNM classification), and tumor grade were also recorded. Participant recruitment was carried out between 2021 and 2024.

For this work, both a discovery cohort and a validation cohort were established

- The validation cohort comprised blood samples from 80 healthy controls and 80 lung cancer patients (68 NSCLC and 12 SCLC). Lung cancer cases were first identified through clinical radiology and subsequently confirmed via histopathological analysis of primary tumor biopsies obtained during bronchoscopy or through image-guided transthoracic biopsy. The healthy volunteers were age-matched to the patient group. From each subject, 5 mL of peripheral blood was collected in EDTA vials. Samples were centrifuged at 2000 rpm for 10 minutes, and the separated plasma was stored at –80 °C until extraction of circulating cell-free DNA (cfDNA).
- The discovery cohort included tumor tissue (TT; n = 40) and tumor-free tissue (TF; n = 40). The TT group comprised 33 NSCLC and 7 SCLC tissue samples, collected at biopsy during bronchoscopy. TF samples were obtained from individuals who underwent interventions for suspected lung cancer but were later confirmed to be disease-free following histopathological examination. Each tissue specimen was divided into two portions: one preserved in PBS at –80 °C for DNA extraction, and the other stored in 1 mL of TRIzol reagent (Invitrogen, Thermo Fisher Scientific, MA, USA) at –80 °C for RNA extraction.

RNA Extraction and Gene Expression

Total RNA was isolated from tissue specimens using TRIzol reagent (Invitrogen, Thermo Fisher Scientific). RNA concentration and purity were assessed on a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) by measuring A260/A280. Equalized RNA inputs were reverse-transcribed with the iScript High-Capacity cDNA Synthesis Kit (Bio-Rad) to generate cDNA for all samples. Equivalent amounts of cDNA from tumor tissue (TT) and tumor-free tissue (TF) were used for quantitative real-time PCR on an Applied Biosystems 7500 Fast system (Thermo Fisher Scientific). Reactions (20 μL) contained cDNA, SYBR Green master mix (Bio-Rad), and gene-specific forward/reverse primers, following the manufacturer's instructions. Primers for MAX, MTURN, HLA-B and CAV1 were designed using the Integrated DNA Technologies (IDT) design tool and sourced from IDT. β-actin served as the internal reference gene. All primer sequences are provided in Supplementary Table (S1). Threshold cycle (Ct) values were recorded as primary outputs.

For each target, relative expression (fold change) between TT and TF was calculated on a per-gene basis. All measurements were performed in technical triplicate for every sample.

DNA Extraction and Quality Assessment

Genomic DNA was purified from equivalent input amounts of tissue and plasma using the PureLink™ Genomic DNA Mini Kit (Invitrogen, Thermo Fisher Scientific) according to the manufacturer's instructions. Yield and purity were determined spectrophotometrically via A260/A280 measurements. DNA integrity was evaluated by electrophoresis on 0.8% agarose gels.

Bisulfite Conversion

From each tissue and plasma specimen, 20 µL of purified DNA was subjected to sodium bisulfite conversion using the EpiTect® Fast Bisulfite Kit (Qiagen, Hilden, Germany), following the denaturation and conversion steps specified by the manufacturer. The reaction chemistry converts unmethylated cytosines to uracil, while methylated cytosines remain unchanged.To verify DNA suitability and normalize input for MSP, each bisulfite-treated sample was first amplified for the β-actin locus as an internal control. The qualified bisulfite-converted DNA then served as template for methylation-specific PCR (MSP) to determine promoter methylation status (methylated vs. unmethylated) of the target genes.

Methylation-Specific PCR (MSP) Analysis:

Promoter methylation of *MAX, MTURN, HLA-B* and *CAV1* was assessed by MSP following established methods.Reaction mix (20 µL). Each MSP contained 2 µL bisulfite-converted DNA, 7 µL EmeraldAmp 1× PCR master mix (Takara Bio, USA), 1.5 µL each of forward and reverse primers (10 pM), and 8 µL nuclease-free water.Amplification was performed on a Bio-Rad T100™ thermal cycler with, Initial denaturation: 95 °C, 15 min. 40 cycles of: denaturation 95 °C, 30 s; gene-specific annealing 30 s at 56 °C (MAX), 52 °C (MTURN), 58 °C (HLA-B); and 50 °C (CAV1) and a final extension at 72 °C, 30 s.Separate reactions were run with methylated (M) and unmethylated (U) primer sets for each gene.Amplicons were resolved on 3% agarose gels, stained with EtBr, and documented on a Gel Doc Go imaging

system (Bio-Rad).Universal human controls for methylated and unmethylated DNA (Qiagen EpiTect) were included alongside no-template controls. Samples were deemed M or U based on band presence in the respective M/U reactions; frequencies were calculated accordingly. β-actin served as an internal control to verify bisulfite DNA suitability and normalization.
Primer design & validation. MSP primers for *MAX, MTURN, HLA-B, CAV1*and *β-actin* (M and U sets) were designed using Methyl Primer from NCBI reference sequences, evaluated with Sequence Manipulation Suite: PCR Primer Stats, and synthesized by Integrated DNA Technologies (IDT). All primer sequences are listed in Supplementary Table S2.

Statistical analysis: Group differences in variable frequencies were tested with Fisher's exact test. Odds ratios (ORs) with 95% confidence intervals (CIs), along with positive predictive value (PPV) and negative predictive value (NPV), were computed using MedCalc (OR calculator). For real-time/qPCR data, one-way ANOVA was applied to parametric data; Kruskal–Wallis tests were used for non-parametric data.Associations between gene expression and promoter methylation in tumor tissue (TT) versus tumor-free tissue (TF) were examined by column-wise analysis followed by Spearman's rank correlation.Unless otherwise noted, analyses were performed in GraphPad Prism 5.0 (GraphPad Software, San Diego, CA), using two-sided $\alpha = 0.05$ as the threshold for statistical significance.Differences in nominal/qualitative methylation frequencies across groups were evaluated with chi-square tests on contingency tables.ROC curves and area under the curve (AUC) were generated in SPSS v29.0.0 (Build 241) to assess diagnostic utility. Significance is annotated as $p< 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***).Tissue–plasma agreement. Concordance of methylation calls between plasma-derived cfDNA and tissue-derived DNA was quantified using Cohen's kappa (κ).

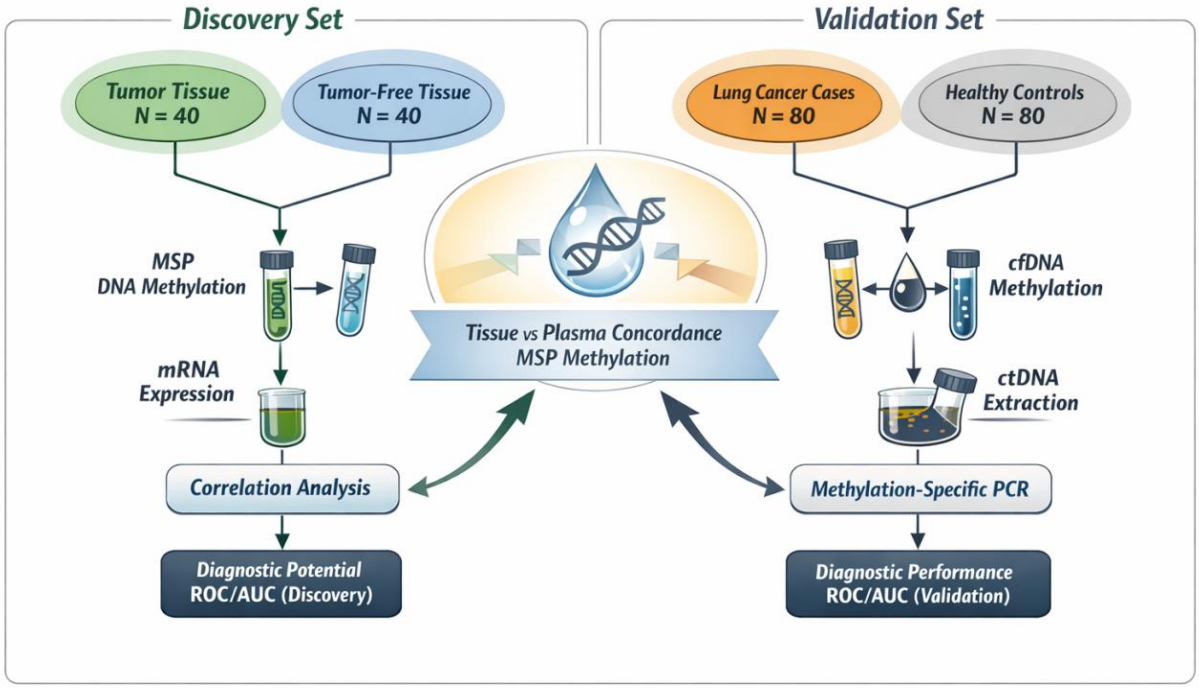


Fig. 1 Workflow of biomarker discovery (tissue) and assay validation (plasma), including methylation/expression analyses, diagnostic evaluation (ROC/AUC), and tissue-plasma concordance (Cohen's κ).

3. Results

3.1 Clinicopathological and Demographic Characteristics of the Study Cohort

The clinicodemographic and pathological characteristics of the study population are summarized in Table 1. In the discovery cohort, comprising tumor tissue (TT, $n = 40$) and tumor-free tissue (TF, $n = 40$), no significant differences were observed between groups with respect to age distribution ($p = 0.96$), sex ($p = 0.96$), or place of residence ($p = 0.99$), indicating that the two groups were broadly comparable for baseline demographic variables. Participants aged >51 years constituted the largest subgroup in both TT and TF, and males predominated in both groups. Histopathological classification of the discovery cases showed that adenocarcinoma (AD) was the most frequent subtype, followed by squamous cell carcinoma (SCC) and small cell lung cancer (SCLC). Most patients presented with advanced disease, with stage IV representing the majority of cases, while stage III and stage II were less frequent. Tumors were further classified according to the International Association for the Study of Lung Cancer (IASLC) criteria, and staging/TNM grading was assigned in accordance with the 7th and 8th IASLC editions. Similarly, in the validation cohort (cases, $n = 80$; controls, $n = 80$), age, sex, and residential distribution did not differ significantly between groups ($p = 0.28$, $p = 0.87$, and $p = 0.30$, respectively). As in the discovery cohort, participants aged >51 years and male subjects formed the majority, and AD remained the predominant histological subtype among cases, followed by SCC and SCLC. The validation cohort was likewise characterized by predominantly advanced-stage disease, with most patients classified as stage IV. In contrast to the comparable demographic variables, smoking status differed significantly between groups in both cohorts. In the discovery cohort, smoking distribution differed between TT and TF ($p = 0.03$), whereas in the validation cohort, smoking status showed a highly significant difference between cases and controls ($p < 0.0001$), with controls being predominantly non-smokers and ex-smokers being more frequent among cases. Collectively, these findings indicate that the study cohorts were demographically comparable, while smoking exposure showed a significant association with lung cancer status.

Table 01: Demographic and clinicopathological characteristics of participants in the discovery set (TT and TF) and Validation set (Lung cancer–case, control).

Characteristics	Discovery set			Validation set		
	TT (N=40) (%)	TF (N=40) (%)	p value	Cases (N=80) (%)	Controls (N=80) (%)	p value
Age (years)						
≤ 40	4 (10.81)	7 (7.14)	0.96	9 (11.25)	12 (15.00)	0.28
41–50	12 (24.32)	14 (28.57)		13 (16.25)	21 (26.50)	
> 51	24 (64.86)	19 (64.28)		57 (71.25)	47 (58.75)	
Sex						
Male	27 (72.9)	10 (71.42)	0.96	60 (75.00)	57 (71.25)	0.87
Female	13 (27.1)	30 (28.57)		20 (25.00)	23 (28.75)	
Residence						
Rural	24 (62.16)	25 (64.28)	0.99	58 (72.50)	51 (63.75)	0.30
Urban	16 (37.83)	15 (35.71)		22 (27.50)	29 (36.25)	
Histology types			—			—
SCLC	7 (16.21)	NA		12 (15.00)	NA	
AD	22 (56.75)	NA		40 (50.00)	NA	
SCC	11 (27.02)	NA		28 (35.00)	NA	
Stage*			—			—
II	2 (2.7)	NA		4 (5.00)	NA	
III	10 (16.21)	NA		10 (12.50)	NA	
IV	28 (81.08)	NA		66 (82.50)	NA	
Smoking status						
Non-smokers	12 (29.72)	11 (14.28)	0.03	30 (37.50)	64 (80.00)	< 0.0001
Ex-smokers	20 (51.35)	16 (50.00)		39 (48.75)	2 (2.50)	
Smokers	8 (18.91)	13 (35.71)		11 (13.75)	14 (17.50)	

*****Data are presented as *n (%)*. In the discovery set, characteristics are compared between tumor tissue (**TT**, N=40) and tumor-free tissue (**TF**, N=40). In the validation set, characteristics are compared between lung cancer **cases** (N=80) and **controls** (N=80). *P* values were calculated using the χ^2 test (or Fisher's exact test where appropriate) for categorical variables. Histology and tumor stage are reported only for lung cancer cases; therefore, *P* values were not computed for these variables. Statistically significant *P* values (<0.05) are highlighted in the manuscript text.

3.2 Promoter methylation frequency (%) of MAX, MTURN, HLA-B, and CAV1 in tumor tissue

In the overall lung cancer cohort, promoter methylation frequencies of MAX, MTURN, HLA-B, and CAV1 in TT were 68.45%, 56.94%, 75.97%, and 29.72%, respectively, which were significantly higher than those observed in TF (20.42%, 15.28%, 29.57%, and 18.42%; $p \leq 0.05$–0.01; (Fig. 2a)

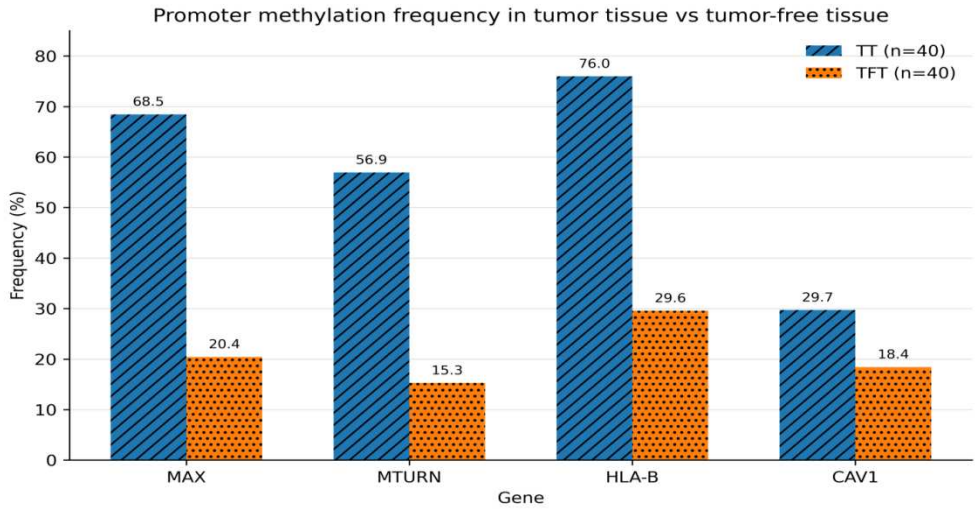


Fig. 2a Promoter methylation frequency (%) of MAX, MTURN, HLA-B, and CAV1 in tumor tissue (TT, n=40) compared with tumor-free tissue (TFT, n=40). Bars represent methylation frequency (percentage of methylated samples) for each gene, with numeric values shown above the bars.

Tissue promoter methylation frequency (%) of MAX, MTURN, HLA-B, and CAV1 across study groups

Subgroup analysis showed that MAX, MTURN, and HLA-B were significantly hypermethylated in NSCLC tumor tissues (n = 31), with frequencies of 63.29% ($p \leq 0.001$), 50.38% ($p \leq 0.05$), and 79.19% ($p \leq 0.001$), and were also elevated in SCLC tumor tissues (n = 6), at 53%, 35.3%, and 65.7%, respectively (all $p \leq 0.05$), compared with TF (20.42%, 15.28%, and 29.57%). In contrast, CAV1 showed pronounced hypermethylation in SCLC TT (84.3%, $p \leq 0.001$), whereas NSCLC TT demonstrated a nonsignificant decrease (21.35%) relative to TF (18.42%) (Fig 2b).

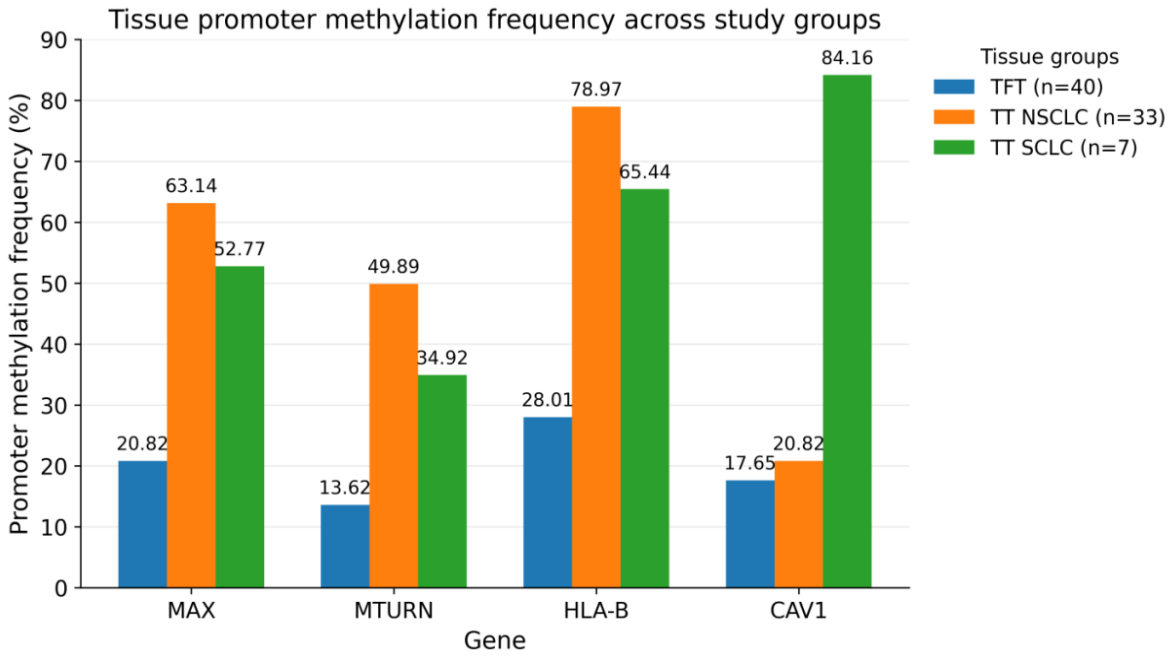


Fig. 2b Tissue promoter methylation frequency (%) of MAX, MTURN, HLA-B, and CAV1 across study groups. Promoter methylation frequencies are shown for tumor-free tissue (TFT, n=40), NSCLC tumor tissue (TT-NSCLC, n=33), and SCLC tumor tissue (TT-SCLC, n=7). Values are expressed as percentages, with exact frequencies displayed above each bar.

3.3 Differential Gene Expression Profile in NSCLC and SCLC Tissue Samples:

Gene-wise comparison of tissue samples showed clear differences in δct values among NSCLC (n=68), SCLC (n=12), and tumor-free tissue (TFT; n=80) (Fig. 3). Because lower δct reflects higher relative gene expression (after normalization to the reference gene), the reduced δct observed in tumor groups indicates increased expression in cancer tissues compared with TFT. For MAX, mean δct was highest in TFT (3.49 ± 0.13) and reduced in NSCLC (1.91 ± 0.18) and SCLC (0.80 ± 0.13), indicating greater MAX expression in tumors, most pronounced in SCLC. A similar pattern was observed for MTURN, with TFT showing the highest δct (5.89 ± 0.16) compared with NSCLC (1.67 ± 0.20) and SCLC (1.09 ± 0.20). HLA-B demonstrated the largest separation between groups, with δct markedly higher in TFT (6.82 ± 0.16) than in NSCLC (1.93 ± 0.18) and SCLC (0.95 ± 0.16). For CAV1, TFT (2.64 ± 0.13) and NSCLC (2.36 ± 0.18) showed comparable δct values, whereas SCLC displayed substantially lower δct (0.73 ± 0.16), suggesting relatively higher CAV1 expression in SCLC. Overall, SCLC consistently exhibited

the lowest δct values across all genes, supporting comparatively higher expression in SCLC relative to NSCLC and TFT is shown in Fig. 3

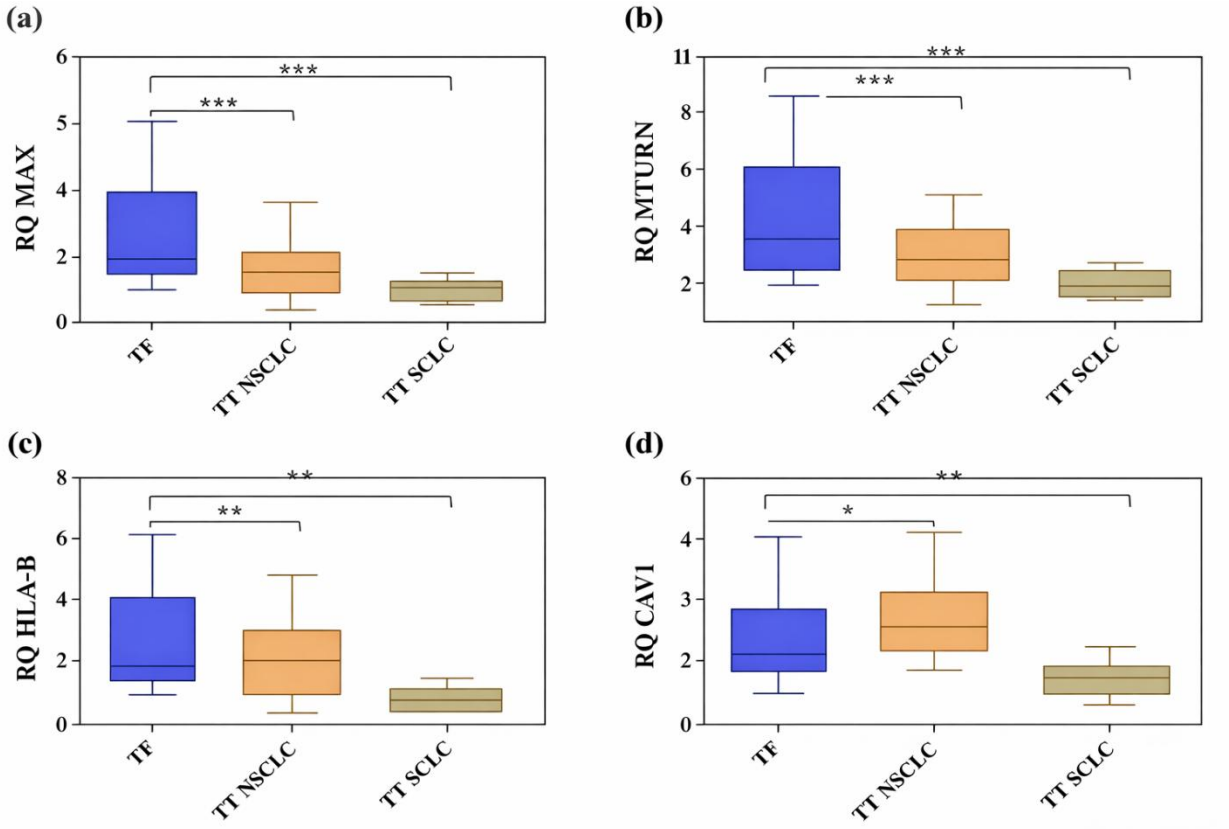


Fig. 3 Comparative relative expression of candidate genes in tumor-free tissue and lung cancer tissue subtypes. Box-and-whisker plots show the relative expression (RQ) of (a) MAX, (b) MTURN, (c) HLA-B, and (d) CAV1 in tumor-free tissue (TF, n = 40), NSCLC tumor tissue (TT-NSCLC, n = 33), and SCLC tumor tissue (TT-SCLC, n = 7). Horizontal brackets indicate statistically significant pairwise comparisons between groups.

Footnotes

Data are presented as box-and-whisker plots, with the central line representing the median, the box indicating the interquartile range (IQR), and the whiskers representing the data range. Statistical analysis was performed using the Kruskal-Wallis test followed by an appropriate post hoc multiple-comparison test.* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Abbreviations: RQ, relative quantification; **TF**, tumor-free tissue; **TT**, tumor tissue; **NSCLC**, non-small cell lung cancer; **SCLC**, small cell lung cancer.

3.4 ROC curve analysis of tissue-based gene expression biomarkers for distinguishing SCLC and NSCLC:

Table 2 indicates that among the evaluated genes, MAX had the best overall diagnostic performance, especially in SCLC, where it achieved the highest AUC and perfect specificity. MTURN and CAV1 also performed reasonably well in SCLC, while HLA-B showed comparatively modest utility. In NSCLC, most markers showed only moderate discrimination, with high specificity but limited sensitivity, suggesting better rule-in than screening performance. CAV1 in NSCLC was highly sensitive but lacked specificity, reducing its reliability as a standalone marker. ROC curve analysis showed that the diagnostic performance of tissue-based gene expression biomarkers differed between SCLC and NSCLC. In SCLC, MAX demonstrated the highest accuracy (AUC = 0.90) with 58% sensitivity and 100% specificity at an optimal cut-off of 0.93 ($p < 0.0001$). MTURN and CAV1 also showed good discriminatory ability (both AUC = 0.80), whereas HLA-B showed moderate performance (AUC = 0.70). In NSCLC, MAX, MTURN, and HLA-B each showed moderate diagnostic accuracy (AUC = 0.70), with high specificity (97–100%) but relatively low sensitivity (25–50%). CAV1 showed lower diagnostic accuracy in NSCLC (AUC = 0.60), despite high sensitivity (84%) and poor specificity (14%). Overall, MAX emerged as the most effective biomarker, particularly for SCLC.

Table 2 Sensitivity and specificity of mRNA expression data in tumor tissue (TT) versus tumor-free tissue (TF)

	SCLC					NSCLC				
Gene	AUC	Sensitivity	Specificity	P value	Optimal Cut off	AUC	Sensitivity	Specificity	P value	Optimal Cut off
MAX	0.9	58	100	<0.0001	0.93	0.7	50	100	<0.0001	0.98
MTURN	0.8	41	100	<0.0001	0.94	0.7	28	97	0.0001	1.0
HLA-B	0.7	33	97	0.0007	0.5	0.7	25	100	0.002	0.52
CAV1	0.8	50	97	0.009	0.6	0.6	84	14	<0.0001	0.5

Receiver operating characteristic (ROC) analysis was performed to evaluate the ability of mRNA expression of MAX, MTURN, HLA-B, and CAV1 in tumor tissue (TT) to discriminate TT from matched tumor-free tissue (TF), analyzed separately for small cell lung cancer (SCLC) and non–

small cell lung cancer (NSCLC). The area under the ROC curve (AUC), sensitivity, specificity, p value, and optimal cut-off value are reported for each gene.

TT, tumor tissue; **TF**, tumor-free tissue, **SCLC**, small cell lung cancer; **NSCLC**, non–small cell lung cancer, **AUC**, area under the ROC curve (higher AUC indicates better discrimination), **Sensitivity (%)** represents the proportion of TT samples correctly classified as tumor, **Specificity (%)** represents the proportion of TF samples correctly classified as non-tumor, **Optimal cut-off** indicates the threshold value that maximized discrimination (typically based on the Youden index, unless otherwise specified), **P values** correspond to the statistical significance of the ROC analysis versus the null hypothesis (AUC = 0.5).

3.5 Correlation between Gene Expression and Promoter Methylation Status in Tumor Tissue:

Spearman's rank correlation analysis demonstrated a significant inverse relationship between tumor tissue–derived promoter methylation and mRNA expression for all studied genes in both NSCLC and SCLC, indicating that increased promoter methylation was associated with reduced transcript levels. In NSCLC, strong negative correlations were observed for HLA-B ($r = -0.8583$, $p < 0.0001$) and MAX ($r = -0.7439$, $p < 0.0001$), while CAV1 ($r = -0.5899$, $p = 0.0005$) and MTURN ($r = -0.5521$, $p = 0.0004$) showed moderate but statistically significant inverse associations. In SCLC, the inverse correlations were consistently stronger, with HLA-B showing the highest correlation ($r = -0.8690$, $p < 0.0001$), followed by MAX ($r = -0.8569$, $p = 0.0004$) and CAV1 ($r = -0.8569$, $p = 0.0004$). MTURN also demonstrated a strong negative correlation ($r = -0.8208$, $p = 0.001$).Overall, these findings support that promoter hypermethylation is closely linked to transcriptional downregulation of MAX, MTURN, HLA-B, and CAV1 in lung cancer tumor tissues, with more pronounced effects in SCLC compared with NSCLC. Fig. 4a-d

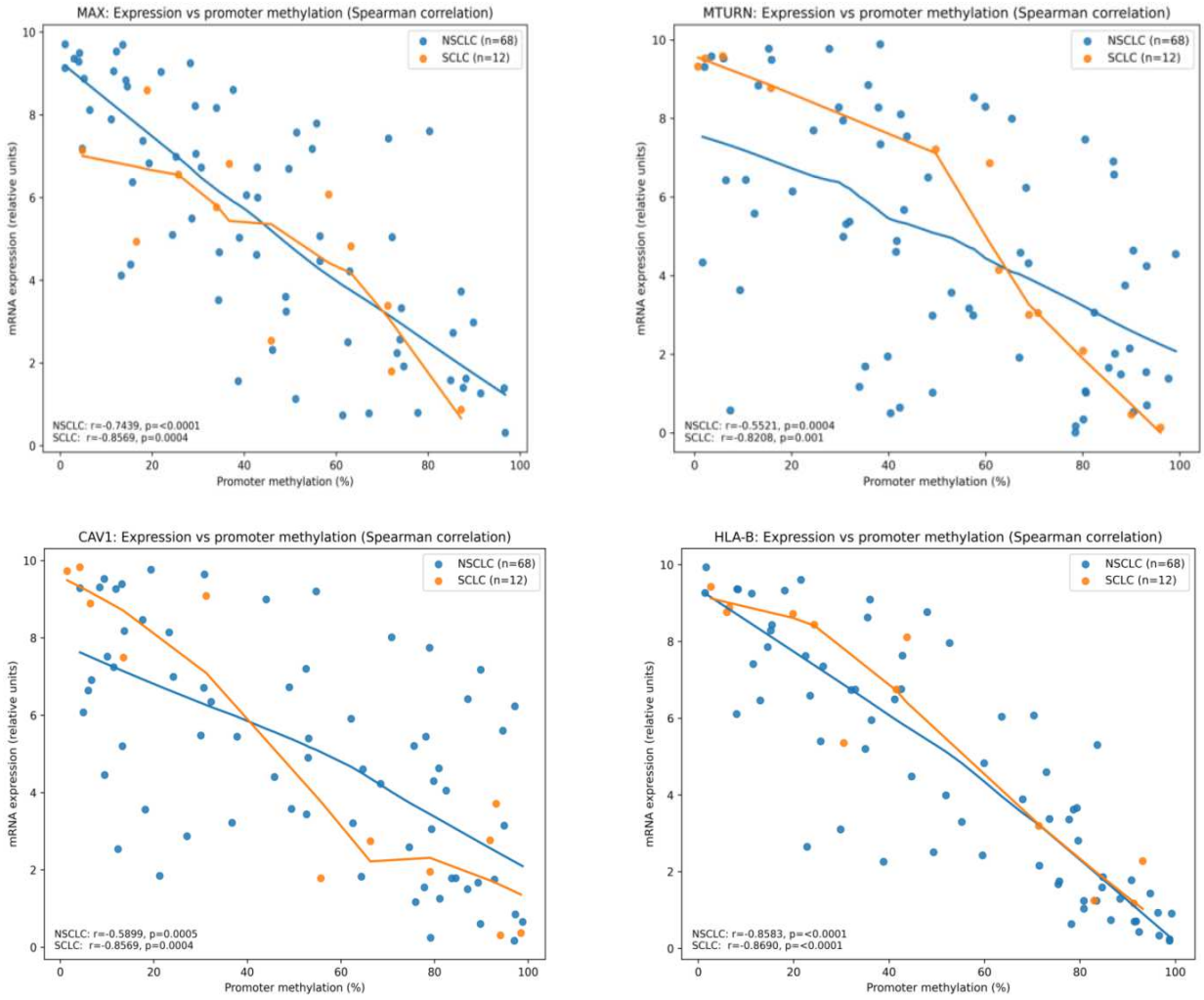


Fig. 4a-d Correlation between gene expression and promoter methylation in tissue-based samples. Spearman correlation analysis was performed to evaluate the association between gene expression levels and promoter methylation status in tumor tissue samples for MAX, MTURN, HLA-B, and CAV1. Correlation plots for NSCLC are shown in panels (a-b), whereas those for SCLC are shown in panels (c-d), respectively.

Footnotes

The correlation between gene expression and promoter methylation was assessed using the **Spearman rank correlation test**. The value of **r** represents the **Spearman correlation coefficient**, indicating the direction and strength of the association between the two variables.

Abbreviations: NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; r, Spearman correlation

3.6 Diagnostic Utility of Tissue-Based Gene Expression Biomarkers:

ROC Curve Analysis: Receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic performance of MAX, MTURN, HLA-B, and CAV1 in distinguishing non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) patients.

In NSCLC, MTURN exhibited the highest diagnostic accuracy with an AUC of 0.7831 (95% CI: 0.688–0.868), followed by HLA-B (AUC = 0.7323, 95% CI: 0.618–0.885) and MAX (AUC = 0.7248, 95% CI: 0.648–0.910). CAV1 showed comparatively lower performance with an AUC of 0.6260 (95% CI: 0.632–0.871).Fig. 5b. In SCLC, all four biomarkers demonstrated improved diagnostic efficiency. CAV1 showed the highest accuracy (AUC = 0.8924, 95% CI: 0.781–0.971), followed by MTURN (AUC = 0.8861, 95% CI: 0.781–0.967), MAX (AUC = 0.8681, 95% CI: 0.762–0.980), and HLA-B (AUC = 0.7963, 95% CI: 0.679–0.935). Fig. 5a. Overall, ROC analysis indicated that these biomarkers possess stronger discriminatory potential in SCLC compared to NSCLC. Among the evaluated markers, MTURN showed the best performance in NSCLC, whereas CAV1 and MTURN exhibited superior diagnostic utility in SCLC. These findings suggest that the selected genes may serve as promising diagnostic biomarkers, particularly for SCLC detection.

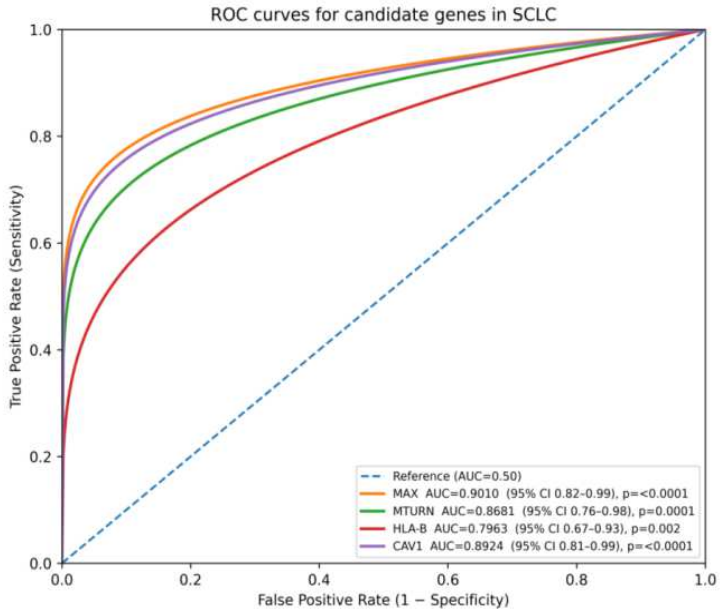


Fig. 5a Receiver operating characteristic (ROC) curves showing diagnostic performance of candidate gene markers in SCLC. ROC curves for MAX, MTURN, HLA-B, and CAV1 were generated to discriminate SCLC cases from controls. The dashed diagonal line represents chance-level classification (AUC = 0.50).

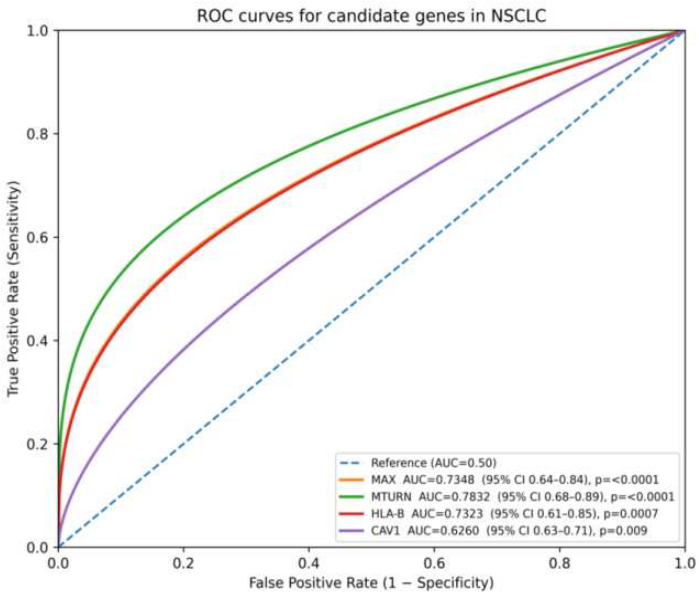


Fig. 5b Receiver operating characteristic (ROC) curves showing diagnostic performance of candidate gene markers in NSCLC. ROC analyses were performed for MAX, MTURN, HLA-B, and CAV1 to discriminate NSCLC cases from controls. The dashed diagonal line represents chance-level classification (AUC = 0.50).

3.7 Plasma-Based cfDNA Extraction and Promoter Methylation Analysis of Candidate Genes:

The concentration of plasma extracted cfDNA was found in the range of 2.87–9.29 ng/µL in healthy control and 13.62–63.04 ng/µL in LC patients. The MSP productsof all genes were run on 3% agarose gel electrophoresis(The methylation frequency in plasma samples in overall LC for genes MAX, MTURN, HLA-B, and CAV1 were observed 43.5, 23.0, 61.7 and 19.5%, a significantly hypermethylated ($p \leq 0.01$) (except CAV1) as compared to healthy control (12.75%, 8.25%, 15.75%, and 10.5%),respectively (Fig. 6). The plasma-derived cfDNA also showed a significant hypermethylation MAX, MTURN and HLA-B, such as 43.64% ($p \leq 0.001$), 29.11% ($p \leq 0.05$)and 62.29% ($p \leq 0.001$) in NSCLC (n = 68) and 42.66%($p \leq 0.05$), 27% ($p \leq 0.05$), and 50% ($p \leq 0.05$) in SCLC(n = 12) as compared to healthy control (n = 80) i.e.12.75, 8.25, & 21.75% respectively. However, for CAV1,a nonsignificant hypomethylation in NSCLC (8.8%) and significant ($p \leq 0.001$) hypermethylation in SCLC (66.66%)were observed as compared to healthy control (12.5%) (Fig. 6).

Fig. 6 Plasma promoter methylation frequency (%) of MAX, MTURN, HLA-B, and CAV1 across study groups. Promoter methylation frequencies are shown for NSCLC (n = 68), SCLC (n = 12), and healthy controls (n = 80). Bars represent the percentage of methylated samples for each gene, with exact frequencies displayed above the bars.

3.8 Association of Plasma-Based Promoter Methylation Biomarkers with Smoking Status:

In plasma-derived cfDNA, promoter methylation was evaluated across smoking strata. Overall, methylation frequencies for the assessed genes increased with tobacco exposure—current and former smokers showed higher percentages than never-smokers ($p \leq 0.05$), indicating a positive association between smoking status and lung cancer–related methylation (Fig. 7). Among individual markers, HLA-B exhibited the strongest smoking-related effect ($p \leq 0.001$), whereas CAV1 showed no significant difference in ex-smokers compared with non-smokers

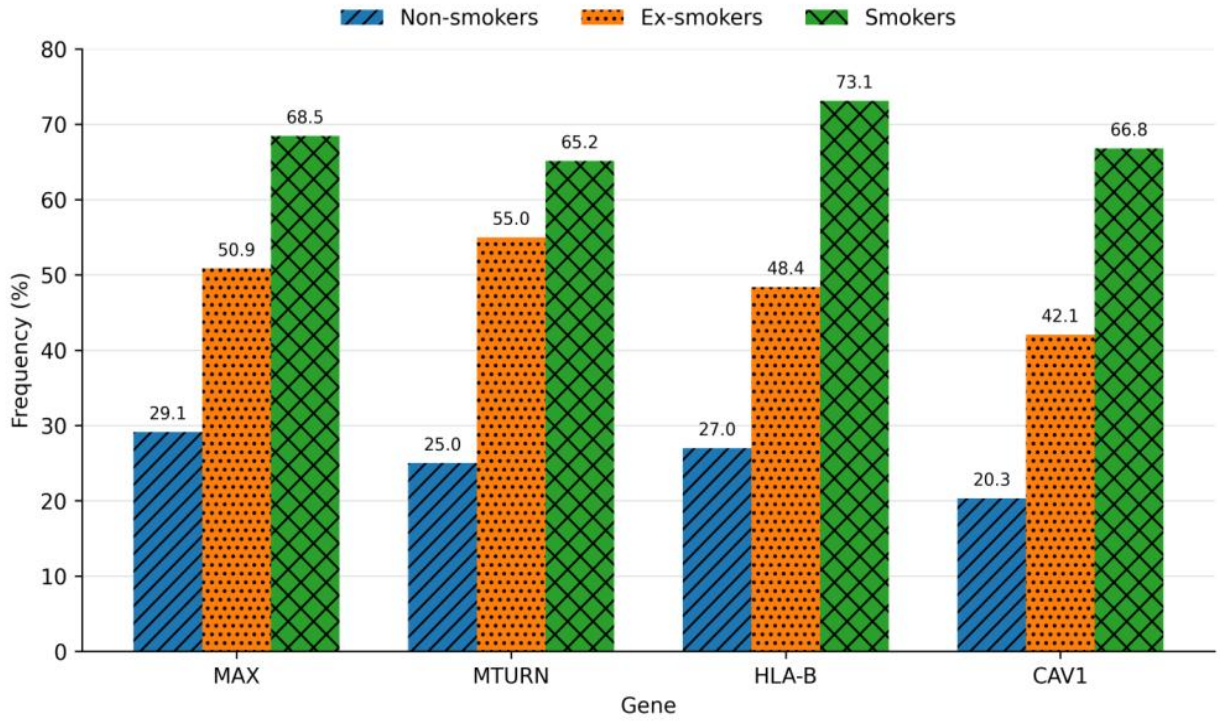


Fig. 7 Smoking status-wise promoter methylation frequency (%) of MAX, MTURN, HLA-B, and CAV1. The bar chart shows the frequency of promoter methylation across non-smokers, ex-smokers, and current smokers for each gene. Values are presented as percentages, with exact frequencies displayed above the bars.

3.9 Association of Plasma Promoter Methylation Patterns with Disease Stage and TNM Classification:

Across 80 lung-cancer plasma samples, promoter methylation was stratified by clinical stage and TNM category (Table 3). By stage, frequencies were: Stage II (n=4) — MAX 25%, MTURN 25%, HLA-B 50%, CAV1 0%; Stage III (n=10) — MAX 40%, MTURN 20%, HLA-B 40%, CAV1 20%; Stage IV (n=66) — MAX 43.9%, MTURN 21.2%, HLA-B 62.9%, CAV1 18.1%. TNM partitions (T1/T2 vs T3/T4; N0/N1 vs N2/N3) showed broadly similar distributions across genes. Aberrant methylation was detectable in earlier categories (T1/T2, N0/N1) and was higher with disease advancement (notably at T3/T4 and N0/N1 as per dataset annotations). Metastatic status mirrored this pattern: compared with M0 (n=14), M1 (n=66) exhibited increased plasma methylation for MAX (35.7%→43.1%) and HLA-B (42.8%→61.1%), with more modest rises for MTURN (14%→21%) and CAV1 (14%→18%). Overall, methylation emerges at early stages and intensifies with progression is shown in Table 3

Table 3 Association of plasma-derived cfDNA promoter hypermethylation with clinical stage and TNM parameters in patients with lung cancer.

Variables	Lung Cancer (LC) N=80								
	n=80	MAX N'= 34 (%)		MTURN N'= 16 (%)		HLA-B N'= 47 (%)		CAV1 N'= 14 (%)	
		No. of pos. of Meth	%	No. of pos. of Meth	%	No. of pos. of Meth	%	No. of pos. of Meth	%
Stage									
II	4	1	25	0	0	2	50	0	0
III	10	4	40	2	20	4	40	2	20
IV	66	29	43.9	14	21.2	41	62.1	12	18.1
TNM Grade									
T									
T1/T2	9	3	33.3	2	22.2	3	33.3	1	11.1
T3/T4	71	31	43.6	14	19.7	44	61.9	13	18.3
N									
N0/N1	7	2	28.5	1	14.2	2	28.5	1	14.2
N2/N3	73	32	45.0	15	20.5	45	61.4	13	17.8
M									
M0	14	5	35.7	2	14.2	6	42.8	2	14.2
M1	66	29	43.9	14	21.2	41	62.1	12	18.1

The frequency of promoter hypermethylation of MAX, MTURN, HLA-B, and CAV1 detected in plasma-derived cell-free DNA (cfDNA) is presented across clinical stage and TNM categories in lung cancer (LC)

patients. For each subgroup, the number of methylation-positive cases and the corresponding percentage are shown.

LC, lung cancer; **cfDNA**, cell-free DNA, **N = 80**, total number of LC patients included in the plasma cohort, **N′**, total number of methylation-positive cases for the respective gene in the overall cohort, **No. of pos. of Meth**, number of methylation-positive cases in the indicated subgroup; **%**, percentage within that subgroup, **T**, primary tumor category (**T1/T2** vs **T3/T4**); **N**, nodal involvement (**N0/N1** vs **N2/N3**); **M**, distant metastasis (**M0** vs **M1**), Staging was performed according to the **8th edition TNM classification system**, Percentages were calculated using the subgroup total as the denominator; minor variation may occur due to rounding.

3.10 Diagnostic Performance of Promoter Methylation Biomarkers in Tissue and Plasma Samples:

In Table 4, tissue samples (40 TT cases and 40 TFT controls), MAX positivity was observed in 22/40 cases and 8/40 controls, showing a statistically significant association with disease status ($p = 0.02$). MAX demonstrated an odds ratio (OR) of 5.6 (95% CI: 1–29), with a sensitivity of 59.46% (95% CI: 42–75), specificity of 78.57% (95% CI: 49–95), positive predictive value (PPV) of 88.00% (95% CI: 72–95), and negative predictive value (NPV) of 42.71% (95% CI: 31–54). Similarly, MTURN positivity in tissue was detected in 17/40 cases and 10/40 controls and was also significantly associated with disease ($p = 0.05$), with an OR of 5.1 (95% CI: 1–26). MTURN showed a sensitivity of 45.95% (95% CI: 29–63), specificity of 85.71% (95% CI: 57–98), PPV of 89.47% (95% CI: 69–97), and NPV of 37.50% (95% CI: 29–46). Among the tissue-based markers, HLA-B showed the strongest association, with positivity in 27/40 cases compared with 11/40 controls ($p = 0.006$), yielding an OR of 6.7 (95% CI: 2–26). HLA-B demonstrated a sensitivity of 62.96% (95% CI: 42–80), specificity of 71.43% (95% CI: 42–91), PPV of 80.95% (95% CI: 64–91), and NPV of 50.00% (95% CI: 35–64). In contrast, CAV1 positivity was detected in 11/40 cases and 9/40 controls and did not differ significantly between groups ($p = 0.5$). Its OR was 1.5 (95% CI: 0.3–6.6), with low sensitivity (29.73%, 95% CI: 15–47) despite moderate specificity (78.57%, 95% CI: 49–95), PPV (78.57%, 95% CI: 54–91), and NPV (29.73%, 95% CI: 23–37).

In plasma samples (80 cases and 80 controls), MAX positivity was present in 34/80 cases and 11/80 controls, showing a highly significant association with disease ($p = 0.0001$). The OR for

MAX was 4.6 (95% CI: 2.1–10), with a sensitivity of 48.75% (95% CI: 37–60), specificity of 85.00% (95% CI: 75–92), PPV of 76.47% (95% CI: 65–85), and NPV of 62.39% (95% CI: 6–68). MTURN positivity in plasma was identified in 16/80 cases and 5/80 controls, also showing a significant association ($p = 0.01$) with an OR of 3.7 (95% CI: 1.3–11). However, MTURN exhibited relatively low sensitivity (20.00%, 95% CI: 12–30) but high specificity (93.75%, 95% CI: 86–98), with PPV of 76.19% (95% CI: 55–89) and NPV of 53.96% (95% CI: 51–57).

Among plasma-based markers, HLA-B again showed the best discriminatory performance. Positivity was observed in 47/80 cases and 15/80 controls, with a highly significant difference ($p < 0.0001$) and an OR of 6.1 (95% CI: 3–12). HLA-B demonstrated a sensitivity of 58.75% (95% CI: 47–70), specificity of 81.25% (95% CI: 71–89), PPV of 75.81% (95% CI: 65–84), and NPV of 66.33% (95% CI: 60–72). In contrast, CAV1 positivity was found in 14/80 cases and 10/80 controls and was not significantly associated with disease ($p = 0.3$), with an OR of 1.4 (95% CI: 0.6–3.5). CAV1 showed poor sensitivity (17.50%, 95% CI: 10–27) despite relatively high specificity (87.50%, 95% CI: 78–93), with PPV of 58.33% (95% CI: 39–74) and NPV of 51.47% (95% CI: 48–54).

Overall, MAX, MTURN, and HLA-B demonstrated significant diagnostic associations in both tissue and plasma, whereas CAV1 did not show significant discriminatory value in either sample type. Among all markers evaluated, HLA-B exhibited the strongest and most consistent diagnostic performance across both tissue and plasma specimens.

The diagnostic utility of MAX, MTURN, HLA-B, and CAV1 was evaluated in both tissue and plasma samples by comparing cases and controls.

Table 4 Diagnostic performance of promoter methylation of MAX, MTURN, HLA-B, and CAV1 in tissue DNA and plasma cfDNA for discriminating lung cancer cases from controls.

Gene	Sample	Cases (N/M)	Controls (N/M)	p value	Odds Ratio (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
MAX	Tissue	40/22	40/8	0.02	5.6 (1 to 29)	59.46 (42 to 75)	78.57 (49 to 95)	88.00 (72 to 95)	42.71 (31 to 54)
MAX	Plasma	80/34	80/11	0.0001	4.6 (2.1 to 10)	48.75 (37 to 60)	85.00 (75 to 92)	76.47 (65 to 85)	62.39 (6 to 68)
MTURN	Tissue	40/17	40/10	0.05	5.1 (1 to 26)	45.95 (29 to 63)	85.71 (57 to 98)	89.47 (69 to 97)	37.50 (29 to 46)
MTURN	Plasma	80/16	80/5	0.01	3.7 (1.3 to 11)	20.00 (12 to 30)	93.75 (86 to 98)	76.19 (55 to 89)	53.96 (51 to 57)
HLA-B	Tissue	40/27	40/11	0.006	6.7 (2 to 26)	62.96 (42 to 80)	71.43 (42 to 91)	80.95 (64 to 91)	50.00 (35 to 64)
HLA-B	Plasma	80/47	80/15	< 0.0001	6.1 (3 to 12)	58.75 (47 to 70)	81.25 (71 to 89)	75.81 (65 to 84)	66.33 (60 to 72)
CAV1	Tissue	40/11	40/9	0.5	1.5 (0.3 to 6.6)	29.73 (15 to 47)	78.57 (49 to 95)	78.57 (54 to 91)	29.73 (23 to 37)
CAV1	Plasma	80/14	80/10	0.3	1.4 (0.6 to 3.5)	17.50 (10 to 27)	87.50 (78 to 93)	58.33 (39 to 74)	51.47 (48 to 54)

For each gene, results are shown separately for **tissue** and **plasma**. "Cases (N/M)" and "Controls (N/M)" denote the total number of samples analyzed (**N**) and the number positive for promoter methylation (**M**). Odds ratios (OR) with 95% confidence intervals (CI) were calculated from 2×2 contingency tables. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are reported as percentages with 95% CIs. *P* values represent between-group comparisons of methylation positivity (χ^2 or Fisher's exact test, as appropriate). **Abbreviations:** cfDNA, cell-free DNA; OR, odds ratio; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

3.11 Concordance between tissue and plasma methylation findings for candidate biomarkers:

To evaluate concordance between tumor tissue (TT) and plasma-derived cfDNA promoter methylation status, agreement was assessed using Cohen's kappa (κ). Overall, the results indicated low-to-moderate concordance between tissue and cfDNA across the studied genes.MAX (κ = 0.367) and MTURN (κ = 0.378) demonstrated fair agreement, suggesting limited but detectable consistency between TT and plasma-derived cfDNA methylation calls. In contrast, HLA-B showed the highest concordance with moderate agreement (κ = 0.479), followed by CAV1 which also exhibited moderate agreement (κ = 0.418). as shown in Table 5. Collectively, these findings support that plasma-derived cfDNA methylation partially reflects tumor tissue methylation patterns, with HLA-B providing the most consistent TT–cfDNA agreement among the evaluated biomarkers.

Table 5 Agreement analysis of gene-specific biomarkers between tissue and plasma samples using Cohen's kappa statistic

S.No.	Gene	Cohen's kappa (κ)	Degree of Agreement
	MAX	0.367	Fair agreement
	MTURN	0.378	Fair agreement
	HLA-B	0.479	Moderate agreement
	CAV1	0.418	Moderate agreement

Cohen's kappa (κ) was used to assess the level of agreement for MAX, MTURN, HLA-B, and CAV1 between the two sample types. Based on κ values, MAX (κ = 0.367) and MTURN (κ = 0.378) showed fair agreement, whereas HLA-B (κ = 0.479) and CAV1 (κ = 0.418) demonstrated moderate agreement. These findings indicate variable but overall acceptable concordance between sample types for the evaluated biomarkers.

Discussion: The present study strengthens the growing evidence that plasma-derived cell-free DNA (cfDNA) methylation may serve as a minimally invasive molecular approach for lung cancer (LC) detection and may, in part, reduce dependence on invasive tissue biopsy. Aberrant promoter methylation is recognized as an early and biologically relevant event in tumorigenesis, and its detection in circulating nucleic acids has emerged as a promising strategy for cancer diagnostics (Belinsky 2004, 2005; Constâncio et al. 2020). Building on this framework, we first identified a tissue-based epigenetic signature by comparing tumor tissue (TT) with tumor-free tissue (TF), integrating promoter methylation status with gene-expression changes in genomic DNA. We subsequently evaluated whether these methylation alterations could also be detected in plasma cfDNA from patients with LC, including both NSCLC and SCLC, to determine whether plasma could function as a reliable surrogate for tumor tissue.

One of the earliest and most notable observations of the study was the markedly increased concentration of cfDNA in patient plasma compared with controls. The nearly tenfold elevation in circulating DNA likely reflects fundamental features of tumor biology, including enhanced apoptosis, necrosis, and active release of nucleic acid fragments by tumor cells and circulating tumor cells into the bloodstream (Moss et al. 2018). This finding further supports the clinical relevance of plasma as a suitable biospecimen for molecular profiling in LC and is consistent with reports showing higher circulating DNA levels in patients with lung cancer than in healthy individuals or those with inflammatory lung disease (Yoon et al. 2009; Szpechcinski et al. 2015).

The internal validity of the molecular findings is reinforced by the broadly comparable distribution of major demographic variables across the discovery and validation cohorts. Age, sex, and residence were similar between groups, suggesting that the observed methylation differences are unlikely to be attributable to major baseline demographic imbalance. By contrast, smoking status differed significantly, particularly in the validation cohort, underscoring the contribution of tobacco exposure to LC-associated epigenetic remodeling. This observation is biologically plausible, as chronic exposure to tobacco carcinogens is known to induce widespread epigenetic alterations, including promoter CpG methylation (Belinsky et al. 2002; Belinsky 2004, 2005). In addition, the predominance of stage IV disease in both cohorts reflects the real-world clinical presentation of LC in many settings and demonstrates the robustness of biomarker detection in clinically established disease. However, this stage distribution also indicates that further validation in genuinely early-stage populations will be necessary before extrapolating these findings to screening applications.

A major strength of this study is the identification of a four-gene panel comprising MAX, MTURN, HLA-B, and CAV1, genes with recognized but distinct roles in cancer biology that have rarely been evaluated together as an integrated epigenetic signature in LC. In tissue, MAX, MTURN, and HLA-B were significantly downregulated in TT relative to TF ($p \leq 0.01$), and this reduced expression was inversely associated with promoter methylation, supporting a methylation-mediated mechanism of transcriptional repression. CAV1, by contrast, displayed a more complex pattern, with significant downregulation in SCLC-TT ($p = 0.0002$) but modest upregulation in NSCLC-TT ($p = 0.03$), suggesting subtype-dependent biological behavior. Receiver-operating characteristic (ROC) analysis further demonstrated that this panel possesses meaningful tissue-level discriminatory ability, indicating that these loci are not only biologically relevant but also diagnostically informative at the tumor source.

These findings are concordant with the wider literature. Dysregulation or inactivation of MAX has been reported in malignancies, including lung cancer (Romero et al. 2014; Wang et al. 2017), while promoter methylation-mediated silencing of HLA class I genes has been described in digestive tract cancers (Nie et al. 2001; Ye et al. 2010). Likewise, MTURN has been implicated in lung cancer-related biomarker studies (Liu et al. 2019; Ying et al. 2022). The findings for CAV1 are also consistent with its recognized context-dependent role in carcinogenesis.

Depending on tumor type, histologic subtype, and microenvironmental context, CAV1 may function either as a tumor suppressor or as an oncogene (Sunaga et al. 2004; Guo et al. 2016). Reduced CAV1 expression and promoter methylation have been reported in lung and other malignancies (Sunaga et al. 2004; Brodie et al. 2014), whereas context-dependent overexpression has also been described in several tumor types. Taken together, the present results support the view that MAX, MTURN, and HLA-B predominantly behave as tumor suppressor-associated loci in LC and that their silencing is at least partly mediated through promoter hypermethylation, whereas CAV1 appears to reflect subtype-specific epigenetic and transcriptional dynamics.

Importantly, the plasma cfDNA methylation findings closely mirrored the tissue observations. Compared with controls, promoter methylation frequencies of MAX, MTURN, and HLA-B were significantly increased in patient plasma. In NSCLC, methylation frequencies were 43.64%, 29.11%, and 62.29%, respectively, whereas in SCLC the corresponding values were 43.66%, 25.0%, and 49.0% (all $p \leq 0.05$ to $p \leq 0.01$), compared with control values of 12.75%, 8.25%, and 15.75%. These frequencies broadly paralleled the hypermethylation patterns observed in tissue. Agreement analysis using Cohen's κ demonstrated fair-to-moderate concordance between tissue and plasma for all genes, and the absence of significant differences in overall methylation-frequency distributions between the two matrices further supports the concept that plasma cfDNA captures biologically relevant tumor-associated epigenetic changes. Collectively, these findings support a "tissue-mirror" model in which plasma cfDNA can serve as a surrogate representation of tumor methylation status (Weiss et al. 2017; Liu et al. 2020; Wei et al. 2021). This interpretation is further strengthened by prior reports describing tumor-associated methylation signals in cfDNA across diverse malignancies (Constâncio et al. 2020; Machado et al. 2025).

From a translational perspective, HLA-B emerged as the most informative individual plasma marker, showing sensitivity of approximately 58.75%, specificity of approximately 81.25%, and favorable predictive values. MAX and MTURN also performed well in cfDNA, particularly in terms of specificity and positive predictive value, with PPV values of at least 75%, NPV values of approximately 52–63%, and specificity ranging from approximately 85% to 93.75%. These metrics compare favorably with earlier methylation-panel studies reporting good diagnostic performance for plasma-based assays and, in some settings, predictive capacity before clinical

diagnosis (Hulbert et al. 2017; Weiss et al. 2017; Shi et al. 2020; Hu et al. 2023). The reproducibility of signals across tissue and plasma, and across multiple loci, reinforces a key principle of liquid-biopsy diagnostics: multi-marker approaches are generally more robust than single-marker assays because they reduce the effects of tumor heterogeneity, subtype variation, and pre-analytical noise. Although CAV1 is less suitable as a universal standalone marker because of its biphasic behavior, its subtype-dependent pattern may still add biological and diagnostic value when incorporated into multi-gene algorithms, particularly for histologic stratification or risk refinement.

The study also offers insight into potential sources of discordance between tissue and plasma findings, as well as apparent false-positive and false-negative results. Tobacco exposure and occupational or environmental insults are plausible contributors to aberrant methylation through repeated epithelial injury, chronic inflammation, and enhanced release of fragmented nucleic acids into circulation. In the present cohort, smokers and ex-smokers showed higher methylation frequencies than never-smokers for most genes ($p \leq 0.05$), consistent with earlier evidence linking tobacco smoke and environmental pollutants to promoter hypermethylation of tumor suppressor genes (Belinsky et al. 2002; Belinsky 2004, 2005). Conversely, false-negative plasma findings may arise from a low circulating tumor fraction, limited tumor shedding, or technical and pre-analytical variability in cfDNA recovery and bisulfite conversion (Moss et al. 2018; Jamshidi et al. 2022). These considerations have practical implications for clinical translation, suggesting that cfDNA methylation results should be interpreted in the context of exposure history and that optimal diagnostic performance will likely depend on multiplexed biomarker panels rather than isolated single-gene readouts.

Stratification according to clinical stage and TNM classification yielded additional biological insight. Aberrant methylation of MAX, MTURN, and HLA-B was detectable even in earlier disease categories, with a progressive increase in frequency across more advanced stages, particularly stage IV and M1 disease. This stage-related gradient suggests that promoter methylation is not merely an incidental event, but rather a biologically progressive phenomenon that intensifies with tumor burden, metastatic spread, and clonal evolution. Previous studies have similarly linked cfDNA methylation patterns with tumor progression, disease burden, and diagnostic discrimination in LC (Weiss et al. 2017; Wei et al. 2021; Hu et al. 2023). For

example, Hu et al. (2023) demonstrated the feasibility of plasma methylation analysis for early-stage lung cancer detection. Although some of the present stage-specific subgroups were relatively small, the directional consistency across analyses supports the potential utility of cfDNA methylation not only for diagnostic discrimination but also for staging support and disease monitoring.

A central message emerging from both the present findings and the existing literature is the importance of multi-gene epigenetic panels in LC. Tumor heterogeneity, subtype-specific molecular behavior, and background methylation related to smoking or environmental exposure all argue against overreliance on any single biomarker. In this context, our results suggest that a three-gene core panel consisting of MAX, MTURN, and HLA-B provides a stable and reproducible signal in both tissue and plasma, while CAV1 offers complementary contextual information and may improve subtype-level interpretation. From an assay-development perspective, this raises the practical possibility of prioritizing the three-gene core for screening or triage strategies, with the addition of CAV1 or other loci in modular fashion as further evidence accumulates.

The study has several notable strengths. First, the matched tissue–plasma design allowed direct comparison between tumor-derived and circulating methylation patterns. Second, the integration of promoter methylation with gene-expression data strengthened the biological interpretation of the findings. Third, the reproducibility of methylation signals across histologic subtypes and sample types supports the robustness of the proposed panel. The fair-to-moderate tissue–plasma concordance observed by Cohen's κ, together with the non-significant differences in methylation-frequency distributions between the two matrices, suggests a biologically meaningful correspondence rather than a purely statistical association.

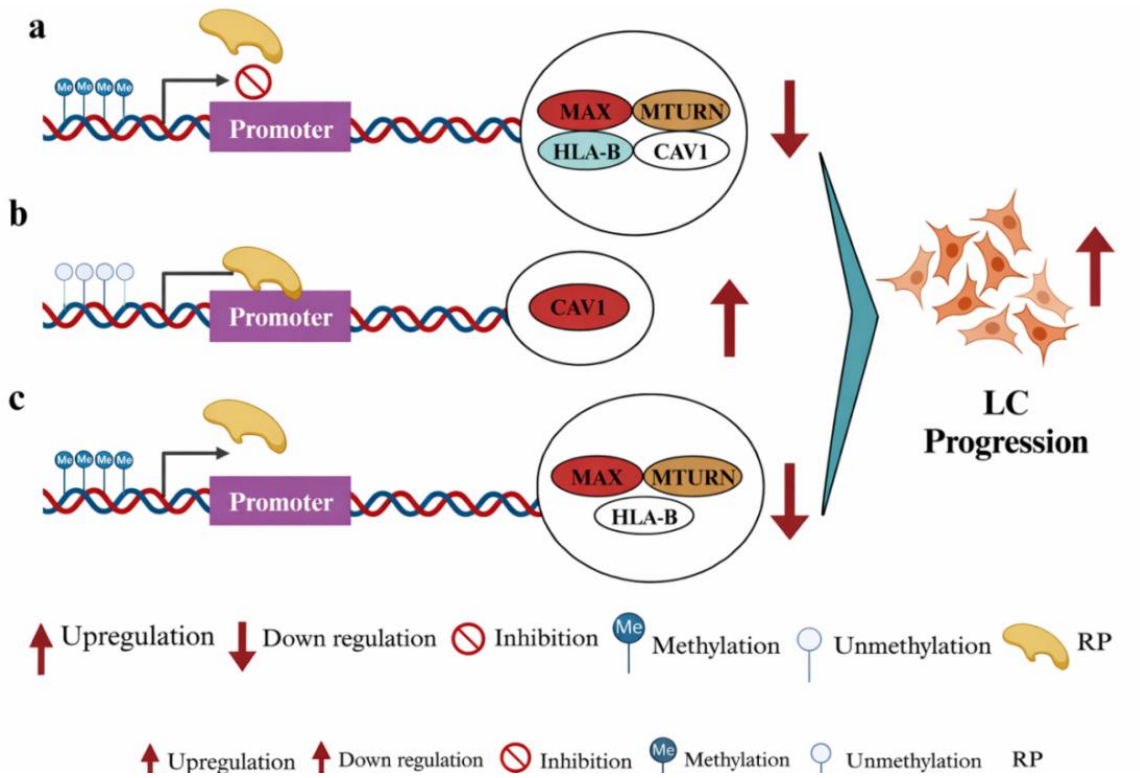
Nevertheless, certain limitations should be acknowledged. Some subgroup sample sizes, particularly for SCLC and early-stage disease, were modest and may have limited statistical precision. In addition, methylation-specific PCR (MSP), although sensitive and cost-effective, is fundamentally locus-specific and largely qualitative or semi-quantitative. Future studies using more quantitative platforms, such as digital PCR or targeted bisulfite sequencing, may allow better threshold refinement and clinical calibration. Finally, despite consideration of exposure-

related variables and comorbid factors, residual environmental or biological confounding cannot be fully excluded.

Overall, the present study demonstrates that promoter methylation of MAX, MTURN, and HLA-B in plasma-derived cfDNA closely reflects tumor-tissue methylation and shows substantial promise as a diagnostic strategy for LC, whereas CAV1 exhibits subtype-dependent behavior consistent with its dual biological role. The high specificity and positive predictive value observed for MAX and MTURN, together with the more balanced sensitivity and specificity of HLA-B, compare favorably with previous reports and support the value of panel-based methylation approaches. The detection of aberrant methylation in earlier-stage disease, alongside its increased frequency in advanced disease, further indicates potential utility in both early detection and longitudinal disease assessment (Fig. 8).

In conclusion, the findings support plasma cfDNA methylation profiling as a practical “tissue-mirror” strategy for LC detection and risk stratification. Large-scale multicenter validation in adequately powered cohorts, with harmonized pre-analytical workflows and integration into composite diagnostic algorithms, now represents the critical next step. Such algorithms should ideally combine epigenetic markers with complementary molecular and clinical variables to improve sensitivity, specificity, and clinical interpretability. With continued validation, cfDNA methylation panels centered on MAX–MTURN–HLA-B, with context-guided inclusion of CAV1, may contribute meaningfully to earlier, less invasive, and more equitable LC detection strategies.

Fig. 8 Schematic representation of the effect of aberrant promoter methylation on gene expression in lung cancer.



Aberrant promoter methylation of oncogenes and tumor suppressor genes plays an important role in the regulation of gene expression in lung cancer (LC). (a) In small cell lung cancer (SCLC), hypermethylation of promoter CpG sites inhibits RNA polymerase (RP) binding, resulting in transcriptional silencing and downregulation of the selected genes, which function as tumor suppressor genes. (b) In non-small cell lung cancer (NSCLC), hypomethylation of promoter CpG sites facilitates RP binding, leading to increased transcription and upregulation of CAV1, which acts as an oncogene. (c) In NSCLC, hypermethylation of promoter CpG sites inhibits RP binding, resulting in reduced transcription and downregulation of MAX, MTURN, and HLA-B, which act as tumor suppressor genes. MAX (MYC-associated factor X), MTURN (maturin, neural progenitor differentiation regulator homolog), and HLA-B (major histocompatibility complex, class I, B), CAV1, caveolin-1; RP, RNA polymerase. Created with BioRender.com.

Conclusion: This study demonstrates that epigenetic repression is tightly linked to lung cancer progression. Promoter hypermethylation of *MAX, MTURN,* and *HLA-B* inversely tracked with their transcript levels in tumor tissue, consistent with methylation-driven gene silencing. By contrast, CAV1 showed a subtype-dependent (biphasic) pattern—hypomethylated in NSCLC but hypermethylated in SCLC highlighting context-specific regulation. In plasma-derived cfDNA,

methylation indices for the panel achieved strong diagnostic performance—high specificity and sensitivity with favorable PPV and NPV—comparable to tissue gDNA. Concordance analyses further revealed fair-to-moderate agreement between plasma and tissue methylation across all genes, supporting cfDNA methylation as a practical "tissue-mirror" readout for LC detection, particularly in NSCLC. Notably, the divergent behavior of CAV1 between SCLC and NSCLC suggests potential value for subtype discrimination, though confirmation in larger SCLC cohorts is needed. Taken together, these findings reinforce promoter hypermethylation as a key mechanism shaping mRNA expression in LC and position plasma cfDNA methylation as a minimally invasive avenue for early screening. Integrating *MAX, MTURN, and HLA-B* into a multi-marker panel can further boost diagnostic sensitivity and specificity beyond single-gene assays.

References:

1. Belinsky SA, Palmisano WA, Gilliland FD, et al (2002) Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res 62:2370–2377
2. Belinsky SA (2004) Gene-promoter hypermethylation as a biomarker in lung cancer. Nat Rev Cancer 4:707–717
3. Belinsky SA (2005) Silencing of genes by promoter hypermethylation: key event in rodent and human lung cancer. Carcinogenesis 26:1481–1487
4. Brodie SA, Lombardo C, Li G, et al (2014) Aberrant promoter methylation of caveolin-1 is associated with favorable response to taxane-platinum combination chemotherapy in advanced NSCLC. PLoS One 9:e107124
5. Constâncio V, Nunes SP, Henrique R, Jerónimo C (2020) DNA methylation-based testing in liquid biopsies as detection and prognostic biomarkers for the four major cancer types. Cells 9:624
6. Ferlay J, Ervik M, Lam F, Laversanne M, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F (2024) Global Cancer Observatory: Cancer Today. International Agency for Research on Cancer, Lyon

7. Guo YL, Zhu TN, Guo W, Dong ZM, Zhou Z, Cui YJ, Zhao RJ (2016) Aberrant CpG island shore region methylation of CAV1 is associated with tumor progression and poor prognosis in gastric cardia adenocarcinoma. Arch Med Res 47:460–470
8. Hu S, Cao Y, Chen R, et al (2023) Accurate detection of early-stage lung cancer using cfDNA methylation signatures. Nat Commun 14:2761
9. Hulbert A, Jusue-Torres I, Stark A, et al (2017) Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. Clin Cancer Res 23:1998–2005
10. Jamshidi A, Liu MC, Klein EA, et al (2022) Evaluation of cell-free DNA approaches for multi-cancer early detection. Cancer Cell 40:1537–1549.e12
11. Lian D, Lian L, Zeng D, Zhang M, Chen M, Liu Y, Ying W, Zhou S (2022) Identification of prognostic values of the transcription factor-CpG-gene triplets in lung adenocarcinoma: a narrative review. Medicine (Baltimore) 101:e32045
12. Liu B, Ricarte Filho J, Mallisetty A, et al (2020) Detection of promoter DNA methylation in urine and plasma aids the detection of non-small cell lung cancer. Clin Cancer Res 26:4339–4348
13. Liu L, Song X, Li X, et al (2019) A three-platelet mRNA set: MAX, MTURN and HLA-B as biomarker for lung cancer. J Cancer Res Clin Oncol 145:2713–2723
14. Machado DI, Brito-Rocha T, Salta S, et al (2025) Circulating cell-free DNA methylation as biomarker for lung cancer detection: a systematic review and meta-analysis of diagnostic studies. Syst Rev 14:120
15. Moss J, Magenheim J, Neiman D, et al (2018) Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. Nat Commun 9:5068
16. Nie Y, Yang G, Song Y, Zhao X, So C, Liao J, Wang LD, Yang CS (2001) DNA hypermethylation is a mechanism for loss of expression of the HLA class I genes in human esophageal squamous cell carcinomas. Carcinogenesis 22:1615–1623
17. Puttick C, Jones TP, Leung MM, et al (2024) MHC Hammer reveals genetic and non-genetic HLA disruption in cancer evolution. Nat Genet 56:2121–2131
18. Romero OA, Torres-Diz M, Pros E, et al (2014) MAX inactivation in small cell lung cancer disrupts MYC-SWI/SNF programs and is synthetic lethal with BRG1. Cancer Discov 4:292–303

19. Shi J, Li J, Guan H, et al (2020) Performance of SHOX2 and RASSF1A methylation assay for lung cancer diagnosis. Front Oncol 10:565780
20. Sunaga N, Miyajima K, Suzuki M, et al (2004) Different roles for caveolin-1 in the development of non-small cell lung cancer versus small cell lung cancer. Cancer Res 64:4277–4285
21. Szpechcinski A, Chorostowska-Wynimko J, Struniawski R, et al (2015) Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease. Br J Cancer 113:476–483
22. Toyooka S, Toyooka KO, Maruyama R, et al (2001) DNA methylation profiles of lung tumors. Mol Cancer Ther 1:61–67
23. Wang D, Hashimoto H, Zhang X, et al (2017) MAX is an epigenetic sensor of 5-carboxylcytosine and is altered in multiple myeloma. Nucleic Acids Res 45:2396–2407
24. Wei B, Wu F, Xing W, et al (2021) A panel of DNA methylation biomarkers improves lung cancer detection in plasma. Sci Rep 11:16782
25. Weiss G, Schlegel A, Kottwitz D, König T, Tetzner R (2017) Validation of the SHOX2/PTGER4 DNA methylation marker panel for plasma-based discrimination of malignant vs nonmalignant lung disease. J Thorac Oncol 12:77–84
26. Ye Q, Shen Y, Wang X, et al (2010) Hypermethylation of HLA class I gene is associated with HLA class I down-regulation in human gastric cancer. Tissue Antigens 75:30–39
27. Yoon K-A, Park S, Lee S-H, et al (2009) Comparison of circulating plasma DNA levels between lung cancer patients and healthy controls. J Mol Diagn 11:182–185

Author Contributions:

Protocol/Project Development: Santosh Kumar

Data Collection or Management: Seema Singh, Santosh Kumar

Data Analysis: Suryakant

Manuscript Writing/Editing: Seema Singh

Supervisor: M K Ahmad, Dileep Verma

Funding: Financial support was obtained from Ministry of Health and Family Welfare (Department of Health Research), New Delhi, India file no. R.12014/09/2021-HR

Financial interests: The authors declare they have no financial interests.

Ethics and Institutional Review Board Statement: This study was reviewed and approved by the by the Institutional Ethics Committee of King George's Medical University, Lucknow, India (Ref. No. 110th ECM IIA/P2). All procedures involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Consent to Participate: Written informed consent to participate in the study was obtained from all individual participants

Data Availability Statement: The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request. All data have been de-identified to protect participant confidentiality, and any shared datasets will not contain information that could directly or indirectly identify individual patients. The original contributions presented in this study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author(s).

Acknowledgments: The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Consent for publication

Participants were informed that the results of the study may be published in scientific journals and presented at scientific meetings. Written informed consent for publication of anonymized data was obtained from all participants. No identifying information of any participant is included in this manuscript.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementryfile.docx