Cancer Management and Research

Elevate Series

Open Access Full Text Article

ORIGINAL RESEARCH

Evaluation of DNA Methylation in TAC1, SOX17, and RASSF1A for the Early Diagnosis of Lung Cancer

Huan Wang[1,*], Teng Zhang[2,*], Liqun Li[3], Feina Lin[1], Xujing Yuan[1], Jing Wang[1], Jiadong Xu[1], Haijun Xu[4], Feng Zhou[4], Weifeng Luo[2], Zhijun Chen[1]

[1]Department of Cardiothoracic Surgery, Zhoushan Hospital, Wenzhou Medical University, Zhoushan, People's Republic of China; [2]Shanghai R & S Biotechnology Co., Ltd, Shanghai, People's Republic of China; [3]Department of Pulmonary and Critical Care Medicine, Zhoushan Hospital, Wenzhou Medical University, Zhoushan, People's Republic of China; [4]Lincheng Street Community Health Service Center, Zhoushan, People's Republic of China

*These authors contributed equally to this work

Correspondence: Zhijun Chen, Email zschenzhijun@126.com

Background: Low-Dose Computed Tomography (LDCT) is commonly used to detect pulmonary nodules; however, it may also contribute to overdiagnosis. DNA methylation shows promise as an approach to discriminate early-stage lung cancer (LC) patients from individuals with benign nodules (hereafter referred to as benign pulmonary nodule group) and healthy individuals.

Methods: This study investigated the performance of DNA methylation in three genes (*TAC1, SOX17*, and *RASSF1A*) in plasma-derived cell-free DNA (cfDNA) for discriminating LC patients from benign and healthy individuals (collectively referred to as benign/healthy individuals). We enrolled 149 LC patients (96 with early-stage [stage IA] and 53 with advanced-stage [non-IA]), 54 benign pulmonary nodule group, and 75 healthy individuals.

Results: Methylation-positive rates for all three genes were significantly higher in LC patients compared to benign/healthy individuals. A combined three-gene model based on the ΔCt values of the three genes demonstrated robust diagnostic performance, achieving a sensitivity of 97.7%, specificity of 96.6%, and an area under the curve (AUC) of 0.99 for discriminating LC patients from benign/healthy individuals. Furthermore, another combined three-gene model based on the ΔCt values of the same genes showed high diagnostic performance for discriminating IA-stage LC patients from benign/healthy individuals, with a sensitivity of 96.9%, specificity of 88.54%, and AUC of 0.95.

Conclusion: This study highlights the robust diagnostic value of a combined three-gene (*TAC1, SOX17*, and *RASSF1A*) methylation model for detecting LC, including early-stage disease, offering high sensitivity and specificity.

Keywords: DNA methylation, diagnosis, lung cancer, early-stage

Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide, primarily due to the late stage at which it is often diagnosed.[1] The implementation of low-dose computed tomography (LDCT) screening has demonstrated a significant reduction in lung cancer mortality by enabling the detection of tumors at earlier, more treatable stages.[2] This underscores the critical importance of effective early detection strategies.

Despite its proven benefit, LDCT screening is hampered by substantial limitations, most notably a high false-positive rate.[3] The majority of indeterminate pulmonary nodules detected by LDCT are ultimately benign, yet their evaluation necessitates repeated radiation exposure, invasive diagnostic procedures (eg, biopsy), and prolonged patient anxiety, placing a significant burden on healthcare systems.[4] Concurrently, classical serum protein biomarkers, such as carcinoembryonic antigen (CEA), have shown insufficient sensitivity and specificity for the early detection of lung cancer, failing to reliably distinguish malignant from benign nodules.[5] Therefore, a pressing clinical need exists for highly accurate, non-invasive biomarkers to complement LDCT and improve the stratification of pulmonary nodules.

https://doi.org/10.2147/CMAR.S573663
Received: 14 October 2025
Accepted: 27 February 2026
Published: 17 March 2026

In this context, cell-free DNA (cfDNA) methylation profiling has emerged as a particularly promising approach for cancer detection and differentiation.[6] DNA methylation alterations are hallmark epigenetic events in carcinogenesis, occurring early and exhibiting cancer-specific patterns.[7] Unlike genetic mutations, epigenetic modifications are reversible and can be quantitatively measured in blood-based assays, offering a stable and clinically actionable readout.[8] The development of methylation biomarkers is increasingly guided by translational frameworks that emphasize rigorous analytical and clinical validation as prerequisites for clinical utility.[9] Furthermore, the robustness of this approach is being shaped by field-wide systematic reviews, which have effectively charted the landscape of methylation biomarkers in other malignancies, such as breast cancer, providing a methodological paradigm for evaluating reproducibility and clinical potential.[10] Our study is situated within this evolving framework.

In lung cancer, several genes, including *SHOX2, RASSF1A, and PTGER4*, have been frequently reported as hypermethylated in tumor tissues and detectable in plasma.[11–13] The SHOX2 gene methylation is associated with cell cycle regulation and has shown promise in distinguishing lung cancer from benign lung diseases.[11] RASSF1A is a well-characterized tumor suppressor gene often silenced in various cancers, including lung cancer.[12] PTGER4 methylation has been linked to inflammatory pathways and may play a unique role in differentiating malignancy from inflammation-driven benign nodules.[13] However, the performance of a combined panel of these three specific markers for differentiating invasive adenocarcinoma (IA) from benign nodules in a screening-like population requires further validation.

Therefore, we hypothesized that a multiplex methylation assay targeting SHOX2, RASSF1A, and PTGER4 could serve as a highly sensitive and specific tool for the non-invasive detection of early-stage lung cancer. The objective of this study was to analytically and clinically validate this methylation signature in a cohort of patients with pulmonary nodules, with the ultimate aim of providing a complementary liquid biopsy tool to reduce unnecessary invasive procedures following LDCT screening.

Materials and Methods

Sample Collection

Blood specimens were collected through the Department of Laboratory Medicine at Wenzhou Medical University Affiliated Zhoushan Hospital. Corresponding clinical and pathological data were retrieved from the Department of Thoracic and Cardiovascular Surgery and the Department of Pathology, respectively. All participants provided written informed consent.

From July 2022 to October 2024, a total of 278 blood samples were collected for this study, including 149 samples from patients with pulmonary nodules (including solid masses or ground-glass nodules), 54 from benign pulmonary nodule group, and 75 from healthy individuals.

Inclusion and Exclusion Criteria

For LC and BPN Groups

Inclusion Criteria: (a) Age > 18 years; (b) Presence of a newly detected pulmonary nodule on computed tomography (CT) scan requiring further pathological diagnosis; (c) Scheduled for surgical resection or biopsy; (d) Willing to provide written informed consent; (e) For the IA stage subgroup, all patients had a maximum tumor dimension of ≤ 3 cm (T1), with no lymph node involvement (N0) or distant metastasis (M0); (f) For BPN groups, patients with pulmonary nodules that were confirmed to be benign by pathological examination (via surgical resection or biopsy) or through clinical follow-up (with CT scan showing stability or regression for at least 12 months).

Exclusion Criteria: (a) History of any prior malignancy; (b) Received neoadjuvant chemotherapy or radiotherapy prior to blood sampling; (c) Inability to provide informed consent.

For Healthy Group

Inclusion Criteria: (a) Age > 18 years; (b) No abnormalities detected on routine chest X-ray or physical examination; (c) No personal history of any cancer; (d) Willing to provide written informed consent.

Exclusion Criteria: (a) Presence of any acute or chronic inflammatory disease at the time of recruitment; (b) History of any chronic respiratory disease (eg, COPD, pulmonary fibrosis). Malignant lesions were staged using the revised TNM classification guidelines.[14]

DNA Isolation and Quantitative Multiplex Methylation-Specific PCR (qMSP)

Plasma (5 mL) was collected from each individual, and cfDNA was extracted using the MagMAX™ plasma-free DNA extraction kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. DNA concentration and purity were assessed using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, USA).

DNA underwent bisulfite conversion using a commercial methylation kit (Shanghai R&S Biotechnology, China). Methylation quantification employed the EpiTect MethyLight Master Mix (Qiagen, USA) with fluorescent probes, using 50 ng bisulfite-converted DNA, 50 nM probe, and 100 nM each primer per reaction. Multiplex qMSP assays simultaneously quantified methylation of *TAC1, SOX17*, and *RASSF1A*, with *β-actin* (*ACTB*) as the reference gene. Amplification was performed in a SLAN-96S PCR system (HONGSHI, China) under the following conditions: 94°C for 30s; 45 cycles of 94°C for 5s and 60°C for 30s (fluorescence acquisition). ΔCt was calculated as Ct (target) – Ct (ACTB). Controls included: Positive control: 1:1:1 DNA mixture from A549, NCI-H209, and NCI-H1975 cell lines; Negative control: PBMC gDNA from healthy donors (Sanger-verified).

Primer/probe sequences (Supplementary Table 1) were synthesized by Sangon Biotech (China). Successful sample detection required the reference gene (*ACTB*) to have a Ct value < 35. For valid samples, methylation of a target gene was defined by a Ct value < 42. For samples where a gene was unmethylated, the Ct value was set to 45, corresponding to the total number of PCR cycles.

Construction of Models for LC Diagnosis and IA Stage Prediction

Logistic regression models based on the combined three-gene were built to discriminate LC patients from benign/healthy individuals, or IA LC patients from benign/healthy individuals. Model performance was evaluated using receiver operating characteristic (ROC) curves, with the area under the curve (AUC) as the assessment metric. Logistic regression analysis and ROC curve generation were performed using GraphPad Prism (version 10.2.1) (Dotmatics, USA). Model A (LC vs Benign/Healthy): pre = 1.414–0.05675 × ΔCt(SOX17) - 0.02257 × ΔCt(TAC1) - 0.03637 × ΔCt(RASSF1A). Model B (IA LC vs Benign/Healthy): pre = 0.37636 × ΔCt(SOX17) + 0.17867 × ΔCt(TAC1) - 0.16980 × ΔCt (RASSF1A). The linear predictors (pre) generated by the logistic regression models (Model A and Model B) represent the log-odds of a sample being classified into the positive class. The final probability was calculated using the logistic function: $P=1/(1+e^{-pre})$.

Statistical Analysis

All statistical analyses were performed using GraphPad Prism (v10.2.1; Dotmatics, USA). All reported *P* values were two-sided, with $P < 0.05$ considered statistically significant. Significance levels are denoted as follows: * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$. We chose logistic regression as the primary classification model due to its high clinical interpretability, low risk of overfitting given the low number of pre-selected features, and its demonstrated sufficient performance in capturing the linear relationship between the methylation markers and malignancy.

Results

Patient Characteristics

A total of 278 blood samples from 278 participants were included: 112 patients with lung masses (83 LC, 29 benign), 91 patients with GGNs (66 LC, 25 benign), and 75 healthy individuals (Figure 1).

As shown in Table 1, among LC patients, males were more prevalent in the mass group (49, 59.0%) than in the GGN group (27, 40.9%) ($P = 0.046$). Conversely, females were more prevalent in the GGN group (39, 59.1%) than in the mass group (34, 41.0%) ($P = 0.046$). Additionally, LC patients with masses were older than those with GGNs (67.5 ± 7.5 vs 62.0 ± 10.2 years, $P = 0.003$). Moreover, IA-stage LC patients were predominantly from the GGN group (51, 77.3% vs

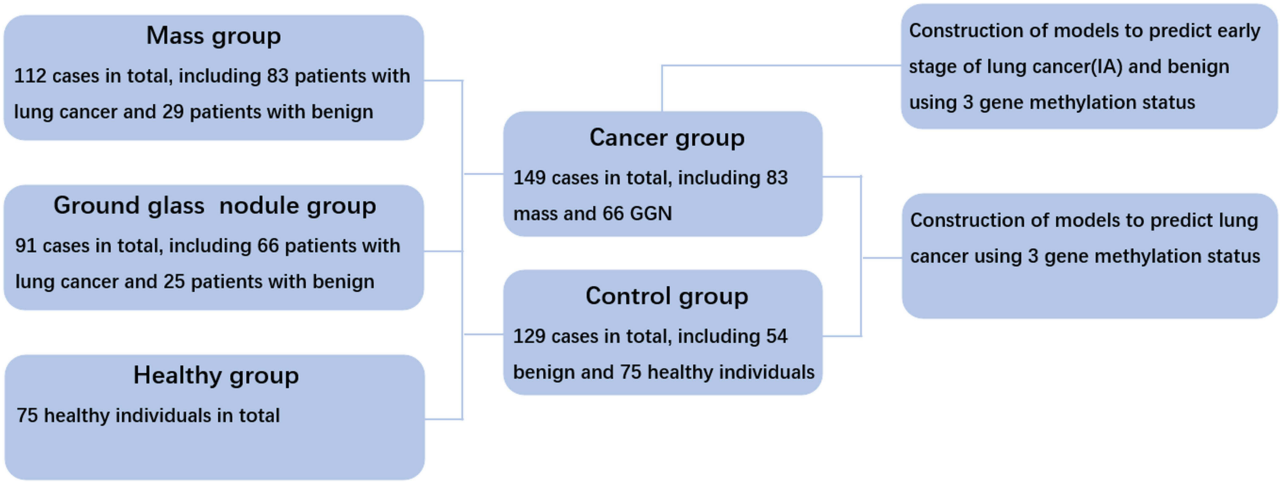

Figure 1 Flowchart depicting participants with different conditions and the subsequent diagnostic model construction based on methylation status.

45, 54.2%; $P = 0.004$), whereas non-IA-stage LC patients were predominantly from the mass group (38, 45.8% vs 15, 22.7%; $P = 0.004$). Histopathological analysis revealed no significant differences in subtypes between mass and GGN LC patients. In summary, patients with masses and GGNs contributed differently to the LC cohort regarding sex, age, and pathological stage.

Diagnostic Value of the 3-Gene Methylation Panel for LC

We first assessed the diagnostic value of *TAC1, SOX17*, and *RASSF1A* methylation in all LC patients. Compared to benign/healthy individuals, LC patients exhibited significantly higher methylation-positive rates for all three genes (Figure 2). Specifically: *SOX17* methylation was positive in 124 (83.22%) LC patients vs 2/54 (3.70%) benign pulmonary nodule group and 1/75 (1.33%) healthy individuals (Figure 2A); *TAC1* methylation was observed in 122 (81.88%) LC patients vs 6/54 (11.11%) benign and 3/75 (4.00%) healthy individuals (Figure 2B); *RASSF1A* methylation was present in 126 (84.56%) LC patients vs 9/54 (16.67%) benign and 4/75 (5.33%) healthy individuals (Figure 2C).

Subsequently, we evaluated the performance of each gene for discriminating LC patients from benign/healthy individuals. ROC curve analysis revealed that *SOX17* methylation achieved a sensitivity of 0.8322, specificity of 0.9767, and AUC of 0.94 (Figure 3A). For *TAC1* methylation, the sensitivity, specificity, and AUC reached 0.8188, 0.9302, and 0.9049, respectively (Figure 3B). Similarly, *RASSF1A* methylation demonstrated a sensitivity of 0.8456,

Table 1 The Clinical Information of the 149 Lung Cancer Patients

Clinicopathological Features	Lung Cancer (n = 149)	Mass Group (n = 83)	GGNs Group (n = 66)	*P*
Sex (n, %)				
Male	76 (51.0%)	49 (59.0%)	27 (40.9%)	0.046
Female	73 (49.0%)	34 (41.0%)	39 (59.1%)	0.046
Age (mean±SD)	65.0±9.1	67.5±7.5	62.0±10.2	0.003
Squamous carcinoma	35 (23.5%)	21 (25.3%)	14 (21.2%)	0.846
Adenocarcinoma	102 (68.5%)	54 (65.1%)	48 (72.7%)	0.622
Adenosquamous carcinoma	7 (4.7%)	5 (6.0%)	2 (3.0%)	0.675
Neuroendocrine neoplasm	3 (2.0%)	2 (2.4%)	1 (1.5%)	1.000
pTNM stage (n, %)				
IA	96 (64.4%)	45 (54.2%)	51 (77.3%)	0.004
Non-IA	53 (35.6%)	38 (45.8%)	15 (22.7%)	0.004

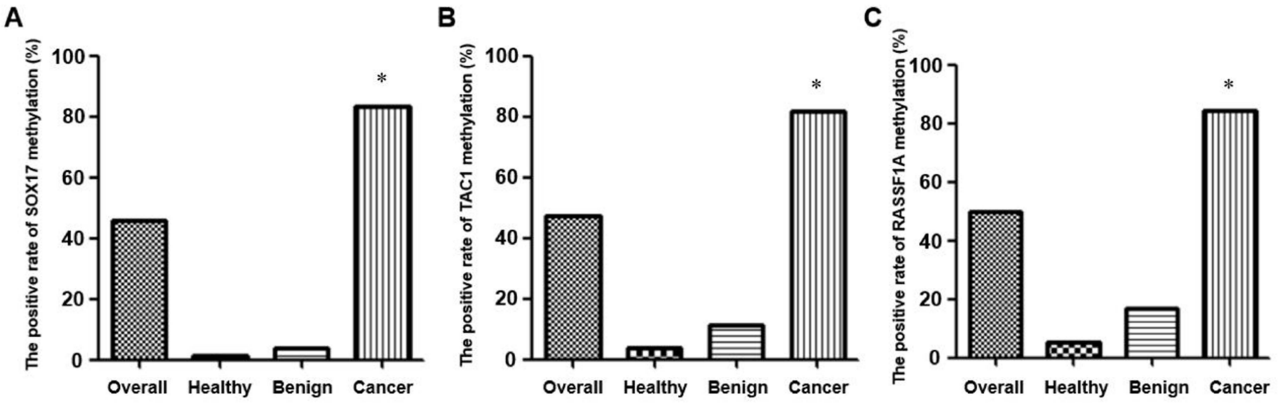

Figure 2 Positive methylation rates for *SOX17* (**A**), *TAC1* (**B**), and *RASSF1A* (**C**) in benign pulmonary nodule group (Benign) and all lung cancer patients (LC).* indicates a significant difference ($p < 0.05$). strict pre-defined Ct cutoffs for positivity (Ct<42 for methylated, Ct=45 for unmethylated).

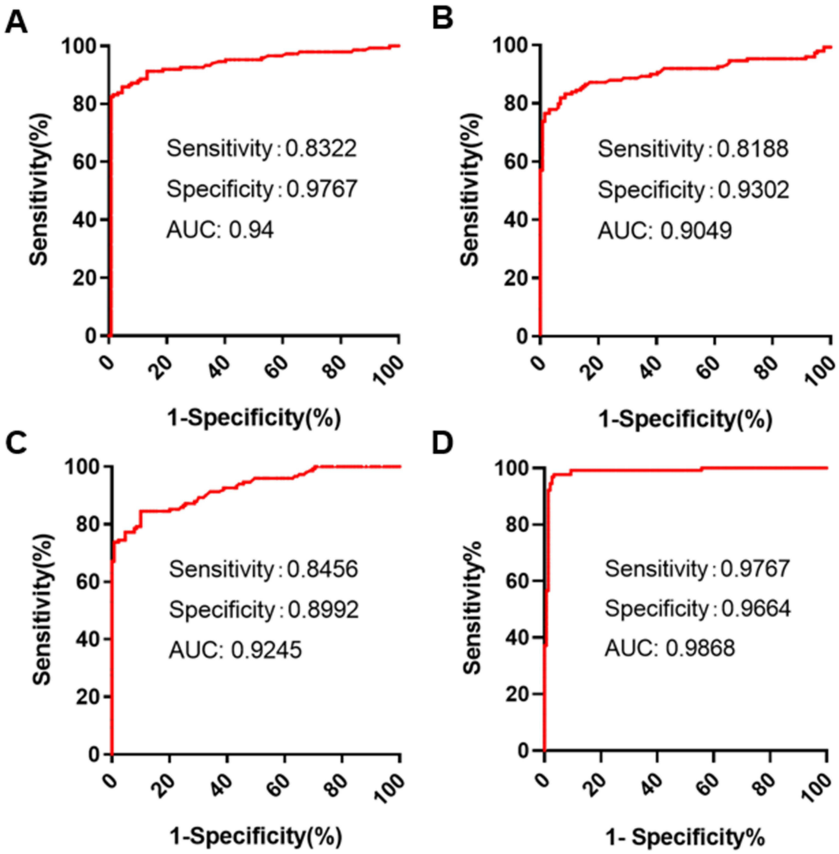

Figure 3 ROC curves showing performance of *SOX17* (**A**), *TAC1* (**B**), *RASSF1A* (**C**) methylation, and the combined three-gene model (**D**) for discriminating all LC patients from benign/healthy individuals. Cutoff values maximize Youden's index.

specificity of 0.8992, and AUC of 0.9245 (Figure 3C). These results highlight the potential diagnostic value of the combined three-gene diagnostic model in LC diagnosis.

And then, we developed a logistic regression model (Model A) incorporating the ΔCt values of all three genes. This combined model demonstrated superior performance: Sensitivity 97.67%, Specificity 96.64%, AUC 0.9868 (Figure 3D), outperforming single-gene analyses.

https://doi.org/10.2147/CMAR.S573663

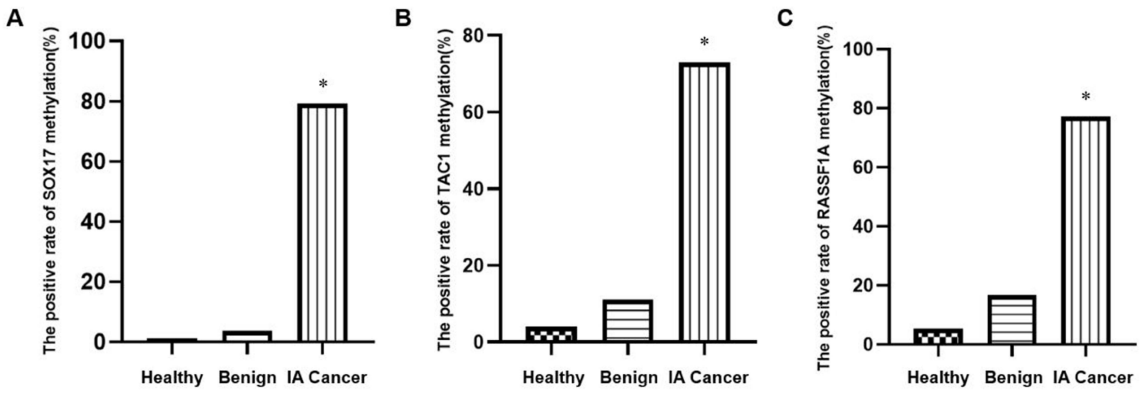

Figure 4 Positive methylation rates for *SOX17* (**A**), *TAC1* (**B**), and *RASSF1A* (**C**) in benign pulmonary nodule group (Benign) and IA-stage lung cancer patients (IA LC).* indicates a significant difference ($p < 0.05$).

Diagnostic Value of the 3-Gene Methylation Panel for Early-Stage (IA) LC

We next evaluated the diagnostic value for early-stage (IA) LC. Compared to benign/healthy individuals, IA LC patients also showed significantly higher methylation-positive rates for all three genes (Figure 4). Specifically, *SOX17* methylation was positive in 76 (79.17%) IA LC patients, while only 2/54 (3.70%) benign pulmonary nodule group and 1/53 (1.33%) healthy individuals showed positive (Figure 4A). Similarly, *TAC1* methylation was detected in 70 (72.92%) IA LC patients, compared to 6/54 (11.11%) benign pulmonary nodule group and 3/53 (4.0%) healthy individuals (Figure 4B). Furthermore, *RASSF1A* methylation was observed in 74 (77.08%) IA LC patients, whereas 9/54 (16.67%) benign pulmonary nodule group and 4/53 (5.33%) healthy individuals tested positive (Figure 4C).

We then evaluated the performance of each gene for discriminating IA LC patients from benign/healthy individual. ROC curve analysis revealed that *SOX17* methylation achieved a sensitivity of 0.8438, specificity of 0.9535, and AUC of 0.9305 (Figure 5A). For *TAC1* methylation, the sensitivity, specificity, and AUC were 0.7396, 0.9147, and 0.8526, respectively (Figure 5B). Similarly, *RASSF1A* methylation demonstrated a sensitivity of 0.7604, specificity of 0.8992, and AUC of 0.8828 (Figure 5C).

A second logistic regression model (Model B) using the ΔCt values of the three genes was constructed for discriminating IA LC patients from benign/healthy individuals. This combined model achieved: Sensitivity 96.9%, Specificity 88.54%, AUC 0.9495 (Figure 5D), highlighting its potential utility for early-stage LC diagnosis.

Discussion

This study demonstrates that a plasma cfDNA methylation panel comprising TAC1, SOX17, and RASSF1A exhibits exceptional diagnostic performance in distinguishing patients with LC, including those with early-stage (IA) disease, from individuals with benign pulmonary nodules (BPNs) and healthy controls. The high sensitivity (97.7%) and specificity (96.6%) achieved by our combined model address a critical unmet need in current LC screening: the reliable stratification of indeterminate pulmonary nodules detected by LDCT.

The poor prognosis of advanced LC reinforces the need for effective early detection strategies.[1,3] While LDCT screening has increased early cancer detection, its high false-positive rate remains a major challenge.[4] Conventional serum tumor markers like CEA and CYFRA 21–1 have shown limited diagnostic value, particularly for distinguishing malignant from benign nodules.[5–7] In contrast, DNA methylation signatures in cfDNA offer a highly specific and sensitive approach, as epigenetic alterations are inherent to the tumor biology.[8]

The widespread adoption of LDCT has led to increased detection of pulmonary nodules. However, >95% of detected nodules are benign, predominantly granulomas or intrapulmonary lymph nodes.[15] As astutely noted, granulomas are inflammatory lesions, and inflammation itself can induce epigenetic alterations, posing a potential risk of false-positive methylation signalsThis is a critical consideration for the clinical utility of any methylation-based assay.

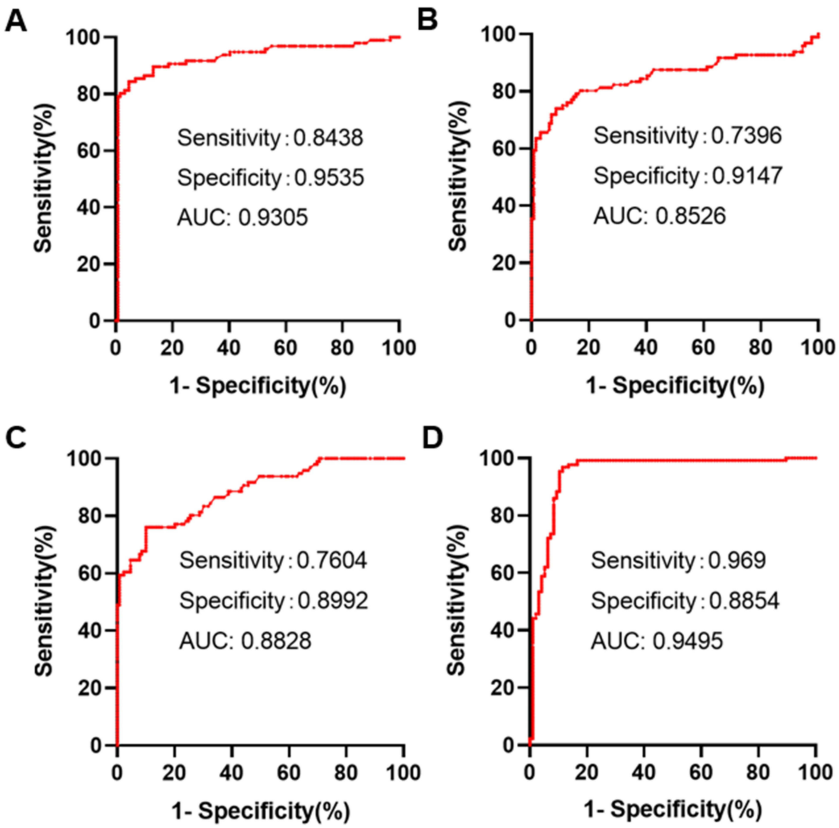

Figure 5 ROC curves showing performance of *SOX17* (**A**), TAC1 (**B**), *RASSF1A* (**C**) methylation, and the combined three-gene model (**D**) for discriminating IA-stage LC patients from benign/healthy individuals. Cutoff values maximize Youden's index.

Conventional tumor markers like SCC-Ag, CYFRA 21–1, CEA, ProGRP, and NSE have been extensively studied,[15] but a 2021 study found no significant differences in these markers between benign and malignant nodules ($P > 0.05$).[6] Individual markers showed limited diagnostic utility (AUCs: CYFRA21-1 0.506, ProGRP 0.503, CEA 0.532, NSE 0.548, SCCA 0.562), and combinations yielded similarly marginal performance (AUC range: 0.502–0.596).[6,15] While some studies explored multi-marker panels or models incorporating clinical factors (eg, smoking history) achieving better performance (eg, AUC 0.86–0.87),[16–18] sole reliance on tumor-associated antigens remains insufficient for reliable benign-malignant discrimination, necessitating more sophisticated approaches.

To address this concern, we specifically analyzed the methylation signals in our benign pulmonary nodule group (which includes inflammatory lesions like granulomas) and compared them to the healthy control group. Reassuringly, the methylation-positive rates for all three genes (TAC1, SOX17, and RASSF1A) in the benign group, while slightly higher than in the healthy controls (SOX17: 3.70% vs 1.33%; TAC1: 11.11% vs 4.00%; RASSF1A: 16.67% vs 5.33%), remained at a very low level. More importantly, the logistic regression model (Model A) achieved a high overall specificity of 96.6% when discriminating LC patients from the combined benign/healthy group, indicating that the model effectively minimizes false positives originating from both healthy individuals and those with inflammatory nodules. Aberrant DNA methylation is an early and prevalent event in carcinogenesis and is extensively investigated for early cancer detection.[19–21] However, tumor-derived DNA in liquid biopsies is scarce and obscured by high background DNA from healthy tissues, often with variant allele frequencies as low as 0.01%,[22] representing only 1–7 copies per mL of plasma.[23] Therefore, identifying highly specific methylation markers and developing ultrasensitive assays are critical for early diagnosis.

We acknowledge that *TAC1, SOX17, and RASSF1A* have been individually investigated in NSCLC, as referenced in our introduction.[11,12,24] The novelty of our study lies not in the initial discovery of these genes' association with lung

https://doi.org/10.2147/CMAR.S573663

cancer, but in the systematic selection and validation of this specific three-gene combination for the critical clinical scenario of differentiating early-stage (IA) lung cancer from benign pulmonary nodules.

When compared to existing methylation panels described in the literature, our *TAC1/SOX17/RASSF1A* combination demonstrates superior performance, particularly for the early-stage cohort which is the primary target of screening. For instance, Hulbert et al reported a panel comprising *CDO1, TAC1, and SOX17* for detecting stage I–IIA NSCLC, achieving a sensitivity of 86% and specificity of 78% (AUC 0.77).[24] While this study validated the diagnostic potential of *TAC1 and SOX17*, our work, by integrating RASSF1A, demonstrates a substantial improvement in performance. Our three-gene panel achieved a markedly higher sensitivity (96.9%) and specificity (88.54%) for IA-stage LC, with an AUC of 0.95. This suggests that the inclusion of RASSF1A provides critical additional discriminatory power.

Similarly, compared to the multi-gene model by Hu et al, which reported 81% sensitivity and 98% specificity,[25] our panel achieves a significantly higher sensitivity for early-stage cancer detection while maintaining high specificity. The performance of our model also compares favorably with studies focusing on individual genes or smaller panels, such as *SHOX2* methylation (60% sensitivity)[11] or *RASSF1A/RARB* (87% sensitivity, 75% specificity).[12]

Beyond the landscape of methylation biomarkers in LC, recent advances in multi-omics and machine learning have identified novel pivotal genes, such as PAK2, as central hubs in the tumorigenesis of other malignancies like head and neck squamous cell carcinoma.[26] Studies like that of Wang demonstrate the power of integrated approaches to uncover genes that orchestrate critical cancer hallmarks, including treatment resistance and immune evasion. While the biological context differs, this emerging paradigm underscores the importance of identifying highly specific molecular drivers for cancer diagnosis. In this context, our study reinforces the value of a focused biomarker strategy. This focused approach may offer advantages in terms of clinical practicality, cost-effectiveness, and easier translation into a standardized clinical assay compared to more complex multi-gene signatures.

This study conclusively establishes the clinical validity of a plasma cfDNA methylation panel targeting TAC1, SOX17, and RASSF1A for the diagnosis of LC. We have developed and validated a combined three-gene model that demonstrates not only high sensitivity and specificity for overall LC detection (AUC 0.99, Sensitivity 97.7%, Specificity 96.6%) but, more critically, exhibits superior performance for early-stage IA disease (AUC 0.95, Sensitivity 96.9%, Specificity 88.54%). The ability of this panel to effectively differentiate malignant from benign pulmonary nodules addresses the most pressing limitation of current LDCT screening programs—the high false-positive rate.

Looking forward, as emerging therapeutic strategies increasingly rely on molecular profiling, the value of precise early-detection biomarkers rises in parallel. Our findings highlight the translational potential of this assay to not only improve screening accuracy but also to facilitate earlier integration with personalized treatment strategies, thereby bridging a critical gap between early diagnosis and the evolving paradigm of precision oncology.

Several limitations warrant consideration in this study. Primarily, the cohort size of the benign pulmonary nodule group (n=54) was comparatively smaller than those of the lung cancer cohort and healthy controls. Additionally, the single-center design necessitates caution in generalizability due to the absence of external validation. To facilitate clinical translation of this assay, future validation through large-scale, prospective, multi-center studies within LDCT screening populations is imperative.

Strengths and Limitations of This Study

Strength: Rigorous Analytical Methodology

Multiplex quantitative Methylation-Specific PCR (qMSP) was employed for simultaneous assessment of TAC1, SOX17, and RASSF1A methylation in plasma cfDNA. The use of ΔCt as the quantitative measure, strict pre-defined Ct cutoffs for positivity (Ct<42 for methylated, Ct=45 for unmethylated), and inclusion of positive/negative controls (cell line mix, PBMC gDNA) enhance analytical precision and reliability.

Strength: Focus on Early-Stage Cohort

The study specifically included a substantial cohort of patients with pathologically confirmed early-stage (IA) lung cancer (n=96), allowing for the methodological development and internal validation of diagnostic models specifically aimed at this clinically crucial target population.

Limitation: Sample Size of Benign Group

The cohort of individuals with benign pulmonary nodules (n=54) is relatively small compared to the lung cancer group (n=149) and the healthy control group (n=75). This imbalance could potentially limit the robustness and generalizability of the model's performance estimates for distinguishing cancer from benign nodules specifically.

Limitation: Single-Center Design & Lack of External Validation

The study utilized samples collected from a single institution, and all model development and validation were performed internally on this cohort. The absence of an external, independent validation cohort limits the assessment of the models' generalizability across different populations and clinical settings.

Ethical Statement

This study was conducted in accordance with the Declaration of Helsinki (as amended in 2013), and approved by the Clinical Ethics Committee of Zhoushan Hospital (Approval No.: [2022-151], Approval Date: 2022.07.20). We strictly adhered to relevant Chinese laws and regulations and internationally recognized ethical guidelines. All participants provided written informed consent after understanding the research content, risks, and benefits. Rigorous data security and privacy protection measures ensured anonymity and confidentiality.

Funding

This work was supported by Zhejiang Provincial Medical and Health Science and Technology Project (2023KY1300), Zhoushan City Science and Technology Program (2020C31045), the Science and Technology Bureau of Putuo District, Zhoushan City (2023YB003), and "Zhou Creates Future" Haina Plan Innovative Talent (Team) ([2023]1-08).

Disclosure

The authors declare no potential conflicts of interest. Huan Wang and Teng Zhang contributed equally to this work.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
2. Siegel RL, Kratzer TB, Giaquinto AN, et al. Cancer statistics, 2025. *CA Cancer J Clin*. 2025;75(1):10–45. doi:10.3322/caac.21871
3. Ettinger DS. Ten years of progress in non-small cell lung cancer. *J Natl Compr Canc Netw*. 2012;10(3):292–295. doi:10.6004/jnccn.2012.0029
4. Gettinger S, Horn L, Jackman D, et al. Five-year follow-up of nivolumab in previously treated advanced non-small-cell lung cancer: results from the CA209-003 study. *J Clin Oncol*. 2018;36(17):1675–1684. doi:10.1200/JCO.2017.77.0412
5. Krist AH, Davidson KW, Mangione CM, et al. Screening for lung cancer: us preventive services task force recommendation statement. *JAMA*. 2021;325(10):962–970. doi:10.1001/jama.2021.1117
6. Xu L, Su Z, Xie B. Diagnostic value of conventional tumor markers in young patients with pulmonary nodules. *J Clin Lab Anal*. 2021;35(9):e23912. doi:10.1002/jcla.23912
7. Kammer MN, Lakhani DA, Balar AB, et al. Integrated biomarkers for the management of indeterminate pulmonary nodules. *Am J Respir Crit Care Med*. 2021;204(11):1306–1316. doi:10.1164/rccm.202012-4438OC
8. Li Z, Lu L, Deng Y, et al. Genetic susceptibility loci of lung cancer are associated with malignant risk of pulmonary nodules and improve malignancy diagnosis based on CEA levels. *Chin J Cancer Res*. 2023;35(5):501–510. doi:10.21147/j.issn.1000-9604.2023.05.07
9. Xu CM, Luo YL, Li S, et al. Multifunctional neuron-specific enolase: its role in lung diseases. *Biosci Rep*. 2019;39(11):BSR20192732. doi:10.1042/BSR20192732
10. Koch A, Joosten SC, Feng Z, et al. Analysis of DNA methylation in cancer: location revisited. *Nat Rev Clin Oncol*. 2018;15(7):459–466. doi:10.1038/s41571-018-0004-4
11. Kneip C, Schmidt B, Seegebarth A, et al. SHOX2 DNA methylation is a biomarker for the diagnosis of lung cancer in plasma. *J Thorac Oncol*. 2011;6(10):1632–1638. doi:10.1097/JTO.0b013e318220ef9a
12. Ponomaryova AA, Rykova EY, Cherdyntseva NV, et al. Potentialities of aberrantly methylated circulating DNA for diagnostics and post-treatment follow-up of lung cancer patients. *Lung Cancer*. 2013;81(3):397–403. doi:10.1016/j.lungcan.2013.05.016

13. Xu W, Lu J, Zhao Q, et al. Genome-wide plasma cell-free DNA methylation profiling identifies potential biomarkers for lung cancer. *Dis Markers*. 2019;2019:4108474. doi:10.1155/2019/4108474
14. Detterbeck FC, Boffa DJ, Kim AW, et al. The Eighth edition lung cancer stage classification. *Chest*. 2017;151(1):193–203. doi:10.1016/j.chest.2016.10.010
15. Mazzone PJ, Lam L. Evaluating the patient with a pulmonary nodule: a review. *JAMA*. 2022;327(3):264–273. doi:10.1001/jama.2021.24287
16. Yang Q, Zhang P, Wu R, et al. Identifying the best marker combination in CEA, CA125, CY211, NSE, and SCC for lung cancer screening by combining ROC curve and logistic regression analyses: is it feasible? *Dis Markers*. 2018;2018:2082840. doi:10.1155/2018/2082840
17. Muller DC, Johansson M, Brennan P. Lung cancer risk prediction model incorporating lung function: development and validation in the UK biobank prospective cohort study. *J Clin Oncol*. 2017;35(8):861–869. doi:10.1200/JCO.2016.69.2467
18. Ajona D, Remirez A, Sainz C, et al. A model based on the quantification of complement C4c, CYFRA 21-1 and CRP exhibits high specificity for the early diagnosis of lung cancer. *Transl Res*. 2021;233:77–91. doi:10.1016/j.trsl.2021.02.009
19. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*. 2011;11(6):426–437. doi:10.1038/nrc3066
20. Ilie M, Hofman V, Long V, et al. Current challenges for detection of circulating tumor cells and cell-free circulating nucleic acids, and their characterization in non-small cell lung carcinoma patients. What is the best blood substrate for personalized medicine? *Ann Transl Med*. 2014;2(11):107. doi:10.3978/j.issn.2305-5839.2014.08.11
21. Baylin SB, Jones PA. A decade of exploring the cancer epigenome-biological and translational implications. *Nat Rev Cancer*. 2011;11(10):726–734. doi:10.1038/nrc3130
22. Crowley E, Di Nicolantonio F, Loupakis F, et al. Liquid biopsy: monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol*. 2013;10(8):472–484. doi:10.1038/nrclinonc.2013.110
23. Bronkhorst AJ, Ungerer V, Holdenrieder S, et al. The emerging role of cell-free DNA as a molecular marker for cancer management. *Biomol Detect Quantif*. 2019;17:100087. doi:10.1016/j.bdq.2019.100087
24. Hulbert A, Jusue-Torres I, Stark A, et al. Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. *Clin Cancer Res*. 2017;23(8):1998–2005. doi:10.1158/1078-0432.CCR-16-1371
25. Hu S, Tao J, Peng M, et al. Accurate detection of early-stage lung cancer using a panel of circulating cell-free DNA methylation biomarkers. *Biomark Res*. 2023;11(1):45. doi:10.1186/s40364-023-00486-5
26. Wang P, Xu S, Guo Q, et al. Discovery of PAK2 as a key regulator of cancer stem cell in head and neck squamous cell carcinoma using multi-omic techniques. *Stem Cells Int*. 2025;12(9):1325262. doi:10.1155/sci/1325262