A Panel of DNA Methylation Biomarkers for Detection and Improving Diagnostic Efficiency of Lung Cancer

Bing Wei

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Fengxin Wu

Excellen Medical Technology Co., Ltd.

Wenqun Xing

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Haibo Sun

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Chi Yan

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Chengzhi Zhao

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Dongqing Wang

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Xiaobing Chen

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Yanli Chen

Excellen Medical Technology Co., Ltd.

Mingming Li

Excellen Medical Technology Co., Ltd.

Jie Ma (✉ majie_fzbl@163.com)

Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital

Research Article

Keywords: DNA methylation, Lung cancer, Circulating tumor DNA, Liquid biopsy

Posted Date: April 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-416830/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on August 18th, 2021.

See the published version at <https://doi.org/10.1038/s41598-021-96242-6>.

Abstract

Lung cancer remains the leading cause of cancer deaths worldwide. Although low-dose spiral computed tomography (LDCT) screening is used for the detection of lung cancer in a high-risk population, false-positive results of LDCT remain a clinical problem. Here, we developed a blood test of a novel three gene methylation panel and validated the diagnostic efficiency from a total of 351 subjects (197 controls, 154 cases). DNA methylation of short stature homeobox 2 gene (*SHOX2*), ras association domain family 1A gene (*RASSF1A*), and prostaglandin E receptor 4 gene (*PTGER4*) could be used as a panel of biomarkers to distinguish between malignant and benign lung disease at high sensitivity and specificity: 87.0% sensitivity [95% CI 80.2-91.5%], 98.0% specificity [95% CI 94.9-99.4%]. Sensitivity in adenocarcinoma, squamous cell carcinoma, small cell lung cancer, and other lung cancer was 89.0%, 87.5%, 85.7%, and 77.8%, respectively. Notably, cancer patients in stage I and II showed high diagnostic sensitivity at 82.5% and 90.5%, respectively. Moreover, the diagnostic efficiency did not show bias toward age, gender, smoking, and the presence of other (nonlung) cancers. These findings clearly showed that this panel of DNA methylation biomarkers was effective in detecting lung cancer noninvasively and may provide clinical utility in stand-alone or in combination with current imaging techniques to improve the diagnosis of lung cancer.

1. Introduction

Lung cancer is the second most common cancer and the leading cause of cancer deaths worldwide [1, 2]. It surpassed breast cancer as the highest cause of cancer-related deaths in women. Lung cancer is expected to account for 22% of all female cancer deaths and 23% of all male cancer deaths in 2020 [2].

Current detection methods of lung cancer include the testing of sputum, x-ray, and computed tomography (CT) scanning of the chest. Studies have shown that using chest X-ray and sputum cytology as screening techniques does not reduce the mortality rate of lung cancer [3, 4]. Low-dose spiral computed tomography (LDCT) has been demonstrated to be more sensitive than chest radiography in high-risk smokers for the detection of early stage lung cancer. It has been shown to reduce the mortality rate [5, 6]. However, the diagnostic accuracy of LDCT screening is limited due to its low specificity causing high downstream healthcare costs [5]. Therefore, novel noninvasive methods are urgently needed to improve lung cancer detection and reduce the mortality rate.

Molecular biomarkers might have the potential for improving the detection and management of lung cancer in the clinical setting [7, 8]. Several biomarkers in gene expression [9], somatic mutations [10], copy number variations [11], differences in methylation [12], or the abundance of plasma proteins [13] have been extensively studied for clinical research. In recent years, aberrant DNA methylation has been observed in various cancers and is considered to play a significant role in carcinogenesis [14, 15]. DNA methylation is a relatively stable biochemical modification carried out by DNA methyltransferases. Methylation biomarkers can be detected not only from tissue but also in serum, fecal, and plasma [16], indicating its potential as a noninvasive method for cancer detection. It has been reported that in vitro

diagnostic tests using DNA methylation-based biomarkers have been approved by the U.S. Food and Drug Administration [17, 18].

Short stature homeobox gene two (*SHOX2*), ras association domain family 1A (*RASSF1A*), and prostaglandin E receptor 4 gene (*PTGER4*) methylation have been separately reported to be highly correlated with lung cancer diagnosis and prognosis. The promoter methylation of *SHOX2* has been identified as a valuable biomarker for lung cancer diagnosis in several research studies [19, 20]. Promotor hypermethylation of *RASSF1A* frequently occurs in lung cancer and is frequently found in small cell lung cancer [21, 22]. Methylated prostaglandin E2 receptor EP4 subtype (*mPTGER4*) was recently reported as a methylation marker for the differential diagnosis of benign and malignant lung diseases [23]. However, the combined detection of *SHOX2*, *RASSF1A*, and *PTGER4* methylation in plasma has hardly been reported. Extensive prospective screening trials have been validated that multiple gene panels could increase sensitivity and improve cancer diagnostic efficiency [17, 18, 24].

In the presented study, a panel of DNA methylation of *SHOX2*, *RASSF1A*, and *PTGER4* was tested in blood plasma based on Real-Time PCR from 351 subjects to distinguish between malignant lung cancer and controls (healthy individuals, benign lung disease, and patients with other cancer).

2. Materials And Methods

2.1 Ethics and sample size estimation

The participants were enrolled in Henan Cancer Hospital, the Affiliated Cancer Hospital of Zhengzhou University. All subjects signed the informed consent before blood collection, and they were informed of the usage of plasma and the test results. The study has been approved by the Medical Ethics Committee of Henan Cancer Hospital (2018157) and all methods were carried out in accordance with the relevant guidelines and regulations, as well as the Declaration of Helsinki.

Based on estimates from previous study, the expected sensitivity and specificity are defined as 85% and 95%, respectively. The estimate's precision is set to 0.07 with a 95% confidence level (two-tailed) for sensitivity and 0.05 with a 95% confidence level (two-tailed) for specificity. We estimated the sample size required for this study through the following equations [25, 26],

$$n = \frac{Z_{\alpha}^2 P(1-P)}{d^2} \quad (1)$$

Where P is the expected value of sensitivity (or specificity), $\frac{Z_{\alpha}}{2}$ is set to 1.96 as corresponding with the 95% confidence interval, and d as the precision of the estimate is pre-determined by the clinical judgment of investigators. By the formula, 100 participants are required based on sensitivity and 73 based on the specificity.

2.2 Study design, subjects, and computed tomography

The clinical study was designed and implemented in Henan Cancer Hospital using the diagnostic kit for lung cancer genes methylation (Excellen Medical Technology Co., Ltd.). Clinical status was not determined before blood draw for methylated genes assay, and blood samples were obtained from all subjects who met the selection criteria. All technicians were blinded to the clinical information of subjects.

Samples from 351 subjects (testing study) passed the sample quality control acceptance criterion and were enrolled for analyzing the DNA methylation (Fig. 1). Among 351 subjects, 154 were diagnosed with lung cancer (LC), including 82 adenocarcinomas, 40 squamous cell carcinomas, 14 small cell lung cancer, and 18 unclassified lung cancer patients. The other 197 cases were controls, including benign lung diseases such as pulmonary fibrosis, infection, sarcoidosis, and bronchiectasis, etc. Nine patients were malignancies in other systems, including one thyroid carcinoma, one osteosarcoma, and seven esophageal cancer. The classification of all conditions was based on diagnosis from computed tomography and subsequent pathological examinations. LC patients were divided into four subgroups based on the TNM guidelines classification criteria [27]. These LC cases covered all major histological types and a broad range of stages. The median age for LC cases and controls was 60 and 53 years, respectively. Subjects with lung cancer smoked more than controls (40 vs. 20 pack-years). All subjects underwent a blood draw before the pathological examination, and subsequent biopsy or surgery was performed. None of the subjects received chemotherapy, radiotherapy, or surgical intervention before the blood draw and computed tomography. More details of these population characteristics (testing study only) were described in Table 1.

Table 1
Characteristics of the patient population

	Total n (%)	Cases n (%)	Controls n (%)
Age (Years)	351 (100)	154 (100)	197 (100)
≤ 50 Years	114 (32.5)	24 (15.6)	90 (45.7)
51–60	109 (31.1)	56 (36.4)	53 (26.9)
61–70	108 (30.8)	62 (40.3)	46 (23.4)
> 70	20 (5.7)	12 (7.8)	8 (4.1)
Median Age	56	60	53
Age Range	17–78	36–78	17–76
Gender			
Male	228 (65.0)	102 (66.2)	126 (64.0)
Female	123 (35.0)	52 (33.8)	71 (36.0)
Smoking status			
Nonsmoker	246 (70.1)	71 (46.1)	175 (88.8)
Exsmoker	37 (10.5)	29 (18.8)	8 (4.1)
Current smoker	68 (19.4)	54 (35.1)	14 (7.1)
Smoking pack years			
Range Packs/Years (Smokers only)	0-120	0-120	0–60
Median Packs/Years (Smoker only)	30	40	20
Mean Packs/Years (Smokers only)	37	41	21
Histology subtype			
Adenocarcinoma	–	82 (53.2)	–
Squamous cell carcinoma	–	40 (26.0)	–
Small cell lung cancer	–	14 (9.1)	–
Other	–	18(11.7)	–
healthy	–	–	43 (21.8)
Benign lung disease	–	–	145 (73.6)
Malignancies in other systems	–	–	9 (4.6)

	Total n (%)	Cases n (%)	Controls n (%)
Stage			
I	–	40 (26.0)	–
II	–	21 (13.6)	–
III	–	46 (29.9)	–
IV	–	44 (28.6)	–
Unknown	–	3 (1.9)	–
Benign lung diseases including pulmonary infection, sarcoidosis and bronchiectasis etc. Malignancies in other systems including thyroid carcinoma, osteosarcoma and esophageal cancer.			

2.3 Sample collection and storage

Samples were collected from outpatients and inpatients, and the sample information was recorded in sample collection forms. A 5 ml peripheral blood sample was collected with 5 ml K₂EDTA anticoagulant tubes (BD biosciences, Franklin Lakes, NJ, USA) to ensure the assay's accuracy. Sample storage and transportation followed the instructions for the use of the Nucleic Acid Extraction Reagent (Excellen Medical Technology Co., Ltd.).

2.4 DNA isolation and bisulfite conversion

Plasma was prepared within 4 hours after the blood draw. After collecting the blood sample, it was centrifuged at 1500 g for 10 min; then, the supernatant was transferred to a 15 mL test tube and recentrifuged at 1500 g for 10 min. The plasma stored at -80°C and avoided repeated freeze-thawing. DNA was obtained from plasma using the Nucleic Acid Extraction Reagent (Excellen Medical Technology Co., Ltd.) according to the respective instructions for use. Briefly, circulating DNA from 2 mL of plasma was digested in 3 mL lysis buffer for 20 min, followed by extracted utilizing magnetic particles. Thereafter, the DNA was converted in a bisulfite reaction, which converts unmethylated cytosine residues to uracil residues. After purification, the bisulfite-converted DNA (bisDNA) was eluted in 35 µL and ready for use in real-time PCR.

2.5 DNA methylation analysis

Bisulfite-modified DNA was used as a template for fluorescence-based Real-Time PCR according to the instructions of the diagnostic kit for methylated gene of lung cancer (Excellen Medical Technology Co., Ltd.). In brief, amplification reactions were carried out in triplicate in a volume of 25 µL that contained 12.5 µL Reaction Buffer, 2.5 µL of Primer Mixture, and 10 µL of bisulfite-modified DNA. Amplifications were carried out using the following profile: 98°C for 5 minutes, followed by 45 cycles at 95°C for 10 seconds and 63°C for 5 seconds to 58°C for 30 seconds. Amplification reactions were carried out in 96-

well plates in an Applied Biosystems 7500 Fast Real-Time PCR Systems (Applied Biosystems). The data was analyzed by Applied Biosystems 7500 Fast Real-Time PCR System Sequence Detection Software v1.4.1. Each plate included patient DNA samples: positive controls (in vitro methylated leukocyte DNA), negative controls (normal leukocyte DNA or DNA from a known unmethylated cell line), and multiple water blanks. All samples were within the range of the assay of sensitivity and reproducibility based on the amplification of the internal reference standard [threshold cycle (Ct) value of ACTB].

In the PCR reaction, the primers and probes can discriminate between methylated and unmethylated DNA sequences, and the methylated sequences are preferentially amplified. Hydrolysis probes specific for the methylated target sequence used in the reaction can exclusively identify methylated target sequences during the PCR amplified reaction, and can be detected on fluorescence channels FAM, HEX, Texas Red and CY5.

The final result of Lung Cancer Analysis was determined utilizing a composite score based on a subject's methylation. The scoring algorithm was calculated by multiplying a subject's methylation results by a constant marker specific weighting factor. The aggregate of these individually weighted marker results determines the composite score (range of 0–10), which was then compared to a cut-off (2.0) to determine a positive or negative result. The final result was positive if the composite score is greater than or equal to 2.0. The final result was negative if the composite score is less than 2.0. All calculations were performed using the Lung Cancer Analysis software v2.2 (Excellen Medical Technology Co., Ltd.). Real-Time PCR data were imported into the software with appropriate settings for assay thresholds and baseline windows, resulting in a cycle threshold (Ct) value per assay for each PCR well.

2.6 Statistical analysis

The performance of the assay was reported by using sensitivity and specificity. Sensitivity was defined as the ratio of correctly assigned positive lung cancer samples in all lung cancer samples. Specificity was defined as the ratio of correctly assigned negative samples in all normal/benign lung and other cancer samples. Positive predictive value (PPV; the probability that the disease is given a positive test result) and negative predictive value (NPV; the probability that the disease is absent given a negative test result) were also calculated. The receiver operating characteristic (ROC) and the area under the ROC curve (AUC) were analyzed using the IBM SPSS Statistics 24. The association of test positivity with demographic characteristics were assessed using the Chi-square test or Fisher exact test when less than 5 individuals were observed. Statistical significance was defined as $P < 0.05$.

3. Results

3.1 Receiver operating characteristic curve analysis and lung cancer diagnostic accuracy

We tested quantitative analysis of promoter methylation in the plasma DNA samples from 154 lung cancer (LC) patients and 197 control subjects using a diagnostic kit for the methylated gene of lung

cancer (Excellen Medical Technology Co., Ltd.) based on Real-Time PCR (RT-PCR). The LC cases covered all major histological types and a broad range of stages and are confirmed with pathological results. More detailed information of the population characteristics is described in methods. We performed receiver operating characteristic (ROC) curve analysis to analyze the diagnostic efficacy of the *SHOX2*, *RASSF1A* and, *PTGER4* methylation in blood plasma. As shown in Fig. 2 and Table 2, the methylation analysis of *SHOX2*, *RASSF1A*, and *PTGER4* panel showed a high diagnostic sensitivity of 87.0% and 98.0% specificity. PPV and NPV are two other necessary measures of diagnostic accuracy. They are related to sensitivity and specificity through disease prevalence (Π). The PPV and NPV values were calculated with the formulas:

$$PPV = \frac{Sensitivity * \Pi}{Sensitivity * \Pi + (1 - specificity) * (1 - \Pi)} \quad (2)$$

$$NPV = \frac{Specificity * (1 - \Pi)}{(Specificity * [1 - \Pi] + [1 - sensitivity] * \Pi)} \quad (3)$$

In high-risk population (e.g., lung cancer prevalence of 2.4% [28]), the PPV is then 51.7%. Correspondingly, the NPV is 99.7%. NPV and PPV values show that the *SHOX2*, *RASSF1A*, and *PTGER4* methylation detection of plasma could be an effective complementary tool in high-risk lung cancer diagnosis.

Table 2

The consistency of RT-PCR and pathological examination in detecting aberrant methylation of the *SHOX2*, *RASSF1A*, and *PTGER4* gene

PCR (<i>SHOX2</i> + <i>RASSF1A</i> + <i>PTGER4</i>)	Pathological examination		Total
	Positive	Negative	
Positive	134	4	138
Negative	20	193	213
Total	154	197	351
AUC	0.938		
95%CI	0.907–0.961		
Kappa	0.86		
Sensitivity	87.0%		
Specificity	98.0%		
Positive predictive value (PPV)	51.7% (cancer prevalence of 2.4%)		
Negative predictive value (NPV)	99.7% (cancer prevalence of 2.4%)		
Abbreviations: <i>SHOX2</i> = short stature homeobox 2 gene; <i>RASSF1A</i> = ras association domain family 1A gene; <i>PTGER4</i> = prostaglandin E receptor 4 gene; AUC = area under the curve			

In addition, we also performed a short-term follow-up test on 28 subjects, of which 26 were lung cancer cases and two benign tumor cases. Blood methylation PCR was performed once before surgery (Included in Table 2) and was retaken five days after the operation (Not included in Table 2). The results showed that 26 cases of lung cancer samples were tested positive by PCR before the operation and negative after the operation. Two cases of benign tumor samples were detected negative before and after the operation. These results indicate that the three methylated gene markers of *SHOX2*, *RASSF1A*, and *PTGER4* originate from lung cancer tissue rather than from healthy tissue or benign tumors.

3.2 Subtype analysis of *SHOX2*, *RASSF1A* and *PTGER4* methylation in plasma

The performance of the *SHOX2*, *RASSF1A* and, *PTGER4* methylation in blood plasma was analyzed with respect to the histological subtypes, including adenocarcinoma, squamous cell carcinoma, small cell lung cancer, large cell carcinoma and other as shown in Fig. 3A and Table 3. The data showed that the sensitivity of the *SHOX2*, *RASSF1A* and *PTGER4* methylation panel in blood plasma was 87.0% in lung cancer group. In a more detailed analysis, the *SHOX2*, *RASSF1A* and, *PTGER4* methylation panel showed 89.0% in adenocarcinoma, 87.5% sensitivity in squamous cell carcinoma, 85.7% in small cell lung cancer and 77.8% sensitivity in other lung cancer, respectively, which means the *SHOX2*, *RASSF1A* and *PTGER4* methylation panel in plasma is a noninvasive biomarker with wide application in almost all histological subtypes of lung cancer. We further evaluated the sensitivity of the *SHOX2*, *RASSF1A*, and *PTGER4* methylation panel in plasma in different tumor stages (Fig. 3B and Table 3). The results showed that the methylation analysis of *SHOX2*, *RASSF1A*, and *PTGER4* in plasma showed high diagnostic ability with a sensitivity of 87.0%. Notably, the *SHOX2*, *RASSF1A*, and *PTGER4* methylation panel in plasma achieved a high sensitivity of 82.5% in stage I and 90.5% in stage II lung cancer patients, respectively.

Table 3
Subtype analysis

Tumor Stages	Histology				
	Adenocarcinoma	Squamous Cell Carcinoma	Small Cell Lung Cancer	Other	All
I	28/33 (84.8%)	3/4 (75%)	0/0	2/3 (66.7%)	33/40 (82.5%)
II	12/14 (85.7%)	2/2 (100%)	1/1 (100%)	4/4 (100%)	19/21 (90.5%)
III	12/12 (100%)	21/23 (91.3%)	6/7 (85.7%)	2/4 (66.7%)	42/46 (91.3%)
IV	21/23 (91.3%)	8/10 (80%)	5/6 (83.3%)	4/5 (80.0%)	38/44 (86.4%)
Unknown	0/0	1/1 (100%)	0/0	2/2 (100%)	3/3 (100%)
All	73/82 (89.0%)	35/40 (87.5%)	12/14 (85.7%)	14/18 (77.8%)	134/154 (87.0%)
Stage and histology-specific performance of the <i>SHOX2</i> , <i>RASSF1A</i> and <i>PTGER4</i> methylation biomarkers using plasma samples from patients with suspected lung cancer.					

To identify possible biases in various demographic characteristics of the population, an overall performance analysis of the 3-gene panel signature was undertaken, including age, gender, and smoking history. Table 4 shows the details of this analysis. In age analysis, sensitivity was above 80% for most age groups, and specificity for LC was consistently high across all age groups. In gender analysis, sensitivity and specificity are similar among males and females. As shown in Table 4, Lung cancer group sensitivity/specificity was 86.3%/96.8% for males, compared with 88.5%/100% for females. In addition, no significant sensitivity/specificity differences were observed among different smoking groups.

Table 4
Overall performance of the panel in diverse demographics characteristics of the population

Clinical characteristics	Number of specimens	% Sensitivity	% Specificity	Chi-square test P value
Diagnosis				
Lung cancer	154	87.0		
healthy	43		97.7	
Benign lung disease	145		98.6	
Other nonlung cancer	9		88.9	0.131
Age				
≤ 50 Years	114	79.2	98.9	
51–60	109	91.1	96.2	
61–70	108	85.5	97.8	
> 70	20	91.7	100	0.665
Gender				
Male	228	86.3	96.8	
Female	123	88.5	100	0.285
Smoking status				
None	246	87.3	98.3	
Former	37	86.2	100	
Current	68	87.0	92.9	0.082

4. Discussion

Low-dose spiral computed tomography (LDCT) can be a reliable screening tool for detecting lung cancer in a high-risk population, which decreases the mortality rate of lung cancer by 20% [5]. However, a severe limitation of LDCT is its low specificity [5, 29], and the rate of surgical resections for the benign disease remains too high (6%- 38%) [30]. DNA methylation has been shown to play an important role in carcinogenesis at an early stage [31–33], which makes DNA methylation alterations promising candidates in cancer biomarker research. Besides, methylation assays are relatively easy to set up, perform, and automate. This study demonstrated that the assessment of DNA methylation markers in blood plasma provides a novel noninvasive method to increase lung cancer diagnostic accuracy.

The 3-gene methylation panel in our study demonstrated high discriminatory power to differentiate patients with LC from healthy subjects and patients with nonmalignant diseases of the lung (AUC = 0.938). This panel was validated in a study incorporating patients with various nonmalignant lung diseases such as pulmonary sarcoidosis and bronchiectasis into the control group. The diagnostic efficiency was not influenced by the presence of benign lung disease and other (nonlung) cancers in the control population. In addition, this discrimination is independent of age, gender, and even smoking status, suggesting that it detects cancer-specific alterations rather than tobacco-related field cancerization. The favorable discrimination ability of the panel in our study indicates that it is probably an applicable and feasible method for the detection and diagnosis of LC.

The combined detection of *SHOX2*, *RASSF1A* and *PTGER4* gene methylation in plasma improved detection sensitivity in lung cancer diagnosis compared to previous studies [20, 23, 34, 35], and achieved high sensitivity across all histological subtypes of lung cancer. Chrisoph et al. performed DNA methylation analysis of the *SHOX2* gene in blood showed a sensitivity of 60% and specificity of 90% in the diagnosis of lung cancer [20]. Zhang et al. detected methylation changes in 322 subjects with both Sanger sequencing and the Methylated Human *SHOX2* and *RASSF1A* Gene Detection Kit (Tellgen Co. Ltd., Shanghai China). The data showed that the *SHOX2* and *RASSF1A* methylation panel in bronchoalveolar lavage fluid (BALF) yielded a diagnostic sensitivity of 81.0% and specificity of 97.4% with an AUC value of 0.892 in lung cancer diagnosis, but it was showed a lower sensitivity in adenocarcinoma of 69.6% [35]. In our study, the methylation panel achieved high diagnostic with a total sensitivity of 87.0% including all histological subtype groups. In subtype analysis, the sensitivity was above 85% for most histological subtypes, especially in adenocarcinoma, the sensitivity detected of 87.5% was much better than previous studies in bronchoalveolar lavage fluid (BALF) samples [35].

In addition, the methylation panel achieved a relatively high sensitivity of 82.5% and 90.5% in stage I and stage II patients, respectively, which demonstrated that the panel detection has potential clinical value in the early diagnosis of lung cancer. The exceptional high mortality rate of lung cancer can be attributed to a high degree of late diagnosis [2]. The 5-year survival rate of lung cancer is only 15–19% for all stages. However, outcomes are significantly better in patients diagnosed at an early stage; the 5-year survival rate could increase to 50–60% [2, 36], especially for stage I ranging from 81–85% [37]. Thus, it is imperative to detect lung cancer at earlier stages. In our study, the 3-gene methylation panel showed good sensitivity in the early diagnosis of lung cancer, indicating that the panel may be used as an early screening tool for lung cancer, thereby increasing lung cancer survival rates.

However, there exist some limitations to our study. *SHOX2* hypermethylation has been reported to occur frequently in a variety of tumors plasma, including lung cancer, head and neck squamous cell carcinoma (HNSCC), renal cell carcinoma (RCC), and colorectal cancer [38]. We realize, therefore, while the combined detection of *SHOX2*, *RASSF1A*, and *PTGER4* gene in plasma achieved a high detection specificity in lung cancer diagnosis, the exact number of patients in other different tumors may be insufficient; to confirm the results, more clinical data would be needed in further studies. Secondly, the validation of this method is limited to plasma free DNA and EDTA anticoagulant blood collection tubes. Other sample types and

blood collection methods have not been performed on the effectiveness. Besides, the free DNA content in plasma samples is low and easily degraded. The samples should be handled and stored strictly according to the requirements; otherwise, it will affect the detection and cause false-negative results.

5. Conclusions

The methylation analysis of *SHOX2*, *RASSF1A*, and *PTGER4* panel in plasma showed an efficient diagnostic ability in lung cancer diagnosis. This novel noninvasive test of three biomarkers has potential clinical value. It could be used in stand-alone or in combination with current imaging detection methods to improve the over-all diagnosis of lung cancer.

Declarations

Acknowledgment

We thank all the subjects who participated in this study.

Funding

This work was supported by Project Construction of Precision Medical Data Platform on Lung Cancer and other Common Malignancy. (Supported by Major Science and Technology Projects of Henan, 161100311500).

Conflict of interest

FW, YC, and ML are employees of Excellen Medical Technology Co., Ltd. All other authors have no conflicts of interest to declare.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request. Subjects' data are not publicly available due to them containing information that could compromise participant consent and confidentiality.

Authors' contributions

JM conceived and designed the paper. BW, CY, CZ, and DW performed the majority of experiments and analyzed data. WX, HS, and XC collected and processed subject samples. FW, YC, and ML provided technical support. All authors read and approved the final manuscript.

References

1. Key Statistics for Lung Cancer. <https://www.cancer.org/cancer/non-small-cell-lung4/cancer/about/key-statistics.html>. Accessed August 9, 2019.

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7-30. doi: 10.3322/caac.21590.
3. Manser R, Lethaby A, Irving LB, Stone C, Byrnes G, Abramson MJ, Campbell D. Screening for lung cancer. *Cochrane Database Syst Rev*. 2013;(6):CD001991. doi: 10.1002/14651858.CD001991.pub3.
4. Marcus PM. Lung cancer screening: an update. *J Clin Oncol*. 2001;19(18 Suppl):83S-86S.
5. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5):395-409. doi: 10.1056/NEJMoa1102873.
6. National Lung Screening Trial Research Team, Church TR, Black WC, Aberle DR, Berg CD, Clingan KL, Duan F, Fagerstrom RM, Gareen IF, Gierada DS, Jones GC, Mahon I, Marcus PM, Sicks JD, Jain A, Baum S. Results of initial low-dose computed tomographic screening for lung cancer. *N Engl J Med*. 2013;368(21):1980-1991. doi: 10.1056/NEJMoa1209120.
7. Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, Pio R, Zulueta JJ, Spira A, Massion PP, Mazzone PJ, Montuenga LM. Biomarkers in Lung Cancer Screening: Achievements, Promises, and Challenges. *J Thorac Oncol*. 2019;14(3):343-357. doi: 10.1016/j.jtho.2018.11.023.
8. Hassanein M, Callison JC, Callaway-Lane C, Aldrich MC, Grogan EL, Massion PP. The state of molecular biomarkers for the early detection of lung cancer. *Cancer Prev Res (Phila)*. 2012;5(8):992-1006. doi: 10.1158/1940-6207.CAPR-11-0441.
9. Silvestri GA, Vachani A, Whitney D, Elashoff M, Porta Smith K, Ferguson JS, Parsons E, Mitra N, Brody J, Lenburg ME, Spira A; AEGIS Study Team. A Bronchial Genomic Classifier for the Diagnostic Evaluation of Lung Cancer. *N Engl J Med*. 2015;373(3):243-251. doi: 10.1056/NEJMoa1504601.
10. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 2008;455(7216):1069-1075. doi: 10.1038/nature07423.
11. Ni X, Zhuo M, Su Z, Duan J, Gao Y, Wang Z, Zong C, Bai H, Chapman AR, Zhao J, Xu L, An T, Ma Q, Wang Y, Wu M, Sun Y, Wang S, Li Z, Yang X, Yong J, Su XD, Lu Y, Bai F, Xie XS, Wang J. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci U S A*. 2013;110(52):21083-8. doi: 10.1073/pnas.1320659110.
12. Belinsky SA. Gene-promoter hypermethylation as a biomarker in lung cancer. *Nat Rev Cancer*. 2004;4(9):707-717. doi: 10.1038/nrc1432.
13. Xiao T, Ying W, Li L, Hu Z, Ma Y, Jiao L, Ma J, Cai Y, Lin D, Guo S, Han N, Di X, Li M, Zhang D, Su K, Yuan J, Zheng H, Gao M, He J, Shi S, Li W, Xu N, Zhang H, Liu Y, Zhang K, Gao Y, Qian X, Cheng S. An approach to studying lung cancer-related proteins in human blood. *Mol Cell Proteomics*. 2005;4(10):1480-1486. doi: 10.1074/mcp.M500055-MCP200.
14. Shen N, Du J, Zhou H, Chen N, Pan Y, Hoheisel JD, Jiang Z, Xiao L, Tao Y, Mo X. A Diagnostic Panel of DNA Methylation Biomarkers for Lung Adenocarcinoma. *Front Oncol*. 2019;9:1281. doi: 10.3389/fonc.2019.01281.

15. Jones PA, Baylin SB. The epigenomics of cancer. *Cell*. 2007;128(4):683-692. doi: 10.1016/j.cell.2007.01.029.
16. Hulbert A, Jusue-Torres I, Stark A, Chen C, Rodgers K, Lee B, Griffin C, Yang A, Huang P, Wrangle J, Belinsky SA, Wang TH, Yang SC, Baylin SB, Brock MV, Herman JG. 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.
17. Food and Drug Administration. Department of Health and Human Services. Letter confirming approval of application for premarket approval of Colorguard.
http://www.accessdata.fda.gov/cdrh_docs/pdf13/p130017a.pdf. Accessed May 24, 2016.
18. Food and Drug Administration. Department of Health and Human Services. Letter confirming approval of application for premarket approval of Epi proColon test.
http://www.accessdata.fda.gov/cdrh_docs/pdf13/p130001a.pdf. Accessed May 24, 2016.
19. Schmidt B, Liebenberg V, Dietrich D, Schlegel T, Kneip C, Seegebarth A, Flemming N, Seemann S, Distler J, Lewin J, Tetzner R, Weickmann S, Wille U, Liloglou T, Raji O, Walshaw M, Fleischhacker M, Witt C, Field JK. SHOX2 DNA methylation is a biomarker for the diagnosis of lung cancer based on bronchial aspirates. *BMC Cancer*. 2010;10:600. doi:10.1186/1471-2407-10-600.
20. Kneip C, Schmidt B, Seegebarth A, Weickmann S, Fleischhacker M, Liebenberg V, Field JK, Dietrich D. 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.
21. Chen H, Suzuki M, Nakamura Y, Ohira M, Ando S, Iida T, Nakajima T, Nakagawara A, Kimura H. Aberrant methylation of RASGRF2 and RASSF1A in human non-small cell lung cancer. *Oncol Rep*. 2006;15(5):1281-1285.
22. Helmbold P, Lahtz C, Herpel E, Schnabel PA, Dammann RH. Frequent hypermethylation of RASSF1A tumour suppressor gene promoter and presence of Merkel cell polyomavirus in small cell lung cancer. *Eur J Cancer*. 2009;45(12):2207-2211. doi:10.1016/j.ejca.2009.04.038.
23. Weiss G, Schlegel A, Kottwitz D, König T, Tetzner R. Validation of the SHOX2/PTGER4 DNA Methylation Marker Panel for Plasma-Based Discrimination between Patients with Malignant and Nonmalignant Lung Disease. *J Thorac Oncol*. 2017;12(1):77-84. doi: 10.1016/j.jtho.2016.08.123.
24. Dochez V, Caillon H, Vaucel E, Dimet J, Winer N, Ducarme G. Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review. *J Ovarian Res*. 2019;12(1):28. doi:10.1186/s13048-019-0503-7.
25. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. *J Biomed Inform*. 2014;48:193-204. doi:10.1016/j.jbi.2014.02.013.
26. Negida A, Fahim NK, Negida Y. Sample Size Calculation Guide - Part 4: How to Calculate the Sample Size for a Diagnostic Test Accuracy Study based on Sensitivity, Specificity, and the Area Under the ROC Curve. *Adv J Emerg Med*. 2019;3(3):e33. Published 2019 May 19. doi:10.22114/ajem.v0i0.158.
27. Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, Cheney RT, Chirieac LR, D'Amico TA, Demmy TL, Dilling TJ, Govindan R, Grannis FW Jr, Horn L, Jahan TM, Komaki R, Kris MG,

- Krug LM, Lackner RP, Lanuti M, Lilenbaum R, Lin J, Loo BW Jr, Martins R, Otterson GA, Patel JD, Pisters KM, Reckamp K, Riely GJ, Rohren E, Schild S, Shapiro TA, Swanson SJ, Tauer K, Yang SC, Gregory K, Hughes M. Non-small cell lung cancer, version 1.2015. *J Natl Compr Canc Netw*. 2014;12:1738-1761. doi: 10.6004/jnccn.2014.0176
28. Chapman CJ, Healey GF, Murray A, Boyle P, Robertson C, Peek LJ, MacDonald IK. EarlyCDT®-Lung test: improved clinical utility through additional autoantibody assays. *Tumour Biol*. 2012;33(5):1319-1326. doi:10.1007/s13277-012-0379-2.
 29. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, Chaturvedi AK, Silvestri GA, Riley TL, Commins J, Berg CD. Selection criteria for lung-cancer screening. *N Engl J Med*. 2013;368(8):728-736. doi: 10.1056/NEJMoa1211776.
 30. Blanchon T, Bréchot JM, Grenier PA, Ferretti GR, Lemarié E, Milleron B, Chagué D, Laurent F, Martinet Y, Beigelman-Aubry C, Blanchon F, Revel MP, Friard S, Rémy-Jardin M, Vasile M, Santelmo N, Lecalier A, Lefébure P, Moro-Sibilot D, Breton JL, Carette MF, Brambilla C, Fournel F, Kieffer A, Fria G, Flahault A; Dépiscan Group. Baseline results of the Depiscan study: a French randomized pilot trial of lung cancer screening comparing low dose CT scan (LDCT) and chest X-ray (CXR). *Lung Cancer*. 2007;58(1):50-58. doi: 10.1016/j.lungcan.2007.05.009
 31. Locke WJ, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KYC, Ross JP. DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Front Genet*. 2019;10:1150. doi: 10.3389/fgene.2019.01150.
 32. Fukushige S, Horii A. DNA methylation in cancer: a gene silencing mechanism and the clinical potential of its biomarkers. *Tohoku J Exp Med*. 2013;229(3):173-185. doi: 10.1620/tjem.229.173
 33. Klutstein M, Nejman D, Greenfield R, Cedar H. DNA Methylation in Cancer and Aging. *Cancer Res*. 2016;76(12):3446-3450. doi: 10.1158/0008-5472.CAN-15-3278.
 34. Nikolaidis G, Raji OY, Markopoulou S, Gosney JR, Bryan J, Warburton C, Walshaw M, Sheard J, Field JK, Liloglou T. DNA methylation biomarkers offer improved diagnostic efficiency in lung cancer. *Cancer Res*. 2012;72(22):5692-5701. doi:10.1158/0008-5472.CAN-12-2309
 35. Zhang C, Yu W, Wang L, Zhao M, Guo Q, Lv S, Hu X, Lou J. DNA Methylation Analysis of the SHOX2 and RASSF1A Panel in Bronchoalveolar Lavage Fluid for Lung Cancer Diagnosis. *J Cancer*. 2017;8(17):3585-3591. doi: 10.7150/jca.21368.
 36. Begum S, Brait M, Dasgupta S, Ostrow KL, Zahurak M, Carvalho AL, Califano JA, Goodman SN, Westra WH, Hoque MO, Sidransky D. An epigenetic marker panel for detection of lung cancer using cell-free serum DNA. *Clin Cancer Res*. 2011;17(13):4494-503. doi: 10.1158/1078-0432.CCR-10-3436.
 37. Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C. Progress and prospects of early detection in lung cancer. *Open Biol*. 2017;7(9). pii: 170070. doi: 10.1098/rsob.170070.
 38. de Vos L, Jung M, Koerber RM, Bawden EG, Holderried TAW, Dietrich J, Bootz F, Brossart P, Kristiansen G, Dietrich D. Treatment Response Monitoring in Patients with Advanced Malignancies Using Cell-Free SHOX2 and SEPT9 DNA Methylation in Blood: An Observational Prospective Study. *J Mol Diagn*. 2020 Apr 30:S1525-1578(20)30299-3. doi: 10.1016/j.jmoldx.2020.04.205.

Figures

Figure 1

Sample disposition, study setup, and polymerase chain reaction (PCR) assay formats. Boxes in the bottom line indicate number of valid results and number of PCR replicates per PCR assay. For more details, see materials and methods section. ACTB, actin, beta gene; SHOX2, short stature homeobox 2 gene; RASSF1A, ras association domain family 1A gene; PTGER4, prostaglandin E receptor 4 gene.

对角段由绑定值生成。

Figure 2

Receiver operating characteristic (ROC) and area under the curve (AUC) analysis of the short stature homeobox 2 gene (SHOX2), ras association domain family 1A gene (RASSF1A), and prostaglandin E receptor 4 gene (PTGER4) methylation panel (AUC=0.938).

Figure 3

Histological subtypes (A) and tumor stages (B) of the patient population.