

RESEARCH

Open Access



Ultrasensitive DNA hypermethylation detection using plasma for early detection of NSCLC: a study in Chinese patients with very small nodules

Chen Chen^{1†}, Xiaojie Huang^{2†}, Wei Yin¹, Muyun Peng¹, Fang Wu³, Xia Wu⁴, Jingqun Tang¹, Mingjiu Chen¹, Xiang Wang¹, Alicia Hulbert⁵, Malcolm V. Brock⁶, Wenliang Liu^{1*}, James G. Herman^{7*} and Fenglei Yu^{1*}

Abstract

Purpose: We had previously developed highly sensitive DNA methylation detection to diagnose lung cancer in patients with pulmonary nodules. To validate this approach and determine clinical utility in Chinese patients with indeterminate pulmonary nodules, we assessed the diagnostic accuracy for early stage lung cancer in plasma samples.

Experimental design: Patients with CT-detected small lung nodules (diameter \leq 3.0 cm) were included. Cases ($n = 163$) had staged IA or IB non-small cell lung cancer (NSCLC), while controls ($n = 83$) had non-cancerous lesions. Promoter methylation of eight lung cancer-specific genes (CDO1, TAC1, SOX17, HOXA7, HOXA9, GATA4, GATA5, and PAX5) was detected using nanoparticle-based DNA extraction (MOB) followed by qMSP.

Results: Methylation detection for CDO1, TAC1, SOX17, and HOXA7 in plasma was significantly higher in cases compared with the benign group ($p < 0.001$). The sensitivity and specificity for lung cancer diagnosis using individual gene was 41–69% and 49–82%. A three-gene combination of the best individual genes has sensitivity and specificity of 90% and 71%, with area under the receiver operating curve (AUC) of 0.88, (95% CI 0.84–0.93). Furthermore, three-gene combinations detected even the smallest lung nodules, with the combination of CDO1, SOX17, and HOXA7 having the overall best performance, while the combination of CDO1, TAC1, and SOX17 was best in tumor sizes less than 1.0 cm.

Conclusions: Using modified MOB-qMSP, high sensitivity and specificity, for the detection of circulating tumor DNA was obtained for early stage NSCLC. This strategy has great potential to identify patients at high risk and improve the diagnosis of lung cancer at an earlier stage.

Keywords: Lung cancer, Biomarker, DNA methylation, Plasma, Early detection

Introduction

Lung cancer remains the leading cause of cancer-related death in the world [1, 2]. The increasing adoption of lung cancer screening has resulted in a rapidly growing

number of patients with small lung nodules requiring management. Most of these nodules are benign but are hard to distinguish from small lung cancers. Thus, the benefits of low-dose computed tomography (CT) screening have been limited due to its low positive predictive value and high false positive rates [3–5]. Distinguishing small malignant nodules in CT scan from benign ones is particularly challenging because of their ambiguous radiographic characteristics [6, 7]. There are intense efforts to improve nodule management by identifying molecular biomarkers, including gene mutation [8–10],

* Correspondence: liuwenliang@csu.edu.cn; hermanj3@upmc.edu; yufenglei@csu.edu.cn

[†]Chen Chen and Xiaojie Huang contributed equally to this work.

¹Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China

⁷UPMC Hillman Cancer Center, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Full list of author information is available at the end of the article



DNA methylation [11, 12], circulating tumor cells [10], and cancer-specific auto-antibodies [13], that could improve the diagnostic accuracy for detection of early stage of lung cancer.

Circulating tumor DNA (ctDNA), which refers to DNA released from lysed or apoptotic tumor cells and circulating freely in blood, is a promising approach for the detection of very early stage of lung cancer [14, 15]. Most previous studies focused on identifying particular gene mutations in ctDNA from lung cancer patients, such as EGFR, p53, and KRAS, as these are frequently mutated in lung cancers and some have implications in targeted therapies [8]. However, the diagnostic sensitivity for these targets is sharply decreased if the tumors were small or lack these mutations [16]. Methylation of cytosine in CpG islands silences hundreds of genes that are involved in the initiation and progression of lung cancer. Several studies have reported cancer-specific DNA methylation changes detectable in plasma, sputum, saliva, and pleural effusions from the patients with lung cancer [9, 11, 17].

Our previous case-control study reported that, using MOB-qMSP, an ultrasensitive DNA methylation detection approach using superparamagnetic nano-beads followed by quantitative methylation-specific real-time PCR, highly prevalent lung cancer-specific gene methylation could be detected in the plasma and sputum from patients with early-stage lung cancer [18–20]. In the present study, we optimized MOB-qMSP and validated its diagnostic utility in a Chinese patient cohort of indeterminate pulmonary nodules, all 3 cm or less in size.

Materials and methods

Study population

From December 2016 to April 2018, 345 patients with small indeterminate lung nodules (< 3 cm on CT scan) that were suspected to be non-small cell lung cancer (NSCLC) were included in this study. Patients who received any pretreatment therapy, including chemotherapy or radiotherapy, or had a history of malignancy were not included. All patients received curative-intent resection. All blood samples were obtained prior to surgery and were immediately processed to isolate plasma. Paired tumor samples were collected immediately after the tumor was removed and stored at -80°C . All patients with pathologically confirmed malignant lesions were staged according to the revised TNM guidelines classification criteria [21]. Patients with NSCLC were included as cancer group, those with histologically benign lesions as the control group. Plasma samples of 20 healthy volunteers (clinical characteristics provided in Supplemental Table S1) were also obtained as a normal group. Nodule size was obtained from the pathologic report. A summary of clinical and pathological characteristics of patients and healthy volunteers included in this

study are presented in Table 1 and Supplemental Table S1.

Among the 345 patients, 34 patients were excluded due to failed ctDNA extraction mainly due to hemolysis or insufficient ctDNA yield. In this study, we used β -Actin (ACTB) as an internal control, with a CT value of ACTB more than 34 indicated insufficient ctDNA yield [18]. Additionally, pathology review revealed 3 cases of small cell carcinoma, 5 carcinoid, 4 unclassified carcinomas, and 53 patients with regional or mediastinal lymph node metastasis, and these patients were excluded from the analysis. In total, 163 patients in the cancer group with early-stage node-negative tumors ($T_{1-2}N_0M_0$) and 83 patients in the control group with benign lung nodules had samples adequate for analysis. This study was approved by the Human Ethics Committee of the Second Xiangya Hospital, Central South University. Written informed consent was obtained from all patients.

DNA isolation and methylation analysis

DNA isolation and bisulfite conversion

DNA extraction from plasma and fresh frozen tissue samples were performed using our previously described MOB approach [18–20], a process that allows DNA extraction and bisulfite conversion in a single tube via the use of silica super magnetic beads. We have optimized the protocol for plasma by adding 4 mL of plasma to 800 μL of proteinase K (800 units/mL, Invitrogen, USA) and 4 mL of Buffer AL (Qiagen, USA), and incubating them together at the same temperature (55°C overnight). After digestion, 4 mL of isopropyl alcohol (IPA) and 200 μL of beads were added. For DNA extraction from fresh frozen tissue sample, 2–5 mg of tissue sample were added to 40–60 μL of proteinase K and 300 μL of Buffer AL. After incubating overnight, 300 μL of IPA and 150 μL of beads were added. Then, the lysate was incubated and rotated for 10 min before adding 5 μL of carrier RNA, and incubating for an additional 5 min. The bisulfite conversion was performed using a thermal cycler, the optimized incubation temperature and time were programmed as showed in Supplemental Table S2.

DNA methylation analysis

The DNA methylation analysis was performed using quantitative real-time methylation-specific PCR and normalized to a control β -Actin (ACTB) assay, as previously described [18]. Amplification reactions were performed using ABI StepOnePlus Real-Time PCR system (Applied Bio.) with all samples being analyzed in triplicate. Thermo cycling conditions were optimized as follows: 95°C for 5 min, 50 cycles at 95°C for 15 seconds, and 60°C for 30 seconds and 72°C for 30 seconds. The primer and hybridization probe sequences for MOB-qMSP analysis are listed in Supplemental Table S3.

Table 1 Clinical characteristics of the patients

Patient characteristics	Cancer group (n = 163)	Control group (n = 83)	p value
Age (year, mean ± SD)	58.79 ± 9.11	52.45 ± 7.27	0.01
Gender (%)			
Male	84 (51.5%)	52 (62.7%)	0.32
Female	79 (48.5%)	31 (37.3%)	
Stage no. (%)			
T1N0	102 (62.6%)	N/A	N/A
T2N0	61 (37.4%)	N/A	
Histologic characteristics no. (%)			
Adenocarcinoma	139 (85.3%)	N/A	N/A
Squamous-cell	22 (13.5%)	N/A	
NOS	2 (1.2%)	N/A	
Granuloma	N/A	35 (42.2%)	
Hamartoma	N/A	10 (12.0%)	
Inflammation	N/A	29 (34.9%)	
Fungal infection	N/A	2 (2.4%)	
Other benign	N/A	7 (8.5%)	
Nodule size (cm, mean ± SD)	1.78 ± 0.67	1.64 ± 0.73	0.33
0–1.0 cm	28 (17.2%)	22 (26.5%)	0.22
1.1–2.0 cm	92 (56.4%)	43 (51.8%)	
2.1–3.0 cm	43 (26.4%)	18 (21.7%)	
Pack-year (IQR)	38 (10–150)	35 (0.75–90)	0.57
COPD no. (%)	2 (1.2%)	2 (2.4%)	0.49
FEV1 (L, mean ± SD)	2.32 ± 0.79	2.65 ± 0.70	0.29
FVC (L, mean ± SD)	3.00 ± 1.02	3.42 ± 0.86	0.23
FEV1/FVC (% , mean ± SD)	73.52 ± 18.53	76.76 ± 14.31	0.11

COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in one second; FVC, forced vital capacity; IQR, interquartile range

As described previously [18], the $2^{-\Delta Ct}$ was calculated for each methylation detection replicate comparing with the mean Ct for ACTB. For replicates which were not detected (ND), a Ct value of 100 was used, creating a near zero value for $2^{-\Delta Ct}$. The mean $2^{-\Delta Ct}$ value was calculated with the formula:

$$\mu 2^{-\Delta Ct} = \frac{(2^{-\Delta Ct \text{ replicate 1}} + 2^{-\Delta Ct \text{ replicate 2}} + 2^{-\Delta Ct \text{ replicate 3}})}{3}$$

Statistical analysis

Demographic and methylation variables were summarized by case-control status with percentages for categorical variables and means and standard deviations for continuous variables. Differences in demographic variables between cases and controls were assessed with Fisher's exact test for categorical variables and the Wilcoxon rank sum test for continuous variables. The association between the methylation and case-control status was expressed as odds ratios and their corresponding 95% confidence intervals

(CI) obtained from logistic regression models with adjustment for the variables including age, gender, smoking status, pack-years of smoking, COPD, and pulmonary function test results. The pack-years of cigarette smoking were defined as the average number of packs smoked per day times the number of years smoked.

To determine the performance of each individual gene, receiver operating curve (ROC) analysis was performed using the $2^{-\Delta Ct}$ values. The area under the curve (AUC) was reported with 95% confidence intervals. On the basis of ROC curves, the three best-performing genes were selected for combined detection in analyses for diagnostic accuracy for lung cancer detection. The best-performing genes as gene panels were also analyzed combined with clinical characteristics [11, 18].

Results

Clinical characteristics of the patients

A total of 246 patients met inclusion criteria, with 163 patients with lung cancer and 83 with non-cancerous lung lesions as the benign group. All lesions were no

larger than 3.0 cm in greatest dimension verified by pathological reports. All cancer cases were histologically confirmed to be NSCLC with negative lymph nodes. According to the 8th edition of National Comprehensive Cancer Network Guidelines for the TNM classification for lung cancer, all of the NSCLC cases were T₁₋₂N₀M₀ (stage Ia-Ib) non-small cell lung cancer (some tumors were T2 based on visceral pleural invasion and not size). Clinical and demographic variables showed no differences between the cancer and benign groups except for age (Table 1). Plasma samples of 20 healthy volunteers were also obtained as a normal group (Supplemental Table S1).

Detection of DNA methylation in plasma and tumor samples

Methylation of these eight genes from plasma and tumor samples was detected using modified MOB-qMSP approach. In plasma samples, compared with cancer and benign group, the healthy group had the lowest methylation rate in all the eight genes ($p < 0.001$). The

methylation detection rate of CDO1, TAC1, SOX17, and HOXA7 were significantly higher in cancer group than in the benign group ($p < 0.001$) (Fig. 1). We first determined the diagnostic sensitivity and specificity according to the presence or absence of detectable methylation, without considering quantitation of DNA methylation (Table 2) [18]. The sensitivity and specificity for lung cancer diagnosis using individual genes from plasma ranged from 41 to 69% and 49 to 82%, respectively, with the best-performing genes being those previously studied. The newly examined genes did not perform as well as these loci. The eight gene methylation status in tumor tissues were also detected using modified MOB-qMSP. Consistent with DNA methylation profiles in plasma, methylation of CDO1, TAC1, SOX17, and HOXA7 were detected more frequently in patients with cancer compared with controls (Supplemental Figure S1).

Gene methylation and lung cancer diagnostic accuracy

We then generated ROC curves for each gene using normalized methylation $2^{-\Delta\text{CT}}$ calculated as described

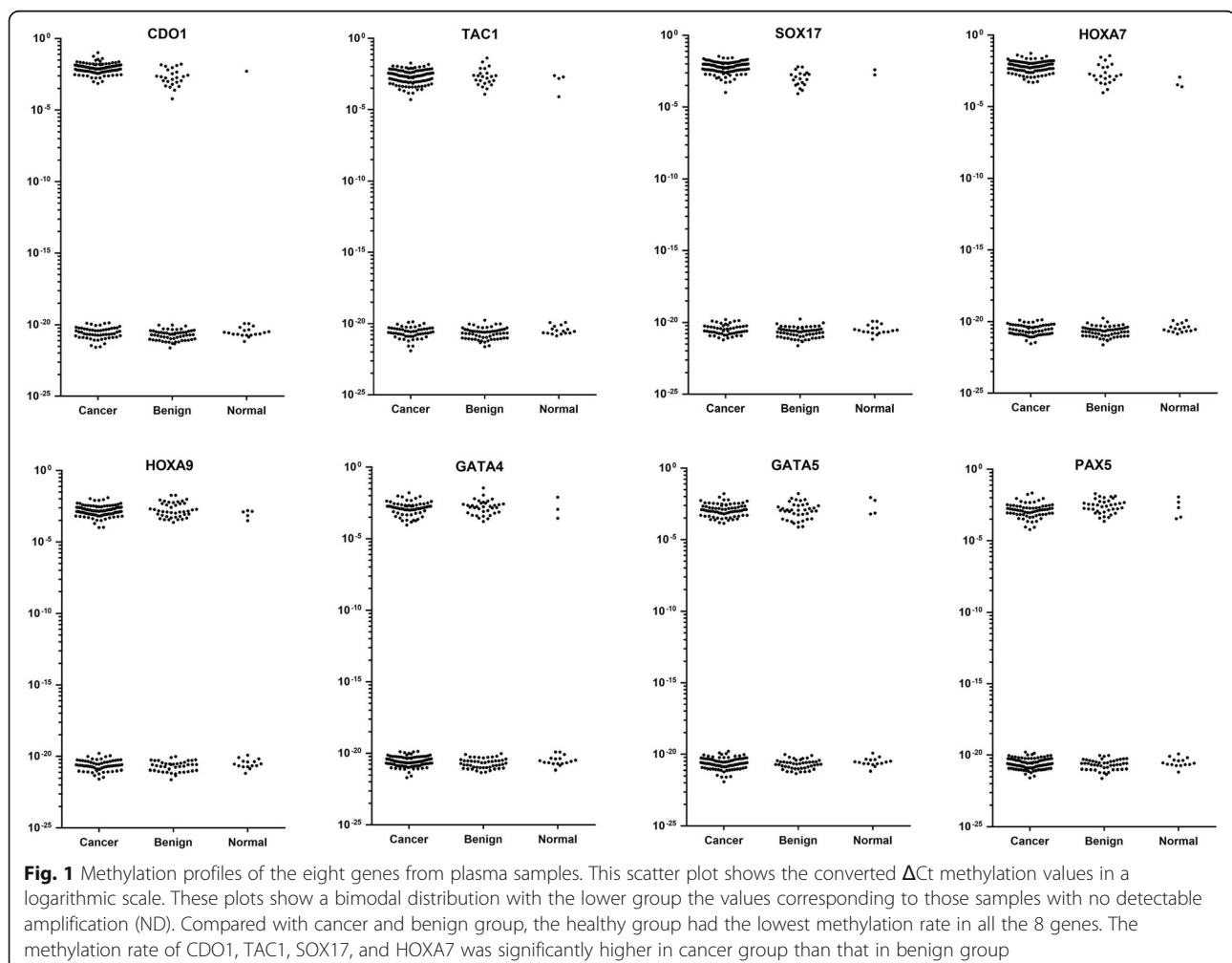


Table 2 Gene methylation detection in plasma samples

Gene	Cancer (n = 163)		Control (n = 83)		PPV	NPV
	n	Sensitivity	n	Specificity		
CDO1	106	65%	17	80%	86%	54%
TAC1	110	67%	26	69%	81%	52%
SOX17	113	69%	15	82%	88%	58%
HOXA7	98	60%	15	82%	87%	51%
HOXA9	101	62%	42	49%	71%	40%
GATA4	68	42%	35	58%	66%	34%
GATA5	72	44%	38	54%	65%	33%
PAX5	67	41%	37	55%	64%	32%

previously [18]. At the best quantitative cutoff, the sensitivity and specificity for lung cancer diagnosis from each single methylated gene ranged from 41 to 68% and from 49 to 87, respectively (Table 3). These sensitivities and specificities were similar to that obtained from the absolute presence of detectable methylation for each gene (Table 2).

The genes with the largest AUC value were as follows: CDO1: AUC 0.78, 95% CI 0.71–0.83; TAC1: AUC 0.71, 95% CI 0.64–0.78; SOX17: AUC 0.82, 95% CI 0.76–0.87; and HOXA7: AUC 0.73, 95% CI 0.67–0.80. The PPV and NPV for these genes were for CDO1, 88% and 53%; for TAC1, 81% and 52%; for SOX17, 90% and 57%; and for HOXA7, 89% and 50%, respectively (Table 3, Supplemental Figure S2). As elevated odds ratios ranged from 2.81 to 7.19, logistic regressions analyses also indicated that the methylation of CDO1, TAC1, SOX17, and HOXA7 were closely related to increasing lung cancer risk (Fig. 2).

We further evaluated the combination of the three best-performing genes. The combination of CDO1, TAC1, and SOX17, which was evaluated as the best combination in our previous study, was examined in the current Chinese cohort. The sensitivity and specificity were 89% and 61% respectively, with AUC of 0.85 (95%

CI, 0.81–0.91). However, different from our previous publication, the combination of best three genes, CDO1, SOX17, and HOXA7, showed the best sensitivity and specificity of 90% and 71%, with a corresponding ROC AUC of 0.88 (95% CI 0.84–0.93; Table 3, Fig. 3a, b, c). This improved performance for this three-gene combination appears to be from a higher prevalence/sensitivity of HOXA7 in this cohort, compared with that previously reported [18], with maintained high specificity. With an adjusted odds ratio of 22.59, (95% CI 11.21–45.53), the combination of CDO1, SOX17, and HOXA7 showed the best association with the diagnosis of early stage NSCLC (Fig. 2). Furthermore, we investigated clinical features (age, pack-year, chronic obstructive pulmonary disease (COPD) status, nodule size, and pulmonary function values) for diagnostic accuracy as clinical predictors alone and in combination with gene panels. Clinical predictors alone had a diagnostic accuracy AUC of 0.70 (95% CI, 0.65–0.79) (Fig. 3d). The best-gene panel combined with the clinical predictors improved the diagnostic accuracy with an AUC of 0.94 (95% CI, 0.91–0.96), slightly better than CDO1, TAC1, and SOX17 plus clinical predictors (Table 4, Fig. 3d, e, f).

Diagnostic accuracy according to tumor size

As the quantity of circulating tumor DNA for detection is related to stage and tumor volume [22], we explored the diagnostic accuracy of the combination of best three genes (CDO1, SOX17, and HOXA7) according to tumor size. In tumors with greatest diameter of 2.1–3 cm, the sensitivity and specificity were 91% and 90%, with the AUC of 0.95, (95% CI 0.90–1.00); in 1.1–2.0 cm tumors, the sensitivity and specificity were 74% and 93%, and the AUC was 0.92, (95% CI 0.87–0.96); in tumors less than 1.0 cm, the sensitivity and specificity decreased slightly to 64% and 82%, with a corresponding AUC of 0.75, (95% CI 0.62–0.89) (Tables 5, 6 and 7). Interestingly, while the combination of CDO1, SOX17, and HOXA7 showed the overall best performance in our patient

Table 3 Sensitivity, specificity, PPV, and NPV at optimal cutoffs with AUC

Gene	Sensitivity	Specificity	PPV	NPV	AUC	95% CI
CDO1	63%	83%	88%	53%	0.78	0.71–0.83
TAC1	68%	70%	81%	52%	0.71	0.64–0.78
SOX17	68%	86%	90%	57%	0.82	0.76–0.87
HOXA7	55%	87%	89%	50%	0.73	0.67–0.80
HOXA9	64%	49%	71%	41%	0.56	0.48–0.64
GATA4	44%	58%	67%	35%	0.53	0.45–0.61
GATA5	43%	63%	70%	36%	0.52	0.44–0.60
PAX5	41%	55%	64%	32%	0.54	0.45–0.62
CDO1, TAC1, SOX17	89%	61%	82%	74%	0.85	0.81–0.91
CDO1, SOX17, HOXA7	90%	71%	86%	78%	0.88	0.84–0.93

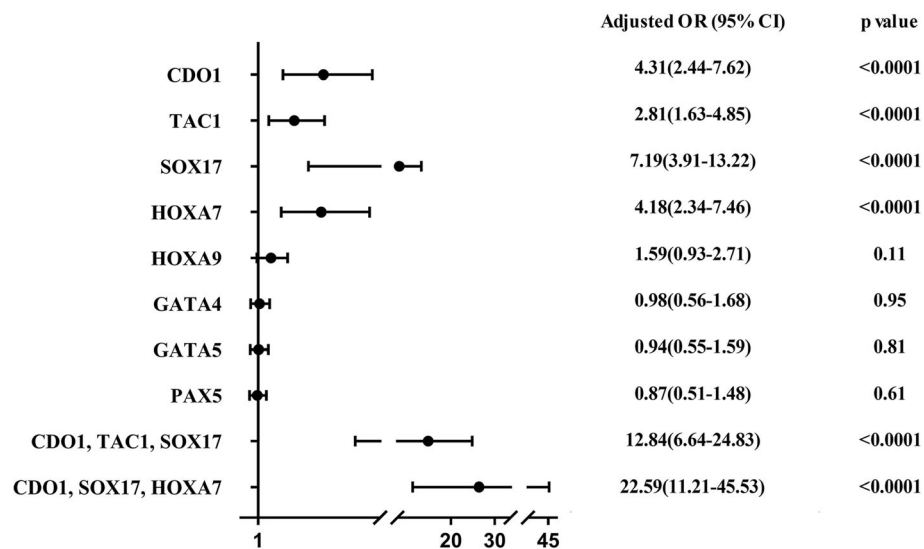


Fig. 2 Performance of gene methylation as predictor for lung cancer. The logistic regression analyses indicated that the methylation of CDO1, TAC1, SOX17, and HOXA7 were closely related to increasing of lung cancer risk. With the best adjusted odds ratio, the combination of CDO1, SOX17, and HOXA7 showed the best performance in the diagnosis of lung cancer

cohort, the combination of CDO1, TAC1, and SOX17 had the best performance in the subgroup with tumor size less than 1.0 cm (Tables 5, 6 and 7).

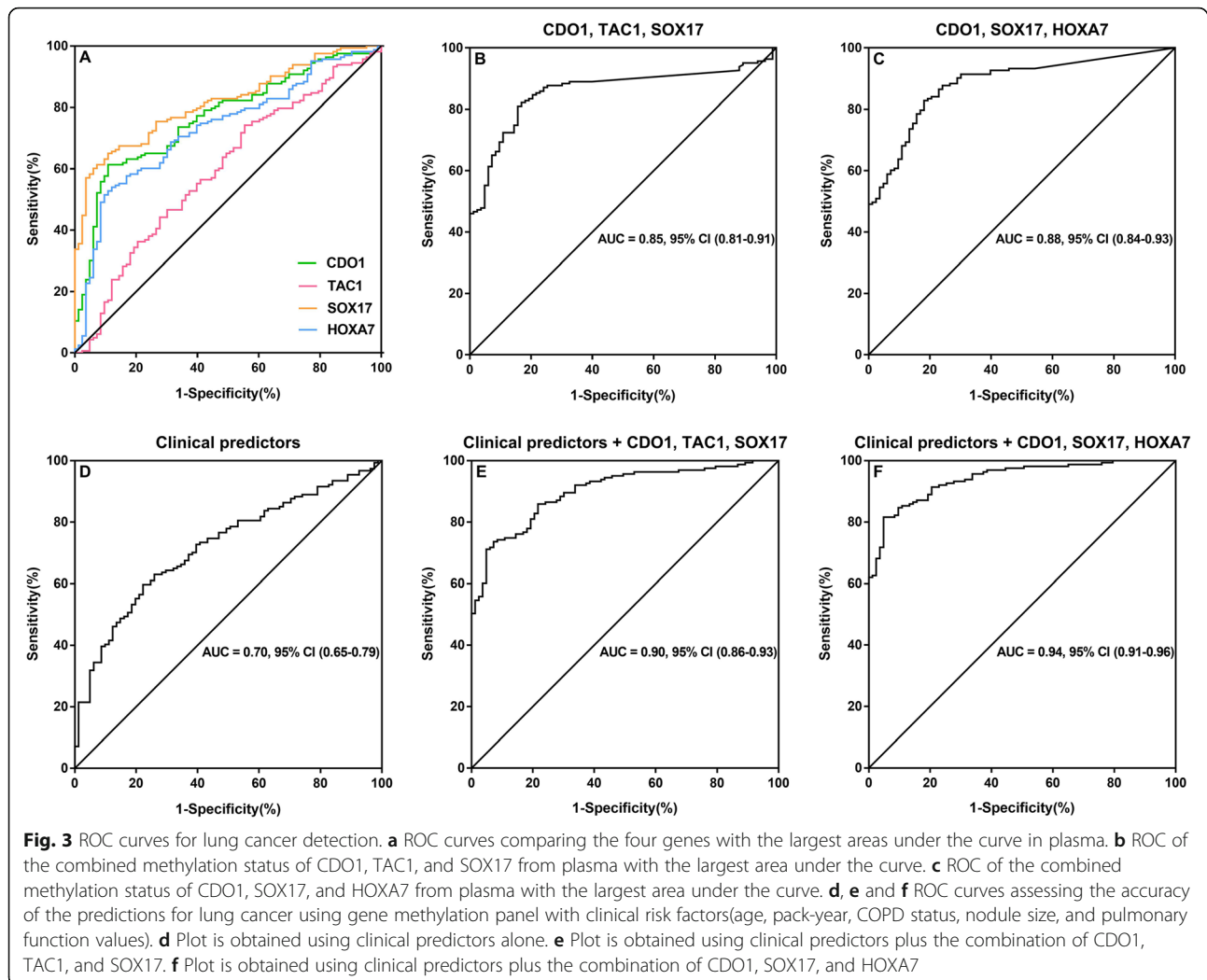
Discussion

In this study, using modified MOB-qMSP, we investigated the detection of promoter hypermethylation of eight genes and one internal control gene in plasma and tumor samples of patients with small lung nodules. This study is a corroboration of our previous study [18], but now examined in a Chinese cohort, suggesting that these detection biomarkers are useful in divergent populations. Although our previous study had demonstrated the high diagnostic sensitivity and specificity of promoter methylation of CDO1, TAC1, HOXA7, HOXA9, and SOX17 in plasma from patients with NSCLC in a Lung Cancer Specialized Program of Research Excellence (SPORE) patient cohort [18, 23], the performance and diagnostic accuracy of these biomarkers still needed validation in another cohort, and might be affected by differences between races, environmental carcinogenic exposure, and smoking status.

In the present study, we evaluated the performance of individual gene biomarkers for the early detection of lung cancer (Tables 2 and 3). This confirmed the utility of CDO1, TAC1, HOXA7, and SOX17, while newly tested genes were not as effective for lung cancer detection. While each gene can detect lung tumor DNA from many patients, given the rarity of these molecules in ctDNA, individual gene sensitivity is somewhat limited. However, by combining the best-performing genes in a panel, we greatly improve the diagnostic sensitivity

without a substantial decline in specificity. From the initial screening of eight genes, four genes were selected as candidates for a panel to distinguish early stage NSCLC from benign lung nodules, which had previously also provided the best cancer sensitivity and specificity. The possible 3 gene combinations (CDO1, TAC1, and SOX17; or CDO1, SOX17, and HOXA7) were able to provide a very high sensitivity ranging from 89 to 90% and a specificity ranging from 61 to 71%. Due to the fact that the variations of DNA methylation patterns in individuals might depend on the alterations of different molecular pathways, the use of a multiple gene panel may provide greater utility for detecting different tumors when compared with a single gene [24]. However, the primary benefit of multigene detection is likely from the additional chances this allows to detect rare ctDNA molecules in plasma samples.

ctDNA in plasma can carry abnormal cellular alterations related to cancer. Several studies have sought to improve early detection of lung cancer by investigating molecular biomarkers in the plasma [25, 26]. However, none of these tests have been widely used in practice due to unsatisfactory sensitivities and specificities. Our previous study demonstrated the high diagnostic accuracy for early-stage lung cancer using a panel of methylated promoter genes in the plasma based on MOB-qMSP strategy [18]. The promoter hypermethylation of CDO1, TAC1, HOXA7, and SOX17 were detected more frequently in the plasma of cancer patients compared with controls. The combination of CDO1, TAC1, and SOX17 showed the highest sensitivity and specificity (93% and 62%). In the present study, higher methylation



frequencies of CDO1, TAC1, HOXA7, and SOX17 were also observed in the plasma of cancer patients. The combination of CDO1, TAC1, and SOX17 showed similar sensitivity and specificity (89% and 61%) which was consistent with our previous results. However, differing from our previous study, the combination of CDO1, SOX17, and HOXA7 showed even better sensitivity and specificity (90% and 71%) in the Chinese cohort, indicating better diagnostic accuracy, which appears to be from the higher prevalence of HOXA7 methylation detection in this population. The performance of these gene methylation markers remains their superiority in

differentiating lung cancer from benign lung nodules, suggesting the potential clinical application for the diagnosis of early stage lung cancer.

Capitalizing on the strengths of highly prevalent DNA methylation biomarkers and ultra-sensitive techniques to detect DNA methylation could facilitate early diagnosis of lung cancer with indeterminate screen-detected pulmonary nodules [27, 28]. Recently, Liang and colleagues reported the high diagnostic accuracy by using high-throughput DNA bisulfite sequencing in tissue and plasma samples from patients with lung cancer [29]. Other studies have also sought to differentiate lung

Table 4 Performance of gene methylation panel in the prediction of early stage lung cancer

Gene	Sensitivity	Specificity	PPV	NPV	AUC	95% CI
Clinical predictors alone	81%	47%	75%	56%	0.70	0.65–0.79
Clinical predictors + CDO1, TAC1, SOX17	93%	63%	83%	82%	0.90	0.86–0.93
Clinical Predictors + CDO1, SOX17, HOXA7	91%	79%	89%	82%	0.94	0.91–0.96

AUC, area under the curve (in the ROC curves); 95% CI, 95% confidence interval

Table 5 Sensitivity, specificity, PPV, and NPV at optimal cutoffs with AUC regarding tumor size (2.1–3 cm)

Gene	Sensitivity	Specificity	PPV	NPV	AUC	95% CI
CDO1	67%	89%	94%	53%	0.77	0.64–0.90
TAC1	65%	72%	85%	46%	0.67	0.51–0.83
SOX17	67%	90%	94%	53%	0.79	0.68–0.90
HOXA7	70%	83%	91%	54%	0.77	0.65–0.89
CDO1, TAC1, SOX17	88%	89%	95%	76%	0.91	0.83–0.99
CDO1, SOX17, HOXA7	91%	90%	96%	81%	0.95	0.90–1.00

Cancer, *n* = 43; control, *n* = 18

cancers from benign lung nodules by investigating ctDNA markers [30–32]. However, these studies mainly used next-generation sequencing technologies which are more costly due to the depth of sequencing required and also needs extensive bioinformatics analyses, which reduce the ease of clinical application. Easy, efficient, and cost-effective detection of regional DNA methylation could reduce testing costs and has enormous potential clinical application.

In our series of studies, we employed a newly developed method, methylation on beads (MOB), which permits DNA extraction and bisulfite conversion in a single-tube cellular processing by using superparamagnetic nano-beads as a DNA carrier. As previously described, it yields on average 1.5- to 5-fold improvement in extraction and conversion efficiency compared with traditional column-based technique, with an even greater improvement in detection sensitivity [19, 20, 33]. Followed by real-time quantitative methylation-specific PCR (qMSP), DNA methylation signals from body fluids, such as plasma and sputum, could be easily and efficiently detected by this assay [18]. In the present study, we slightly modified this MOB-qMSP protocol, by shortening the bisulfite converting time but at a higher reaction temperature (Supplemental Table S2), which could minimize DNA damage during bisulfite conversion and increase the efficiency of the detection of ctDNA methylation.

The demographic characteristics of the cancer group and control group in this study differed slightly (Table 1). Similar to our previous study, methylation detection of these genes was not associated with gender, smoking

status, and pulmonary function. One factor that might relate to diagnostic accuracy was tumor size. Several studies have reported lower detection rates with decreasing tumor size, especially for biomarkers detected in plasma [14, 34]. In the previous study based on mainly American patients, the diagnostic accuracy of single gene or gene panels of combined genes showed no differences among subgroups of different sizes, but few tumors were sub-centimeter in size. In the present study, the gene panel CDO1, SOX17, and HOXA7 showed the best diagnostic accuracy in the subgroup with tumor size 2.1–3 cm and decreased in subgroups with smaller tumor size, especially when the tumor size was less than 1 cm. On the other hand, in the subgroup with tumor size 0–1 cm, the gene panel CDO1, TAC1, and SOX17 (sensitivity 71%, specificity 82%, and AUC 0.81, respectively) was slightly better than gene panel CDO1, SOX17, and HOXA7 (64%, 82%, and 0.75, respectively), indicating this panel may have better sensitivity and diagnostic accuracy in the detection of very small lung cancer lesions (Tables 5, 6 and 7). While the sensitivity for tumors < 1 cm is slightly lower than larger tumors, previous studies have not been able to detect this tumors even this well, and in most cases did not attempt to detect such early-stage lung cancer [29, 35, 36].

Promoter region hypermethylation is an important mechanism of gene silencing involved in several physiological and pathological processes, especially in the initiation and progression of cancer [37, 38]. Several studies have reported that the presence of promoter hypermethylation of tumor suppressor genes could be observed in control populations as a random or a physiologic

Table 6 Sensitivity, Specificity, PPV, and NPV at optimal cutoffs with AUC regarding tumor size (1.1–2 cm)

Gene	Sensitivity	Specificity	PPV	NPV	AUC	95% CI
CDO1	72%	88%	93%	59%	0.80	0.73–0.88
TAC1	70%	70%	83%	52%	0.69	0.60–0.79
SOX17	65%	93%	95%	55%	0.82	0.75–0.89
HOXA7	61%	77%	85%	48%	0.72	0.63–0.81
CDO1, TAC1, SOX17	73%	93%	96%	62%	0.91	0.87–0.97
CDO1, SOX17, HOXA7	74%	93%	96%	63%	0.92	0.87–0.96

Cancer, *n* = 92; control, *n* = 43

Table 7 Sensitivity, Specificity, PPV, and NPV at optimal cutoffs with AUC regarding tumor size (0–1 cm)

Gene	Sensitivity	Specificity	PPV	NPV	AUC	95% CI
CDO1	82%	46%	66%	67%	0.64	0.49–0.80
TAC1	61%	78%	78%	61%	0.68	0.56–0.81
SOX17	61%	77%	77%	61%	0.68	0.56–0.80
HOXA7	82%	59%	72%	72%	0.73	0.60–0.87
CDO1, TAC1, SOX17	71%	82%	83%	69%	0.81	0.69–0.93
CDO1, SOX17, HOXA7	64%	82%	82%	64%	0.75	0.62–0.89

Cancer, *n* = 28; control, *n* = 22

event related to age, smoking status, or environmental carcinogenic exposures, which decreased the diagnostic accuracy of cancer when using DNA methylation as a biomarker [39, 40]. This consideration brings up the necessity of validation of the performance of DNA methylation biomarkers within different populations, and emphasizes the importance that the control group should be of similar age and exposure. While in the present study, DNA methylation of single genes could be detected in plasma of some patients with benign lung nodules, and more rarely in healthy controls, there remains a high sensitivity and specificity in this screen-detected lung nodule cohort. This suggests this approach has diagnostic accuracy and suggests the potential for clinical application in the evaluation of screen-detected nodules.

The National Lung Screening Trial (NLST) reported a sensitivity of 93% in baseline scans but a false positive rate of 26.3% (specificity of 74%) and a false discovery rate of 96%. This would make our plasma-based detection to have a similar performance to CT detection. However, a direct comparison is not possible, since we only examined nodules < 3 cm, and it is likely our sensitivity would be greater if larger nodules were included. The primary utility would not be an alternative to CT screening, but as a complimentary test to enhance the management of detected nodules. With this level of sensitivity and specificity, a simple non-invasive test could potentially reduce the need for invasive confirmation tests needed to establish or rule out the diagnosis of cancer in patients with indeterminate pulmonary nodules detected by CT screening.

Conclusion

Taken together, this study demonstrates that, with modified MOB-qMSP assay, detection of early stage NSCLC with high sensitivity and specificity could be obtained using a panel of methylated promoter genes in plasma, even extending this detection to sub-centimeter nodules. This gene panel and detection strategy have great potential for an adjunct to CT screening, identifying patients at high risk for lung cancer, reducing false positive

results, and improving the diagnosis of lung cancer at an earlier stage.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13148-020-00828-2>.

Additional file 1. Supplemental Figure S1. Methylation profiles of the 8 genes from tissue samples. Consistent with DNA methylation profiles in plasma, methylation of CDO1, TAC1, SOX17, and HOXA7 were detected more frequently in patients with cancer compared with benign controls. Supplemental Figure S2. Receiver operator classification curves for lung cancer detection for the 8 genes obtained from Plasma. Supplemental Table S2. Bisulfite conversion thermal cycler conditions. Supplemental Table S3. Primers and probes of qMSP for plasma and tissue samples.

Acknowledgements

Not applicable.

Authors' Contributions

CC, WL, JH, and FY contributed to the conception and design. CC, XH, WY, AH, and JH contributed to the development of the methodology. CC, XH, FW, XWu, MP, JT, MC, and XWa contributed to the acquisition of the data. CC, XH, WY, MP, MB, and JH contributed to the analysis and interpretation of data. CC, XH, AH, WL, JH, and FY contributed to the writing, review, and/or revision of the manuscript. XH, MP, XW, FW, and MB contributed to administrative, technical, or material support. CC, XH, WL, JH and FY contributed to the study supervision. All authors read and approved the final manuscript.

Funding

This study was funded by the China National Natural Science Foundation no.81000905 and no.81972638 (to C Chen and WL Liu), Hunan Provincial Natural Science Foundation no.14JJ4014 and Hunan Provincial Key Area R&D Program no. 2019SK2253 (to C Chen, XJ Huang, MY Peng, WL Liu, and FL Yu), CSCO Y-young2019-034 (to C Chen), and NIH/NCI U01CA214165 (to JG Herman).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This study was approved by the Human Ethics Committee of the Second Xiangya Hospital, Central South University. Written informed consent was obtained from all patients.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. ²Department of Cardiovascular Surgery, The Second Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. ³Department of Oncology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. ⁴Department of Pathology, The Second Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China. ⁵Department of Surgery, University of Illinois at Chicago School of Medicine, Chicago, IL, USA. ⁶Department of Surgery, The Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁷UPMC Hillman Cancer Center, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.

Received: 15 November 2019 Accepted: 11 February 2020

Published online: 05 March 2020

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019; 69(1):7–34 PubMed PMID: 30620402.
- Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. 2016;66(2):115–32 PubMed PMID: 26808342.
- The Lancet Respiratory M. Lung cancer screening in Europe: hurdles to overcome. *Lancet Respir Med*. 2018;6(12):885 PubMed PMID: 30629929.
- Ruano-Ravina A, Perez-Rios M, Casan-Clara P, Provencio-Pulla M. Low-dose CT for lung cancer screening. *Lancet Oncol*. 2018;19(3):e131–e2 PubMed PMID: 29508756.
- Ten Haaf K, van der Aalst CM, de Koning HJ. Clinically detected non-aggressive lung cancers: implications for overdiagnosis and overtreatment in lung cancer screening. *Thorax*. 2018;73(5):407–8 PubMed PMID: 29440512.
- Liu Y, Wang H, Li Q, McGettigan MJ, Balagurunathan Y, Garcia AL, et al. Radiologic features of small pulmonary nodules and lung cancer risk in the National Lung Screening Trial: a nested case-control study. *Radiology*. 2018; 286(1):298–306 PubMed PMID: 28837413. Pubmed Central PMCID: 5738292.
- Huang P, Park S, Yan R, Lee J, Chu LC, Lin CT, et al. Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study. *Radiology*. 2018; 286(1):286–95 PubMed PMID: 28872442. Pubmed Central PMCID: 5779085.
- Chabon JJ, Simmons AD, Lovejoy AF, Esfahani MS, Newman AM, Haringsma HJ, et al. Circulating tumour DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients. *Nat Commun*. 2016;7:11815 PubMed PMID: 27283993. Pubmed Central PMCID: 4906406.
- Peng M, Chen C, Hulbert A, Brock MV, Yu F. Non-blood circulating tumor DNA detection in cancer. *Oncotarget*. 2017;8(40):69162–73 PubMed PMID: 28978187. Pubmed Central PMCID: 5620327.
- Alix-Panabieres C, Schwarzenbach H, Pantel K. Circulating tumor cells and circulating tumor DNA. *Annu Rev Med*. 2012;63:199–215 PubMed PMID: 22053740.
- Leng S, Wu G, Klinge DM, Thomas CL, Casas E, Picchi MA, et al. Gene methylation biomarkers in sputum as a classifier for lung cancer risk. *Oncotarget*. 2017;8(38):63978–85 PubMed PMID: 28969046. Pubmed Central PMCID: 5609978.
- Belinsky SA, Leng S, Wu G, Thomas CL, Picchi MA, Lee SJ, et al. Gene methylation biomarkers in sputum and plasma as predictors for lung cancer recurrence. *Cancer Prev Res*. 2017;10(11):635–40 PubMed PMID: 28904059. Pubmed Central PMCID: 5668184.
- Ren S, Zhang S, Jiang T, He Y, Ma Z, Cai H, et al. Early detection of lung cancer by using an autoantibody panel in Chinese population. *Oncoimmunology*. 2018;7(2):e1384108 PubMed PMID: 29308305. Pubmed Central PMCID: 5749655.
- Szpechcinski A, Chorostowska-Wynimko J, Struniawski R, Kupis W, Rudzinski P, Langfort R, et al. Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease. *Br J Cancer*. 2015; 113(3):476–83 PubMed PMID: 26125447. Pubmed Central PMCID: 4522634.
- Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct detection of early-stage cancers using circulating tumor DNA. *Sci Translat Med*. 2017;9:403 PubMed PMID: 28814544.
- Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, et al. Biomarkers in lung cancer screening: achievements, promises, and challenges. *J Thorac Oncol*. 2019;14(3):343–57 PubMed PMID: 30529598.
- Shi YX, Wang Y, Li X, Zhang W, Zhou HH, Yin JY, et al. Genome-wide DNA methylation profiling reveals novel epigenetic signatures in squamous cell lung cancer. *BMC Genom*. 2017;18(1):901 PubMed PMID: 29169318. Pubmed Central PMCID: 5701423.
- Hulbert A, Jusue-Torres I, Stark A, Chen C, Rodgers K, Lee B, et al. Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. *Clin Cancer Res*. 2017;23(8):1998–2005 PubMed PMID: 27729459. Pubmed Central PMCID: 6366618.
- Bailey VJ, Zhang Y, Keeley BP, Yin C, Pelosky KL, Brock M, et al. Single-tube analysis of DNA methylation with silica superparamagnetic beads. *Clin Chem*. 2010;56(6):1022–5 PubMed PMID: 20360128. Pubmed Central PMCID: 3000409.
- Keeley B, Stark A, Pisanic TR 2nd, Kwak R, Zhang Y, Wrangle J, et al. Extraction and processing of circulating DNA from large sample volumes using methylation on beads for the detection of rare epigenetic events. *Clin Chim Acta*. 2013;425:169–75 PubMed PMID: 23911908. Pubmed Central PMCID: 3963364.
- Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The Eighth Edition Lung Cancer Stage Classification. *Chest*. 2017;151(1):193–203 PubMed PMID: 27780786.
- Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med*. 2014;20(5):548–54 PubMed PMID: 24705333. Pubmed Central PMCID: 4016134.
- Wrangle J, Machida EO, Danilova L, Hulbert A, Franco N, Zhang W, et al. Functional identification of cancer-specific methylation of CD01, HOXA9, and TAC1 for the diagnosis of lung cancer. *Clin Cancer Res*. 2014;20(7): 1856–64 PubMed PMID: 24486589. Pubmed Central PMCID: 4019442.
- Carvalho AL, Jeronimo C, Kim MM, Henrique R, Zhang Z, Hoque MO, et al. Evaluation of promoter hypermethylation detection in body fluids as a screening/diagnosis tool for head and neck squamous cell carcinoma. *Clin Cancer Res*. 2008;14(1):97–107 PubMed PMID: 18172258.
- Ostrow KL, Hoque MO, Loyo M, Brait M, Greenberg A, Siegfried JM, et al. Molecular analysis of plasma DNA for the early detection of lung cancer by quantitative methylation-specific PCR. *Clin Cancer Res*. 2010;16(13):3463–72 PubMed PMID: 20592015. Pubmed Central PMCID: 2899894.
- Belinsky SA, Klinge DM, Dekker JD, Smith MW, Bocklage TJ, Gilliland FD, et al. Gene promoter methylation in plasma and sputum increases with lung cancer risk. *Clin Cancer Res*. 2005;11(18):6505–11 PubMed PMID: 16166426.
- Koch A, Joosten SC, Feng Z, de Ruijter TC, Draht MX, Melotte V, et al. Analysis of DNA methylation in cancer: location revisited. *Nat Rev Clin Oncol*. 2018;15(7):459–66 PubMed PMID: 29666440.
- Duruiseaux M, Esteller M. Lung cancer epigenetics: from knowledge to applications. *Semin Cancer Biol*. 2018;51:116–28 PubMed PMID: 28919484.
- Liang W, Zhao Y, Huang W, Gao Y, Xu W, Tao J, et al. Non-invasive diagnosis of early-stage lung cancer using high-throughput targeted DNA methylation sequencing of circulating tumor DNA (ctDNA). *Theranostics*. 2019;9(7):2056–70 PubMed PMID: 31037156. Pubmed Central PMCID: 6485294.
- Zhang YC, Zhou Q, Wu YL. The emerging roles of NGS-based liquid biopsy in non-small cell lung cancer. *J Hematol Oncol*. 2017;10(1):167 PubMed PMID: 29061113. Pubmed Central PMCID: 5654124.
- Pecuchet N, Zonta E, Didelot A, Combe P, Thibault C, Gibault L, et al. Base-position error rate analysis of next-generation sequencing applied to circulating tumor DNA in non-small cell lung cancer: a prospective study. *PLoS Med*. 2016;13(12):e1002199 PubMed PMID: 28027313. Pubmed Central PMCID: 5189949 following competing interests: VT received honorarium from Boehringer Ingelheim and is a consultant for RainDance Technologies. HB is an occasional consultant for AstraZeneca, Boehringer Ingelheim and Pfizer. PLP received honorarium from RainDance Technologies, Life Technologies, ThermoFisher Diagnosis, AstraZeneca, and Boehringer Ingelheim. EF received honorarium from Boehringer Ingelheim and Roche. LG received honorarium from Boehringer Ingelheim. The other authors have declared that no competing interests exist.
- Abbosh C, Birkbak NJ, Swanton C. Early stage NSCLC - challenges to implementing ctDNA-based screening and MRD detection. *Nat Rev Clin Oncol*. 2018;15(9):577–86 PubMed PMID: 29968853.
- Wang Z, Kambhampati S, Cheng Y, Ma K, Simsek C, Tieu AH, et al. Methylation biomarker panel performance in EsophaCap Cytology samples for diagnosing Barrett's esophagus: a prospective validation study. *Clin Cancer Res*. 2019;25(7):2127–35 PubMed PMID: 30670490.

34. Chen KZ, Lou F, Yang F, Zhang JB, Ye H, Chen W, et al. Circulating tumor DNA detection in early-stage non-small cell lung cancer patients by targeted sequencing. *Sci Rep.* 2016;6:31985 PubMed PMID: 27555497. Pubmed Central PMCID: 4995492 Valley Biotechnology Inc. in Beijing, China. All other authors declare no conflict of interest.
35. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature.* 2019; PubMed PMID: 31142840.
36. Cree IA, Uttley L, Buckley Woods H, Kikuchi H, Reiman A, Harnan S, et al. The evidence base for circulating tumour DNA blood-based biomarkers for the early detection of cancer: a systematic mapping review. *BMC Cancer.* 2017; 17(1):697 PubMed PMID: 29061138. Pubmed Central PMCID: 5654013.
37. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science.* 2017;357:6348 PubMed PMID: 28729483. Pubmed Central PMCID: 5940341.
38. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet.* 2018;392(10149):777–86 PubMed PMID: 30100054.
39. Vaz M, Hwang SY, Kagiampakis I, Phallen J, Patil A, O'Hagan HM, et al. Chronic cigarette smoke-induced epigenomic changes precede sensitization of bronchial epithelial cells to single-step transformation by KRAS mutations. *Cancer Cell.* 2017;32(3):360–76 e6. PubMed PMID: 28898697. Pubmed Central PMCID: 5596892.
40. van Dongen J, Nivard MG, Willemsen G, Hottenga JJ, Helmer Q, Dolan CV, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat Commun.* 2016;7:11115 PubMed PMID: 27051996. Pubmed Central PMCID: 4820961.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

