Original Research

Autoantibodies against tumor-associated antigens in sputum as biomarkers for lung cancer ☆

Ning Li, Van K. Holden, Janaki Deepak, Nevins W. Todd, Feng Jiang*

Departments of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA

A R T I C L E   I N F O

A B S T R A C T

Tumor antigens (TAs) can initiate host immune responses and produce TA-associated autoantibody (TAAbs), potential cancer biomarkers. Sputum is directly generated from the upper and lower airways, and thus can be used as a surrogate sample for the diagnosis of lung cancer based on molecular analysis. To develop sputum TAAb biomarkers for the early detection of lung cancer, the leading cause of cancer death, we probed a protein microarray containing more than 9,000 antigens with sputum supernatants of a discovery set of 30 lung cancer patients and 30 cancer-free smokers. Twenty-eight TAs with higher reactivity in sputum of lung cancer cases vs. controls were identified. The diagnostic significance of TAAbs against the TAs was determined by enzyme-linked immunosorbent assays (ELISAs) in sputum of the discovery set and additional 166 lung cancer patients and 213 cancer-free smokers (validation set). Three sputum TAAbs against DDX6, ENO1, and 14–3-3ζ were developed as a biomarker panel with 81% sensitivity and 83% specificity for diagnosis of lung cancer, regardless of stages, locations, and histological types of lung tumors. This study provides the first evidence that sputum TAAbs could be used as biomarkers for the early detection of lung cancer.

Introduction

Lung cancer is the leading cause of cancer-related deaths, accounting for 28% of all cancer deaths [1]. 85% of lung tumors are non-small cell lung cancers (NSCLCs) [1], which comprise two major histological subtypes: squamous cell carcinoma (SCC) and adenocarcinoma (AC). Given that the 5-year survival rate depends on the stage of lung cancer, the early detection of NSCLC is clinically important [1, 2]. Efforts have been made to develop tumor cells-derived proteins as circulating cell-free (e.g., serum or plasma) biomarkers for lung cancer. For instance, analysis of CEA, CYFRA-21-2, SFTBP57, EGFR, ProGRP58, and Ciz159 proteins in plasma showed promise for the diagnosis of lung cancer [3, 4]. However, since the circulating tumor proteins that are released from the cancer cells are few and diluted in a large serum or plasma volume, the sensitivity of the circulating cell-free molecular biomarkers in blood is very low, particularly for the early stages of NSCLC [5].

Self-proteins (antigens) of tumors can be affected by specific point mutations, misfolding, overexpression, aberrant glycosylation, truncation, or aberrant degradation [6,7]. The tumor antigens (TAs) elicit immune responses and stimulate TA-associated autoantibodies (TAAbs) before or during tumor formation [8], which may provide biomarkers for the diagnosis of cancer [9]. Compared with the TAs, TAAbs have sev-

eral benefits as biomarkers. First, although the TAs are usually at an undetectable level in body fluids, the TAAbs response to the TAs are often amplified in the immune responses, and hence more easily detectable [7]. Second, as opposed to TAs that are frequently transiently elevated in body fluids, the corresponding TAAbs are relatively lasting, and thus are more stable and readily detected [8]. Third, TAAbs are highly specific and handily detectable in small volumes of specimens with well-established secondary reagents. Indeed, the potential use of determination of TAAbs in serum or plasma for the diagnosis of lung cancer was demonstrated [10]. For example, a serum based-earlyCDT-lung assay that consisted of six TAAbs against p53, NY-ESO-1, CAGE, GBU4–5, Annexin 1, and SOX2 could detect lung cancer with 38% sensitivity and 88% specificity [11]. However, the sensitivity of the serum TAAb biomarkers was not sufficient for the early diagnosis of lung cancer in the clinical settings.

Sputum is coughed up from the lower airways and deep lungs, and hence could act as a mirror to pulmonary diseases [12,13]. Sputum essentially consists of cellular and liquid fractions [14,15]. The cellular portion contains exfoliated airway epithelial cells from the bronchial tree [12, 13]. We and others have shown that genetic and epigenetic analyses of exfoliated respiratory epithelial cells could provide an approach for the early detection of lung cancer [16–20,12,13,21–37]. The

liquid fraction of sputum is complex dilute aqueous solution of lipids, glycoconjugates, and plasma-derived mediators and proteins [12,13]. In contrast with the epithelial cells exfoliated from local respiratory tract sites, *sputum supernatant comprises the liquid portion that* is produced by the cells lining the bronchi and airways where NSCLCs are immersed in the bio-fluid [12,13]. Furthermore, *sputum supernatant* may comprise local- and systemic-derivative lung TAs and the corresponding TAAbs, which raises the rationale for using sputum as a diagnostic medium for detection of NSCLC with TAAb biomarkers. Therefore, we hypothesized that the analysis of TAAbs in sputum may provide an alternative means for the noninvasive diagnosis of lung cancer. The objective of this study was to validate the hypothesis by identifying and developing TAAb biomarkers in sputum for the early detection of NSCLC.

Materials and methods

Study population

The study was performed under a research protocol (HP-40666) approved by the Institutional Review Board of University of Maryland Baltimore. Our enrollment criteria were individuals between the ages of 50–75 who had more than a 30 pack-year smoking history. The clinical diagnosis of lung cancer was made with histopathologic examinations of specimens obtained by CT-guided transthoracic needle biopsy, transbronchial biopsy, videotape-assisted thoracoscopic surgery, or surgical resection. The surgical pathologic staging was determined according to the TNM classification of the International Association for the Study of Union Against Cancer with the American Joint Committee on Cancer and the International Staging System for Lung Cancer. Histopathological classification was determined according to the World Health Organization classification. When tumors were located in the bronchus (cartilage bearing airway), they were classified as central lung cancers. When tumors did not arise from the bronchus and instead were associated with bronchioles, they were classified as peripheral lung cancer. A total of 196 lung cancer patients and 243 cancer-free smokers were enrolled. The cases and controls were divided into two sets. The first one consisted of 30 stage I NSCLC lung cancer patients and 30 cancer-free smokers (Table 1) and used as a discovery set. The second one included 68 stage I NSCLC, 46 stage II, and 52 stage III-IV patients, as well as 213 cancer-free smokers (Table 2) and was used as a validation set. To mitigate the confounding effects, the cases and controls were matched 1:1 by sex, age, cigarette smoking, and chronic obstructive pulmonary disease (COPD) in both sets of specimens. Information of lung tumor stages, locations, and histological types were included in the tables.

Table 1

Characteristics of NSCLC patients and cancer-free smokers in cohort 1*.

	NSCLC cases (n = 30)	Controls (n = 30)
Age	66.32 (SD 11.37)	65.79 (SD 11.33)
Sex		
Female	12	12
Male	18	18
Race		
White	10	10
African American	20	20
Pack-years (median)	36.29	35.78
Stage	all are stage I	
Histological type		
Adenocarcinoma	16	
Squamous cell carcinoma	14	
Locations		
Central tumors	13	
Peripheral tumors	17	

Abbreviations: NSCLC, non-small cell lung cancer.

* All *P*>0.05.

Table 2

Characteristics of NSCLC patients and cancer-free smokers in cohort 2 *.

	NSCLC cases (n = 166)	Controls (n = 213)
Age	67.25 (SD 10.94)	66.18 (SD 11.15)
Sex		
Female	64	82
Male	102	131
Race		
White	122	155
African American	44	58
Pack-years (median)	36.37	35.49
Stage		
Stage I	68	
Stage II	46	
Stage III-IV	52	
Histological type		
Adenocarcinoma	86	
Squamous cell carcinoma	80	
Locations		
Central tumors	78	
Peripheral tumors	88	

Abbreviations: NSCLC, non-small cell lung cancer.

* All *p* > 0.05.

Sputum collection and preparation

We collected sputum from the participants as described in our previous reports [12,13,21–23,25,27, 28,34–39]. Sputum specimens were mucoid and of lower respiratory origin as indicated by the presence of bronchial epithelial cells [12,13,21,35]. Only sputum with a squamous epithelial cell count of <10 cells/high-power field was used in this study [12,35]. Each of the sputum samples was diluted with 3 ml of phosphate buffered saline (PBS) (Thermo Fisher Scientific, Grand Island, NY) per 1 gram of sputum. The diluted samples were passed through an 18-gage needle for mechanical disruption [12], and then centrifuged for 10 min. The supernatant was collected and further centrifuged at 3500 rpm for 20 min and stored at −80 °C until use.

Protein microarray profiling

The ProtoArray human protein microarrays (Invitrogen, Carlsbad CA) were used for profiling TAs in the sputum samples of discovery set (Table 1). Briefly, the microarray slides were treated with the blocking buffer (Invitrogen) at 4 °C for 1 h. We then probed the arrays with a 1:10 dilution of each sample. We incubated the microarrays at 4 °C for 2 h, which were washed five times with PBS buffer. We diluted Alexa Fluor®647-conjugated goat anti-human IgG antibody into probe buffer (Invitrogen) to a 1 μg/mL final concentration, which was added to each array and incubated for 1.5 h at 4 °C. We removed the secondary antibody, and washed the arrays, which were scanned using an Axon GenePix 4000B fluorescent microarray scanner. We analyzed the results by using the Prospector Imager and Analyzer (Invitrogen). We probed the microarrays with goat anti-GST antibody to check the uniformity of the proteins spotted in the protein microarrays followed by incubation with Alexa Fluor 555-labeled antigoat IgG. We also incubated other control microarrays only with the secondary antibody Alexa Fluor 647-labeled anti-human IgG for background determination.

Enzyme-linked immunosorbent assay (ELISA)

We purified recombinant fusion proteins for the selected TAs using UltimateTM ORF Clones (Invitrogen) according to the manufacturer's instructions. We examined expression of recombinant protein in SDS–PAGE and Coomassie blue staining. We used Western blot analysis to confirm if the bands were reactive with reference antibodies [40]. ELISA analysis of TAAbs against the corresponding TAs was carried out using a protocol as previously described [41,42]. Briefly, we coated pu-

Table 3

28 TAs showed significantly differential signal intensities between lung cancer patients and cancer-free smokers.

GenBank ID	Protein Description	Ratio*	*p*value
BC039826.1	**DEAD (Asp-Glu-Ala-Asp) box polypeptide 6 (DDX6)**	**43.6**	**3.36E–03**
NP_001419.1	**Enolase 1 (ENO1)**	**33.8**	**7.36E–03**
BC050424.1	**F-box protein 38 (FBXO38)**	**26.2**	**2.76E–03**
NP_001129173.1	**14–3–3 protein zeta (14–3–3ζ)**	**19.7**	**5.60E–03**
NP_640,343.1	**Cancer/testis antigen 1 (CTAG1A)**	**15.9**	**4.50E–03**
PV3665	**Casein kinase 1, delta (CSNK1D)**	**12.1**	**7.32E–03**
CAA56734.1	**Cell surface associated Mucin 1 (MUC1)**	**10.7**	**1.78E–03**
BC026100.2	Testis-specific transcript, Y-linked 8 (TTTY8)	8.4	1.80E–03
BC059404.1	B-cell CLL/lymphoma 6, member B (BCL6B)	7.8	4.01E–02
NM_002742.1	Serine/threonine-protein kinase D1 (PRKD1)	7.5	2.37E–02
NP_000691.1	Annexin 1 (ANXA1)	6.9	2.86E–02
BC001497.2	RAN binding protein 5 (RANBP5)	5.1	2.35E–02
NM_005118.2	tumor necrosis factor superfamily, member 15 (TNFSF15)	4.6	3.21E–03
BC005220.1	Chaperonin containing TCP1, subunit 8 (theta) (CCT8)	45.1	8.33E–03
AAH04364.1	Transmembrane protein 39B (TMEM39B)	3.2	3.65E–02
NM_006792.2	Mortality factor 4 (MORF4)	43.3	2.33E–02
NM_203,454.	Putative C->*U*-editing enzyme APOBEC-4 (APOBEC4)	28.3	1.67E–02
XM_936,115.1	HLA class II histocompatibility antigen, DQ(3) alpha chain (HLA-DQA3)	22.6	2.13E–02
PV3825	Casein kinase 1, gamma 1 (CSNK1G1)	21.9	1.67E–02
NM_002622.3	Prefoldin subunit 1 (PFDN1)	17.3	2.63E–02
NM_031966.4	Cyclin B1 (CCNB1)	4.7	2.67E–02
BC033708.1	Ral GEF with pH domain and SH3 binding motif 1 (RALGPS1)	4.8	4.50E–03
AAF66640.1	Heat shock protein HSP60 (HSP60)	40.3	2.38E–02
NM_005517.2	High-mobility group nucleosomal binding domain 2 (HMGN2)	3.5	8.54E–03
BC041157.1	Thromboxane A synthase 1 (TBXAS1)	3.6	9.34E–03
BC096097.2	Histidine ammonia-lyase (HAL)	4	6.23E–03
NM_172,160.1	**Methylcrotonoyl-Coenzyme A carboxylase 1 (alpha) (MCCC1)**	**10.6#**	**2.53E–03**
BC013119.1	Signal-induced proliferation-associated 1 like 2 (SIPA1L2)	3#	3.33E–02

* signal intensity ratio of two groups.

\# the proteins' signal intensity ratio is calculated by mean of signal intensity of cancer-free smokers divided by that lung cancer patients, suggesting the proteins has a low level in lung cancer patients vs. cancer-free smokers.The proteins in bold are chosen for subsequent ELISA validation based on criteria: 1) having differential signals between two the populations with a $p \le 0.01$ and, 2) having signal intensity ratio between the two populations being >10.0.

rified proteins to a 96-well microliter plate (Fisher Scientific) at concentration of 0.5 μg/mL. We diluted the supernatant of sputum at 1:100 and horseradish peroxidase (HRP)—conjugated goat anti-human IgG at 1:4000, which were used as the primary and secondary antibodies, respectively. The reaction was sopped with 5% SDS, and the optical density (OD) was measured in an ELISA reader (Thermo Fisher Scientific) at 405 nm. We calculated the ELISA relative absorbance of each sputum sample–antigen reaction by the OD450 of expressed antigens divided by median OD450 of all antigens tested. We used the median value to normalize the background of each sputum specimen. Each sample was analyzed in triplicate.

Statistical analysis

To analyze microarray data, we compared the signal intensity from each protein with the signal from the negative control features among the array and assigned a CI-p value for each protein. We identified the significant signals and calculated the Z-scores, which reflected the signal strength relative to all protein features. We compared the two groups and identified the proteins that had increased signal values in both groups. We calculated a p value for each protein across the hypothesis that there was no signal increase in one group compared with another. We compared clusters using Multi Experiment Viewer (Dana-Farber Cancer Institute, Boston, MA) based on Pearson's correlation distant metrics using normalized values in $\log_{10}$ to picture the discrimination between the two groups.

To analyze ELISA data, we used a one-tailed Student's *t*-test to evaluate whether the means of normal and cancer cases were statistically different from each other. We used Pearson's correlation analysis to determine the relationship of each targets with demographic characteristics of the patients and cancer-free controls. We used final clinical and pathologic diagnoses as reference standards to estimate sensitivity and

specificity of the analysis of each biomarker. Furthermore, we analyzed the significantly associated TAAbs by using logistic regression models with constrained parameters as least absolute shrinkage and selection operator (LASSO) to eliminate the less relevant variables [33]. We estimated functions of the combined biomarkers by logistic regression with or without adjustment for known risk factors for NSCLC. Receiver-operator characteristic (ROC) curve analysis was undertaken using expression level for each biomarker from cancer patients and cancer-free controls by analyze-it software (analyze-it Software, Leeds, UK). We also generated a 95% confidence interval for the difference in the area under the ROCs (AUCs) by using the bootstrap. We established the optimal cut-off value by using the Youden index. Moreover, to compare different biomarker panels, we computed their AUCs to determine the sensitivity and specificity as previously described [43].

Results

TAs of lung cancer in sputum identified by protein microarrays

The lowest average background in sputum was 482 relative fluorescence units (RFUs), and the highest average background was 517 RFUs. Mean background signal values on the arrays were within the expected range for all sputum samples. Furthermore, signal used values for protein (antigen) features reached the scanner maximum of ∼65,000 RFU, indicating a dynamic range >2 logs. Therefore, the experiments and data acquisition of the microarray analysis were performed under optimal conditions. Altogether, 28 TAs (Table 3) were identified that showed significantly differential signal intensities between lung cancer patients and cancer-free smokers with a *p*< 0.05 and signal intensity ratio of two groups being greater than 3.0. Among the 28 antigens, 26 or two displayed increased or decreased levels in lung cancer cases *vs.* cancer-free smokers, respectively.
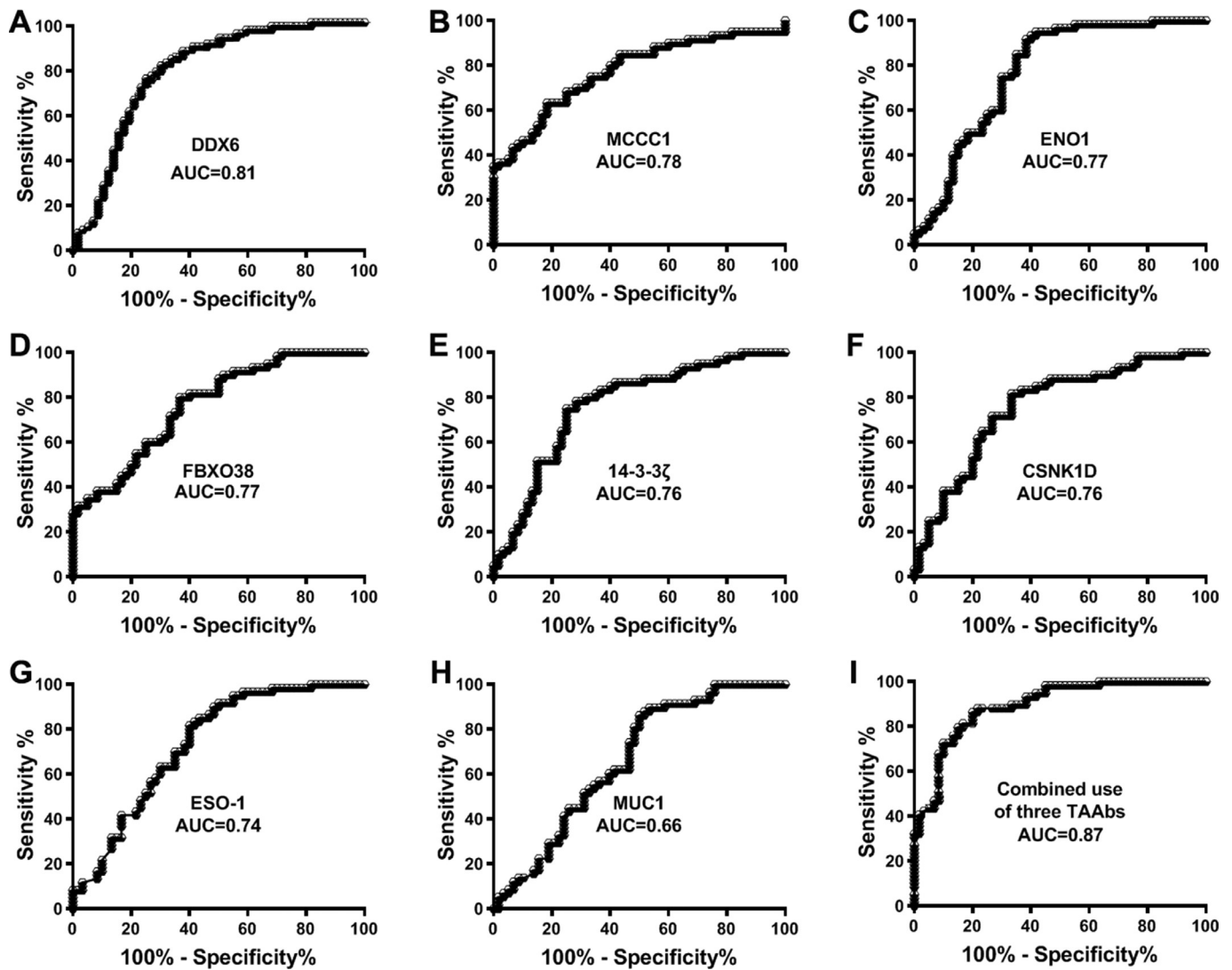
Fig. 1. Receiver-operator characteristic (ROC) curves of 8 TAAbs in sputum of 166 patients diagnosed with NSCLC and 213 cancer-free smokers.

Table 4
ELISA validation of eihgt lung cancer-antigens in cohort 1.

Protein Description	Relative OD in cancer-free smokers	Relative OD in cancer patients	pvalues
DDX6	0.2125 (0.4131)	1.112 (1.1276)	6.92E−08
ENO1	0.0046 (0.0186)	0.15655 (0.3968)	0.00399
FBXO38	0.03077 (0.0496)	0.3220 (0.5037)	0.00002
14–3–3ζ	0.00652 (0.0168)	0.09751 (0.2493)	0.006
ESO-1	0.00831 (0.0303)	0.2315 (0.6001)	0.0052
CSNK1D	0.0967 (0.2643)	0.4955 (0.8259)	0.0005
MUC1	0.05822 (0.1446)	0.76187 (1.5741)	0.00095
MCCC1	4.2286 (0.6328)	1.7976 (1.6222)	9.41E−09

Validation of the identified lung TAs by ELISA in sputum

From the 28 antigens, 8 antigens (Tables 3 and 4) were chosen for subsequent ELISA validation based on criteria: 1) having differential signals between the two populations with a $p \leq 0.01$ and, 2) having signal intensity ratio between the two populations being >10.0. ELISA was first performed in sputum samples of the discovery set that was used in the protein microarray analysis. Of the 8 TAs tested, seven or one had augmented or reduced interaction levels with TAAbs in sputum of lung cancer cases vs. cancer-free smokers, respectively (All $p < 0.01$).

The tested TAs by ELISA displayed the same significant direction as by protein array (All $p < 0.01$) (Table 4). To further validate the TAs, we analyzed TAAbs against the 8 TAs in a validation set with a relatively large sample size that consisted of 166 patients diagnosed with lung cancer in different stages, locations, and histological types and 213 cancer-free smokers. Consistent with data observed in discovery set, the TAAbs in validation set displayed a considerably different level in sputum of lung cancer patients vs. control subjects (All $p < 0.01$).

Development of sputum TAAb biomarkers for lung cancer

We evaluated the potential of using TAAbs against the 8 TAs for diagnosis of NSCLC in the 166 cases and 213 controls. The TAAbs exhibited AUC values of 0.66–0.81 in distinguishing NSCLC patients from cancer-free smokers with 60–74% sensitivities and 60–78% specificities (Fig. 1 and Table 5). We used a stepwise logistic regression model to select the optimal panel of biomarkers. Three TAAbs, which were against DDX6, ENO1, and 14-3-3ζ respectively, were selected as the best biomarkers in a panel. The optimal cutoff for the combined biomarkers was $U = 0.3628$, where $U = -6.2573 + 1.3529*\log$ (DDX6)-1.9647*log (ENO1)+2.3165*14-3-3ζ. Any subject with U≥0.3628 was classified as a NSCLC case. The combined use of the three TAAbs produced 0.88 AUC in distinguishing NSCLC patients from the controls, which was sig-

Table 5
diagnostic performance of eight lung cancer-antigens in cohort 2.

Protein Description	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
DDX6	0.8062 (0.7367 to 0.8905)	73.68% (62.32% to 83.13%)	75.44 (62.24% to 85.87%)
MCCC1	0.7762 (0.6912 to 0.8593)	70.00% (56.79% to 81.15%)	71.67% (58.56% to 82.55%)
ENO1	0.7747 (0.6883 to 0.8611)	75.00% (62.14% to 85.28%)	70.00% (56.79% to 81.15%)
FBXO38	0.7686 (0.6864 to 0.8508)	61.67% (48.21% to 73.93%)	70.00% (56.79% to 81.15%)
14–3–3ζ	0.7656 (0.6789 to 0.8522)	58.33% (44.88% to 70.93%)	78.62% (67.67% to 89.22%)
CSNK1D	0.7615 (0.6744 to 0.8473)	71.67% (58.56% to 82.55%)	73.33% (60.34% to 83.93%)
ESO-1	0.7358 (0.6457 to 0.8259)	70.00% (56.79% to 81.15%)	65.00% (51.60% to 76.87%)
MUC1	0.6636 (0.5634 to 0.7639)	60.34% (46.64% to 72.95%)	60.34% (46.64% to 72.95%)
Combined use of three TAAbs	0.8667 (0.7935 to 0.9499)	81.36% (69.09% to 90.31%)%	83.33% (71.48% to 91.71%)

Abbreviations: AUC, the area under receiver operating characteristic curve; CI, confidence interval.

nificantly higher than that of any single one ($p < 0.05$) (Fig. 1). Furthermore, the analysis of the three biomarkers produced 81% sensitivity and 83% specificity for diagnosis of NSCLC (Table 5). Interestingly, combined analysis of the three TAAbs did not display a statistical difference between stages, locations, and histological types of lung cancer (All $p > 0.05$). Furthermore, the prevalence of the biomarkers was related with pack-years of smoking ($p < 0.05$), however, not associated with patient age and gender.

Discussion

Since sputum is directly secreted from the lower airways and deep lungs where lung tumors exist, it has the advantages as surrogate material over serum or plasma for specifically diagnosing lung cancer. First, sampling of sputum is simple and does not require trained health care professionals. Second, sputum is microbiologically stabilized after the collection. Third, the collection can be done anywhere, including at home. Fourth, large numbers of sputum samples can be collected rapidly, and therefore, are particularly useful for screening a large number of high-risk individuals or for usage in decentralized regions. Finally, and most importantly, since sputum originates from the deep lungs where lung tumors are immersed in the biofluid, sputum may constitute a direct source of TAAb biomarkers for the early detection of NSCLC.

In this study, we first comprehensively profiled proteins in sputum using microarray and identified TAs in sputum of lung cancer patients. Sputum TAAbs against TAs (DDX6, ENO1, and 14-3-3ζ) were further developed as a biomarker panel that had 81% sensitivity and 83% specificity for diagnosis of lung cancer. In addition, the observed TAAb responses were independent of patients' characteristics, including race, gender, and tumor stage and type, suggesting that the sputum TAAb biomarkers could be used for equally detecting the major types of NSCLC at the early stages, regardless of locations of the primary tumors.

ENO1 exists on the surface of cells and has important functions in cell migration and tissue invasion. Activated ENO1 has a direct coupling to glycolytic activity. Furthermore, ENO1 is upregulated by HIF1α in response to hypoxia [44]. We have found that *ENO1* gene has increased levels in both genomic copy number and transcript in lung cancer cells compared with normal cells, and its dysregulation is involved in lung tumorigenesis [45]. Furthermore, we developed a panel of six genomic probes including one targeting *ENO1* gene that could be analyzed in the exfoliated airway epithelial cells of sputum for lung cancer detection [22]. Moreover, our previous study revealed that ELISA analysis of ENO1 protein in sputum supernatant had a diagnostic value in distinguishing patients with stage I lung cancer from cancer-free individuals [12].

14–3–3ζ is a family of cellular proteins and could interact with other crucial cellular proteins involved in the tumor development and progression [38]. We previously found that increased 14-3-3ζ protein level was positively associated with stage and grading of NSCLC and inversely related to poor outcomes of the patients [38]. Furthermore, lung cancer cells with down-regulation of 14-3-3ζ were more sensitive to cisplatin

in vitro that was associated with the inhibition of cell proliferation, additive G2-M cell cycle arrest, and increased apoptosis [38]. In addition, 14-3-3ζ might act as a proto-oncogene and have oncogenic function in the development of NSCLC tumorigenesis [38]. Interestingly, Qiu et al. previously showed that the occurrence serum TAAb to 14-3-3ζ was one of the three biomarkers that could precede onset of symptoms and diagnosis of lung cancer [46].

We are interested to find that TAAb against DDX6 (also named RCK/p54) is one of the biomarkers in the panel. DDX6 is a member of the family of DEAD/H-box RNA helicases and associated with t(11;14)(q23;q32) chromosomal translocation in B-cell lymphoma cell line RC-K8 [47–49]. DDX6 has numerous functions, including translation initiation, pre-mRNA splicing, ribosome assembly by acting as an RNA-binding protein. It also contributes to the proliferation and differentiation of the stem and progenitor cells. However, functions of DDX6 in lung carcinogenesis are largely unknown. Herein we show that DDX6 might be a novel TA and detection of the corresponding TAAb in sputum could help diagnose NSCLC at the early stage. Furthermore, the TAAbs are not only noble candidates for the early detection of NSCLC but potential targets for therapeutic intervention of the malignancy.

Although showing promise, the TAAb biomarkers developed in this study are still limited in their sensitivity (81%) and specificity (83%) for diagnosis of lung cancer. We are extensively evaluating all the 28 TAs defined by the protein microarray in a large case-control study to identify additional TAAb biomarkers that can be added to the three so that the diagnostic efficacy of the approach could be improved. Furthermore, it will be useful in the future to combine the TAAb biomarkers with other types of molecular biomarkers, such as DNA and microRNAs, in sputum to create a combined panel with better diagnostic characteristics. In addition, given that the TAs in sputum and blood might have diverse roles in the carcinogenesis of NSCLC through various mechanisms, further comparing sputum and serum ATTb biomarkers for the diagnostic values and determining if their integration would have a synergistic influence on detection of NSCLC is warranted.

The overall 5-year survival rate for stage I NSCLC patients who are typically treated with surgery remains up to 83%. In contrast, only 5–15% and less than 2% of patients with stage III and IV NSCLC are alive after five years [1,2]. These statistics provide the primary rationale to develop biomarker that can improve NSCLC early detection [1,2]. Furthermore, the National Lung Screening Trial (NLST) shows that low-dose CT (LDCT) could detect NSCLC at the early stage and significantly reduce the mortality [2]. However, 25% of smokers screened by LDCT have indeterminate pulmonary nodules (PNs), of which 96% are benign growths. LDCT screening for lung cancer often produces a high false-positive rate and over-diagnosis. Therefore, many patients with benign PNs will undergo harmful diagnostic and therapeutic procedures, presenting a major clinical challenge. The ability to immediately and precisely distinguish malignant from benign PNs at baseline LDCT screening is an unmet need. The TAAb biomarkers might have two important applications. First, they could be used for the early detection of lung cancer patients in high-risk populations (e.g., heavy smokers). Second, the use of the biomarkers in LDCT-based lung cancer screening would i), spare

individuals with benign growths from radiation exposure and unnecessary surgical resections or biopsies, and ii), lead to more personalized therapy by enabling effective treatments to be immediately initiated for lung cancer, and hence reduce lung cancer-associated mortality.

In conclusion, using protein microarray together with ELISA, we developed three TAAbs as a small panel of sputum biomarkers for the early detection of lung cancer. The findings from this study may lay the basis to perform a next step to a new diagnostic approach that might have future clinical utility.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank the University of Maryland Marlene and Stewart Greenebaum Comprehensive Cancer Center Biostatistics Shared Service to provide statistical analysis services.

Ethical approval and consent to participate

All work was approved by the Ethics Committee of University of Maryland Baltimore

Consent for publication

We obtained consent to publish this paper from all the participants of this study.

Authors' contributions

NL and VKH conducted the experiments and participated in data interpretation. VKH, JD, and NWT participated in coordination and acquisition of data. VKH, NWT and FJ participated in study design, coordination, and data analysis and interpretation, and prepared the manuscript. All authors read and approved the final manuscript.

References

[1] R. Siegel, D. Naishadham, A. Jemal, Cancer statistics, 2013, CA Cancer J. Clin. 63 (2013) 11–30.
[2] D.R. Aberle, A.M. Adams, C.D. Berg, W.C. Black, J.D. Clapp, R.M. Fagerstrom, I.F. Gareen, C. Gatsonis, P.M. Marcus, J.D. Sicks, Reduced lung-cancer mortality with low–dose computed tomographic screening, N. Engl. J. Med. 365 (2011) 395–409.
[3] M. Okada, W. Nishio, T. Sakamoto, K. Uchino, T. Yuki, A. Nakagawa, N. Tsubota, Prognostic significance of perioperative serum carcinoembryonic antigen in non-small cell lung cancer: analysis of 1,000 consecutive resections for clinical stage I disease, Ann. Thorac. Surg. 78 (2004) 216–221.
[4] S.M. Hanash, E.J. Ostrin, J.F. Fahrmann, Blood based biomarkers beyond genomics for lung cancer screening, Transl. Lung Cancer Res. 7 (2018) 327–335.
[5] I. Belczacka, A. Latosinska, J. Metzger, D. Marx, A. Vlahou, H. Mischak, M. Frantzi, Proteomics biomarkers for solid tumors: current status and future prospects, Mass Spectrom. Rev. 38 (2019) 49–78.
[6] C. Desmetz, A. Mange, T. Maudelonde, J. Solassol, Autoantibody signatures: progress and perspectives for early cancer detection, J. Cell. Mol. Med. 15 (2011) 2013–2024.
[7] P. Zaenker, E.S. Gray, M.R. Ziman, Autoantibody production in cancer–the humoral immune response toward autologous antigens in cancer patients, Autoimmun. Rev. 15 (2016) 477–483.
[8] M.A. Murphy, J.J. O'Leary, D.J. Cahill, Assessment of the humoral immune response to cancer, J. Proteomics 75 (2012) 4573–4579.
[9] C.J. Chapman, A. Murray, J.E. McElveen, U. Sahin, U. Luxemburger, O. Tureci, R. Wiewrodt, A.C. Barnes, J.F. Robertson, Autoantibodies in lung cancer: possibilities for early detection and subsequent cure, Thorax 63 (2008) 228–233.
[10] B. Yang, X. Li, T. Ren, Y. Yin, Autoantibodies as diagnostic biomarkers for lung cancer: a systematic review, Cell. Death Discov. 5 (2019) 126.
[11] S. Lam, P. Boyle, G.F. Healey, P. Maddison, L. Peek, A. Murray, C.J. Chapman, J. Allen, W.C. Wood, H.F. Sewell, J.F. Robertson, EarlyCDT-Lung: an immunobiomarker test as an aid to early detection of lung cancer, Cancer Prev. Res. 4 (2011) 1126–1134.
[12] L. Yu, J. Shen, K. Mannoor, M. Guarnera, F. Jiang, Identification of ENO1 as a potential sputum biomarker for early-stage lung cancer by shotgun proteomics, Clin. Lung Cancer 15 (2014) 372–378 e371.
[13] Q. Qiu, N.W. Todd, R. Li, H. Peng, Z. Liu, H.G. Yfantis, R.L. Katz, S.A. Stass, F. Jiang, Magnetic enrichment of bronchial epithelial cells from sputum for lung cancer diagnosis, Cancer 114 (2008) 275–283.
[14] S.C. Bishwal, M.K. Das, V.K. Badireddy, D. Dabral, A. Das, A.R. Mahapatra, S. Sahu, D. Malakar, I.I. Singh, H. Mazumdar, S.J. Patgiri, T. Deka, W. Kapfo, K. Liegise, R.U. Kupa, S. Debnath, R. Bhowmik, R. Debnath, R.K. Behera, M.G. Pillai, P. Deuri, R. Nath, K.P. Khalo, W.A. Sing, B. Pandit, S. Bhattacharya, D. Behera, L. Saikia, V. Khamo, R.K. Nanda, Sputum proteomics reveals a shift in vitamin D-binding protein and antimicrobial protein axis in tuberculosis patients, Sci. Rep. 9 (2019) 1036.
[15] M.C. Rose, J.A. Voynow, Respiratory tract mucin genes and mucin glycoproteins in health and disease, Physiol. Rev. 86 (2006) 245–278.
[16] F.B. Thunnissen, Sputum examination for early detection of lung cancer, J. Clin. Pathol. 56 (2003) 805–810.
[17] S.A. Belinsky, W.A. Palmisano, F.D. Gilliland, L.A. Crooks, K.K. Divine, S.A. Winters, M.J. Grimes, H.J. Harms, C.S. Tellez, T.M. Smith, P.P. Moots, J.F. Lechner, C.A. Stidley, R.E. Crowell, Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers, Cancer Res. 62 (2002) 2370–2377.
[18] S.A. Belinsky, Gene-promoter hypermethylation as a biomarker in lung cancer, Nat. Rev. Cancer 4 (2004) 707–717.
[19] M. Varella-Garcia, J. Kittelson, A.P. Schulte, K.O. Vu, H.J. Wolf, C. Zeng, F.R. Hirsch, T. Byers, T. Kennedy, Y.E. Miller, R.L. Keith, W.A. Franklin, Multi-target interphase fluorescence in situ hybridization assay increases sensitivity of sputum cytology as a predictor of lung cancer, Cancer Detect. Prev. 28 (2004) 244–251.
[20] M. Varella-Garcia, A.P. Schulte, H.J. Wolf, W.J. Feser, C. Zeng, S. Braudrick, X. Yin, F.R. Hirsch, T.C. Kennedy, R.L. Keith, A.E. Baron, S.A. Belinsky, Y.E. Miller, T. Byers, W.A. Franklin, The detection of chromosomal aneusomy by fluorescence in situ hybridization in sputum predicts lung cancer incidence, Cancer Prev. Res. 3 (2010) 447–453.
[21] R. Li, N.W. Todd, Q. Qiu, T. Fan, R.Y. Zhao, W.H. Rodgers, H.B. Fang, R.L. Katz, S.A. Stass, F. Jiang, Genetic deletions in sputum as diagnostic markers for early detection of stage I non-small cell lung cancer, Clin. Cancer Res. 13 (2007) 482–487.
[22] F. Jiang, N.W. Todd, R. Li, H. Zhang, H. Fang, S.A. Stass, A panel of sputum-based genomic marker for early detection of lung cancer, Cancer Prev. Res. 3 (2010) 1571–1578.
[23] N. Anjuman, N. Li, M. Guarnera, S.A. Stass, F. Jiang, Evaluation of lung flute in sputum samples for molecular analysis of lung cancer, Clin. Transl. Med. 2 (2013) 15.
[24] C. Gupta, J. Su, M. Zhan, S.A. Stass, F. Jiang, Sputum long non-coding RNA biomarkers for diagnosis of lung cancer, Cancer Biomark. 26 (2019) 219–227.
[25] F. Jiang, N.W. Todd, Q. Qiu, Z. Liu, R.L. Katz, S.A. Stass, Combined genetic analysis of sputum and computed tomography for noninvasive diagnosis of non-small-cell lung cancer, Lung Cancer 66 (2009) 58–63.
[26] R.L. Katz, T.M. Zaidi, R.L. Fernandez, J. Zhang, W. He, C. Acosta, M. Daniely, L. Madi, M.A. Vargas, Q. Dong, X. Gao, F. Jiang, N.P. Caraway, A.A. Vaporciyan, J.A. Roth, M.R. Spitz, Automated detection of genetic abnormalities combined with cytology in sputum is a sensitive predictor of lung cancer, Mod. Pathol. 21 (2008) 950–960.
[27] N. Li, J. Ma, M.A. Guarnera, H. Fang, L. Cai, F. Jiang, Digital PCR quantification of miRNAs in sputum for diagnosis of lung cancer, J. Cancer Res. Clin. Oncol. 140 (2014) 145–150.
[28] J. Shen, J. Liao, M.A. Guarnera, H. Fang, L. Cai, S.A. Stass, F. Jiang, Analysis of MicroRNAs in sputum to improve computed tomography for lung cancer diagnosis, J. Thorac. Oncol. 9 (2014) 33–40.
[29] J. Su, N. Anjuman, M.A. Guarnera, H. Zhang, S.A. Stass, F. Jiang, Analysis of lung flute-collected sputum for lung cancer diagnosis, Biomark Insights 10 (2015) 55–61.
[30] J. Su, Q. Leng, Y. Lin, J. Ma, F. Jiang, C.J. Lee, H. Fang, Integrating circulating immunological and sputum biomarkers for the early detection of lung cancer, Biomark Cancer 10 (2018) 1179299.
[31] J. Su, J. Liao, L. Gao, J. Shen, M.A. Guarnera, M. Zhan, H. Fang, S.A. Stass, F. Jiang, Analysis of small nucleolar RNAs in sputum for lung cancer diagnosis, Oncotarget 7 (2016) 5131–5142.
[32] Y. Su, H. Fang, F. Jiang, Integrating DNA methylation and microRNA biomarkers in sputum for lung cancer detection, Clin Epigenetics 8 (2016) 109.
[33] Y. Su, M.A. Guarnera, H. Fang, F. Jiang, Small non-coding RNA biomarkers in sputum for lung cancer diagnosis, Mol. Cancer 15 (2016) 36.
[34] Y. Xie, N.W. Todd, Z. Liu, M. Zhan, H. Fang, H. Peng, M. Alattar, J. Deepak, S.A. Stass, F. Jiang, Altered miRNA expression in sputum for diagnosis of non-small cell lung cancer, Lung Cancer 67 (2010) 170–176.
[35] L. Xing, J. Su, M.A. Guarnera, H. Zhang, L. Cai, R. Zhou, S.A. Stass, F. Jiang, Sputum microRNA biomarkers for identifying lung cancer in indeterminate solitary pulmonary nodules, Clin. Cancer Res. 21 (2015) 484–489.
[36] L. Xing, N.W. Todd, L. Yu, H. Fang, F. Jiang, Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers, Mod. Pathol. 23 (2010) 1157–1164.
[37] L. Yu, N.W. Todd, L. Xing, Y. Xie, H. Zhang, Z. Liu, H. Fang, J. Zhang, R.L. Katz, F. Jiang, Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers, Int. J. Cancer 127 (2010) 2870–2878.
[38] T. Fan, R. Li, N.W. Todd, Q. Qiu, H.B. Fang, H. Wang, J. Shen, R.Y. Zhao, N.P. Caraway, R.L. Katz, S.A. Stass, F. Jiang, Up-regulation of 14-3-3zeta in lung cancer and its implication as prognostic and therapeutic target, Cancer Res. 67 (2007) 7901–7906.
[39] F. Jiang, Q. Qiu, A. Khanna, N.W. Todd, J. Deepak, L. Xing, H. Wang, Z. Liu, Y. Su, S.A. Stass, R.L. Katz, Aldehyde dehydrogenase 1 is a tumor stem cell-associated marker in lung cancer, Mol. Cancer Res. 7 (2009) 330–338.

[40] L. Dai, P. Ren, M. Liu, H. Imai, E.M. Tan, J.Y. Zhang, Using immunomic approach to enhance tumor-associated autoantibody detection in diagnosis of hepatocellular carcinoma, Clin. Immunol. 152 (2014) 127–139.

[41] K. Csorba, S. Schmidt, F. Florea, N. Ishii, T. Hashimoto, M. Hertl, S. Karpati, L. Bruckner-Tuderman, W. Nishie, C. Sitaru, Development of an ELISA for sensitive and specific detection of IgA autoantibodies against BP180 in pemphigoid diseases, Orphanet. J. Rare. Dis. 6 (2011) 31.

[42] L. Dai, J.C. Tsay, J. Li, T.A. Yie, J.S. Munger, H. Pass, W.N. Rom, Y. Zhang, E.M. Tan, J.Y. Zhang, Autoantibodies against tumor-associated antigens in the early detection of lung cancer, Lung Cancer 99 (2016) 172–179.

[43] J. Shen, Z. Liu, N.W. Todd, H. Zhang, J. Liao, L. Yu, M.A. Guarnera, R. Li, L. Cai, M. Zhan, F. Jiang, Diagnosis of lung cancer in individuals with solitary pulmonary nodules by plasma microRNA biomarkers, BMC Cancer 11 (2011) 374.

[44] T. Ito, K. Funamoto, N. Sato, A. Nakamura, K. Tanabe, T. Hoshiai, K. Suenaga, J. Sugawara, S. Nagase, K. Okamura, N. Yaegashi, Y. Kimura, Maternal undernutrition induces the expression of hypoxia-related genes in the fetal brain, Tohoku J. Exp. Med. 226 (2012) 37–44.

[45] R. Li, H. Wang, B.N. Bekele, Z. Yin, N.P. Caraway, R.L. Katz, S.A. Stass, F. Jiang, Identification of putative oncogenes in lung adenocarcinoma by a comprehensive functional genomic approach, Oncogene 25 (2006) 2628–2635.

[46] J. Qiu, G. Choi, L. Li, H. Wang, S.J. Pitteri, S.R. Pereira-Faca, A.L. Krasnoselsky, T.W. Randolph, G.S. Omenn, C. Edelstein, M.J. Barnett, M.D. Thornquist, G.E. Goodman, D.E. Brenner, Z. Feng, S.M. Hanash, Occurrence of autoantibodies to annexin I, 14-3-3 theta and LAMR1 in prediagnostic lung cancer sera, J. Clin. Oncol. 26 (2008) 5060–5066.

[47] A. Iio, T. Takagi, K. Miki, T. Naoe, A. Nakayama, Y. Akao, DDX6 post-transcriptionally down-regulates miR-143/145 expression through host gene NCR143/145 in cancer cells, Biochim. Biophys. Acta 1829 (2013) 1102–1110.

[48] K. Taniguchi, A. Iwatsuki, N. Sugito, H. Shinohara, Y. Kuranaga, Y. Oshikawa, T. Tajirika, M. Futamura, K. Yoshida, K. Uchiyama, Y. Akao, Oncogene RNA helicase DDX6 promotes the process of c-Myc expression in gastric cancer cells, Mol. Carcinog. 57 (2018) 579–589.

[49] Y. Tokumaru, T. Tajirika, N. Sugito, Y. Kuranaga, H. Shinohara, T. Tsujino, N. Matsuhashi, M. Futamura, Y. Akao, K. Yoshida, Synthetic miR-143 inhibits growth of HER2-positive gastric cancer cells by suppressing KRAS networks including DDX6 RNA helicase, Int. J. Mol. Sci. 20 (2019).