



Identification and validation of novel circulating biomarkers for early diagnosis of lung cancer



Li Zhang^{b,e,1}, Dan Pu^{c,1}, Dan Liu^a, Ye Wang^a, Wenxin Luo^a, Huairong Tang^d, Yan Huang^d, Weimin Li^{a,e,*}

^a Department of Respiratory Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan, China

^b Lab. of Pathology, Key Lab. of Transplantation Engineering and Immunology, Ministry of Health, West China Hospital, Sichuan University, Chengdu, Sichuan, China

^c Clinic Skill Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China

^d Health Management Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China

^e Precision Medicine Center, West China Hospital, Sichuan University, Chengdu, Sichuan, China

ARTICLE INFO

Keywords:

Lung cancer
Novel biomarkers
Combination analysis
Early diagnosis
Antibody array

ABSTRACT

Objectives: This study aimed to identify novel circulating biomarkers in lung cancer.

Materials and methods: Expression of 1000 secreted proteins in 15 early lung cancer patients and 10 healthy controls' plasma were examined by RayBiotech antibody array. Candidate biomarkers were identified by SPSS analysis ($p < 0.05$ between lung cancer and healthy controls was considered significant) and validated by ELISA in 371 lung cancer patients and 388 healthy controls. Receiver operating characteristic curve and binary logistic regression were performed to evaluate the diagnosis efficacy and to establish diagnostic models.

Results and conclusion: Twenty-two molecules expressed aberrantly in early lung cancer patients compared to healthy controls. Large sample validation showed significantly up-regulated levels of brain angiogenesis inhibitor 1, E-Cadherin, integrin-binding sialoprotein, and down-regulated expression of thrombospondin-1 (all $p < 0.0001$) in lung cancer patients. Receiver operating characteristic curve analysis indicated E-Cadherin, brain angiogenesis inhibitor 1 and thrombospondin-1 had higher sensitivity and specificity than classical biomarkers carcinoembryonic antigen, carbohydrate antibody 19-9 and cytokeratin 19 fragments. The high sensitivity and specificity of E-Cadherin brain angiogenesis inhibitor 1 and thrombospondin-1 were also confirmed in early lung cancer analysis. Combination analysis showed brain angiogenesis inhibitor 1, E-Cadherin and thrombospondin-1 had better diagnostic efficacy than classic lung cancer biomarkers. Our findings demonstrated the potential status of BAI-1, E-Cadherin and TSP-1 in diagnosis of lung cancer.

1. Introduction

Lung cancer is the leading malignancy associated with death, accounting for 86,380 male and 71,660 female deaths in 2015 in United States [1]; in China the mortality is 610.2 per 100,000 [2]. The high mortality of lung cancer is always closely associated with late diagnosis. The 5-year survival rate for lung cancer patients diagnosed at advanced stages is less than 15%, while for those diagnosed at early stages the survival rate can reach 70–80%, which highlights the importance of early diagnosis to improve the overall survival [3–5]. However, early lung cancer screening is difficult due to lack of obvious symptoms [6]. The National Lung Screening Trial (NLST) in United States enrolled a total of 53,454 high risk participants and randomly divided them into

low-dose computed tomography (LDCT) (26,722 participants) and chest X radiography (26,732 participants) screening groups. The results of 6.4 years' screening by the two methods showed that LDCT reduced almost 20% mortality of lung cancer compared with x-ray. However, LDCT screening has a problem of over-diagnosis, because it cannot confirm whether early stage lesions detected in asymptomatic participants will finally progress into symptoms and affect long term survival status. Later analysis of NLST indicated that 18% of lung cancer and 22% of non-small cell lung cancer (NSCLC) cases screen detected by LDCT were cases of over-diagnosis [7–9]. Furthermore, the radiation injury and high cost associated with CT scan have been points of controversy [10]. Other diagnostic methods such as bronchoscopy and needle biopsy are also painful and time-consuming [11,12].

* Corresponding author at: Department of Respiratory Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan, China.

E-mail address: weimi003@scu.edu.cn (W. Li).

¹ These authors contributed equally to this paper.

In contrast with these diagnosis methods, biomarkers from peripheral blood are a widely accepted new way to detect primary lung cancer and its metastases for the properties of non-invasive, convenient and low cost [13]. Circulating biomarkers are various components in blood, urine or other fluids, including circulating tumor cells (CTC), secreted proteins, metabolic products, microRNA, auto-antibody and circulating tumor DNA (ctDNA) [14]. A variety of circulating biomarkers are now used for lung cancer diagnosis and prognosis, including carcinoembryonic antigen (CEA) [15], carbohydrate antibody 19-9 (CA199) [16], carbohydrate antibody 12-5 (CA125) [17], cytokeratin 19 fragments (Cyfra21-1) [18] and neuron-specific enolase (NSE) [19]. The circulating level of these markers usually indicates the tumor risk [20]. However, most of these markers are sensitive only in advanced lung cancer patients (III + IV) and have no benefit for the early diagnosis [21]. Research efforts are in urgent need to identify novel circulating biomarkers that are specific and sensitive enough for clinical utility.

Antibody arrays from RayBio tech have been widely used in biological researches to identify secret molecules in culture medium or body fluid, which can enhance drug-resistance of tumor cells [22], promote tumor metastasis [23] and increase stemness of cancer stem cells [24]. In this study, antibody arrays were performed to screen candidate biomarkers by comparing the expression variation of 1000 secreted molecules in plasma from 15 early lung cancer patients and 10 healthy controls. Twenty two aberrantly expressed molecules were found in lung cancer plasma. Larger samples validation by ELISA in 371 lung cancer patients and 388 healthy controls confirmed E-cadherin, brain angiogenesis inhibitor 1 (BAI-1), integrin-binding sialoprotein (IBSP) and thrombospondin-1 (TSP-1) were more sensitive and specific than current clinical biomarkers of CEA, CA199 and Cyfra21-1, which suggested that E-cadherin, BAI-1, IBSP and TSP-1 could function as biomarkers in early diagnosis of lung cancer.

2. Materials and methods

2.1. Study design

This study was approved by the Medical Ethics Committee of the West China Hospital (Approval number: 2017-114). The procedure of this study included two stages: antibody array screening and ELISA validation for the candidates. The antibody array screening dealt with plasma samples of 15 early lung cancer patients (I + IIA) and 10 healthy controls to identify biomarkers which could be used in early diagnosis of lung cancer. ELISA was used for validation due to its low cost and high reliability. In the validation period, we analyzed plasma samples from 371 lung cancer patients, including all four classical tumor stages, and 388 healthy volunteers.

2.2. Study participants

All together 371 lung cancer patients and 388 healthy volunteers met the eligibility criteria, and were included in this study with signed informed consent. Blood samples of lung cancer patients were collected from patients before therapy in Department of Respiration Thoracic Surgery and the Center of Lung Cancer in West China Hospital of Sichuan University between March 2014 and February 2016. Blood samples were collected within two weeks after the first biopsy-proven lung cancer diagnosis and prior to treatment of surgical procedure, or radiotherapy or chemotherapy. Healthy volunteers were all from the Physical Examination Centre of West China Hospital, with their blood samples collected during the routine physical examination including CT, liver function tests and complete blood tests (blood routine examinations, blood biochemical assays and tumor markers tests). The healthy controls were excluded from any disease including malignancy or benign tumors by the routine physical examination results. All plasmas were prepared according to the standard protocols by centrifuge of the anticoagulant whole blood (5 ml) at 800 g and 4 °C for

20 min. The isolated plasma was then centrifuged (800 g, 20 min, 4 °C) to remove the remaining red blood cells and leukocytes. The plasma was then transferred to 1.5 ml sterile eppendorf tubes and refrigerated at –80 °C.

2.3. Screening of candidate biomarkers

A total of 25 plasma samples from 15 lung cancer patients at early stages (diagnosed with CT scanning; I and IIA) and 10 healthy volunteers, were subjected to RayBiotech antibody array with 1000 secreted proteins for the candidate biomarker screening. All included plasma samples had detailed records of the clinical characteristics, and met the eligibility criteria to ensure the reliability of the identified biomarkers.

2.4. Validation with large samples by ELISA

The plasma level of secreted proteins was validated in large samples by ELISA. The assay kit included BAI-1 (BIOMATIK, EKU02798), E-Cadherin (R&D system, DY648 and DY008), IBSP (LifeSpan, LS-F4424), and TSP-1 (R&D system, DTSP10). All plasma samples and kit components were equilibrated to room temperature before the assay. The detection procedure was in accordance with the manufacturer's instructions. The plasma sample was diluted by sample diluents, added into each well and then incubated with plate sealed. The liquids were removed from wells, added with detection antibody solutions, and then incubated for a second time. After 3–5 washes with buffer, the liquids were added with hydrogen peroxide solution and incubated for 30 min. Color development was achieved by adding of the mixed substrate. Finally, 50 ul ulfuric acid was conducted to stop reaction and the optical density (O.D.) was tested at 450 nm on a plate reader (Biotek). The concentrations were calculated according to the standard curves. Standard samples containing recombinant proteins, plasma samples and blank controls were all assayed in duplicate to reduce variation.

2.5. Statistical analysis

All standard curves for each biomarker were created according to the instruction manual of ELISA kits. The concentration of all tested samples was calculated according the formula from standard curves. *T* test in SPSS19.0 was used to analyze the protein level differences between lung cancer patients and healthy volunteers. The sensitivity and specificity of all biomarkers for lung cancer diagnosis were evaluated by receiver operating characteristic (ROC) curves and areas under the curves (AUC) with 95% confidence interval (CI). And the comparison ROC was conducted by MedCalc (version 15.2.2). The best cut-off value for diagnosis was determined by maximizing the specificity and sensitivity at 95% CI. A two-tailed *P* value less than 0.05 was considered significant. All figures were completed by GraphPad Prism version 5 for Windows.

3. Results

3.1. The RayBiotech antibody array screened 22 candidate biomarkers in lung cancer

The result of RayBiotech C series antibody array for 1000 secreted proteins in plasmas from 15 early lung cancer patients (I and IIA) and 10 healthy volunteers identified 22 aberrantly expressed secreted proteins (Supplement Fig. 1). Among the 22 proteins, 19 were enhanced in lung cancer plasma, including GRP78 (glucose regulating protein 78), HSP (heat shock protein) 60 and 70, Ras, FABP1 (fatty acid binding protein 1), Cerberus 1 (CER1), FGFR1 (fibroblast growth factor receptor 1), TNFRSF (tumor necrosis factor receptor superfamily factor), APJ (apelin receptor J), TSH (thyroid stimulating hormone), CXCL11 (chemokine 11), Visfatin, MMP7 (matrix metalloproteinase 7), BAI1 (brain angiogenesis inhibitor 1), E-cadherin, IBSP (induced bone sialo-

protein), Nell2, Cathepsin and thrombin- γ . Three of the proteins including SDC3 (syndecan-3), TSP1 (Thrombospondin-1) and C3a decreased in lung cancer patients compared with normal volunteers. Most of these proteins are important in tumor biological functions, such as induction of apoptosis (TSP-1), metastasis suppression (BAI1, E-Cadherin), metastasis promotion (Apelin, MMP7), tumorigenesis initiation (Ras) and inhibition of tumor growth (TNFRSF).

3.2. The clinical characteristics of all participants

Since the most reliable way to confirm the candidate biomarkers is to validate in large cohort samples with clear and sufficient clinical records, we applied ELISA testing to validate 22 candidate biomarkers in a total of 759 plasma samples (371 lung cancer patients and 388 healthy controls). The lung cancer patients included 227 male and 144 female patients, and healthy control group included 181 male and 208 female. All patients and controls were divided into three age groups (< 45 years, 45–65 years, and > 65 years). The lung cancer cases were mainly adenocarcinoma (ADC, 69.5%) and squamous carcinoma (SCC, 22.9%), plus some cases of small cell lung cancer, large cell cancer and adeno-squamous carcinoma. The other important characteristics of lung cancer included stages (I: 142 cases, II: 43 cases, III: 84 cases, IV: 87 cases), smoking history (No: 195 cases, Yes: 176 cases) and metastasis (No: 239 cases, Yes: 132 cases). All clinical characteristics of participants are listed in Supplement Table 1.

3.3. Sensitivity and specificity of the novel biomarkers in lung cancer diagnosis

The results of large sample validation showed significantly higher correlation between lung cancer and four potential markers: E-cadherin, BAI-1, IBSP and TSP-1. The levels of E-cadherin, IBSP and BAI-1 were dramatically up-regulated (all $p < 0.0001$), while TSP-1 was down-regulated ($p < 0.0001$) in lung cancer patients (Fig. 1A).

To further explore the diagnosis accuracy, ROC analysis was introduced to define the sensitivity (SN) and specificity (SP) of novel biomarkers. The highest SN/SP was found in E-Cadherin (SN: 91.1%, SP: 89.7%, AUC: 95.8%, Cut-off: 10448 pg/ml). The other three had

lower measures: TSP-1 (SN: 82.6%, SP: 92.8%, AUC: 93.6%, Cut-off: 10458 pg/ml); BAI-1 (SN: 79.2%, SP: 84.8%, AUC: 91.0%, Cut-off: 1096 pg/ml); IBSP (SN: 80.9%, SP: 59.5%, AUC: 79.2%, Cut-off: 366 pg/ml) (Fig. 1B).

Three classical biomarkers (CA199, CEA and Cyfra21.1) were analyzed to compare with the four novel ones for diagnostic value. The results showed lower AUC and SN/SP in the classic biomarkers than in the four novel ones (CA199: SN 80.7%, SP 41.3%, AUC 67.2%, Cut-off: 6.03 ng/ml; CEA: SN 80.5%, SP 62.1%, AUC 80.0%, Cut-off: 1.63 ng/ml; Cyfra21-1: SN 70.2%, SP 54.2%, AUC 72.9%, Cut-off: 1.85 ng/ml) (Supplement Fig. 2). The novel biomarkers BAI-1, E-Cadherin, IBSP and TSP-1 were more sensitive and specific than CA199, CEA and Cyfra21-1 in lung cancer diagnosis.

3.4. The diagnostic value of novel biomarkers in ADC and SCC patients

Since ADC and SCC are the two major subtypes (ADC: 50%, SCC: 35%) of NSCLC [25], high specificity to them is very important for potential novel biomarkers. All together 258 ADC and 85 SCC patients were in our cohort. The ELISA assay result confirmed the significant difference of BAI-1, E-Cadherin, IBSP and TSP-1 ($p < 0.0001$ for all) in both ADC and SCC patients versus healthy controls (Figs. 2A and 3A).

ROC analysis was also used to compare the importance of novel biomarkers in ADC and SCC diagnosis. As shown in Figs. 2B and 3B, BAI1 was more sensitive in SCC (SN: 84.7%, SP: 89.9%, AUC: 94.5%, Cut-off: 1178 pg/ml) than in ADC (SN: 77.9%, SP: 84.0%, AUC: 89.4%, Cut-off: 1079 pg/ml); E-cadherin was more important in ADC (SN: 92.6%, SP: 89.2%, AUC: 96.3%, Cut-off: 10316 pg/ml) than SCC (SN: 91.8%, SP: 76.6%, AUC: 93.7%, Cut-off: 8929 pg/ml). We also found that IBSP was indicative in ADC (SN: 81.0%, SP: 61.1%, AUC: 80.4%, Cut-off: 371 pg/ml) but TSP1 showed no significance between ADC and SCC patients.

3.5. The value of novel biomarker for early diagnosis of lung cancer

As early diagnosis of lung cancer is important in promoting 5-year survival rate, the diagnosis value in early lung cancer is of great importance for novel biomarkers. In total 185 early lung cancer patients

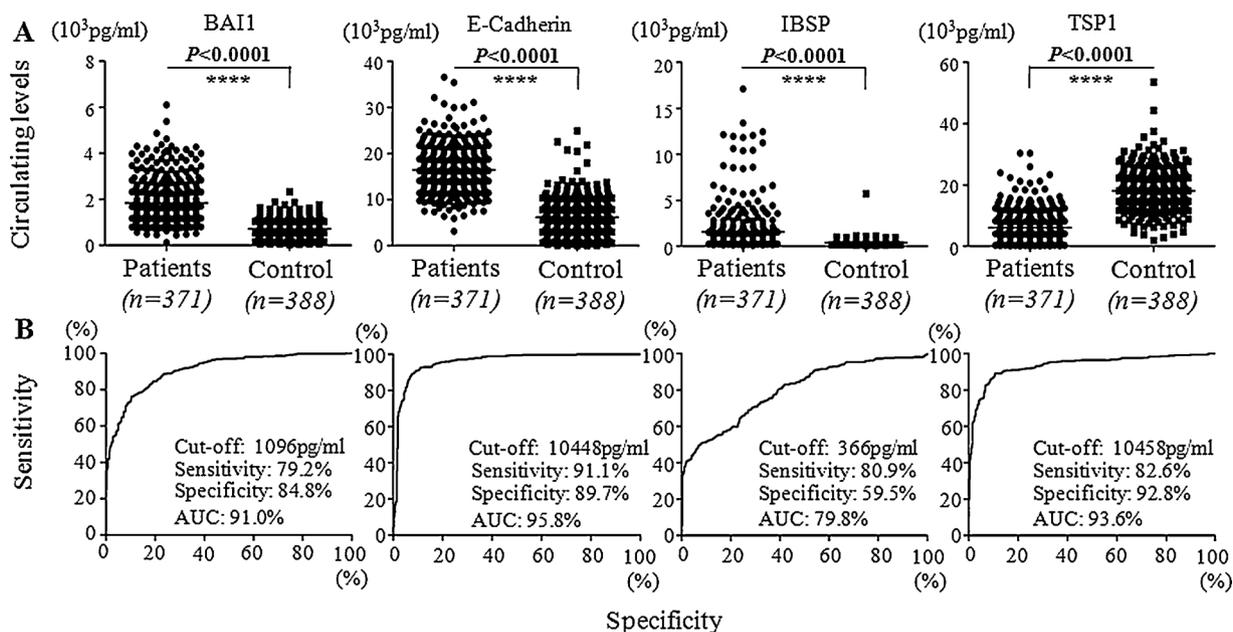


Fig. 1. The concentrations and ROC analysis of BAI1, E-Cadherin, IBSP, TSP1 in all lung cancer patients.

A: concentrations of BAI1, E-Cadherin, IBSP, TSP1 in all lung cancer patients. The black horizontal lines are median values. P values were determined by the Chi-square test, **** $P < 0.0001$.

B: ROC curve analyses of BAI1, E-Cadherin, IBSP, TSP1 in all lung cancer patients versus controls.

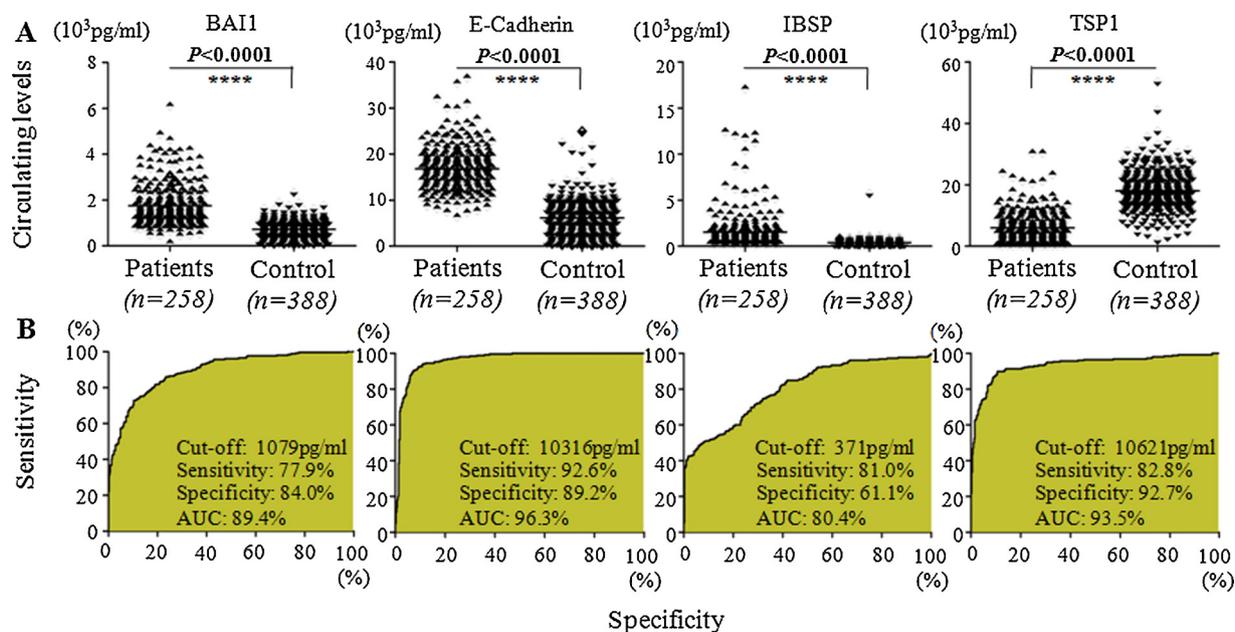


Fig. 2. The concentrations and ROC analysis of BAI1, E-Cadherin, IBSP, TSP1 in ADC patients.

A: concentrations of BAI1, E-Cadherin, IBSP, TSP1 in ADC patients. The black horizontal lines are median values. P values were determined by the Chi-square test, **** P < 0.0001.

B: ROC curve analyses of BAI1, E-Cadherin, IBSP, TSP1 in ADC patients versus controls.

(stage I and II) were enrolled in this study and the results confirmed significant difference in novel biomarkers between early lung cancer patients and healthy controls (P < 0.001 for all) (Fig. 4A). The ROC analysis results were BAI: SN 70.2%, SP 89.5%, AUC 89.6%, Cut-off: 1162 pg/ml; E-Cadherin: SN 95.6%, SP 84.3%, AUC 96.3%, Cut-off: 9715 pg/ml; IBSP: SN 77.3%, SP 61.2%, AUC 77.5%, Cut-off: 372 pg/ml; TSP-1: SN 88.9%, SP 91.3%, AUC 96.2%, Cut-off: 10761 pg/ml (Fig. 4B), which confirmed that E-Cadherin and TSP-1 were more valuable in early diagnosis of lung cancer than BAI and IBSP.

3.6. The relationship between novel biomarker and clinical characteristics

Through observation we found higher concentration of BAI1 in male than in female patients (P < 0.01, Supplement Fig. 3A) and in patients with smoking history than those without (P < 0.01, Supplement Fig. 3B). Histological types analysis showed lower BAI1 levels in ADC than in SCC (P < 0.05) and higher E-cadherin levels in ADC (P < 0.01, Supplement Fig. 3C). However, the age analysis found no significant differences in biomarker levels among the three groups (< 45 years, 45–65 years and > 65 years) (Supplement Fig. 4A). The

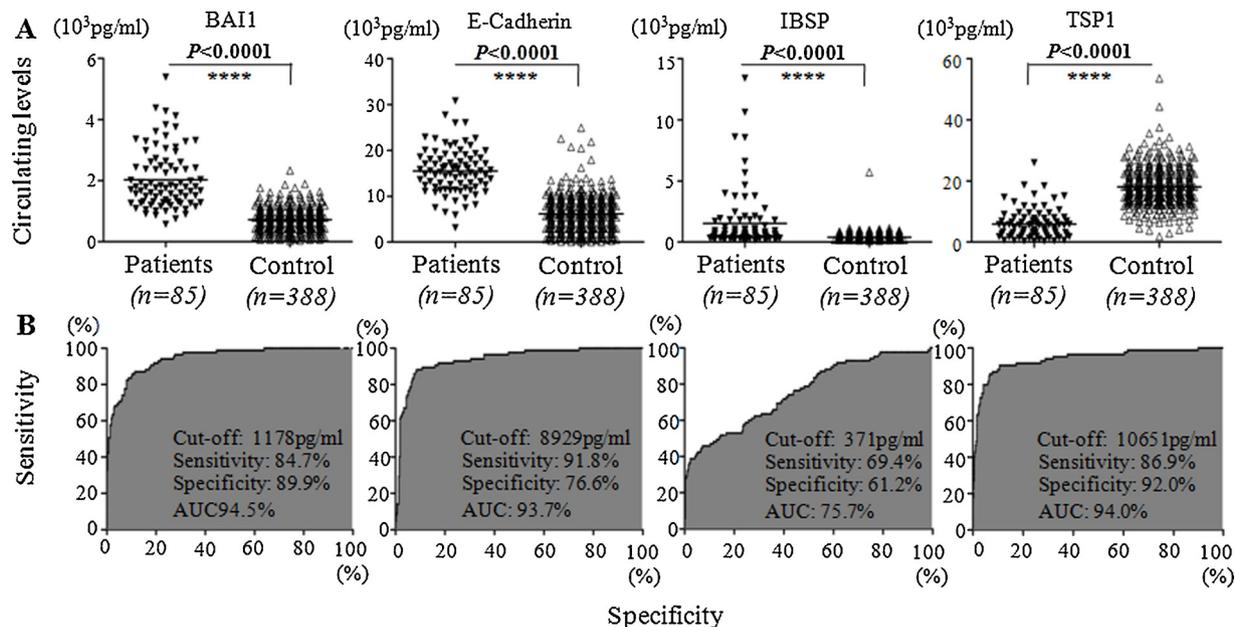


Fig. 3. The concentrations and ROC analysis of BAI1, E-Cadherin, IBSP, TSP1 in SCC patients.

A: concentrations of BAI1, E-Cadherin, IBSP, TSP1 in SCC patients. The black horizontal lines are median values. P values were determined by the Chi-square test, **** P < 0.0001.

B: ROC curve analyses of BAI1, E-Cadherin, IBSP, TSP1 in SCC patients versus controls.

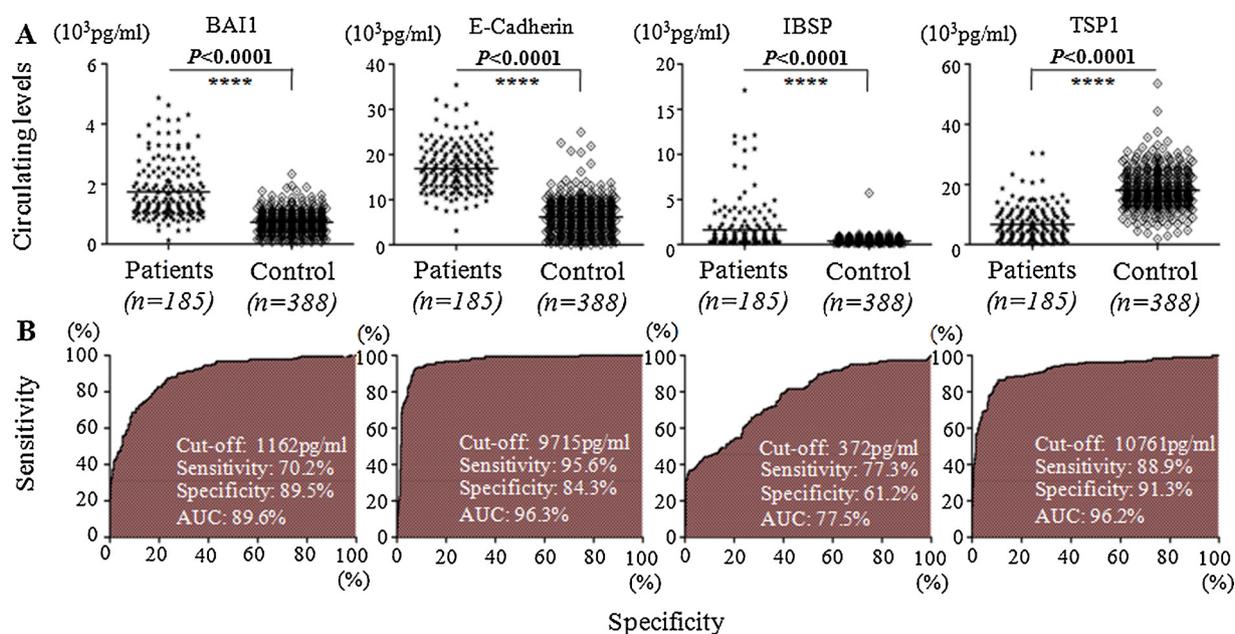


Fig. 4. The concentrations and ROC analysis of BAI1, E-Cadherin, IBSP, TSP1 in early lung cancer patients.

A: concentrations of BAI1, E-Cadherin, IBSP, TSP1 in early lung cancer patients. The black horizontal lines are median values. P values were determined by the Chi-square test, **** $P < 0.0001$.

B: ROC curve analyses of BAI1, E-Cadherin, IBSP, TSP1 in early lung cancer patients versus controls.

level of the novel biomarkers showed no association with the occurrence of metastasis (Supplement Fig. 4B).

3.7. Comparison ROC analyses of novel biomarkers and construction of diagnostic models for lung cancer

Comparison ROC curves based on the results of ELSA were constructed to confirm diagnostic efficiency of novel biomarkers for lung cancer. Compared to the existing clinical markers (CEA, CA199 and Cyfra21-1), BAI-1, E-Cadherin and TSP-1 showed high AUC (Fig. 5A, BAI-1: 0.911, 95% CI 0.891–0.930; E-Cadherin: 0.959, 95% CI 0.945–0.972; TSP-1: 0.930, 95% CI: 0.911–0.949), while IBSP showed no advantage (0.797, 95% CI: 0.767–0.827). We further found that even the combination of CEA, CA199 and Cyfra21-1 (0.803 95% CI: 0.718–0.888) could not compare with the individual diagnostic efficacy of BAI-1 (0.899, 95% CI: 0.842–0.956, $p = 0.33$), E-Cadherin (0.981, 95% CI: 0.963–1.000, $p < 0.01$) and TSP-1 (0.905, 95% CI: 0.841–0.968, $p < 0.05$) (Table 1A, Fig. 5B). Binary logistic regression analysis showed that combination of BAI-1, E-Cadherin, IBSP and TSP-1 had higher diagnostic accuracy (AUC: 0.995, 95% CI: 0.992–0.998) (Fig. 5C left, yellow line), as well as higher sensitivity (ADC: 98.8%, SCC: 98.5%) and specificity (ADC: 97%, SCC: 97.6%) in both ADC and SCC patients (Fig. 5B middle and right, red line), indicating the diagnosis value of the novel biomarkers in both ADC and SCC (Table 1B).

4. Discussion

Compared with 388 healthy volunteer controls, 371 lung cancer patients in this study showed significantly higher levels of BAI-1, E-Cadherin and IBSP, but lower TSP-1 ($p < 0.0001$ for all). The ROC analysis showed that BAI-1, E-Cadherin and TSP-1 were more sensitive and specific than classic biomarkers CA199, CEA and Cyfra21-1. The combination diagnostic efficacy of the four markers was also much better than that of CEA, CA199 and Cyfra21-1. These results demonstrated that BAI-1, E-Cadherin, TSP-1 and IBSP, as well as their combination, should be promising novel biomarkers in lung cancer diagnosis.

Such results were in agreement with previous literature that these

four biomarkers either suppress tumor (BAI1, E-Cadherin and TSP-1) or promote carcinogenesis (IBSP). BAI1 is a member of G-protein-coupled receptors (GPCRs) that acts as a suppressor in carcinogenesis [26]. Hatanaka et al. and Fujii et al. demonstrated that BAI1 dramatically inhibited tumor angiogenesis and suppressed metastasis in malignant glioma and kidney cell lines [27,28]. E-Cadherin is an anti-tumor factor that can inhibit cancer metastasis by suppressing epithelial-mesenchymal transition [29]. Increased E-Cadherin in tumor tissues is also found associated with better overall survival in lung adenocarcinoma patients [30]. IBSP is a structural protein of the bone matrix that promotes tumorigenesis and tumor metastasis. Increased IBSP expression is associated with more advanced tumor grade and poorer prognosis in high-grade glioma patients [31]. TSP-1 is an anti-angiogenesis glycoprotein which inhibits the *Kras*-driven carcinogenesis of lung cancer in mouse models [32].

Among all possible biomarkers, CTCs are indicators only limited to advances progression and therapy response in patients with metastasis status. The problem of CTC in early lung cancer diagnosis is low signal-noise ratio [33]. The ctDNA represents somatic mutations in growth factor receptors, tumor suppressor genes and oncogenes, so it is useful in predicting relapse of anti-tumor but not efficient in early diagnosis [34]. Seidel et al analyzed modified nucleotides in urine from 92 miners, who are high-risk population for lung cancer, to evaluate the variation of nucleosides levels by multilayer perceptron (MLP) classifier and reversed-phase high-pressure liquid chromatography (RP-HPLC) assay. Their results indicated that 18 nucleotides and metabolites (pseudouridine, m1A, uric acid, m1G, m2G, N2, N2Gua et al.) were significantly higher in urine from uranium miners than in healthy controls, which showed 97% sensitivity and 85% specificity in prediction of lung cancer [35]. MicroRNAs (miRNAs) are a large family of tissue specific small non-coding RNAs, which are 19–24 nucleotides long and conserved in evolution. They improve development and differentiation of tissues and organs, and are aberrantly expressed in many malignant tumors. Zhu et al. detected significant variations of miR-29c and miR-429 levels in serum of NSCLC patients (70 patients and 48 healthy controls). The ROC analysis results were 72.3% for miR-29c and 72.7% for miR-429, both higher than that for carcinoma embryonic antigen (53.4%) in diagnosis of stage I NSCLC [36]. Combinations of

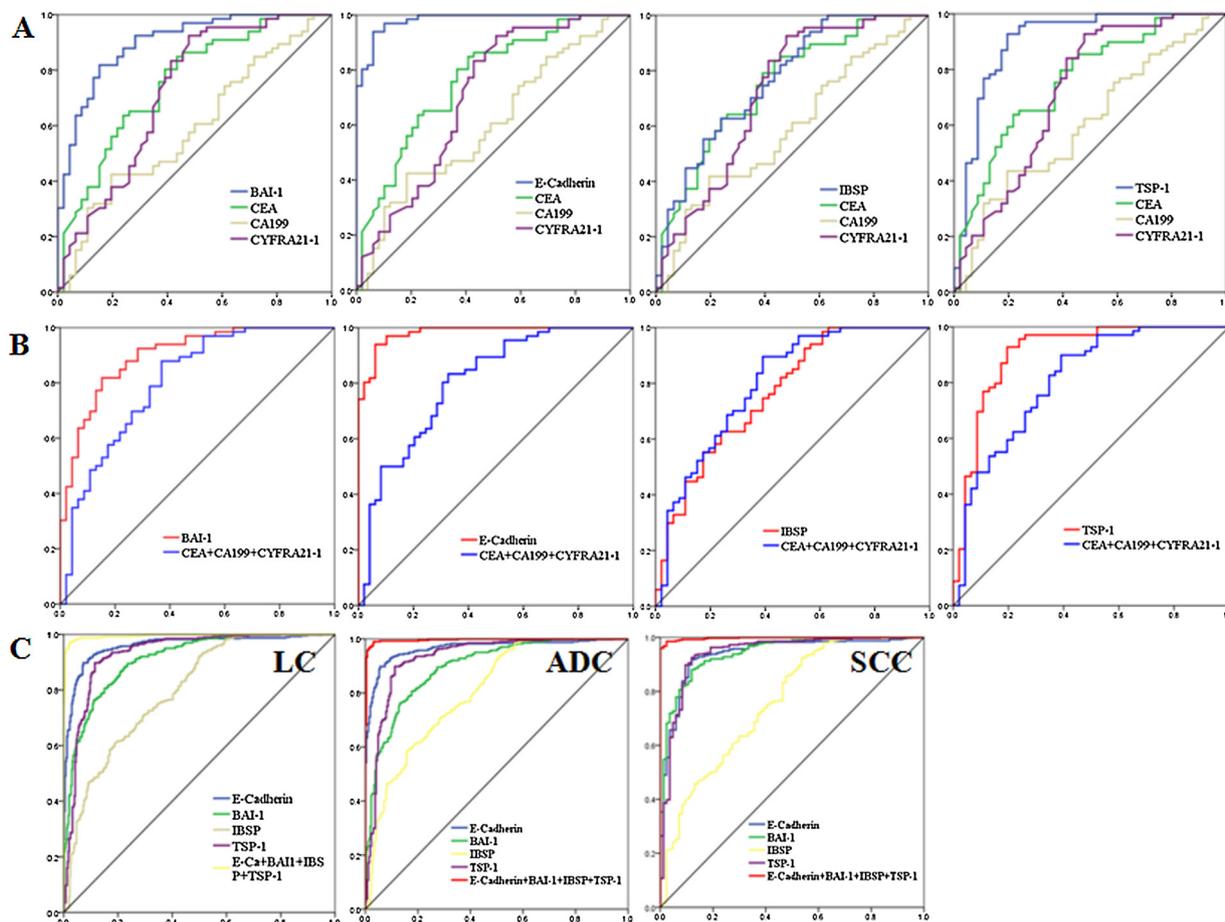


Fig. 5. ROC curve of different models of lung cancer patients.

A: comparison of novel and classical biomarkers, B: comparison of novel biomarkers with combination of classical biomarkers; C: diagnosis efficiency of combination of novel biomarkers. LC: all lung cancer, ADC: adenocarcinoma, SCC: squamous carcinoma.

circulation biomarkers are also regarded to help lung cancer diagnosis and predict prognosis. Isocitrate dehydrogenase 1, together with CA125, Cyfra21-1 and CEA, was found to be greatly helpful in NSCLCs diagnosis with AUC 89.6%, SN60.4% and SP95.6% [37]. A panel including CEA, retinol binding protein (RBP), 1-antitrypsin (ATT) and squamous cell carcinoma antigen (SCC) showed 89.3% SN and 84.7% SP [38]. Yee and others indicated that connective tissue-activating peptide III/neutrophil activating protein-2 (CTAPIII/NAP-2) combined with haptoglobin improved the accuracy of a lung cancer risk prediction model including age, smoking and lung function [39]. The auto-antibody detection was also reported to complement the diagnostic utility with blood protein biomarkers [40,41]. The major limitations of these researches were lack of sufficient SN/SP, and the sample sizes were usually small (< 300 cases).

Compared with former studies, our study had several important advantages. Firstly, the RayBio antibody array was used to screen lung cancer plasma including 1000 secreted molecules to identify 22 potential diagnostic biomarkers. Then the identified biomarkers were checked in a sample cohort (759 cases covering all stages of lung cancer) large enough to make possible validation of the screened novel biomarkers. The second advantage of this study was that it reported, to the best of our knowledge, for the first time the enhanced circulating level of BAI-1 in lung cancer patients. Similar with our results, former studies also indicated elevated E-Cadherin level in lung cancer serum [42,43], increased expression of IBSP [44] and inhibited level of TSP-1 in blood of lung cancer patients [45,46], but these studies were of smaller cohorts (< 100 cases). Thirdly, the four novel biomarkers showed significant differences in early lung cancer patients (stages

I + II). All these results suggested that these novel biomarkers, in combination with imaging and clinical examination information, may provide a better panel for rapid and economic diagnosis for early lung cancer.

The major limitation of this study was the lack of relationship between biomarker levels and the survival status of lung cancer patients. This was because most patients (about 300 cases) enrolled in this study was within one year of diagnosis of lung cancer, therefore no prognostic information was yet available. The second limitation was that the majority of participants in our study were from Southwest China. To encompass the complex causes of lung cancer and the differences in the genomic variations, future studies should include patients from different regions.

One strange result in this study was that the tumor suppressors BAI1 and E-Cadherin were found to be elevated in lung cancer plasma. We speculated the reason was that BAI1 and E-cadherin were excluded from tumor cells and released into the circulating system to create an ideal microenvironment for tumor growth or invasion. But this hypothesis needs confirmation by further mechanism research.

In conclusion, our results indicated the four biomarkers panels will be helpful for early lung cancer diagnosis and high risk population screening of lung cancer.

Funding

Supported by Science and Technology Support Program of Sichuan Province China (2014SZ0148 and 2014SZ0126).

Table 1
Diagnostic efficiency analysis.

A. Comparison of diagnostic efficiency between novel biomarkers and classical biomarkers				
	AUC(95%CI)	SN (%)	SP (%)	P Value
BAI-1	0.899(0.842–0.956)	81.8	84.7	0.33
E-Cadherin	0.981(0.963–1.000)	93.9	93.8	< 0.01
IBSP	0.767(0.677–0.856)	62.6	76.0	0.09
TSP-1	0.905(0.841–0.968)	92.7	80.4	< 0.05
CEA + CA199 + Cyfra21-1 ref.	0.803(0.718–0.888)	87.8	63.0	–
CEA	0.775(0.664–0.847)	84.8	56.5	–
CA199	0.589(0.482–0.696)	42.4	80.4	–
Cyfra21-1	0.721(0.620–0.823)	92.7	52.1	–
B. Diagnostic efficiency of novel biomarkers in all lung cancer, ADC and SCC patients				
	AUC(95%CI)	SN (%)	SP (%)	
All LC vs Control				
BAI-1 + E-Cadherin + IBSP + TSP-1	0.992(0.987–0.997)	97.0	97.7	
BAI-1	0.911(0.891–0.930)	79.2	84.5	
E-Cadherin	0.959(0.945–0.972)	91.1	89.7	
IBSP	0.797(0.767–0.827)	80.9	59.5	
TSP-1	0.930(0.911–0.949)	82.6	93.6	
ADC vs. control				
BAI-1 + E-Cadherin + IBSP + TSP-1	0.993(0.991–0.996)	97.9	97.0	
BAI-1	0.895(0.871–0.919)	77.9	84.0	
E-Cadherin	0.965(0.952–0.978)	92.6	89.2	
IBSP	0.803(0.769–0.837)	81.0	61.1	
TSP-1	0.930(0.908–0.952)	82.8	92.7	
SCC vs. control				
BAI-1 + E-Cadherin + IBSP + TSP-1	0.981(0.967–0.992)	95.0	96.1	
BAI-1	0.896(0.820–0.931)	70.2	89.5	
E-Cadherin	0.963(0.911–0.978)	95.6	84.3	
BSP	0.760(0.700–0.820)	77.3	61.2	
TSP-1	0.962(0.920–0.974)	88.9	91.3	

LC: lung cancer; ADC: adenocarcinoma; SCC: squamous carcinoma.

Declaration of Competing Interest

The authors declare no conflicts of interest

Acknowledgement

We thanks medical record department for providing information of patients.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.lungcan.2019.06.019>.

References

- [1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2015, *CA Cancer J. Clin.* 65 (2015) 5–29.
- [2] W.C. Chen, R.S. Zheng, P.D. Baade, S.W. Zhang, H.M. Zeng, F. Bray, A. Jemal, X.Q. Yu, Jie He, Cancer statistics in China, 2015, *CA Cancer J. Clin.* 66 (2016) 115–132.
- [3] C.I. Henschke, D.F. Yankelevitz, D.M. Libby, M.W. Pasmantier, J.P. Smith, O.S. Miettinen, Survival of patients with stage I lung cancer detected on CT screening, *N. Eng. J. Med.* 355 (2006) 1763–1771.
- [4] P. Goldstraw, J. Crowley, K. Chansky, D.J. Giroux, P.A. Groome, R. Rami-Porta, P.E. Postmus, V. Rusch, L. Sobin, The IASLC lung cancer staging project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumors, *J. Thorac. Oncol.* 2 (2007) 706–714.
- [5] R. Siegel, D. Naithadham, A. Jemal, Cancer statistics, 2013, *CA Cancer J. Clin.* 63 (2013) 11–30.
- [6] A. McWilliams, B. Lam, T. Sutedja, Early proximal lung cancer diagnosis and treatment, *Eur. Respir. J.* 33 (2009) 656–665.
- [7] J.L. Mulshine, T.A. D'Amico, Issues with implementing a high-quality lung cancer screening program, *CA Cancer J. Clin.* 64 (2014) 351–363.
- [8] D.R. Alberle, 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.
- [9] P.M. Boiselle, Computed tomography screening for lung cancer, *JAMA* 309 (2013) 1163–1170.
- [10] D.R. Alberle, C.D. Berg, W.C. Black, T.R. Church, R.M. Fagerstrom, B. Galen, I.F. Gareen, C. Gatsonis, J. Goldin, J.K. Gohagan, B. Hillman, C. Jaffe, B.S. Kramer, et al., The national lung screening trials: overview and study design, *Radiology* 258 (2011) 243–253.
- [11] P.B. Bach, J.N. Mirki, T.K. Oliver, C.G. Azzoli, D.A. Berry, O.W. Brawley, T. Byers, G.A. Colditz, M.K. Gould, J.R. Jett, A.L. Sabichi, R. Smith-Bindman, D.E. Wood, et al., Benefit and harms of CT screening for lung cancer: a systematic review, *JAMA* 307 (2012) 2418–2429.
- [12] J. Vansteenkiste, C. Dooms, C. Mascaux, K. Nackaerts, Screening and early detection of lung cancer, *Ann. Oncol. (Suppl. 10)* (2012) x320–x327.
- [13] Y. Sakao, S. Tomimitus, Y. Takeda, M. Natsuaki, T. Itoh, Carcinoembryonic antigen as a predictive factor for postoperative tumor relapse in early-stage adenocarcinoma, *Eur. J. Cardio-Thorac. Surg.* 25 (2004) 520–522.
- [14] K.J. Martin, M.V. Fournier, G.P. VeerReddy, A.B. Pardee, A need for basic research on fluid-based early detection biomarkers, *Cancer Res.* 70 (2010) 5203–5206.
- [15] M.K. Showe, A. Vachani, A.V. Kossenkov, M. Youself, C. Nichols, E.V. Nikonova, C. Chang, J. Kucharczuk, B. Tran, E. Wakeam, T.A. Yie, D. Speicher, W.N. Rom, et al., Gene expression profiles in peripheral blood mononuclear cells can distinguish patients with non-small cell lung cancer from patients with nonmalignant lung diseases, *Cancer Res.* 69 (2009) 9202–9210.
- [16] F. Alatas, O. Alatas, M. Metintas, O. Colak, E. Harmanci, S. Demir, Diagnostic value of CEA, CA153, CA199, Cyfra21-1, NSE and TSA assay in pleural effusion, *Lung Cancer* 31 (2001) 9–16.
- [17] R. Molina, J.M. Auge, J.M. Escudero, R. Marrades, N. Vinolas, E. Carcereny, J. Ramirez, X. Filella, CA125, CA199, CA153 and TAG-72.3 as tumor markers in patients with lung cancer: comparison with Cyfra21-1, CEA, SCC and NSE, *Tumour Biol.* 29 (2008) 371.
- [18] P. Iyengar, M.S. Tsao, Clinical relevance of molecular markers in lung cancer, *Surg. Oncol.* 11 (2002) 167–169.
- [19] M.D. Seemann, T. Beinert, H. Furst, U. Fink, An evaluation of the tumor markers carcinoembryonic antigen (CEA), cytokeratin marker (Cyfra21-1) and neuron-specific enolase (NSE) in the differentiation of malignant from benign solitary pulmonary lesions, *Lung Cancer* 26 (1999) 149–155.
- [20] R. Sargia, D. Harpole, J.E. Herndon, E. Pisick, A. Elias, A.T. Skarin, Role of serum tumor markers CA125 and CEA in non-small cell lung cancer, *Anticancer Res.* 21 (2001) 1241–1246.
- [21] E. Jantus-Lewintre, M. Uso, E. Sanmartin, S. Camps, Update on biomarkers for the detection of lung cancer, *Lung Cancer Targets Ther.* 3 (2012) 21–29.
- [22] H. Korkaya, G. Kim, A. Davis, F. Malik, N.L. Henry, S. Ithimakin, A.A. Quraishi,

- N. Tawakkoi, R. D'Angelo, A.K. Paulson, S. Chung, T. Luther, H.J. Patholak, et al., Activation of an IL6 inflammatory loop mediates trastuzumab resistance in HER2+ breast cancer by expanding the cancer stem cell population, *Mol. Cell* 47 (2012) 570–584.
- [23] M.Y. Kim, T. Oskarsson, S. Acharyya, D.X. Nguyen, X.H. Zhang, L. Norton, J. Massague, Tumor self-seeding by circulating cancer cells, *Cell* 139 (2009) 1315–1326.
- [24] W.J. Chen, C.C. Ho, Y.L. Chang, H.Y. Chen, C.A. Lin, T.Y. Ling, Yuan S.S. Yu SL, Y.J. Chen, C.Y. Lin, S.H. Pan, H.Y. Chou, Y.J. Chen, et al., Cancer-associated fibroblasts regulate the plasticity of lung cancer stemness via paracrine signaling, *Nat. Commun.* 25 (2014) 3472.
- [25] D. Johnson, J.H. Schiller, Recent clinical advances in lung cancer management, *J. Clin. Oncol.* 32 (2014) 973–982.
- [26] S.M. Cork, E.G. Van Meir, Emerging roles for the BAI1 protein family in the regulation of phagocytosis, synaptogenesis, neurovasculature, and tumor development, *J. Mol. Med. (Berl.)* 89 (2011) 743–752.
- [27] H. Hatanaka, Y. Oshika, Y. Abe, Y. Yoshida, T. Hashimoto, A. Handa, H. Kijima, H. Yamazaki, H. Inoue, Y. Ueyama, M. Nakamura, Vascularization is decreased in pulmonary adenocarcinoma expressing brain-specific angiogenesis inhibitor 1 (BAI1), *Int. J. Mol. Med.* 5 (2000) 181–183.
- [28] K. Fujii, K. Kurozumi, T. Ichikawa, M. Onishi, Y. Shimazu, J. Ishida, E.A. Chiocca, B. Kaur, I. Date, The integrin inhibitor cilengitide enhances the anti-glioma efficacy of vasculostatin-expressing oncolytic virus, *Cancer Gene Ther.* 20 (2013) 437–444.
- [29] Q.P. Hu, J.Y. Kuang, Q.K. Yang, X.W. Bian, S.C. Yu, Beyond a tumor suppressor: soluble E-cadherin promotes the progression of cancer, *Int. J. Cancer* 138 (2016) 2804–2812.
- [30] L. Wen, L. Sun, Y. Xi, X. Chen, Y. Xing, W. Sun, Q. Meng, L. Cai, et al., Expression of calcium sensing receptor and E-cadherin correlated with survival of lung adenocarcinoma, *Thorac. Cancer* 6 (2015) 754–760.
- [31] T. Xu, R. Qin, J. Zhou, Y. Yan, Y. Lu, X. Zhang, D. Fu, Z. Lv, W. Li, C. Xia, G. Hu, X. Ding, J. Chen, High bone sialoprotein (BSP) expression correlates with increased tumor grade and predicts a poorer prognosis of high-grade glioma patients, *PLoS One* 7 (2012) e48415.
- [32] K.H. Baek, D. Bhang, A. Zaslavsky, L.C. Wang, A. Vachani, C.F. Kim, S.M. Albelda, G.I. Evan, S. Ryeom, Thrombospondin-1 mediates oncogenic Ras-induced senescence in premalignant lung tumors, *J. Clin. Invest.* 123 (2013) 4375–4389.
- [33] K.J. Medicine, Cancer's circulation problem, *Science* 327 (2010) 1072–1074.
- [34] K. Schmidt, F. Diehl, A blood-based DNA test for colorectal cancer screening, *Discov. Med.* 7 (2007) 7–12.
- [35] A. Seidel, P. Seidel, O. Manuwal, O. Herbarth, Modified nucleosides as biomarkers for early cancer diagnose in exposed populations, *Environ. Toxicol.* 30 (2015) 956–967.
- [36] W.Y. Zhu, J.Y. He, D.D. Chen, B. Zhang, L. Xu, H. Ma, X. Liu, Y. Zhang, H. Le, Expression of miR-29c, miR-93 and miR-429 as potential biomarkers for detection of early stage non-small cell lung cancer, *PLoS One* 19 (2014) e87780.
- [37] N. Sun, Z.L. Chen, F.W. Tan, B. Zhang, R. Yao, C. Zhou, J. Li, Y. Gao, Z. Liu, X. Tan, F. Zhou, M.Y. He, K. Shao, et al., Isocitrate dehydrogenase 1 is a novel plasma biomarker for the diagnosis of non-small cell lung cancer, *Clin. Cancer Res.* 19 (2013) 5136–5145.
- [38] E.F. Patz, M.J. Campa, E.B. Gottlin, I. Kusmartseva, X.R. Guan, J.E. Herndon, Panel of serum biomarker for early lung cancer detection, *J. Clin. Oncol.* 25 (2007) 5578–5583.
- [39] J. Yee, M.D. Sadar, D.D. Sin, M. Kuzyk, L. Xing, J. Kondra, A. McWilliams, S.F. Man, S. Lam, Connective tissue-activating peptide III: a novel blood biomarkers for early lung cancer detection, *J. Clin. Oncol.* 27 (2009) 2787–2792.
- [40] 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. Thorngquist, G.E. Goodman, et al., Occurrence of autoantibodies to annexinI, 14-3-3 theta and LAMR1 in prediagnostic lung cancer sera, *J. Clin. Oncol.* 26 (2008) 5060–5066.
- [41] L. Wu, W. Chang, J. Zhao, Y. Yu, X. Tan, T. Su, L. Zhao, S. Huang, S. Liu, G. Cao, Development of autoantibody signatures as novel diagnostic biomarkers of non-small cell lung cancer, *Clin. Cancer Res.* 16 (2010) 3760–3768.
- [42] A. Gogali, K. Charalabopoulos, I. Zampira, A.K. Konstantinidis, F. Tachmazoglou, G. Daskalopoulos, Soluble adhesion molecules E-cadherin, intercellular adhesion molecule-1, and E-selectin as lung cancer biomarkers, *Chest* 138 (2010) 1173–1179.
- [43] K. Charalopoulos, A. Gogali, Y. Dalavaga, G. Daskalopoulos, M. Vassilioiu, G. Bablekos, The clinical significance of soluble E-cadherin in nonsmall cell lung cancer, *Exp. Oncol.* 28 (2006) 83–85.
- [44] N.S. Fedarko, A. Jain, A. Karadag, M.R. van Eman, L.W. Fisher, Elevated serum bone sialoprotein and osteopontin in colon, breast, prostate and lung cancer, *Clin. Cancer Res.* 7 (2001) 4060–4066.
- [45] A.Z. Dudek, H. Mahaseth, Circulating angiogenic cytokines in patients with advanced non-small cell lung cancer: correlation with treatment response and survival, *Cancer Invest.* 23 (2005) 193–200.
- [46] W.L. Bigbee, V. Gopalakrishnan, J.L. Weissfeld, D.O. Wilson, S. Dacic, A.E. Lokshin, J.M. Siegfried, A multiplexed serum biomarker immunoassay panel discriminates clinical lung cancer patients from high-risk individuals found to be cancer-free by CT screening, *J. Thorac. Oncol.* 7 (2012) 698–708.