Proteomic analysis of bronchoalveolar lavage fluid (BALF) from lung cancer patients using label-free mass spectrometry

Abduladim Hmmier [a,b,1], Michael Emmet O'Brien [c,1], Vincent Lynch [d], Martin Clynes [d], Ross Morgan [c,2], Paul Dowling [a,*,2]

[a] Department of Biology, Maynooth University, Maynooth, Co. Kildare, Ireland
[b] BioNano Integration Research Group, Biotechnology Research Centre, Tripoli, Libya
[c] Department of Respiratory Medicine, Beaumont Hospital, Dublin 9, Ireland
[d] National Institute for Cellular Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland

ARTICLE INFO

ABSTRACT

Background: Lung cancer is the leading cause of cancer-related mortality in both men and women throughout the world. The need to detect lung cancer at an early, potentially curable stage, is essential and may reduce mortality by 20%. The aim of this study was to identify distinct proteomic profiles in bronchoalveolar fluid (BALF) and plasma that are able to discriminate individuals with benign disease from those with non-small cell lung cancer (NSCLC).
Methods: Using label-free mass spectrometry analysis of BALF during discovery-phase analysis, a significant number of proteins were found to have different abundance levels when comparing control to adenocarcinoma (AD) or squamous cell lung carcinoma (SqCC). Validation of candidate biomarkers identified in BALF was performed in a larger cohort of plasma samples by detection with enzyme-linked immunoassay.
Results: Four proteins (Cystatin-C, TIMP-1, Lipocalin-2 and HSP70/HSPA1A) were selected as a representative group from discovery phase mass spectrometry BALF analysis. Plasma levels of TIMP-1, Lipocalin-2 and Cystatin-C were found to be significantly elevated in AD and SqCC compared to control.
Conclusion: The results presented in this study indicate that BALF is an important proximal biofluid for the discovery and identification of candidate lung cancer biomarkers.
General significance: There is good correlation between the trend of protein abundance levels in BALF and that of plasma which validates this approach to develop a blood biomarker to aid lung cancer diagnosis, particularly in the era of lung cancer screening. The protein signatures identified also provide insight into the molecular mechanisms associated with lung malignancy.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Lung cancer is the most common cancer worldwide, in 2012 it contributed to 13% of the total number of new cases diagnosed and is the most common cause of cancer-related mortality, accounting for more than 1.4 million deaths per year globally [1]. The overall prognosis remains poor with just over one in eight lung cancer patients living for five-years or more after their diagnosis. Approximately 85% to 90% of patients with lung cancer have had direct exposure to tobacco [2,3], other risk factors include environmental smoke exposure and occupational exposure to agents such as asbestos [4].

Lung cancer is categorised into two groups, small cell (SCLC) and non-small cell (NSCLC). The majority of cases are NSCLC (85%) of which 40% are adenocarcinoma (AD), 25–30% squamous cell carcinoma (SqCC), 10–15% large cell carcinoma (LCC), and then rarer variants such as mixed/undifferentiated pulmonary carcinomas [5]. Substantial progress has been made in our understanding of the biological processes and mutations that cause lung cancer, this has led to the development of targeted therapy for lung cancer with some marked improvement in survival in selected groups. Advances in stereotactic radiotherapy for early stage disease are also a promising new treatment modality for those who are unfit for surgery, which has demonstrated excellent control of local disease [6–8]. Despite this array of new targeted treatments, cure remains elusive for the vast majority of patients diagnosed with lung cancer [9].

Approximately 85% of patients with lung cancer are symptomatic at presentation, the remainder are detected by interpreting chest radiographs performed for an unrelated health issue [10]. Computed

tomography (CT) screening of high-risk individuals, particularly smokers, is helping to detect the disease in its early, more-curable stages [11]. Screening of asymptomatic individuals has been shown to decrease mortality by 20% in those aged 55–74 years who are current smokers or quit within the last 15 years [12]. Lung cancer screening has now been endorsed by a number of different medical societies and is available to many in the United States [13–15]. However concerns remain in relation to the high false-positive rate of screening, unnecessary interventions, and its overall cost effectiveness despite being implemented [16]. Efforts to further risk stratify patients clinically beyond those criteria employed in the National Lung Cancer Screening, along with technological developments in lung nodule volume measurement, may ultimately reduce the number and frequency of low dose CT scans required for population screening [17,18]. Developing a lung cancer biomarker to complement radiological imaging is a strategy that may reduce false-positive and false-negative screening rates, improve cost-effectiveness, and aid the earlier detection of this disease [19].

Biomarkers circulating in blood could constitute the gold standard for non-invasive cancer diagnostics [20]. Blood is the ideal final test medium, however due to the high dynamic range, identification of candidate protein biomarkers remains difficult. This high dynamic range, spanning 10–12 orders of magnitude, is a result of abundant proteins, such as albumin, representing more than 99% of the total bulk mass of protein content [21]. Fractionation techniques, such as immunodepletion, can reduce this dynamic range, however biomarker discovery in blood remains a challenge [22,23].

One approach to overcome this problem is to analyse more proximal biofluids such as saliva, urine, exhaled breath condensate and bronchoalveolar lavage fluid (BALF) [24–28]. Proximal biofluids have a much reduced dynamic range of protein abundances and in some cases are in direct contact with the site of the disease [29]. BALF is routinely sampled during bronchoscopy in individuals with suspected lung cancer, it contains a wide variety of cellular material such as macrophages and neutrophils, a large number of proteins produced by epithelial and inflammatory cells, and tumour cells if present [30].

In the present study, our aim was to identify altered levels of candidate protein biomarkers in BALF samples from individuals with lung cancer compared to a control group diagnosed with benign nodules or sarcoidosis. Using quantitative mass spectrometry, we identified significantly changed proteins and catalogued their biological processes and molecular functions. With this information, a number of candidate biomarkers were selected for verification and initial validation in blood samples from lung cancer and control groups. This study may enhance our knowledge of the proteomic profile induced in the lung microenvironment by the presence of tumour cells, and may help elucidate the protein signature created by the tumour cells themselves. Furthermore, this data may facilitate the identification of a useful clinical biomarker panel for the diagnosis and monitoring of lung cancer in the future.

2. Material and methods

2.1. Patient information and sample collection

Patient samples were collected according to standard phlebotomy procedures following informed consent. Institutional ethical approval was granted from the Beaumont Hospital, Ethics Committee (REC reference: 10/61). BALF and blood samples were collected from patients who were undergoing diagnostic bronchoscopy in Beaumont Hospital, Dublin, Ireland. All samples were immediately placed on ice prior to processing and stored at −80 °C until analysis. In total, 42 BALF and 72 plasma samples were analysed (Table 1). After being extracted from the respiratory airways, the majority of BALF samples appeared to be clear and any samples with a slightly reddish appearance due to blood contamination were excluded. Post centrifugation, any BALF samples forming a distinct layer of red blood cells were excluded. Post mass spectrometry analysis, all samples were searched individually to

Table 1
Patient data.

BALF	Plasma
Control (benign & sarcoid)	Control (benign & sarcoid)
n = 16	n = 26
Average age = 56	Average age = 49
8 male/8 female	18 male/8 female
NSCLC	NSCLC
n = 26	n = 46
Average age = 65	Average age = 67
13 male/13 female	23 male/23 female
Adenocarcinoma, n = 13	Adenocarcinoma, n = 28
7 stage I/II; 6 stage III/IV	14 stage I/II; 14 stage III/IV
Squamous cell carcinoma, n = 13	Squamous cell carcinoma, n = 18
8 stage I/II; 5 stage III/IV	7 stage I/II; 11 stage III/IV

Patient data table including number of samples, average age, male/female and disease staging for controls and lung cancer (adenocarcinoma and squamous cell carcinoma).

identify associated proteins. Any samples presenting with a large number of blood based proteins as the top hits during protein identification (with a similar profile to digested serum/plasma samples) were also excluded from the main analysis.

2.2. Mass spectrometry

To remove debris and cells, BALF samples were centrifuged at $4000 \times g$ for 15 min at 4 °C. Proteins from the resulting supernatants were concentrated using 5 kDa Amicons (Millipore) by centrifugation at $4000 \times g$ for 45 min at 4 °C. 200 μl of the resulting retenate was treated using the ReadyPrep 2-D clean-up kit (Bio-Rad) to precipitate protein. The isolated BALF proteins were resuspended in buffer containing 8 M urea/50 mM NH4HCO3/0.1% ProteaseMax. The protein amount was estimated using an RC/DC protein assay from Bio-Rad [31]. Mass spectrometry analysis was performed according to a previously optimised [32].

2.3. Label-free analysis

Progenesis label-free LC-MS software version 3.1 from Non-Linear Dynamics (Newcastle upon Tyne, UK) was used to process the raw data generated from LC-MS/MS analysis. Data alignment was based on the LC retention time of each sample, allowing for any drift in retention time given and adjusted retention time for all runs in the analysis. A reference run was established with the sample run that yielded most features (i.e. peptide ions). The retention times of all of the other runs were aligned to this reference run and peak intensities were then normalized [33]. A number of criteria were applied to ensure proper identification of BALF derived proteins, including an ANOVA *p*-value between experimental groups of ≤0.05, fold change ≥2 and proteins with ≥2 peptides matched [34].

2.4. ELISA analysis

Human HSP70/HSPA1A DuoSet ELISA; DY1663-05, Human Cystatin C DuoSet ELISA; DY1196, Human TIMP-1 DuoSet ELISA; DY970-05 & Human Lipocalin-2/NGAL DuoSet ELISA; DY1757 were used for the verification study (R&D Systems, Oxon, UK). Each of these ELISA assays was performed according to their individual manufacturer's protocol and guidelines. The concentration of each protein in the serum samples was measured by comparing the optical density (OD) using a microplate reader (Bio-Tek). Standard curves were calculated for each ELISA kit.

2.5. Bioinformatics analysis of potential protein interactions

The freely available software package PANTHER (http://pantherdb.org) was used for the bioinformatics analysis of the mass spectrometric findings generated in this study [35].

3. Results

3.1. Proteomic profiling and identification

In total, 267 proteins were found to be statistically significant when comparing control to AD BALF samples (fold-change ≥2, number of match peptides ≥2, p-value ≥0.05). Of these 267 significant proteins (Table 2 – top 30), 39 were found to have a higher abundance in the control cohort, mainly involved in cellular processes (27.90%), metabolic processes (29.40%) and cellular component organization (22.10%) (Fig. 1). 228 proteins were of higher abundance in the AD samples and involved in cellular (24.30%) and metabolic (23.10%) processes. 261 proteins were found to be statistically significant when comparing control to SqCC BALF samples (fold-change ≥2, number of match peptides ≥2, p-value ≥0.05). Of these 261 significant proteins (Table 3 – top 30), 62 were found to have a higher abundance in the control cohort, with cellular (20.70%) and metabolic (29.30%) processes extensively represented (Fig. 1). 199 proteins were of higher abundance in the SqCC samples (cellular process, 24.40%; metabolic process, 22.70%). 292 proteins were found to be statistically significant when comparing AD to SqCC BALF samples (fold-change ≥2, number of match peptides ≥2, p-value ≥0.05). Of these 292 significant proteins (Table 4 – top 30), 183 were found to have a higher abundance in the AD cohort (cellular process, 22.60%; metabolic process, 22.90%) (Fig. 1). 109 proteins were of higher abundance in the SqCC samples (cellular process, 30.70%; metabolic process, 25.50%).

3.2. ELISA data

Four proteins (Cystatin-C, TIMP-1, Lipocalin-2 and HSP70/HSPA1A) were selected as a representative group from the discovery phase experiments in BALF to verify in patient plasma samples. Commercial available ELISA kits were available to evaluate the abundance of these proteins in the verification plasma sample cohort. Cystatin-C, TIMP-1, Lipocalin-2 and HSP70/HSPA1A passed the following criteria for selection: p-value ≤0.05, protein fold change >2 and proteins with ≥2 peptides matched. Label-free mass spectrometry analysis of BALF samples found that Cystatin-C was elevated in AD compared to SqCC (4.8-fold; p-value = 4.89E-08) and elevated in SqCC compared to control (4.5-fold; p-value = 0.01) (Tables – supplemental). TIMP-1 was discovered to be elevated in AD compared to SqCC (2.2-fold; p-value = 0.0001), elevated in SqCC compared to control (2.2-fold; p-value = 0.02) and elevated in AD compared to control (2-fold; p-value = 2.39E-05) (Tables – supplemental). Lipocalin-2 was found to have an increased abundance in SqCC compared to control (2.2-fold; p-value = 0.006) (Tables – supplemental). HSP70/HSPA1A was increased in AD compared to SqCC (2.2-fold; p-value = 0.001) and in control compared to SqCC (2.1-fold; p-value = 0.002) (Tables – supplemental).

Verification of these four proteins (Cystatin-C, TIMP-1, Lipocalin-2 and HSP70/HSPA1A) was performed in plasma samples using sandwich ELISAs to measure natural and recombinant protein levels. The mean concentration for Cystatin-C was 716 ng/ml, 8027 ng/ml and 6981 ng/ml in control, AD and SqCC respectively with associated Kruskal-Wallis one-way ANOVA p-values of 1.53E-07, 7.13E-06 and 0.0002 (Fig. 2). TIMP-1 was discovered to have mean concentration values of 2004 ng/ml, 3927 ng/ml and 8280 ng/ml in control, AD and SqCC respectively with associated p-values of 2.20E-06, 3.81E-07 and 4.86E-06 (Fig. 2). Mean concentration values for Lipocalin-2 were found to be 980 ng/ml, 1404 ng/ml and 1537 ng/ml in control, AD and SqCC respectively with associated p-values of 0.02, 0.03 and 0.1 (Fig.

Table 2
Control v AD proteomics data.

Gene name	Peptide count	Confidence score	Anova (p)	Max fold change	Highest mean condition	Lowest mean condition	Description
KV402	2	10	1.31E-14	7.7	AD	Control	Ig kappa chain V-IV region Len
KV101	3	20	5.22E-12	7.9	AD	Control	Ig kappa chain V-I region AG
KV121	2	9	7.17E-12	6.5	AD	Control	Ig kappa chain V-I region Ni
HV302	2	7	2.24E-10	7.6	AD	Control	Ig heavy chain V-III region WEA
GNAI1	3	8	3.24E-10	7.3	AD	Control	Guanine nucleotide-binding protein G(i) subunit alpha-1
HV103	3	10	7.06E-10	5.7	AD	Control	Ig heavy chain V-I region V35
CALM	6	42	8.76E-10	9.6	AD	Control	Calmodulin
KV110	2	9	6.59E-09	3.6	AD	Control	Ig kappa chain V-I region HK102 (Fragment)
KV104	3	17	9.96E-09	9.5	AD	Control	Ig kappa chain V-I region CAR
KV201	2	10	3.60E-08	11.1	AD	Control	Ig kappa chain V-II region Cum
ANXA2	9	41	3.98E-08	3.4	AD	Control	Annexin A2
SLPI	10	63	4.95E-08	7.6	AD	Control	Antileukoproteinase
PRDX6	4	11	6.76E-08	4.1	AD	Control	Peroxiredoxin-6
S10A2	2	14	7.38E-08	15.5	AD	Control	Protein S100-A2
IGLL5	8	44	8.12E-08	2.5	AD	Control	Immunoglobulin lambda-like polypeptide 5
PSA4	2	6	8.31E-08	8.2	AD	Control	Proteasome subunit alpha type-4
HV320	3	9	8.64E-08	6.2	AD	Control	Ig heavy chain V-III region GAL
B2MG	3	17	1.10E-07	7.2	AD	Control	Beta-2-microglobulin
KV302	5	25	1.29E-07	2.1	AD	Control	Ig kappa chain V-III region SIE
LDHA	6	24	1.43E-07	3.5	AD	Control	L-lactate dehydrogenase A chain
DCD	2	13	1.68E-07	3.6	AD	Control	Dermcidin
CD59	3	19	2.24E-07	11.3	AD	Control	CD59 glycoprotein
DMBT1	20	101	2.36E-07	10.9	AD	Control	Deleted in malignant brain tumors 1 protein
KV309	2	9	2.65E-07	3.2	AD	Control	Ig kappa chain V-III region VG (Fragment)
LDHB	8	28	2.70E-07	5.0	AD	Control	L-lactate dehydrogenase B chain
ACTN4	9	28	3.11E-07	3.1	AD	Control	Alpha-actinin-4
KV118	2	10	3.32E-07	10.2	AD	Control	Ig kappa chain V-I region WEA
HV303	5	21	4.12E-07	3.6	AD	Control	Ig heavy chain V-III region VH26
LV103	2	6	5.63E-07	13.2	AD	Control	Ig lambda chain V-I region NEW
ACBP	6	29	6.27E-07	4.9	AD	Control	Acyl-CoA-binding protein

List of differentially expressed proteins (top 30 hits) when comparing control to AD BALF patient samples. The table includes information on gene name, peptide count, confidence score (XCorr), ANOVA (p-value), fold-change, highest/lowest condition and protein description. See supplemental tables for full list of significant proteins.
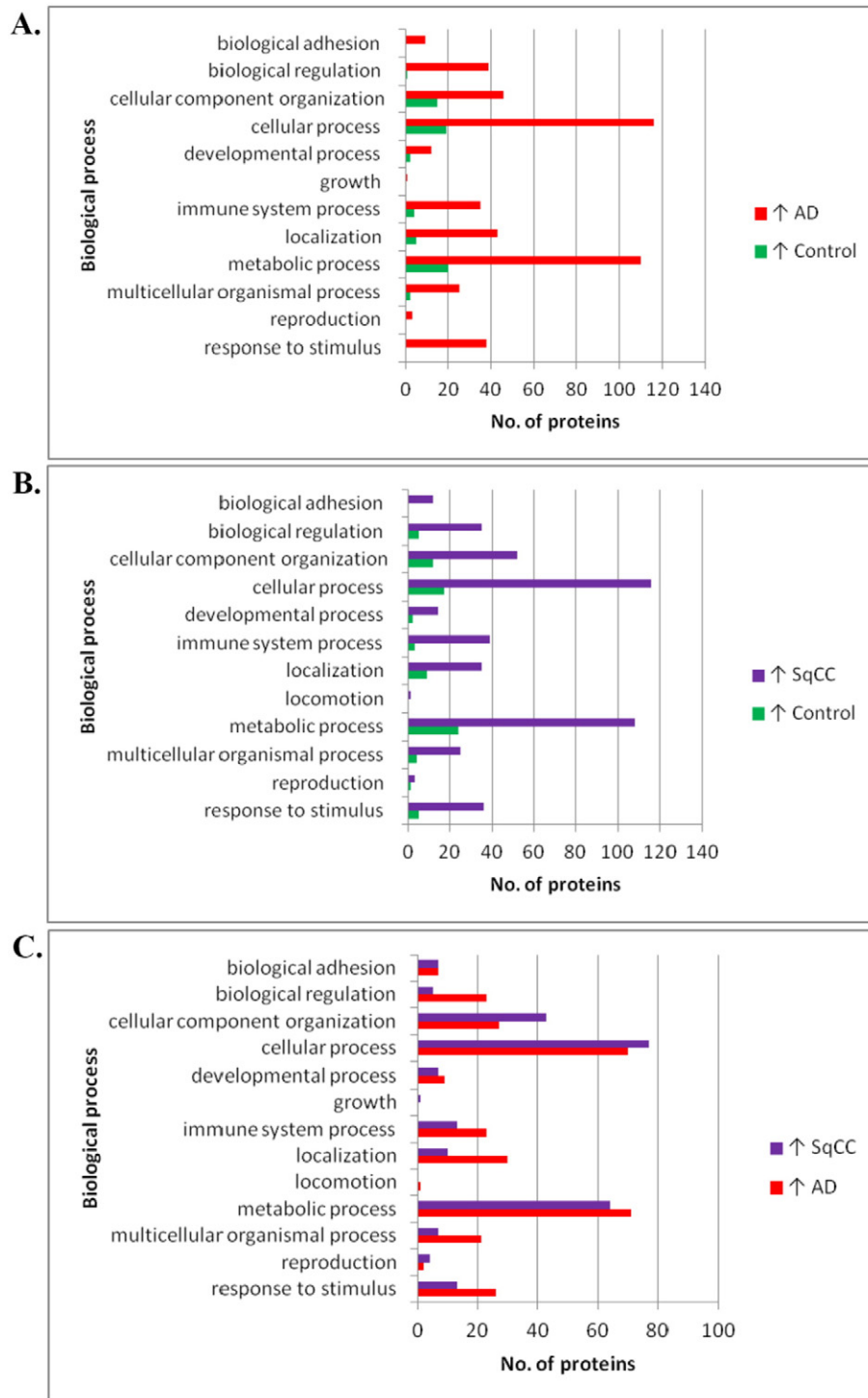
Fig. 1. Biological process analysis. Biological process analysis (Panther) for (A) control v AD, (B) control v SqCC and (C) AD v SqCC.

2). HSP70/HSPA1A was not found to be significant in any of the comparisons (control v AD; control v SqCC or AD v SqCC) with mean concentration values of 5.4 ng/ml, 4.3 ng/ml and 11.2 ng/ml in control, AD and SqCC respectively (Fig. 2).

4. Discussion

Blood is an accessible, acceptable, and widely used biological sample for identification of biomarkers of disease [36]. However, one of the greatest challenges in proteomic analysis of both serum and plasma samples is the wide range of concentration of different proteins [37].

The dynamic range of proteins in blood limits the ability to examine the blood proteome for discovery due to the presence of a few proteins at very high concentrations [38,39].

Arising from the inherent analytical challenges related to blood proteomics, proximal fluids have gained increasing attention for conducting candidate protein biomarker discovery [40]. Compared to distal biofluids (e.g. blood), proximal biofluids are less complex and are likely to be significantly enriched in potential biomarker candidates due to their vicinity to the site of disease [29,41]. Once a candidate biomarker is discovered in a proximal biofluid, targeted verification and validation can then be performed in blood.

Table 3
Control v SqCC proteomics data

Gene name	Peptide count	Confidence score	Anova (p)	Max fold change	Highest mean condition	Lowest mean condition	Description
A1AG2	4	35	1.96E-12	8.2	SqCC	Control	Alpha-1-acid glycoprotein 2
A1AT	31	273	3.45E-08	4.0	SqCC	Control	Alpha-1-antitrypsin
KV101	3	18	1.16E-07	6.0	SqCC	Control	Ig kappa chain V-I region AG
IGHG1	11	124	1.30E-07	4.7	SqCC	Control	Ig gamma-1 chain C region
KV302	2	14	5.33E-07	2.5	SqCC	Control	Ig kappa chain V-III region SIE
A2GL	4	39	5.47E-07	4.8	SqCC	Control	Leucine-rich alpha-2-glycoprotein
FIBA	5	21	7.91E-07	4.8	SqCC	Control	Fibrinogen alpha chain
KNG1	4	15	1.00E-06	4.2	SqCC	Control	Kininogen-1
CBG	4	12	1.02E-06	2.9	SqCC	Control	Corticosteroid-binding globulin
K2C5	8	39	1.50E-06	4.0	SqCC	Control	Keratin, type II cytoskeletal 5
C1S	2	7	2.06E-06	7.7	SqCC	Control	Complement C1s subcomponent
PNPH	3	9	2.31E-06	3.6	SqCC	Control	Purine nucleoside phosphorylase
HPTR	8	38	2.56E-06	8.7	SqCC	Control	Haptoglobin-related protein
MYH9	12	53	3.10E-06	5.0	SqCC	Control	Myosin-9
NUCB1	2	5	5.04E-06	4.4	Control	SqCC	Nucleobindin-1
KV401	2	6	5.39E-06	3.0	SqCC	Control	Ig kappa chain V-IV region (Fragment)
PSB10	2	10	7.40E-06	3.5	SqCC	Control	Proteasome subunit beta type-10
S10A2	2	7	7.44E-06	32.9	SqCC	Control	Protein S100-A2
IGKC	10	72	7.67E-06	2.3	SqCC	Control	Ig kappa chain C region
ACBP	2	7	9.17E-06	4.5	SqCC	Control	Acyl-CoA-binding protein
H4	6	17	9.43E-06	3.6	SqCC	Control	Histone H4
CFAH	12	64	1.01E-05	4.0	SqCC	Control	Complement factor H
K1C13	13	81	1.05E-05	31.2	SqCC	Control	Keratin, type I cytoskeletal 13
APOH	9	45	1.54E-05	3.4	SqCC	Control	Beta-2-glycoprotein 1
LDHA	7	31	1.62E-05	4.3	SqCC	Control	L-lactate dehydrogenase A chain
IGHM	7	38	1.87E-05	3.5	SqCC	Control	Ig mu chain C region
K2C8	5	21	1.92E-05	2.2	Control	SqCC	Keratin, type II cytoskeletal 8
TGM3	8	53	1.98E-05	8.4	SqCC	Control	Protein-glutamine gamma-glutamyltransferase E
ALBU	59	464	2.03E-05	4.5	SqCC	Control	Serum albumin
ANT3	8	40	2.23E-05	2.8	SqCC	Control	Antithrombin-III

List of differentially expressed proteins (top 30 hits) when comparing control to SqCC BALF patient samples. The table includes information on gene name, peptide count, confidence score (XCorr), ANOVA (p-value), fold-change, highest/lowest condition and protein description. See supplemental tables for full list of significant proteins.

Table 4
AD v SqCC proteomics data.

Gene name	Peptide count	Confidence score	Anova (p)	Max fold change	Highest mean condition	Lowest mean condition	Description
CYTC	4	22	4.89E-08	4.9	AD	SqCC	Cystatin-C
ZA2G	19	111	1.58E-07	4.2	AD	SqCC	Zinc-alpha-2-glycoprotein
TBB1	4	15	1.80E-07	6.2	AD	SqCC	Tubulin beta-1 chain
CATC	5	24	9.74E-07	3.3	AD	SqCC	Dipeptidyl peptidase 1
NHRF1	4	17	1.35E-06	3.6	AD	SqCC	Na(+)/H(+) exchange regulatory cofactor NHE-RF1
MDHC	4	13	3.96E-06	3.7	AD	SqCC	Malate dehydrogenase, cytoplasmic
K2C5	8	34	4.58E-06	4.5	SqCC	AD	Keratin, type II cytoskeletal 5
S10A8	5	29	6.35E-06	3.9	SqCC	AD	Protein S100-A8
CRIP1	3	17	7.03E-06	7.0	AD	SqCC	Cysteine-rich protein 1
ABHEB	2	6	7.28E-06	3.9	AD	SqCC	Alpha/beta hydrolase domain-containing protein 14B
CD59	3	19	9.49E-06	6.7	AD	SqCC	CD59 glycoprotein
CASPE	4	12	1.08E-05	4.6	SqCC	AD	Caspase-14
CLIC1	3	13	1.24E-05	5.0	AD	SqCC	Chloride intracellular channel protein 1
H90B2	2	10	1.46E-05	4.0	AD	SqCC	Putative heat shock protein HSP 90-beta 2
TRFE	45	295	1.48E-05	3.0	SqCC	AD	Serotransferrin
HNRPD	2	9	1.64E-05	3.0	AD	SqCC	Heterogeneous nuclear ribonucleoprotein D0
NUCL	4	16	1.93E-05	2.2	AD	SqCC	Nucleolin
ELNE	6	41	2.31E-05	11.8	SqCC	AD	Neutrophil elastase
KV118	2	7	2.60E-05	7.1	AD	SqCC	Ig kappa chain V-I region WEA
TTHY	9	75	3.13E-05	4.8	AD	SqCC	Transthyretin
MYG	2	7	4.59E-05	7.7	AD	SqCC	Myoglobin
APOH	3	12	5.12E-05	3.6	SqCC	AD	Beta-2-glycoprotein 1
PARK7	3	19	5.42E-05	4.6	AD	SqCC	Protein DJ-1
KLK11	4	17	6.09E-05	3.5	AD	SqCC	Kallikrein-11
HSP72	5	18	7.06E-05	3.1	AD	SqCC	Heat shock-related 70 kDa protein 2
PEBP1	7	35	7.48E-05	3.2	AD	SqCC	Phosphatidylethanolamine-binding protein 1
A1AG2	4	33	7.56E-05	3.1	SqCC	AD	Alpha-1-acid glycoprotein 2
HGB1A	2	11	7.65E-05	3.1	AD	SqCC	Putative high mobility group protein B1-like 1
CAP7	4	33	7.66E-05	5.8	SqCC	AD	Azurocidin
NUCB2	6	27	7.76E-05	3.4	AD	SqCC	Nucleobindin-2

List of differentially expressed proteins (top 30 hits) when comparing AD to SqCC BALF patient samples. The table includes information on gene name, peptide count, confidence score (XCorr), ANOVA (p-value), fold-change, highest/lowest condition and protein description. See supplemental tables for full list of significant proteins.
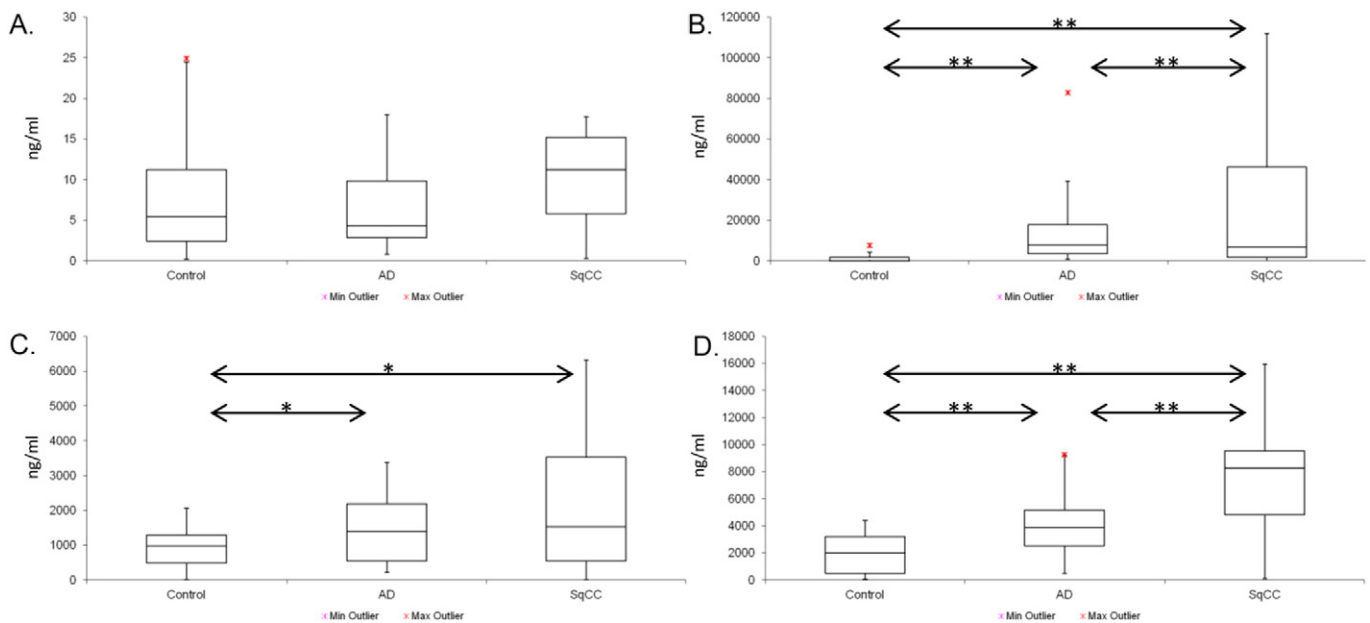
Fig. 2. ELISA verification. Box-and-Whisker plots (A) HSP70, (B) Cystatin-C, (C) Lipocalin-2 and (D) TIMP-1 levels in control, AD and SqCC plasma samples. Box-and-Whisker plots represent data with boxes ranging from the 25th to the 75th percentile of the observed distribution of values. The bold line of the box displays the median value for plasma levels of HSP70, Cystatin-C, Lipocalin-2 and TIMP-1. Whiskers span minimum to maximum observed values. Statistical significance was tested by Kruskal-Wallis one-way ANOVA using SPSS 15.0 for Windows software. ($^*p \leq 0.05$: significant; $^{**}p \leq 0.01$: very significant).

BALF is a mixture of different cell types as well as a wide variety of soluble components such as phospholipids, nucleic acids, peptides and proteins [42]. The utility of BALF has been exploited for many years in clinical research, with recent technological advances permitting detailed proteomic profiling of the protein/peptide signature in particular lung diseases. This approach provides a rich source of candidate biomarkers, in addition to providing insight into the complex pathological mediators associated with lung diseases at the molecular level [43,44].

The data presented in this study demonstrates a marked increase in the abundance of proteins involved in cellular and metabolic processes in patients with lung cancer compared to controls. In agreement with our findings, Almatroodi et al. published a list of proteins found to be increased in BALF from primary lung adenocarcinoma compared to controls using liquid chromatography-mass spectrometry [45]. Of the 33 proteins consistently overexpressed in lung adenocarcinoma samples compared to controls presented in that study, ACTN4, ANXA2, CLIC1, GRP78, H4, LKHA4, S10A8, SAMH1 were also found to be of greater abundance in AD compared to control groups in this study.

Following selection of four candidate proteins for verification in plasma, TIMP-1, Lipocalin-2 and Cystatin-C were all found to be significantly elevated in NSCLC compared to control. This indicates that there is a good level of agreement between proteins found to have different abundance levels in BALF (cancer v control) and plasma samples (cancer v control), furthermore the increase was a consistent finding irrespective of cancer subtype, i.e. AD or SqCC.

`TIMP-1 was discovered to be elevated in SqCC compared to control and AD compared to control when analysing BALF samples by mass spectrometry. Verification of these results in plasma samples found that TIMP-1 had significantly higher levels in AD and SqCC compared to the control group as measured by ELISA. TIMPs are natural inhibitors of matrix metalloproteinases (MMPs) present in most tissues and body fluids [46]. By inhibiting MMPs activities, they participate in tissue remodelling of the extracellular matrix. The balance between MMPs and TIMPs activities is involved in both normal and pathological processes such as wound healing, tissue remodelling, angiogenesis, tumour development and metastasis [47], [48]. In cancer, MMP over-expression is thought to play an important role in tumour invasion and metastasis. The major function of TIMP-1 is as an

inhibitor of MMPs, but several independent roles in tumour development are likely [49], [50].

Neutrophil gelatinase-associated lipocalin (NGAL), also known as Lipocalin-2, is a 178-amino acid protein which exists in three molecular forms, a 25-kDa monomer, a 45-kDa homodimer, and a 135-kDa heterodimer. A number of malignant tumours consistently overexpressed NGAL with increased concentration in blood, urine, and other biologic fluids [51]. In this study, BALF levels of Lipocalin-2 were increased in SqCC compared to control with mean concentration values significantly elevated in AD and SqCC compared to control during verification analysis of plasma samples. NGAL is also commonly associated with tumour size, stage, and invasiveness [52]. Ricci et al. found that there was an association between NGAL serum levels and bladder cancer stage. The author suggests that serum NGAL may be a useful non-invasive biomarker to provide clinical information for bladder cancer disease management [53]. Studies investigating the role of NGAL in lung cancer have discovered high levels of NGAL associated with lung adenocarcinoma at both the protein and mRNA level [54], [55].

Cystatin C is a small 13 kDa protein that is a member of the cysteine proteinase inhibitor family. Cystatin-C was elevated in SqCC compared to control during discovery phase analysis of BALF, this was confirmed by ELISA in plasma as it was significantly elevated in AD and SqCC compared to control. Cystatin-C is likely to be biologically relevant as an imbalance between cysteine proteases and their inhibitors arises in malignancy, influencing tumour cell invasion and metastasis [56]. Microarray analysis has revealed that cystatin C was one of the most highly upregulated genes in multiple myeloma [57]. Additionally, Kos et al. demonstrated that a significant correlation between increased serum cystatin C and malignant progression in melanoma and colorectal cancer exists [58].

While the main focus of this investigation was to identify proteins with different abundance levels in BALF and verify these results in plasma samples, a significant number of proteins were found to be changed in BALF samples when comparing AD to SqCC. A finding of particular interest was that folate receptor alpha was increased 14-fold in AD compared to SqCC. Previous work has shown that folate receptor alpha expression had a high discriminatory capacity for lung adenocarcinomas versus squamous cell carcinomas as determined by

immunohistochemistry [59,60]. The utility of MS-based quantitative proteomics as an additional piece of information to facilitate clinicians in the classifications of tumours is a significant area in development, as more treatments are specifically aimed at squamous and non-squamous cell carcinoma. Also, a better understanding of the NSCLC histological subtypes may provide valuable insight in support of developing targeted therapeutics that are specific for precisely classified lung tumours. Several new targeted therapies have been recently approved for non-squamous NSCLC that inhibit Vascular Endothelial Growth Factor (VEGF), Epidermal Growth Factor Receptor (EGFR), Anaplastic Lymphoma Kinase (ALK), and Reactive Oxygen Species (ROS-1) [61]. As more driver mutations are discovered and therapeutic compounds developed to target these abnormal pathways, precise diagnosis of lung cancer and its subtypes will be crucial [62]. Protein expression and signalling in either BALF or plasma may have clinical relevance in the future through its ability to precisely differentiate subtypes and facilitate lung cancer diagnosis.

The variety of patients treated in the clinic, tumour heterogeneity and treatment-related factors will significantly influence the protein constituents of BALF. This is an important limitation in the identification of accurate biomarker profiles, and necessitates the development of biomarker panels and stratification of patients based on previous medical history in order to facilitate the most accurate results based on the measurement of protein levels. Proteins are dynamic and constantly turned over in cells, characteristics that make them very sensitive to factors such as disease severity and previous/current medication history. A single BALF-based biomarker is unlikely to have sufficient sensitivity or specificity for use as a stand-alone test, therefore a panel of unrelated biomarkers may be more effective. When ambiguous results are encountered from biomarker panel testing it may indicate technical difficulties with measuring multiple biomolecules. It is also likely that some complex diseases are not amenable to monitoring with biomarkers and therefore other approaches such as scanning-based procedures may be more suitable.

The BALF proteome is a very complex mixture with many of the top protein hits by mass spectrometry found to be well-established blood-based proteins, namely albumin, haptoglobin and immunoglobin proteins. However, when compared to un-fractioned plasma/serum digests, BALF does contain more low abundant protein identifications and therefore may facilitate the analysis of more cellular secreted/shed proteins associated with disease. Watson and co-workers presented data showing the differences in the proteolytic fingerprints of serum and BALF in guinea pigs [63]. This is very important in the context of protein degradation and subsequent measurement of candidate biomarkers in serum compared to BALF and the interpretation of data from these two pools of bodily fluids.

In conclusion, the findings of this study indicate that the BALF proteome is significantly altered in NCSLC samples compared to control. In addition, significant differences in the protein profile between AD and SqCC in BALF could also be used to help distinguish lung cancer sub-types in this population. The elevation of specific proteins in patients diagnosed with AD and SqCC may be used to develop a specific biomarker panel associated with each NSCLC sub-type. Certain candidate proteins, particularly Cystatin-C, TIMP-1 and Lipocalin, had altered abundance levels in BALF and correlated well with increased abundance in plasma. Proteins expressed at different levels in BALF or plasma may also provide insight into the molecular mechanisms associated with malignancy, particularly in relation to metabolism, cellular processing, and the immune response.

Overall the findings of this study support the use of BALF as a robust proximal biofluid for the discovery of candidate biomarkers in lung cancer. Our results demonstrate that BALF proteomic profiles are transferrable into blood, which is a promising finding in the search for non-invasive biomarkers to diagnose and monitor lung cancer.

Funding

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbacli.2017.03.001.

References

[1] R. Siegel, D. Naishadham, A. Jemal, Cancer statistics, 2012, CA Cancer J. Clin. 62 (2012) 10–29.
[2] R. Meza, C. Meernik, J. Jeon, M.L. Cote, Lung cancer incidence trends by gender, race and histology in the United States, 1973–2010, PLoS One 10 (2015), e0121323. .
[3] S.A. Kenfield, E.K. Wei, M.J. Stampfer, B.A. Rosner, G.A. Colditz, Comparison of aspects of smoking among the four histological types of lung cancer, Tob. Control. 17 (2008) 198–204.
[4] L.A. Torre, F. Bray, R.L. Siegel, J. Ferlay, J. Lortet-Tieulent, A. Jemal, Global cancer statistics, 2012, CA Cancer J. Clin. 65 (2015) 87–108.
[5] R.S. Herbst, J.V. Heymach, S.M. Lippman, Lung cancer, N. Engl. J. Med. 359 (2008) 1367–1380.
[6] L. Yang, L. Wang, Y. Zhang, Immunotherapy for lung cancer: advances and prospects, Am. J. Clin. Exp. Immunol. 5 (2016) 1–20.
[7] H. Lemjabbar-Alaoui, O.U. Hassan, Y.W. Yang, P. Buchanan, Lung cancer: biology and treatment options, Biochim. Biophys. Acta 1856 (2015) 189–210.
[8] B. Parashar, S. Arora, A.G. Wernicke, Radiation therapy for early stage lung cancer, Semin. Interv. Radiol. 30 (2013) 185–190.
[9] M. Snee, M.Q. Hatton, Longstanding challenges in lung cancer: are we meeting them? Clin. Oncol. (R. Coll. Radiol.) (2016).
[10] S.G. Spiro, M.K. Gould, G.L. Colice, A.C.o.C. Physicians, Initial Evaluation of the Patient with Lung Cancer: Symptoms, Signs, Laboratory Tests, and Paraneoplastic Syndromes: ACCP Evidenced-Based Clinical Practice Guidelines, second ed., Chest, 132, 2007 149S–160S.
[11] M. Usman Ali, J. Miller, L. Peirson, D. Fitzpatrick-Lewis, M. Kenny, D. Sherifali, P. Raina, Screening for lung cancer: a systematic review and meta-analysis, Prev. Med. 89 (2016) 301–314.
[12] 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, N.L.S.T.R. Team, Reduced lung-cancer mortality with low-dose computed tomographic screening, N. Engl. J. Med. 365 (2011) 395–409.
[13] V.A. Moyer, U.S.P.S.T. Force, screening for cognitive impairment in older adults: U.S. Preventive Services Task Force recommendation statement, Ann. Intern. Med. 160 (2014) 791–797.
[14] R.S. Wiener, M.K. Gould, D.A. Arenberg, D.H. Au, K. Fennig, C.R. Lamb, P.J. Mazzone, D.E. Midthun, M. Napoli, D.E. Ost, C.A. Powell, M.P. Rivera, C.G. Slatore, N.T. Tanner, A. Vachani, J.P. Wisnivesky, S.H. Yoon, A.A.C.o.L.-D.C.L.C.S.i.C, Practice, an official American Thoracic Society/American College of Chest Physicians policy statement: implementation of low-dose computed tomography lung cancer screening programs in clinical practice, Am. J. Respir. Crit. Care Med. 192 (2015) 881–891.
[15] R. Wender, E.T. Fontham, E. Barrera, G.A. Colditz, T.R. Church, D.S. Ettinger, R. Etzioni, C.R. Flowers, G.S. Gazelle, D.K. Kelsey, S.J. LaMonte, J.S. Michaelson, K.C. Oeffinger, Y.C. Shih, D.C. Sullivan, W. Travis, L. Walter, A.M. Wolf, O.W. Brawley, R.A. Smith, American Cancer Society lung cancer screening guidelines, CA Cancer J. Clin. 63 (2013) 107–117.
[16] P.F. Pinsky, Cost-effectiveness of CT screening in the National Lung Screening Trial, N. Engl. J. Med. 372 (2015) 387.
[17] N. Horeweg, J. van Rosmalen, M.A. Heuvelmans, C.M. van der Aalst, R. Vliegenthart, E.T. Scholten, K. ten Haaf, K. Nackaerts, J.W. Lammers, C. Weenink, H.J. Groen, P. van Ooijen, P.A. de Jong, G.H. de Bock, W. Mali, H.J. de Koning, M. Oudkerk, Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening, Lancet Oncol. 15 (2014) 1332–1341.
[18] J.K. Field, M. Oudkerk, J.H. Pedersen, S.W. Duffy, Prospects for population screening and diagnosis of lung cancer, Lancet 382 (2013) 732–741.
[19] T. Atwater, P.P. Massion, Biomarkers of risk to develop lung cancer in the new screening era, Ann. Transl. Med. 4 (2016) 158.
[20] G. Sozzi, M. Boeri, Potential biomarkers for lung cancer screening, Transl. Lung Cancer Res. 3 (2014) 139–148.
[21] S. Ray, P.J. Reddy, R. Jain, K. Gollapalli, A. Moiyadi, S. Srivastava, Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead, Proteomics 11 (2011) 2139–2161.
[22] S. Camerini, P. Mauri, The role of protein and peptide separation before mass spectrometry analysis in clinical proteomics, J. Chromatogr. A 1381 (2015) 1–12.

[23] A. Kumar, D. Baycin-Hizal, J. Shiloach, M.A. Bowen, M.J. Betenbaugh, Coupling en-
richment methods with proteomics for understanding and treating disease, Proteo-
mics Clin. Appl. 9 (2015) 33–47.

[24] P. Dowling, R. Wormald, P. Meleady, M. Henry, A. Curran, M. Clynes, Analysis of the
saliva proteome from patients with head and neck squamous cell carcinoma reveals
differences in abundance levels of proteins associated with tumour progression and
metastasis, J. Proteome 71 (2008) 168–175.

[25] J. Beretov, V.C. Wasinger, E.K. Millar, P. Schwartz, P.H. Graham, Y. Li, Proteomic anal-
ysis of urine to identify breast cancer biomarker candidates using a label-free LC-
MS/MS approach, PLoS One 10 (2015), e0141876. .

[26] A. Farina, J.M. Dumonceau, P. Antinori, I. Annessi-Ramseyer, J.L. Frossard, D.F.
Hochstrasser, M. Delhaye, P. Lescuyer, Bile carcinoembryonic cell adhesion molecule
6 (CEAM6) as a biomarker of malignant biliary stenoses, Biochim. Biophys. Acta
1844 (2014) 1018–1025.

[27] F. Shen, Y. Zhang, Y. Yao, W. Hua, H.S. Zhang, J.S. Wu, P. Zhong, L.F. Zhou, Proteomic
analysis of cerebrospinal fluid: toward the identification of biomarkers for gliomas,
Neurosurg. Rev. 37 (2014) 367–380 (discussion 380).

[28] M.D. Pastor, A. Nogal, S. Molina-Pinelo, A. Quintanal-Villalonga, R. Meléndez, I.
Ferrer, B. Romero-Romero, M.J. De Miguel, J.L. López-Campos, J. Corral, R. García-
Carbonero, A. Carnero, L. Paz-Ares, IL-11 and CCL-1: novel protein diagnostic bio-
markers of lung adenocarcinoma in bronchoalveolar lavage fluid (BALF), J. Thorac.
Oncol. 11 (2016) 2183–2192.

[29] T.B. Schaaij-Visser, M. de Wit, S.W. Lam, C.R. Jimenez, The cancer secretome, current
status and opportunities in the lung, breast and colorectal cancer context, Biochim.
Biophys. Acta 1834 (2013) 2242–2258.

[30] W. Ding, W. Yue, Y. Fu, The current research feature and prospect of bronchoalveolar
lavage in diagnosing lung cancer, Zhongguo Fei Ai Za Zhi 13 (2010) 370–374.

[31] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram
quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem.
72 (1976) 248–254.

[32] L. Manzanares-Miralles, Ö. Sarikaya-Bayram, E.B. Smith, S.K. Dolan, Ö. Bayram, G.W.
Jones, S. Doyle, Quantitative proteomics reveals the mechanism and consequence of
gliotoxin-mediated dysregulation of the methionine cycle in Aspergillus niger, J. Pro-
teome 131 (2016) 149–162.

[33] P. Dowling, C. Hayes, K.R. Ting, A. Hameed, J. Meiller, C. Mitsiades, K.C. Anderson, M.
Clynes, C. Clarke, P. Richardson, P. O'Gorman, Identification of proteins found to be
significantly altered when comparing the serum proteome from multiple myeloma
patients with varying degrees of bone disease, BMC Genomics 15 (2014) 904.

[34] S. Murphy, M. Zweyer, M. Henry, P. Meleady, R.R. Mundegar, D. Swandulla, K.
Ohlendieck, Label-free mass spectrometric analysis reveals complex changes in
the brain proteome from the mdx-4cv mouse model of Duchenne muscular dystro-
phy, Clin. Proteomics 12 (2015) 27.

[35] H. Mi, A. Muruganujan, P.D. Thomas, PANTHER in 2013: modeling the evolution of
gene function, and other gene attributes, in the context of phylogenetic trees,
Nucleic Acids Res. 41 (2013) D377–D386.

[36] J.M. Jacobs, J.N. Adkins, W.J. Qian, T. Liu, Y. Shen, D.G. Camp, R.D. Smith, Utilizing
human blood plasma for proteomic biomarker discovery, J. Proteome Res. 4
(2005) 1073–1085.

[37] T.D. Veenstra, T.P. Conrads, B.L. Hood, A.M. Avellino, R.G. Ellenbogen, R.S. Morrison,
Biomarkers: mining the biofluid proteome, Mol. Cell. Proteomics 4 (2005) 409–418.

[38] W.J. Qian, D.T. Kaleta, B.O. Petritis, H. Jiang, T. Liu, X. Zhang, H.M. Mottaz, S.M.
Varnum, D.G. Camp 2nd, L. Huang, X. Fang, W.W. Zhang, R.D. Smith, Enhanced de-
tection of low abundance human plasma proteins using a tandem IgY12-SuperMix
immunoaffinity separation strategy, Mol. Cell. Proteomics 7 (2008) 1963–1973.

[39] R. Schiess, B. Wollscheid, R. Aebersold, Targeted proteomic strategy for clinical bio-
marker discovery, Mol. Oncol. 3 (2009) 33–44.

[40] W.J. Qian, J.M. Jacobs, T. Liu, D.G. Camp, R.D. Smith, Advances and challenges in liq-
uid chromatography-mass spectrometry-based proteomics profiling for clinical ap-
plications, Mol. Cell. Proteomics 5 (2006) 1727–1744.

[41] C.R. Jimenez, H.M. Verheul, Mass Spectrometry-Based Proteomics: From Cancer Bi-
ology to Protein Biomarkers, Drug Targets, and Clinical Applications, Am Soc Clin
Oncol Educ Book, 2014 e504–e510.

[42] A. Plymoth, C.G. Löfdahl, A. Ekberg-Jansson, M. Dahlbäck, H. Lindberg, T.E. Fehniger,
G. Marko-Varga, Human bronchoalveolar lavage: biofluid analysis with special em-
phasis on sample preparation, Proteomics 3 (2003) 962–972.

[43] J. Wu, M. Kobayashi, E.A. Sousa, W. Liu, J. Cai, S.J. Goldman, A.J. Dorner, S.J. Projan,
M.S. Kavuru, Y. Qiu, M.J. Thomassen, Differential proteomic analysis of bronchoalve-
olar lavage fluid in asthmatics following segmental antigen challenge, Mol. Cell. Pro-
teomics 4 (2005) 1251–1264.

[44] E.V. Nguyen, S.A. Gharib, L.M. Schnapp, D.R. Goodlett, Shotgun MS proteomic analy-
sis of bronchoalveolar lavage fluid in normal subjects, Proteomics Clin. Appl. 8
(2014) 737–747.

[45] S.A. Almatroodi, C.F. McDonald, A.L. Collins, I.A. Darby, D.S. Pouniotis, Quantitative
proteomics of bronchoalveolar lavage fluid in lung adenocarcinoma, Cancer Geno-
mics Proteomics 12 (2015) 39–48.

[46] R. Chirco, X.W. Liu, K.K. Jung, H.R. Kim, Novel functions of TIMPs in cell signaling,
Cancer Metastasis Rev. 25 (2006) 99–113.

[47] E. Lambert, E. Dassé, B. Haye, E. Petitfrère, TIMPs as multifacial proteins, Crit. Rev.
Oncol. Hematol. 49 (2004) 187–198.

[48] K. Brew, H. Nagase, The tissue inhibitors of metalloproteinases (TIMPs): an ancient
family with structural and functional diversity, Biochim. Biophys. Acta 1803 (2010)
55–71.

[49] H.J. An, Y.J. Lee, S.A. Hong, J.O. Kim, K.Y. Lee, Y.K. Kim, J.K. Park, J.H. Kang, The prog-
nostic role of tissue and serum MMP-1 and TIMP-1 expression in patients with non-
small cell lung cancer, Pathol. Res. Pract. 212 (2016) 357–364.

[50] M. Pesta, V. Kulda, R. Kucera, M. Pesek, J. Vrzalova, V. Liska, L. Pecen, V. Treska, J.
Safranek, M. Prazakova, O. Vycital, J. Bruha, L. Holubec, O. Topolcan, Prognostic sig-
nificance of TIMP-1 in non-small cell lung cancer, Anticancer Res. 31 (2011)
4031–4038.

[51] D. Bolignano, V. Donato, A. Lacquaniti, M.R. Fazio, C. Bono, G. Coppolino, M. Buemi,
Neutrophil gelatinase-associated lipocalin (NGAL) in human neoplasias: a new pro-
tein enters the scene, Cancer Lett. 288 (2010) 10–16.

[52] C. Li, Y.R. Chan, Lipocalin 2 regulation and its complex role in inflammation and can-
cer, Cytokine 56 (2011) 435–441.

[53] S. Ricci, D. Bruzzese, A.D.I. Carlo, Evaluation of MMP-2, MMP-9, TIMP-1, TIMP-2,
NGAL and MMP-9/NGAL complex in urine and sera from patients with bladder can-
cer, Oncol. Lett. 10 (2015) 2527–2532.

[54] A. Friedl, S.P. Stoesz, P. Buckley, M.N. Gould, Neutrophil gelatinase-associated
lipocalin in normal and neoplastic human tissues. Cell type-specific pattern of ex-
pression, Histochem. J. 31 (1999) 433–441.

[55] S. Candido, R. Maestro, J. Polesel, A. Catania, F. Maira, S.S. Signorelli, J.A. McCubrey,
M. Libra, Roles of neutrophil gelatinase-associated lipocalin (NGAL) in human can-
cer, Oncotarget 5 (2014) 1576–1594.

[56] J.L. Cox, Cystatins and cancer, Front. Biosci. (Landmark Ed) 14 (2009) 463–474.

[57] E. Terpos, E. Katodritou, E. Tsiftsakis, E. Kastritis, D. Christoulas, A. Pouli, E. Michalis,
E. Verrou, K. Anargyrou, K. Tsionos, M.A. Dimopoulos, K. Zervas, G.M.S. Group,
Cystatin-C is an independent prognostic factor for survival in multiple myeloma
and is reduced by bortezomib administration, Haematologica 94 (2009) 372–379.

[58] J. Kos, B. Stabuc, N. Cimerman, N. Brünner, Serum cystatin C, a new marker of glo-
merular filtration rate, is increased during malignant progression, Clin. Chem. 44
(1998) 2556–2557.

[59] D.J. O'Shannessy, G. Yu, R. Smale, Y.S. Fu, S. Singhal, R.P. Thiel, E.B. Somers, A.
Vachani, Folate receptor alpha expression in lung cancer: diagnostic and prognostic
significance, Oncotarget 3 (2012) 414–425.

[60] M.I. Nunez, C. Behrens, D.M. Woods, H. Lin, M. Suraokar, H. Kadara, W. Hofstetter, N.
Kalhor, J.J. Lee, W. Franklin, D.J. Stewart, I.I. Wistuba, High expression of folate recep-
tor alpha in lung cancer correlates with adenocarcinoma histology and EGFR
[corrected] mutation, J. Thorac. Oncol. 7 (2012) 833–840.

[61] P. Bansal, D. Osman, G.N. Gan, G.R. Simon, Y. Boumber, Recent advances in targetable
therapeutics in metastatic non-squamous NSCLC, Front. Oncol. 6 (2016) 112.

[62] R. Caparica, G. de Castro, I. Gil-Bazo, C. Caglevic, R. Calogero, M. Giallombardo, E.S.
Santos, L.E. Raez, C. Rolfo, BRAF mutations in non-small cell lung cancer: has finally
Janus opened the door? Crit. Rev. Oncol. Hematol. 101 (2016) 32–39.

[63] D.S. Watson, K. Jambunathan, D.S. Askew, K. Kodukula, A.K. Galande, Robust sub-
strate profiling method reveals striking differences in specificities of serum and
lung fluid proteases, BioTechniques 51 (2011) 95–104.