ROYAL SOCIETY
OF CHEMISTRY

PAPER

Check for updates

Alterations in serum protein glycopatterns related to small cell lung cancer, adenocarcinoma and squamous carcinoma of the lung†

Liyuan Jia,‡[a] Tianran Ma,‡[a] Yiqian Liang,[b] Haoqi Du,[a] Jian Shu,[a] Xiawei Liu,[a] Zhiwei Zhang,[a] Hanjie Yu, [iD][a] Mingwei Chen*[b] and Zheng Li*[a]

Background: The main reason why lung cancer has maintained a high rate of morbidity and mortality is that its early diagnosis is difficult. No current lung cancer screening is recommended by any major medical organization due to the lack of sensitive and specific screening technologies. Thus, this study aimed to systematically investigate the correlation between the alterations in serum glycosylation and three main types of lung cancers (SCLC, ADC and SqCC). *Materials and methods*: We investigated the protein glycopatterns in sera from 333 subjects (65 healthy volunteers, 38 benign lung disease patients, 49 small cell lung cancer patients, and 181 NSCLC patients) using a lectin microarray. A serum microarray was produced to evaluate and verify the terminal carbohydrate moieties of the glycoproteins in individual serum samples from 30 cases simultaneously. *Results*: There were 16 lectins (*e.g.*, RCA120, BS-I, and UEA-I), 24 lectins (*e.g.*, HHL, PTL-I, and MAL-II), and 18 lectins (*e.g.*, GSL-I, LEL, and ACA) that exhibited significant differences in serum protein glycopatterns in the patients with SCLC, ADC and SqCC compared with the controls (HV and BPD). There were 6 lectins (*e.g.*, EEL, NPA, and LEL) that exhibited significantly increased NFIs in ADC and SqCC compared with SCLC. Also, there were 5 lectins (*e.g.*, Jacalin, BS-I, and UEA-I) that exhibited significantly decreased NFIs in ADC compared with SCLC and SqCC. *Conclusions*: This study can facilitate the discovery of potential biomarkers for the differential diagnosis of lung cancer based on the precise alteration in serum protein glycopatterns.

1 Introduction

Lung cancer is the most common malignant tumor and the first leading cause of cancer-related deaths among men and the second leading cause of cancer-related deaths among women worldwide.[1–4] The main reason why lung cancer has maintained a high rate of morbidity and mortality is that its early diagnosis is difficult. Most lung cancer patients have metastatic primary lung cancer at the time of diagnosis, and their five-year survival rate is only 3.5%. Therefore, the low survival rate of lung cancer is closely related to the abovementioned reasons.[3,5] In contrast, early-stage lung cancer has relatively favorable survival.[6] However, currently, there is no lung cancer screening recommended by any major medical organization due to the lack of sensitive and specific screening technologies.[7] Lung cancers are broadly classified into two types, namely, small cell lung cancers

(SCLC) and non-small cell lung cancers (NSCLC), which is based upon the microscopic appearance of the tumor cells. SCLC comprises about 10–15% of lung cancers, which metastasize rapidly to many sites within the body and are most often discovered after they have spread extensively. The clinical stage of SCLC can be divided into limited diseases and extensive diseases. However, NSCLC is the most common lung cancer, accounting for about 85% of all cases, which has several main types: squamous cell lung cancer (SqCC), lung adenocarcinoma (ADC), large cell carcinoma (LCC) and lung adenosquamous carcinoma (ASC). ADC forms the largest group (approximately 60%) among all NSCLC and is thereby the most frequent cause of tumor-related deaths.[8–10]

Tissues and blood were used extensively in early attempts to detect lung cancer. However, it is difficult to acquire enough tissue samples for molecular marker detection because most lung cancer patients are often diagnosed late, and they are not suitable for surgery. In contrast, serum samples are easy to obtain *via* minimally invasive methods, and they are suitable and convenient for molecular marker detection. The serum isocitrate dehydrogenase 2 (IDH2) can be a new and reliable biomarker for the diagnosis and prognosis of NSCLC.[11] Currently, some tumor markers are used to detect lung cancer, such as carcinoembryonic antigen (CEA), serum cytokeratin 19 fragment (Cyfra 21-1) and progesterone-releasing

[a]*Laboratory for Functional Glycomics, College of Life Sciences, Northwest University, Xi'an 710069, P. R. China. E-mail: zhengli@nwu.edu.cn*

[b]*Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of School of Medicine of Xi'an, Jiaotong University, Xi'an 710061, China*

‡ These authors contributed equally to this work.

peptide (pro-GRP). However, the relatively low sensitivity and specificity of these markers in the detection of cancer cells make their application unsatisfactory in the early diagnosis and monitoring of diseases.[12,13] Furthermore, some target prodrugs have been designed and synthesized for the precision therapy of NSCLC. The prodrug polyamine analog gefitinib (PPG) offers reliable and rapid data for the diagnosis of NSCLC within 4 h.[14,15]

There is significant correlation between aberrant glycosylation modification and many diseases. Many types of genetic diseases are result of congenital disorder of glycosylation. Alterations in glycan structures are closely related to the mechanism of numerous diseases, suggesting that a change in these glycans may play a critical role in tumor formation and development.[16] For example, most of the proteins in the serum are secreted by the liver, thus many studies have identified N-glycan alterations in the development of liver disease, which is the basis of many investigations. Alterations in core fucosylation, outer-arm fucosylation, sialylation and glycan branching have been found in the serum of patients with liver disease.[17] Similarly, there are also many groups that give greater insight into the underlying biological mechanisms of aberrant glycosylation in lung cancer.[18,19] For example, the role of α-1,3-fucosyltransferase (Fut3) in the synthesis of sLex and prognosis of its physiological function have been assessed in lung cancer. The accumulation of tumor-associated antigens, such as sLex, was mostly studied in the sera of patients with lung cancer. The level of core fucosylation increased in ADC compared with nonmalignant tissue. However, although the pathogenic mechanism is not completely clear, glycosylation modification is closely related to many diseases. Changes in the glycan structures of serum proteins are an indication when a disease is present, which can provide important information in the pathogenesis and progression for various lung diseases.[20] The differential glycosylation patterns of lung tumor tissue and nonmalignant tissue at the level of individual glycan structures were examined using nLC-chip-TOF-MS.[21] Also, a serum mass profile-based signature identifying patients with early lung cancer was developed,[22,23] which profiled the low-molecular-weight serum fraction by MALDI-MS and revealed several components with abundances discriminating patients with early lung cancer from healthy high-risk smokers.[24] Additionally, 38 glycopeptides from 22 different proteins were significantly differentially abundant in the serum of lung cancer using LC-MS/MS, and the differential expression levels of serum proteins harboring selective glycopeptide candidates were confirmed using sandwich-based ELISAs.[25] Lectins are also carbohydrate–binding proteins that can distinguish glycans based on structural differences.[26] Researchers were able to distinguish between patients with ADC, SQLC and SqCC and patients with benign conditions using lectin arrays to detect the A1AT-specific glycosylation changes in individual serum glycoproteins.[27] The lectin microarray can quickly detect many different types of glycosylation forms simultaneously. Thus, the emergence of this high-throughput sugar group technology has become the main method to study the glycosylation of various protein samples.[28]

SCLC, ADC and SqCC are the most common lung cancers in clinic. LCC and ASC are unusual types, which account for less than 7% of lung cancers. In this study, we aimed to systematically investigate the correlation between the alterations in serum glycosylation and the three main types of lung cancers (SCLC, ADC and SqCC) using lectin microarrays, provide new basic insight into serum protein glycopatterns, and compare different or similar alterations in glycoprotein glycopatterns between SCLC, ADC and SqCC.

2 Materials and methods

2.1 Study approval

The collection and use of all human serum for research presented here were approved by the Ethical Committee of Northwest University and the First Hospital of Medicine College of Xi'an Jiaotong University (Xi'an, China). Written informed consent was received from participants for the collection of their whole serum. This study was conducted in accordance with the Ethical Guidelines of the Declaration of Helsinki.

2.2 Serum collection

Serum samples were collected from 65 healthy volunteers (median age 69.7 years [range 60.5–84.5], 60% male), and 268 patients with lung diseases (66.7% male). These patients had two main types, SCLC and NSCLC. SCLC patients ($n = 87$) including (1) benign pulmonary disease (BPD, $n = 38$, median age 60.1 years [range 42.0–83.0]), (2) limited disease small cell lung cancer (LD-SCLC, $n = 23$, median age 69 years [range 65.5–76.0]), and (3) extensive disease small cell lung cancer (ED-SCLC, $n = 26$, median age 59.5 years [range 46.5–73.5]). The NSCLC patients ($n = 181$) included: (1) lung adenocarcinoma early stage (ADC-ES, $n = 17$, median age 59.0 years [range 41.5–71.0]), (2) lung adenocarcinoma advanced stage (ADC-AS, $n = 73$, median age 56.0 years [range 45.5–63.5]), (3) squamous cell lung cancer early stage (SqCC-ES, $n = 21$, median age 45.0 years [range 37.0–58.5]), and (4) squamous cell lung cancer advanced stage (SqCC-AS, $n = 70$, median age 62.5 years [range 52.0–72.5]). The baseline characteristics of HVs and patients with lung diseases are summarized in Table 1.

Serum was separated immediately *via* centrifugation at $13\,000 \times g$ for 10 min at 4 °C. An EDTA-free inhibitor cocktail (Halt protease inhibitor; Thermo Scientific Pierce Protein Research Products, Rockford, IL, USA) was added immediately at a concentration of 10 μL mL^{-1} serum. The produced serum was aliquoted into small portions, immediately frozen on dry ice and stored at −80 °C. The protein concentration was determined using the BCA assay.

2.3 Lectin microarrays

To normalize the differences between subjects and to tolerate individual variation, the serum samples from HV and patients with SCLC, ADC and SqCC were pooled. Total proteins in the sera were labeled using Cy3 fluorescent dye and purified using a Sephadex-G25 column. Lectin microarrays were utilized as previously described.[29,30] Briefly, 37 lectins with different binding preferences covering N- and O-linked glycans were spotted on homemade epoxysilane-coated slides. Each lectin

Table 1 Baseline characteristics of healthy volunteers and patients of benign lung diseases and lung cancer-induced small cell lung cancer, adenocarcinoma and squamous carcinoma

Subject group	Healthy volunteers	BPD	SCLC		ADC					SqCC				
			LS	ES	I	II	IIIa	IIIb	IV	I	II	IIIa	IIIb	IV
n	65	38	23	26	7	10	20	25	28	10	11	15	25	30
Males, (n)%	35 (67.7)	11 (55.8)	16 (69.6)	16 (69.7)	2 (28.6)	5 (50)	8 (40)	7 (28)	9 (32.1)	5 (50)	10 (91.0)	13 (86.7)	17 (11.3)	24 (80)
Age, years	69.7 (60.5–84.5)	60.1 (42.0–83.0)	69 (65.5–76.0)	59.5 (46.5–73.5)	59.6 (44–77)	59.2 (40–77)	62.3 (42–74)	53.4 (41–68)	59.8 (41–74)	56 (39–72)	65.1 (44–82)	52 (39–72)	65.1 (44–82)	62.85 (40–65)
Pack/year, mean (±SD)			35.91 (±26.05)	38.24 (±30.45)	2.75 (±2.76)	16.43 (±20.85)	28.93 (±27.29)	36.48 (±25.94)	29.18 (±28.17)	35.92 (±23.24)	40.06 (±35.58)	31.45 (±22.21)	36.48 (±25.94)	29.18 (±28.17)
Current smoker, N (%)			2 (8.70%)	1 (3.85%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (9.10%)	2 (13.3%)	8 (32%)	6 (20%)
Total serum protein (g L⁻¹)	63.22 ± 3.87	61.58 ± 1.16	58.42 ± 3.53	63.98 ± 2.15	64.01 ± 1.25	70.18 ± 1.19	58.82 ± 2.64	66.9 ± 3.04	57.19 ± 3.01	62.85 ± 2.53	68.17 ± 1.85	58.24 ± 2.34	63.92 ± 1.53	66.24 ± 2.43

was spotted in triplicate per block, with quadruplicate blocks on one slide. After immobilization, the slides were blocked with blocking buffer containing 2% BSA in 1 × PBS (0.01 M phosphate buffer containing 0.15 M NaCl, pH 7.4) for 1 h and rinsed twice with 1 × PBS. Then the blocked slide was incubated with Cy3-labeled salivary proteins diluted in 0.6 mL of incubation buffer for 3 h at room temperature. After incubation, the microarray was rinsed twice with 1 × PBST (0.2% Tween 20 in 1 × PBS) for 5 min each, and finally rinsed in 1 × PBS before drying. The microarrays were scanned using a Genepix 4000B confocal scanner (Axon Instruments, Foster City, Calif., USA) set at 70% photomultiplier tube and 100% laser power. The acquired images were analyzed at 532 nm for Cy3 detection using the Genepix 6.0 software.

2.4 Data acquisition and analysis

The acquired images were analyzed at 532 nm for Cy3 detection using the Genepix 3.0 software. The average background was subtracted, and values less than the average background ± 1.5 standard deviations (SD) were removed from each data point. The median of the effective data points for each lectin was globally normalized to the sum of medians of all the effective data points for each lectin in a block. Each pooled sample was observed consistently by three repeated slides and the normalized medians of each lectin from 9 repeated blocks were averaged and its SD was counted. The normalized data of the parallel groups were compared with each other based on fold change, according to the following criteria: fold change ≥ 1.5 or ≤0.67 in pairs indicated up-regulation or down-regulation, respectively. Differences between two arbitrary data sets or multiple data sets were tested by Student's t test or one-way ANOVA for each lectin signal using SPSS statistics. The original data was further analyzed using Expander 6.0 (http://acgt.cs.tau.ac.il/expander/) to perform hierarchical clustering analysis.[31–33]

2.5 Serum microarrays

A sera microarray was produced by 120 individual serum samples from HVs ($n = 30$) and patients with BPD ($n = 30$), SCLC ($n = 30$), ADC ($n = 30$) and SqCC ($n = 30$) according to our previous protocols.[30,34] These serum samples were randomly chosen from the 333 samples we collected before. The baseline characteristics of HVs and patients with lung diseases are summarized in Table 1. Briefly, the serum samples were dissolved in spotting buffer containing 0.5 mg mL⁻¹ BSA in 1 × PBS, pH 7.4, to a concentration of 1 mg mL⁻¹ before spotting on homemade epoxysilane-coated slides with Stealth micro spotting pins (SMP-10B). After immobilization, the slides were blocked with the blocking and incubation buffers (ESI Appendix†) for 1 h and rinsed twice with 1 × PBS. Cy3-labeled lectin diluted in 0.5 mL of buffer (ESI Table 7†) was incubated on the blocked slide for 3 h at room temperature in the dark. The slide was washed and dried via centrifugation at 600 rpm for 5 min. The slide was scanned using a Genepix 4000B confocal scanner at 70% photomultiplier tube and 100% laser power settings, and the acquired images were analyzed at 532 nm for Cy3 detection. All values less than average

background $\pm$ 2SD were removed from each data point, and the median of the effective data points of each sample was counted. The medians from one group were averaged and expressed as the mean $\pm$ SD. Differences between two arbitrary groups or multiple groups of medians were tested using Student's *t*-test or one-way ANOVA for each serum sample using SPSS version 19.

3 Results

3.1 Alterations of serum glycopatterns from patients with SCLC

The subjects were chosen at the same or similar age stage and the proportionality subjects in each group was similar to avoid differences between subjects and tolerate individual variations.

The layout of the lectin microarrays and glycopatterns of the Cy3-labeled pooled serum proteins from HV, BPD, LD-SCLC, and ED-SCLC bound to the lectin microarrays are shown in Fig. 1A and B. The normalized fluorescent intensities (NFIs) for each lectin are summarized as the mean value $\pm$ SD in ESI Table 1.† Specifically, (1) results showing a significant increase in NFIs (fold change $\geq$ 1.50, $p < 0.05$), (2) results showing a significant decrease in NFIs (fold change $\leq$ 0.67, $p < 0.05$), and (3) results showing an almost even level in NFIs (fold change range from 0.67 to 1.50, no significant difference). All the results based on fold change in pairs (with *p*-values lower than 0.05) with the NFIs of each lectin from HV, BPD, LD-SCLC, and ED-SCLC are shown in ESI Table 2.† The results showed that there were 16 lectins giving a significantly differential signal in pooled serum
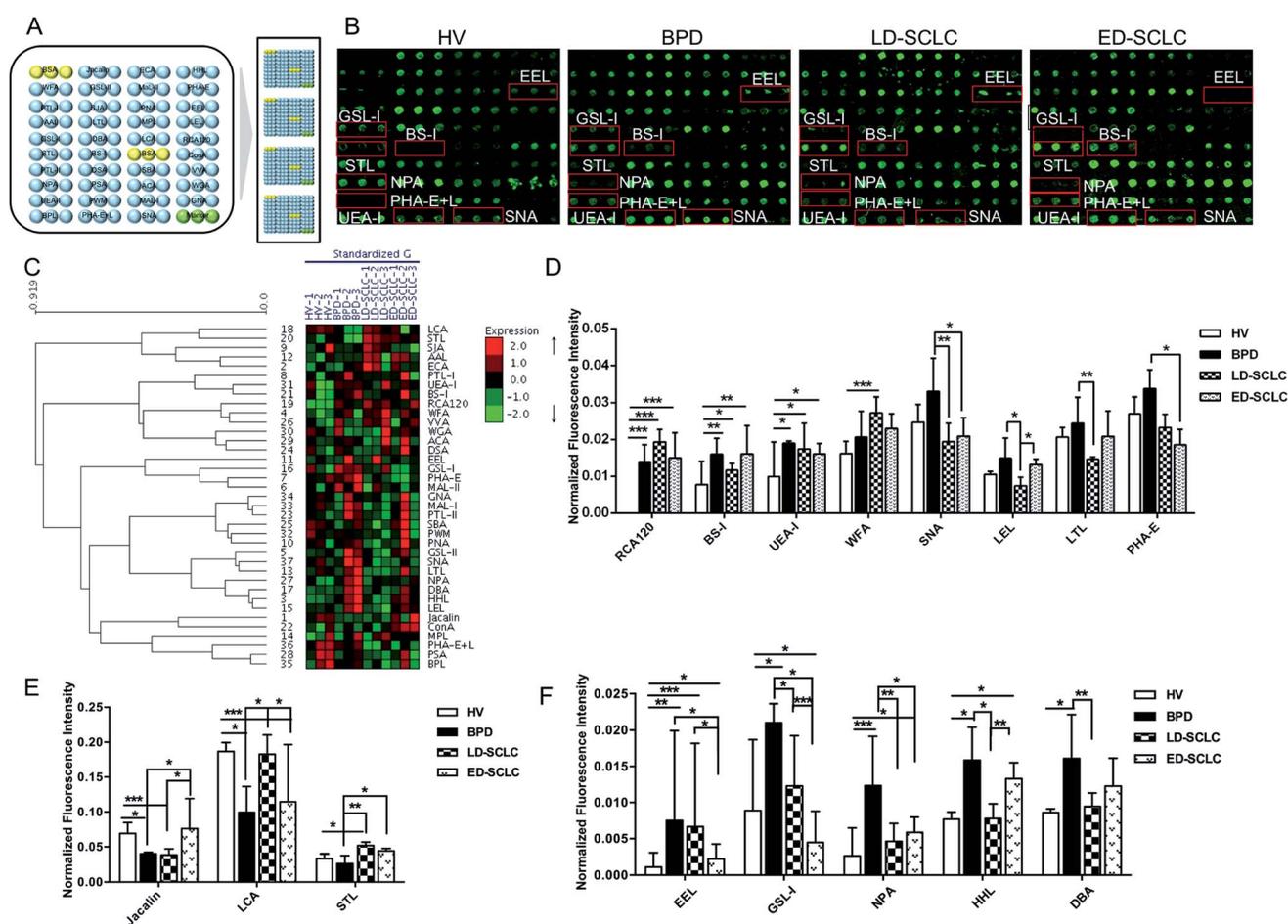

Fig. 1 Glycopatterns of serum glycoproteins from male and female HVs and lung disease patients with BPD, LD-SCLC and ED-SCLC using a lectin microarray. (A) Layout of the lectin microarray. Each lectin was spotted in triplicate per block, with triplicate blocks on one slide. Cy3-labeled BSA was spotted as a location marker and BSA as a negative control. (B) Profiles of Cy3-labeled serum proteins from HV and lung disease patients with BPD, LD-SCLC and ED-SCLC bound to the lectin microarrays. A portion of the slide with three replicates of the lectin arrays is shown. The lectin microarrays revealed a significant difference, as marked with red frames. (C) Heat map and hierarchical clustering analysis of the 37 lectins with three biological replicates. The samples are listed in the columns, and the lectins are listed in the rows. The color and intensity of each square indicate the expression levels relative to the other data in the row. Red, high; green, low; and black, medium. (D) Three lectins revealed significant differences between HV and all groups of patients with SCLC, and five lectins revealed significant differences between BPD and SCLC patients, lectins showing decreasing trends of NFIs ($p < 0.05$). (E) Three lectins showing a decreasing trend of NFIs (fold change $\leq$ 0.67, $p < 0.05$) in parts of groups with SCLC compared with HV and BPD according to one-way ANOVA (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$). (F) Five lectins showing increasing trend of NFIs (fold change $\geq$ 1.5, $p < 0.05$) in some of the groups with SCLC compared with HV and BPD according to one-way ANOVA (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

from HV, BPD, LD-SCLC, and ED-SCLC. The generated data from three biological replicates were imported into EXPANDER 6.0 to perform a hierarchical clustering analysis (Fig. 1C).

The results showed that the β-Gal and Galβ-1 binder RCA120, the α-Gal and α-GalNAc binder BS-I, and the Fucα1-2Galβ1-4Glc(NAc) binder UEA-I exhibited significantly increased NFIs in BPD, LD-SCLC, and ED-SCLC compared with HV (all fold change ≥ 1.51, p ≤ 0.037); however, there was no significant difference between BPD, LD-SCLC, and ED-SCLC. The terminal GalNAcα/β1-3/6Gal binder WFA had significantly increased NFIs in LD-SCLC compared with HV (fold change = 1.69, p < 0.001). Also, the Siaα2-6Gal/GalNAc binder SNA, (GlcNAc)$_n$ and high mannose-type N-glycans binder LEL and the Fucα1-2Galβ1-4GlcNAc binder LTL were associated with decreased NFIs in LD-SCLC compared with BPD (all fold change ≤ 0.60, p ≤ 0.021); however, SNA showed significantly decreased NFIs in ED-SCLC compared with BPD (fold change = 0.55, p = 0.016), and an increase in the NFIs of LEL was observed in ED-SCLC compared with LD-SCLC (fold change = 1.76, p = 0.014). The bisecting GlcNAc binder PHA-E was associated with decreased NFIs in ED-SCLC compared with BPD (fold change = 0.63, p = 0.049) (Fig. 1D).

The Galβ1-3GalNAcα-Ser/Thr (T) binder Jacalin had significantly decreased NFIs in BPD and LD-SCLC compared with HV (all fold change ≤ 0.58, p ≤ 0.043), and significantly increased NFIs in ED-SCLC compared with BPD and LD-SCLC (all fold change ≥ 1.91, p ≤ 0.042). The α-D-Man and Fucα-1,6GlcNAc binder LCA had a decrease in NFIs in BPD and ED-SCLC compared with HV (all fold change ≤ 0.63, p ≤ 0.028), but an increase in NFIs was observed in LD-SCLC compared with BPD and ED-SCLC (all fold change ≥ 1.84, p ≤ 0.032). The trimers and tetramers of the GlcNAc binder STL had significantly increased NFIs in LD-SCLC and ED-SCLC compared with BPD, and had significantly increased NFIs in LD-SCLC compared with HV (all fold change ≥ 1.55, p ≤ 0.046) (Fig. 1E).

The Galα1-3(Fucα1-2) Gal binder EEL showed an increase in NFIs in BPD, LD-SCLC and ED-SCLC compared with HV (all fold change ≥ 1.99, p ≤ 0.23), but a decrease in NFIs was observed in ED-SCLC compared with BPD and LD-SCLC (all fold change ≤ 0.34, p ≤ 0.047). Also, the αGalNAc/αGal binder GSL-I was associated with decreased NFIs in LD-SCLC and ED-SCLC compared with BPD (all fold change ≤ 0.58, p ≤ 0.039) and a decrease in NFIs was observed in ED-SCLC compared with HV and LD-SCLC (all fold change ≤ 0.58, p ≤ 0.039). The high-mannose binder NPA, high-mannose binder HHL, and αGal-NAc and GalNAcα1-3(Fucα1-2) Gal binder DBA showed a significant increase in NFIs in BPD compared with HV and LD-SCLC (all fold change ≥ 1.87, p ≤ 0.010) (Fig. 1F).

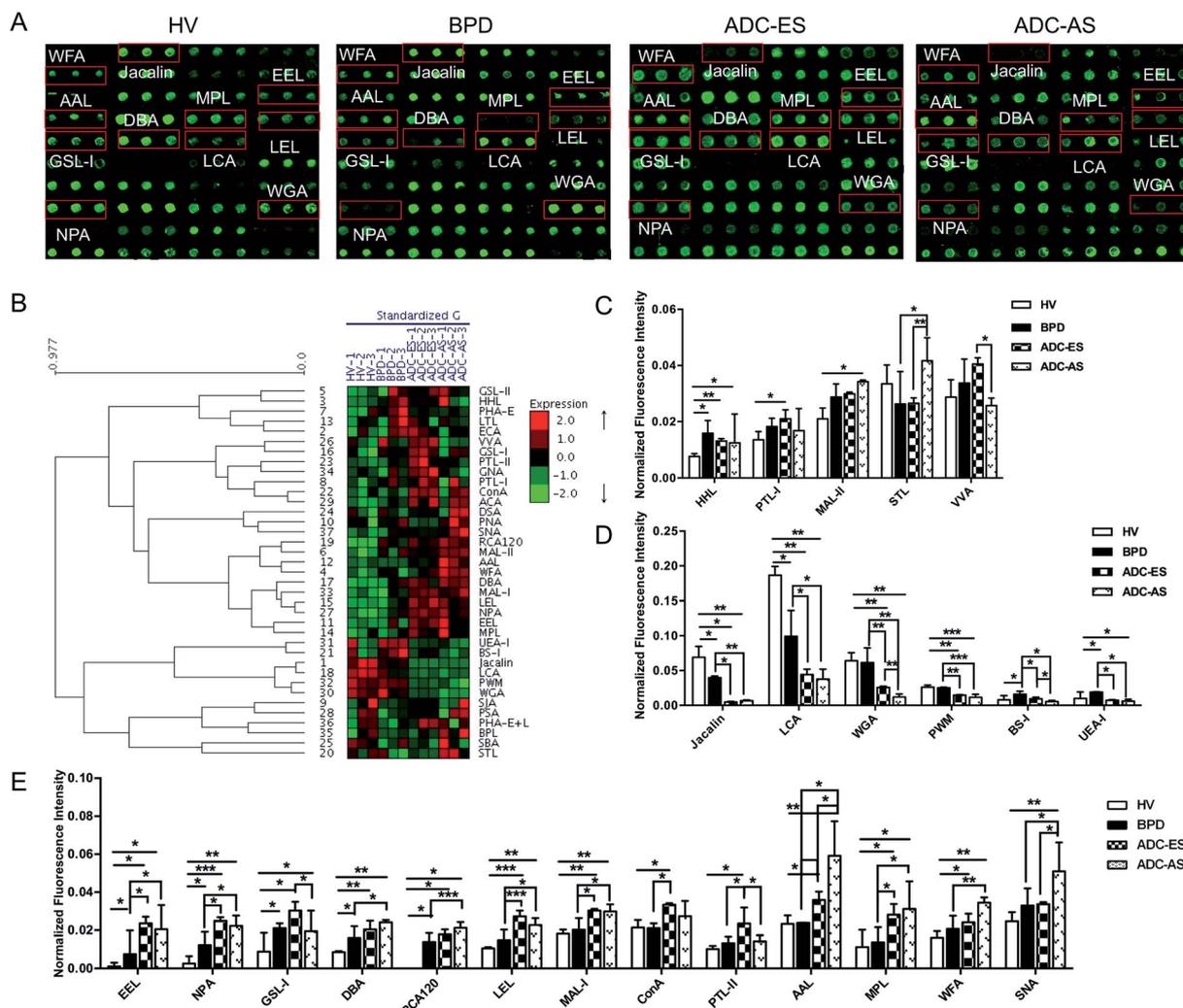## 3.2 Alterations of serum glycopatterns from patients with ADC

The glycopatterns of the Cy3-labeled pooled serum proteins from HV, BPD, ADC-ES, and ADC-AS bound to the lectin microarrays are shown in Fig. 2A. The NFIs for each lectin are summarized as the mean values ± SD in ESI Table 3.† All the results based on fold change in pairs (with p-values lower than

0.05) with the NFIs of each lectin from HV, BPD, ADC-ES, and ADC-AS are shown in ESI Table 4.† The results showed that there were 24 lectins showing a significantly differential signal in the pooled serum from HV, BPD, ADC-ES, and ADC-AS. The generated data from three biological replicates was imported into EXPANDER 6.0 to perform hierarchical clustering analysis (Fig. 2B). The results showed that HHL exhibited significantly increased NFIs in all the patients compared with HV (all fold change ≥ 1.64, p ≤ 0.041). The GalNAc binder PTL-I had significantly increased NFIs in ADC-ES compared with HV (fold change = 1.55, p = 0.016); however, the Siaα2-3Galβ1-4Glc(NAc)/Glc binder MAL-II exhibited increased NFIs in ADC-AS compared with HV (fold change = 0.64, p < 0.001). STL was associated with increased NFIs in ADC-AS compared with BPD and ADC-ES (fold change = 1.60 and 1.57, p = 0.017 and 0.009). The terminal GalNAc binder VVA exhibited significantly decreased NFIs in ADC-AS compared with ADC-ES (fold change = 0.64, p = 0.011) (Fig. 2C).

Both Jacalin and LCA exhibited significantly decreased NFIs in all the patients compared with HV (all fold change ≤ 0.58, p ≤ 0.043) and in ADC-ES and ADC-AS compared with BPD (all fold change ≤ 0.45, p ≤ 0.032). The multivalent Sia and (GlcNAc)$_n$ binder WGA and the (LacNAc)$_n$ binder PWM exhibited significantly decreased NFIs in ADC-ES and ADC-AS compared with HV and BPD (all fold change ≤ 0.59, p ≤ 0.008), and WGA showed a significant decreasing trend from ADC-ES to ADC-AS (fold change = 0.46, p = 0.020). BS-I and UEA-I exhibited significantly increased NFIs in BPD compared with HV (fold change = 2.07 and 1.91, p = 0.043 and 0.045), and decreased NFIs in ADC-ES and ADC-AS compared with BPD (all fold change ≤ 0.59, p ≤ 0.030), and BS-I also showed a significant decreasing trend from BPD and ADC-ES to ADC-AS (all fold change ≤ 0.56, p ≤ 0.038), and UEA-I exhibited a significant decrease in ADC-AS compared with HV (fold change = 0.60, p = 0.012) (Fig. 2D).

EEL, NPA, GSL-I, DBA, and RCA120 were associated with significantly increased NFIs in all the patients compared with HV (all fold change ≥ 1.87, p ≤ 0.049), and the NFIs of EEL and NPA in ADC-ES and ADC-AS were significantly increased compared with BPD (all fold change ≥ 1.82, p ≤ 0.040). The NFIs of DBA and RCA120 showed a significant increasing trend from BPD and ADC-ES to ADC-AS (all fold change ≥ 1.52, p ≤ 0.033). GSL-I, LEL, the Galβ-1/4GlcNAc/Siaα2-3Gal binder MAL-I, the high-mannose binder ConA, and the Gal binder PTL-II exhibited significantly increased NFIs in ADC-ES compared with HV and BPD (all fold change ≥ 1.50, p ≤ 0.049), and decreased NFIs in ADC-AS compared with ADC-ES (all fold change ≤ 0.64, p ≤ 0.022). The Fucα1-6 GlcNAc (core fucose) binder AAL, the Galβ1-3GalNAc/GalNAc binder MPL, the terminating in GalNAcα/β1-3/6Gal binder WFA, and SNA showed significantly increased NFIs in ADC-AS compared with HV, BPD and ADC-ES (all fold change ≥ 1.51, p ≤ 0.015), except for the NFIs of MPL, which were not significantly different between ADC-ES and ADC-AS, and the NFIs of WFA and SNA were not significantly different between ADC-ES and BPD (Fig. 2E).
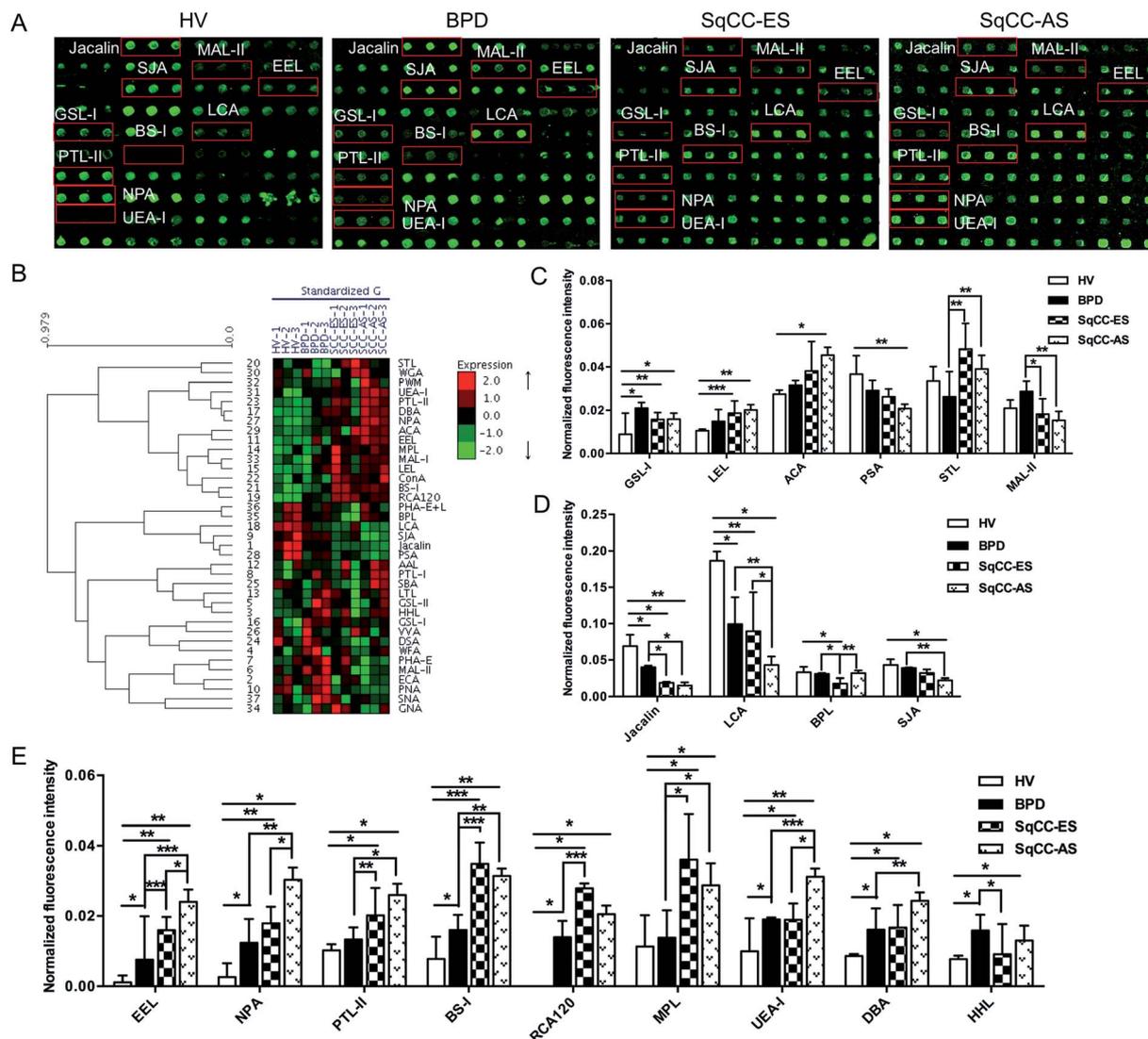
Fig. 2 Glycopatterns of the serum glycoproteins from male and female HVs and lung disease patients with BPD, ADC-ES and ADC-AS using a lectin microarray. (A) Layout of the lectin microarray. Each lectin was spotted in triplicate per block, with triplicate blocks on one slide. Cy3-labeled BSA was spotted as a location marker and BSA as a negative control. The profiles of Cy3-labeled serum proteins from HV and lung disease patients with BPD, ADC-ES and ADC-AS bound to the lectin microarrays. A portion of the slide with three replicates of the lectin arrays is shown. The lectin microarrays revealed a significant difference, as marked with red frames. (B) Heat map and hierarchical clustering analysis of the 37 lectins with three biological replicates. The samples are listed in the columns, and the lectins are listed in the rows. The color and intensity of each square indicate the expression levels relative to the other data in the row. Red, high; green, low; and black, medium. (C) Three lectins revealed significant differences between the HV and ADC patients, and two lectins revealed significant differences between the BPD and ADC patients, with the lectins showing decreasing trends of NFIs ($p < 0.05$). (D) Six lectins showed a decreasing trend of NFIs (fold change ≤ 0.67, $p < 0.05$) in parts of groups with ADC compared with HV and BPD according to one-way ANOVA (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$). (E) Eleven lectins showed increasing trend of NFIs (fold change ≥ 1.5, $p < 0.05$) in some of the groups with ADC compared with HV and BPD according to one-way ANOVA (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

3.3 Alterations of serum glycopatterns from patients with SqCC

The glycopatterns of the Cy3-labeled pooled serum proteins from HV, BPD, SqCC-ES, and SqCC-AS bound to the lectin microarrays are shown in Fig. 3A. The NFIs for each lectin are summarized as the mean value ± SD in ESI Table 5.† All the results based on fold change in pairs (with *p*-values lower than 0.05) with the NFIs of each lectin from the HV, BPD, SqCC-ES, and SqCC-AS groups are shown in ESI Table 6.† The results showed that there were 18 lectins to give a significantly differential signal in the pooled serum from HV, BPD, SqCC-ES, and SqCC-AS. The generated data

from three biological replicates were imported into EXPANDER 6.0 to perform hierarchical clustering analysis (Fig. 3B). The results showed that both GSL-I and LEL significantly increased the NFIs in the patients compared with HV (all fold change $\geq$ 1.80, $p \leq 0.037$), except for the NFIs of LEL had no difference between BPD and HV. The Galβ1-3GalNAcα-Ser/Thr (T) binder ACA had significantly increased NFIs in SqCC-AS compared with HV (fold change = 1.68, $p = 0.004$) and the α-D-Man binder PSA had significantly decreased NFIs in SqCC-AS compared with HV (fold change = 0.58, $p = 0.019$). STL exhibited significantly increased NFIs in SqCC-ES and SqCC-AS compared with BPD

Fig. 3 Glycopatterns of the serum glycoproteins from male and female HVs and lung disease patients with BPD, SqCC-ES and SqCC-AS using a lectin microarray. (A) Layout of the lectin microarray. Each lectin was spotted in triplicate per block, with triplicate blocks on one slide. Cy3-labeled BSA was spotted as a location marker and BSA as a negative control. The profiles of Cy3-labeled serum proteins from HV and lung disease patients with BPD, SqCC-ES and SqCC-AS bound to the lectin microarrays. A portion of the slide with three replicates of the lectin arrays is shown. The lectin microarrays revealed significant differences, as marked with red frames. (B) Heat map and hierarchical clustering analysis of the 37 lectins with three biological replicates. The samples are listed in the columns, and the lectins are listed in the rows. The color and intensity of each square indicate the expression levels relative to the other data in the row. Red, high; green, low; and black, medium. (C) Four lectins revealed significant differences between the HV and SqCC patients, and two lectins revealed significant differences between the BPD and SqCC patients, with the lectins showing decreasing trends of NFIs ($p < 0.05$). (D) Four lectins showed decreasing trend of NFIs (fold change ≤ 0.67, $p < 0.05$) in some of the groups with SqCC compared with HV and BPD according to one-way ANOVA (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$). (E) Nine lectins showed an increasing trend of NFIs (fold change ≥ 1.5, $p < 0.05$) in some of the groups with SqCC compared with HV and BPD according to one-way ANOVA (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

(fold change = 1.85 and 1.51, p = 0.006 and 0.002). MAL-II exhibited significantly decreased NFIs in the SqCC-ES and SqCC-AS groups compared with BPD (fold change = 0.64 and 0.54, p = 0.011 and 0.008) (Fig. 3C).

Both Jacalin and LCA exhibited significantly decreased NFIs in the patients compared with HV (all fold change $\leq$ 0.58, $p \leq$ 0.043), and the NFIs of Jacalin in SqCC-ES and SqCC-AS were significantly lower than that of BPD (all fold change $\leq$ 0.48, $p \leq$ 0.026), and the NFIs of LCA in SqCC-AS were significantly lower

than that of BPD and SqCC-ES (fold change = 0.44 and 0.48, p = 0.008 and 0.013). The Galβ1-3GalNAc binder BPL exhibited significantly decreased NFIs in SqCC-ES compared with HV and BPD (fold change = 0.54 and 0.58, p = 0.038 and 0.015), but an increase in the NFIs was observed in SqCC-AS compared with SqCC-ES (fold change = 1.79, p = 0.008). The terminal in Gal-NAc and Gal binder SJA exhibited significantly decreased NFIs in SqCC-AS compared with HV and BPD (fold change = 0.52 and 0.58, p = 0.012 and 0.003) (Fig. 3D).

The NFIs of EEL, NPA, and PTL-II exhibited a significant increasing trend in HV, BPD, and SqCC-ES to SqCC-AS (all fold change $\geq$ 1.52, $p \leq$ 0.030); however, BS-I, RCA120, and MPL exhibited significantly increased NFIs in SqCC-ES and SqCC-AS compared with HV and BPD (all fold change $\geq$ 2.00, $p \leq$ 0.028). UEA-I and DBA exhibited significantly increased NFIs in SqCC-AS compared with HV, BPD, and SqCC-ES (all fold change $\geq$ 1.87, $p \leq$ 0.047) and their NFIs in BPD and SqCC-ES were significantly higher than that of HV (all fold change $\geq$ 1.54, $p \leq$ 0.006). The high-mannose binder HHL exhibited significantly increased NFIs in BPD and SqCC-AS compared with HV (fold change = 2.06 and 1.71, p = 0.017 and 0.015), but a decrease in the NFIs was observed in SqCC-ES compared with the BPD (fold change = 0.58, p = 0.015) (Fig. 3E).

3.4 Individual validation of the different serum glycopatterns in patients with lung cancer

A serum microarray was fabricated to validate the differential expression levels of serum glycopatterns among HV, BPD, LD-SCLC, and ED-SCLC by spotting 120 individual serum samples (20 male and 10 female samples randomly selected from HV, BPD, LD-SCLC, and ED-SCLC) in spotting buffer to a concentration of 1 mg mL^{-1} on the surface of an epoxy slide according to our previous publication.[26] Each serum sample was spotted in triplicate, and the layout of the serum microarrays is shown in Fig. 4A. The three lectins (BS-I, UEA-I and GSL-I) that revealed significant differences (p < 0.05) in the serum glycopatterns according to the results of the respective lectin microarrays were selected to validate the differential expression levels of the targeted glycan structures in the individual serum (Fig. 4B). BS-I staining showed increased fluorescence intensities (FIs) ($p \leq$ 0.002) in the patients with LD-SCLC and ED-SCLC compared with individual HVs. UEA-I staining showed increased FIs in the individuals with LD-SCLC and ED-SCLC compared with the individuals with HVs ($p \leq$ 0.045). GSL-I staining showed decreased FIs in the individuals with ED-SCLC compared with the individuals with HVs (p < 0.001). These results are generally consistent with the results from the lectin microarrays.

Another serum microarray was fabricated to validate the differential expression levels of serum glycopatterns among HV, BPD, ADC and SqCC by spotting 120 individual serum samples (20 male and 10 female samples randomly selected from HV, BPD, ADC and SqCC). Each serum sample was repeatedly spotted three times, and the layout of the serum microarray is shown in Fig. 5A. Based on the results of the previous lectin microarrays, three different lectins (p < 0.05) (PTL-II, UEA-I and DBA) were selected for serum glycosylation changes in lung cancer, and further verified by the serum microarray to confirm if there were differences in the expression level of the target glycan structure in the individual serum samples. As a result, PTL-II had increased FIs in ADC-ES and ADC-AS compared with the HV group (p < 0.05). Moreover, it had a significant increase in alteration in SqCC-ES and SqCC-AS compared with the HV group. UEA-I had decreased FIs in ADC-AS compared with the HV group (p < 0.05). DBA had increased FIs in SqCC-AS compared with the HV group

(p = 0.001) (Fig. 5B). These results are generally consistent with the results from the lectin microarrays.

3.5 Systematic comparison of the serum protein glycopatterns among SCLC, ADC and SqCC

There were 16 lectins (e.g., RCA120, BS-I, and UEA-I), 24 lectins (e.g., HHL, PTL-I, and MAL-II), and 18 lectins (e.g., GSL-I, LEL, and ACA) that revealed significant differences in the serum protein glycopatterns from the patients with SCLC, ADC and SqCC compared with the controls (HV and HPD), respectively. A systematic comparative analysis of the serum protein glycopatterns among SCLC, ADC and SqCC was performed based on the above data, which showed that there were 6 lectins (EEL, NPA, LEL, MPL, MAL-I, and GSL-I) that exhibited significantly increased NFIs in ADC and SqCC compared with SCLC (all fold change $\geq$ 1.5, $p \leq$ 0.01), and almost an even level in NFIs (fold change range from 0.67 to 1.50, no significant difference) between ADC and SqCC (Fig. 5A). However, there were 5 lectins (Jacalin, BS-I, UEA-I, WGA, and PWM) that exhibited significantly decreased NFIs in ADC compared with SCLC and SqCC (all fold change $\geq$ 1.5, $p \leq$ 0.01), and Jacalin showed significantly increased NFIs in SCLC compared with SqCC-ES (all fold change $\geq$ 1.54, $p \leq$ 0.54), and BS-I exhibited significantly decreased NFIs in SCLC compared with SqCC (all fold change $\geq$ 1.5, $p \leq$ 0.01).

4 Discussion

Glycosylation is one of the most common post-translational modifications (PTMs) of secreted proteins and plays a critical role in signal recognition and conduction, cell–cell interaction, cell adhesion, malignant transformation and metastasis in the development of many cancers.[35] In the diagnosis of diseases such as lung cancer, gastric cancer, breast cancer, and hepatocellular carcinoma, serum glycobiomarker identification has led to big breakthroughs and improvements in their detection, and thus has become a topic of increasing interest because of its high sensitivity and specificity.[36–39] Moreover, the growing practice of glycobiology provides great promise in the discovery of glycosylated biomarkers for various diseases. As high-throughput glycan technology, lectin microarrays can be used for the rapid and accurate detection of multiple different carbohydrate structures without the need to release N- or O-linked glycans before analysis. Consequently, real information can be obtained about protein glycosylation.[28,40]

In the present study, 333 serum samples were collected for screening candidate lectins and constructing diagnostic models in retrospective cohort. SCLC with a high rate of metastasis and recurrence, is more aggressive and malignant than several types of lung cancer. More than 70% of patients are diagnosed with metastatic disease. Therefore, there is an urgent need for effective SCLC biomarkers to improve its diagnosis and achieve early detection. There were 16 lectin (e.g., PHA-E + L, BS-I, and UEA-I) alterations in the serum glycopatterns between HV, BPD, LD-SCLC and ED-SCLC using lectin microarrays, serum microarrays, and statistical analysis. ADC is the most common type of lung cancer, which has obvious cellular and molecular
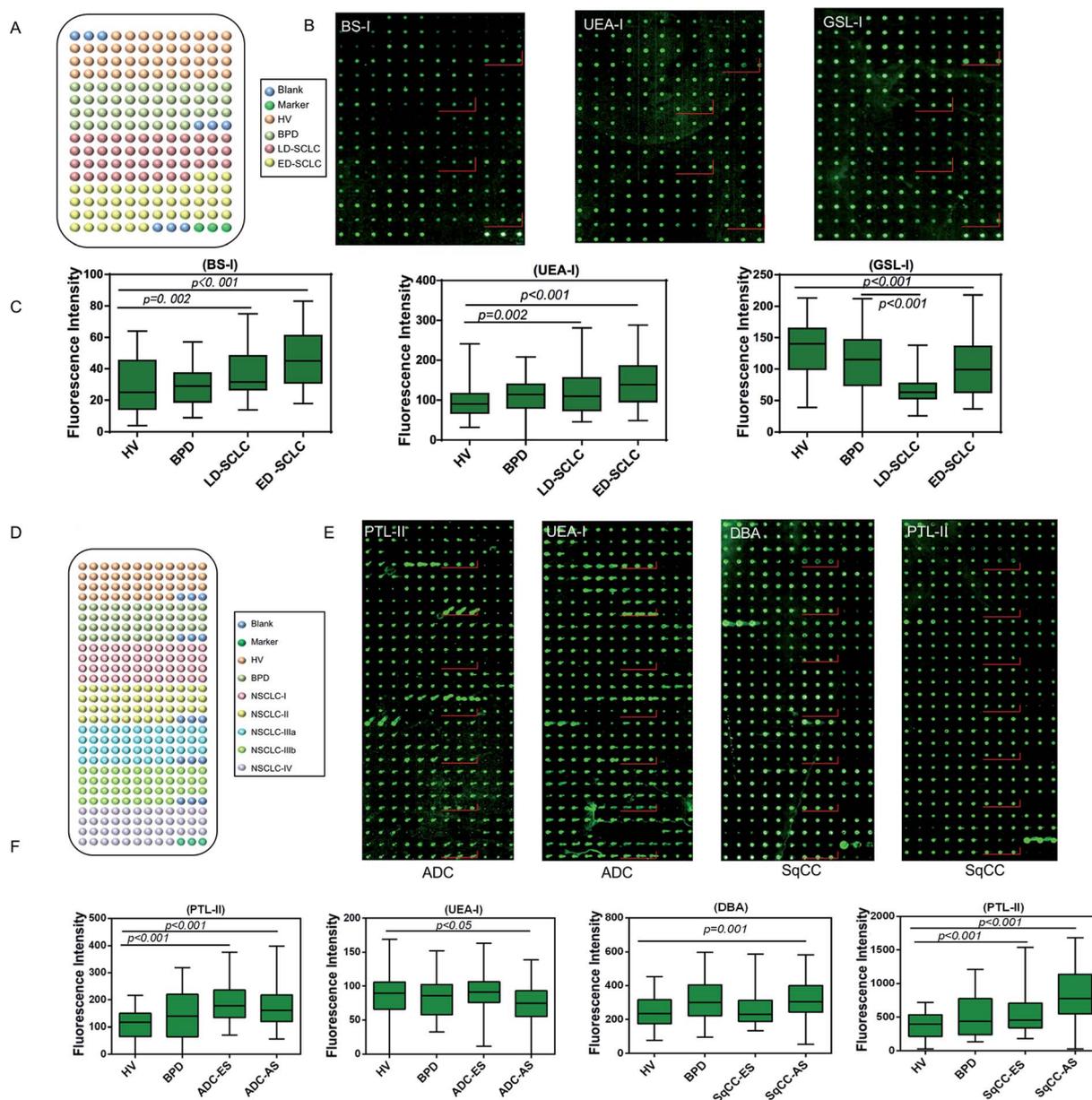
Fig. 4 Validation of the differential expressions of the glycopatterns in the serum associated with SCLC, ADC and SqCC. (A) Layout of the serum microarrays. A total of 120 serum samples ranging from HV ($n = 30$) and from patients with BPD ($n = 30$), LD-SCLC ($n = 30$), and ED-SCLC ($n = 30$) were dissolved in spotting buffer to a concentration of 1 mg mL^{-1} and spotted on epoxysilane-coated slides. Each sample was spotted in triplicate per block, with two continuous blocks on one slide. (B) Scan images of Cy3-labeled lectins binding to a serum microarray. Here, a block of the slide with two sequential blocks is shown. (C) Box plot analysis of the original data obtained from the serum microarrays. Error bars represent 95% confidence intervals for the means. The statistical significance of the differences between groups is indicated by the p-value. (D) Layout of the serum microarrays. A total of 210 serum samples ranging from HV ($n = 30$) and from patients with BPD ($n = 30$), ADC/SqCC-I ($n = 30$), ADC-II/SqCC ($n = 30$), ADC-IIIa/SqCC ($n = 30$), ADC-IIIb ($n = 30$) and ADC/SqCC-IV ($n = 30$) were dissolved in spotting buffer to a concentration of 1 mg mL^{-1} and spotted on epoxysilane-coated slides. Each sample was spotted in triplicate per block, with two continuous blocks on one slide. (E) Scan images of Cy3-labeled lectins binding to a serum microarray. Here, a block of the slide with two sequential blocks is shown. (F) Box plot analysis of the original data obtained from the serum microarrays. Error bars represent 95% confidence intervals for the means. The statistical significance of the differences between groups was indicated by the p-value.

characteristics, and its glands or ducts have a large amount of mucus. Because of its different pathological features and prognosis, ADC is classified as a type of NSCLC, which is totally different from SCLC. There were 24 lectin (e.g., LEL, AAL, MAL-I, and PWM) alterations in the serum glycopatterns between HV,

BPD, ADC-ES and ADC-AS using lectin microarrays, serum microarrays, and statistical analysis. SqCC is one form of NSCLC, which begins in the tissues that line the air passages in the lungs. It is also known as epidermoid carcinoma. Most SqCC of the lungs are located centrally, usually in the larger
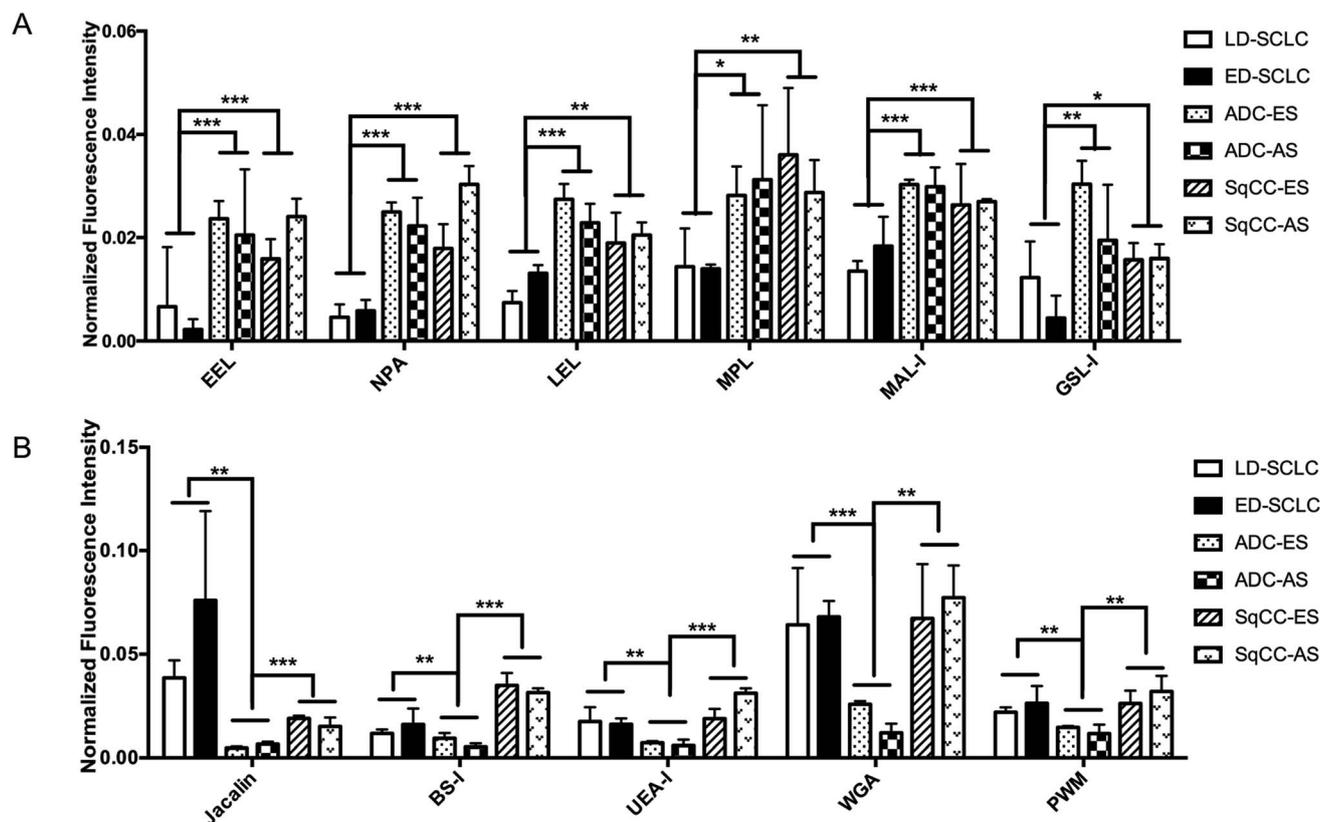
Fig. 5 Systematic comparison of the glycoprotein glycopatterns between SCLC, ADC and SqCC sera with the lectin microarray. (A) Six lectins showed significant differences between the SCLC and NSCLC patients (all fold change ≥ 1.5, $p < 0.05$). (B) Five lectins showed significant differences between ADC and SCLC, SqCC patients (all fold change ≤ 0.67, $p < 0.05$).

bronchi that join the trachea to the lung. There were 18 lectins (e.g., MPL, BS-I, UEA-I) showing significant alterations in NFIs in the pooled serum among HV, BPD, SqCC-ES, and SqCC-AS.

The serum of patients with different cancers has its own specific glycoprotein glycans. When different types of cancer occur, the corresponding glycan type or expression level changes, and these changes can be quickly detected by the lectin microarray. In our previous studies, the results showed that in different cancers, the alterations in glycopatterns are totally different. In breast cancer, the glycans recognized by MAL-I, ECA, NPA, BPL, BS-I, PTL-II, DBA, PNA, PHA-E + L, UEA-I, and PWM showed differential expression.[41] In gastric cancer, the glycans recognized by ECA, PSA, PHA-E + L, HHL, PNA, EEL, MPL, and GSL-I showed differential expression.[31] In hepatic carcinoma, the glycans recognized by MAL-II, SNA, WFA, PNA, and AAL showed differential expression.[42] In lung cancer, the alteration in the glycopattern is different from breast cancer, gastric cancer and hepatic carcinoma. Furthermore, different types lung cancer show different glycopatterns. The changes in the serum glycosylation levels are correlated with tumor pathological characteristics, and this correlation can be amplified by the lectin microarray signals, which provides a good reference for the rapid diagnosis of different diseases.

The fold changes in the serum glycopatterns based on the ratio of the NFIs of each lectin from the HV, BPD, SCLC, ADC, and

SqCC patients are summarized in Table 2. We found that the blood group B antigen and high-mannose recognized by EEL and NPA increased in the patients with BPD, SCLC, ADC, and SqCC compared with HV, and SCLC, ADC, and SqCC compared with BPD. Thus, EEL and NPA have potential to become markers for the diagnosis of lung diseases. The biantennary complex-type N-glycan with outer Gal and Sia2-6Gal/GalNAc recognized by PHA-E + L and SNA decreased in the patients with SCLC compared with HV and BPD, but increased in the patients with SqCC-AS compared with HV and BPD. Therefore, SCLC and SqCC may be distinguished by PHA-E + L and SNA. Siaα2-3Galβ1-4Glc(NAc)/Glc recognized by MAL-II only reflected the change in ADC and core fucose recognized by AAL, and (GlcNAc)$_n$ recognized by WGA, and BraBLDhed (LacNAc)$_n$ recognized by PWM only reflected the change in SqCC. Thus, these lectins can be used to distinguish ADC and SqCC (Table 2). According to the research results, different combinations of lectins can be used to detect the type of lung cancer, and even its pathological stage.

Untargeted abnormal growth of lung tissue often leads to a benign lung tumor, and benign lung tumors may be derived from many different structures of lung. Currently, very little is known about the causes of benign lung tumors and nodules. However, in general, benign lung tumors are usually caused by inflammation caused by infection, tuberculosis, lung abscess and non-infectious causes. Our results indicated that there were 8

Table 2 Fold change of the serum glycopatterns based on the ratio of the NFIs of each lectin from HVs, BPD, SCLC, ADC, and SqCC patients

| | Compared with HV | | | | | | | | Compared with BPD | | | | | |
| | BPD | SCLC | | ADC-ES/HV | | SqCC | | | SCLC | | ADC | | SqCC | |
Lectin	BPD/HV	LD/HV	ED/HV	ES/HV	AS/HV	ES/HV	AS/HV	Lectin	LD/BPD	ED/BPD	ES/BPD	AS/BPD	ES/BPD	AS/BPD
EEL	6.707↑	5.940↑	1.994↑	14.210↑	21.478↑	21.144↑	18.316↑	EEL	—	0.297↑	2.119↑	3.202↑	3.153↑	2.731↑
NPA	4.655↑	1.743↑	2.218↑	6.751↑	11.439↑	9.438↑	8.404↑	NPA	0.374↑	0.476↑	—	2.457↑	2.027↑	1.805↑
MPL	—	—	—	3.193↑	2.541↑	2.495↑	2.764↑	MPL	—	—	2.619↑	2.084↑	2.046↑	2.267↑
LEL	—	—	—	1.776↑	1.915↑	2.593↑	2.158↑	LEL	0.501↑	—	—	—	1.836↑	1.527↑
GSL-I	2.371↑	—	0.504↓	1.775↑	1.794↑	3.421↑	2.194↑	GSL-I	0.585↑	0.213↑	—	—	—	—
DBA	1.870↑	—	—	1.938↑	2.829↑	2.376↑	2.815↑	DBA	0.588↓	—	—	1.513↑	—	1.505↑
LCA	0.532↓	—	0.614↓	0.480↓	0.231↓	0.237↓	0.201↓	LCA	1.840↓	—	—	0.434↑	0.445↑	0.378↑
BS-I	2.069↑	1.511↑	2.081↑	4.522↑	4.074↑	—	—	BS-I	—	—	2.185↑	1.969↑	0.586↓	0.329↓
UEA-I	1.914↑	1.754↑	1.625↑	1.901↑	3.143↑	—	0.597↓	UEA-I	—	—	—	1.642↑	0.380↑	0.312↑
PHA-E + L	—	0.649↓	0.660↓	—	—	—	—	SNA	0.590↓	0.634↓	—	—	—	1.545↑
SNA	—	—	—	—	—	—	2.066↑	HHL	0.491↓	—	0.574↓	—	—	—
HHL	2.059↑	—	1.726↑	—	1.689↑	1.712↑	1.630↑	MAL-II	—	—	0.633↓	0.534↓	—	—
MAL-II	—	—	—	—	—	—	1.633↑	AAL	—	—	—	—	1.501↑	2.471↑
AAL	—	—	—	—	—	1.539↑	2.533↑	MAL-I	0.670↓	—	—	—	—	—
MAL-I	—	—	—	—	—	1.668↑	1.646↑	WGA	—	—	—	—	0.421↓	0.195↓
WGA	—	—	—	—	—	0.403↓	0.187↓	PWM	—	—	—	—	0.585↓	0.463↓
PWM	—	—	—	—	—	0.562↓	0.445↓							

lectins (*e.g.* BS-I, UEA-I, and RCA120) to show a significant increase in NFIs (fold change ≥ 1.87, $p < 0.043$) and 2 lectins (Jacalin and LCA) to show a significant decrease in NFIs (fold change ≤ 0.58, $p < 0.043$) in BPD compared with HV. Moreover, there were 6 lectins (*e.g.* BS-I, EEL, and HHL) to show a significant increase in NFIs (fold change ≥ 2.06, $p < 0.043$) in BPD compared with HV.

One limitation for this study is that our investigation does not refer to the mechanisms that cause the alterations in serum protein glycosylation in patients with LC. Past research revealed that fucosylation of E-cadherin activated the downstream gene to be translated, which promoted the dimerization and phosphorylation of EGFR, and this can ultimately lead to lung cancer cell proliferation and metastasis.[43] Thus, it is necessary to deeply study the mechanism of glycosylation and the development of lung cancer. The second limitation was the lack of enough clinical samples from patients of LCC or ASC. The sample size was statistically non-significant. Therefore, this study aimed to study three common types of lung cancer. Another concern is the lack of verification about the different glycopatterns among SCLC, ADC and SqCC. These findings were calculated based on previous research results and are of great significance for the construction of a lung cancer diagnostic model. Therefore, we have collected more clinical samples for further research, with the aim of identifying the lung cancer types and stages more accurately. Studies in larger cohorts enrolling more patients with the different subtypes also need to be carried out in the future.

In conclusion, we systematically investigated the correlation between the alterations in serum glycosylation related to three main types of lung cancer (SCLC, ADC and SqCC) using lectin microarrays, which helped us gain new insight into the serum protein glycopatterns. Furthermore, the different or similar alterations of glycoprotein glycopatterns among SCLC, ADC and SqCC were systematically compared. There were 16 lectins (*e.g.*, RCA120, BS-I, and UEA-I), 24 lectins (*e.g.*, HHL, PTL-I, and MAL-II), and 18 lectins (*e.g.*, GSL-I, LEL, and ACA) that revealed significant differences in serum protein glycopatterns from the patients with SCLC, ADC and SqCC compared with the controls (HV and HPD), respectively. There were 6 lectins (*e.g.*, EEL, NPA, and LEL) that exhibited significantly increased NFIs in ADC and SqCC compared with SCLC. Also, there were 5 lectins (*e.g.*, Jacalin, BS-I, and UEA-I) that exhibited significantly decreased NFIs in ADC compared with SCLC and SqCC. This study can facilitate the discovery of potential biomarkers for the differential diagnosis of lung cancer based on the precise alterations in serum protein glycopatterns.

Abbreviations

SCLC	Small-cell lung cancer
NSCLC	Non-small-cell carcinoma
ADC	Lung adenocarcinoma
SqCC	Squamous cell lung cancer
HV	Healthy volunteer
BPD	Benign pulmonary disease
LD-SCLC	Limited disease small cell lung cancer
ED-SCLC	Extensive disease small cell lung cancer
ADC-ES	Lung adenocarcinoma early stage
ADC-AS	Lung adenocarcinoma advanced stage
SqCC-ES	Squamous cell lung cancer early stage
SqCC-AS	Squamous cell lung cancer advanced stage

Author contributions

All authors participated in literature research and data classification. LJ, TM, XL and YZ: design of study and manuscript

preparation; YL, HY and MC: serum collection; HD, JS and ZZ: data statistics; ZL: critical revision of the manuscript.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

References

1 R. Siegel, J. Ma, Z. Zou and A. Jemal, *Ca-Cancer J. Clin.*, 2014, **64**, 9–29.

2 R. L. Siegel, K. D. Miller and A. Jemal, *Ca-Cancer J. Clin.*, 2015, **65**, 5–29.

3 L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent and A. Jemal, *Ca-Cancer J. Clin.*, 2015, **65**, 87–108.

4 F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal, *Ca-Cancer J. Clin.*, 2018, 1–31.

5 W. Chen, R. Zheng, P. D. Baade, S. Zhang, H. Zeng, F. Bray, A. Jemal, X. Q. Yu and J. He, *Ca-Cancer J. Clin.*, 2016, 1–18.

6 L. A. Torre, R. L. Siegel and A. Jemal, *Adv. Exp. Med. Biol.*, 2016, **893**, 1–19.

7 B. J. Addario, *Ann. Transl. Med.*, 2015, **3**, 26.

8 R. Siegel, D. Naishadham and A. Jemal, *Ca-Cancer J. Clin.*, 2013, **63**, 11–30.

9 J. H. Cheng, W. S. Liu, Z. M. Li and Z. G. Wang, *Chin. J. Integr. Med.*, 2007, **13**, 269–274.

10 World Health Organization, http://who.int/gho/database/en/, accessed June 21, 2018, DOI: 10.1111/his.12157.

11 J. J. Li, R. Li, W. Wang, B. Zhang, X. Song, C. Zhang, Y. Gao, Q. Liao, Y. He, S. You, Z. Tan, X. Luo, Y. Li, M. Tang, X. Weng, W. Yi, S. Peng, S. Liu, Y. Tan, A. M. Bode and Y. Cao, *Mol. Oncol.*, 2018, **12**, 602–610.

12 T. Shinkai, N. Saijo, K. Tominaga, K. Eguchi, E. Shimizu, Y. Sasaki, J. Fujita, H. Futami, H. Ohkura and K. Suemasu, *Cancer*, 1986, **57**, 1318–1323.

13 J. L. Pujol, J. Grenier, J. P. Daures, A. Daver, H. Pujol and F. B. Michel, *Cancer Res.*, 1993, **53**, 61–66.

14 X. Song, X. Han, F. Yu, X. Zhang, L. Chen and C. Lv, *Theranostics*, 2018, **8**, 2217–2228.

15 X. Zhang, N. He, Y. Huang, F. Yu, B. Li, C. Lv and L. Chen, *Sens. Actuators, B*, 2019, **282**, 69–77.

16 S. R. Stowell, T. Z. Ju and R. D. Cummings, *Annu. Rev. Pathol.: Mech. Dis.*, 2015, **10**, 473–510.

17 A. Mehta, H. Herrera and T. Block, *Adv. Cancer Res.*, 2015, **126**, 257–279.

18 K. Ueda, Y. Fukase, T. Katagiri, N. Ishikawa, S. Irie, T. A. Sato, H. Ito, H. Nakayama, Y. Miyagi, E. Tsuchiya, N. Kohno,

M. Shiwa, Y. Nakamura and Y. Daigo, *Proteomics*, 2009, **9**, 2182–2192.

19 J. A. Vasseur, J. A. Goetz, W. R. Alley and M. V. Novotny, *Glycobiology*, 2012, **22**, 1684–1708.

20 R. Kannagi, Y. Fukushi, T. Tachikawa, A. Noda, S. Shin, K. Shigeta, N. Hiraiwa, Y. Fukuda, T. Inamoto, S. Hakomori, *et al.*, *Cancer Res.*, 1986, **46**, 2619–2626.

21 L. R. Ruhaak, S. L. Taylor, C. Stroble, U. T. Nguyen, E. A. Parker, T. Song, C. B. Lebrilla, W. N. Rom, H. Pass, K. Kim, K. Kelly and S. Miyamoto, *J. Proteome Res.*, 2015, **14**, 4538–4549.

22 D. Zhang, X. Li, X. Liu, Y. Wang, M. Zhang, Q. Wang, T. Chen and Z. Li, *Proteomics: Clin. Appl.*, 2020, **14**, e1900016.

23 T. Chen, C. He, M. Zhang, X. Li, X. Liu, Y. Liu, D. Zhang and Z. Li, *J. Cancer*, 2019, **10**, 5628–5637.

24 P. Widlak, M. Pietrowska, J. Polanska, M. Marczyk, M. Ros-Mazurczyk, R. Dziadziuszko, J. Jassem and W. Rzyman, *Lung Cancer*, 2016, **99**, 46–52.

25 X. M. Zeng, B. L. Hood, M. Sun, T. P. Conrads, R. S. Day, J. L. Weissfeld, J. M. Siegfried and W. L. Bigbee, *J. Proteome Res.*, 2010, **9**, 6440–6449.

26 J. Hirabayashi, M. Yamada, A. Kuno and H. Tateno, *Chem. Soc. Rev.*, 2013, **42**, 4443–4458.

27 Y. Liang, T. Ma, A. Thakur, H. Yu, L. Gao, P. Shi, X. Li, H. Ren, L. Jia, S. Zhang, Z. Li and M. Chen, *Glycobiology*, 2015, **25**, 331–340.

28 G. Gupta, A. Surolia and S. G. Sampathkumar, *OMICS*, 2010, **14**, 419–436.

29 Y. Zhong, Y. Qin, L. Dang, L. Jia, Z. Zhang, H. Wu, J. Cui, H. Bian and Z. Li, *Proteomics*, 2015, **15**, 3283–3295.

30 Y. Qin, Y. Zhong, M. Zhu, L. Dang, H. Yu, Z. Chen, W. Chen, X. Wang, H. Zhang and Z. Li, *J. Proteome Res.*, 2013, **12**, 2742–2754.

31 J. Shu, H. Yu, X. Li, D. Zhang, X. Liu, H. Du, J. Zhang, Z. Yang, H. Xie and Z. Li, *Oncotarget*, 2017, **8**, 35718–35727.

32 Y. Zhong, J. Zhang, H. Yu, J. Zhang, X. X. Sun, W. Chen, H. Bian and Z. Li, *Biochem. Biophys. Res. Commun.*, 2015, **468**, 906–912.

33 H. Zhu, M. Liu, H. Yu, X. Liu, Y. Zhong, J. Shu, X. Fu, G. Cai, X. Chen, W. Geng, X. Yang, M. Wu, Z. Li and D. Zhang, *J. Diabetes Res.*, 2017, **2017**, 5728087.

34 H. A. Badr, D. M. Alsadek, A. A. Darwish, A. I. Elsayed, B. O. Bekmanov, E. M. Khussainova, X. Zhang, W. C. Cho, L. B. Djansugurova and C. Z. Li, *Expert Rev. Proteomics*, 2014, **11**, 227–236.

35 H. Nie, X. Liu, Y. Zhang, T. Li, C. Zhan, W. Huo, A. He, Y. Yao, Y. Jin, Y. Qu, X. L. Sun and Y. Li, *Sci. Rep.*, 2015, **5**, 16007.

36 L. Veillon, C. Fakih, H. Abou-El-Hassan, F. Kobeissy and Y. Mechref, *ACS Chem. Neurosci.*, 2018, **9**, 51–72.

37 L. Oliveira-Ferrer, K. Legler and K. Milde-Langosch, *Semin. Cancer Biol.*, 2017, **44**, 141–152.

38 G. Vecchio, A. Parascandolo, C. Allocca, C. Ugolini, F. Basolo, M. Moracci, A. Strazzulli, B. Cobucci-Ponzano, M. O. Laukkanen, M. D. Castellone and N. Tsuchida, *Oncotarget*, 2017, **8**, 27075–27092.

39 Q. Leng, Y. Lin, M. Zhan and F. Jiang, *Oncotarget*, 2018, **9**, 24684–24692.

40 Y. Q. Shi, Q. He, Y. J. Zhao, E. H. Wang and G. P. Wu, *Cytotechnology*, 2013, **65**, 355–362.

41 X. W. Liu, H. J. Yu, Y. Qiao, J. J. Yang, J. Shu, J. X. Zhang, Z. W. Zhang, J. J. He and Z. Li, *Ebiomedicine*, 2018, **28**, 70–79.

42 Y. Qin, Y. Zhong, T. Ma, F. Wu, H. Wu, H. Yu, C. Huang and Z. Li, *Glycoconjugate J.*, 2016, **33**, 125–136.

43 L. Jia, J. Zhang, T. Ma, Y. Guo, Y. Yu and J. Cui, *Front. Oncol.*, 2018, **8**, 565.