Cent. Eur. J. Biol. • 7(2) • 2012 • 343-353
DOI: 10.2478/s11535-012-0016-5

VERSITA

Central European Journal of **Biology**

Proteomic analysis of differentially expressed serum proteins in lung cancer

Research Article

Shilpa Bhatnagar[1], Naveen Chaudhary[1], Shabnam Malik[1], Deepshikha Pande Katare[2], Swatantra K. Jain[1,*]

[1]*Department of Biotechnology, Hamdard University, 110062 New Delhi, India*

[2]*Amity Institute of Pharmacy, Amity University, 201303 Noida, Uttar Pradesh, India*

Received 11 August 2011; Accepted 28 December 2011

Abstract: Lung cancer continues to represent a major public health concern with high morbidity and mortality worldwide. Early detection of lung cancer is problematic due to a lack of diagnostic markers with high sensitivity and specificity. To determine the differently expressed proteins in the serum of lung cancer and identify the function of such proteins, two-dimensional electrophoresis (2DE) and liquid chromatography mass spectrometry (LC-MS) were used to screen the serum of lung cancer model induced by 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK). A total of 25 protein spots were qualitatively different and 6 were quantitatively different in the serum from rats bearing induced lung cancer when compared with normal controls. Two of the proteins that showed major changes in concentration in sera were identified to be Immunoglobulin γ 2A chain C region (heavy chain) and Transferrin by LC-MS/MS.

Keywords: *NNK • 2D Electrophoresis • LC-MS/MS • Transferrin • Immunoglobulin γ 2A chain C region*

1. Introduction

Lung cancer was considered to be a rare disease in the mid 1800s, representing only 1% of all cancers. However, it has now become one of the most common cancers worldwide, in terms of both incidence and mortality (1.35 million new cases per year and 1.18 million deaths), with the highest rates in Europe and North America [1]. In 2006, the most recent year for which statistics are currently available, lung cancer accounted for more deaths than breast cancer, prostate cancer and colon cancer combined. In India, in contrast to the mortality rate in men which began declining more than 20 years ago, the mortality rate in women due to lung cancer has been rising over the past decade [2]. Besides the reasons of environmental pollution, smoking and heredity, a lack of early diagnostic tools is also an important factor contributing to the rising incidence of lung cancer. It is also difficult to screen a large population for a disease, especially because of a lack of awareness regarding the usefulness of regular blood tests and routine check-ups. In recent years with the development of biological and imaging techniques such as spiral CT, cytological examination of sputum and endoscopy of bronchus, the positive rate of early diagnosis is increasing [3]. However, so far this has not resulted in a significant reduction in disease specific mortality, and the overall five year survival rate for lung cancer remains only 14%. Therefore, there is an urgent need to develop tools for early diagnosis and monitoring of disease progression. Among these tools, validated biomarkers are viewed as the most important, thus necessitating the discovery of new specific biomarkers for lung cancer.

Proteomic study is currently considered a powerful tool for evaluation of large scale protein expression, and has been widely applied in analysis of various diseases, especially cancer research [4-7]. Biomarker discovery and validation is a central application in current proteomic research and it facilitates the early diagnosis, treatment, monitoring and prognosis of many kinds of cancers.

* *E-mail: skjain@jamiahamdard.ac.in, sk608@rediffmail.com*

Protein based biomarkers have great potential as cancer biomarkers due to the fact that mRNA levels may not always indicate directly the levels of proteins, the molecules that are actually involved in biological functions [8]. Post-translational modifications such as glycosylation, phosphorylation and methylation can represent specific tumor functions and interactions with the tumor microenvironment that can be more informative and advantageous than genomics, or transcriptomics alone.

Two-dimensional electrophoresis is a well established technique used for both qualitative and quantitative analysis of proteins, and has been applied to proteomic studies in several human cancers such as colorectal, gastric, breast, lung, prostate and pancreatic cancer [9-14]. Despite a number of limitations, the resolution of 2D gels is impressive, rendering this technology still a preferred tool in many proteomics studies [15]. Along with 2DE, two technical methods often applied in the research of serum proteomics include: Surface enhanced laser desorption and ionization time-of-flight mass spectrometry (SELDI-TOF-MS) and Matrix assisted laser desorption/ ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Despite various improvements of the 2-DE technique, the main inherent limitation of the method still remains that many proteins are expressed at such low levels that they escape detection. Final identification of protein requires spot removal from the gel, digestion, and analysis of peptides by mass spectrometry. SELDI-TOF-MS is less sensitive to high molecular weight protein (>20 kDa) and may have instrument-to-instrument variation. Despite great advantages and insights provided to cancer biology, several pre-analytical, analytical, and post-analytical limitations are still associated with MS methodologies. Pre-analytical limitations are related to poor patient selection; non-standardized sample collection, storage, and processing; and poor instrument calibration, all of which introduced important bias in the analyses. Some MS techniques such as MALDI MS applied to fresh tissue or blood samples do not readily allow direct identification of proteins. Post-analytical limitations are mainly due to bioinformatics/biostatistics artifacts.

Blood serum is most commonly used in biomarker studies as it has the advantage of easy access through routine blood chemistry from patients. Potential biomarkers are most often found in blood and include biomarkers existing in biopsied cancer tissue and many circulating protein fragments originating in the diseased tissue microenvironment [16-18]. To date, there are no validated serum biomarkers available for early diagnosis of lung cancer. Hence, the search is ongoing to improve early and specific detection and disease monitoring. In the present study, 2DE and LC-MS/MS were employed to investigate protein expression alterations in the sera of animals bearing lung cancer, to identify the role of tumor associated molecules as potential biomarkers.

2. Experimental Procedures

2.1 Animals

Pathogen free, adult male Wistar rats (n=24) were procured from the Central Animal House facility of Jamia Hamdard University, New Delhi. Animals ranged in weight from 100-120 g and were housed in a temperature controlled room (25+/-2°C) in polypropylene cages, usually in groups of three. They were exposed to alternate light/dark cycles of 12 hours per day with access to water and chow *ad libitum* (Hindustan Lever Ltd.).

2.2 Lung cancer induction in animal models

An induced lung cancer model was developed in rats by modifying the protocol of Belinsky *et al*. [19]. Rats were divided into two groups namely, control and treated, each group contained 12 rats. A single high dose (2.5 mg/kg b.wt.) of 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK; Sigma Chemical Co., USA) was administered subcutaneously. This was followed by NNK (1.5 mg/kg b.wt.) administration on alternate days for one month. The progression of disease was followed continuously. Animals were sacrificed each month and lungs collected. Lung tissue was fixed in formalin and embedded in paraffin wax for sectioning and staining with hematoxylin and eosin (H&E). Islands of tumor cells in alveolar septa could be seen in treated animals after eight months of NNK treatment. At this stage blood samples were collected and serum proteins analyzed from both NNK treated and controls.

2.3 Preparation of serum sample

Rats were anaesthesized with diethyl-ether and blood was drawn from their tail vein. Studies were conducted in accordance with the guidelines of the “Committee for the Purpose of Control and Supervision of Experiments on Animals” (CPCSEA, Project No: 503). Blood drawn from both control and treated rats was coagulated at room temperature for 30 minutes followed by centrifugation at 1685x*g* for 20 minutes [20]. Serum was then aliquoted and stored at -80°C.

2.4 High-abundance protein depletion in serum

Serum samples from both control and treated rats were processed to selectively deplete one of the most abundant proteins, albumin which account for nearly

half of the total serum protein content [21]. A dye based proteoprep kit for selectively depleting albumin was used according to the manufacturer's instructions (Bangalore Genei). Depleted serum samples were recovered and stored at -80°C for long term storage. Gradient SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed using 12% separating gel with the discontinuous buffer system as described by Laemmli [22] to observe the protein profiles of control and treated serum samples.

2.5 Two-dimensional gel electrophoresis

2.5.1 Acetone/TCA precipitation

Briefly, 80-100 µl of the albumin depleted serum was mixed with 500 µl of 10% TCA in acetone. The protein was precipitated overnight at -20°C and centrifuged at 8800x*g*, 4°C for 15 minutes. The supernatant was discarded and pellet was washed with 500 µl of 90% ice cold acetone and was air dried for approximately 5 minutes taking care not to over dry it. For two-dimensional gel electrophoresis, the protein pellet was suspended in 100 µl of lysis buffer containing 7 M urea, 2.5 M thiourea, 40 mM Tris- HCl, 4% w/v CHAPS and 0.5% v/v Igepal CA-630. The protein sample was stored frozen at -80°C until further analysis. Protein concentration was measured by Bradford method [23].

2.5.2 Isoelectric focusing (IEF)

In the first dimension, proteins were resolved by isoelectric focusing using immobolized linear pH gradient (IPG) strips (GE Healthcare, pH 3-10, 11 cm). 100-120 µg of total serum protein diluted in rehydration buffer containing 8 M urea, 2% CHAPS, 0.25% v/v pH 3-10 ampholytes, 30 mM dithiothreitol (DTT) and 0.1% bromophenol blue (BPB) was loaded to obtain a total volume of 200 µl per IPG strip. The strip was allowed to rehydrate to its original thickness (0.5 mm) overnight at room temperature for 12-16 hrs. Rehydrated IPG strips were then focused using Ettan IPG phor II isoelectric focusing system (GE Healthcare) according to following protocol:

500 V-1 hr, 1000 V-1 hr, 6000 V-2 hrs, 6000 V for a total of 40,000 Vhrs.

Following IEF, IPG strips were washed with MilliQ water and incubated in equilibration buffer (30% glycerol, 6 M urea, 2% SDS, 75 mM Tris-HCl pH 8.8, 0.002% BPB) containing 1% DTT on a rocking bed for 15 minutes, followed by a 15 minute incubation in the same equilibration buffer (without DTT) containing 25% iodoacetamide while rocking.

2.5.3 SDS-PAGE

The second vertical SDS-PAGE was performed using SE 600 Ruby electrophoresis unit (GE Healthcare). In the 2nd dimension, proteins were separated on the basis of their molecular weight. The equilibrated IPG strips were placed on top of a 12% homogenous SDS polyacrylamide gel (14x16 cm) and sealed with 0.5% agarose. 12% SDS-PAGE gels 40 ml contained 30% ACR/0.8% BIS 16 ml, 1.5M Tris-HCl 10 ml (pH 8.8), Milli-Q water 13.2 ml, 10% APS 0.400 ml, 10% SDS 0.400 ml, TEMED 0.016 ml. Gels were run in parallel at constant voltage (100-120 V) in 190 mM Tris-glycine running buffer (pH 8.3) until the bromophenol blue dye reached the bottom of each gel.

2.6 Protein staining

The gels were incubated in fixative (50% Methanol, 5% Glacial Acetic Acid) for 1-2 h and were then transferred to a 50% methanol solution for 10 minutes. After washing the gels with distilled water the gels were incubated with constant shaking in 0.02% sodium thiosulphate for 2 minutes. Following incubation the gels were rinsed twice with distilled water for 1 min each. The gels were later stained in 0.1% silver nitrate solution (chilled) for 20 min with constant shaking. The solution was discarded and gels were rinsed gently with distilled water for 5 min. For visualization of the protein bands, the gels were placed in a freshly prepared developing solution containing 2% sodium carbonate and 80% formaldehyde. The reaction was stopped by adding 5% glacial acetic acid to the gels after the desirable intensity of the protein bands was obtained.

2.7 Image analysis

Gel images were scanned with an image scanner (Biorad Versa doc 4000 MP). Spot detection, quantification, matching and comparison of 2DE protein patterns of both control and treated serum samples were performed with PD Quest 2D analysis software version Advanced 8.0 (Biorad, Hercules, CA). For protein matching, serum protein gel maps of normal controls were regarded as a reference, and artificial matched spots were constructed and the analysis of differential expression was performed by applied software. Spots that were either present in only one of the groups or showed statistically significant changes in expression intensity were selected for LC-MS analysis.

2.8 Protein identification by liquid chromatography mass spectrometry (LC-MS)

The chosen protein spots in 2DE maps were excised manually using a clean ethanol wiped scalpel followed by washing three times with 1 ml milliQ water for 10 minutes each. The spots were then destained by adding 15 mM potassium ferricyanide and 50 mM sodium thiosulphate for 5-10 minutes until the gel pieces

start to appear yellow in color. The supernatant was discarded and 250 µl of 200 mM ammonium bicarbonate was added for 15 minutes until the gel pieces turned colorless. 200 µl of 50:50 (v/v) 200 mM ammonium bicarbonate: acetonitrile was added to the gel pieces kept on rocker for 15-20 minutes. Further supernatant was discarded and the gel pieces were dehydrated with HPLC grade acetonitrile. Upon complete evaporation of acetonitrile, each gel piece was rehydrated at 4°C in 20 ng/µl trypsin prepared in 50 mM ammonium bicarbonate. The gel pieces were incubated overnight at 37°C. To stop the action of trypsin, 1% formic acid was added. The sample was then spun down and the supernatant containing tryptic peptides was collected and stored at -20°C. The nano LC was performed with an Agilent 1100 NanoLC-1100 system combined with a microwell-plate sampler and thermostatted column compartment for pre-concentration (LC Packings, Agilent). Approximately 6 µl of the sample was loaded on the column (Zorbax 300SB-C18, 150 mm x 75 µm, 3.5 µm) using a pre-concentration step in a micro pre-column cartridge (Zorbax 300SB-C18, 5 mm x 300 µm, 5 µm) at a flow rate of 5 µl/min. After 5 min, the pre-column was connected with the separating column, and multistep gradient was initiated (3% for 5min, 15% for 5-8 min, 45% for 8-50 min, 90% for 50-55 min, 90% for 55-70 min, then 3% for 71 min). The buffers used were 0.1% formaldehyde in water (A) and 0.1% formaldehyde in 90% ACN (B). Calibration of the instrument was performed using the standard Agilent tune mix. An LC/MSD Trap XCT with a nano-electrospary interface (Agilent) operated in the positive ion mode was used for MS. Ionization (1.5 kV ionization potential) was performed with a liquid junction and a non-coated capillary probe (New Objective, Cambridge, USA). Peptides are ionized in the liquid phase in the Electrospray ionizer and enter the ion trap, are fragmented (MS/MS) and then detected. Peptide ions were analyzed by the data-dependent method as follows: The scan sequence consisted of 1 full MS scan followed by 4MS/MS scans of the most abundant ions. Each MS/MS spectrum (corresponding to a specific peptide sequence) generated was then used to search for matched peptides with the help of MASCOT software (Matrix Science) using Agilent Ion Trap Analysis software version 5.2. Next, the National Center for Biotechnology non-redundant (NCBInr) database was searched with rat as the taxonomy. A protein hit was identified often by multiple independently sequenced peptides from the same protein and the data were checked for consistent error distribution. In Mascot, the ion score for an MS/MS match is based on the calculated probability, P, that the observed match between experimental data and the database sequence is a random event. The reported score is -10log(P).

2.9 Western blot analysis

To confirm the results of LC-MS/MS, an equal amount of protein (20 µg) was loaded and resolved by 10% SDS-PAGE. Proteins were then electrotransferred to nitrocellulose membranes (Whatman) and then blocked for 1 hour at room temperature in a blocking solution containing 3% BSA (Bovine Serum Albumin) and 0.1% Tween 20 in Tris-buffered saline (TBS) (50 mM Tris-HCl and 150 mM NaCl). The membrane was then incubated with primary antibody (Rabbit anti rat transferrin, 1:500; Santa Cruz and Mouse anti rat IgG2A, 1:2000; Sigma) in blocking solution at 4°C overnight. The membrane was rinsed with 0.1% Tween 20 in TBS and incubated with horse radish peroxidase-conjugated secondary antibodies (Goat anti-rabbit IgG; 1:5,000; AbD Serotec and goat antimouse IgG; 1:10,000; Santa cruz) for 1 hour at room temperature. The immunoreaction was detected using 10 mL of DAB (3,3-diaminobenzidine tetrahydrochloride) solution (0.6 mg/ml) in 1X TBS buffer and stopping the reaction with 30% H_2O_2 when the desired intensity was obtained.

3. Results

Lung cancer was induced in rats successfully after administering subcutaneous injections of NNK. Approximately 95-98% rats developed tumors and the latency period between the administration and induction of tumors was consistent among all rats. After H&E staining and pathological examination, we observed that test group animals showed an island of tumor cells in their alveolar septa. Many atypical binucleate cells and a mitotic figure could be identified in the island of tumor cells under 400X high power magnification At the time of sacrificing the the tumor size was 122x62 microns (Figure 1).

3.1 Differential protein expression in serum of treated and control rats

The 1D protein expression profiles in serum of NNK treated and control rats are shown (Figure 2). Several protein bands at (~66 kDa and ~49 kDa) in the treated sample have been found to be up-regulated in comparison to the control.

Gel images and 2D patterns of normal and treated serum samples revealed well resolved and reproducible profiles (Figure 3). Triplicate sets of 2D gel electrophoresis experiments were performed to increase the reliability of the results. The experiments were repeated on serum from three individual rats

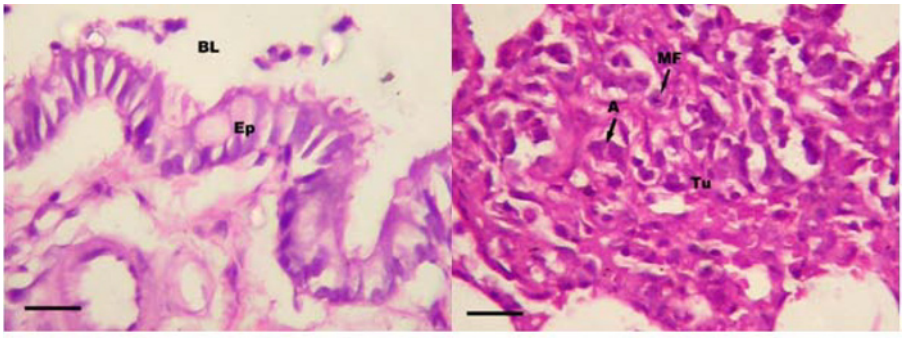

Figure 1. Histological analysis of lung tissue A) Normal lung specimen, showing the wall of bronchiole with normal bronchial epithelium lining. BL = Bronchial Lumen, Ep = Epithelium. (HE x 400). Scale bar = 20 microns. B) 400x high power photomicrograph from test group animal (after 8 months of treatment) showing an island of tumour cells in the alveolar septa. Many atypical cells are seen in the island. A binucleate form and a mitotic figure can be identified. Tu-Tumour cells, A-Atypical Binucleate cell, MF -Mitotic Figure. (HE x 400). Scale bar = 20 microns. Tumor size is 122x62 microns.

and a representative gel is shown in each case. Approximately 130 and 73 protein spots in 2D gels of control and treated serum samples, respectively, were detected by silver staining. After gel to gel matching and normalization, using PD Quest software, the differentially expressed proteins were defined on the basis of change in expression intensities. Qualitative differences were considered meaningful when one spot was detected in one gel and not in another. Each matched protein spot was assigned a unique *SSP* (sample spot protein) *number* by the *PD Quest* software. Based on these criteria, it was found that 15 protein spots were present exclusively in treated animals that were absent in controls. Similarly, 9 proteins that were present in control animals were found to be totally absent in treated animals. Around 6 protein spots were found to be differentially expressed between control and treated samples. Comparison of the 6 differentially expressed protein spots (A-F) before and after lung cancer formation revealed a decrease in intensity in all the six spots (A-F) in treated samples over control. The data are expressed as the mean +/- standard deviation, with P<0.001(***) as limits of significance (Figure 4). Spots shown to be differentially expressed were then subjected to further analysis. We present the LC-MS/MS analysis data of two such spots that were found to have expressed in only one sample, either control or treated. These proteins have been identified as transferrin (control) and Immunoglobulin γ 2A chain C region (treated), using NCBInr database search by MASCOT software (Table 1). Redundancy of proteins appearing in the database under different names and accession numbers was eliminated. The apparent molecular weight of spot 2 (identified as transferrin) is less than 29 kDa. It is possible that this

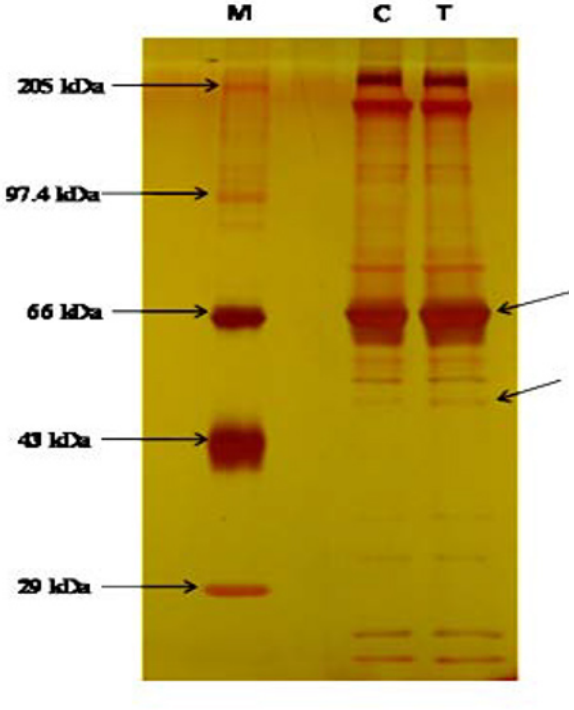

Figure 2. 1-D protein profiles of control (C) and NNK treated (T) serum samples; (M-Broad range protein marker). Few protein bands (~66 kDa and ~49 kDa, marked with arrows) in treated serum sample are little up-regulated as compared to the control serum sample.

may be a truncated form of transferrin. Further analysis of this protein is ongoing. An image of the master gel automatically produced by PD quest software showed all the proteins from both control and NNK treated serum samples (Figure 5). The red spots are the matched proteins that are quantitatively up-regulated in treated samples *versus* control, while green are the matched proteins that are quantitatively down-regulated in

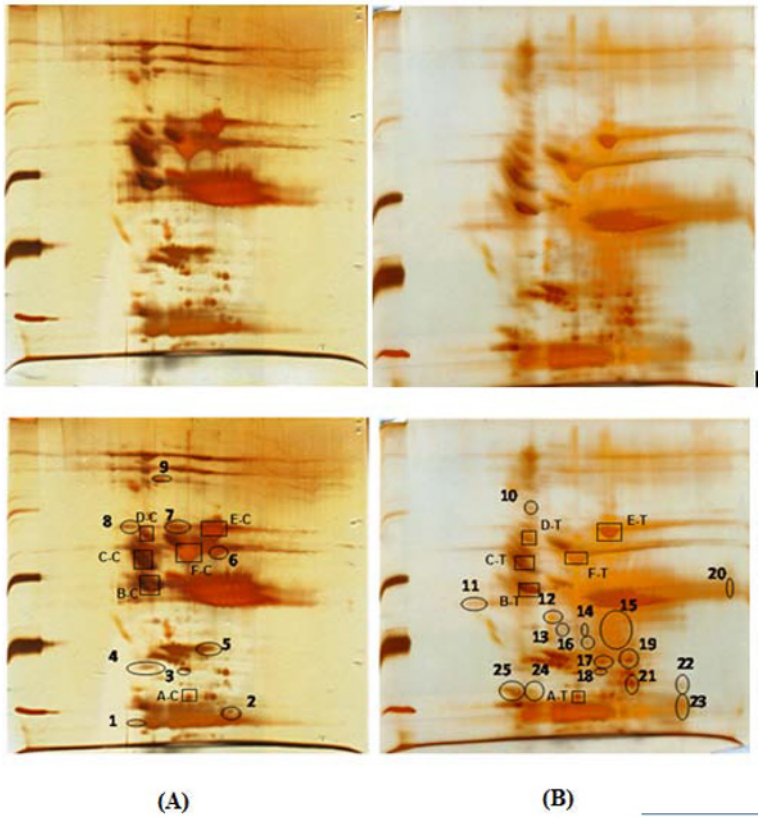

Figure 3. 2D protein profiles of (A) Control and (B) NNK treated serum samples, after lung cancer formation, acquired using IPG strips (11 cm, pH 3-10). Spots 1-9 are present only in control while spots 10-25 are present only in NNK treated serum samples. Spots A-F are present in both the gels but are differentially expressed. These A-F spots have been denoted as A,B,C,D,E,F-C in serum samples from control animals while as A,B,C,D,E,F-T in serum samples from NNK treated animals.

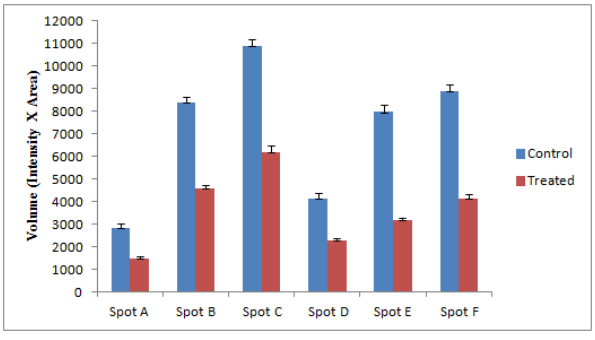

Figure 4. Comparison of content (spot intensity x area) of 6 differentially expressed proteins before and after lung cancer development. The protein content in spots is decreased in spots A-F. The data are expressed as mean +/- standard deviation. Comparisons of the means of protein content in sera of control and treated rats were made by student's Tukey t-test, using GraphPad Instat 3 software, with, $P<0.001$ (***) as limits of significance.

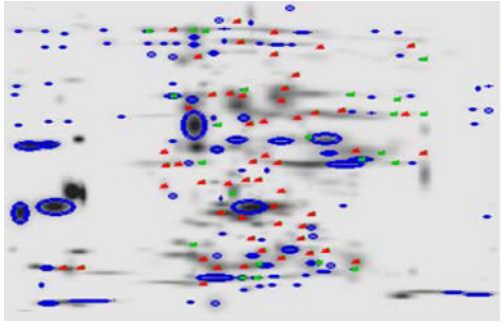

Figure 5. Master gel. The image is automatically produced by PD quest software, having all the proteins from both control and NNK treated serum samples. The red spots are the matched proteins that are quantitatively up-regulated in treated samples *versus* control, while green are the matched proteins that are quantitatively down-regulated in treated samples *versus* control, blue spot on the other hand show qualitative differences between control and treated.

S.No	Spot number	Protein name	Theoretical MW (kDa)	pI	MOWSE Score
1	2 (from Control)	Transferrin	78.5	6.94	167
2	19 (from Treated)	Immunoglobulin γ 2A chain C region (Heavy)	35.67	7.72	194

Table 1. Qualitatively Newly Expressed Proteins Identified by LC-MS in Serum of Lung Cancer Induced and Normal Healthy Control Rats.

treated samples *versus* control, and blue spots indicate qualitative differences between control and treated. The representative data of mass spectrometry of these two proteins have been shown in Figures 6A and 7A, respectively. The Mascot search histograms for both the proteins (Transferrin & Immunoglobulin chain C region) are shown in Figures 6B and 7B, respectively.

Data analysis was carried out using the database search engine Mascot (www.matrixscience.com). NCBInr which is a comprehensive, non-identical protein database was used for the sequence search. The taxonomy search was confined to the genus, Rattus, thus ensuring that the hit list only contained entries from the selected species. The significant hit (gi|1854476) identified Spot 2 (CONTROL) as Transferrin in Rattus *norvegicus* and the significant hit (gi|121052) identified Spot 19 (TREATED) as Immunoglobulin γ-2A chain C region in Rattus *norvegicus*.

3.2 Probability based Mowse score

A histogram illustrates the protein score distribution. The 50 highest protein scores are divided into 16 bins according to their score, and the heights of the bars show the number of matches in each bin. For a search of MS/MS data, protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits. The green shaded area extends up to the average identity threshold for an individual peptide match. The protein score in the result report from an MS/MS search is derived from the ions scores as a non-probabilistic basis for ranking protein hits. In Mascot, the ions score for an MS/MS match is based on the calculated probability, P, that the observed match between the experimental data and the database sequence is a random event. The reported score is -10*Log(P), where P is the probability that the observed match is a random event. In the MASCOT score histogram (Figure 6B), the individual ion scores >40 indicate extensive homology (P<0.05) and are considered to be the significant scores while scores <40 (in shaded green region) are random hits and not significant.

In the MASCOT score histogram (Figure 7B), the individual ion scores >39 indicate extensive homology (P<0.05) and are considered to be the significant scores while scores <39 (in shaded green region) are random

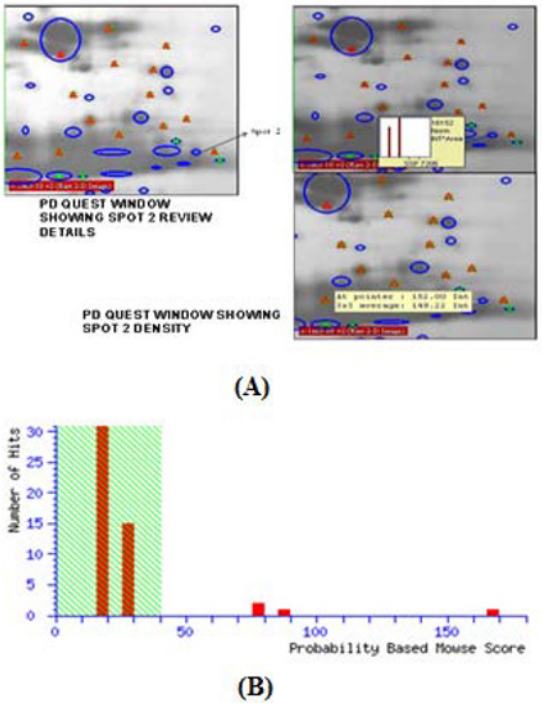

Figure 6. (A) Cropped/enlarged 2D gel images for analysis of spot 2 (SSP 7206) from control serum (transferrin) showing the spot density. Graph Partial Matches; each bar represents the spot's quantity in a gel. Spot 2 density is 152.00. Missing bars indicate that the spot in control is not matched in treated sample. (B) Mascot search results of protein spot 2 are Transferrin. Mowse Score is 167 (>40, P<0.05).

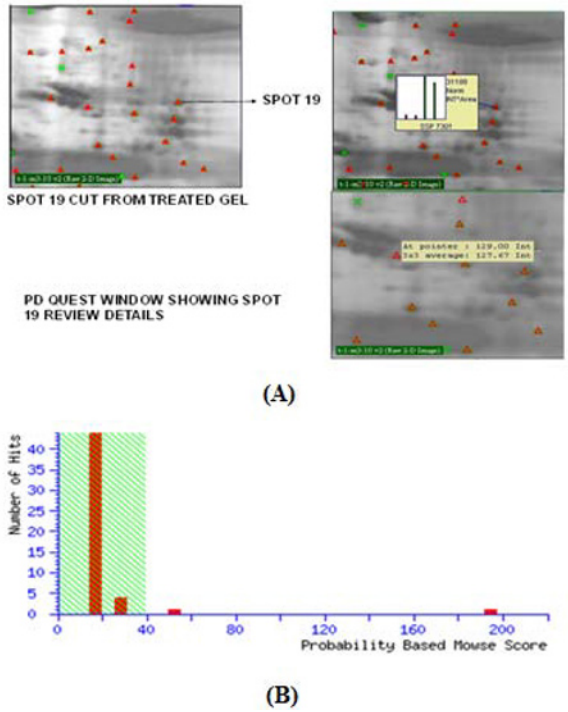

Figure 7. (A) Cropped 2D gel images of for analysis of spot 19 (SSP 7301) from treated serum showing the spot density. Graph Partial Matches; each bar represents the spot's quantity in a gel. Spot 19 density is 129.00. Histogram bars reveal variation in spot quantity between gels. (B) Mascot search results of protein spot 19 are Immunoglobulin γ 2A chain C region. Mowse Score is 194 (>39, P<0.05).

hits and not significant.

Peptide sequence of transferrin (spot 2 from control serum sample)

MRFAVGALLA CAALGLCLAV PDKTVKWCAV
SEHENTKCIS FRDHMK**TVLP** **ADGPR**LACVK
K**TSYQDCIK**A ISGGEADAIT LDGGWVYDAG
LTPNNLKPVA AEFYGSLEHP QTHYLAVAVV
KK**GTDFQLNQ** **LQGK**KSCHTG LGRSAGWIIP
IGLLFCNLPE PRKPLEKAVA SFFSGSCVPC
ADPVAFPQLC QLCPGCGCSP TQPFFGYVGA
FKCLR**DGGGD** **VAFVK**HTTIF EVLPQKADRD
QYELLCLDNT RKPVDQYEDC YLARIPSHAV
VARNGDGKED LIWEILKVAQ EHFGKGKSKD
FQLFGSPLGK DLLFKDSAFG CYGVPPRMDY
RLYLGHSYVT AIRNQREGVC PEASIDSAPV
KWCALSHQER AKCDEWSVTS NGQIECESAE
STEDCIDKIV NGEADAMSLD GGHAYIAGQC
GLVPVMAENY DISSCTNPQS DVFPKGYYAV
AVVKASDSSI NWNNLKGKKS CHTGVDRTAG
WNIPMGLLFS RINHCKFDEF FSQGCAPGYK
KNSTLCDLCI GPAKCAPNNR EGYNGYTGAF
QCLVEKGDVA FVKHQTVLEN TNGKNTAAWA
KDLKQEDFQL LCPDGTKKPV TEFATCHLAQ
APNHVVVSRK EKAARVSTVL TAQKDLFWKG
DKDCTGNFCL FRSSTKDLLF RDDTKCLTKL
PEGTTYEEYL GAEYLQAVGN IRKCSTSRLL
EACTFHKS

The matched peptides are highlighted in bold.

Peptide sequence of IgG 2A chain C region (spot 19 from treated serum sample)

AETTAPSVYP LAPGTALKSN **SMVTLGCLVK**
GYFPEPVTVT WNSGALSSGV HTFPAVLQSG
LYTLTSSVTV PSSTWSSQAV TCNVAHPASS
TKVDKKIVPR **ECNPCGCTGS** **EVSSVFIFPP**
KTK**DVLTITL** **TPK**VTCVVVD ISQNDPEVRF
SWFIDDVEVH TAQTHAPEKQ SNSTLRSVSE
LPIVHRDWLN GKTFKCKVNS GAFPAPIEKS
ISKPEGTPRG PQVYTMAPPK EEMTQSQVSI
TCMVK MNG QPQENYKNTP PTMDTDGSYF
LYSKLNVKKE TWQQGNTFTC SVLHEGLHNH
HTEKSLSHSP GK

The matched peptides are highlighted in bold.

3.3 Confirmation of the expression of transferrin and immunoglobulin γ 2A chain C region

Western blots confirmed that the transferrin protein was decreased/ depleted in treated rats' sera samples in comparison to control (Figure 8A). The IgG-2A levels were increased in treated rats' sera in comparison to control (Figure 8B).

4. Discussion

Blood serum contains a large number of secreted or shed low abundance proteins that are critical for signaling cascades and regulatory events. During necrosis, apoptosis and hemolysis, contents of cells may be released into the serum. The presence of these components in the blood reinforces the benefits of using a proteomic approach to identify biomarkers for disease states.

In the present study, we compared the protein profiles between the serum of control and lung cancer induced rats using 2-DE and LC-MS approaches. Using these approaches we have observed a number of proteins differentially expressed in control and treated rat sera. Two of these proteins have been characterized; one has been identified as transferrin (found in control sera sample and depleted in treated sample), while the other protein, immunoglobulin γ 2A chain C region (found to be upregulated in treated serum sample). Tumor markers are secreted, released, or leaked into the interstitial fluids, and thus into the lymph, and finally (or directly) into the bloodstream, where they become detectable in serum samples. To be able to enter the bloodstream directly, larger molecules, often proteins, are cleaved into truncated forms or fragments, which are sometimes specific to the protease micro-environment of the tumor. In our study, based on apparent molecular weight of putative transferrin, it is possible that its truncated derivative is present in the spot.

Transferrin is a plasma protein that transports iron through the blood to the liver, spleen and bone

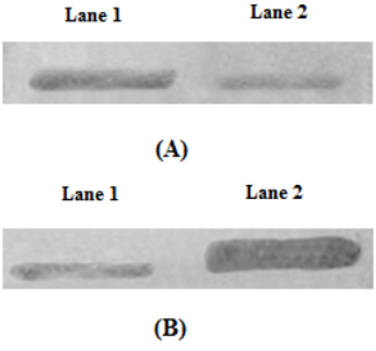

Figure 8. Western blot (A) Analysis with transferrin specific antibody confirms the down-regulation of the protein in NNK treated rats' sera in comparison to control. Lane 1- Control, Lane 2- NNK treated. (B) Analysis with Immunoglobulin γ 2A chain C region specific antibody confirms the up-regulation of protein in NNK treated rats' sera in comparison to control. Lane 1- Control, Lane 2- NNK treated.

marrow. It is a glycoprotein that binds iron very tightly but reversibly. It is an 80 kDa protein with homologous N-terminal and C-terminal iron binding domains [24]. The liver is the main source of transferrin, but other sources such as the brain also produce this molecule. The main role of transferrin is the transportation of iron from the sites of absorption (duodenum) and red blood cell recycling macrophages to virtually all tissues, particularly the tissues involved in erythropoiesis and those with actively dividing cells. Transferrin predominantly plays a key role where erythropoiesis and active cell division occur. Transferrin is also associated with the innate immune system, creating an environment low in free iron, where few bacteria are able to survive. The serum transferrin receptor (sTfR) is a sensitive indicator of iron deficiency erythropoiesis which is not affected by inflammation. Concentrations of this molecule are inversely correlated with body iron-stores [25]. In fact, increases in body iron stores have been reported in patients with liver and lung cancers. Specifically, various investigators have shown that tissue and serum ferretin levels are abnormally high, while serum iron and serum transferrin levels [26,27] are low in breast cancer and other types of cancers. These latter conclusions were made by assessing the saturation of transferrin with iron and by examining the serum levels of the iron-storage protein, ferretin [28]. Most recently transferrin and its receptor have been tested to diminish tumor cells by using the receptor to attract antibodies. NNK treatment has been demonstrated to enhance the expression of fatty acid synthase, transketolase, pulmonary surfactant-associated protein C, L-plastin, annexin A1, and haptoglobin, while the expression of transferrin, α-1-antitrypsin, and apolipoprotein A-1 was decreased [29].

Another protein that was identified in the sera of treated rats was Immunoglobulin 2A chain C region. Studies by Gianazza *et al.* on rat serum identified 34 proteins with human homologues and concluded that even abundant proteins could be markers for disease states [30-32]. Traditionally, it was believed that the only source of Immunoglobulins (IgGs) was mature B lymphocytes, but recently, some researchers have reported that IgG could also be detected in carcinoma cells derived from epithelium [33]. Elevated levels of IgG, IgM or IgA antibodies are frequently observed in patients with cancers of epithelial origin, including carcinomas of breast, colon and liver [34]. IgG-specific monoclonal or polyclonal antibodies were used for immunohistochemistry to determine the expression of IgG in malignant epithelial tumors. In all evaluated cancer tissues, including those from breast cancer (n=10), liver cancer (n=14), colon cancer (n=6) and lung cancer (n=12), positive staining was detected in the cytoplasm of all cancer tissues. In contrast, IgG was not detected in similar number of normal tissues of breast, colon, liver and lung [35].

IgG with unidentified specificity can also be secreted directly from cancer cells, and these antibodies are involved in their survival and growth [36]. In the serum and tissue of lung squamous cell carcinoma, different levels of Ig protein fragments were detected which may present evidence for IgG secretion [35,37]. This characteristic pattern of IgG expression in cancerous tissue may serve as a potential marker for malignant cell transformation. However, as of yet, because of the scarcity of reports on the gene expression of IgG in cancer cell lines, cancer tissues and serum, there is still a controversy regarding the expression of IgG protein by them.

5. Conclusions

In our experiment, the focus of observation and comparison was on the sera from rats bearing induced lung cancer matched to healthy controls. By regulating and optimizing all conditions of the 2D technique, we finally obtained satisfactory patterns in which approximately 203 serum protein spots were detected from the treated and control animals, collectively. In addition, 25 qualitatively different protein spots were acquired by image analysis. Using the LC-MS technique, we identified two differentially regulated proteins, transferrin and immunoglobulin γ 2A chain C region, from control and treated sera respectively. The depletion of transferrin in the serum samples of NNK treated animals suggests its role and involvement of iron tumorogenesis of lungs. The overexpression of immunoglobulin γ 2A chain C region in NNK induced rats suggests that being a member of immunoglobulins, IgG 2A chain C region may regulate a series of changes of immunity in the process of forming lung cancer. In conclusion, i) the serum proteins were separated successfully for proteomic analysis between the normal and treated rats' sera by the optimized 2DE technique. ii) There were a number of differently expressed proteins between the normal control and treated rats' sera. iii) The differentially expressed proteins, transferrin and immunoglobulin 2A chain C region in serum of rats bearing induced lung cancer *versus* control rats were confirmed through western blotting. These proteins may be helpful in improving lung cancer diagnosis and therapy. Years of clinical validation studies and their subsequent approval may

reveal a protein, or combination of proteins, that can be used to reliably assess all stages of lung cancer. Ultimately such biomarkers would aid clinicians in diagnosing lung cancer during the early stages, eliminating the need for lung biopsy and allowing early treatment, thereby preventing the progression of fibrosis.

Acknowledgements

This work was supported by a research project from Defense Research and Development Organization, Government of India (ERIP/ER/0703641/M/01/1050) to SKJ & DPK. SM is an ICMR SRF. We thank Dr. Ashok Mukherjee for interpretation of histopathological slides.

References

[1] Commonly Diagnosed Cancers Worldwide, Cancer Research UK, 2005
[2] Jemal A., Tiwari R.C., Murray T., Ward E., Samuels A., Ghafoor A., et al., Cancer statistics, CA Cancer J. Clin., 2005, 54, 8-29
[3] Li M., Ye B., Zhang Y., Chen H., Xia D., Liu M., et al.,Proteomic analysis of serum in lung cancer induced by 3-methylcholanthrene, J. Biomed. Biotechnol., 2009, 1-12
[4] Pandey A., Mann M., Proteomics to study genes and genomes, Nature, 2000, 405, 837–846
[5] Alessandro R., Fontana S., Kohn E., De Leo G., Proteomic strategies and their application in cancer research, Tumori., 2005, 91, 447–455
[6] Koomen J.M., Haura E.B., Bepler G., Sutphen R., Remily-Wood E.R., Benson K., et al., Proteomic contributions to personalized cancer care, Mol. Cell. Proteomics, 2008, 7, 1780–1794
[7] Peng X.Q., Wang F., Geng X., Current advances in tumor proteomics and candidate biomarkers for hepatic cancer, Expert Rev. Proteomics, 2009, 6, 551-561
[8] Simpson R.J., Bernhard O.K., Greening D.W., Moritz R.L., Proteomics-driven cancer biomarker discovery: looking to the future, Curr. Opin. Chem. Biol., 2008, 12, 72-77
[9] Cho H.J., Baek K.E., Park S.M., Kim K.I., Choi Y.L., Cho H.J., et al., RhoGDI2 expression is associated with tumor growth and malignant progression of gastric cancer, Clin. Cancer Res., 2009, 15, 2612–2619
[10] Hongsachart P., Huang-Liu R., Sinchaikul S., Pan F.M., Phutrakul S., Chuang Y.M., et al., Glycoproteomic analysis of WGA-bound glycoprotein biomarkers in sera from patients with lung adenocarcinoma, Electrophoresis, 2009, 30, 1206–1220
[11] Ono M., Matsubara J., Honda K., Sakuma T., Hashiguchi T., Nose H., et al., Prolyl 4-hydroxylation of alpha-fibrinogen: A novel protein modification revealed by plasma proteomic, J. Biol. Chem., 2009, 284, 29041–29049
[12] Schulz D.M., Böllner C., Thomas G., Atkinson M., Esposito I., Hofler H., et al., Identification of differentially expressed proteins in triple-negative breast carcinomas using DIGE and mass spectrometry, J. Proteome Res., 2009, 8, 3430–3438
[13] Vellaichamy A., Sreekumar A., Strahler J.R., Rajendiran T., Yu J., Varambally S., et al., Proteomic interrogation of androgen action in prostate cancer cells reveals roles of aminoacyl tRNA synthetases, PLoS One, 2009, 4, e7075
[14] Mathivanan S., Lim J.W., Tauro B.J., Ji H., Moritz R.L., Simpson R.J., Proteomic analysis of A33-immunoaffinity-purified exosomes released from the human colon tumor cell line LIM1215 reveals a tissue-specific protein signature, Mol. Cell. Proteomics, 2010 9, 197–208
[15] Sagynaliev E., Steinert R., Nestler G., Lippert H., Knoch M., Reymond M.A., Web-based datawarehouse on gene expression in human colorectal cancer, Proteomics, 2005, 5, 3066–3078
[16] Sung H.J., Cho J.Y., Biomarkers for the lung cancer diagnosis and their advances in proteomics, BMB reports, 2008, 41, 615-625
[17] Zhang Y., Li M., Liu Y., Han N., Zhang K., Xiao T., et al., ENO1 protein levels in the tumor tissues and circulating plasma samples of non-small cell lung cancer patients, Chinese Journal of lung cancer, 2010, 13, 1089-1093
[18] Ugurel S., Bell N., Sucker A., Zimpfer A., Rittgen, W., Schadendorf D., Tumor type M2 pyruvate kinase (TuM2-PK) as a novel plasma tumor marker in melanoma, Int. J. Cancer, 2005, 117, 5, 825-830
[19] Belinsky A., White M., Boucheron A., Richardson C., Swenberg A., Anderson M., Accumulation and persistence of DNA adducts in respiratory tissue of rats following multiple administrations of the tobacco specific carcinogen 4-(N-Methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone, Cancer Res., 1986, 46, 1280-1284

[20] Tuck M.K., Chan D.W., Chia D., Godwin A.K., Grizzle W.E., Krueger K.E., et al., Standard operating procedures for serum and plasma collection: early detection research network consensus statement Standard Operating Procedure Integration Working Group, J. Proteome Res., 2009, 8, 113–117
[21] Anderson N.L., Anderson N.G., The human plasma proteome: history, character, and diagnosis prospects, Mol. Cell. Proteomics, 2002, 1, 845-67
[22] Laemmli U.K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature, 1970, 227, 680–685
[23] Bradford M.M., Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem., 1976, 72, 248-254
[24] Huebers H.A., Finch C.A., The physiology of transferrin and transferrin receptors, Physiol. Rev., 1987, 67, 520-582
[25] Kuvibidila S., Gauthier T., Warrier R.P., Rayford W., Increased levels of serum transferrin receptor and serum transferrin receptor/log ferritin ratios in men with prostate cancer and the implications for body-iron stores, J. Lab. Clin. Med., 2004, 144, 176-182
[26] Faulk W.E., His B.L., Stevens P.J., Transferrin and transferrin receptors in carcinoma of the breast, Lancet, 1980, 2, 390-392
[27] Lamoureux G., Mandeville R., Poisson R., Poisson S.L., Jolicoeur R., Biological markers and breast cancer. A multiparametric study, increased serum protein levels, Cancer, 1982, 49, 502-512
[28] Stevens G., Graubard I., Micozzi S., Neriishi K., Blumberg B.S., Moderate elevation of body iron level and increase risk of cancer occurrence and death, Int. J. Cancer, 1994, 56, 364-369
[29] Kassie F., Anderson L.B., Scerber R., Yu N., Lahti D., Upadhyaya P., Indole-3-carbinol Inhibits 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone Plus Benzo(a)pyrene–Induced Lung Tumorigenesis in A/J Mice and Modulates Carcinogen-Induced Alterations in Protein Levels, Cancer Res., 2007, 67, 6502-6511
[30] Knekt P.M., Reunanen A., Takkunen H., Body iron stores and risk of cancer, Int. J. Cancer, 1994, 56, 379-82
[31] Eberini I., Miller I., Zancan V., Bolego C., Puglisi L., Gemeiner M., et al., Proteins of rat serum IV. Time-course of acute-phase protein expression and its modulation by indomethacine, Electrophoresis, 1999, 20, 846–853
[32] Eberini I., Agnello D., Miller I., Villa P., Fratelli M., Ghezzi P., et al., Proteins of rat serum V: adjuvant arthritis and its modulation by nonsteroidal anti-inflammatory drugs, Electrophoresis, 2000, 21, 2170–2179
[33] Haynes P., Miller I., Aebersold R., Gemeiner M., Eberini I., Lovati M.R., et al., Proteins of rat serum: I establishing a reference two-dimensional electrophoresis map by immunodetection and microbore high performance liquid chromatography electrospray mass spectrometry, Electrophoresis, 1998, 19, 1484-1492
[34] Kimoto Y., Expression of heavy-chain constant region of immunoglobulin and T-cell receptor gene transcripts in human non-hematopoietic tumor cell lines, Genes Chromosomes Cancer, 1998, 22, 83–86
[35] Qiu X., Zhu X., Zhang L., Mao Y., Zhang J., Hao P., et al., Human epithelial cancers secrete immunoglobulin G with unidentified specificity to promote growth and survival of tumor cells, Cancer Research, 2003, 63, 6488-6495
[36] Olubuyide I.O., Salimonu L.S., Adeniran S.O., Soluble immune complexes and immunoglobulin (IgG, IgA and IgM) levels in Nigerians with primary liver cell carcinoma, Afr. J. Med. Sci., 1993, 22, 57-62
[37] Ouyang G., Chen Y., Ruan L., Differential phosphoproteome analysis of nasopharyngeal carcinoma and normal nasopharyngeal epithelial tissues, Int. J. Path. Clin. Med., 2007, 27, 461-469