# **RESEARCH ARTICLE**

# **Comparative proteomics and global genome-wide expression data implicate role of ARMC8 in lung cancer**

Asif Amin<sup>1</sup>, Shoiab Bukhari<sup>1</sup>, Taseem A Mokhdomi<sup>1</sup>, Naveed Anjum<sup>1</sup>, Asrar H Wafai<sup>1</sup>, Zubair Wani<sup>1</sup>, Saima Manzoor<sup>1</sup>, Aabid M Koul<sup>1</sup>, Basit Amin<sup>1</sup>, Qurat-ul-Ain<sup>1</sup>, Hilal Qazi<sup>1</sup>, Sumira Tyub<sup>2</sup>, Ghulam Nabi Lone<sup>3</sup>, Raies A Qadri<sup>1\*</sup>

# Abstract

Background: Cancer loci comprise heterogeneous cell populations with diverse cellular secretions. Therefore, disseminating cancer-specific or cancer-associated protein antigens from tissue lysates could only be marginally correct, if otherwise not validated against precise standards. <u>Materials and Methods</u>: In this study, 2DE proteomic profiles were examined from lysates of 13 lung-adenocarcinoma tissue samples and matched against the A549 cell line proteome. A549 matched-cancer-specific hits were analyzed and characterized by MALDI-TOF/MS. <u>Results</u>: Comparative analysis identified a total of 13 protein spots with differential expression. These proteins were found to be involved in critical cellular functions regulating pyrimidine metabolism, pentose phosphate pathway and integrin signaling. Gene ontology based analysis classified majority of protein hits responsible for metabolic processes. Among these, only a single non-predictive protein spot was found to be a cancer cell specific hit, identified as Armadillo repeat-containing protein 8 (ARMC8). Pathway reconstruction studies showed that ARMC8 lies at the centre of cancer metabolic pathways. <u>Conclusions</u>: The findings in this report are suggestive of a regulatory role of ARMC8 in control of proliferation and differentiation in lung adenocarcinomas.

Keywords: Lung adenocarcinoma - ARMC8 - MALDI-TOF/MS

Asian Pac J Cancer Prev, 16 (9), 3691-3696

#### Introduction

Lung cancer is the single most devastating cause of cancer-related deaths in developed as well as developing countries (Jemal, 2010). The high mortality rate associated with this disease is due to the late diagnosis in majority rate of the cases (Herbst, 2008; Youlden et al., 2008). Smoking is the major risk factor in lung cancer. Other risk factors include exposure to radon and asbestos, occupational exposures, hormones imbalance and genetic factors (Darby et al., 2001; Ganti et al., 2006; Ferlay et al., 2010; Bouchardy et al., 2011). Lung cancer remains a major health concern in Kashmir valley and constitutes about 10% of all cancers (Wani et al., 2014). In 2010, the number of lung cancer cases registered at SKIMS, the main oncology facility in the valley was at record high and surpassed the number of esophageal cancer cases the most prevalent cancer here. Despite advances in diagnostic and treatment modalities, there has been little improvement in survival rates over the past three decades (Siegel et al., 2012; Luqman et al., 2014).

Proteomic profiling is one of the most important strategies for cancer biomarker discovery, with obvious advantages of being functional mediators and thus the true representative of a biological phenotype. Proteins impart a huge complexity to the biological system and are functionally more diverse than nucleic acids. Furthermore, most functional events are mediated post-transcriptionally and cannot be predicted from the levels of DNA or RNA (Anderson and Seilhamer, 1997).

However, most proteomic studies rely upon heterogeneous and complex tumor tissue, containing besides neoplastic epithelial cancer cells, numerous other components such as immune inflammatory infiltrates, fibroblasts, endothelial cells and other stromal cell populations. Therefore, a mere change in protein expression may not necessarily reflect altered expression within the cancer cells but may possibly be due to altered intra or extracellular tissue microenvironment. Thus the heterogeneity of cell types in cancer tissue presents a problem concerning the identification of cancer specific proteins. Therefore in this study, we used comparative 2DE proteomic profiling of lung adenocarcinoma tissue samples (Simsek et al., 2013) against model lung adenocarcinoma cell line A549 to evaluate cancer cellspecific changes in protein expression.

<sup>1</sup>Department of Biotechnology, <sup>2</sup>Centre of Research for Development, University of Kashmir, Srinagar (J&K)-190006, <sup>3</sup>Department of Cardiovascular and Thoracic Surgery, Sher-i-Kashmir Institute of Medical Sciences, Srinagar (J&K)-190011, India. \*For correspondence: raies@kashmiruniversity.ac.in.

# Asif Amin et al Materials and Methods

#### Clinical specimens

Tissue specimens from clinically confirmed fresh tumors as well as case-controls were obtained fresh at the time of surgery from Department of Cardio Vascular and Thoracic Surgery, Sher-i-Kashmir Institute of Medical Sciences, Srinagar after written informed consent from patients and as per approved guidelines by the institutional ethics committee. Samples were transported on ice to the lab and immediately stored at -80°C until further use. After obtaining histopathological data, only adenocarcinoma samples were selected for the study. Samples with mixed histology, metastatic origin other than lung, or prior exposure to chemo/ radiotherapy were excluded.

#### Sample preparation

The frozen tissues were washed three times with chilled PBS. A total of 200 mg of tissue was ground to powder in liquid nitrogen with a precooled mortar and pestle. The powder was left in lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 50 mM DTT, 1% ampholyte) for 30 minutes on ice. The resulting lysates were then vortexed and incubated at room temperature for 10 minutes. After centrifugation at 12000 rpm at 4°C for removal of particulate matter, the protein solutions was collected and quantified for protein content. Protein concentrations were determined using the Micro BCA<sup>TM</sup> (bicinchoninic acid) protein assay kit (Pierce, USA).

#### Two-dimensional gel electrophoresis

Protein samples ( $200 \mu g$ ) were poured along the back edge of IEF tray channel (Biorad) and IPG strips (pH 4-7 NL, GE Amersham) placed carefully avoiding air bubbles. After initial rehydration of 12 hours, first dimensional separation was performed using 100 V for 30 minutes, 500 V for 3 hours and 3,500 V for 12 hours. At the end of the run the total volt hours was~29000 volts. The IPG strips were equilibrated in equilibration buffer in two steps, each of 15 minutes. The first step employed 50 mM Tris HCl buffer (pH 8.8), 6 M urea, 30% glycerol, 2% SDS, 0.002% bromophenol blue and 1% DTT, while 2.5% iodoacetamide replaced DTT in the second step. Second dimensional separation was performed by using 10% SDS polyacrylamide gel. Electrophoresis was performed in a Hoefer system at 40 mA at room temperature for 4 hours.

#### Image analysis

Following the second dimension, the gels were subsequently stained with silver nitrate (Mortz et al., 2001) or colloidal commassie brilliant blue. The gels were scanned on calibrated was scanning densitometer (Biorad,GS 800) and analysis carried out using PD Quest image analysis software (Biorad). One or two representative gels were used to create a match-set. Spots were detected and matched automatically to a master gel selected by the software. The wherever required spot detection and matching were edited manually. The spot boundary tool was applied to detect large spots. The patterns in sections of the gels in appropriate magnification were checked and spots were added manually to the master gel to allow matching unique spots present in the individual gels. The spot quantity table containing all matched spots was generated. The gel images were normalized according to the total quantity in the analysis set. The differential spots between tumor and tumoradjacent normal were tagged alphabetically after autodetection and matching by the software.

### MALDI-TOF analysis and database search

Differential protein spots were manually excised from the gels and were sent to Centre for Cellular and Molecular platforms (CCAMP), National Centre for Biological Sciences, India for protein characterization using Mass Spectrometry. Briefly, manually excised spots were subjected to in-gel trypsination (Promega) as described (Shevchenko et al., 1996). Peptides were desalted as described (Rappsilber et al., 2003) and mixed with a-cyano-4-hydroxycinnamic acid (CHCA) prior to MALD-TOF MS (Ultraflex III, Bruker). MS and MS/MS data were used in subsequent searches by using the MSDB protein sequence database for human proteins.

#### Gene expression signature and pathway analysis

All predicted proteins were analysed for their biological process and pathway annotations via Panther Gene analysis (http://www.pantherdb.org/). Interacting partners of selected proteins and their co-expressed genes were analysed by using Cytoscape and Genevestigator. *In-silico* expression profile of genes were analyzed for various neoplasm datasets retrieving the expression values from Agilent Humone whole genome array database from Genevestigator response viewer (https://www. genevestigator.com/gv/).

# Results

# Two-dimensional gel electrophoresis of lung adenocarcinoma and matched controls

Two dimensional gel electrophoresis was used to analyze the differential proteome in lung adenocarcinoma and adjacent normal samples. Two dimensional gel electrophoresis was repeated three times for the proteins obtained from tumor and normal tissues from the same patient respectively to validate the reproducibility. The image analysis showed that these patterns were reproducible. Figure 1 depicts the representative two dimensional gel electrophoresis profiles of lung adenocarcinoma and adjacent normal tissue respectively. An average of 600 spots was detected by the PD Quest image analysis software. The comparative spot analysis of tumor and adjacent normal showed 13 spots of differential expression, among which 5 spots were found to be upregulated in tumor, 6 were tumor specific and 2 were specific to normal tissue (Figure 1C). Differentially expressed proteins were defined as significant only if the intensity alterations were greater than 1.5 fold and showed recurrence more than two times in the thirteen pairs of tissue samples examined. Accordingly these proteins were selected and identified from 2D PAGE Swiss-Prot repository of A549 cell line based on their molecular weight and isoelectric point. Data from these



Figure 1. Representative Two-dimensional Gel Electrophoresis Profile of Lung Adenocarcinoma (A) Tumor (B) Adjacent Normal Tissue. Proteins were separated on 4 to 7 non-linear IPG strips and 10% SDS-PAGE gels. Protein spots were visualized by comassie blue staining and analyzed by densitometer. Protein Spots annotated U are present in both tumor and adjacent normal samples but up-regulated in tumor tissue, while as spots marked marked T and N are only present in tumor and normal samples respectively. (C) Closeup view of differentially expressed protein spots in tumor and adjacent normal samples



Figure 2. Representative Two-dimensional Gel Electrophoresis Profile of Lung Adenocarcinoma Cell Line A549. (A) Spot marked A1 represent the protein spot present in A549 and also found to be up- regulated in the tumor sample (U2) (B) Close-up view of the matched spot in A549 and lung adenocarcinoma samples. Proteins were separated on 4 to 7 non-linear IPG strips and 10% SDS-PAGE gels. Protein spots were visualized by comassie blue staining and analyzed by densitometer

identified proteins is listed in the Table I. These proteins were found to be mainly cancer related pathway regulators e.g., pentose phosphate pathway, integrin signalling and pyrimidine metabolism.

#### DOI:http://dx.doi.org/10.7314/APJCP.2015.16.9.3691 Proteomic profiling of human lung adenocarcinoma

The spots that showed upregulation in tumor samples include Tumor rejection antigen (Gp96) 1(U1), Armidillo repeat-containing protein 8 (ARMC8), Cytochrome b-c1 complex subunit 1 (U3), mitochondrial, Protein arginine N-methyltransferase 1 (U4), Reticulocalbin-1 (U5). The tumor specific proteins were Heat shock protein HSP 90-alpha (T1), Alpha-actinin-4 (T2), Dihydropyrimidinase-related protein 2 (T3). Transaldolase (T4), Proteasome subunit alpha type-1 (T5), Oncogene DJ1 (T6). Two proteins Peroxiredoxin-6 (N1), Plateletactivating factor acetylhydrolase IB subunit gamma (N2) were specific to normal samples (Table 1).

# Two-dimensional gel electrophoresis of model lung adenocarcinoma cell line A549

Purified cell lysates from cultured human lung adenocarcinoma cell line (A549) Model to develop a comparative proteome profile. Figure 2 depicts the representative two dimensional gel electrophoresis profile of human lung adenocarcinoma cancer cell line A549. The two dimensional electrophoresis of the proteins from A549 was repeated three times to ensure reproducibility. The representative two dimensional gel electrophoresis pattern is shown in the Figure 2. When two dimensional gel electrophoresis profiles of A549, cell line and lung adenocarcinoma tissue samples were analyzed by the PD quest software, a match of 24 spots was obtained. Among these, only one spot (A1) of molecular weight ~80 kDa and pI of ~4.8 respectively coincided with a differentially expressed spot (U2) found up-regulated in the tumor

## (A) MATRIX Mascot Search Results



Figure 3. Identification of Armadillo repeat- containing protein 8 (ARMC8) in Lung adenocarcinoma by MALDI -TOF MS. (A) Mascot search results of protein spot A1 from A549 cell line and (B) Protein spot U2 from human lung adenocarcinoma tissue both identified as ARMC8. MOWSE score is 66 (> 56, P<0.05) for A1 and 56((> 56, P<0.05) for U2 respectively

Asif Amin et al	
Table 1. Identification of Protein Spots Found to be Differentially E	spressed

Spot No.	Experimental MW (kDa)/PI	Protein Description	Accession No. (NCBI)	Expression Status
U1	97/4.7	Tumor rejection antigen (Gp96) 1	CAI64497	Upregulated in Tumor
U2	80/4.8	Armadillo repeat-containing protein 8; ARMC8	Q8IUR7.2	Upregulated in Tumor
U3	52/5.93	Cytochrome b-c1 complex subunit 1, mitochondrial	NP_003356	Upregulated in Tumor
U4	43/5.63	Protein arginine N-methyltransferase 1; PRMT1	NP_001527	Upregulated in Tumor
U5	39/4.94	Reticulocalbin-1; RCN1	NP_002892	Upregulated in Tumor
T1	83/4.5	Heat shock protein HSP 90-alpha; HSP90AA1	NP_005339	Tumor specific
T2	110/4.85	Alpha-actinin-4; ACTN4	NP_004915.2	Tumor specific
T3	62/5.87	Dihydropyrimidinase-related protein 2; DPYSL2	NP_001377.1	Tumor specific
T4	38/6.43	Transaldolase; TALDO1	NP_006746.1	Tumor specific
T5	29/6.37	Proteasome subunit alpha type-1; PSMA1	NP_002777.1	Tumor specific
T6	20/6.45	Oncogene DJ1; PARK7	Q99497.2	Tumor specific
N1	25/6.11	Peroxiredoxin-6; PRDX6	NP_004896.1	Normal specific
N2	25/6.25	Platelet-activating factor acetylhydrolase IB subunit gamma; PAFAH1B3	NP_001139412.1	Normal specific
A1	80/4.8	Armadillo repeat-containing protein 8; ARMC8	Q8IUR7.2	A549 protein coinciding with spot U2



Figure 4. Classification of the Differentially Expressed Proteins as Assessed by (A) Gene Ontology Search and (B) Biological Function



**Figure 5. Protein-Protein Physical/Functional interaction.** (A) All the predicted query proteins indicated by black circles directly interacts with energy regulatory switch. (B) PARMC8 lies on the signal axis responsible for regulating



(B) Dataset: 746 neoplasm categories (sample selection: HS-SAMPLES-0)



Figure 6. Differential Expression of ARMC8 Gene: (A) Heatmap analysis of ARMC8 across the different anatomical features showing the differential expression. (B) Across the neoplasms of organs

samples. This protein was found to be Armadillo repeatcontaining protein (ARMC8) as revealed by MALDI-TOF MS analysis (Figure 3).

#### Protein identification by MALDI-TOF MS analysis

Protein spot was cut manually and subjected to trypsin digestion as discussed in materials and methods. The peptide mass fingerprinting map was obtained successfully by analysis of MALDI-TOF MS after digesting the chosen protein spot. Finally the protein identification was made by correlating the spectra with the entries in the Swiss-Prot



Figure 7. Perturbations Employed to Assess the Role of ARMC8 Gene in a Specific Cellular Function Depict its Role in Colorectal Cancers at Significant Log<sup>2</sup> Ratio, Indicated by Arrows

(536789 sequences; 190518892 residues) using Mascot search engine. These spots were excised and subjected to in-gel tryptic digestion and the resulting peptides were analyzed by MALDI-TOF-MS. The acquired spectra were processed and searched against a non redundant SwissProt protein sequence database using the Mascot and other search engine (Figure 3). Finally the protein was identified as Armadillo repeat-containing protein (ARMC8) by correlating the spectra with the entries in the Swiss-Prot (536789 sequences; 190518892 residues) using the Mascot search engine.

#### Discussion

There are various issues that need to be considered when employing tissue proteomics for biomarker discovery. Solid tumors contain a plethora of stromal cell populations besides malignant cells including mesenchymal supporting cells like fibroblasts and adipocytes, cells of the vasculature, and a variety of immune infiltrates. The proteomic profiles obtained may therefore reflect the protein features from malignant cells, accessory stromal cells and even the proteins from blood plasma that remains present in tumor associated vasculature. Such heterogeneity thus limits the identification of cancer cell specific biomarkers. Microdissection by isolation of neoplastic cells from histologically identified spots offers a possible solution but its use is limited by the amount of sample processed for subsequent proteomic analysis (Poschmann et al., 2009). Since cancer cell lines accurately represent tumor cells in vivo without the complex in vivo microenvironment, a comparative analysis of clinical tissue samples and cell lines could prove useful in the identification of the proteins dysregulated in the malignant cell populations of tumors.

In the first instance, we performed two-dimensional gel electrophoresis of 13 adenocarcinomas and paired normal tissues samples. A total of 13 proteins were found

to be differentially expressed most of them involved in cancer associated metabolic processes. The differential proteins among lung adenocarcinoma and adjacent normal pairs were identified by searching Swiss-Prot 2D PAGE repository of A549 cell line based on their molecular weight and isoelectric point. The spots that showed upregulation in tumor samples include Tumor rejection antigen (Gp96) 1, Armidillo- repeat containing Protein 8 (ARMC8), Cytochrome b-c1 complex subunit 1, mitochondrial, Protein arginine N-methyltransferase 1, Reticulocalbin-1. The tumor specific proteins were Heat shock protein HSP 90-alpha, Alpha-actinin-4, Dihydropyrimidinase-related protein 2. Transaldolase, Proteasome subunit alpha type-1, Oncogene DJ1. Two proteins Peroxiredoxin-6, Platelet-activating factor acetylhydrolase IB subunit gamma were specific to normal samples.

The functional based cataloguing of all the protein hits predominantly sorted into cancer related pathway regulators e.g., pentose phosphate pathway, integrin signaling pathway and pyrimidine metabolism (Figure 4a). Gene ontology based analysis classified majority of protein hits responsible for metabolic processes (Figure 4b). We observed several clusters obtained through pathway reconstructions that seem to be involved in mechanisms suggestive of neoplastic transformation through the participation of Tumor specific HSP90AA1 (Table I), exhibiting its effect via MARCK5 regulations, responsible for cellular energy extension (Figure 5a). Cooperative functioning of Oncogene DJ1 with TALDO1, evident from their physical interaction, makes another cluster that could set the cellular configuration active for cancer supporting metabolic events (Figure 5a). These proteins define the crucial signal transduction axis corresponding to metabolic regulatory switch via glucose catabolic processes and predominantly through NAPD metabolism as well as regeneration.

A comparison of proteomic profiles established from tissue samples with that of model lung adenocarcinoma cell line A549, identified a match-set of atleast 24 spots, however, only one protein spot corresponding to Armadillo repeat-containing protein (ARMC8) (as identified by MALDI-TOF/MS) showed differential expression. A centralized pathway reconstruction with ARMC8 formed a functional cluster showing a coordinative dependency between the regulators of meiotic division and catenin dependent cellular attachment (Figure 5b), that if deregulated can lead to invasiveness. Genevestigator was utilized as a meta analysis database to examine the expression pattern of ARMC8 in all the neoplasms. All those probes having high probability that correspond to the actual gene of interest were searched for their expression profiles across all the cancers. Heatmap based visualized expression of ARMC8 in various human cancers was constructed based on Human 47K-microarray data. It was observed that ARMC8 was highly expressed in Chronic Myelogenous Leukaemia though with a slight variability based on the origin of tumor mostly representing blood malignancies (Figure 6). Heatmaps generated from perturbation data sets for ARMC8 selective probe reflected the net output result with

#### Asif Amin et al

a log2 ratio of 1.57 for colorectal cancer with an increase in expression by 3.69 folds. However ARMC8 has not yet been reported with significant score in lung pathologies (Figure 7). Thus these observations indicate the adapter proteins sassociated with ARMC8 play key roles in the control of proliferation and differentiation. These associations are suggestive of interconnected pathways that are responsible for deregulation and alterations in the signal procedures that directly or indirectly may expedite the onset of oncogenic signals in lung adenocarcinomas

## Acknowledgements

Funding support from Department of Science and Technology, Ministry of Science and Technology, Govt. of India Grant No.: (SR/SO/HS-90/2009 dated 24/09/2010) is highly acknowledged.

# References

- Anderson L, Seilhamer J (1997). A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis*, **18**, 533-7.
- Bechtel JJ, Petty TL, Saccomanno G (2000). Five year survival and later outcome of patients with X-ray occult lung cancer detected by sputum cytology. *Lung Cancer*, **30**, 1-7.
- Bouchardy C, Benhamou S, Schaffar R, et al (2011). Lung cancer mortality risk among breast cancer patients treated with anti-estrogens. *Cancer*, **117**, 1288-95.
- Darby S, Hill D, Deo H, et al (2006). Residential radon and lung cancer--detailed results of a collaborative analysis of individual data on 7148 persons with lung cancer and 14,208 persons without lung cancer from 13 epidemiologic studies in Europe. Scand J Work Environ Health, **32**, 1-83.
- Ferlay J, Shin HR, Bray F, et al (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*, 127, 2893-917.
- Gabrielson E (2006). Worldwide trends in lung cancer pathology. *Respirology*, **11**, 533-8.
- Ganti AK, Sahmoun AE, Panwalkar AW, et al (2006). Hormone replacement therapy is associated with decreased survival in women with lung cancer. *J Clin Oncol*, **24**, 59-63.
- Herbst RS, Heymach JV, Lippman SM (2008). Lung cancer N. Engl J Med, **359**, 1367-80.
- Jemal A, Bray F, Center MM, et al (2010). Global cancer statistics. CA Cancer J Clin, 61, 69-90.
- Kadara H, Kabbout M, Wistuba II (2011). Pulmonary adenocarcinoma: A renewed entity in 2011. *Respirology*, 17, 50-65.
- Luqman M, Javed MM, Daud S, et al (2014). Risk factors for lung cancer in the Pakistani population. Asian Pac J Cancer Prev, 15, 3035-9.
- Mortz E, Krogh TN, Vorum H, et al (2001). Improved silver staining protocols for high sensitivity protein identification using matrix-assisted laser desorption/ionization-time of flight analysis. *Proteomics*, 1, 1359-63.
- Poschmann G, Sitek B, Sipos B, et al (2009). Identification of proteomic differences between squamous cell carcinoma of the lung and bronchial epithelium. *Mol Cell Proteomics*, 8, 1105-16.
- Rappsilber J, Ishihama Y, Mann M (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem*, **75**, 663-70.
- Shevchenko A, Wilm M, Vorm O, et al (1996). Mass

spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem*, **68**, 850-8.

- Siegel R, Naishadham D, Jemal A (2012). Cancer statistics, 2012. CA Cancer J Clin, **62**, 10-29.
- Simsek C, Sonmez O, Keyf AI, et al (2013). Importance of serum SELDI-TOF-MS analysis in the diagnosis of early lung cancer. *Asian Pac J Cancer Prev*, **14**, 2037-42.
- Sun DS, Hu LK, Cai Y, et al (2014). A systematic review of risk factors for brain metastases and value of prophylactic cranial irradiation in non-small cell lung cancer. *Asian Pac J Cancer Prev*, **15**, 1233-39.
- Wani MA, Jan FA, Khan NA, et al (2014). Cancer trends i**100.0** Kashmir; common types, site incidence and demographic profiles: national cancer registry 2000-2012. *Indian J Cancer*, **51**, 133-7.
- Youlden DR, Cramb SM, Baade PD (2008). The international 75.0 epidemiology of lung cancer: geographical distribution and secular trends. J Thorac Oncol, 3, 819-31.

50.0

25.0

0

6.3

56.3

31.3