IDENTIFICATION OF DIFFERENTIALLY EXPRESSED SECRETORY PROTEINS IN LUNG CANCER

Porangana Nayar 101553

Under the supervision of

Dr. Aklank Jain

May - 2014

Submitted in fulfillment of the degree of

4 year Degree Programme B.Tech

DEPARTMENT OF BIOTECHNOLOGY

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT

TABLE OF CONTENTS

CERTIFICATE

This is to certify that the work titled "**Identification of differentially expressed secretory proteins in lung cancer**" submitted by "**Porangana Nayar"** in fulfillment for the award of degree of 4 Year B.Tech Degree Program of Jaypee University of Information Technology, Waknaghat has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Dr. Aklank Jain

Associate Professor (Biotechnology)

ACKNOWLEDGEMENT

I take this opportunity to express my profound gratitude and deep regards to my guide **Dr. Aklank Jain** for his exemplary guidance ,monitoring and constant encouragement throughout the course of this thesis. The blessing, help and guidance given by him time to time and the support of my parents shall carry me a long way on which I am about to embark.

I owe my special debt of gratitude and heart full thanks to Mr. Ankit Srivastava and Ms. Akanksha Khandelwal for devoting their time and guiding me with their valuable suggestions and for bearing my mistakes and shortcomings.

Porangana Nayar

SUMMARY

Cancer is one such disease which has enormously spread like wild fire in the community in the past decade. This disease has quite an interesting irony related to it, the irony being that even the smallest mutation in the gene can be fatal. The randomness and a-symptomatic nature of this disease is what has made it such a hot topic of research. There is a saying "Prevention is better than Cure", but when it comes to cancer, more often there is neither a method of prevention nor of a cure and that is the reason that there are innumerable studies being conducted across the world to identify and determine various methods to prevent and cure this disease. Using this as my driving force, I began working on the research project titled – "Identification of Differentially Expressed Proteins in Lung Cancer." The aim of this project is to be able to come up with a set of proteins which are uniquely or differentially expressed in the patient samples as compared to healthy individuals. Of all the studies that have been conducted, there is not even a single protein which has been identified to show relevance to cancer in all populations and the reason behind this is twofold. Firstly, is the genetic differences from population to population and secondly, the environment is also a major factor. Culminating these genetic differences and the varying environments, the protein profile of the individual differs and therefore it has not yet been possible to find a common protein with relevance to lung cancer. Therefore, I am looking forward to identify a set of proteins which work well for the North Indian population and that is the gist of the focus of my report.

Porangana Nayar **Dr. Aklank Jain.**

List of Figures and Tables

Figures:

Tables:

Chapter 1

Introduction

The speed at which the field of medicine is advancing is exponential, whether it is the diagnostic imaging technology, surgical management or therapeutic strategies. Despite this fact, the ironical part is Cancer; one of the most deadly diseases continues to be undefeated. The validation of the alarming spread of this disease is clearly reflected in the latest world cancer report published by the International Agency for Research on Cancer (IARC), which is a part of the World Health Organization. According to IARC cancer report; in 2012, the worldwide burden of cancer rose to an estimated 14 million new cases per year, a figure expected to rise to 22 million annually within the next two decades. Over the same period, cancer deaths are predicted to rise from an estimated 8.2 million annually to 13 million per year. Globally, in 2012 the most common cancers diagnosed were those of the lung (1.8 million cases, 13.0% of the total), breast (1.7 million, 11.9%), and large bowel (1.4 million, 9.7%). The most common causes of cancer death were cancers of the lung (1.6 million, 19.4% of the total), liver (0.8 million, 9.1%), and stomach (0.7 million, 8.8%). And, if we see in the Indian scenario, lung cancer has an extremely high fatality rate of ~9.3 % of all cancer (1). These figures coherently point towards the fact that lung cancer cases are escalating by the day and spreading like grapevines into the society and therefore it is the need of the hour to come up with various diagnostic and prognostic tools for the early detection and prevention of this disease before it creates havoc.

To combat such a scenario where cancer seems to be on a rise, there is an urgent need of techniques and more so of certain markers which can help us detect and diagnose the cancer at earlier stages and then develop an appropriate treatment and prognosis. This can be done by studying the protein profile of an individual and try and associate the levels of presence or absence of a particular protein(s) with the disease.

This way one can aim to approach the cancer at earlier stages. This approach may sound very easy but has lot of underlying problems associated with it. One of the major problems is in respect to the immense genetic variations in populations across the globe, due to which no specified set of proteins have been claimed to be unanimously present in cancer patients(of a specific type) for the entire population. Along with the genetic variation, the environmental conditions also play a major contribution towards the protein profile of an individual. Therefore, due to this highly variable combination of genetic makeup and environment, there has been no success in developing a common marker for cancer.

Chapter 2

Review of literature

Lung cancer is the uncontrolled growth of abnormal cells in one or both lungs. These abnormal cells do not carry out the functions of normal lung cells and do not develop into healthy lung tissue. As they grow, the abnormal cells can form tumors and interfere with the functioning of the lung, which provides oxygen to the body via the blood. Cancer begins with a single mutation in the DNA and subsequently a series of mutations creates a lung cancer cell. Before becoming fully cancerous, cells can be precancerous, in that they have some mutations but still function normally as lung cells. With each new mutation, the lung tissue cell becomes more mutated and may not be as effective in carrying out its function as a lung cell. At a later stage of disease, some cells may travel away from the original tumor and start growing in other parts of the body. This process is call metastasis and the new distant sites are referred to as metastases.

Primary lung cancer originates in the lung, i.e. the lung is the primary organ. When cancer cells from other primary organs travel to the lung, this is then known as secondary lung cancer.

Signs and symptoms

In case of lung cancer, any sort of unusual feelings need to be diagnosed immediately, although many times these unusual feelings are attributed to other causes, like Bronchitis. As most of the other cancers, the signs and symptoms of lung cancer rarely show up in the earlier stages and in most of the cases can take years to develop and they may not even appear until the disease is advanced.

Symptoms of lung cancer that are predominant in the chest majorly include: coughing(persistent and intense), pain in the chest, change in color or volume of sputum, shortness of breath, changes in voice or hoarsening of voice, stridor (harsh sounds with

each breath, recurrent lung problems like bronchitis or pneumonia, coughing up phlegm or mucus and coughing up blood.

If the lung cancer has spread to other places in the body; those commonly being : lymph nodes, bones, brain, liver and adrenal glands; the symptoms may be very different which are: loss of appetite or unexplained weight loss, muscle wasting (cachexia), fatigue, headaches, bone or joint pain, bone fractures unrelated to accidental injuries, neurological symptoms like memory loss or unsteady gait, neck or facial swelling, general weakness, bleeding and blood clots.

Causes and risk factors

Since the lungs are the core of our respiratory system, essentially everything that we breathe has a potential risk. One of the most important risk factor of lung cancer is tobacco smoking. Not only active, but passive smoking also is to a large extent responsible for lung cancer development. Apart from smoking the environment is also a reservoir of many substances and pollutants which have an adverse effect on the lung cells. Some such substances are: asbestos, radon gas, industrial substances including arsenic, uranium, beryllium, vinyl chloride, nickel chromates, coal products, mustard gas, chloromethyl ethers, gasoline and diesel exhaust, radiation exposure such as X rays, air pollutants, tuberculosis and genetics.

Diagnosis

If lung cancer is suspected as a result of a screening procedure, a small piece of tissue from the lung must be examined under a microscope to look for cancer cells. This procedure is called a biopsy and can be performed in two different ways, one way is that the doctor passes a needle through the skin into the lungs to remove a small piece of tissue; this is known as a needle biopsy. The second way is doing a biopsy during bronchoscopy. It is done by inserting a small tube which has a light at the end through the mouth or nose and into the lungs. This light allows the doctor to see inside the lungs and subsequently extract a small tissue sample.

Other techniques which can be used for the purpose of diagnosis include: imaging tests such as X rays, MRIs, ultrasounds, PET scans, CT scans, sputum cytology, thoracoscopy, mediastinoscopy and thoracotomy.

Types and staging

Lung cancer is of two major types, non **small cell lung cancer (NSCLC)** and **small cell lung cancer (SCLC).** Staging of cancer is based on the whether the cancer is local or has spread from the lungs to the nearby lymph nodes and other organs. Since the lungs are large, the tumors can grow in them for a long time before they are found. Due to this reason early- stage lung cancer (stages I and II) is difficult to detect and is therefore diagnosed in the later stages (stages III and IV).

NSCLC accounts for approximately 85% of lung cancers. Among them there are these types of tumors: adenocarcinoma, squamous cell carcinoma and large cell carcinoma. In order to come up with an effective treatment strategy for curing the cancer, the physician needs to be aware of the stage and the extent of the cancer, this is the reason that almost all cancers including lung cancer are categorically segregated in to different stages. NSCLC has been divided into four stages:

- Stage I - the cancer is located only in one of the lungs and has not yet spread to the any lymph nodes
- Stage II- the cancer is in the lung and is now present in the nearby lymph nodes.
- Stage III – cancer is found in the lung and also in the lymph nodes in the middle of the chest (locally advanced disease). It has further two subtypes:
    - If cancer has spread only to the lymph nodes on the same side of the chest where the cancer started, it is known as stage IIIA.
    - If the cancer has spread to the lymph nodes on the opposite side of the chest, or above the collar bone, it is known as stage IIIB.
- Stage IV – it is the most advanced stage of lung cancer. It is when the cancer has spread to both the lungs to the fluid in the area surrounding the lungs and to other parts of the body such as the liver and bones.

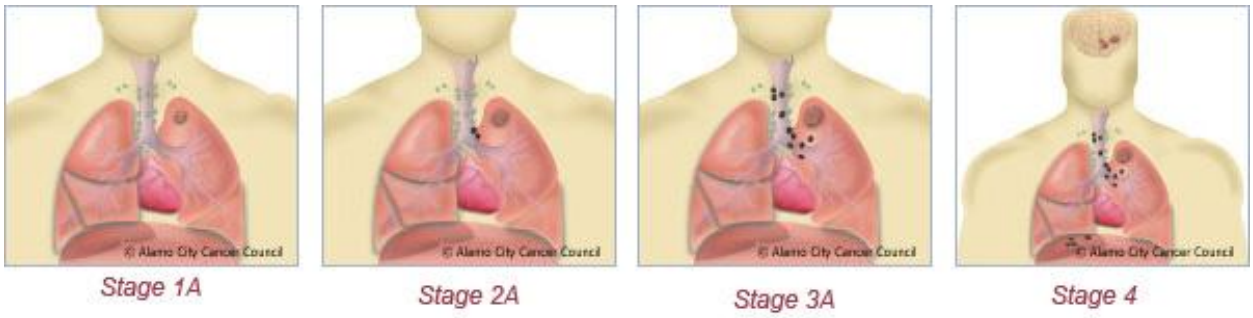


Fig. 1

SCLC accounts for the residual 15% of lung cancers. It results from smoking and grows more rapidly and spreads to other parts of the bodies faster as compared to NSCLC. It is also more responsive to chemotherapy. It is divided broadly into two major stages, namely:

- Limited stage – cancer is found on one side of the chest, involving just one part of the lung and nearby lymph nodes.
- Extensive stage – cancer has spread to other regions of the chest or other parts of the body.

Prevention

Taking precautions is an action to lower the chance of getting cancer. Half the battle is won if careful precautions are taken by people towards protecting themselves from the cancer and in turn also dramatically lower the number of cancer related deaths. As discussed above preventive measures are taken against the possible known risk factors, but the problem lies in the fact that some risk factors may be avoided, while others cannot. For instance, inheriting certain genes cannot be avoided whereas getting rid of habits like smoking is something we can easily avoid. In general, different ways to prevent lung cancer are: changing lifestyle and eating habits, regular exercise, refrain from use of tobacco and smoking.

Screening

Unlike mammography for breast cancer or colonoscopy for colon cancer, a widely accepted screening tool for early- stage lung cancer has not been available till recently.

Regular chest X rays are not a very reliable source in order to detect lung tumors at the early stages. Recent guidelines from ASCO (American Society of Clinical Oncologists) suggests that annual screening with low- dose computed tomography (LDCT) or smokers and former smokers at high risk for developing lung cancer. High risk factors include being between the age of 55 to 74, having smoked for 30 years or more, and either continuing to smoke or having quit within the past 15 years. At this time, yearly screening with LDCT is recommended for high-risk individuals after careful discussion with their physicians.

Treatment

NSCLC

Surgery, radiation, chemotherapy, and targeted treatments—alone or in combination—are used to treat lung cancer. Each of these types of treatments may cause different side effects.

Surgery – Stage I and stage II NSCLC are treated with surgery to remove the tumor. This is done by removing the lobe, or section of the lung containing the tumor. The rate of success of surgeries at this stage is very high as at these stages the cancer is still localized and has been removed before spreading to other parts of the body. Video assisted thoracoscopic surgery (VATS) is also used for this purpose. This is done by making a small incision in the chest and inserting a thoracoscope, which at its end has a light and a tiny camera connected to a video monitor which makes it possible for the surgeon to see inside the chest. This way the infected lung lobe can be removed through the scope without making a large incision in the chest.

Chemotherapy and Radiation – for the patients with NSCLC, who have undergone successful surgery, chemotherapy given after surgery, known as "adjuvant chemotherapy", may help prevent the cancer from returning. This fact holds true for patients with stage II and IIIA disease. For people with stage III lung cancer, who cannot undergo surgery, chemotherapy in combination with definitive (high dose) radiation treatments. In stage IV lung cancer, chemotherapy is the main typical treatment. The chemotherapy treatment plan usually consists of a combination of drugs such as cisplatin

(Platinol) or carboplatin (Paraplatin) plus docetaxel (Taxotere), gemcitabine (Gemzar), paclitaxel (Taxol), vinorelbine (Navelbine) or pemetrexed (Alimta).

Neoadjuvant Chemotherapy – in this mode of treatment, chemotherapy is administered before radiation or surgery, so that the drugs help reduce the size of the tumor so that it shrinks to size small enough to make it easier for removal during surgery, increasing the effectiveness of radiation and destroying hidden cancer cells at the earliest possible time.

Targeted treatments – unlike chemotherapy drugs, which cannot differentiate between normal cells and cancer cells, targeted therapies are designed specifically to attack cancer cells by attaching to or blocking targets that appear on the surfaces of those cells. People who have advanced lung cancer with certain molecular biomarkers may receive treatment with a targeted drug alone or in combination with chemotherapy. These treatments for lung cancer include:

Erlotinib (Tarceva) – it is a targeted treatment which has shown to benefit some people with NSCLC. This drug blocks a specific kind of receptor on the cell surface—the epidermal growth factor receptor (EGFR). Receptors such as EGFR act as doorways by allowing substances in that they can encourage a cancer cell to grow and spread. Lung cancer cells that have a mutation on the EGFR are likely to respond to treatment with erlotinib instead of chemotherapy.

Bevacizumab (Avastin) – like other tissues of the body, tumors too need a blood supply for their growth and survival. This drug works by stopping VEGF (Vascular endothelial growth factor), which stimulates the growth of blood vessels. When combined with chemotherapy, bevacizumab has been shown to improve survival in people with certain types of non-small lung cancer, such as adenocarcinoma and large cell carcinoma.

Crizotinib (Xalkori) – is a newly available treatment which works on people with advanced NSCLC who have the ALK biomarker. Mutations in the way cells program ALK result in changes to the way it functions, leading to increased tumor cell growth. Crizotinib works by blocking ALK and stopping the growth of the tumor.

SCLC

Chemotherapy and Radiation therapy – for patients with SCLC regardless of stage, chemotherapy is an essential part of treatment. Radiation treatment may be used as well depending on the stage of cancer. For people with limited-stage lung cancer, a combination of chemotherapy plus radiation therapy given at the same time is the recommended treatment. The most commonly used initial chemotherapy regimen is etoposide (Toposar or Vepesid) plus cisplatin (Platinol), known as EP. For people with extensive-stage small cell lung cancer, chemotherapy alone using the EP regimen is the standard treatment. However, another regimen that may be used is carboplatin (Paraplatin) plus irinotecan (Camptosar).

Preventive Radiation Therapy to the Brain - In more than half of the people with small cell lung cancer, the cancer also spreads to the brain. For people whose lung cancer has responded to chemotherapy, doctors may prescribe radiation therapy to the brain to help prevent the cancer from spreading to the brain. This procedure is known as prophylactic cranial irradiation (PCI). This can benefit patient with both limited-stage and extensive-stage small cell lung cancers.

Surgery - A very small percentage of people who have limited-stage small cell lung cancer and no lymph node tumors may benefit from surgery, after which adjuvant chemotherapy is given.

With such intense treatment procedures along with chemotherapeutic drugs and radiation, the patient's body is bound to experience certain side effects. Therefore it is of extreme importance that these side effects are monitored so that they not attain the level of lethality. The body's reaction to chemotherapy, radiation, or targeted treatments depends on a number of factors such as length of treatment, dosage prescribed, and a person's health history. The following are the most frequent and common side effects that are experienced by patients:

- Blood clots
- Bone issues
- Chemobrain (problems with memory and concentration)

- Dental issues (tooth decay, dry mouth and mouth sores)
- Diarrhea
- Fatigue
- Hair loss
- Lymphedema (painful swelling that happens in the lymph node when your body's lymphatic fluid is unable to circulate properly and builds up in your soft tissues instead.)
- Nausea and vomiting
- Neuropathy (sense of tingling or numbness in hands and feet)
- Pain
- Rash
- Weight loss or gain

Biomarkers of Lung Cancer

A big advancement in the field of cancer is the identification of certain proteins or molecules that give doctors a more detailed picture of the tumor. These markers are called biomarkers because they are found on the surface of the cells, or in the genes that program cells. Some biomarkers include:

- Epidermal growth factor receptor (EGFR)
- Anaplastic lymphoma kinase (ALK)
- K-ras mutations (KRAS 1)

When a person is diagnosed with lung cancer, doctors may perform tests to determine if one of these biomarkers is present in the tumor. These biomarkers help the doctor decide which treatment options will work the best. These individualized treatments may help patients avoid treatments that are unlikely to work and help to decrease the side effects of treatment

Table 7

Twelve biomarker classifier proteins[§].

Biomarker	UniProt ID	Direction*	Description
Cadherin-1	P12830	down	cell adhesion, transcription regulation
CD30 Ligand	P32971	up	cytokine
Endostatin	P39060	up	inhibition of angiogenesis
HSP 90α	P07900	up	chaperone
LRIG3	Q6UXM1	down	protein binding, tumor suppressor
MIP-4	P55774	up	monokine
Pleiotrophin	P21246	up	growth factor
PRKCI	P41743	up	serine/threonine protein kinase, oncogene
RGM-C	Q6ZVN8	down	iron metabolism
SCF sR	P10721	down	decoy receptor
sL-Selectin	P14151	down	cell adhesion
YES	P07947	up	tyrosine kinase, oncogene

[§]Up or down regulation in NSCLC cases relative to controls.

Fig 2.

Prognosis

Unfortunately, compared with some other types of cancer, the outlook for lung cancer is not very good. Overall, of all types of lung cancer, about 32 out of every 100 people (32%) will live for at least 1 year after they are diagnosed. Around 10 out of every 100 people (10%) will live for at least 5 years. And about 5 out of every 100 people (5%) will live for at least 10 years. As with many other types of cancer, the outcome depends on how advanced your cancer is when it is diagnosed. Lung cancer is one of the most difficult cancers to treat. It is often diagnosed at a late stage. It tends to occur in older people who may also have other medical conditions.

NSCLC

Of all the people with stage 1A non small cell lung cancer, between 58 and 73 people out of every 100 (58% to 73%) will live for at least 5 years.Of all the people with stage 1B non small cell lung cancer, between 43 and 58 people out of every 100 (43% to 58%) will live for at least 5 years. For stage 2A lung cancer, between 36 and 46 out of every 100 people diagnosed (36% to 46%) will live for at least 5 years with treatment. For

stage 2B non small cell lung cancer, between 25 and 36 out of every 100 people diagnosed (25% to 36%) will live for at least 5 years. For stage 3A non small cell lung cancer, between 19 and 24 out of every 100 people diagnosed (19% to 24%) will live for at least 5 years. For stage 3B, between 7 and 9 out of every 100 people diagnosed (7% to 9%) will live for at least 5 years. Stage 4 is the most advanced stage, where the cancer has spread. Understandably, the survival statistics are very low for this stage. Unfortunately, lung cancer is often diagnosed late and for many people the cancer has already spread when they are diagnosed. Only between 2 and 13 out of every 100 people diagnosed with stage 4 non small cell lung cancer (2% to 13%) will live for at least 5 years. It can seem illogical for stage 3B cancer to have 5 year survival rates from 7% to 9% and stage 4 from 2% to 13%. This is because the staging system only looks at the extent of the cancer. It does not look at the specific types of cancer. So the stage 4 group may include more people who have slowly growing cancers or cancer that responds very well to particular treatments than the stage 3 group.

SCLC

Of all the people diagnosed with small cell lung cancer, around 1 in 3 have limited disease at the time of diagnosis. With treatment about 25 out of every 100 people (25%) will live for at least 2 years. 2 out of 3 people with small cell lung cancer already have extensive disease at the time of diagnosis. Unfortunately the survival rate is very low. With treatment, fewer than 5 out of every 100 people (5%) will live for at least 5 years.

Secretory Proteins

Serum or plasma based assays are generally the most popular choice for clinical screening and diagnosis and one of the major reasons behind is that obtaining blood samples is simple and non invasive as compared to other sources of samples. The blood circulation extends throughout the entire human body and therefore acts as a repository

of signals released from any tissue in the body. Since there is a flip side of every coin, one of the major problem associated with secretory proteins is that, the presence of high abundance proteins like albumin, haptaglobin, transferrins and immunoglobulins hinders the detection of the tumor specific biomarkers, which unlike the former, are found in scarce concentrations.

"Secretome", was coined by Tjalsma and colleagues in order to describe the total proteins that are released by a cell, tissue or organism. It is known as a fact that the Secretome constitutes about 10- 15 % of the entire pool of proteins encoded by the human genome and have been identified for playing key roles such as immune defense, blood coagulation, matrix remodeling and cell signaling. Studies have revealed of two broad mechanisms via which proteins are secreted into the extracellular space: the classical secretory pathway and the non- classical secretory pathway. In the former pathway, proteins targeted for extracellular release are synthesized as protein precursors which most often contain signal peptides located at the N terminus. The presence of these signal peptides direct the proteins to the rough endoplasmic reticulum and then further to the Golgi apparatus, after which the proteins are released into the extracellular environment in the secretory vesicles. On the other hand, in the non classical pathway proteins are exported by targeting endosomes recycling back to the plasma membrane, directly translocating across the plasma membrane.

In the scenario of cancer, the tumor cells are in constant contact with the extracellular environment and interact with it in order to create favorable conditions for tumor progression, which include angiogenesis or metastasis. All these interactions are mediated by a variety of proteins that are secreted by the tumor cells, including growth factors, adhesion molecules, cytokines, chemokines, shed receptors and proteases. . thus in a nutshell, the cancer Secretome can be described as consisting of proteins that are secreted only by the cancer cells along with those proteins which are secreted by the cancer associated stromal cells. This is the step that has made it evident that secreted proteins are a promising source for the development of serological biomarkers, with the most valid hypothesis that they will eventually enter into the blood circulation.

Proteomics techniques for analyzing secretomes:

Exponential growth and advancement in the field of proteomics technology has paved the way for increasing the numbers of secretory proteins being identified. In addition, the current proteomic approaches not only aid in simple protein identification, but are advanced enough to portray the reliable quantification of protein abundance in any sample at any given point of time. The traditional 2-DE has been used to compare protein abundance by determining the differences in the staining intensity and has therefore been employed in lots of secretome studies of various cancers including lung cancer. However, one of the disadvantages of using 2-DE is that it is capable of detecting only a certain classes of proteins such as those with extreme isoelectric points or molecular weights. Therefore, to negate this fact out, a more comprehensive and thorough analysis of protein mixtures is carried out through shotgun proteomic techniques such as multidimensional protein identification.

Another method that is gaining immense popularity for this purpose is the label- free MS based quantitation method, which has been found to be simpler and faster as compared to its labeling based counterpart. The label- free quantitation method can be broadly categorized into 2 subsets: area under the curve (AUC) and spectral counting. AUC basically involves measuring changes in chromatographic peak area from precursor ion spectra, while on the other hand spectral counting pertains to counting the number of peptides assigned to a protein in an MS/MS experiment.

One of the most commonly used and well received strategies is the shotgun proteomics strategy of one- dimensional gel electrophoresis combined with LC-MS/MS. For this technique, samples are pre- separated by SDS- PAGE so as to reduce the sample complexity and the appropriated gel pieces are sliced and excised from the gel lane. Theses excised gel slices are then subjected to reduction, alkylation and in- gel digestion (mostly with trypsin) before it is subjected to LC-MS/MS analysis. The efficiency of this technique has been validated by a recent comprehensive study of the secretome of 23 human cancer cell lines from 11 cancer types- nasopharyngeal carcinoma, colorectal carcinoma, epidermoid carcinoma, T cell lymphocytoma, breast, bladder, cervical, liver, lung, oral an pancreatic cancer. A total of 4,584 non redundant proteins were identified

with an average of 1300 proteins detected per cell line. Of these, 55.8% of the proteins were predicted to have been secreted, based on the signaling secretomeP and TMHMM prediction. Some of the several promising proteins identified are monocyte differentiation antigen CD14, stromal cell- derived factor 1, cathepsin L1 and interferon induced 17kDa protein were validated as potential serological cancer markers using ELISA.

Recently, there has been a lot of interest in the post translational modification related to secretory proteins. Glycosylation is one of the most prevalent and biologically important post translational modifications in proteins, and an interesting fact is that most of the secretory proteins as well as serum proteins are believed to be glycosylated. The presence of this glycosylation as such helps in targeting the carbohydrate moiety and therefore could be used as an efficient method to enrich for the secretory proteins required and do away with the other intercellular contaminants. Hydrazide chemistry is applied to isolate glycoproteins as a means to selectively capture N-linked glycoproteins in human serum, in this process, the carbohydrate moieties are first oxidized to aldehydes, which then are further covalently bound to Hydrazide groups which are immobilized on resins. Subsequently, proteolytic digestion followed by PNGase F treatment then releases the peptides from the bound carbohydrate groups. This strategy was successfully employed to enrich for glycoproteins from the secretome of breast cancer cells. These glycoproteins may also be captured by hydrophilic affinity interactions, where the hydrophilic glycopeptides would be retained

On the hydrophilic interaction liquid chromatography (HILIC) columns and thus separated from the hydrophobic unglycosylated peptides. Thus, a combination of these two techniques can be efficiently used for the study of a contamination free pool of desired secretory proteins.

In the field of secretomics and lung cancer, Yu and colleagues generated a malignant pleural effusion (MPE) proteome dataset of 482 proteins by removing high abundance proteins from (MPE) fluids, obtained from lung adenocarcinoma patients, prior to GeLC-MS/MS analysis. This MPE proteome dataset was combined with secretome datasets from three different adenocarcinoma cell lines and analyzed to identify

biomarkers of malignancy. Among the 107 potential pleural effusion (PE) biomarkers identified, alpha- 2- HS glycoprotein and insulin like growth factor binding protein 2 levels were shown to be elevated in lung cancer patients with MPE than those with the non malignant pleural effusion.

CHAPTER 3

Materials and Methods

Sample Collection

The lung cancer patient blood samples were collected from Indira Gandhi Medical College (IGMC), Shimla

Protein Isolation

Reagents used:

Table 1: RBC Lysis Buffer

Reagents	Stock	Working Conc.	Vol. for 1000 ml	Vol. for 100 ml
Tris HCl (8.0)	10 mM	1 M	10 ml	1 ml
EDTA	1 mM	0.5 M	2 ml	0.2 ml
NH4Cl (8.0)	125 mM	1 M	125 ml	12.5 ml
Water	-	-	863 ml	86.3 ml

Table 2: Cell Lysis Buffer

Reagents	Stock	100 ml	Final Conc.	Vol. for 10ml
Tris HCl	0.5 M	10 ml	50 mM	1 ml
EDTA	0.5 M	400 µl	2 mM	0.04 ml
EGTA	0.5 M	200 µl	1 mM	0.02 ml
Glycerol	-	10 ml	10%	1 ml
Tween 80	-	100 µl	0.1%	0.01 ml
DTT	0.5 M	200 µl	1 mM	0.02 ml
PMSF	0.1 M	1 ml	1 mM	0.1 ml
PIC	-	200 µl	-	20 µl
NaCl	-	3.522 g	500 mM	0.2992 g

Protocol

1. Take 500 µl of blood sample and add three times RBC lysis buffer.
2. Mix it well by inverting for 10 minutes.
3. Centrifuge at 13,000 rpm for 1 minute to obtain WBC pellet.
4. Add cell lysis buffer (50 µl) according to pellet sixe obtained.
5. Keep on ice for 30 minutes, vortexing after every 10 minutes.
6. Carry out protein estimation by Bradford Assay.

Protein estimation by Bradford assay

1. The proteins of unknown concentrations were added to different wells of the microtitre plate and the volume with distilled/ milliQ water in each well.
2. To each well, 200 µl of Bradford reagent was added using a micropipette.
3. The plate was allowed to stand in an incubator at 37 °C for 30 minutes.

SDS – PAGE

Materials:

Reagents used:

5X Sample Buffer

10 % w/v	SDS
10 mM	beta-mercapto-ethanol
20% w/v	Glycerol
0.2 M	Tris – HCl, pH 6.8
0.05% w/v	Bromophenolblue

1X Running Buffer

25 mM	Tris – HCl
200 mM	Glycine
0.1% w/v	SDS

Resolving Gel

Table 3: Resolving Gel (12%)

	10 ml	20 ml	30 ml	50 ml	75 ml
Water	3.175 ml	6.35 ml	9.525 ml	15.875 ml	24.5 ml
30% Acrylamide	4 ml	8 ml	12ml	20 ml	30 ml
1.5 M Tris (pH 8.8)	2.503 ml	5.06 ml	7.509 ml	12.515 ml	19 ml
10% SDS (w/v)	0.1 ml	200 µl	300 µl	500 µl	750 µl
10% APS (w/v)	0.1 ml	200 µl	300 µl	500 µl	750 µl
TEMED	4 µl	8 µl	12 µl	20 µl	30 µl

Table 4: Stacking Gel (5%)

	15 ml	4 ml	5 ml	10 ml
Water	10.2 ml	2.66 ml	3.4 ml	6.8 m l
30 % Acrylamide	24.9 ml	0.67 ml	0.83 ml	1.66 ml
IM Tris (6.8)	1.89 ml	0.5 ml	0.63 ml	1.26 ml
SDS (10%)	150 µl	0.04 ml	0.05 ml	100 ml
APS (10%)	150 µl	0.04 ml	0.05 ml	100 ml
TEMED	15 µl	0.004 ml	0.005 ml	10 µl

Protocol

1. **Assembling the glass plates:**

   The glass plates were cleaned with detergent (Tween 20) and water then with ethanol. They were then assembled with spacers. For the mini- gel, the glass plate assembly was attached to the apparatus and the whole assembly was then placed in a small, flat Tupperware container.

2. **Casting of the gel:**

   The gels were cast beginning with the resolving gel followed by the stacking gel. The degassed resolving gel was transferred to a beaker followed by addition of APS/TEMED and then mixed. Quickly the resolving gel solution was added with the help of pipette to the centre of the plates to a height of 4 cm from the top of the large plate. Isobutanol is quickly added to the top till the plate's level and allowed to stand till the resolving gel polymerized. This prevents the gel's contact with air/oxygen that will react with the free radicals' generated by APS in the process and delay polymerization.

   The stacking gel reagents were added to the flask and degassed for 10 minutes. Meanwhile, the iso- butanol was poured off and water was used several times to rinse it within the gel plate space. The space was then air dried in an incubator at 37 °C. To polymerize the stacking gel, APS/TEMED was added and mixed followed by pouring on top of the polymerized resolving gel. The comb was inserted straight down into the stacking gel while it was allowed to polymerize undisturbed for 20 – 30 minutes. Any bubbles in the gel must be removed carefully.

3. **Loading the gel:**

   Tris – Glycine electrophoresis buffer was poured into the upper and lower any unpolymerized polyacrylamide that may have seeped in. when loading the gel,

1X loading buffer was loaded in every lane so that the dye front will migrate more evenly.

4. Sample preparation:

The samples were prepared with BSA as standard. 15 µl of BSA with 6 µl of loading buffer was used as marker. 27 µl of each sample along with 6 µl of loading buffer was prepared. The samples and marker were heated at 100°C for 10 – 12 minutes in a water bath then vortexed for 10 seconds. The samples were then loaded onto the gel which was run at 50 V till the stacking gel was crossed followed by 100 V till the end.

5. Staining the gel:

The gel was stained using commassie blue. 12 mg of Commassie blue was used from the stock of 60 mg/liter and added to the staining solution. Staining solution was prepared using 0% acetic acid and 40% methanol in a total volume of 200 ml. the volume was raised using distilled water. The gel was then submerged in staining solution for 30 minutes.

6. Destaining the gel:

The gel was then destained using destaining solution and the bands were observed. The destaining solution constituted of 50% methanol, 10% acetic acid and the rest was distilled water in a total volume of 100 ml. the gel was immersed in destaining solution till the solution turned blue. The step was repeated till clear bands became visible and the gel was washed with distilled water and stored.

CHAPTER 4

Results and discussions

1. **Protein Estimation by Bradford assay:**

   The readings recorded are as follows:

Table 5: Average Absorbance for Bradford's Standard Plot

Sample ID	LC 3A	LC 4A	LC 5A	LC 6A	LC 7A	LC 8A	healthy control
Absorbance	1.376	1.337	1.307	1.303	1.341	1.276	1.354

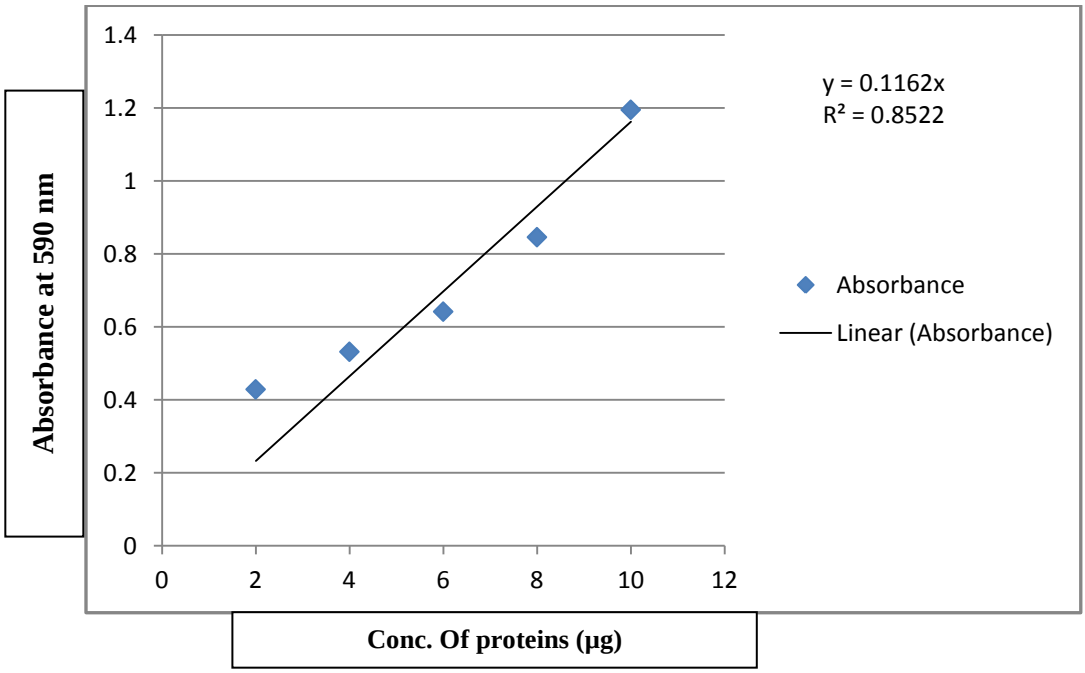


Table 5.1: Standard plot for Bradford's Assay

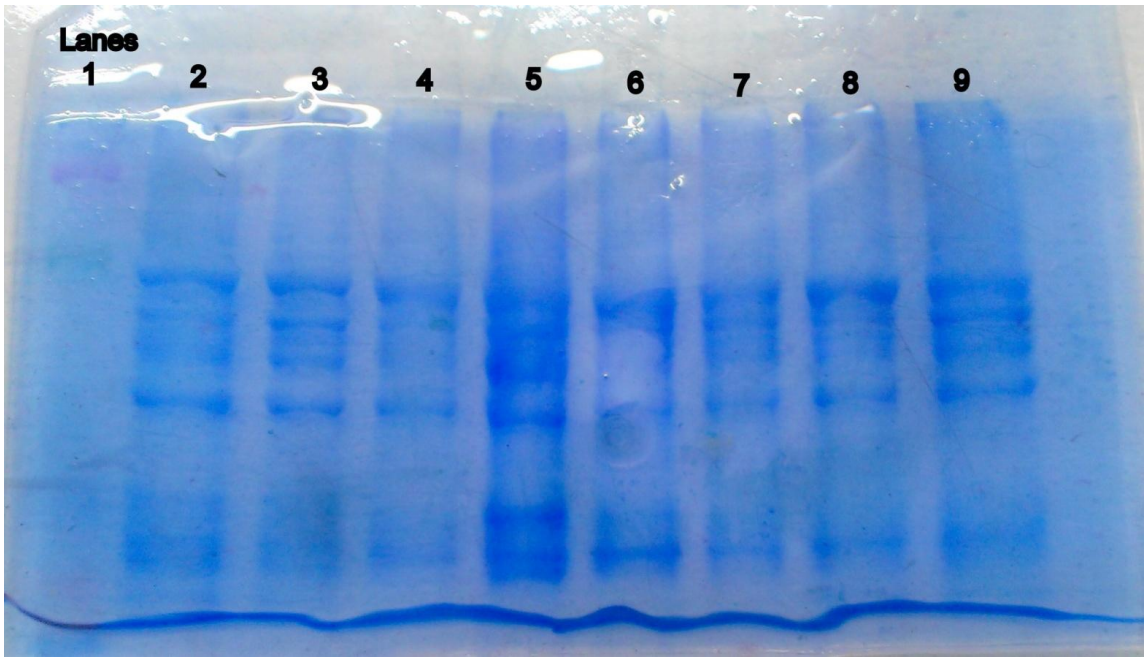
Recommended Standard proteins for Bradford Assay include bovine gamma globulin (IgG), lysozyme and ovalbumin because they have closer to the average or typical number of amino acids that bind the dye. The color development in Bradford assay is complete within 2 minutes and the color remains stable for upto an hour. With the

micro- assay procedure, Bradford assay can be used to determine proteins in the range of 1 to 20 µg.

2. **SDS PAGE**

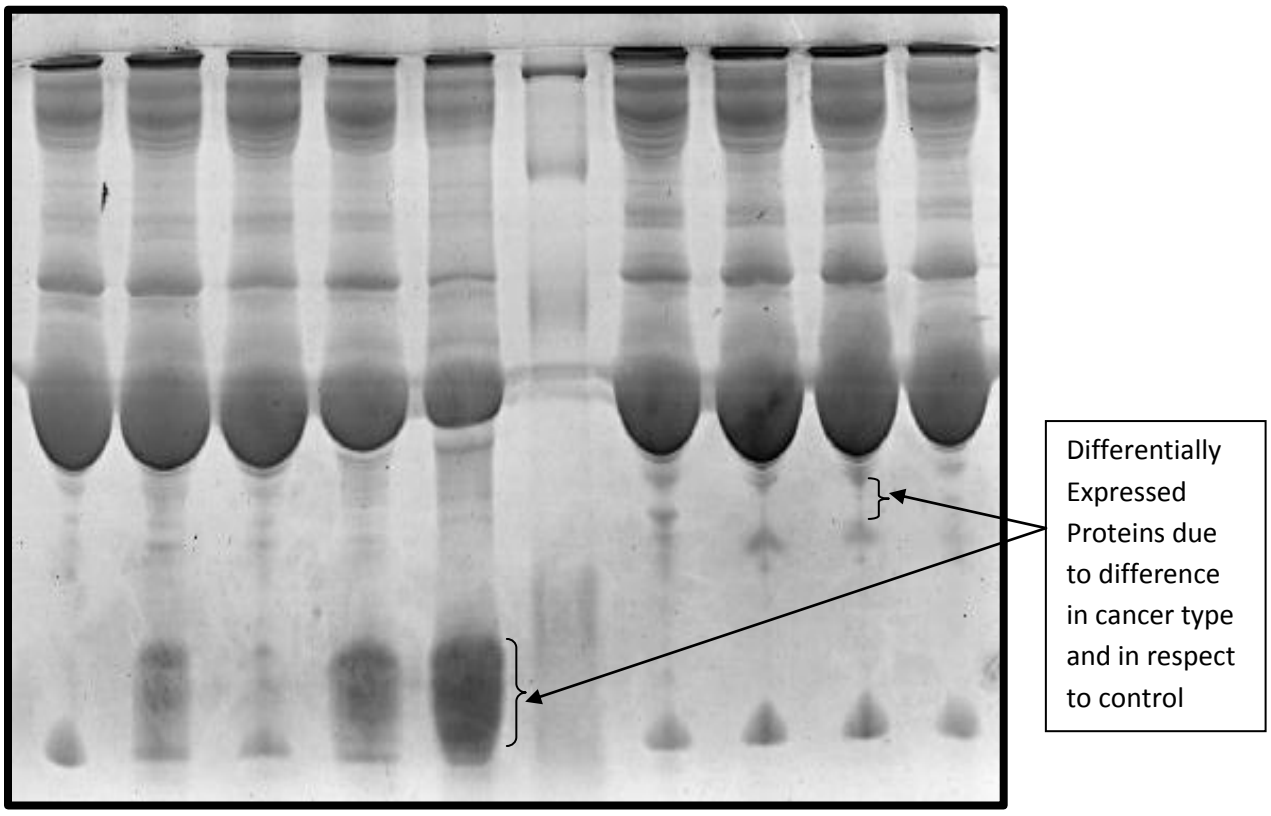
Initially, in SDS-PAGE no desired bands were observed due to low concentration of protein in samples. There is also a possibility that the protein was degraded or there might have been some problem with the reagents used. Therefore, standardization and optimization procedures were carried to obtain optimum results.

SDS- PAGE using different control samples



Lane 1 consists of Biorad Kladioscope marker and the rest of the lanes from 2 to 9 consist of different samples of healthy controls. Bands can be clearly seen after commassie blue staining.

SDS- PAGE gel picture for lung cancer samples showing differential expression

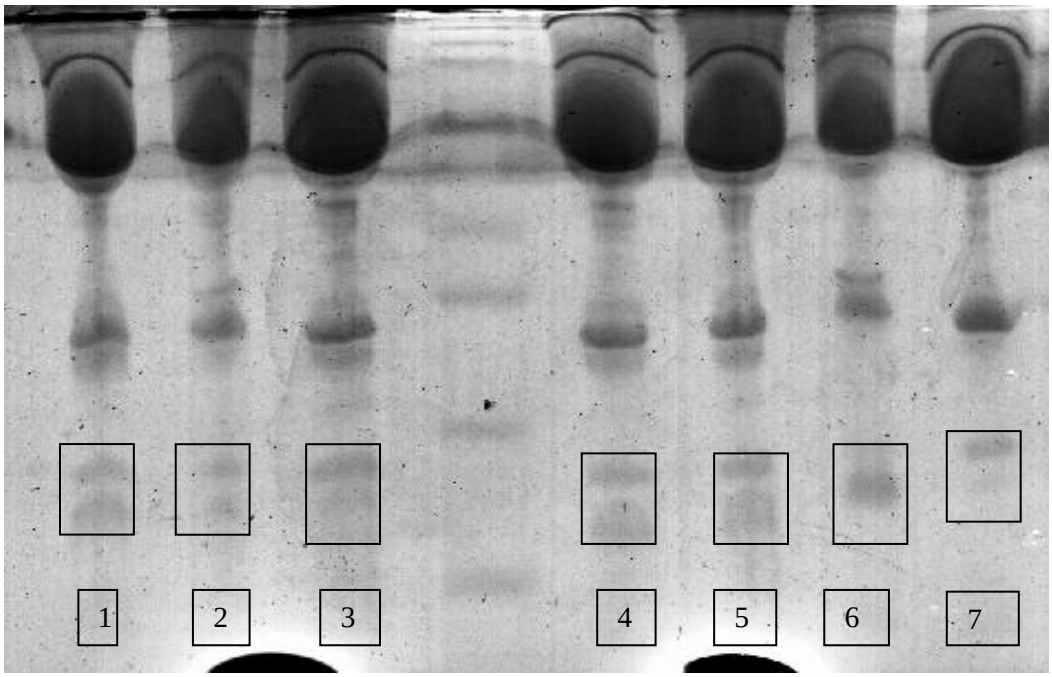


Lane	Sample
1	1st Control
2	LC 1
3	LC 2
4	LC 3
5	LC 4
6	Molecular Weight Marker
7	LC 5
8	LC 6
9	LC 7
10	2nd Control

Several differentially expressed proteins were obtained as can be seen from the figure. These regions were excised and sent for mass spectrometry.

Mass spectrometry results of lung cancer samples

Lane	1	2	3	4	5	6	7	8



Bands Cut	Sample ID	Protein Identified
1	L3 (IIIA)	Ras- related protein RABB1a
2	L4 (IIIB)	tRNA(guanine(10)-N2-methyltransferase homolog)
3	L5 (IIIB)	Ras- related protein RABB1a
4	L6 (IIIB)	Protein NrdI
5	L7 (IIIA)	Peptide chain release factor I
6	L8 (IIIA)	WASH complex subunit CCDC53 homolog
7	L9 (healthy control)	Ras- related protein RABA4c

From the mass spectrometry results it can be seen that there is a considerably varying expression of proteins among the patient samples. Also it can be seen that the common proteins are very few which is what we wish to obtain. Following are the functions of the above mention differentially expressed proteins:

1. **Ras- related protein RABB1a**: They are small GTPases and key regulators of intracellular membrane trafficking, from the formation of transport vesicles to their fusion with membranes. They have an inactive GDP-bound form and an active GTP-bound form that is able to recruit to membranes different sets of downstream effectors directly responsible for vesicle formation, movement,

tethering and fusion. RAB1A regulates vesicular protein transport from the endoplasmic reticulum (ER) to the Golgi compartment and on to the cell surface, and plays a role in IL-8 and growth hormone secretion.

2. **tRNA(guanine (10)-N2-methyltransferase homolog):** Catalytic subunit of an S-adenosyl-L-methionine-dependent tRNA methyltransferase complex that mediates the methylation of the guanosine nucleotide at position 10 (m2G10) in tRNAs/

3. **Protein NrdI:** It encodes a zinc dependent endopeptidase that cleaves the peptide substrates at the N terminus of arginine residues and is a member of the peptidase M16 family. It also interacts with the heparin binding EGF- like growth factor which is known to play a role in cell migration and proliferation.

4. **Peptide chain release factor:** It directs the termination of translation in response to the peptide chain termination codons UAG and UAA.

5. **WASH complex subunit CCDC53 homolog:** The WASH (WASP and Scar homologue) complex is present at the surface of endosomes and recruits and activates the Arp2/3 complex to induce mediated actin nucleation. The WASH complex plays a key role in the fission of tubules that serve as transport intermediates during endosome sorting.

6. **Ras- related protein RABA4c:** Intracellular vesicle trafficking and protein transport.

Conclusion

With the abundantly escalating cases of lung cancer, no fool proof treatment strategy nor an efficient detection and screening procedure has been developed till date. With this concern in mind, researchers all across the globe aim to come up with various strategies so that this deadly disease can be tackled and brought under control at the earliest.

Apart from using secretory proteins various other approaches like SNP detection, marker discovery and others are being tries so that one way or the other our goal is achieved. The ultimate aim of my project has been to be able to successfully come up with a set of proteins which are common for the lung cancer patients of this particular area. Through the duration of this project, I came across quite a few road blocks and failed results, but inspite of that I have been able to gain considerable amount of success in this project. Hopefully my results can be of clinical use some day.

Appendices

The reagents were obtained from the following manufacturers:

Sr. No.	Reagents	Manufacturers
1	SDS PAGE assembly	Biorad
2	Bradford assembly	Nunc
3	TEMED	Biorad
4	Sodium Carbonate	Sisco Research Labs Pvt Ltd
5	Potassium Ferricyanide	S.D. Fine chemical Ltd.
6	Potassium Thiosulphate	Merck
7	Formaldehyde	Merck
8	Ethanol	Changshu Yangyuan chemical, China
9	Acetic Acid	Fisher Scientif
10	Methanol	Merck
11	Commassie RT50	Biorad
12	Loading Buffer	Biorad

References

1. Unlocking Biomarker Discovery: Large Scale Application of Aptamer Proteomic Technology for Early Detection of Lung Cancer; Rachel M. Ostroff et al. PLoS One, December 2010.
2. Secretomic Analysis Identifies Alpha-1 Antitrypsin (A1AT) as a Required Protein in Cancer Cell Migration, Invasion, and Pericellular Fibronectin Assembly for Facilitating Lung Colonization of Lung Adenocarcinoma Cells; Ying-Hua Chang et al. Mol Cell Proteomics, 2012 November
3. Sieving through the cancer secretome; Lin et al, Elsevier, 2013
4. Lung Cancer Proteomics: Recent Advances in Biomarker Discovery, Indovina et al, International Journal of Proteomics, 2011
5. Biomarkers of Lung-Related Diseases: Current Knowledge by Proteomic Approaches, Lau et al, Journal of cellular physiology
6. Cancer mortality in India: a nationally representative survey
7. Cancer Statistics, 2014
8. A proteomic approach to identification of lung cancer biomarkers.

Brief description of the student

I am currently on the verge of the completion of my B.Tech degree in Biotechnology in June 2014 from Jaypee University of Information Technology, Waknaghat, Solan (H.P). My current CGPA is 8.1 and I have a keen interest in pursuing my higher studies in the field of Cancer Biology. I have applied for my Masters in US and Canada and have received offers of admission from Florida Institute of Technology and McGill University, Canada. I aim to achieve success in my field with my diligence, perseverance and hard work.

Porangana Nayar