Comparing the Proteome of Squamous Cell Carcinoma versus Normal Esophagus to Find Molecular Markers for Recognition of the Disease

Zahra Najafi¹*, Ferdous Rastgar Jazii¹, Omid Reza Noorain¹.

¹National Institute of Genetic Engineering and Biotechnology, Tehran-Iran

* Corresponding Author: email address: znajafee@yahoo.com (Z. Najafi)

ABSTRACT

Since esophageal cancer is among the most fatal cancers all over the world and our country has the most frequent number of patients, identification of squamous cell carcinoma associated proteins would be essential. Finding molecular markers for recognition of the disease could be beneficial for efficient and early treatment of the disease. As protein expression in cancerous cells is different from normal ones, identification of protein expression pattern could be helpful for recognition of the disease. Proteomics is a novel method for recognition of protein collection in a specific tissue. In this method 2D gel electrophoresis is performed to separate all proteins, and then spot analysis will be done through mass spectrometry and bioinformatics software would help in the recognition of proteins. In comparison to normal tissue in cancerous cells, we would have up-regulation, down-regulation, appearance or disappearance of a specific protein. So, protein extraction was performed from healthy and cancerous mucosal tissue of esophagus and all peptides were separated through 2D gel electrophoresis. After spot analysis, proteins with different expression were identified with mass spectrometry and bioinformatics. In comparison to normal tissue, 14 proteins were found to be over expressed or have no expression in tumors. The main objective of this study is that proteomics is an ideal method to find the molecular basis of squamous cell carcinoma in esophageal cancer.

Keywords: Squamous Cell Carcinoma; Esophagus; Esophageal cancer; Proteomics; Two dimensional electrophoresis; Polypeptide marker

INTRODUCTION

Cancer is defined as uncontrolled proliferation of cells. This phenomena is arises from genetic changes and would cause normal tissue damage, metastasis, migration and establishing new tumors in different tissues [1]. Cancer results from genetic damages [2].

In cancer, different chromosomal disorders happen which result in genetic changes. In addition, many environmental factors can increase genetic damage and cancer risk, such as inflammatory reactions, chemical stimulations, and physical blows and infectious microorganisms.

Two essential groups of genes are very important in genetic changes and damages:

1) Cellular oncogenes (or proto-oncogenes), and
2) Antioncogenes (or tumor suppressor genes) [3].

Oncogenes are effective in cell growth and proliferation and tumor suppressor genes regulate or limit cell proliferation. These genes can effect cell proliferation, induction of differentiation regulation of cell to cell contacts and senescence induction or apoptosis [4, 5].

In all kinds of cancers, tumor suppressors are inactivated so inactivation of suppressors has effective role in initiation and progress of the disease [6].

Activation of oncogenes and deactivation of suppressor genes is the effect of chemical carcinogens [7], chromosomal changes [8], viruses [9] and their products. For example, activation of myc or ras, myb or erb-B oncogenes or inactivation of tumor suppressor genes as p53, Rb, MCC and DCC are the consequence of chemical carcinogens [10]. Tumorogenesis is a multistage process that includes initiation [11], promotion [12], progression [13], and metastasis [14]. Every stage has its specific genetic and epigenetic damages and leads to make malignancy and invasion.
Metastasis includes growth, angiogenesis, and invasion. Without access to blood, tumors outer cellular spaces, enter to circulation, connect to basal endothelial membranes and exit from lymphatic system and vessels, and to complete metastasis they digest extracellular matrix and immigrate to mesenchymal matrix and deposit angiogenesis factors [15]. Tumor made vascular system is not like normal vessels and it helps for better invasion to vascular system. During transmission in vascular system, tumor cells can connect to each other through adhesive specificity and clump up. Sometimes WBCs are also added to this collection to protect it. On the other hand, this clump will initiate metastasis.

Esophagus cancer is the sixth cancer in the world and the fourth in developing countries [16]. It usually happens in middle ages or older but it could occur in people younger than 25 years old. The disease is more common in men than women and more in black people than whites. Incident frequency is quite different all over the world. Environmental factors could have an impact on disease occurrence. Previous studies showed that China and Iran have high risk in disease incidence. Torkman Sahra and Gonbad in Iran have the most frequency of esophagus cancer and China with 65% of occurrence is in the second stage. Also 70% of all cases in the world belong to Iran [17]. In molecular studies, proteomics is a method for proteome recognition in different organisms. Proteome is the collection of proteins that are produced by each organism.

Each specific species has its own constant genome, but proteome changes from time to time in different stages of life. So, proteome is the collection of specific proteins in a specific life stage of each organism. On the other hand, genome is only the recipe of producing proteins but proteins perform special cell functions. Also, different cells have different proteins. All of the cells in a body have the same genome sequence but activation of different genes in various sorts of tissues produce broad kinds of proteins. Also tumor cells produce proteins other than normal tissue, so recognition of the entire human proteome and protein – protein interactions would guide us to know molecular basis of the diseases and produce more effective drugs.

In human genome project, 3 billion DNA base pairs, especially protein coding sequences were recognized. To some extent it is possible to predict protein sequence from genome sequence, but proteome is more complicated can not grow more than several millimeters, so tumor cells produce angiogenesis factors, digest than genome. On the other hand, proteins have different functions in different environments because some proteins are active in aqueous environment but others are hydrophobous. Researchers believe that human genome has 40 thousand genes but 100 thousand proteins are being built. Proteome provides us with information about specificities of all these proteins in translation stage and post translational modification level and also about protein – protein interactions [18].

So, in conclusion:
1. DNA sequences are constant map of biochemical pathways.
2. DNA sequence can not show changes and post translational modifications such as splicing and translational modifications as phosphorylation, glycosilation and other essential changes that make a protein functional.
3. Most of the time, genome remains intact but proteins can change due to induction or suppression of different genes [19].
4. Our knowledge about genome sequence can not lead to protein recognition, we can use genome only for ORF prediction, but it's not accurate.
5. We can not predict protein action by its DNA sequence.
6. We can not predict number of proteins which may be being built by a gene.

MATERIALS AND METHODS

Patients and tissue sampling
Cancerous and normal tissue specimens were collected from 45 patients with SCCE who underwent surgery. Tissue samples were collected immediately after surgery, wrapped in aluminum foil, snap frozen in liquid nitrogen and maintained at -70°C. The age of the patients at the time of diagnosis ranged from 27 to 86 years (63% males, 37% females) with a mean of 55 years.

Protein preparation
100-150 mg of tissue was sliced on ice and pulverized under liquid nitrogen using a microdismembrator (Braun, Germany). Subsequently 600 μl homogenization buffer was added to the pulverized tissues, mixed and 10 μl of the following protease inhibitors were added: Pepstatin (1mg/mL in isopropanol), benzamidine (16mg/mL in H2O), phenylmethylsulphonyl fluoride (PMSF at 25mg/mL in isopropanol). To this homogenate,
10μl of RNase A (10mg/ml in homogenization buffer) and DNase I (1mg/ml in homogenization buffer) were added and incubated on ice for 20min. Subsequently, urea at 7mol/l, thiourea at 2mol/L, 5% β-mercaptoethanol and 0.5% SDS were gradually added and the volume of solution was adjusted to 1.5ml with the homogenization buffer. Samples were centrifuged at high speed for removal of insoluble particles [20] and 5μl of each was used for protein concentration assessment using the Bradford assay. [21]

Two-dimensional electrophoresis
Samples were subjected to isoelectrofocusing (IEF) following adaptations and slight modification [22]. The first dimension gel was composed of 4.2% acrylamide, 0.27% (V/V) Nonidet P-40 (NP-40), 5% sucrose and 6% ampholytes (pH 4-6, 5-7 and 6-8 at a ratio of 2:1:2, respectively) and 0.05% TEMED (N, N, N’, N’- Tetramethylethylene diamine). The solution was degassed and 0.04% ammonium persulfate was added, mixed and poured to a height of 130 mm in cylindrical glass tubes with a 1.5 mm internal diameter. A volume of sample equal to 75 μg total protein was mixed with 0.33 volume of neutralizing buffer (9 mol/l urea, 8% NP-40 and 5% ampholytes pH 3.5-10), loaded on IEF gel, overlaid with 10 μl sample buffer (4 mol/l urea, 1% ampholytes pH 3.5-10) and filled with catholyte. The upper chamber buffer or catholyte was composed of extensively degassed 0.02mol/l NaOH and the lower chamber buffer or anolyte; 0.01mol/L phosphoric acid. Iselectric focusing was applied, without prefocusing, at 300 V for 1 h, afterward at 600 V for 10 h and 800 V for 1 h in order to complete focusing. Gels were removed and equilibrated for 20 min at room temperature in equilibration solution (60mmol/l Tris-HCl, pH 6.8, 2% SDS, 5% β-mercaptoethanol (V/V), 10% glycerol (V/V), and 0.002% bromphenol blue). The second dimension gel consisted of 33.3 ml of 30% stock acrylamide solution and N, N’-methylene bisacrylamide(29.2% and 0.8% W/V respectively), 41.7 ml de-ionized water and 25 ml separation gel buffer (1.5mol/l Tris- HCl, pH 8.8, 0.4% SDS [Sodium Dodecyl Sulfate]), 0.034% W/V) ammonium persulfate and 0.05% TEMED. The equilibrated first dimension gel was layered on the second dimension gel and fixed in place with 1% agarose and electrophoresis was carried out at 30 mA/plate at 10°C constant temperature by applying a cooling system. [23]

Protein detection
Proteins were detected using a slight modification of the previously reported method [24]. The gel was fixed (methanol, water, acetic acid and formaldehyde: 50/38/12/0.05 per volume) for at least 1h with constant shaking and followed by 3 20 min washes with 50% ethanol, pretreated with sodium thiosulfate (Na2S2O5,5H2O: 0.2g/l) for 1 min, and washed three times each for 20s with ddH2O. Impregnation of the gel with AgNO3 (1.9 g/l and 0.075% (V/V) of 37% formaldehyde) was carried out and the residual AgNO3 was removed by 3 × 20s successive washes with de-ionized water. The gel was developed by soaking it in developing solution containing Na2CO3 (60g/L), 0.05% (V/V) of 37% formaldehyde and 4mg/l of Na2S2O5,5H2O for 10min. up until it appeared to develop yellowish brown spots. The gel was then rinsed twice, each for 2min, with ddH2O. Further development was stopped by immersing the gel in stop solution (50% methanol and 12% acetic acid) and stored in 30% ethanol at 4°C until scanning.

Mass spectrometry
Silver stained protein spots containing the proteins of interest were de-stained thoroughly with 1% H2O2 (typically 1min) and lyophilized to dryness. Silver stain removal by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI). The dehydrated gel bands were hydrated with 15μg/l (Promega, Madison, WI) of porcine trypsin in 25mmol/l NH4HCO3, pH8.2 on ice for 45min. Excess trypsin was removed; gel bands were covered with 25mmol/l NH4HCO3, pH8.2 and incubated at 37°C overnight. Tryptic peptides were extracted from the gel bands with 70% acetonitrile and 0.1% trifluoroacetic acid. Samples were desalted with C18 Zip Tips (Millipore, Bedford, MA) as per manufacturer’s protocols. 0.5μl of sample was co-crystallized with0.5μl of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 1% trifluoroacetic acid and spotted directly on a stainless steel MALDI target plate. Mass spectra were acquired using a MALDI-TOF/TOF mass spectrometer (Voyager 4700, Applied Biosystems, Foster City, CA). MALDI-TOF/TOF spectra were internally calibrated (<20ppm) using trypsin autolysis products. Post-acquisition baseline correction and smoothing was carried out using software provided with the TOF/TOF instrument. Spectra were submitted to Mascot
RESULTS
Identification of poly peptides
For every sample 2 different 2D gel electrophoresis were performed and tumoral tissue was compared with its own normal tissue. Position, size and intensity of every spot were compared in tumoral and normal tissue, and also with other samples. Making comparison based on existence or elimination of spots and also increase or decrease of spot intensity was performed and differences between normal and malignant tissue of the same origin were recorded in table.

In addition, as a final record, spots which had been changed (elimination, existence or change in size) in more than 70% of gels were recorded. Studies have been performed on 30 samples of normal and tumoral tissues from 11 women and 19 men aged between 23-73 (mean 58 years old).

All tumors were metastatic and included different tissues from mucosa to serous. Cellular differentiation in different patients was recognized as, 7 patients were well, 4 were moderate and 14 were poor.

Comparing protein expression in normal and tumoral tissues
There are two methods for the study of polypeptides and changes in expression levels, one, using software to count and compare every spot in normal and tumoral tissue of the same specimen, which is more accurate and correct. The second method is finding different spots and their changes by watching them and recording changes. In comparison with tumoral and normal tissues, spots with more than 70% changes were recognized and recorded. 14 spots in tumoral tissue were recognized which had changed drastically, from these counts, 7 spots existed or had increased expression levels and 7 spots were eliminated or had decreased expression level (Tables 1).

Table 1. Identification and characteristics of the 14 proteins whose expression were subjected to change in SCCE

<table>
<thead>
<tr>
<th>Polypeptide</th>
<th>PI/MW</th>
<th>Expression Profiling or level</th>
<th>Tissue specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5.5/40.3</td>
<td>Disappearance</td>
<td>Normal</td>
</tr>
<tr>
<td>B</td>
<td>5.45/17.6</td>
<td>Down Regulation</td>
<td>Normal</td>
</tr>
<tr>
<td>C</td>
<td>5.6/17.6</td>
<td>Down Regulation</td>
<td>Normal</td>
</tr>
<tr>
<td>D</td>
<td>5.8/17.6</td>
<td>Down Regulation</td>
<td>Normal</td>
</tr>
<tr>
<td>E</td>
<td>6.08/17.6</td>
<td>Down Regulation</td>
<td>Normal</td>
</tr>
<tr>
<td>F</td>
<td>6.6/76</td>
<td>Down Regulation</td>
<td>Normal</td>
</tr>
<tr>
<td>H</td>
<td>7.3/35.5</td>
<td>Down Regulation</td>
<td>Normal</td>
</tr>
<tr>
<td>1</td>
<td>6/64</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
<tr>
<td>2</td>
<td>5.5/32.8</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
<tr>
<td>3</td>
<td>5.4/20.1</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
<tr>
<td>4</td>
<td>6.9/42.4</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
<tr>
<td>5</td>
<td>6.1/27.2</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
<tr>
<td>6</td>
<td>6.1/27.2</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
<tr>
<td>7</td>
<td>6.1/27.2</td>
<td>Up Regulation</td>
<td>Tumor</td>
</tr>
</tbody>
</table>
In figures 1 and 2, there are 2 gels from normal and tumoral tissues. In comparing the gels, we can find 7 spots in normal tissue which are eliminated or decreased from tumoral sample, these spots include A, B, C, D, E, F, and H. Spot A is completely eliminated but it could be present to some extent in some samples due to having traces of normal tissue accompanying with tumoral ones. And 6 spots, B to H have decreased expression in tumoral tissue in comparison with normal tissue. According to figure 2, there are 7 spots in tumoral sample which existed or increased expression level in comparison with normal sample. The newly existed proteins in tumoral tissue nominated as tumor associated proteins.

These peptides are specific for tumoral tissues and have critical role in cancer development. These spots that are numbered from 1 to 4 are common in 70% of cases and spots 5 to 7 had increased expression level in majority of tumoral samples (Figures 1 and 2). Some polypeptides which are expressed more in esophagus tissue could be seen as large and intense color spots. These kinds of spots have high expression levels and may cover other tiny spots in this area and we may lose some information about tiny expressing proteins in these points.

Also some times tumoral tissue is accompanied with normal ones and we may recognize some normal proteins through tumoral gels.

![Figure 1. A representative 2DE gel of a normal tissue. Proteins that become down-regulated in corresponding tumor are shown with arrows and capital letters](image1)

![Figure 2. A representative 2DE gel of tumor tissue. Arrows and numbers indicate up-regulated proteins in comparison with their matched normal tissue](image2)

**Protein recognition by mass spectrometry**

Silver stained protein spots containing the proteins of interest were separated completely from gel and after adding acid acetic 5%, samples were used for mass spectrometry. By using MALDI/TOF/TOF system, amino acid sequences of the protein were recognized and similar spots were distinguished.

**Bioinformatic studies**

Raw data for mass spectrometry were analyzed according to data banks' information. Information about a single protein included biological function, protein interactions, cellular position, first to third protein structure, specific ligands and its homology with different proteins (Table 2).

Different data banks as listed below were used for this purpose:

- www.ncbi.nih.gov
- www.expasy.ch/tools
- www.ebi.c.uk
- www.pdb.bni.gov
- www.malsoft.com
- www.ncbi.nlm.nih.gov/BLAST
- www.google.com/profound
DISCUSSION

In esophagus cancer, less than 5% of patients survive more than 5 years of disease occurrence [25]. Two types of esophagus cancer are recognized: Adenocarcinoma, and squamous epithelial tissue. Etiology of adenocarcinoma is barrette phenomena which is caused by stomach acid reflection in esophagus and PH changes would lead to transformation of squamous cells to cylindrical ones in the last one third length of esophagus. Most esophagus cancer cases in the world occur in Iran [26]. So recognition of molecular etiology of the disease is very important. Also finding molecular markers for early distinguish of the disease would be helpful in effective treatment and patients' surveillance.

Most of the studies have been performed on esophagus adenocarcinoma, and researchers could recognize some molecular markers such as increased expression level of P53, PCNA, growth factors like epidermal growth factor (EGF) and its receptor (EGFR) and TGFα [27]. But they cannot find any special marker for squamous epithelial cell carcinoma. These days, due to advances in technology, through using 2D electrophoresis, recognition of specific markers related to this kind of disease is possible. Occurring in limited areas in the world, esophagus cancer has not been fully studied as other kinds of cancers have been. According to multistage nature of cancer, knowing initial changes would help a lot, for example, recognition of Barrette phenomena is a critical sign for initiation of esophagus cancer. In addition, besides barrette, some mutations in tumor suppressor genes like P53 and APC would happen [28]. Also, before cancer dysplasia in esophagus tissue would be an alarm for a serious case. Cytopathological studies have been performed in Iran and also in Linxian area of China but molecular studies are undoubtedly essential in disease studies.

2D electrophoresis technique has been performed for lung [29], liver [30], and colon [31] cancers, and changes in protein expression,
elimination and existence of proteins in cancerous versus normal tissue were mentioned. These changes would occur in skeletal cell structure (morphology) protein function, cell proliferation and differentiation [32]. All of the studies have been performed for not only finding molecular etiology but also molecular markers for recognition of the disease (in polypeptide level) and follow up the progress of treatment. Among all these, decrease or elimination of E Cadherin is performed in esophage cancer [33]. This protein has an essential role in morphogenesis, differentiation (especially in fetus), establishing polarity in fetus evolution and induction of early genes and connection and interaction between cells [34]. Decrease or deletion of Trp-B or its isoforms are being studied in this research and is categorize in changes which could be the subject of future studies.

This protein as a good molecular marker candidate preserves cytoskeleton cell structure and establishes involuntary esophage motions with contraction mechanism [35]. Novel Progresses in proteomics and mass spectrometry will lead to find more specific and sensitive molecular markers for early distinguishing critical diseases as cancers.

ACKNOWLEDGEMENTS
The authors would like to thank Professor Reza Malekzadeh at Tehran University of Medical Sciences for his invaluable help to have NIH (in the USA) to conduct mass spectrometry on the samples.

REFERENCES