Proteomic analysis of gene expression during human esophagus cancer

Hakimeh Zali[1], Ghasem Ahmadi[2], Roza Bakhshandeh[2] , Mostafa Rezaei-Tavirani[3,*]

[1]Faculty of Paramedical science, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2] Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3]Faculty of Medicine, Ilam University of Medical Sciences, Ilam,Iran

*Corresponding Author: email address: rezaei.tavirani@ibb.ut.ac.ir(M. Rezaei-Tavirani)

ABSTRACT

Esophagus cancer is the eighth most common cancer worldwide and particularly high in an area extending from the southern border of the Caspian Sea in Iran across central Asia to China. Since information about this mysterious disease is poor, proteomics may be solving this enigma. Altering gene expression in cancer cell is a remarkable indicator can be detected by proteomics techniques and bioinformatic analysis. In this study, normal and cancerous cells were obtain from patients, total proteins were purified by standard methods, and proteins separated by two dimensional electrophoresis (2DE). Some of proteins were identified by Mass spectrometry (MS-MALDI method). By using bioinformatic analysis illustrate molecular mechanism in this disease. Analysis of gels base on Flicker software and Mass Spectrometry led the same result. 61 protein spots detected in both gels that 21 spots have down regulated and 12 spots have up regulated in cancerous cell than normal. About 14 spots were disappeared in cancer cell while 14 new spots expressed. By using flicker detected 8 Protein that refer to TRFE, SZ07, C1 TC, Kininogen, anexin, keratin, fructosebisphosphate aldolase A and heat shock. Mass spectrometry (MS-MALDI method) identified anexin, keratin, fructosebisphosphate aldolase A and heat shock. Identified proteins were functionally categorized based on Gene Ontology (GO) annotation terms using the DAVID program package. The major molecular functions that annotated with PIR include phosphoprotein, disease mutation while annotated by GO include response to organic substance, response to wounding and cellular homeostasis. The cellular component and molecular function presenting the greatest enrichment that concluded two clusters that the two most importants are cellular homeostasis and extracellular region part. Results reveal that the most of molecular function in cancerous tissue maintenance cellular homeostasis, cell regeneration and repair, so tissues undergo stress try to survive. It can be also concluded that aldolase A, fructose-bisphosphate, keratin 14, formyltetrahydrofolate synthetase and transferrin can be some diagnostic biomarkers and also drug targets in esophagus cancer.

Keywords: Proteomics; Esophagus Cancer; Biomarker

INTRODUCTION

Esophagus cancer is the eighth most common cancer worldwide and varies several-hundred-fold between nations and between geographic regions within nations [1]. The incidence is particularly high in an area extending from the southern border of the Caspian Sea in Iran across central Asia to China. Esophagus cancer is kind of multifactorial disease that principal genetic and epigenetic factors not exactly well known so far. Studies indicate that the major etiologic risk factors in most developed nations are ethanol and cigarette smoking [1, 2]. These two factors may convey a significant additive effect which is thought to predispose individuals to greater risk. For example, alcohol may interact with folate, vitamin B12, and methyl group metabolism to modulate risk [1, 3]. A number of studies suggest an inverse relationship between risk of esophageal cancer and the consumption of fresh fruits and vegetables [1, 4, 5]. The indigenous diet which is low in fresh fruits, vegetables, animal products, vitamins and several trace elements may be susceptible reason for esophageal cancer and parts of Asia where alcohol consumption does not the high risk[1,6,7]. In high-risk area between Iran and China, micronutrient deficiencies coupled with the exposure to carcinogenic substances in salt-pickled vegetables or moldy foods may be contributing factors [1]. Tumor suppressor genes, oncogenes, and apoptotic genes are involved in the initiation

and development of esophageal cancer, but to date no gene directly related to esophageal cancer has been identified [2]. Mutations in different codons of p53 cause loss of suppressor function p53 could be susceptibly various esophagus cancers [1, 8]. Over expression of p73mRNA was found in 51.8% esophageal tumor tissues [9]. The APC is tumor suppressor gene like MCC genetic loci occurs in the majority of human esophageal cancers and is involved in the development and/or progression of the disease [10]. P16INK4a and p15INK4b, two tumor suppressor genes found to occur less frequently in human esophageal cancer in Lixian county, China [11]. The oncogenes most frequently activated in esophageal cancer are cyclin D1, c-erbB1 and 2, FRAT1, c-myc, c-ras, Int-2/hst-1, and EGFR [12, 13]. By the way, the procedure that is used to analyze changes in expression of proteins in a cell is proteomics. In proteomic methods by using two dimensional electrophoresis and mass spectrometry, gene expression can be analyzed. Here the amount of expression of proteins in esophageal cancer by proteomic techniques were measured and compared to the normal.

MATERIALS AND METHODS

Sampling:

Normal and cancer esophagus tissues samples were taken from patient. Normal tissues sample belong to the cancerous patient that give cancerous sample. Characteristics of normal and cancerous sample were confirm by histopathological tests

Protein Purification:

Fresh tissue samples of esophagus were snap frozen and kept in liquid nitrogen until use. Tissue samples were powdered by microdismembrator at maximum speed for 60 seconds under liquid nitrogen conditions. Each powdered tissue sample was added to an appropriate amount of lysis buffer containing 10 mM Tris-HCl pH=7.5, 1 mM MgCl2, 1 mM EGTA, 0.1 mM Phenylmethylsulfonylfluoride (PMSF), 5 mM betamercapto ethanol, 0.5% CHAPS and 10% glycerol.25 After 30 minutes incubation on ice, the lysate was centrifuged at 16000g for 30 minutes at 4°C. Protein concentration of all samples was estimated using a Bradford based microassay [14].

Two Dimensional SDS-PAGES:

The first dimension of 2D electrophoresis was performed on the PROTEAN IEF Cell system (Bio-Rad). Next, gels were equilibrated for 15 min in equilibration buffer I (6M urea, 2% SDS, 0.375 M Tris Hcl pH8.8, 20% glycerol, 130mM DTT). A 12% SDS-Polyacrylamide slab gel was used for the second dimension gel electrophoresis. Equilibrated IPG strips were placed on the surface of the second dimension gels and then sealed with 0.5% agarose in SDS electrophoresis buffer (25mM Tris base, 192mM glycine, 0.1%SDS) and were run vertically [15].

Coomassie Brilliant Blue Staining:

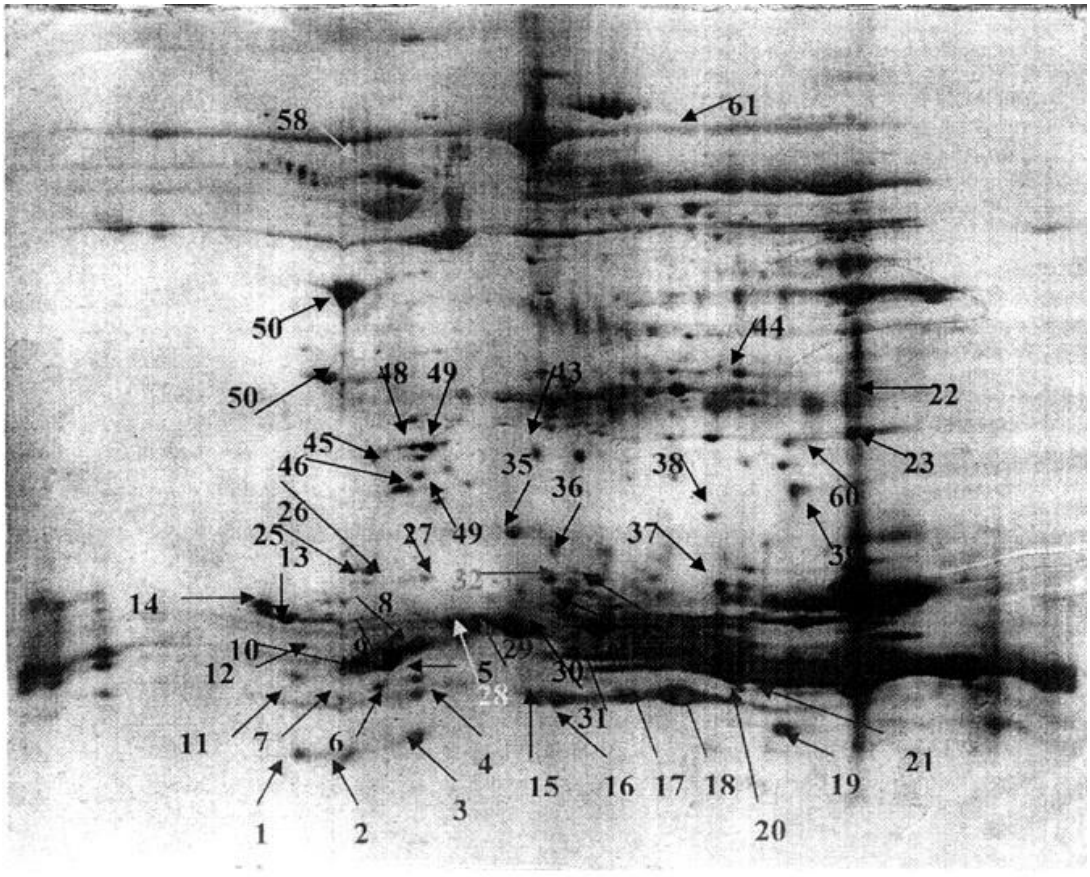
After electrophoresis, the gels were stained with Coomassie Brilliant Blue staining [16].

Mass spectrometry:

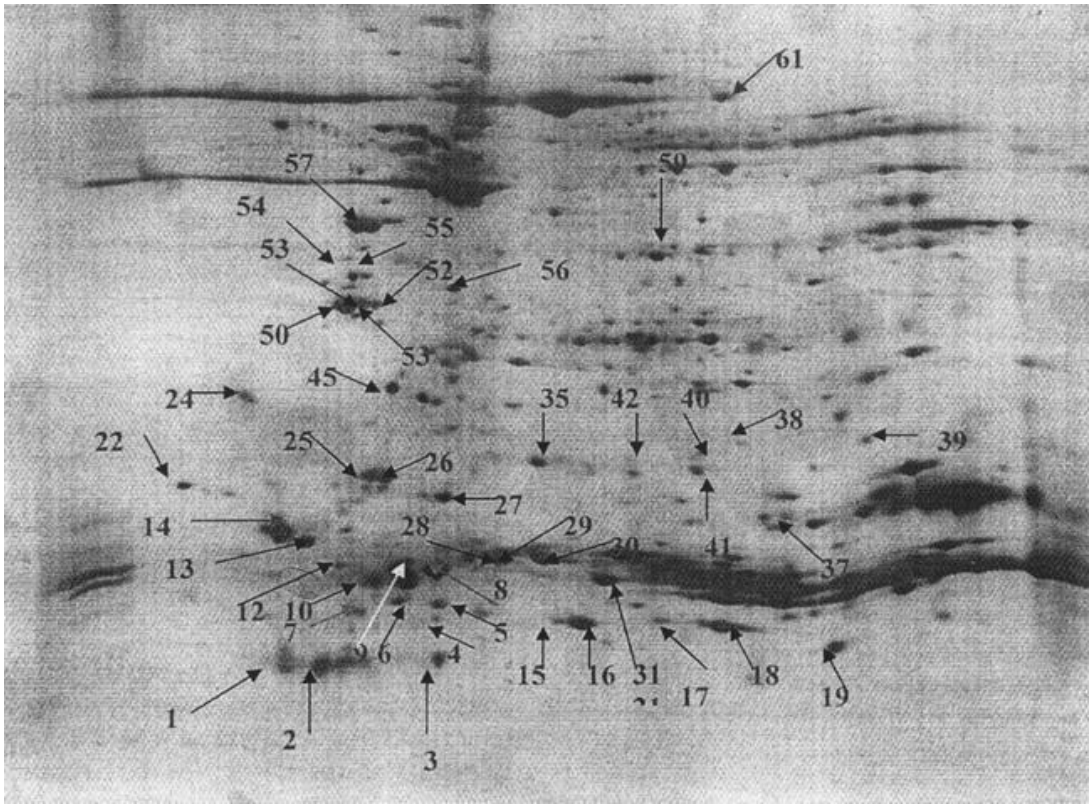
Coomassie Brilliant Blue protein spot containing the interested protein was performed to enhance peptide adsorption by matrix-assisted laser desorption/ionization time-of- flight mass spectrometry (MALDI) [17]. The dehydrated gel bands were hydrated with 15 μg/L (Promega, Madison, WI) of porcine trypsin in 25 mmol/L NH4HCO3, pH8.2 on ice for 45 min. Excess trypsin was removed; gel bands were covered with 25 mmol/L NH4HCO3, pH8.2 and incubated at 37℃ overnight. Tryptic peptides were extracted from the gel bands with 70% acetonitrile and 0.1% trifl uoroacetic acid. Sample was desalted with C18 Zip Tips (Millipore, Bedford, MA) as per manufacturer's protocols. 0.5 μL of sample was co-crystallized with 0.5 μL 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 (< 20 ppm) 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 (http://matrixscience.com) for peptide mass fingerprinting.

Bioinformatic Analysis:

2DE gels are scanned and gels are analyzed by flicker software to compare gels together and compare the spots in one statement in gels and get the density of same spot in each of gel. Then the spots compared to data banks to detection the spots in one statement in every experiment gel and data bank references gels.. The here identified proteins were functionally categorized based on Gene Ontology (GO) annotation terms using the DAVID program package.



a



b

Figure1: 2DE gel image of a) normal tissue, b) esophagus cancer tissue. The spots correspond to the proteins that identified by flicker software

RESULTS

The separated proteins appear as single spots on 2DE gels. The analyzed 2DE gels of (a) normal and (b) cancerous human esophagus tissues are represented in figure 1. As it is shown in figure 1, 61 spots are labeled in both gels that 21 spots have down regulated and 12 spots have up regulated in cancer tissue than normal tissue. 14 spots were disappeared in cancer tissue while 14 spots newly expressed. The next step for analysis the separated protein is identification of desired spots. There are two ways for protein identification based on mass spectrometry and using of protein databases [17]. As shown in table 1, analysis by flicker lead to identification of TRFE, SZ07, C1 TC, kininogen, annexin, keratin, fructose bisphosphate aldolase A and heat shock (9, 21, 34, 35, 50, 53, 58 and 61). So Mass spectrometry is a potent method for protein identification, four desired spots corresponded to annexin, keratin, fructose bisphosphate aldolase A and heat shock, picked and treated for Mass spectrometry (MS-MALDI method). For confirming flicker findings the two methods confirm each other. Table 2 briefly describes classification of detected proteins base on Gene Ontology (GO) that categorized according to cellular component, molecular function and biological process. For functional classification of proteins, DAVID datasets analysis is done. The results are tabulated in table 3 including Functional Annotation Clustering and table 4 Functional Annotation Chart of the proteins.

Table 1: The volume of the detected spots in normal and cancer tissues that calculated by flicker software

ID	Normal	cancer	flicker	Mass Spectrometry
9	118096	60872	TRFE	
21	22249		SZ07	
35	645333	47961	C1 TC	
37	41367	26980	Kininogen	
50	67428	105726	Anne xin	Anne xin
53		14571	Keratin	Keratin
58	9202		aldolase A	aldolase A
61	96783	31098	Heat s hock	Heat s hock

DISCUSSION

Esophageal cancer is one of the most common and highly fatal cancers. To identify diagnostic or therapeutic biomarkers for this cancer, investigators are nowadays performing proteomic analyses of cancer tissues and cells and revealing a large number of molecules which are diagnostic, prognostic and informative for carcinogenesis [18]. The aim of this study is understanding proteomics perspective of esophagus cancer in early stages to elucidate their tumor biomarkers. Proteomic analysis of cancerous and normal esophagus tissues as it is displayed in figure 1 depicted some alteration in gene expression for both tissues. However, some of proteins in cancerous tissue are up regulated (12 spots) the more proteins are down regulated (21 spots). The new express proteins (14 spots) and disappeared proteins (also 14 spots) were detected as esophagus cancerous biomarker. It is reported in previous study that there are

newly expressed protein and disappear protein related to esophagus cancer [19].

Table 2: Protein classification of detected spots base on gene ontology

Gene name	Description	Cellular Components	Biological Processes	Molecular Functions
TRFE	Serotransferrin	apical plasma membrane, extracellular region	cellular iron ion homeostasis, transferrin transport, transmembrane transport	ferric iron binding
C1TC	C-1-tetrahydrofolate synthase, cytoplasmic	cytosol, mitochondrion	folic acid metabolic process, histidine biosynthetic process, methionine biosynthetic process, oxidation-reduction process	formate-tetrahydrofolate ligase activity, methenyltetrahydrofolate cyclohydrolase activity, ATP binding
Kininogen	Kininogen-1	extracellular space, plasma membrane	inflammatory response, positive regulation of apoptosis, negative regulation of cell adhesion	cysteine-type endopeptidase inhibitor activity, receptor binding, zinc ion binding
Annexin	Annexin A5	cytoplasm	anti-apoptosis, signal transduction, negative regulation of coagulation	calcium ion binding, phospholipase inhibitor activity
keratin	Keratin, type I cytoskeletal 14	cytosol, mitochondrion, nucleus	epidermis development, hemidesmosome assembly, intermediate filament bundle assembly	protein binding, structural constituent of cytoskeleton
aldolase A	aldolase A, fructose-bisphosphate	cytosol, actin cytoskeleton	fructose metabolic process , glycolysis , metabolic process , striated muscle contraction	fructose-bisphosphate aldolase activity , lyase activity , protein binding
heat shock	Heat shock-related 70 kDa protein 2	cell surface	Stress response	Chaperone

Previous studies also demonstrated that there are several protein biomarkers such as annexin and heat shock families related to esophagus cancer. These biomarkers also are seen in the other gastrointestinal cancer and their expressions were altering in similar manner [20, 21]. Flicker software helps to introduce the proteins by linking to the related data banks. Here as shown in table 1, by using Flicker it was detected that the 8 spots refer to TRFE, SZ07, C1 TC, Kininogen, anexin, keratin, fructosebisphosphate aldolase A and heat shock. Mass spectrometry (MS-MALDI method) is a powerful instrument for protein identification. The mass results confirm that the Flicker finding is real and the considered proteins are anexin, keratin, fructosebisphosphate aldolase A and heat shock. Considering table 2 most of proteins are cytoskeleton proteins and also the proteins that are presented in stress conditions such as diseases and etc. Function of some proteins known to associate with susceptibility to some cancer i.e. genetic variations in C1TC can susceptibility to colorectal cancer [22]. It also can increase probability of C1TC to be an important key in carcinogenesis of esophagus cancer. Reports of proteomic analyses of cancerous tissues and noncancerous tissues classified the proteins into digestive enzymes, growth factors, cell adhesion molecules, calcium-binding proteins, proteases, protease inhibitors, transporter proteins, structural molecules, apoptosis inhibitor, molecular chaperone, as well as proteins related to cell growth, cell differentiation, cell transformation, tumor invasion, carcinogen metabolism, and others [18].

Table 3: Functional Annotation Chart by using DAVID software

Category	Term	Genes	%	P-Value
SP_PIR_KEYWORDS	duplication		42.9	2.0E-3
SP_PIR_KEYWORDS	disease mutation		57.1	9.3E-3
GOTERM_BP_FAT	negative regulation of coagulation		28.6	1.1E-2
SP_PIR_KEYWORDS	blood coagulation		28.6	1.4E-2
GOTERM_BP_FAT	cellular homeostasis		42.9	1.6E-2
GOTERM_BP_FAT	regulation of coagulation		28.6	1.8E-2
GOTERM_BP_FAT	response to wounding		42.9	2.1E-2
UP_SEQ_FEATURE	glycosylation site:O-linked (GalNAc...)		28.6	2.9E-2
SP_PIR_KEYWORDS	plasma		28.6	2.9E-2
SP_PIR_KEYWORDS	phosphoprotein		85.7	3.2E-2
GOTERM_BP_FAT	response to organic substance		42.9	3.7E-2
GOTERM_BP_FAT	homeostatic process		42.9	4.0E-2

Here identified proteins were functionally categorized based on Gene Ontology (GO) annotation terms using the DAVID program package. Amongst other, we found that 8 proteins were annotated with GO cellular component terms. As depicted in table 3, the major molecular functions that annotated with PIR include phosphoprotein (85.7%), disease mutation(57.1%) while annotated by GO include response to organic substance(50.0%), response to wounding (42.9), purine nucleotide biosynthetic process(33.3) and cellular homeostasis (42.9). Disease mutation categorized Protein for which at least one variant, responsible for a disease, so proteins responsible alteration in esophagus cancer may be aldolase A, fructose-bisphosphate, keratin 14, formyltetrahydrofolate synthetase and transferrin .

In order to determine whether the cellular component and the molecular function mentioned above represent functions enriched in the esophagus data set relative to the "theoretical human proteome", we used DAVID to calculate probable overrepresentations of protein classifications relative to the annotated human proteome and thus excluded common protein functions from skewing the functional analysis of the esophagus cell sub cellular functions. DAVID furthers assigns a statistical significance indicator to protein functions and classifications. The cellular component and molecular function showing the greatest enrichment that concluded two clusters in table 4. Cluster1 contains cellular homeostasis (1.6E-2) and extracellular region part (4.8E-2) that are greatly enriched in our dataset. Furthermore, response to wounding and different kind of binding was demonstrating the greatest enrichment in cluster 2. Results disclose that most of molecular function in cancer tissue trying to maintenance cellular homeostasis, keeping cell regeneration and repair, so tissues undergo stress try to survive. It can be also concluded that proteins that are detected in this study; aldolase A, fructose-bisphosphate, keratin 14, formyltetrahydrofolate synthetase and transferrin can be diagnostic or therapeutic biomarkers. In order to prove these results, it is necessary to have further experiments. To be more accurate, it is suggested that these biomarkers should be analyzed in the mucous secretions and blood.

Table 4: Functional Annotation Clustering by using DAVID software

Annotation Cluster 1		Enrichment Score: 1.31	Count	P_Value
	GOTERM_BP_FAT	cellular homeostasis	3	1.6E-2
	GOTERM_CC_FAT	extracellular region part	3	4.8E-2
Annotation Cluster 2		**Enrichment Score: 1**	**Count**	**P_Value**
	GOTERM_BP_FAT	response to wounding	3	2.1E-2
	GOTERM_MF_FAT	metal ion binding	3	6.2E-1
	GOTERM_MF_FAT	cation binding	3	6.3E-1
	GOTERM_MF_FAT	ion binding	3	6.4E-1

ACKNOWLEDGEMENT

This research resulted from MSc thesis of Ghasem Ahmadi.

REFERENCES

1.Lepage C, Bouvier AM, Manfredi S, Coatmeur O, Cheynel N, Faivre J. Trends in incidence & management of esophageal adenocarcinoma in a well-defined population. Gastroenterol Clin Biol 2005; 29(12):1258-63.

2.Kwong KF. Molecular biology of esophageal cancer in the genomics era. Surg Clin North Am 2005; 85(3): 539- 53.

3.Jaskiewicz K, Marasas WF, Lazarus C. Association of esophageal cytological abnormalities with vitamin and lipotrope deficiencies in populations at risk for esophageal cancer. Anticancer Res 1988;8:711-5.

4.Graham S, Marshall J, Haughey B. Nutritional epidemiology of cancer of the esophagus. Am J Epidemiol 1990;131:454-67.

5.Block G, Patterson B, Subar A. Fruit, vegetables, and cancer prevention: a review of the epidemiological evidence. Nutr Cancer 1992;18:1-29.

6.Van Rensburg SJ. Epidemiologic and dietary evidence for a specific nutritional predisposition to esophageal cancer. J Natl Cancer Inst 1981;67:243-51.

7.Cook-Mozaffari PJ, Azordegan F, Day NE. Oesophageal cancer studies in the Caspian Littoral of Iran: results of a case-control study. Br J Cancer 1979;39:293-309.

8.**Schrump** DS, Nguyen DM. Novel molecular targeted therapy for esophageal cancer. J Surg Oncol 2005 ; 92(3): 257-61.

9.Feng B, Jiang YG, Wang RW, Fan SZ. Transcriptional expression of p73 gene in esophageal cancer tissue sand its significance. China Onc 2002;12(3): 201–3.

10.Nair KS, Naidoo R, Chetty R. Microsatellite analysis of the APC gene and immunoexpression of E-cadherin, catenin, and tubulin in esophageal squamous cell carcinoma. Hum Pathol 2006;37(2): 125-34.

11.Xing EP, Nie Y, Wang LD, Yang GY, Yang CS. Aberrant methylation of p16INK4a and deletion of p15INK4b are frequent events in human esophageal cancer in Liinxian, China. Carcinogenesis 1999;20(1):77–84.

12.Schrump DS, Nguyen DM. Novel molecular targeted therapy for esophageal cancer. J Surg Oncol 2005 ; 92(3): 257-61.

13.Jiang W, Zhang YJ, Kahn SM, Hollstein MC, Santella RM, Lu SH, Harris CC, Montesano R, Weinstein IB. Altered expression of the cyclin D1 and retinoblastoma genes in human esophageal cancer. Proc Natl Acad Sci USA 1993;90:9026–30.

14.Kazemi Noureini S, Ziaee AA, Rastegar-Jazii F, Yazdanbod M. Telomerase activity in Iranian patients with esophageal squamous cell carcinoma, Med J of the Islamic Republic of Iran 2004; 18:147-52.

15.Zhou G, Hongmei L, Decamp D. 2d differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers, Molec and Cell Proteomics 2002; 1: 117-23.

16.Sasse J, Gallagher SR., Staining proteins in gels. Curr Protoc Immunol 2004;Chapter 8:Unit 8.9.

17. Larsen MR, Cordwell SJ, Roepstroff P. Graphite powder as an alternative or reversed-phase material for desalting and concentration of peptide mixtures prior to matrix-assisted laser desorption/ionization-

mass spectrometry. Proteomics 2002;2: 1277-87.

18.Kuramitsu Y, Nakamura K. Proteomic analysis of cancer tissues: shedding light on carcinogenesis and possible biomarkers. Proteomics 2006;6(20):5650-61.

19.Zhou G, Li H, DeCampD, Chen S, Shu H, Gong Y, Flaig M, Gillespie J W, Hu N, Taylor PR, Michael R. Differential In-gel Electrophoresis for the Identification of Esophageal Scans Cell Cancer-specific Protein Markers. Molecular & Cellular Proteomics 2002; 1, 117-23.

20.Zhang J, Wang K, Zhang J, Liu SS, Dai L, Zhang JW. Using Proteomic Approach to Identify Tumor-Associated Proteins as Biomarkers in Human Esophageal Squamous Cell Carcinoma. J. Proteome Res 2011, 10 (6), pp 2863–2872

21.Nomura F, Byori R. Clinical proteomics in laboratory medicine. Rinsho Byori 2006;54(4):413-20.

22.Osei-Boateng M. Nutrition rehabilitation in the promotion of child health in Ghana. Int Nurs Rev 1979;26(5):133-6.