FINAL PROJECT REPORT FOR
IRPA EA GRANT

PROJECT TITLE:

STUDY OF THE PROTEIN PROFILE OF COLON CANCER CELLS.

PROJECT NUMBER:

06-02-05-3197EA12

PROJECT LEADER:

DR. GAM LAY HARN

CONTENT

Summary of the Project

Colorectal cancer is a worldwide public health concern and the frequent cause of mortality and morbidity in developed, developing and industrialized countries. In Malaysia, colorectal cancer ranks second most common cancers. The vast majority of colorectal cancers are adenocarcinomas, which arise from preexisting adenomatous polyps that develop in the normal colonic mucosa. In the present study, the sequential protein extraction, separation and identification were applied in order to study the proteome of colorectal cancer proteins. In addition, the unique and differentially proteins that were isolated from the human colon (normal and cancer) tissues were analyzed using LC/MS/MS. In the sequential protein extraction technique, three types of protein extraction buffers were used to extract proteins from different cellular compartment of colon tissue. They were T1 buffer, S2 buffer and S3 buffer. The extraction buffers were always used in the orderly sequence of T1 buffer followed by S2 buffer and then S3 buffer. The proteins extracts were separated by means of electrophoresis and the protein bands were visualized after being stained with Coomassie Blue or Silver. The image of gels was then captured by high resolution scanner. It was found that the expression of lower molecular weight proteins (below 37 kDa) in cancer tissues is greater than that of the normal tissue. We noticed the hydrophilic proteins are relatively easier to digest by trypsin as compared to the hydrophobic protein. In T1 buffer, there were a total of thirty-six of the down-regulated proteins, thirty-one of the up-regulated proteins were identified. In addition, six unique proteins from cancer tissues and six unique proteins from

normal tissues were also identified in T1 buffer. In S2 buffer, a total of two down-regulated protein and eight up-regulated proteins were identified. In S3 buffer, eight of the up-regulated proteins from colon cancer tissues were identified. The functions of the proteins identified in this study that were extracted from both normal and cancer tissues consisted of protein-binding proteins, which made up 26 % and 22 % of the identified proteins from normal and cancer tissues, respectively. These proteins are 14-3-3 protein family and immunoglobulin family; Structural proteins, which made up 18 % and 36 % of the identified proteins from normal and cancer tissues, respectively. These proteins are actin, desmin, vimentin, and cytokeratins family; Transport proteins, which made up 16 % and 15 % of the identified proteins from normal and cancer tissues, respectively. These proteins are serum albumin precursor and hemoglobin alpha, beta and epsilon chain; Hypothetical protein groups, which made up 16 % and 4.3 % of identified proteins from normal and cancer tissues, respectively; Enzymes or catalytic proteins, which made up 16 % and 17 % of identified proteins from normal and cancer tissues, respectively. In addition, a small number of proteins identified were cytoskeleton-binding proteins, endopeptidase inhibitor proteins, motor activity proteins, channels protein, defense response protein, DNA-RNA binding protein, lipoprotein and ATP-binding protein.

Due to time constrain, not all the differentially expressed and unique proteins that were identified in this study was subjected to LC/MS/MS analysis. However, the presently developed method demonstrates its capability to extract, separate,

analyze and identify proteins from human colon tissue. The proteins that were identified may serve as candidates for future development of vaccine or drug-target therapeutic against colon cancer. Nevertheless, in order to achieve this, much larger specimen number is required to define the significance of the target proteins in human colon cancer.

Introduction

Cancer that begins in the colon is called colon cancer whilst cancer that begins in the rectum is called rectal cancer. Cancers affecting either of these organs may also be called colorectal cancer (National Cancer Institute, 2004). Colorectal cancer is a worldwide public health concern and it is among the best characterized cancers with regard to genetic progression of the disease (Fearon, 1990). Environmental factors and host immunological characteristics could also contribute to initiation and progression of this cancer (Harkins *et al.*, 2002).

Colorectal cancer is also a frequent cause of mortality and morbidity in developed, developing and industrialized countries (Fearon, 2001). In Malaysia, cancer of colorectal ranks second most common cancers and this has also been the case for both male and female. The Chinese had the highest number, followed by Indians and Malays had the lowest number of cases. It was reported that 9.6 % of the patients were at stage I, 31.7% were at stage II, 33.7% were at stage III and 25% were at stage IV (Zarihah *et al.*, 2000). Below age of 60 years, the disease occurred almost equally among male and female. Thereafter, the frequency in males rose rapidly. The incidence of colon cancer is increased exponentially with age (Chan *et al.*, 1994; Hooi and Devaraj, 1997; Penang Hospice, 2003).

Although early detection of colon cancer is important, for many individuals the detection does not occur. Almost one-half of all patients with colorectal cancer die of metastasis disease (Brunagel *et al.*, 2002). Patients with stage II (Dukes' stage B) colorectal cancer have a 5-year survival of 70-80% after curative surgery. For stage III (Dukes' stage C), which involves the regional lymph nodes, the 5-year survival rate decreases to 30-50%. Patient who with stage III disease can be treated with adjuvant chemotherapy. Some patients with stage II (Dukes' stage B) disease are at high risk for recurrence and receive adjuvant chemotherapy (O'Connell e*t al.*, 1992).

A screening policy for early diagnosis of colorectal cancer has not been definitely adopted. Moreover, at the time of diagnosis, some 25% of patients already have liver metastases (Ponz de Leon *et al.*, 2000; Blumgart and Fong, 1995). Advances in early detection and surgery have been largely responsible for reducing mortality and morbidity of colon cancer.

Colorectal cancers develop through a series of histological distinct stages from adenoma to carcinoma (Muto *et al.*, 1995). Recently, the multi-step model of colon cancer development was proposed by Fearon and Vogelstein (Fearon and Vogelstein, 1990). The mechanism was revised and presented in greater detail to include the interdependence of different pathways and involvement of many more gene mutations than before (Bodmer, 1996 and Iiyas *et al.*, 1999). It is believed that mutations in certain genes set the stage for initiation and transformation of normal colonic epithelial cells. Further accumulation of mutations in other genes contributes to the progression of cancer through adenoma – carcinoma – metastasis stages (Natayan and Roy, 2003). Each gene codes certain proteins, which were either expressed or suppressed when the cell turns cancerous.

The vast majority of colorectal cancers are adenocarcinomas, which arise from preexisting adenomatous polyps that develop in the normal colonic mucosa. This

adenoma-carcinoma sequence is a well-characterized clinical and histopathologic series of events with which discrete molecular genetic alterations have been associated. DNA methylation changes are a relatively early event and have been detected at the polyp stage. Colorectal cancers and polyps have an imbalance DNA methylation, which are global hypomethylation and regional hypermethylation. Hypomethylation can lead to oncogene activation, whereas hypermethylation can lead to silencing of tumor suppressor genes. *ras* gene mutations are observed commonly occur in larger polyps but not in smaller polyps, suggesting a role for this oncogene in polyp growth (El-Deiry, 2005). Carcinomas are found in as many as 4% of neoplastic polyps. Cells must accumulate 4 to 5 molecular defects, including activation of oncogenes and inactivation of tumor suppressor genes, to undergo malignant transformation. In normal mucosa, the surface epithelium regenerates approximately every 6 days. Crypt cells migrate from the base of the crypt to the surface, where they undergo differentiation, maturation, and, ultimately, lose the ability to replicate. In adenomas, several genetic mutations alter this process, starting with inactivation of the adenomatous polyposis coli (*APC*) gene, allowing unchecked cellular replication at the crypt surface. With the increase in cell division, further mutations occur resulting in activation of the K-*ras* oncogene in the early stages and *p53* mutations in later stages. These cumulative losses in tumor suppressor gene function prevent apoptosis and give the cell eternal life (Cirincione and Cagir 2004).

The stages of colorectal cancer

Staging is an assessment of the extent of the spread of a particular tumor. One of the more commonly used systems is that established by the International Union Against Cancer (Franks and Teich, 1986). Stage 0 cancer is found only in the innermost lining of the colon or rectum, it is known as Carcinoma *in situ*. For Stage 1 or Dukes' A colorectal cancer is assigned when the cancer has grown into the inner wall of the colon or rectum. However, the tumor has not reached the outer wall of the colon or extended outside the colon. During the stage 2 or

Dukes' B, the tumor extends more deeply into or through the wall of the colon or rectum, it may have invaded nearby tissue but cancer cells have not spread to the lymph nodes. Stage 3 or Dukes' C is the stage where cancer has spread to the nearby lymph nodes, but not to other parts of the body. Stage 4 or Dukes'D is most critical cancer staging because the cancer has spread to other parts of the body, such as the liver or lungs. In certain situation, the cancer has been treated however it is returned after a period of time when the cancer could not be detected. The disease may return in the colon or rectum, or in another part of the body. In this case, it is determined as recurrent cancer (National Cancer Institute, 2004).

Tumor (cancer) grade

Tumor grade is a system used to classify cancer cells in terms of how abnormal they look under a microscope and how quickly the tumor is likely to grow and spread. Many factors are considered when determining tumor grade, including the structure and growth pattern of the cells. The specific factors used to determine tumor grade vary with each type of cancer. Histological grade, also called differentiation, refers to how much the tumor cells resemble normal cells of the same tissue type. In cancer, refers to how mature (developed) the cancer cells are in a tumor. Differentiated tumor cells resemble normal cells and tend to grow and spread at a slower rate than undifferentiated or poorly differentiated tumor cells, which lack the structure and function of normal cells and grow uncontrollably. Based on the microscopic appearance of cancer cells, tumor grade is classified into four degrees of severity, which are Grades 1, 2, 3, and 4. The cells of Grade 1 tumors resemble normal cells, and tend to grow and multiply slowly. Grade 1 tumors are often well-differentiated or low-grade tumors, and are generally considered the least aggressive in behavior. Conversely, the cells of Grade 3 or Grade 4 tumors are usually poorly differentiated or undifferentiated high-grade tumors, and are generally the most aggressive in behavior. In addition, they do not look like normal cells of the same type. Thus, Grade 3 and

4 tumors tend to grow rapidly and spread faster than tumors with a lower grade (AJCC, 2002).

Table : The guidelines for grading tumors (AJCC, 2002)

Cancer Grade	Cell description
GX	Grade cannot be assessed (Undetermined grade)
G1	Well-differentiated (Low grade)
G2	Moderately differentiated (Intermediate grade)
G3	Poorly differentiated (High grade)
G4	Undifferentiated (High grade)

Proteomics

The proteome, first formalized in 1995, designs all the proteins expressed by the genome of a cell, tissue, or organism. Complementing the field of genomics, proteomics is designed to elucidate both protein levels and posttranslational modifications in different cell types under different physiological conditions (Pandey and Mann, 2000). In recent years, proteome analysis has gained great interest and is presently use as a modern tool in drug discovery to determine the biochemical processes involved in diseases (Fung *et al.*, 2000), to monitoring cellular processes (Blackstock and Weir, 1999) and for characterizing protein expression levels (Glocker *et al.*, 2000). The comparative characterization of protein patterns in tissues or body fluids of healthy individuals and patients has the potential to serve as basis for new diagnostic tools and the design of disease specific therapies. Gel eletrophoresis (2-DE) (O'Farrell, 1975) followed by in-gel proteolytic digestion and mass spectrometric analysis (Mann and Talbo, 1996; Jensen *et al.*, 1999) has become a powerful method for the identification of proteins present in complex mixtures.

The progress in genomic sequencing, the availability of Internet databases, and the development of mass spectrometry-based strategies for protein identification,

heralded the use of proteomic analysis in biomedicine (Banks *et al.*, 2000). The use of proteomics for analyzing signal transduction pathways that lead to colon cancer (Friedman *et al.*, 2004) and breast cancer cell development has now become a reality and provides the knowledge base for the identification of therapeutic targets and the development of new anti-cancer drugs (Hondermarck *et al.*, 2001).

Objectives of the projects:

In the recent years, protein complement of the proteome has evolved significantly as a result of improved technologies for electrophoresis and mass spectrometry for protein identification. Through these technologies, information on prognosis of colon cancer (Stierum *et al.*, 2003) can be obtained. In fact of the advancement of LC/MS/MS technology, proteomic analysis on colon cancer is one of the leading technologies that applied in cancer research. These protein identifications provide the basis for future comparisons with the proteomes of purified colon cancer or tumor cells. There is a great lack of information on the proteome for Malaysian colorectal cancer patients. Thus, this study is aim to analyze the proteome of colorectal cancerous and normal tissues obtained from Malaysian colorectal cancer patients with the objectives of:

1. To develop a sequential protein extraction technique that is able to extract proteins from normal and cancer tissues.
2. To determine the protein maps of human colon normal tissues and human colon cancer tissues.
3. To identify differentially expressed protein between normal and cancerous tissue.
4. To analyze the target proteins using LC/MS/MS analysis.

Methodology

Samples collection

Human colon normal tissues and colon cancer tissues were provided by Hospital Pulau Pinang. The normal tissues collected are the tissues adjacent to the tumor tissue, which is the tissue of the mucosa layer. The normal tissues were confirmed normal by the hospital's pathologist. The tissues were collected after informed consent from the patients was obtained. The tissues were cut into similar sizes, weighed and immediately stored at –70°C before analyzed. A total of 20 sets of tissues comprising of normal and cancerous tissues from 20 Malaysian patients were analyzed.

Tissue lysis with sequential extraction

Equal weight of tissues (normal and cancer) was subjected to analysis. The deep-frozen colon tissue specimens were disrupted by grinding in a liquid nitrogen-cooled mortar until became powder-form. The powder form-like tissues were aliquot in separated eppendoft tubes. Sequential extraction of protein using three protein extraction buffers (OPE buffers A) was carried out. Whereby, proteins were sequentially extracted with buffer in increasingly protein solubilizing strength. These extraction buffers were prepared according to BIO-RAD Manual (BIO-RAD, 2002) with modification. The extraction procedures are described in the following sections.

Tris Extraction Buffer:

The concentration of Tris Extraction Buffer (T1) is 40mM Tris. 1000 µl of T1 extraction buffer was added to 500 mg of homogenized tissue. The mixture was vortexed for 30 seconds and centrifuged at 20°C with the speed of 12,000 x g for 8 minutes. The supernatant was collected in aliquots (50 µl) and kept in –20°C. The pellet was recovered for the second step of the sequential extraction.

Solution Two Extraction Buffer

Solution Two Extraction Buffer (S2) contains 8 M Urea, 50 mM DTT, 4% (w/v) CHAPS, 0.2% (v/v) Carrier ampholytes (pH 3-10) and 0.0002% (w/v) Bromophenol Blue. Composition of S2 solution is shown in table 2.2. The recovered pellet was washed once with the T1 solution, it was then vortexed, centrifuged and the supernatant was discarded. After washing, a volume of 1000 µl S2 solution was added to the pellet. The mixture was then vortexed for 30 seconds and then centrifuged at 20°C with the speed of 12,000 x g for 8 minutes. The supernatant was collected in aliquots (50 µl) and kept at –20°C. The pellet was again recovered for the third step of the sequential extraction.

Solution Three Extraction Buffer

Solution Three Extraction Buffer (S3) contains 5 M Urea, 2 M Thiourea, 50 mM DTT, 2% (w/v) CHAPS, 2% (v/v) Triton, 0.2% Carrier ampholytes (pH 3-10), 40 mM Tris and 0.0002% Bromophenol Blue. The composition of S3 solution is shown in table 2.3. The recovered pellet from S2 extraction was washed twice with the S2 solution, it was then vortexed and centrifuged, and the supernatant was discarded. After washing, a 1000 µl of S3 solution was added to the pellet. The mixture was again vortexed for 30 seconds and then centrifuged at 20°C at the speed of 12,000 x g for 8 minutes. The supernatant was collected in aliquots (50 µl) and kept at –20°C. The pellet was then discarded or kept in -20 °C for another experiment. A total of 20 set of specimens made up of normal and cancerous colonic tissues was collected from 20 patients.

Protein Concentration Determination

Ultraviolet spectrophotometer (MODEL***) was used in the determination of protein concentration. The absorbance (*A*) at 280 nm and 260 nm wavelengths were measured and the following formula is applied;

$$\textit{Protein } (mgcm^{-3}) = 1.55\, A_{280} - 0.76\, A_{260}$$

According to the above equation, A_{280} and A_{260} stand for total protein and total nucleic acid respectively. A 1000 times dilution was carried out on the protein samples, whereby every 1 μl of sample was diluted in 999 μl of 0.01 M phosphate buffer. The absorbance (zero against blank) for each sample was taken three times and then an average of the readings was calculated.

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed as described by the modified method of Laemmli (Laemmli, 1970). The gel percentages were 7.5%, 10% and 15% polyacrylamide gel in a vertical slab gel apparatus. Two sizes of polyacrylamide gel were used. They were the small gel with the size of 8.3 x 7.4 cm and the big gel with the size of 20 x 16 cm.

The preparation of protein samples for SDS-PAGE separation was carried out by adding 20% (v/v) of sample buffer (0.5M Tris-HCl, 10% (v/v) glycerol, 0.02% (w/v) SDS, 0.1% (w/v) Coomassie blue) to each extracted protein sample (5 μl) and vortexed for 30 seconds. The samples (normal and cancer) were then loaded into individual wells of SDS-polyacrylamide gel using a syringe. Meanwhile, 5 μl of unstained protein molecular weight markers was boiled at 100°C for 15 minutes and then loaded into a separate well.

For 8.3 x 7.4 cm gel, electrophoresis was performed using vertical electrophoresis slab gel apparatus at a constant voltage of 200 volts through out the electrophoresis process. For 20 x 16 cm gel, the gel was run at a constant voltage of 200 volts when the samples were in the stacking gel. When the dye front reached the resolving gel, voltage was increased to 245 volts. Electrophoresis was terminated when the dye front was 2 to 3 mm away from the bottom edge of the gel for both experiments.

2-D Gel Electrophoresis

Protein concentration of the Thiourea buffer extracts was determined using *RC DC* protein assay kit (BIORAD), and 500 µg of each extracts in 185 µl of rehydration buffer (same compositions as Thiourea buffer) was allowed to rehydrate into each IPG strips (4-7 pH range, 11 cm) for 15 hours and focused using IEF Cell (BIORAD) starting from 0 to 250 V within 15 min, followed by 250 to 8000 V within 2.5 hours and maintained at 8000 V until 35000 V-hr was achieved. Subsequently, the IPG strips were equilibrated for 15 min with gentle shaking in SDS-PAGE Equilibration Buffer I (6 M urea, 0.375 M Tris pH 8.8, 2% SDS, 20% glycerol, 2% DTT and a trace amount of bromophenol blue), followed by another 15 min of gentle shaking in SDS-PAGE Equilibration Buffer II (same composition as SDS-PAGE Equilibration Buffer, but 2.5% iodoacetamide was used instead of 2% DTT). Second dimension separation was carried out under constant voltage of 200 V for approximately 3 hours in 10% SDS-PAGE.

Gel washing and staining

At the completion of electrophoresis, the gel sandwich was disassembled. The stacking gel was discarded and the resolving gel was washed and soaked twice in distilled water for 10 minutes each.

Coomassie blue staining

The polyacrylamide gel was stained in coomassie blue staining solution consisting of 0.2 g Coomasie Blue R250, 5% (v/v) methanol and 2% (v/v) glacial acetic acid. The gel was stained for 2 hours (for 8.3 x 7.4 cm) or overnight (for 20 x 16 cm). During destaining, the coomassie blue stain solution was discarded and the destaining solution consists of 50% (v/v) methanol and 10% (v/v) acetic acid was added until sufficiently cover the gel. After 20 minutes, the solution was replaced by the new destaining solution and then the gel was continued to soak overnight in the destaining solution.

Silver staining

Polyacrylamide gel was stained with Silver Stain according to the method provided by the manufacturer (BIO-RAD). The gel was fixed twice with fixative solution (10% v/v ethanol, 5% v/v acetic acid). After 15 minutes, the fixative solution was discarded and oxidizer solution (10% v/v) was added, the gel was soaked in the solution for 5 minutes. The polyacrylamide gel was then washed with deionized water for 5 minutes, which was repeated for 5 times. Silver reagent (10% v/v) was added to stain the polyacrylamide gel for 20 minutes. Again, the gel was washed with deionized water. This was followed by adding of the developer solution (3.2% w/v) to the gel until the developer solution turned into brown "smokey" color. It was then replaced by another additional of fresh developer solution. Finally, the stopping solution (5% v/v acetic acid) was added to polyacryalamide gel.

Imaging

After staining, the image of the gel was captured using High Resolution Scanner (Canon). All digitized captured images were then stored as JPEG format

In-Gel Digestion

The SDS-Polyacrylamide gel was washed thoroughly in 100 mM NH_4HCO_3 and the differentially expressed protein bands from either normal or cancer colon tissues were then excised from the gel. In-gel digestion using trypsin was performed according to Gam and Aishah, (2002) with slight modification. The gel pieces were first excised and shrunk by dehydration in acetonitrile. The solvent was then discarded and the gel pieces were dried in a vacuum centrifuge. A volume of 10 mM dithiotreitol (DTT) in 100 mM NH_4HCO_3 sufficient to cover the gel pieces was added and the protein was reduced for 1 hour at 56 °C. After cooling to room temperature, the DTT solution was replaced with a same volume of 55 mM iodoacetic acid in 100 mM NH_4HCO_3. After 45 minutes incubation at ambient temperature in the dark with occasional vortexing, the gel pieces were washed with 50-100 μl of 100 mM NH_4HCO_3 for 10 minutes, dehydrated with

acetonitrile, rehydrated in 100 mM NH_4HCO_3 and dehydrated in the same volume of acetonitrile. The liquid phase was removed and the gel pieces were dried in a vacuum centrifuge. The gel pieces were swollen in digestion buffer containing 50 mM NH_4HCO_3, 5 mM $CaCl_2$, and 12.5 ng/μl of tripsin in an ice-cold bath. After 45 minutes, the supernatant was removed and replaced with 10 μl of the same buffer but without trypsin to keep the gel pieces wet during enzymatic cleavage at 37 °C overnight. Peptides were extracted from the gel matrix by adding 15 μl of 20 mM NH_4HCO_3, vortexed and incubated at room temperature for 10 minutes and the supernatant was recovered after a brief spin. This was followed by adding (1 to 2 times the volume of gel pieces) 5% (v/v) formic acid in acetonitrile:water mixture (70:30), vortexed and incubated for 20 minutes at room temperature. It was then spun down and the supernatant was recovered. These steps were repeated 3 times. Pooled extracts were dried down in a vacuum centrifuge and stored at -20 °C. Each of the pooled extracts (peptides) was added with 30 μl of deionized water. Then, they were centrifuged at 500 X g for 5 minutes at 15 °C. The extracts were then subjected to LC/MS/MS analysis or stored at -20 °C.

Mass spectrometric analysis

A volume of 5 μl of the sample was injected into a RPC-column (C_{18} 300 A, 5 μm, 1 mm X 150 mm) connected to a Hewlett Packard HPLC. A capillary pump was used to pump the mobile phase at 20 μl/min flow rate, the linear gradient used was 5% B to 95% B in 65 minutes. Mobile phase A was 0.05% Formic acid in deionized water and B was 0.05% formic acid in ACN. The HPLC was interfaced to an ion trap mass spectrometer.

The dry gas temperature was 300°C, dry gas flow rate was of 8.0 L/min, nebulizer pressure of 30.0 psi. The above parameters were used subsequently for acquiring MS data. The peptides were ionized using electrospray soft ionization technique (ESI).

The experimental method was made up of one experimental protocol consisting of 2 scan events. The first scan event was a full scan MS and the second was the data dependent MS/MS scan which is dependent on the results of the first scan event. The most intense ion in a MS scan will be automatically isolated and excited to MS/MS scan. This linkage is known as data dependent scan.

The parameters set for data dependent scan (MS/MS scan) were default collision energy (voltage) = 1.15V, charge state = 2, minimum threshold = 3000 counts, and the isolation width = 2 m/z.

Mascot Protein Identification

The protein was identified through Mascot Protein Database Search engine. The Peptide Mass Tolerance was set as ± 2 Da and ± 0.8 Da was set for the Fragment Mass Tolerance. Only 1 missed cleavage was allowed and the instrument type that was used in the analysis was ESI-TRAP meaning Electrospray Ionization and ion Trap Mass Spectrometer. This software was directly downloaded from the Internet Website, which is *www.matrix-science.com*. Matches were computed using a probability-based Mowse score defined as -10 × log (*P*), where *P* is the probability that the observed match was a random event (Perkin *et. al.*, 1999). To understand the functions and the particulars about those proteins, SwissProt (*www.expaxy.com*) and Pubmed (*www.ncbi.com*) protein database search were carried out.

Results and Discussion

Sequential protein extraction

In the sequential protein extraction technique, three types of protein extraction buffers were used to extract proteins from different cellular compartment of colon tissue. The sequential extraction technique buffers was named as OPE buffers A. OPE buffers A consists of three extraction buffers, namely Tris Extraction Buffer (T1), Solution Two Extraction Buffer (S2) and Solution Three Extract Buffer (S3). The main ingredient in T1 is 40 mM Tris. Tris extraction buffer was used to extract soluble proteins, such as cytosolic proteins and nuclear proteins (Biorad, 2002). The main compositions of S2 are Urea, DTT, CHAPS and carrier ampholytes. This extraction buffer extracts the protein with intermediate solubility (Biorad, 2002). The contents of S3 are almost similar with S2, with addition of Thiourea, Triton and Tris. This solution is used to extract the proteins (Biorad, 2002) that are insoluble in T1 and S2. The extraction buffers are always used in the orderly sequence of T1 followed by S2 and then S3.

Sequential extraction technique

In general, the usage of sequential extraction buffers is benefited to this study. This is because proteins from different cellular compartments were extracted in fractions. This method not only allows the analysis of proteins according to their localization in the cells, it also reduces the number of protein extracted in single fraction, which allows better isolation of proteins and their visualization in SDS-PAGE.

Figure 1A shows the protein profiles of normal and cancer colon tissues from the four different patients, which were extracted using Tris Extraction Buffer (T1). Figure 1A shows that there are consistent protein bands that were expressed in all the four normal tissues tested. Furthermore, there are also protein bands that are only found in three out of four tissues analyzed. The same observation is also seen in cancer tissues. These consistently expressed bands are represented in diagram 1A. In our present study, only the consistently expressed

bands are targeted for further analysis. The distinct bands that were only detected in one particular tissue sample are ignored, as it may not represent the proteins that involved in cancer formation. Diagram 1A represents the protein bands consistently detected in all the four tissues analyzed. The black colored band represents the protein bands that were detected in all the tissues analyzed. Whereas, the orange colored band represents the protein bands that were found in at least 75% of the tissues. In colon cancer tissues, the black-colored bands were detected at 190 kDa, 177 kDa, 160 kDa, 48.8 kDa, 39 kDa, 35 kDa, 27 kDa, 26 kDa and 24kDa. However, the orange-colored band was detected at 11 kDa. In colon normal tissues, the black-colored bands were detected at 190 kDa, 110 kDa, 48.8 kDa, 43 kDa, 36.3 kDa, 35 kDa, 26 kDa and 17 kDa, while the orange-colored bands are found at 78 kDa, 27 kDa and 24 kDa. When comparing between cancer and normal, it is clear that amongst these proteins, there are some proteins that were differently expressed in either normal or cancer. These differently expressed proteins include unique proteins (expressed only in cancer or in normal) and up-regulated proteins (higher intensity in cancer) and down-regulated protein (lower intensity in cancer). The protein bands that were subjected for further analysis are at 190 kDa, 110 kDa, 48.8 kDa, 36.3 kDa, 35 kDa, 26 kDa and 17 kDa in normal and at 177 kDa, 160 kDa, 48.8 kDa, 27 kDa, 26 kDa and 24 kDa in cancer tissues. . The consistent protein bands in T1, the 48.8 kDa protein bands were frequently found in both normal and cancer tissues. However, the band is more intense in normal as compared to those in cancer tissue.

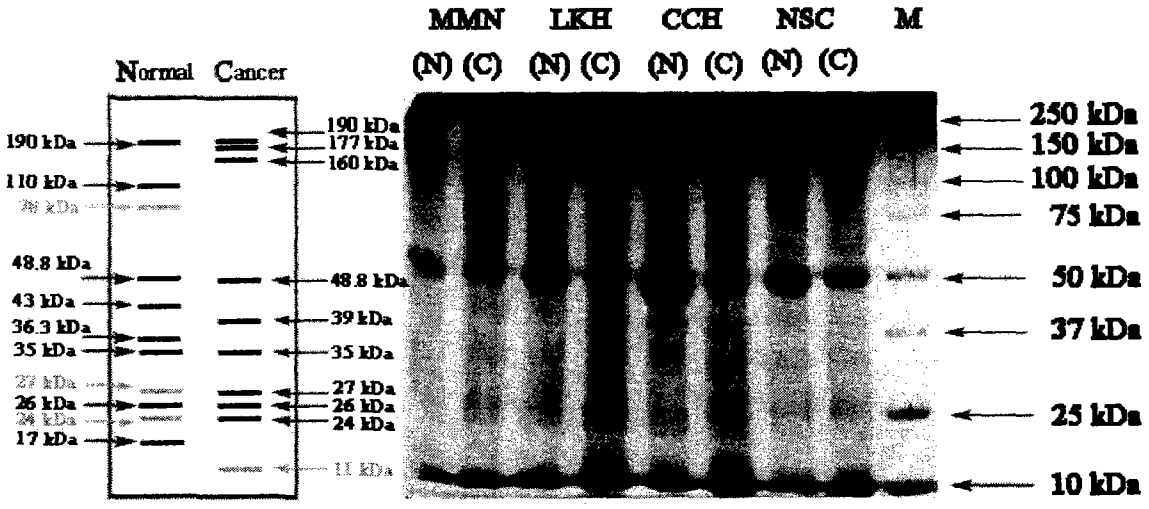

Diagram 1A

Figure 1A : Proteins profiles of normal and cancer colon tissues from four different patients. The proteins were extracted using in Tris Extraction Buffer (T1). N: Normal colon tissues; C: Cancer colon tissues; M: Marker; Lanes 1 and 2: MMN sample; Lanes 3 and 4: LKH sample; Lanes 5 and 6: CCH sample and Lanes 7 and 8: NSC sample. Diagram 1A shows the consistent protein bands found in the normal and the cancer tissues of the four patients.

Figure 1B shows the proteins profiles of proteins extracted using S2 buffer from normal and cancer colon tissues of the four patients. Diagram 1B represents the consistent protein bands expressed in the tissues analyzed, where black bands show the consistent protein bands that were expressed in all the tissues analyzed. Whereas, the orange bands show the consistent protein bands that were expressed in at least 75% of tissues analyzed. For the cancer tissues, the black-bands are at 76 kDa, 38 kDa, 32 kDa and 25 kDa. Furthermore, orange-bands are at 65 kDa, 18 kDa and 12 kDa. For the normal tissues, the black-

bands are at 180 kDa, 65 kDa, 38 kDa, 32 kDa and 25 kDa. Only one orange-band detected at 12 kDa. Amongst the bands, only two bands were targeted at 180 kDa and 32 kDa from normal and four bands at 76 kDa, 65 kDa, 38 kDa and 25 kDa from cancer were randomly choose for further analysis. Some of these bands appeared in a very low intensity, such bands are at 180 kDa for normal and band at 76 kDa for cancer. In S2 buffer, most of the proteins from cancer were found at the low molecular weight region, in the range of 35 kDa to 10 kDa. In contrast, very few protein bands from normal were detected in this region. Furthermore, these bands exist in a very low intensity. From figure 1B, it was found that the proteins extracted in S2 buffer are different from those extracted in T1 buffer. This suggests that proteins in T1 buffer and S2 buffer were originated from different cellular compartment of the tissue. The constituent of S2, which made up of urea and other chaotropic agents may disturb the rugged protein-protein interaction found among structural and cytoskeletal proteins. CHAPS has proved especially useful for the solubilization of aggregating proteins (Hjelmeland, 1990).

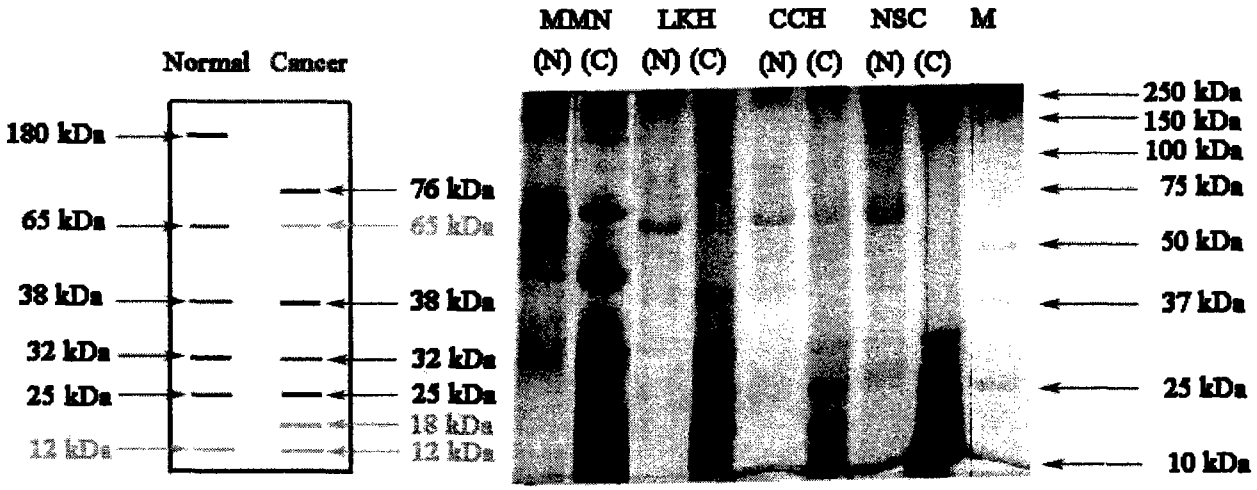

Diagram 1B

Figure 1B : Proteins profiles of normal and cancer colon tissues from four different patients. The proteins were extracted using in Solution Two Extraction Buffer (S2). N: Normal colon tissues; C: Cancer colon tissues; M: Marker; Lanes 1 and 2: MMN sample; Lanes 3 and 4: LKH sample; Lanes 5 and 6: CCH sample and Lanes 7 and 8: NSC sample. Diagram 1B shows the consistent protein bands found in the normal and the cancer tissues of the four patients.

Figure 1C shows the proteins profiles of proteins extracted using S3 buffer from normal and cancer colon tissues of the four patients. Diagram 1 C represents the consistency protein bands expressed in the tissues analyzed. For the cancer tissues, the consistent black-bands are at 38 kDa, 27 kDa, 23 kDa 20 kDa and 17 kDa. The consistent orange-band is at 65 kDa. For the normal tissues, the consistent black-bands are at 38 kDa, 27 kDa and 20 kDa. Furthermore, the consistent orange-bands are at 150 kDa, 100 kDa, 65 kDa, 46 kDa and 35 kDa. When comparing the protein profiles of cancer and normal, it is obvious that the expression of lower molecular weight proteins (below 37 kDa) in cancer tissues is greater than that of the normal tissue. Thus, in order to understand the identity of these proteins, the protein bands at 17 kDa, 20 kDa, 23 kDa and 27 kDa were targeted for further analysis. The constituent of S3

buffer was different with S2 buffer because of the presence of thiouea in S3 buffer. A mixture of urea and thiourea was recommended to enhance the protein solubilization for extraction of more stringent proteins from the tissue (Rabilloud, 1998). In other words, S3 buffer was used to extract the insoluble protein in T1 buffer and S2 buffer.

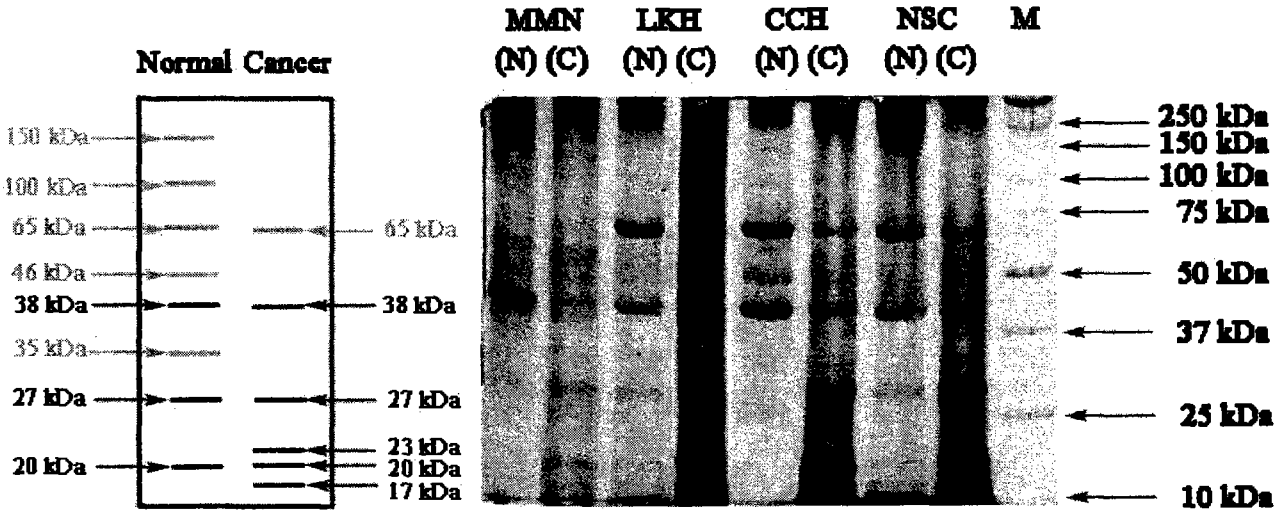

Diagram 1C

Figure 1C : Proteins profiles of normal and cancer colon tissues from four different patients. The proteins were extracted using in Solution Three Extraction Buffer (S3). N: Normal colon tissues; C: Cancer colon tissues; M: Marker; Lanes 1 and 2: MMN sample; Lanes 3 and 4: LKH sample; Lanes 5 and 6: CCH sample and Lanes 7 and 8: NSC sample. Diagram 3.2C shows the consistent protein bands found in the normal and the cancer tissues of the four patients.

Besides the SDS-PAGE analysis, S3 protein extracts was also subjected to 2D gel analysis, the image was captured and analyzed using PDQuest software (BIORAD) and the differentially expressed proteins was subjected to LC/MS/MS analysis, the result is shown in Table 2. Figure 1D shows the 2D gel protein

profiles for cancer and normal, respectively. It is clear that more protein spots was detected in 2D gel separation because it provides a more comprehensive of protein separation where the proteins were separated according to their pI and MW whereas proteins were only separated by difference in MW in SDS-PAGE.

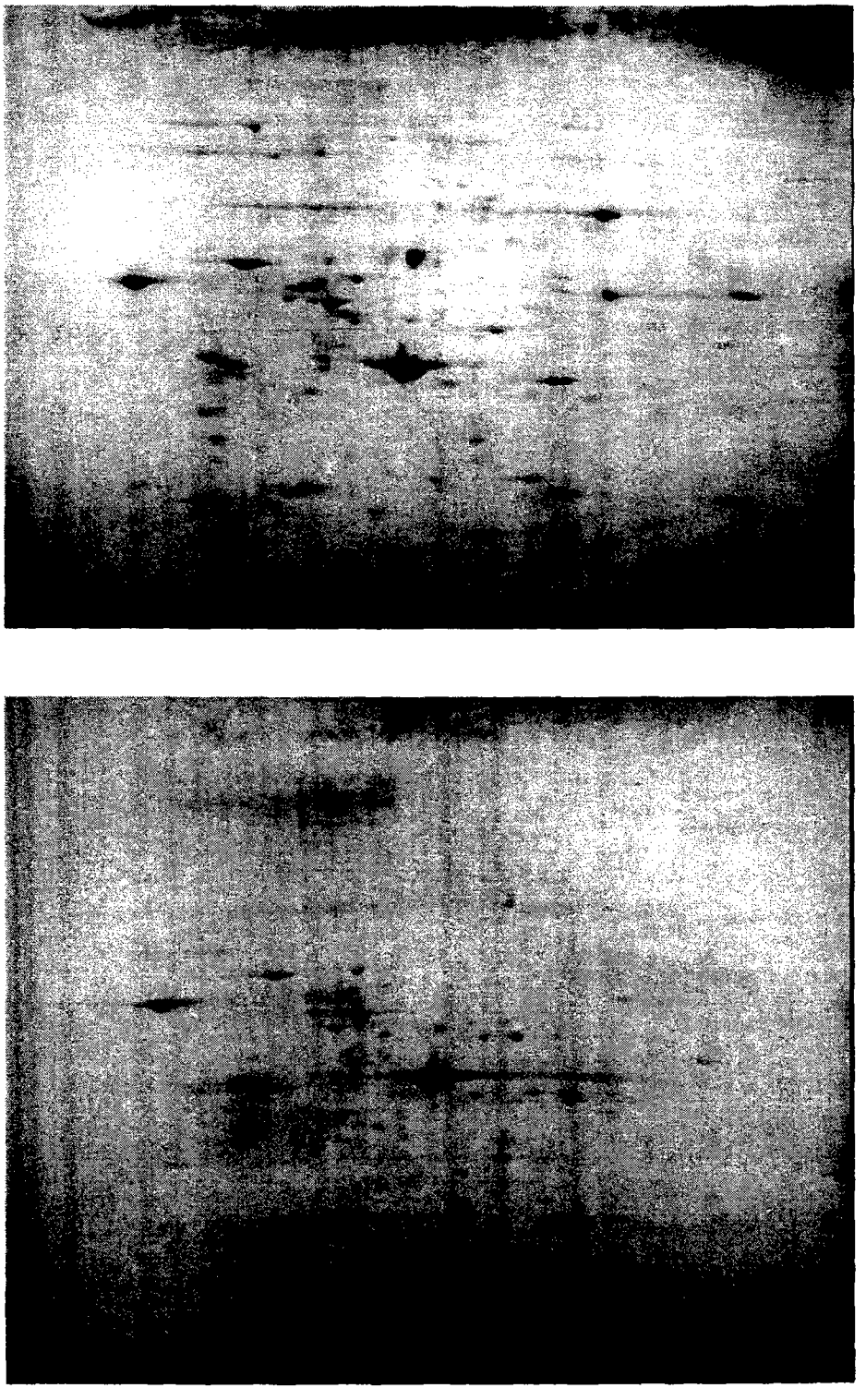

Figure 1D: 2D gel protein profiles for one of the S3 protein extracts; Upper, cancer; Lower, normal

Identification of proteins using LC-MS/MS analysis

Recently, the most common implementation of proteome analysis is the separation of proteins expressed in a tissue by means of gel electrophoresis and the identification of the separated protein bands by MS or MS/MS (Figeys *et al.*, 1998). Generally, protein identification is performed by applying a combination of protein separation, protein digestion, tryptic peptide analysis via tandem mass spectrometry and protein database search engine (Schlosser *et al.*, 2003). In this study, the separation of proteins was employed by means of SDS-PAGE. As a consequent, proteins with different molecular weight could be characterized effectively. The separation of proteins by SDS-PAGE posses certain limitations, such as the low resolving ability of SDS-PAGE as compared to 2-D electrophoresis, which led to overlapping of protein bands in SDS-PAGE. However, the un-resolve proteins that were presence in a single band in SDS-PAGE can be identified individually by means of LC/MS/MS. Whereby, in our present study, some of the Mascot protein database search results indicate the presence of more than one protein in single protein band, which show indicating the overlapping of proteins in the band. Moreover, several limitations inherent by 2-D electrophoresis, such as the difficulty in the detection and identification of the proteins, which are hydrophobic, basic (pI > 9) and large (> 200 kDa) can be analyzed using SDS-PAGE. (Becamel *et al.*, 2002). Therefore, combination efficient separation of proteins from SDS-PAGE with high performance liquid chromatography (LC) coupled to MS/MS (LC-MS/MS) was carried out in this study (Mann *et al.*, 2001).

LC-MS/MS is a highly automated (Ducret *et al.*, 1998), fast and reliable, powerful and sensitive instrument for detection of proteins and peptides (Govorukhina, 2003). In principle, the tryptic peptides generated by proteolytic digestion target protein were separated by liquid chromatography (LC) and analyzed by tandem mass spectrometry (MS/MS). The application of LC separation technique, such as high performance liquid chromatography (HPLC) has been successfully used

in peptides separation (Isaaq *et al.*, 2002). Mass spectrometry, with the development of 'soft' ionization method such as electrospray (ESI), has become important technique for protein characterization (Fenn *et al.*, 1989; Carr *et al.*, 1991). Furthermore, tandem mass spectrometry (MS/MS) was utilized to provide mass and sequence information (Roepstorff and Fohlman, 1984) as well as (partial) primary structures of peptides by a process called collision-induced dissociation (CID) (Hunt *et al.*, 1986). Through MS/MS analysis, a series of product ions were generated providing the amino acid sequence information of the peptide. Using this approach, the identity of the proteins could be determined (Ducret *et al.*, 1996).

In this study, the target protein bands were excised and enzymatically digestion of the protein using an in-gel digest method was carried out (Shevchenko *et al.*, 1996). Following which the tryptic peptides were introduced into a mass spectrometer via HPLC-based separation techniques. A program was created to control the mass spectrometer to switch from automated MS to MS/MS switching mode in order to generate CID spectra of the most intense ion in a MS scan. Once the necessary spectra have been generated, the instrument reverted back to MS mode. This cycle was repeated until the end of the separation protocol. The CID spectra generated during an experiment was used to identify the protein by matching it to the known protein sequences (Gibson and Biemann, 1984) existing in the NCBI non-redundancy protein sequence database using Mascot software (Perkins *et al.*, 1999), which can be found in the following Website (http://www.matrixscience.com). In this study, the identification of proteins was performed by matching experimental MS/MS spectra of its peptides against calculated spectra in the NCBI Human protein database (Becamel *et al.*, 2002). The structural and functional information of the identified proteins were determined via available SWISS-PROT protein database World Wide Web (http://au.expasy.org) and PubMed (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?) protein database search engine.

The Base Peak Chromatogram (BPC) of the tryptic peptides of protein at band 48.8 kDa (T1, normal) is shown in figure 2A. An example of the MS spectrum of the peptide fragment that eluted at 31.6 min retention time is shown in figure 2B. A doubly charged precursor ion with *m/z* value of 917.9 was obtained from the ionized peptide fragment (figure 2B). The product ions spectrum (MS/MS spectrum) for the precursor ion is shown in figure 2D. A peptide with amino acid sequence of VFSNGADLSGVTEEAPLK (figure 2C) was derived from the product ions spectrum shown in figure 3.4D. This peptide sequence was confirmed via a series of b and y product ions. The b product ions are b_7, b_9 and b_{15} ions at *m/z* 685.3, 899.9 and 1460.4, while the y production ions are y_3, y_4, y_5, y_7, y_9, y_{10}, y_{11}, y_{12}, y_{13} and y_{14} ions at m/z 357.3, 428.2, 557.3, 787.4, 944.6, 1030.5, 1143.6, 1258.7, 1329.7 and 1386.6. In the interpretation of peptide sequence, the mass of every b- and y- product ions are important information to determine its sequence, whereby its series of products ions whose mass differences matched amino acid residue masses (Mann and Wilm, 1994). Together with other peptides that were sequenced in the analysis, 5 peptides were found match alpha-1-antitrypsin precursor (NCBI accession number: ITHU) from *Homo sapiens* taxonomy when subjected to Mascot search engine (figure 3A and B). The total score gathered from the five matched peptide were 246, where score > 41 indicates identity or extensive homology. The total sequence coverage of the protein by these peptides was 12 %. The theoretical molecular weight of the protein is 46881, which is relatively close to the experimental molecular weight.

Besides alpha-1-antitrypsin precursor, two other proteins, namely serum albumin precursor and serum albumin were identified from the 48.8 kDa band (figure 4). The scores gathered by the peptides that led to the identification of serum albumin and serum albumin precursor are 1054 and 1087, respectively. The sequence coverage by these peptides is 36% (figure 5 B) and 37 % (figure 6 B)

for serum albumin precursor and serum albumin. Using the same method, identification of other proteins were carried out.

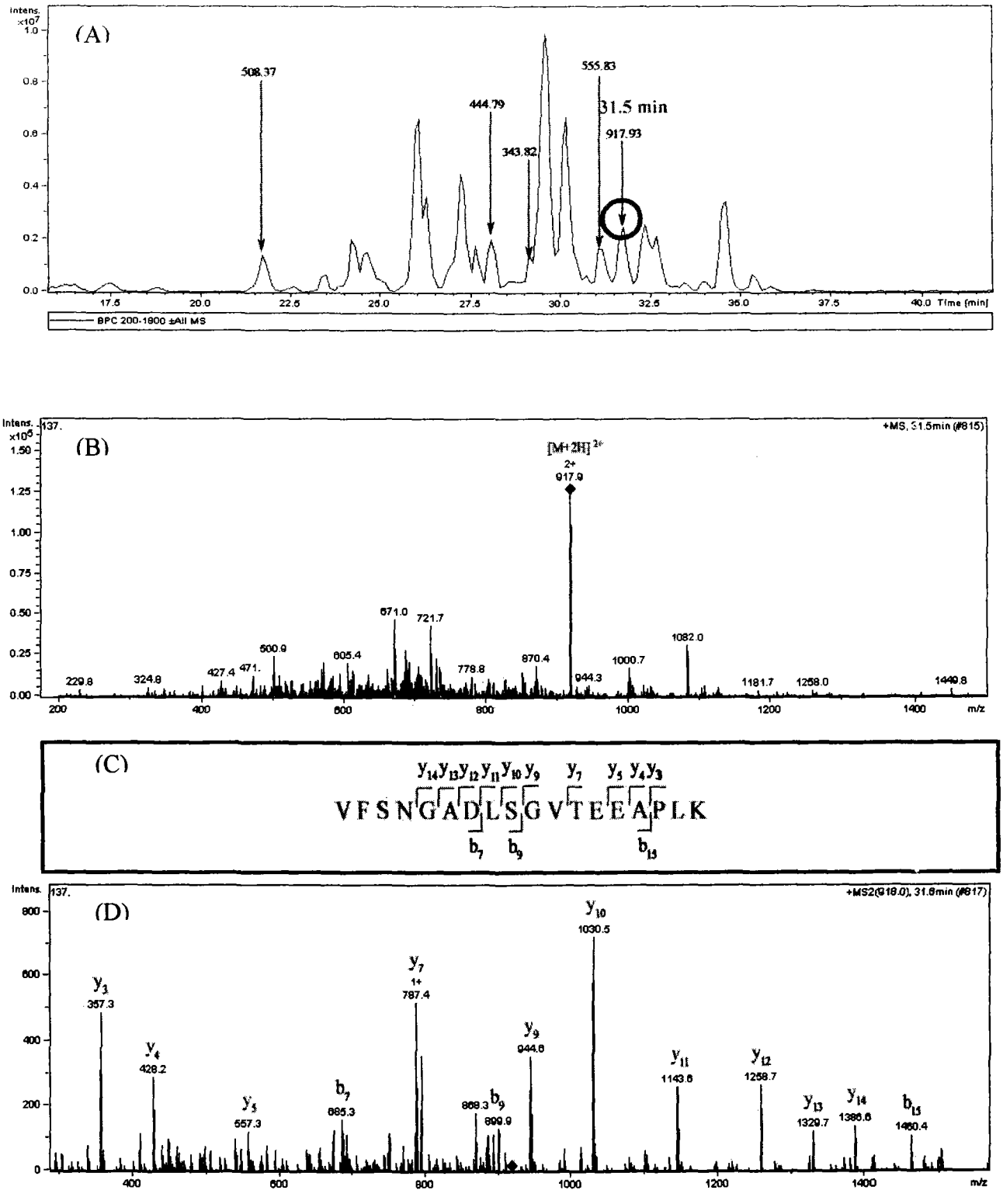

Figure 2: Identification of alpha-1-antitrypsin precursor. (A) represents the base peak chromatogram. (B) represents the full scan MS. (C) represents the amino acid sequence divided from the MS/MS spectra in (D). (D) represents the MS/MS spectrum of 917.93 precursor ion.

ITHU Mass: 46881 Score: 246 Queries matched: 5

alpha-1-antitrypsin precursor [validated] - human

Check to include this hit in error tolerant search

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
☐	36	343.82	685.63	685.44	0.19	0	27	2.1	1	IVDLVK
☐	75	444.79	887.57	887.50	0.07	0	33	0.42	1	AVLTIDEK
☐	98	508.37	1014.73	1014.61	0.12	0	60	0.00074	1	SVLGQLGITK
☐	106	555.83	1109.65	1109.60	0.05	0	58	0.00086	1	LSITGTYDLK
☐	153	917.93	1833.85	1832.92	0.93	0	69	5e-05	1	VFSNGADLSGVTEEAPLK

Figure 3A: Five peptides that confirmed that identity of alpha-1-antitrypsin precursor

{MATRIX} {SCIENCE} **Mascot Search Results**

Protein View

Match to: **ITHU** Score: **246**
alpha-1-antitrypsin precursor [validated] - human
Found in search of E:\HERNG YEE 270204\NT1-53-67.mgf

Nominal mass (M_r): **46881**; Calculated pI value: **5.37**
NCBI BLAST search of ITHU against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens
Links to retrieve other entries containing this sequence from NCBI Entrez:
A1AT_HUMAN from Homo sapiens

Fixed modifications: Carboxymethyl (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **12%**

Matched peptides shown in **Bold Red**

```
   1 MPSSVSWGIL LLAGLCCLVP VSLAEDPQGD AAQKTDTSHH DQDHPTFNKI
  51 TPNLAEFAFS LYRQLAHQSN STNIFFSPVS IATAFAMLSL GTKADTHDEI
 101 LEGLNFNLTE IPEAQIHEGF QELLRTLNQP DSQLQLTTGN GLFLSEGLKL
 151 VDKFLEDVKK LYHSEAFTVN FGDTEEAKKQ INDYVEKGTQ GKIVDLVKEL
 201 DRDTVFALVN YIFFKGKWER PFEVKDTEEE DFHVDQVTTV KVPMMKRLGM
 251 FNIQHCKKLS SWVLLMKYLG NATAIFFLPD EGKLQHLENE LTHDIITKFL
 301 ENEDRRSASL HLPKLSITGT YDLKSVLGQL GITKVFSNGA DLSGVTEEAP
 351 LKLSKAVHKA VLTIDEKGTE AAGAMFLEAI PMSIPPEVKF NKPFVFLMIE
 401 QNTKSPLFMG KVVNPTQK
```

Figure 3B:Mascot search result for the analysis of protein at band 48.8 (T1, normal colon tissue). The protein identified is alpha-1-antitrypsin precursor. Five peptides were found matched the amino acid sequence of alpha-1-antitrypsin. The total sequence coverage of the identified peptides is 12 %.

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 41 indicate identity or extensive homology ($p<0.05$).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

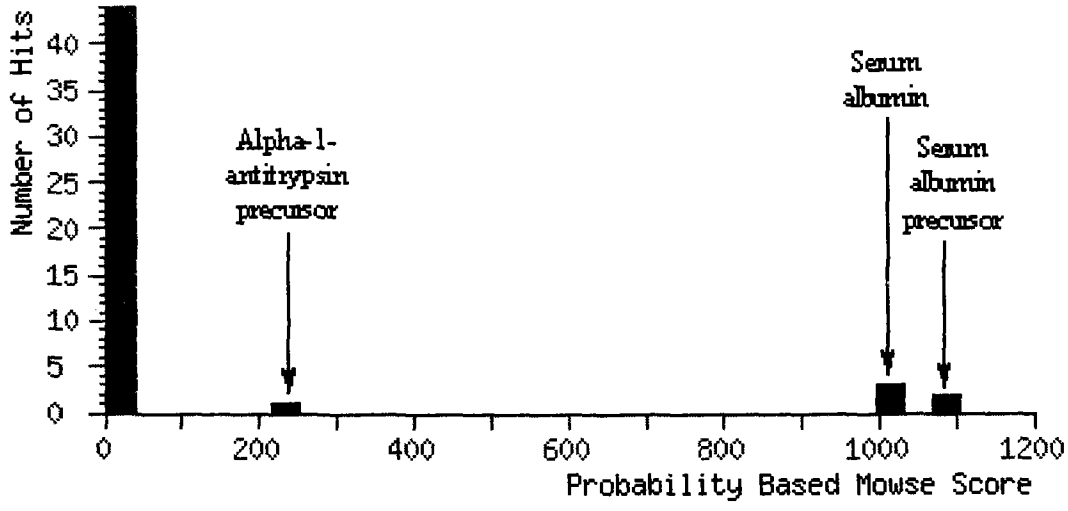

Figure 4: The identification of three different proteins, namely alpha-1-antrypsin precursor, serum albumin precursor and serum albumin from the protein band at 48.8 kDa (T1, normal).

ABHUS **Mass:** 71352 **Score:** 1087 **Queries matched:** 19
serum albumin precursor [validated] - human

☐ Check to include this hit in error tolerant search

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
	45	337.24	672.47	672.37	0.09	0	19	13	3	AWAVAR
	58	348.23	694.45	694.33	0.12	0	10	48	9	NYAEAK
☑	74	387.26	772.51	772.42	0.09	0	45	0.033	1	AACLLPK
☑	78	395.29	788.57	788.46	0.10	0	37	0.18	1	LVTDLTK
	81	401.77	801.53	801.47	0.05	1	2	5.2e+02	7	NLGKVGSK
☑	84	440.75	879.49	879.43	0.05	0	22	3.9	1	AEFAEVSK
☑	87	464.29	926.57	926.49	0.08	0	21	5.3	1	YLYEIAR
☑	88	470.75	939.49	939.44	0.04	0	33	0.26	1	DDNPNLPR
☑	89	476.25	950.49	950.43	0.05	0	22	4.4	1	DLGEENFK
☑	90	480.83	959.65	959.56	0.09	0	49	0.0081	1	FQNALLVR
☑	91	480.84	959.67	959.56	0.11	0	(45)	0.023	1	FQNALLVR
☑	95	492.79	983.57	983.48	0.08	0	14	25	1	TYETTLEK
☑	97	500.85	999.69	999.60	0.09	0	25	2.3	1	QTALVELVK
☑	99	507.35	1012.69	1012.59	0.09	0	54	0.0025	1	LVAASQAALGL
☑	102	528.34	1054.67	1054.58	0.08	1	22	3.7	1	KYLYEIAR
	104	537.79	1073.57	1073.54	0.03	1	5	2.1e+02	3	LDELRDEGK
☑	115	564.88	1127.75	1127.69	0.05	1	18	9.6	1	KQTALVELVK
☑	116	575.33	1148.65	1148.61	0.04	0	23	2.5	1	LVNEVTEFAK
☑	207	829.86	1657.71	1657.73	-0.02	0	17	8.5	1	QNCELFEQLGEYK

Figure 4A: Nineteen peptides that confirmed that identity of serum albumin precursor

MATRIX SCIENCE

Mascot Search Results

Protein View

Match to: ABHUS Score: 1087
serum albumin precursor [validated] - human
Found in search of E:\HERNG YEE 270204\NT1-53-67.mgf

Nominal mass (M_r): 71352; Calculated pI value: 5.92
NCBI BLAST search of ABHUS against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens
Links to retrieve other entries containing this sequence from NCBI Entrez:
ALBU_HUMAN from Homo sapiens

Fixed modifications: Carboxymethyl (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 36%

Matched peptides shown in Bold Red

```
  1 MKWVTFISLL FLFSSAYSRG VFRRDAHKSE VAHRFKDLGE ENFKALVLIA
 51 FAQYLQQCPF EDHVKLVNEV TEFAKTCVAD ESAENCDKSL HTLFGDKLCT
101 VATLRETYGE MADCCAKQEP ERNECFLQHK DDNPNLPRLV RPEVDVMCTA
151 FHDNEETFLK KYLYEIARRH PYFYAPELLF FAKRYKAAFT ECCQAADKAA
201 CLLPKLDELR DEGKASSAKQ RLKCASLQKF GERAFKAWAV ARLSQRFPKA
251 EFAEVSKLVT DLTKVHTECC HGDLLECADD RADLAKYICE NQDSISSKLK
301 ECCEKPLLEK SHCIAEVEND EMPADLPSLA ADFVESKDVC KNYAEAKDVF
351 LGMFLYEYAR RHPDYSVVLL LRLAKTYETT LEKCCAAADP HECYAKVFDE
401 FKPLVEEPQN LIKQNCELFE QLGEYKFQNA LLVRYTKKVP QVSTPTLVEV
451 SRNLGKVGSK CCKHPEAKRM PCAEDYLSVV LNQLCVLHEK TPVSDRVTKC
501 CTESLVNRRP CFSALEVDET YVPKEFNAET FTFHADICTL SEKERQIKKQ
551 TALVELVKHK PKATKEQLKA VMDDFAAFVE KCCKADDKET CFAEEGKKLV
601 AASQAALGL
```

Figure 4B:Mascot search result for the analysis of protein at band 48.8 (T1, normal colon tissue). The protein identified is serum albumin precursor. Nineteen peptides were found matched the amino acid sequence of serum albumin precursor. The total sequence coverage of the identified peptides is 36 %.

1BJ5 **Mass:** 68160 **Score:** 1054 **Queries matched:** 19

serum albumin [validated] - human

☐ Check to include this hit in error tolerant search

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
	45	337.24	672.47	672.37	0.09	0	19	13	3	AWAVAR
	58	348.23	694.45	694.33	0.12	0	10	48	9	NYAEAK
☑	74	387.26	772.51	772.42	0.09	0	45	0.033	1	AACLLPK
☑	78	395.29	788.57	788.46	0.10	0	37	0.18	1	LVTDLTK
	81	401.77	801.53	801.47	0.05	1	2	5.2e+02	7	MLGKVGSK
☑	84	440.75	879.49	879.43	0.05	0	22	3.9	1	AEFAEVSK
☑	87	464.29	926.57	926.49	0.08	0	21	5.3	1	YLYEIAR
☑	88	470.75	939.49	939.44	0.04	0	33	0.26	1	DDNPNLPR
☑	89	476.25	950.49	950.43	0.05	0	22	4.4	1	DLGEENFK
☑	90	480.83	959.65	959.56	0.09	0	49	0.0081	1	FQNALLVR
☑	91	480.84	959.67	959.56	0.11	0	(45)	0.023	1	FQNALLVR
☑	95	492.79	983.57	983.48	0.08	0	14	25	1	TYETTLEK
☑	97	500.85	999.69	999.60	0.09	0	25	2.3	1	QTALVELVK
☑	99	507.35	1012.69	1012.59	0.09	0	54	0.0025	1	LVAASQAALGL
☑	102	528.34	1054.67	1054.58	0.08	1	22	3.7	1	KYLYEIAR
	104	537.79	1073.57	1073.54	0.03	1	5	2.1e+02	3	LDELRDEGK
☑	115	564.88	1127.75	1127.69	0.05	1	18	9.6	1	KQTALVELVK
☑	116	575.33	1148.65	1148.61	0.04	0	23	2.5	1	LVNEVTEFAK
☑	207	829.86	1657.71	1657.73	-0.02	0	17	8.5	1	QNCELFEQLGEYK

Figure 5A: Nineteen peptides that confirmed that identity of serum albumin

Mascot Search Results

Protein View

Match to: 1BJ5 Score: 1054
serum albumin - human
Found in search of E:\HERNG YEE 270204\NT1-53-67.mgf

Nominal mass (M_r): 68160; Calculated pI value: 5.73
NCBI BLAST search of 1BJ5 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens

Fixed modifications: Carboxymethyl (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 37%

Matched peptides shown in Bold Red

```
  1 HKSEVAHRFK DLGEENFKAL VLIAFAQYLQ QCPFEDHVKL VNEVTEFAKT
 51 CVADESAENC DKSLHTLFGD KLCTVATLRE TYGEMADCCA KQEPERNECF
101 LQHKDDNPNL PRLVRPEVDV MCTAFHDNEE TFLKKYLYEI ARRHPYFYAP
151 ELLFFAKRYK AAFTECCQAA DKAACLLPKL DELRDEGKAS SAKQRLKCAS
201 LQKFGERAFK AWAVARLSQR FPKAEFAEVS KLVTDLTKVH TECCHGDLLE
251 CADDRADLAK YICENQDSIS SKLKECCEKP LLEKSHCIAE VENDEMPADL
301 PSLAADFVES KDVCKNYAEA KDVFLGMFLY EYARRHPDYS VVLLLRLAKT
351 YETTLEKCCA AADPHECYAK VFDEFKPLVE EPQNLIKQNC ELFEQLGEYK
401 FQNALLVRYT KKVPQVSTPT LVEVSRNLGK VGSKCCKHPE AKRMPCAEDY
451 LSVVLNQLCV LHEKTPVSDR VTKCCTESLV NRRPCFSALE VDETYVPKEF
501 NAETFTFHAD ICTLSEKERQ IKKQTALVEL VKHKPKATKE QLKAVMDDFA
551 AFVEKCCKAD DKETCFAEEG KKLVAASQAA LG
```

Figure 5B: Mascot search result for the analysis of protein at band 48.8 (T1, normal colon tissue). The protein identified is serum albumin. Nineteen peptides were found matched the amino acid sequence of serum albumin. The total sequence coverage of the identified peptides is 37 %.

Identification of proteins from normal and cancer colon tissues

Using the same approach, the proteins that were identified from both normal and cancer colonic tissues are listed in table 1. The Swiss-Prot accession number, protein description, theoretical molecular weight (MW) and isoelectric point (pI) of the identified proteins are shown in the same table. In our present study, a total of eighty-two proteins were identified from the normal and cancer colonic tissues tested. Herein, seventy-three identified proteins or approximately 89% of the total identified proteins were found at the theoretical masses between 10 kDa to 100 kDa. Only one protein was found at theoretical masses less than 10 kDa, while, eight proteins were at higher than 100 kDa theoretical molecular masses.

In this study, we noticed the hydrophilic proteins are relatively easy to digested by trypsin as compared to the hydrophobic protein. This phenomenon lead to sufficient number of peptides for identification of hydrophilic proteins, and therefore, their identities can always be assigned. However, the gene products that were less frequently detected may be more interesting in proteomics studies because they may carry more biological information that involved in disease-related changes (Fountoulakis and Suter, 2002).

The proteins that were represented by more than one band are described as heterogeneity. Based on the results, the heterogeneity proteins were mainly derived from structural, transport and enzyme subunit such as Actin, gamma-enteric smooth muscle (P12718), Desmin (P17661), Vimetin (P08670), Keratin, type I cytoskeletal 10 (P13645), Cytoskeletal 9 (P35527), Hemoblobin beta chain (P68871), Hemoglobin beta chain (P02023), Hemoglobin (P01922) and Triosephosphate isomerase EC 5.3.1.1 (P60174). Most of the pathways and systems of these proteins were regulated by post-translational modifications with the attachment of specific ligands. The heterogeneity of the proteins can be influenced by protein phosphorylation. Phosphorylation of specific proteins had been recognized as a key reversible regulatory of many proteins (Figeys *et al.*,

1998). In addition, glycosylation (Duvet *et al.*, 2002) and limited proteolytic cleavage are also been reported to be responsible for the heterogeneity of some proteins (Fountoulakis *et al.*, 2004). Phospholylation, glycosylation and limit proteolytic activity has been accounted on the major modification of proteins and therefore will affect the mobility of the modified proteins in SDS-PAGE

Table 1: Up-regulated, down-regulated proteins and unique proteins extracted in T1 buffer from colon tissues

SWISS-PROT accession number	Down-regulated proteins (Normal Tissues)	SWISS-PROT accession number	Up-regulated proteins (Cancer Tissues)
P12718	Actin, gamma-enteric smooth muscle	P04262	Type II, keratin subunit protein
P17661	Desmin	P35900	Keratin 20
P08670	Vimetin	P35908	Keratin, 67k type II epidermal
Q9DFD0	Hypothetical protein	P13647	Keratin k5
Q7Z5W1	Hypothetical protein	Q86Y46	Keratin 6 irs 3
CAD19027	Hypothetical protein	P04264*	Keratin 1, type II cytoskeletal
BAC87457	Hypothetical protein	P35527	Cytoskeletal 9
CAC08866	Hypothetical protein	P04264	Keratin, type II cytoskeletal I
CAC08835	Hypothetical protein	P35527	Keratin 9, type I, cytoskeletal
Q8WTU6	Hypothetical protein	P08779	Keratin, type 1, cytoskeletal 16
Q16167	Serum albumin (Fragment)	P48669	Keratin, type II, cytoskeletal 6F
AAN17825	Serum albumin	P02538	Keratin 6A, type II
Q6LEH2	Serum albumin (Fragment)	P19013	Keratin, type II, cytoskeletal 4
Q01995	Transgelin	P08729	Keratin, type II, cytoskeletal 7
Q91706	Skeletal muscle beta-tropomyosin	P12035	Keratin, type II, cytoskeletal 3
P12277	Creatine kinase	P05783	Keratin, type I, cytoskeletal 18
Q9NPP6	Immunoglobulin heavy chain variant	Q8NF17	Hypothetical protein
CAC10242	Immunoglobulin heavy chain variant	P02100	Hemoglobin epsilon chain
P02760	Ig light chain Fab fragment, antibody a5b7 chain A	1CH4A	Module-substitued chimera hemoglobin beta-alpha mutant
BAA33580	Anti-Hbs AG immunoglobulin Fab Kappa chain	P01857	Ig gamma-1 chain region
Q96SA9	Immunoglobulin kappa light chain	P01859	Ig gamma-2 chain region
P01860	Ig gamma-3 heavy chain disease proteins	AAR32505	Immunoglobulin heavy chain variable region

BAC01749	Immunoglobulin kappa light chain VLJ region		
		P01842	Ig lambda chain
P01782	Ig heavy chain V-II region [ART]	P13645	Keratin, type I, cytoskeleton 10

Table 1 (continued): Up-regulated, down-regulated proteins and unique proteins extracted in T1 buffer from colon tissues

SWISS-PROT accession number	Down-regulated proteins (Normal Tissues)	SWISS-PROT accession number	Up-regulated proteins (Cancer Tissues)
AAR32470	Immunoglobulin heavy chain variable region	P02023	Hemoglobin beta chain
AAR32435	Immunoglobulin heavy chain variable region	P68871	Hemoglobin beta chain
Q9UL88	Immunoglobulin heavy chain Fd region	P01922	Hemoglobin alpha chain
AAS85988	Immunoglobulin heavy chain Fd region	P02768	Serum albumin [Precursor]
P13645	Keratin, type I, cytoskeleton 10	P60174	Triosephosphate isomerase EC. 5.3.1.1
P02023	Hemoglobin beta chain	P00915	Carbonate dehydratase EC 4.2.1.1
P68871	Hemoglobin beta chain	BAB71634	Ig gamma-2 chain region
P01922	Hemoglobin alpha chain		
P02768	Serum albumin [Precursor]		
P60174	Triosephosphate isomerase EC. 5.3.1.1		
P00915	Carbonate dehydratase EC 4.2.1.1		
BAB71634	Ig gamma-2 chain region		
SWISS-PROT accession number	Unique proteins (Normal Tissues)	SWISS-PROT accession number	Unique proteins (Cancer Tissues)
P32119	Preoxiredoxin 2	Q9BX84	LTRPC6 channel kinase6
Q03001	Bullous pemphigoid antigen 1 precursor	Q9AVW8	DNA-directed RNA polymerase II largest chain EC 2.7.7.6
P01009	Alpha-1-antitrypsin precursor	P29312	14-3-3 protein zeta

P20848	Alpha-1-antitrypsin related protein presursor	P31946	14-3-3 protein beta/alpha
1QMBA	Alpha-1-antitrypsin mutant YES, chain A	S38532	Protein 14-3-3 eta chain
P02647	Apolipoprotein	S31975	14-3-3 protein epsilon

Table2: Up-regulated and down-regulated proteins extracted in S2 and S3 buffers from colon tissues

SWISS-PROT accession number	Down-regulated proteins (Normal Tissues)	SWISS-PROT accession number	Up-regulated proteins (Cancer Tissues)
P00915	Actin-gamma	Q8NF17	Hypothetical protein
Q96QE3	ATP[GTP]-binding protein	P02768	Serum albumin [Precursor]
		AAN17825	Serum albumin
		P02675	Fibrinogen beta chain
		P00915	Carbonate dehydratase EC 4.2.1.1
		Q9PTK9	Kinesin-like protein 2
		P35527	Cytoskeletal 9
		P19013	Keratin, type II, cytoskeletal 4
		P05783	Keratin, type I, cytoskeletal 18
		CAE99935	Hypothetical protein
		Q13733	Sodium/potassium-transporting ATPase alpha-4 chain EC 3.6.3.9

Distribution of the functional proteins identified in sequential extraction buffers

Proteins were identified using Mascot protein search engine on the basis of peptide mass matching (Henzel *et al.*, 1993) with redundancy of post-translational modification and proteolysis. Ji *et al.*, (1994) reported that protein patterns of normal human colonic crypts isolated from different regions of large intestine were almost 95 % identical when compared using SDS-PAGE. Proteomic study conducted on the biopsies specimens is not representable as colon cancer biopsies are tissue heterogeneity. Stulik *et al.*, (1999) reported that epithelial biopsies sample in the colorectum are highly complex, containing not only epithelial cells, but also endothelium, stromal tissue, bacteria, necrotic cells, blood, inflammatory cells, stools, which made the analysis not possible (Stulik *et al.*, 1999). However, proteomic study on colon carcinoma cell line may also not be representative as proteins in a cell are expressed different according to their growth environment. Thus, colonic normal and cancer tissues were used in our study. The proteins identified were group according to these definitions, the definition of "up-regulated" proteins is those that were more abundantly in cancer but less in normal. However, the "down-regulated" protein is defined as those proteins that are more abundant in normal as compared to cancer. Furthermore, the definition of "unique" proteins is those that either expressed only in normal tissues or cancer tissues.

The definition of "up-regulated", "down-regulated" and unique proteins are solely dependent on their existence in SDS-PAGE. The up-regulated proteins are defined as the proteins that were identified from the protein bands (in SDS-PAGE), which were more intense in cancer as compared to normal. The "down-regulated" proteins are defined as the proteins that were identified from the protein bands (in SDS-PAGE), which were more intense in normal as compared to cancer.

However, we noticed that an intense protein band normally contained more than one type of proteins. Therefore, for the up-regulated and down-regulated protein bands, we will excise the same molecular weight bands in normal and cancer. The proteins that only present in the more intense band will be grouped as unique protein band at 48.8 kDa in T1. Both normal and cancer tissue were found to expressed their protein band. However, the intensity of the band was greater in normal compared to cancer. Thus, we group it as a down-regulated protein band. Nevertheless, when we performed in-gel digestion on both the 48.8 kDa bands from normal and cancer. We found that in normal, there are three proteins registered, namely, serum albumin, serum albumin precursor and alpha-1-antitrypsin precursor. However, in the cancer, only two proteins registered as serum albumin and serum albumin precursor. Therefore, we group the alpha-1-antitrypsin precursor as unique protein whereas serum albumin and serum albumin precursor as down-regulated protein. The same methods of grouping are applied to other proteins. Finally, the unique proteins are defined as the proteins identified from the protein bands (in SDS-PAGE) that only found in either cancer or normal. In our present study, we noticed that many of the unique proteins identified are belong the serum albumin or hemoglobin proteins. These proteins are heterogenous proteins, which displayed variable mobility in SDS-PAGE and appeared as unique protein bands. However, for ease of definition for unique proteins, these protein will be defines as unique proteins.

The functions of the identified proteins in T1 buffer from normal and cancer tissues are shown in figure 6. There were a total of thirty-six of the down-regulated proteins and thirty-one of the up-regulated proteins were identified. In addition, six unique proteins from cancer and six unique proteins from normal were identified (Table 2). Among these proteins, the structural proteins made up approximately 9 % and 44 % of the identified proteins from normal and cancer tissues, respectively. The protein-binding proteins made up about 30 % and 24 % of the identified proteins from normal and cancer tissues, respectively. Moreover, about 21% and 18 % were identified from normal and cancer tissues

as transport proteins, respectively. Hypothetical proteins made up approximately 16 % and 3 % of the identified proteins from normal and cancer tissues, respectively. The enzyme made up about 7 % and 8 % of the identified proteins from normal and cancer tissues, respectively. Both cytoskeleton-binding proteins and endopeptidase inhibitor proteins were detected as unique proteins in the colon normal tissues. These proteins covered about 7 % of the total identified proteins from normal tissues. The proteins presence in minor quantity were lipoprotein-binding protein and DNA-RNA binding protein. In T1, lipoprotein-binding protein was found as unique protein in colon normal tissues. However, DNA-RNA binding protein was only found as unique protein in colon cancer tissues. Each of them was only covered 2 % of the total identified proteins from normal and cancer tissues, respectively.

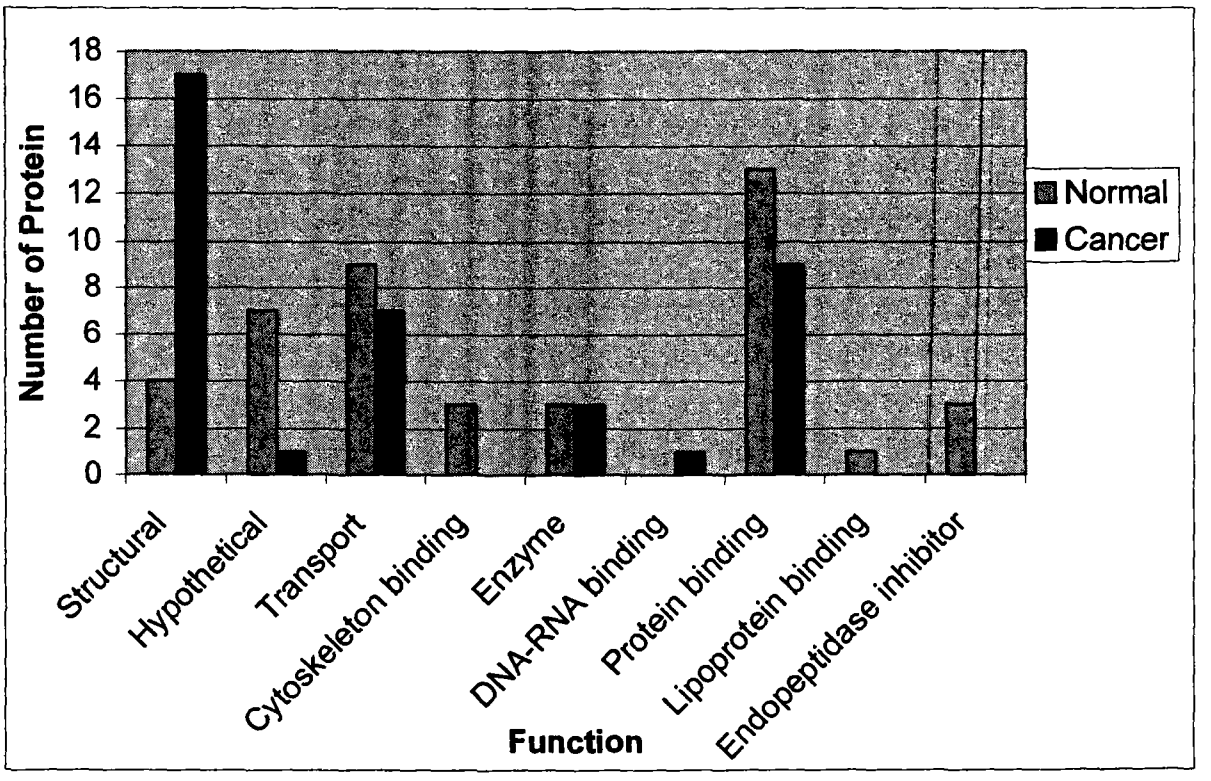

Figure 6: T1 Extraction Buffer

Figure 7 shows the functions of the identified proteins in S2 buffer from normal and cancer tissues according to their functions. A total of two down-regulated protein and eight up-regulated proteins were identified (Table 2). The two up-regulated proteins were as structural protein (actin-gamma; P00915) and ATP-binding protein (ATP[GTP]-binding protein; Q96QE3). In contrast, two of the up-regulated proteins that fowere identified as transport proteins, they are serum albumin (AAN17825) and serum albumin (precursor) (P02768). Meanwhile, another four proteins were grouped in hypothetical protein, channel protein (fibrinogen beta chain; P02675), enzyme (carbonate dehydratase EC 4.2.1.1; P00915) and motor activity protein (kinesin-like protein 2; Q9PTK9).

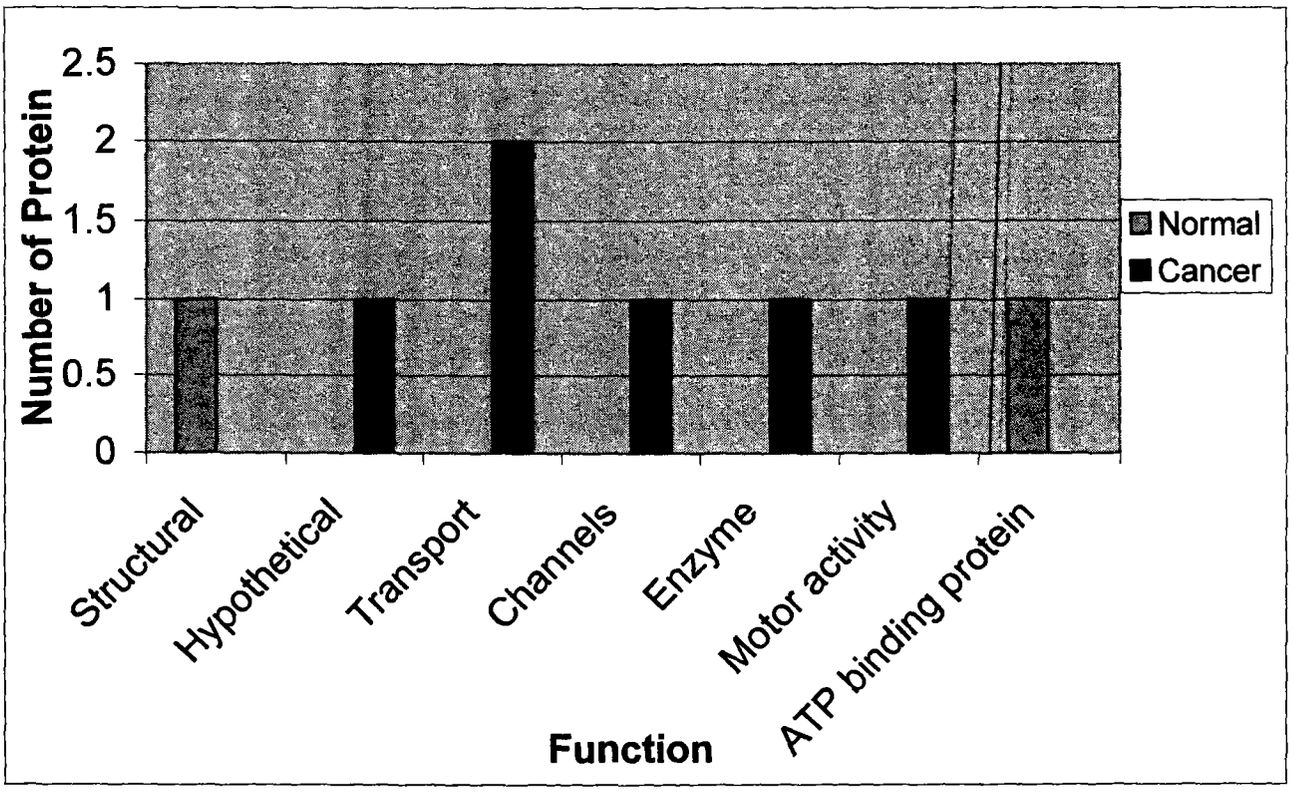

Figure 8: Solution Two Extraction Buffer

The functions of the identified proteins in S3 buffer from cancer tissues are shown in the figure 9. Eight of the up-regulated proteins from colon cancer tissues were identified (Table 2). Amongst these proteins, three were structural proteins and another three were transport proteins. The remaining proteins were hypothetical protein and enzyme.

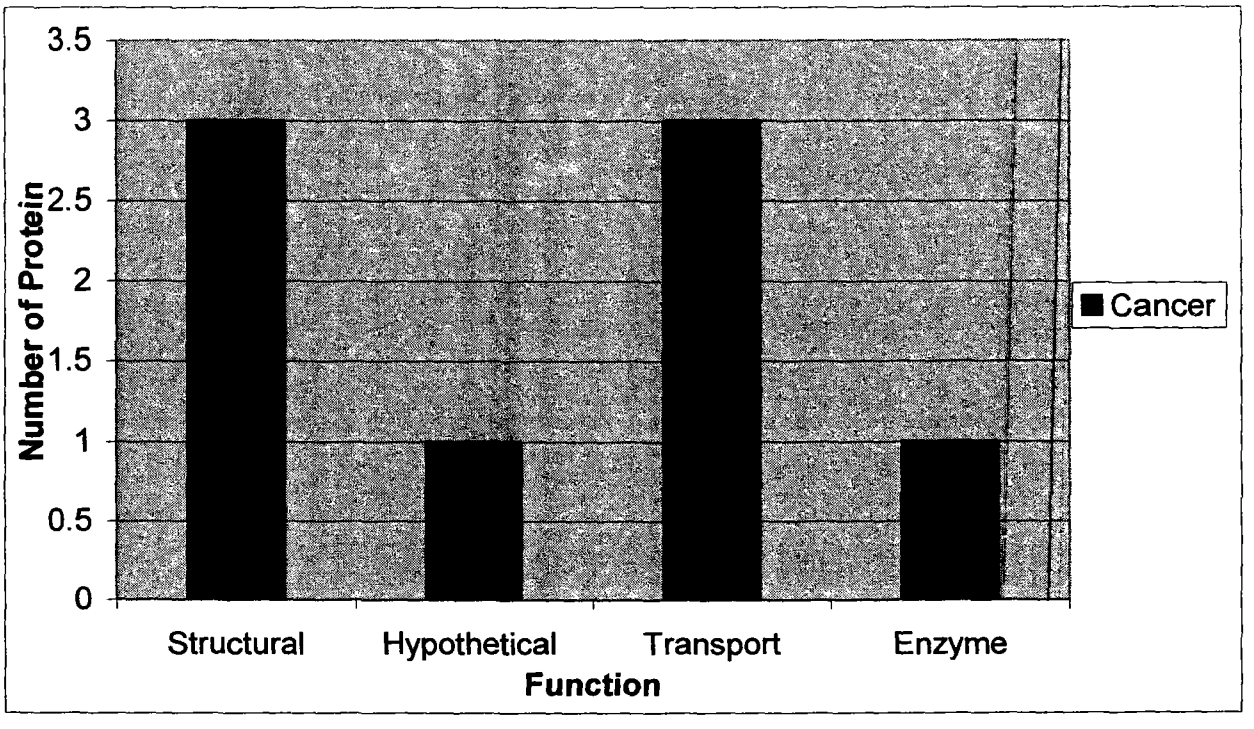

Figure 9: Solution Three Extraction Buffer

Conclusion

In this study, we have successfully isolated and analyzed the colonic cancerous and normal proteins. The protein maps for cancerous and normal proteins were identified where only the proteins that were consistently detected were considered for building of protein maps. The differentially expressed proteins between the normal and cancerous tissues were then targeted for further analysis using tandem mass spectrometric analysis. The products ions spectra of the peptides were then subjected to protein databases search for protein identification. Due to financial constraint, we only analyzed the S3 protein extracts with 2D gel analysis and the results shows that more protein spots were detected as compared with the SDS-PAGE, which shows that it is a promising separation technique for the continual of this project.

The identity of the differentially expressed proteins will grand us better understanding of the development the prognosis of colorectal cancer, furthermore more specimens needs to be analyzed in order to determine the suitability of the identified proteins as biomarker for detection and therapeutic invention for colorectal cancer amongst Malaysian patients.

References

AJCC. (2002). *American Joint Committee on Cancer: Cancer Staging Manual.* 6th Edition. New York: Springer Press: pp. 113-124.

Banks, R. E., Dunn, M. J., Hochstrasser, D. F., Sanchez, J. C., Blackstock, W., Pappin, D. J. and Selby, P. J. (2000). Proteomics: new perspectives, new biomedical opportunities. *Lancet.* **356**: 1749-1756.

Becamel, C., Galeotti, N., Poncet, J., Jouin, P., Dumuis, A., Bockaert, J. and Marin, P. (2002). A proteomic approach based on peptide affinity chromatography, 2-dimensional electrophoresis and mass spectrometry to identify multiprotein complexes interacting with membrane-bound receptors. *Biol. Proced. Online.* **4(1)**: 94-104.

Bio-Rad Laboratories INC. (2002). The ProteomeWorks system. United Stated patent application 5,616,502.

Bio-Rad Laboratories INC. (2002). The ProteomeWorks system. United Stated patent application 5,616,502.

Blackstock, W. P. and Weir, M. P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol.* **17**: 121-127.

Blumgart, L. H. and Fong, Y. (1995). Surgical options in the treatment of hepatic metastases from colorectal cancer. *Curr. Probl. Surg.* **32**: 333-421.

Bodmer, B. F. (1996). The somatic evolution of cancer: the Harveian Oration of 1996. *J. Royal Coll. Physicians Lond.* **31**: 82-89.

Brunagel, G., Schoen, R. E., Bauer, A. J., Vietmeier, B. N. and Getzenberg, R. H. (2002). Nuclear matrix protein alterations associated with colon cancer metastasis to the liver. *Clinical Cancer Research.* **8**: 3039-3045.

Carr, S. A., Hemling, M. E., Bean, M. F. and Roberts, G. D. (1991). Integration of mass spectrometry in analytical biotechnology. *Anal. Chem.* **63**: 2802-2824.

Chan, C. K., Singh, J., Rasid, B.K. and Devaraj, T. (1994). Penang cancer cases reported to the National Cancer Registry of Malaysia, 1987-1990: An Epidemiological Analysis. *Med. Journal Malaysia.* **49**: 122-131.

Cirincione, E. and Cagir, B. (2004). Rectal Cancer [Online]. [Accessed 20th December 2004]. Available from World Wide Web: http://www.emedicine.com/med/topic1994.htm

Ducret, A., Bruun, C. F., Bures, E. J., Marhaug, G., Husby, G. and Aebersold, R. (1996). Characterization of human serum amyloid A protein isoforms separated by two-dimensional electrophoresis by liquid chromatography / electrospray ionization tandem mass spectrometry. *Electrophoresis.* **17**: 866-876.

Ducret, A., Van-Oostveen, I., Eng, J. E., Yates-III, J. R. and Aebersold, R. (1998). High throughput protein characterization by automated reverse-phase chromatography / electrospray tandem mass spectrometry. *Protein Sci.*, **7**: 706-719.

Duvet, S., de Beeck, A. O., Cocquerel, L., Wychowski, C., Cacan, R. and Dubuisson, J. (2002). Glycosylation of the hepatitis C virus envelope protein E1 occurs posttranslationally in a mannosylphosphoryldolichol-deficient CHO mutant cell line. *Glycobiology.* **12**: 95-101.

El-Deiry, W. S. (2005). Colon Cancer, Adenocarcinoma [Online]. [Accessed 15th May 2005]. Available from World Wide Web: http://www.emedicine.com/med/topic413.htm

Fearon, E. R. (2001). Molecular biology of gastrointestinal cancers. *In: V. T. De Vita, Jr,, S. Hellman, S. A. Rosenberg, (eds). Cancer: principles and practice of oncology.* 6th edition. Philadelphia: Lippincott, Williams, and Wilkins Press. pp. 1037-1049.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell.* **61**: 759-767.

Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell.* **61**: 759-767.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989). Electrospray ionization for mass soectrometry of large biomolecules. *Science.* **246**: 64-71.

Figeys, D., Gygi, S. P., Zhang, Y., Watts, J., Gu, M. and Aebersold, R. (1998). Electrophoresis combined with novel mass spectrometry techniques: Powerful tools for the analysis of proteins and proteomes. *Electrophoresis.* **19**: 1811-1818.

Fountoulakis, M. and Suter, L. (2002). Proteomics analysis of the rat liver. *Journal of Chromatography B.* **782**: 197-218.

Fountoulakis, M., Tsangaris, G., Oh, J., Maris, A. and Lubec, G. (2004). Protein profile of the HeLa cell line. *Journal of Chromagraphy A.* **1038**: 247-265

Franks, L. M. (1986). What is cancer? *In: L. M. Franks and N. M. Teich, (eds). Introduction to the cellular and molecular biology of cancer.* Oxford – New York – Tokyo: Oxford University Press. pp. 1-26.

Friedman D. B., Hill, S., Keller, J. W., Merchant, N. B., Levy, S. E., Coffey, R. J. and Caprioli, R. M. (2004). Proteome analysis of human colon cancer by two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics.* **4**: 793-811.

Fung, E. T., Thulasiraman, V., Weinberger, S. R. and Dalmasso, E. A. (2001). Protein biochips for differential profiling. *Curr. Opin. Biotechnol.* **12**: 65-69.

Gam, L. H. and Aishah, L. (2002). Identification and confirmation of urinary human chorionic gonadotropin using in-gel digestion and HPLC / MS / MS analysis. *Jurnal Biosains.* **13 (1)**: 27-34.

Govorukhina, N. I., Keizer-Gunnink, A., van der Zee, A. G. J., de Jong, S., de Bruijin, H. W. A. and Bischoff, R. (2003). Sample preparation of human serum for the analysis of tumor markers: comparison of different approaches for albumin and γ-globulin depletion. *Journal of Chromatography A.*

Gibson, B. W. and Biemann, K. (1984). Strategy for the mass spectrometric verification and correction of the primary structures of proteins deduced from their DNA sequences. *Proc. Natl. Acad. Sci. USA.* **81**: 1956-1960.

Harkins, L., Volk, A. L., Samanta, M., Mikolaenko, I., Britt, W. J., Bland, K. I. and Cobbs, C. S. (2002). Specific localization of human cytomegalovirus nucleic acids and proteins in human colorectal cancer. *Lancet*: **16**: 1557-1563.

Hjelmeland, L. M. (1990). Solubilization of native membrane proteins. *Methods in enzymology.* **182**: 253-264.

Hooi, L. N. and Devaraj, T. (1998). A hospital based study of cancer admissions, *Med. Journal Malaysia.* **53**: 22-28.

Hunt, D. F., Yates, J. R., Shabanowitz, J., Winston, S. and Hauer, C. R. (1986). Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci.* **83**: 6233-6237.

Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C. and Watanabe, C. (1993). Identifying proteins from two-dimensional gels by molecular mass searching of peptides fragments in protein sequence databases. *Proc. Natl. Acad. Sci. USA.* **90**: 5011-5015.

Ilyas, M., Straub, J., Tomlinson, I. P. and Bodmer, W. F. (1999). Genetic pathways in colorectal and other cancers. *Eur. J. Cancer.* **35**: 1986-2002.

Issaq, H. J., Conrads, T. P., Janini, G. M. and Veenstra, T. D. (2002). Methods for fractionation, separation and profiling of proteins and peptides. *Electrophoresis.* **23**: 3048-3061.

Jensen, O. N., Wilm, M., Shevchenko, A. and Mann, M. (1999). Sample preparation methods for mass spectrometric peptide mapping directly from 2–DE gels. *Methods Mol. Biol.* **112**: 513-530.

Ji, H., Whitehead, R. H., Reid, G. E., Moritz, R. L., Ward, L. D. and Simpson, R. J. (1994). Two-dimensional electrophoretic analysis of proteins expressed by normal and cancerous human crypts: application of mass spectrometey to peptide-mass fingerprinting. *Electrophoresis.* **15(3-4)**: 391-405.

Laemmli, U. K. (1970). Clevage of structural proteins during the assembly of the head of bactriophage T4. *Nature.* **227**: 680-685.

Mann, M. and Wilm, M. (1994). Error-Tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **66(24)**: 4390-4399.

Mann, M., Hendrickson, R. C. and Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem.* **70**: 437-473.

Mann, M. and Talbo, G. (1996). Developments in matrix-assisted laser desorption / ionization peptide mass spectrometry. *Curr. Opin. Biotechnol.* **7**: 11-19.

Muto, T., Bussey, H. J. and Morson, B. C. (1995). The evolution of cancer of the colon and rectum. *Cancer.* **36**: 2251–2270.

Narayan, S. and Roy, D. (2003). Role of APC and DNA mismatch repair genes in the development of colorectal cancers. *Molecular cancer.* **2**: 41-56.

National Cancer Institute: Cancer of colon and rectum (2004) [Online]. [Accessed 19th February 2004). Available from World Wide Web: http://www.cancer.gov/cancerinfo/wyntk/colon-and-rectum

O' Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**: 4007-4021.

O' Connell, M. J., Schaid, D. J., Ganju, V., Cunningham, J., Kovach, J. S. and Thibodeau, S. N. (1992). Current status of adjuvant chemotherapy for colorectal cancer. Can molecular markers play a role in predicting prognosis? *Cancer (Phila.).* **70**: 1732-1739.

PO Box 2345, Beijing 100023, China
www.wjgnet.com
wjg@wjgnet.com

World J Gastroenterol 2006 August 21; 12(31): 4973-4980
World Journal of Gastroenterology ISSN 1007-9327

Analysis of differentially expressed proteins in cancerous and normal colonic tissues

Lay-Harn Gam, Chiuan-Herng Leow, Che Nin Man, Boon-Hui Gooi, Manjit Singh

Lay-Harn Gam, Chiuan-Herng Leow, School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
Che Nin Man, National Poison Centre, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
Boon-Hui Gooi, Manjit Singh, Department of Surgery, Penang General Hospital, Penang, Malaysia
Supported by the Ministry of Science, Technology and Innovation Malaysia (06-02-05-3197EA12)
Correspondence to: Dr. Lay-Harn Gam, School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia. layharn@usm.my
Telephone: +60-4-6533888-2208 Fax: +60-4-6570017
Received: 2006-01-23 Accepted: 2006-02-18

Abstract

AIM: To identify and analyze the differentially expressed proteins in normal and cancerous tissues of four patients suffering from colon cancer.

METHODS: Colon tissues (normal and cancerous) were homogenized and the proteins were extracted using three protein extraction buffers. The extraction buffers were used in an orderly sequence of increasing extraction strength for proteins with hydrophobic properties. The protein extracts were separated using the SDS-PAGE method and the images were captured and analyzed using Quantity One software. The target protein bands were subjected to in-gel digestion with trypsin and finally analyzed using an ESI-ion trap mass spectrometer.

RESULTS: A total of 50 differentially expressed proteins in colonic cancerous and normal tissues were identified.

CONCLUSION: Many of the identified proteins have been reported to be involved in the progression of similar or other types of cancers. However, some of the identified proteins have not been reported before. In addition, a number of hypothetical proteins were also identified.

Key words: Colon cancer; Tissue specimens; Sequential protein extraction; Proteomics

Gam LH, Leow CH, Man CN, Gooi BH, Manjit S. Analysis of differentially expressed proteins in cancerous and normal colonic tissues. *World J Gastroenterol* 2006; 12(31): 4973-4980

http://www.wjgnet.com/1007-9327/12/4973.asp

INTRODUCTION

Colorectal cancer is a worldwide public health concern and also a frequent cause of mortality and morbidity in developed, developing and industrialized countries[1]. The incidence of colon cancer increases exponentially with age. Environmental factors and host immunological characteristics could contribute to initiation and progression of the disease[2]. Colorectal cancer is among the best characterized cancers with regard to the genetic progression of the disease[3].

The ability to obtain maps of the protein content of cells provides a basis for comparison of the protein expression in normal and cancer cells. This capability becomes especially important for mapping the changes that occur during cancer progression. In the transformation process, changes in protein expression may result in expression of proteins at elevated or reduced levels[4,5,6]. When a cell transforms from normal to malignant, changes in protein expression ultimately are reflected in the phenotype of the cell. It is also important to map the differences between different stages of cancer and between different cancer samples at the same stage of progression in order to understand the pathways involved in these processes. Ultimately, identification of important proteins involved in the transformation process may lead to the identification of early markers for detection of specific types of cancers and their treatments[7].

In the present study, proteins from normal and cancerous colon tissues were extracted in three separate fractions. The mapping of a protein profile in each fraction was carried out individually. This approach eases the comparison of protein expression in normal and cancerous tissue as it reduces the total proteins present in each fraction. Using this method, we believe that proteins from different compartments of the cells were extracted in sequential order, where the aqueous soluble proteins were first extracted followed by non-aqueous soluble proteins consisting of proteins with moderate to extreme hydrophobicity. The proteins analyzed in this study were extracted from human tissue specimens (both cancerous and normal). In proteomics study, analysis of cancerous

tissue has advantages over analysis in established cancer cell lines, as the latter may not represent the actual expression of proteins.

MATERIALS AND METHODS

Tissue specimen collection

Normal and cancerous colonic tissues were collected from patients with colon cancer. The specimens were provided by Hospital Pulau Pinang. Normal tissue was taken from the mucosal layer of the normal colon adjacent to the tumor tissue. Both the normal and cancerous tissues were confirmed by hospital pathologists. Ulceration or infected colon cancers were exclusion criteria for the study. The tissues were collected after informed consent had been obtained from the patients. The patients had different grades of colorectal cancer as shown in Table 1. The tissues were cut into similar sizes, weighed and immediately stored at -70℃ until analyzed.

Well-differentiated adenocarcinoma refers to the less aggressive type of cancer where the cancerous tissues more closely resemble the normal tissue under histopathological examination, whereas moderately differentiated adenocarcinoma refers to moderately aggressive cancer cells.

Tissue lysis and sequential extractions

The tissues were defatted and an equal weight of tissues (normal and cancerpis) was determined and subjected to analysis. The deep-frozen colon tissue specimens were disrupted by grinding in a liquid nitrogen-cooled mortar until the specimens became powder-form. The powder form-like tissues were divided into aliquots in separate eppendoft tubes. Sequential extraction of protein using three protein extraction buffers was carried out. These extraction buffers were prepared according to Molloy *et al*[8] with modifications. The extraction procedures are described in the following sections.

Tris Extraction Buffer

The constituent of Extraction Buffer 1 (S1 buffer) is 40 mmol/L Tris. 1000 μL of S1 buffer was added to 500 mg of homogenized tissue. The mixture was vortexed for 30 seconds and centrifuged at 20℃ with the speed of 12000 r/min for 8 min. The supernatant was collected in aliquots (50 μL) and kept at -70℃. The pellet was thoroughly washed with S1 buffer before being subjected to the second step in the sequential extraction.

Solution Two Extraction Buffer

The extraction Buffer 2 (S2 buffer) contained 8 mol/L Urea, 50 mmol/L DTT, 40 g/L CHAPS, 2 mL/L Carrier ampholytes (pH 3-10) and 2 mg/L Bromophenol Blue. The recovered pellet was washed twice again with the T1 solution, and was then vortexed, centrifuged and the supernatant was discarded. After washing, a volume of 1000 μL S2 solution was added to the pellet. The mixture was then vortex for 30 s and then centrifuged at 20℃ with the speed of 12000 r/min for 8 min. The supernatant was collected in aliquots (50 μL) and kept at -70℃. The pellet was thoroughly washed with the same extraction buffer before being subjected to the third step in the sequential extraction.

Table 1 Clinical and pathological data of patient tissue specimens

Patient #	Age	Gender	Race	Cancer stage	Cancer grade
MMN	77	Female	Malay	Stage IV	Well differentiated adenocarcinoma
LKH	66	Male	Chinese	Stage II	Moderately differentiated adenocarcinoma
CCH	55	Male	Chinese	Stage IV	Moderately differentiated adenocarcinoma
NSC	55	Female	Chinese	Stage III	Moderately differentiated adenocarcinoma

Solution three Extraction Buffer

The extraction Buffer 3 (S3 buffer) contained 5 M Urea, 2 mol/L Thiourea, 50 mM DTT, 20g/LCHAPS, 20 mL/L Triton, 2 mL/L Carrier ampholytes (pH 3-10), 40 mmol/L Tris and 2 mg/L Bromophenol Blue. The recovered pellet from S2 extraction was washed twice with S2 buffer, it was then vortexed, centrifuged and the supernatant was discarded. After washing, 1000 μL of S3 buffer was added to the pellet. The mixture was vortexed for 30 s and then centrifuged at 20℃ with the speed of 12000 r/min for 8 min. The supernatant was collected in aliquots (50 μL) and kept at -70℃. The pellet was then discarded.

Sample preparation and Electrophoresis

The preparation of protein samples for SDS-PAGE separation was carried out by adding 200 mL/L of sample buffer (0.5 mol/L Tris-HCl, 100 mL/L glycerol, 20 mg/L SDS, 100 mg/L Coomassie blue) to each fraction (5 μL) and vortexed for 30 s. The samples (normal and cancerous) and the protein molecular weights markers were then loaded into individual wells of SDS-polyacrylamide gel using a syringe. Electrophoresis was performed using a vertical electrophoresis slab gel apparatus at a constant voltage of 200 volts throughout the electrophoresis process. Electrophoresis was terminated when the dye front was 2 to 3 mm away from the bottom edge of the gel for both experiments. The gel was stained with Coomassie blue.

In-Gel Digestion

The SDS-Polyacrylamide gel was washed thoroughly in 100 mmol/L NH_4HCO_3 and the differentially expressed protein bands from either normal or cancerous colonic tissues were then excised from the gel. In-gel digestion was performed using trypsin according to Gam and Aishah[9] with slight modification. The gel pieces were first excised and shrunk by dehydration in acetonitrile. The solvent was then discarded and the gel pieces were dried in a vacuum centrifuge. A volume of 10 mmol/L dithiotreitol (DTT) in 100 mmol/L NH_4HCO_3 that was sufficient to cover the gel pieces was added and the protein was reduced for 1 h at 56℃. After cooling to room temperature, the

DTT solution was replaced with the same volume of 55 mmol/L iodoacetic acid in 100 mmol/L NH_4HCO_3. After 45 min incubation at ambient temperature in the dark with occasional vortexing, the gel pieces were washed with 50-100 μL of 100 mmol/L NH_4HCO_3 for 10 min, dehydrated with acetonitrile, rehydrated in 100 mM NH_4HCO_3 and dehydrated in the same volume of acetonitrile. The liquid phase was removed and the gel pieces were dried in a vacuum centrifuge. The gel pieces were swollen in digestion buffer containing 50 mmol/L NH_4HCO_3, 5 mmol/L $CaCl_2$, and 12.5 ng/μL of trypsin in an ice-cold bath. After 45 min, the supernatant was removed and replaced with 10 μL of the same buffer but without trypsin to keep the gel pieces wet during enzymatic cleavage at 37℃ overnight. Peptides were extracted from the gel matrix by adding 15 μL of 20 mmol/L NH_4HCO_3, vortexed and incubated at room temperature for 10 min and the supernatant was recovered after a brief spin. This was followed by adding (1 to 2 times the volume of gel pieces) 50 mL/L formic acid in acetonitrile:water mixture (70:30), vortex and incubated for 20 min at room temperature. It was then spun down and the supernatant was recovered. These steps were repeated 3 times. Pooled extracts were dried in a vacuum centrifuge and stored at -70℃. Each of the pooled extracts (peptides) was added with 30 μL of deionized water. Then, they were centrifuged at 500 r/min for 5 min at 15℃. The extracts were then subjected to LC/MS/MS analysis.

Mass spectrometric analysis

A volume of 5 μL of the sample was injected into a RPC-column (C_{18} 300 A, 5 μm, 1 mm X 150 mm) connected to a Hewlett Packard HPLC. A capillary pump was used to pump the mobile phase at 20 μL/min flow rate, the linear gradient used was 5% B to 95% B in 65 min. Mobile phase A was 500 μL/L formic acid in deionized water and B was 500 μL/L formic acid in ACN. The HPLC was interfaced to an ion trap mass spectrometer. The dry gas temperature was 300℃, dry gas flow rate was of 8.0 L/min, nebulizer pressure of 30.0 psi. The above parameters were used subsequently for acquiring MS data. The peptides were ionized using electrospray soft ionization technique (ESI).

The experimental method was made up of one experimental protocol consisting of 2 scan events. The first scan event was a full scan MS and the second was the data dependent MS/MS scan which is dependent on the results of the first scan event. The most intense ion in a MS scan will be automatically isolated and excited to MS/MS scan. The parameters set for data dependent scan (MS/MS scan) were default collision energy (voltage) = 1.15V, charge state = 2, minimum threshold = 3000 counts, and the isolation width = 2 m/z.

Mascot Protein Identification

The protein was identified using Mascot Protein Database Search engine. The Peptide Mass Tolerance was set as ± 2 u and ± 0.8 u was set for the Fragment Mass Tolerance. Only 1 missed cleavage was allowed. This software is available at *www.matrix-science.com*. Matches were computed using a probability-based Mowse score defined as -10 × log (*P*), where *P* is the probability that the observed match was a random event.

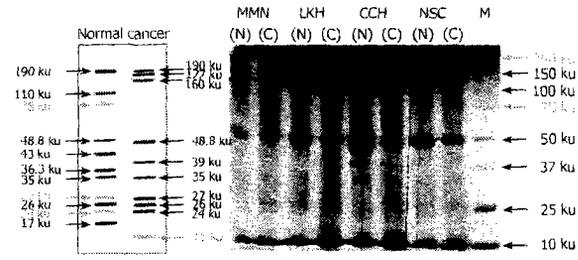

Figure 1 Proteins profiles of normal and cancer colon tissues from four different patients. The proteins were extracted using in S1 buffer. N: Normal colon tissues; C: Cancer colon tissues; M: Marker; Lanes 1 and 2: MMN samples; Lanes 3 and 4: LKH samples; Lanes 5 and 6: CCH samples and Lanes 7 and 8: NSC samples. Diagram 1 shows the consistent protein bands found in the normal and the cancer tissues of the figure.

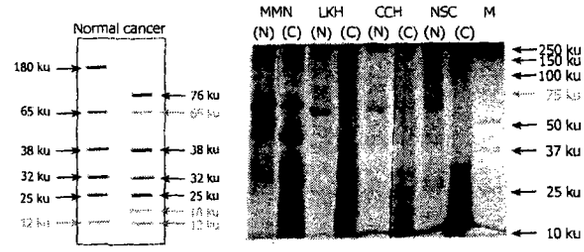

Figure 2 Proteins profiles of normal and cancer colon tissues from four different patients. The proteins were extracted using in S2 buffer. N: Normal colon tissues; C: Cancer colon tissues; M: Marker; Lanes 1 and 2: MMN samples; Lanes 3 and 4: LKH samples; Lanes 5 and 6: CCH samples and Lanes 7 and 8: NSC samples. Diagram 2 shows the consistent protein bands found in the normal and the cancer tissues of the figure.

RESULTS

Figure 1 shows the protein profiles of protein extracted using S1 buffer from normal and cancerous colon tissues of four different patients. Diagram 1 shows the consistent protein bands for both normal and cancerous tissues. The thick lines represent the protein bands that were expressed in all four tissues and the faint lines represent the protein bands that were expressed in three of the four patients. The distinct bands that were only detected in one particular tissue are ignored, as it may not represent the proteins that are involved in cancer formation. In colon cancer tissues, the protein bands that were expressed in all the four patients were at 190 ku, 177 ku, 160 ku, 48.8 ku, 39 ku, 35 ku, 27 ku, 26 ku and 24 ku. However, a protein band at 11 ku was detected in three of the four patients. In normal colon tissues, protein bands at 190 ku, 110 ku, 48.8 ku, 43 ku, 36.3 ku, 35 ku, 26 ku and 17 ku were found in all the patients' tissues while protein bands at 78 ku, 27 ku and 24 ku were found expressed only in three patients. When comparing cancerous and normal tissue, it is clear that some of the proteins were differently expressed between the two types of tissues.

Figure 2 shows the profiles of proteins extracted using S2 buffer from both normal and cancerous colonic tissues. Diagram 2 represents the consistent protein bands that

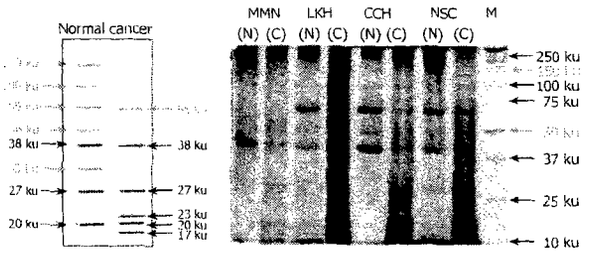

Figure 3 Proteins profiles of normal and cancer colon tissues from four different patients. The proteins were extracted using in S3 buffer. N: Normal colon tissues; C: Cancer colon tissues; M: Marker; Lanes 1 and 2: MMN samples; Lanes 3 and 4: LKH samples; Lanes 5 and 6: CCH samples and Lanes 7 and 8: NSC samples. Diagram 3 shows the consistent protein bands found in the normal and the cancer tissues of the figure.

Table 2 Up-regulated, down-regulated proteins extracted in S1, S2, S3 buffers from colon tissues

SWISS-PROT accession number	Down-regulated proteins (Normal Tissues)	SWISS-PROT accession number	Up-regulated proteins (Cancer Tissues)
P12718	Actin, gamma-enteric smooth muscle	P04262	Type II, keratin subunit protein
P17661	Desmin	P35900	Keratin 20
P08670	Vimetin	P35908	Keratin, 67k type II epidermal
Q9DFD0	Hypothetical protein	P13647	Keratin k5
Q7Z5W1	Hypothetical protein	Q86Y46	Keratin 6 irs 3
CAD19027	Hypothetical protein	P04264*	Keratin 1, type II cytoskeletal
BAC87457	Hypothetical protein	P35527	Cytoskeletal 9
CAC08866	Hypothetical protein	P04264	Keratin, type II cytoskeletal I
CAC08835	Hypothetical protein	P35527	Keratin 9, type I, cytoskeletal
Q8WTU6	Hypothetical protein	P08779	Keratin, type I, cytoskeletal 16
Q01995	Transgelin	P48669	Keratin, type II, cytoskeletal 6F
Q91706	Skeletal muscle beta-tropomyosin	P02538	Keratin 6A, type II
P12277	Creatine kinase	P19013	Keratin, type II, cytoskeletal 4
P00915	Actin-gamma	P08729	Keratin, type II, cytoskeletal 7
Q96QE3	ATP[GTP]-binding protein	P12035	Keratin, type II, cytoskeletal 3
		P05783	Keratin, type I, cytoskeletal 18
		P13645	Keratin, type I, cytoskeleton 10
		Q8NF17	Hypothetical protein
		CAE99935	Hypothetical protein
		P02675	Fibrinogen beta chain
		P00915	Carbonate dehydratase EC 4.2.1.1
		Q9PTK9	Kinesin-like protein 2
		P60174	Triosephosphate isomerase EC 5.3.1.1
		Q6PJ43	ACTG1 protein
		Q13733	Sodium/potassium-transporting ATPase alpha-4 chain EC 3.6.3.9

were expressed in the tissues analyzed. For the cancerous tissues, the protein bands at 76 ku, 38 ku, 32 ku and 25 ku were consistently expressed in all the cancerous tissues analyzed while protein bands at 65 ku, 18 ku and 12 ku were only found in the tissues of three of the four patients. For the normal tissues, the bands at 180 ku, 65 ku, 38 ku, 32 ku and 25 ku were found in the tissues of all four patients while only one band at 12 ku was detected in the tissues of three patients. Most of the proteins extracted from cancerous tissues using S2 buffer were found at the low molecular weight region, in the range of 35 ku to 10 ku. In contrast, relatively few protein bands from normal tissues were detected in this region and those that were detected were present at low protein intensity.

Figure 3 shows the protein profiles of proteins extracted using S3 buffer from normal and cancer colon tissues of the four patients. Diagram 3 represents the consistent protein bands expressed in the four tissues analyzed. Protein bands at 38 ku, 27 ku, 23 ku 20 ku and 17 ku from cancerous tissues were found in all four patients whereas only the band at 65 ku was found in three of the four patients. Normal tissues of the corresponding patients displayed bands at 38 ku, 27 ku and 20 ku whereas tissues from three patients were found to express consistent bands at 150 ku, 100 ku, 65 ku, 46 ku and 35 ku. The protein band at 65 ku was detected at the highest intensity compared to the other bands and it was uniquely expressed in all the patients except for patient #MMN, which was the only patient diagnosed with well differentiated adenocarcinoma cancer. Thus, its expression is probably related to the cancer grade. Nevertheless, this may just be an ambiguous assumption as only one patient with well differentiated cancer grade was analyzed. When comparing the protein profiles of protein extracted from tissues using S3 buffer, it is obvious that the expression of low molecular weight proteins (below 37 ku) in cancerous

Table 3 Unique proteins extracted in S1, S2, S3 buffers from colon tissues

SWISS-PROT accession number	Unique proteins (Normal Tissues)	SWISS-PROT accession number	Unique proteins (Cancer Tissues)
P32119	Peroxiredoxin 2 precursor	Q9BX84	LTRPC6 channel kinase 6
Q03001	Bullous pemphigoid antigen 1 precursor	Q9AVW8	DNA-directed RNA polymerase II largest chain EC 2.7.7.6
P01009	Alpha-1-antitrypsin precursor	P29312	14-3-3 protein zeta
P20848	Alpha-1-antitrypsin related protein presursor	P31946	14-3-3 protein beta/alpha
1QMBA	Alpha-1-antitrypsin mutant YES, chain A	S38532	Protein 14-3-3 eta chain
		S31975	14-3-3 protein epsilon

tissues was greater than that of the normal tissues.

Table 2 shows the list of proteins detected, with a total of 15 and 24 down-regulated and up-regulated proteins identified from normal and cancerous tissues, respectively. Table 3 shows the unique proteins that were identified form normal and cancerous tissues.

DISCUSSION

The sequential protein extraction technique consists of three types of protein extraction buffers, namely the S1 buffer, S2 buffer and S3 buffer. The S1 buffer is used mainly to extract aqueous soluble proteins, such as cytosolic proteins and nuclear proteins while the S2 and S3 buffers were used to extract proteins with intermediate to extreme hydrophobicity. The constituent of S2, which is made up of chaotropic agents and other reagents may disturb rugged protein-protein interaction found among structural, cytoskeletal proteins and aggregating proteins[10]. However, the combination of urea and thioureas in S3 was used to enhance the solubility of membrane proteins[8]. The extraction buffers were always used in the orderly sequence of S1 followed by S2 and then S3. The use of sequential protein extraction buffers was first suggested by Molloy *et al*[8] for separation of protein mixtures using 2D-gel electrophoresis. However, in our present study, the extracted proteins were separated using the SDS-PAGE technique, whereby direct comparison of protein profiles from tissues of four different patients was carried out.

In general, use of sequential extraction buffers is beneficial to this study because proteins were extracted according to their increasing hydrophobic strength, which may imply that proteins from different cellular compartments were extracted in different fractions. This method not only allows the analysis of proteins according to their localization in the cells but also reduces the number of proteins extracted in each fraction, which provides better resolution of proteins and their visualization in SDS-PAGE.

The protein profiles of the proteins extracted using the S1 and S2 buffers are significantly different, however the protein profiles of proteins extracted using S2 and S3 showed only slight differences. The significant variation between the protein profiles of S1 and S2 can be explained by the vast difference in the property of the proteins extracted, where S1 buffer extracts proteins only solubilize in aqueous solution while S2 buffer extracts the non-aqueous soluble proteins. The protein profiles displayed by proteins extracted using S2 and S3 also showed a certain degree of variation as S3 buffer was used to enhance the solubility of more stringent proteins from the tissue[11].

The differentially expressed proteins in each of the gels were then subjected to further analysis and the identity of each protein was analyzed using LC/MS/MS. Proteins were identified using the Mascot protein search engine on the basis of peptide mass matching[12] with redundancy of post-translational modification and proteolysis. The definition of "up-regulated", "down-regulated" and unique proteins in this study are solely dependent on their existence in SDS-PAGE. When analyzed using Quantity one software (BioRad), the intensity of the bands for up-regulated proteins was more intense in cancerous as compared to normal tissues while the "down-regulated" proteins were the proteins that were more intense in normal as compared to cancerous tissues. However, some of the intense protein bands in SDS-PAGE were found to contain more than one type of protein. Therefore, for the up-regulated and down-regulated protein bands, both the protein bands of the identical molecular weight from normal and cancerous tissues were excised and subjected to LC/MS/MS analysis. The proteins that were detected in both tissues were then quantified by their peptides' peak areas in selected ion chromatogram analysis. The proteins present in one tissue and not the other are reported as unique proteins.

A total of ninety-five proteins were identified from the protein bands indicated earlier for both the normal and cancerous tissues. However, some of these proteins were of serum origin, due to the embedded blood vessels in the tissues, and these proteins were excluded from the protein list. Furthermore, some of the proteins were represented by more than one band as described by heterogeneity. The heterogeneous proteins were mainly derived from structural, transport and enzyme subunits such as Actin, gamma-enteric smooth muscle, Desmin, Vimetin, Keratin, type I cytoskeletal 10, Cytoskeletal 9, and Triosephosphate isomerase EC 5.3.1.1. Phospholylation, glycosylation and limited proteolytic activities have been considered as major modification of proteins and they therefore affect the mobility of the modified proteins in SDS-PAGE[13,14,15].

The structural proteins were mostly extracted from the cancerous tissues using the S3 buffer. The majority of the structural proteins identified belongs to the cytokeratin family. Williams *et al*[16] reported that cytokeratin expressing tumor cells contain cytokeratins 8 and 18, but not cytokeratin 19 or vimentin. The roles of cytokeratins

are commonly known to encompass both structural and signaling capabilities. In the present study, cytokeratin 18 was found in cancerous colon tissues in S3 buffer extracts. The expression of cytokeratin 18 in colonic cancerous tissues is consistent with those reported by Leong *et al*[17]. In contrast, the expression of cytokeratin 18 was found reduced in malignant tumors of prostate carcinoma compared to benign prostatic hyperplasia[18]. The expression of cytokeratins is sometimes used as an indicator of the behavioral changes of a tissue. For instance, in breast tissue, cytokeratin 8 has been reported to be the major receptor for plasminogen on breast cells[18].

Vimentin is expressed at higher levels in the low-secreting variant. It is present in many cell lines, but normally absent from differentiated cells. Thus, lower vimentin expression in the high secretors may be indicative of a more differentiated phenotype[19]. In our present study, vimentin is a down-regulated protein found in S1 buffer extracts. Williams *et al*[16] has reported that the absence of vimentin in colon cancer tissue is due to its correlation with cytokeratin 19 or other cytokeratin expressions. However, Birchmeier *et al*[20] and Seshadri *et al*[21] reported no unequivocal correlation between a gain of vimentin and poor prognosis with primary tumors.

Hypothetical proteins were extracted in all three extraction buffers from both normal and cancerous colonic tissues. Generally, hypothetical proteins are defined as the proteins that were not described at the protein level, but were predicted from cDNA sequences[22]. As shown in Table 2, a series of hypothetical proteins were identified. Afjelhi-Sadat, *et al*[23] suggested that the hypothetical proteins may serve as marker or protein vaccine candidates. Particularly when the hypothetical protein is found in cancerous tissue exclusively, it can be considered as novel and of pivotal importance. Some of the hypothetical proteins were detected as heterogeneous proteins suggesting the post-translational nature of the proteins.

Peroxiredoxin 2 is a type transport protein and it was extracted in the S1 buffer as a unique protein in normal tissues. Peroxiredoxin 2 is part of the peroxidase family, which are proteins that have been known to connect with cell proliferation, differentiation and apoptosis.

Fibrinogen is a type of channel protein and its role was reported to be supporting the binding of growth factors and promoting the cellular responses of adhesion, proliferation, migration during angiogenesis and tumor cell growth[24]. Fibrinogen was detected in the S2 buffer of cancerous tissues. The presence of fibrinogen in cancer cells was demonstrated to affect the progression of tumor cell growth and metastasis on the basis that fibrinogen alters the ability of breast cancer cells to migrate[24]. In addition, fibrinogen was also found present in lung cancer following the detection of elevated plasma fibrinogen levels in advanced stages of the disease[25] and the localization of fibrinogen in tumor tissue[26,27]. There are also reports of the deposition of fibrinogen on the surfaces of tumor cells and lymphocytes[28,29]. In primary and metastatic tumor[30], the presence of fibrinogen was regarded as an adverse event leading to tumor development. Yamaguchi, *et al*[31] reported that lung cancer was developed through induction of hyperfibrinogenemia.

Transgelin (SM22-alpha) is a cytoskeletal-binding protein that was isolated in S1 buffer as a down-regulated protein. Maguire *et al*[32] has reported that loss of transgelin expression is an indication of early tumor progression and the authors suggested that transgelin may serve as a diagnostic marker for breast and colon cancer. We found that the expression of transgelin in cancerous tissue is much lower than that of normal tissue. Other cytoskeletal-binding proteins that were detected in S1 buffer extract were Bullous pemphigoid antigen 1 precursor and Skeletal muscle beta-tropomyosin.

Some of the enzymes that play important roles in metabolic pathways and are essential for energy production were detected as up-regulated protein in S1 buffer. These enzymes include triosephosphate isomerase and LTRPC6 channel kinase 6. LTRPC6 channel kinase 6 is an essential ion channel and serine/threonine-protein kinase is crucial for magnesium homeostasis. It plays an important role in epithelial magnesium transport and in the active magnesium absorption in the gut and kidney.

Sodium/potassium-transporting ATPase alpha-4 chain was detected as the up-regulated protein in S3 buffer. It is an integral membrane protein that is located in the cell membrane[33]. It catalyzes the hydrolysis of ATP coupled with the exchange of sodium and potassium ions across the plasma membrane.

Carbonate dehydratase is also the up-regulated protein detected in S1 and S2 buffer, it is a zinc-containing enzyme that catalyses the reversible hydration of carbon dioxide, which is significant in the transport of CO_2 from the tissues to the lungs. It was reported that over expression of the zinc enzyme carbonic dehydratase is observed in certain human cancers[34]. Its expression was reported to be elevated in the renal cancer cells as compared to the surrounding normal kidney tissue[35].

DNA-directed RNA polymerase II largest chain is the DNA-RNA binding protein that was detected as unique protein in S1 buffer extract of cancer tissues. It was also reported in human brain tumors[36].

Proteins 14-3-3 are the protein-binding proteins that were isolated in the S1 buffer as unique proteins from colon cancer tissue. Protein 14-3-3 plays a role in the regulation of signal transduction protein pathways implicated in the control of cell proliferation, differentiation and survival[37]. This protein was reported to play multiple roles in maintaining cell survival[38,39]. Friedman, *et al*[40] has reported that protein 14-3-3 in three different isoforms was expressed up to 1.7 fold in the colon tumors as compared to normal colon tissue.

Alpha-1-antitrypsin is an endopeptidase inhibitor protein that was isolated as a unique protein in S1 buffer from normal tissue only. In individuals who lack this inhibitor protein, the protease destroys the membrane system, leaving the colon and rectum vulnerable to colorectal cancer development[41]. The main role of alpha-1-antitrypsin is in defense against elastase damage that occurs in the lung under normal physiology conditions[42]. Alpha-1-antitrypsin was detected in patients with lung cancer[43] and was used as a cancer marker for cervical cancer[44]. Contrary to our finding, Friedman, *et al*[40]

reported the detection of alpha-1-antitrypsin in colon cancer.

ATP (GTP)-binding protein is ATP-binding protein that was detected as down-regulated protein in S2 buffer. It is a mitochondrial membrane protein.

Kinesin is a motor activity protein that was detected as an up-regulated protein in S2 buffer extract. It is a microtubule-associated force-producing protein that may play a role in organelle transport[45]. This protein was also reported in breast cancer tissue[46]. Kinesin-like protein 2 is an immunogenic breast cancer antigen[47]. In addition, kinesin-like protein 2 was also reported in pancreatic cancer[48].

ACTG1 is another type of motor activity protein that was detected as an up-regulated protein. It is a protein involved in the formation of filaments, which is the major component of the cytoskeleton[49]. Its existence in colorectal cancer was also reported by Vadlamudi and Shin[50]. Alteration of cytoskeletal proteins may have an important role in cancer initiation or progression. Mutations affecting four major cytoskeletal components have now been identified in human neoplastic cells. Chou, *et al*[51] postulated that mutated cytoskeletal genes may be members of a class of oncogenes, which are fundamentally different from both the nuclear-acting and growth factor/receptor/protein kinase-related types of oncogenes.

In this study, a number of the differentially expressed proteins extracted from the normal and cancerous colonic tissues were identified. Amongst the proteins identified, some of the proteins have their roles identified in similar or other types of cancers. There are also a series of hypothetical proteins with unknown functions. Some of these proteins have their primary location at the cell membrane, which may serve as potential antigens for drug-targeted therapy or as candidates for vaccines against colon cancer. These proteins are under further investigation in our laboratory.

ACKNOWLEDGMENTS

We would like thank the Ministry of Science, Technology and Innovation in Malaysia for providing the grant to carry out this project. We also want to extent our gratitude to the Ministry of Health in Malaysia for providing tissue specimens for the analysis.

REFERENCES

1 **Fearon ER**. Molecular biology of gastrointestinal cancers. In: De Vita VT Jr, Hellman S, Rosenberg SA. (eds). Cancer: principles and practice of oncology. 6th edition. Philadelphia: Lippincott, Williams, and Wilkins Press, 2001: 1037-1049
2 **Harkins L**, Volk AL, Samanta M, Mikolaenko I, Britt WJ, Bland KI, Cobbs CS. Specific localisation of human cytomegalovirus nucleic acids and proteins in human colorectal cancer. *Lancet* 2002; **360**: 1557-1563
3 **Fearon ER**, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990; **61**: 759-767
4 **Alaiya AA**, Franzen B, Auer G, Linder S. Cancer proteomics: from identification of novel markers to creation of artifical learning models for tumor classification. *Electrophoresis* 2000; **21**: 1210-1217
5 **Stulik J**, Hernychova L, Porkertova S, Knizek J, Macela A, Bures J, Jandik P, Langridge JI, Jungblut PR. Proteome study of colorectal carcinogenesis. *Electrophoresis* 2001; **22**: 3019-3025
6 **Chong BE**, Hamler RL, Lubman DM, Ethier SP, Rosenspire AJ, Miller FR. Differential screening and mass mapping of proteins from premalignant and cancer cell lines using nonporous reversed-phase HPLC coupled with mass spectrometric analysis. *Anal Chem* 2001; **73**: 1219-1227
7 **Emmert-Buck MR**, Gillespie JW, Paweletz CP, Ornstein DK, Basrur V, Appella E, Wang QH, Huang J, Hu N, Taylor P, Petricoin EF 3rd. An approach to proteomic analysis of human tumors. *Mol Carcinog* 2000; **27**: 158-165
8 **Molloy MP**, Herbert BR, Walsh BJ, Tyler MI, Traini M, Sanchez JC, Hochstrasser DF, Williams KL, Gooley AA. Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. *Electrophoresis* 1998; **19**: 837-844
9 **Gam LH**, Aishah L. Identification and confirmation of urinary human chorionic gonadotropin using in-gel digestion and HPLC / MS / MS analysis. *Jurnal Biosains* 2002; **13** (1): 27-34
10 **Hjelmeland LM**. Solubilization of native membrane proteins. *Methods Enzymol* 1990; **182**: 253-264
11 **Rabilloud T**. Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis. *Electrophoresis* 1998; **19**: 758-760
12 **Henzel WJ**, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci USA* 1993; **90**: 5011-5015
13 **Figeys D**, Gygi SP, Zhang Y, Watts J, Gu M, Aebersold R. Electrophoresis combined with novel mass spectrometry techniques: powerful tools for the analysis of proteins and proteomes. *Electrophoresis* 1998; **19**: 1811-1818
14 **Duvet S**, Op De Beeck A, Cocquerel L, Wychowski C, Cacan R, Dubuisson J. Glycosylation of the hepatitis C virus envelope protein E1 occurs posttranslationally in a mannosylphosphoryldolichol-deficient CHO mutant cell line. *Glycobiology* 2002; **12**: 95-101
15 **Fountoulakis M**, Tsangaris G, Oh JE, Maris A, Lubec G. Protein profile of the HeLa cell line. *J Chromatogr A* 2004; **1038**: 247-265
16 **Williams K**, Chubb C, Huberman E, Giometti CS. Analysis of differential protein expression in normal and neoplastic human breast epithelial cell lines. *Electrophoresis* 1998; **19**: 333-343
17 **Leong, ASY**, Cooper K, Leong FJM. Manual of diagnostic antibodies for immunohistology. London: Greenwich Medical Media Pub, 1999: 25-35
18 **Hembrough TA**, Kralovich KR, Li L, Gonias SL. Cytokeratin 8 released by breast carcinoma cells in vitro binds plasminogen and tissue-type plasminogen activator and promotes plasminogen activation. *Biochem J* 1996; **317** (Pt 3): 763-769
19 **Ludowyke RI**, Warton K, Scurr LL. Inhibition of antigen and calcium ionophore induced secretion from RBL-2H3 cells by phosphatase inhibitors. *Cell Biol Int* 1998; **22**: 855-865
20 **Birchmeier C**, Birchmeier W, Brand-Saberi B. Epithelial-mesenchymal transitions in cancer progression. *Acta Anat (Basel)* 1996; **156**: 217-226
21 **Seshadri R**, Raymond WA, Leong AS, Horsfall DJ, McCaul K. Vimentin expression is not associated with poor prognosis in breast cancer. *Int J Cancer* 1996; **67**: 353-356
22 **Solomon A**, Murphy CL, Weaver K, Weiss DT, Hrncic R, Eulitz M, Donnell RL, Sletten K, Westermark G, Westermark P. Calcifying epithelial odontogenic (Pindborg) tumor-associated amyloid consists of a novel human protein. *J Lab Clin Med* 2003; **142**: 348-355
23 **Afjehi-Sadat L**, Shin JH, Felizardo M, Lee K, Slavc I, Lubec G. Detection of hypothetical proteins in 10 individual human tumor cell lines. *Biochim Biophys Acta* 2005; **1747**: 67-80
24 **Simpson-Haidaris PJ**, Rybarczyk B. Tumors and fibrinogen. The role of fibrinogen as an extracellular matrix protein. *Ann N Y Acad Sci* 2001; **936**: 406-425
25 **Edwards RL**, Rickles FR, Moritz TE, Henderson WG, Zacharski LR, Forman WB, Cornell CJ, Forcier RJ, O'Donnell JF, Headley E. Abnormalities of blood coagulation tests in

patients with cancer. *Am J Clin Pathol* 1987; **88**: 596-602
26 **Ogura T**, Hirao F, Tsubura E, Yamamura Y. Autoradiographic study on localization of fibrinogen in rabbit lung cancer. *Gann* 1971; **62**: 367-372
27 **Wojtukiewicz MZ**, Zacharski LR, Memoli VA, Kisiel W, Kudryk BJ, Rousseau SM, Stump DC. Abnormal regulation of coagulation/fibrinolysis in small cell carcinoma of the lung. *Cancer* 1990; **65**: 481-485
28 **Cardinali M**, Uchino R, Chung SI. Interaction of fibrinogen with murine melanoma cells: covalent association with cell membranes and protection against recognition by lymphokine-activated killer cells. *Cancer Res* 1990; **50**: 8010-8016
29 **Atagi S**, Sone S, Fukuta K, Ogura T. Inhibition by fibrin coagulation of lung cancer cell destruction by human interleukin-2-activated killer cells. *Jpn J Cancer Res* 1992; **83**: 1088-1094
30 **Nagy JA**, Masse EM, Herzberg KT, Meyers MS, Yeo KT, Yeo TK, Sioussat TM, Dvorak HF. Pathogenesis of ascites tumor growth: vascular permeability factor, vascular hyperpermeability, and ascites fluid accumulation. *Cancer Res* 1995; **55**: 360-368
31 **Yamaguchi T**, Kimura H, Yokota S, Yamamoto Y, Hashimoto T, Nakagawa M, Ito M, Ogura T. Effect of IL-6 elevation in malignant pleural effusion on hyperfibrinogenemia in lung cancer patients. *Jpn J Clin Oncol* 2000; **30**: 53-58
32 **Maguire PB**, Moran N, Cagney G, Fitzgerald DJ. Application of proteomics to the study of platelet regulatory mechanisms. *Trends Cardiovasc Med* 2004; **14**: 207-220
33 **Broude NE**, Modyanov NN, Monastyrskaya GS, Sverdlov ED. Advances in Na+,K+-ATPase studies: from protein to gene and back to protein. *FEBS Lett* 1989; **257**: 1-9
34 **Whittington DA**, Waheed A, Ulmasov B, Shah GN, Grubb JH, Sly WS, Christianson DW. Crystal structure of the dimeric extracellular domain of human carbonic anhydrase XII, a bitopic membrane protein overexpressed in certain cancer tumor cells. *Proc Natl Acad Sci USA* 2001; **98**: 9545-9550
35 **Tureci O**, Sahin U, Vollmar E, Siemer S, Gottert E, Seitz G, Parkkila AK, Shah GN, Grubb JH, Pfreundschuh M, Sly WS. Human carbonic anhydrase XII: cDNA cloning, expression, and chromosomal localization of a carbonic anhydrase gene that is overexpressed in some renal cell cancers. *Proc Natl Acad Sci USA* 1998; **95**: 7608-7613
36 **Cuatico W**, Cho JR, Spiegelman S. Particles with RNA of high molecular weight and RNA-directed DNA polymerase in human brain tumors. *Proc Natl Acad Sci USA* 1973; **70**: 2789-2793
37 **Fu H**, Subramanian RR, Masters SC. 14-3-3 proteins: structure, function, and regulation. *Annu Rev Pharmacol Toxicol* 2000; **40**: 617-647
38 **Ichimura T**, Isobe T, Okuyama T, Takahashi N, Araki K, Kuwano R, Takahashi Y. Molecular cloning of cDNA coding for brain-specific 14-3-3 protein, a protein kinase-dependent activator of tyrosine and tryptophan hydroxylases. *Proc Natl Acad Sci USA* 1988; **85**: 7084-7088
39 **Masters SC**, Subramanian RR, Truong A, Yang H, Fujii K, Zhang H, Fu H. Survival-promoting functions of 14-3-3 proteins. *Biochem Soc Trans* 2002; **30**: 360-365
40 **Friedman DB**, Hill S, Keller JW, Merchant NB, Levy SE, Coffey RJ, Caprioli RM. Proteome analysis of human colon cancer by two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics* 2004; **4**: 793-811
41 **Yang P**, Cunningham JM, Halling KC, Lesnick TG, Burgart LJ, Wiegert EM, Christensen ER, Lindor NM, Katzmann JA, Thibodeau SN. Higher risk of mismatch repair-deficient colorectal cancer in alpha(1)-antitrypsin deficiency carriers and cigarette smokers. *Mol Genet Metab* 2000; **71**: 639-645
42 **Rosenberg S**, Barr PJ, Najarian RC, Hallewell RA. Synthesis in yeast of a functional oxidation-resistant mutant of human alpha-antitrypsin. *Nature* 1984; **312**: 77-80
43 **Harris CC**, Cohen MH, Connor R, Primack A, Saccomanno G, Talamo RC. Serum alpha1-antitrypsin in patients with lung cancer or abnormal sputum cytology. *Cancer* 1976; **38**: 1655-1657
44 **Govorukhina NI**, Keizer-Gunnink A, van der Zee AG, de Jong S, de Bruijn HW, Bischoff R. Sample preparation of human serum for the analysis of tumor markers. Comparison of different approaches for albumin and gamma-globulin depletion. *J Chromatogr A* 2003; **1009**: 171-178
45 **Bloom GS**, Endow SA. Motor proteins 1: kinesins. *Protein Profile* 1995; **2**: 1105-1171
46 **Venkatesh LK**, Gettemeier T, Chinnadurai G. A nuclear kinesin-like protein interacts with and stimulates the activity of the leucine-rich nuclear export signal of the human immunodeficiency virus type 1 rev protein. *J Virol* 2003; **77**: 7236-7243
47 **Scanlan MJ**, Gout I, Gordon CM, Williamson B, Stockert E, Gure AO, Jager D, Chen YT, Mackay A, O' Hare MJ, Old LJ. Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. *Cancer Immunity* 2001; **1**: 4-21
48 **Fontijn RD**, Goud B, Echard A, Jollivet F, van Marle J, Pannekoek H, Horrevoets AJ. The human kinesin-like protein RB6K is under tight cell cycle control and is essential for cytokinesis. *Mol Cell Biol* 2001; **21**: 2944-2955
49 **Kabsch W**, Vandekerckhove J. Structure and function of actin. *Annu Rev Biophys Biomol Struct* 1992; **21**: 49-76
50 **Vadlamudi RK**, Shin J. Genomic structure and promoter analysis of the p62 gene encoding a non-proteasomal multiubiquitin chain binding protein. *FEBS Lett* 1998; **435**: 138-142
51 **Chou CC**, Davis RC, Fuller ML, Slovin JP, Wong A, Wright J, Kania S, Shaked R, Gatti RA, Salser WA. Gamma-actin: unusual mRNA 3'-untranslated sequence conservation and amino acid substitutions that may be cancer related. *Proc Natl Acad Sci USA* 1987; **84**: 2575-2579

S- Editor Wang J **L- Editor** Lutze M **E- Editor** Ma WH

Asian Conference on Clinical Pharmacy

The Future of Asian Clinical Pharmacy

Capitalizing on Diversity

23 - 26 July 2005 · Penang, Malaysia

Editors

Syed Azhar Syed Sulaiman
Ab Fatah Ab Rahman
Mohamed Izham Mohamed Ibrahim
Rosman Ahmad

Published by

National Poison Centre of Malaysia, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia
Tel: +604 657 0099 Fax: +604 656 8417 e-mail: pmnet@pm.usm.my website: http://www.pm.usm.my

Learning Objective(s)

1) Identify opportunities for collaboration among schools and colleges of pharmacy and local health departments.

2) Describe an interdisciplinary process for student education in travel clinics and mass influenza immunization clinics.

3) Identify possible long-term benefits to public health departments that may result from collaboration with a school or college of pharmacy

MSP20

Method Development for Proteomic Analysis of Human Colorectar Cancer Tissue

Gam Lay Harn[1], Leow Chiuan Herng[1], Che Nin Man[2], Gooi Boon Hui[3], Manjit Singh a/l Karam Singh[3]
[1]School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia, [2]National Poison Centre, Universiti Sains Malaysia, Penang, Malaysia, [3]Department of Surgery, Hospital Pulau Pinang, Malaysia

Colon and rectum are parts of the digestive system. Cancer affecting either of these organs may also be called colorectal cancer. Colorectal cancer is worldwide public health cancer and it is among the best characterized cancers with regard to the genetic progression of the disease. Colorectal cancer is also a frequent cause of mortality and morbidity in developed, developing and industrialized countries. In Malaysia, cancer of colorectal ranks second most common cancers and this has also been the case for both male and female. In Malaysia, the Chinese had the higher number of cases, followed by the Indians and the Malays. In this study, a method was developed for analysis of proteins from human colon tissues, whereby proteomics approach was applied for extraction, separation and identification of the proteins. The colon tissues were extracted using sequential and non-sequential protein extraction techniques. In both the techniques, six types of protein extraction buffers were developed. The proteins extracted were separated according to their molecular masses by using 10% SDS-PAGE. The gels were then stained with either Commassie blue or silver. The image of the gel was captured using high resolution scanner. The target protein bands were excised from the gel and proteins were digested In-gel with trypsin. The tryptic peptides were then analyzed using LC/MS/MS, which allows sequencing of the analyzed peptides. The MS/MS product ions spectrum were then search against Mascot protein database search engine. A total of eighty two proteins were identified from the normal and cancer tissues. Amongst these proteins, the abundance proteins groups are protein binding proteins, structural proteins, transport pritens, hypothetical proteins and enzymes. In addition, a minority of proteins detected were cytoskeleton-binding proteins, endopeptidase inhibitor proteins, motor activity proteins, channels proteins, defense response proteins, DNA-RNA binding protein, lipoproteins and ATP-binding protein.

29TH ANNUAL CONFERENCE OF THE MALAYSIAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY

28TH - 29TH SEPTEMBER 2004

NIKKO HOTEL KUALA LUMPUR

Organised by

BABBIC

THE MALAYSIAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY

PROTEOMICS ANALYSIS OF DIFFERENTIALLY EXPRESSED PROTEINS EXTRACTED FROM MALIGNANT COLON TISSUES AS COMPARED TO NORMAL COLON TISSUES

Gam Lay Harn[1], Leow Chiuan Herng[2] & Manjit Singh A/L Karam Singh[1]

[1]School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang;
[2]Department of Surgery, Hospital Pulau Pinang, Penang

Colorectal cancer continues to be the second leading cause of death for both men and women (Williams et al., 1996), despite major advances in basic research over the last two decades, response rate to therapy remain poor. Prognostic uncertainty could be overcome, in part by defining the changes that occur in the malignant tumor at either the gene (genomics) or protein (proteomics) levels and much attention has naturally focused on comparisons of the biochemistry of normal and malignant cells. The roles of proteins as the functional components of a cell are ultimately critical in the understanding of the development of cancer. The knowledge of the differentially expressed proteins of malignant colon cells compared to normal colon cells may lead to the understanding of the changes in the biochemical reactions of cancer cells. Such tumor-associated changes may identify new markers and new target for therapeutic intervention.In our present study, proteomics approach was applied in order to detect and analyze the differentially expressed proteins in malignant colon tissues compared to normal colon tissues. The tissues were collected after informed consents were obtained from the patients. The comparison study was performed on the normal and cancerous tissues specimens that were from the same patients in order to minimize the variation amongst patients. Cancerous and normal tissues from four patients, which were graded as moderately to poorly differentiated stage of cancer were analyzed. The tissues were homogenized and the proteins were extracted using an aqueous buffer. Protein concentrations were determined and the protein mixtures were then subjected to SDS-PAGE separation. The protein profiles obtained from cancerous and normal tissues were compared in order to identify the target proteins for further analysis. The results show that the proteins that were expressed amongst the normal tissues and amongst the cancerous tissues, respectively are very consistence. The protein bands that were consistently present in all the four cancerous tissues but absent in all the normal tissues (or vice versa) were excised and in-gel digestion of proteins were carried out. The digested fragments of proteins were then analyzed using LC/MS/MS analysis. The MS/MS data obtained were subjected to Mascot proteins database search program and from which the identities of the target proteins were identified. Using this method, we were able to identify 10 proteins that were differentially expressed in cancerous tissues compared to normal tissues. Amongst the proteins identified is serum albumin precursor.