International Journal of Pharmacology

ISSN 1811-7775

science alert

ansinet
Asian Network for Scientific Information

OPEN ACCESS International Journal of Pharmacology

ISSN 1811-7775
DOI: 10.3923/ijp.2017.425.437

Research Article
Identification of Glycobiomarker Candidates for Breast Cancer Using LTQ-Orbitrap Fusion Technique

[1,2]Abdullah Saleh Al-wajeeh, [2]Mohd Nazri Ismail, [3]Salizawati Muhamad Salhimi, [4]Imran Abdul Khalid, [3]Sultan Ayesh Mohammed Saghir, [1]Mohammed Alsayrafi, [1]Thomas Michael Harvey and [1]Aishah Binti A. Latiff

[1]Anti-Doping Lab Qatar, Doha, Qatar
[2]Analytical Biochemistry Research Centre, Universiti Sains Malaysia, 11800 Penang, Malaysia
[3]School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia
[4]Seberang Jaya Hospital, Penang, Malaysia

Abstract

Background and Objective: Proteomics has been developed remarkably and become an intensive field of research; especially protein glycosylation which is considered one of the most common post-translational modifications. Glycosylated proteins play vital roles in several biological and molecular pathways and it represents over 50% of human proteins. The aim of this study was to examine differentially expressed glycoprotein profiles of breast cancer and its adjacent tissues among breast cancer women in Malaysia. **Materials and Methods:** A total of 48 pairs of normal and cancerous tissues were collected from breast cancer women at two hospitals in Penang, Malaysia: Seberang Jaya Hospital (SJH) and Kepala Batas Hospital (KBH) from January 2010- January 2013. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) was employed to map protein expression profiles. Gel-Eluted Liquid Fractionation Entrapment Electrophoresis (GELFREE) system was used for separation and fractionation of protein extracts. The protein extracts were subjected to molecular weight fractionation via electrophoretic separation through SDS PAGE material in tubular environment, all in the liquid phase. The enrichment of GELFREE fractions was carried out using glycopeptide enrichment strategy. After that, they were fragmented by tandem mass spectrometry (LC-MS/MS) analysis using LC/MS LTQ-Orbitrap Fusion. The protein profiling was done using PEAKS Client (7.5) software. **Results:** In this study, 11 glycoproteins were detected only in tumour tissues, but not in normal ones. Extracted protein concentrations, number of proteins and glycopeptides were found to be higher in tumour tissue samples compared to normal tissue samples. **Conclusion:** The findings of this study provided useful insights into the glycoproteins identification and could establish fundamental insights for further differential glycoproteomics profiling for breast cancer biomarker discovery.

Key words: Proteomics, glycoprotein biomarker, post-translational modifications, LTQ-Orbitrap Fusion, breast cancer

Received: February 12, 2017 Accepted: April 24, 2017 Published: June 15, 2017

Citation: Abdullah Saleh Al-wajeeh, Mohd Nazri Ismail, Salizawati Muhamad Salhimi, Imran Abdul Khalid, Sultan Ayesh Mohammed Saghir, Mohammed Alsayrafi, Thomas Michael Harvey and Aishah Binti A. Latiff, 2017. Identification of glycobiomarker candidates for breast cancer using LTQ-Orbitrap fusion technique. Int. J. Pharmacol., 13: 425-437.

Corresponding Authors: Abdullah Saleh Al-wajeeh, Anti-Doping Lab Qatar, Doha, Qatar Tel: +97444132908/+97433591803
Mohd Nazri Ismail, Analytical Biochemistry Research Centre, Universiti Sains Malaysia, 11800, Penang, Malaysia Tel: 04-653 2694

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Proteomics approaches using Mass Spectrometry (MS) are considered the most widely approaches used to study the protein expression among normal and abnormal tissues. Proteomics has the capacity to identify the biomarkers and find out the signatures for disease diagnosis or prognosis[1,2]. It can also help in explaining the molecular mechanism behind disease development and progression[3-5].

Breast cancer is known as the most common type of cancers affecting women. Globally, breast cancer has been reported to be the critical common health problem and one of the major causes of cancer deaths[6,7]. About 1,384,155 million women are affected with breast cancer, with almost 459,000 reported deaths, which indicates nearly 14% of deaths caused by cancer[8,9].

Based on the previous statistics, the occurrence of breast cancer in women all over the world will reach approximately 3.2 million new cases per year by 2050[6]. Taking into account that reports in 2003-2005 that have mentioned 47.3 per 100.00 Age-Standardised Rates (ASR), the Malaysian National Cancer Registry (NCR) has put breast cancer as the most widely identified cancerous disease among Malaysian women. The occurrence of breast cancer among the main ethnic groups in Malaysia is distributed as follows: Indians (54.2 per 100, 000) and Malays (34.9 per 100, 000). Specifically, 48 cases per 1000,000 were reported by the Penang Cancer Registry in 2004-2008. Furthermore, the International Agency for Research in Cancer (GLOBOCAN) 2012 estimated the ASR of breast cancer in Malaysia as 38.7 per 100,000 with 5410 new cases in 2012[10].

Protein technologies are useful as they can be employed in recognizing protein or protein profiles in order to improve the classification of the subtypes of breast cancer or uncovering novel protein biomarkers, which can indicate significant metastatic potential or therapy resistance therapy for each definite subtype[2]. Most proteins undergo some form of Post-Translational Modifications (PTMs) through the addition of phosphorylation, acetylation, glycosylation, methylation and ubiquitylation or the deletion of parts to make a finished protein[6].

According to Pagel *et al.*[11], some functions of PTMs can be detected in the generation of active forms of proteins, cell signalling, coding proteins for transport to specific compartments and making polypeptides for degradation.

The onset of modern technique such as the LTQ-Orbitrap Fusion technique offers several advantages compared with the previous used methods[12]. It has the capability to produce high resolution, mass accuracy and MSn of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS) with a reasonable price. The LTQ-Orbitrap can also used to generate top-down data using protein standards ranging from 10 to 25 kDa and characterize recombinant monoclonal antibodies[13-15].

Carbohydrates are essential for cell metabolism and energy production and are building blocks of the extracellular matrix. Among post-translational modifications of proteins, glycosylation is the most common one[16]. Furthermore, it was reported that almost one-half of all mammalian proteins are glycosylated[17]. Glycans (N-glycosylation and O-glycosylation) are the major types of glycosylation forms which are included in cellular identification and signalling progressions[18]. In human cells, the most common covalent modification of proteins is N-glycosylation[18,19]. Among all PTMs proteins, glycosylation either O-linked or N-linked is the significant in the extracellular and on the cell surface as it shows high sensitivity to biochemical environment[19-21]. This study designed to characterize the differences in protein expressions between normal and malignant tissue. As well as, to examine the glycoproteomic profiles of breast cancer and its adjacent normal tissues using proteomics techniques such as GELFREE fractionation system, LTQ-Orbitrap and bioinformatic software PEAKS.

MATERIALS AND METHODS

Tissue collection: This study was conducted after obtaining the Ethical approval from Human Ethics Committee of the Ministry of Health Malaysia and Human Ethical Clearance Committee of Universiti Sains Malaysia (USM) (USMKK/PPP/JEPeM [(211.3.(15)]). Furthermore, consent forms were obtained from patients or their close relatives before sample collection (tissue breast cancer). From January 2010- January 2013, 48 pairs of normal and cancerous tissues from breast cancer were gathered from two hospitals: Seberang Jaya Hospital (SJH) and Kepala Batas Hospital (KBH) in the mainland of Penang. Collection of these samples was done after hospital pathologist had confirmed that these samples were as normal and cancerous, respectively. The proportion that was taken from the breast cancer tissue was almost 10 mm^3 of tumour tissue. Nevertheless, it should be noted that there was a change in the location of taking samples of breast cancer tissue for the control group tissue. Specifically, the location of the samples in the control group was decided to be within 10 cm of the breast cancer tissue.

Tissue processing: The tissue specimens were stored at -80°C until analysis. Prior to the extraction of protein, the samples underwent two essential consecutive things: first, they were unfrozen at room temperature for a definite portion of time and second, they were washed off with cold distilled water. This was followed by getting rid of fatty layers on the tissues using a scalpel. After this, the tissues were chopped into tiny pieces which were weighed, labelled and put into micro centrifuge tubes to make them ready for the analysis.

Tissue homogenization and protein extraction: Tissues were first weighed, ground in a mortar using liquid nitrogen and homogenized in ice using a glass homogenizer. Then, they were subjected to protein extraction using lysis buffer which was composed of 25 mM Tris, 150 M NaCL, 5 mM EDTA and 1% CHAPS, with an adjusted pH 7.4[22]. After that, each homogenized tissue was subjected to two things: Vortexing for a short period of 30 sec and centrifugation at 1,500 rpm at 4°C for 10 min. This was followed by transferring supernatant into extremely good quality dialysis tubes (7,000 MWCO) and sealing it. A cool dialysis buffer (50 mM ammonium bicarbonate pH 7.5) was used to keep homogenate tissue at 4°C. The dialyses buffer was changed with a fresh buffer every 48 h. Then, the homogenate sample was collected in a microcentrifuge tube and kept at -80°C until the analysis started.

Bradford assay: The homogenate sample was placed in a microcentrifuge tube and the total protein concentration was determinate using Bradford protein assay (Bio-Rad, USA). Total protein quantification for the each tissue sample was carried out according to Bradford[23]. Each sample of 5 µL was mixed with 250 µL of Bradford reagent in a 96 well plate. The solution was then incubated at room temperature for 15 min. A Bovine Serum Albumin (BSA) as standard with wavelength absorbance at 595 nm was conducted and the standard curve was ranging from 0.0-2 mg mL^{-1}. The total protein concentration was measured and averaged by the comparison of the absorbance value obtained for the sample and the standard curve.

GELFREE 8100 fractionation system for protein fractionation: As explained by Witkowski and Harkins[24], fractionation of the protein was performed on GELFREE 8100 fractionation system (Expedeon, CA, USA). For this study, 10% Tris-Acetate cartridge was selected. Considering the complexity of the protein extracts found in tissue specimens, reduction of this complexity was done through the application of molecular weight fractionation on 150 µg protein extracts via electrophoretic separation through SDS-PAGE material in tubular environment. All fractions were created in the liquid phase Gel-Eluted Liquid Fraction Entrapment Electrophoresis (GELFREE), resulting in 12 fractions (MW) in the liquid phase for each examined protein extract. All 12 GELFREE fractions were combined into four groups, with 3 fractions in each. The total protein concentration for each group was measured. A 50 µg protein from the combined fractions was frozen for 30 min at -80°C and freeze-dried overnight to prepare it for the next analysis

Image analysis: The determination of GELFREE fractionation of proteins and the quality of separation were done through the application of SDS-PAGE using 12.5% resolving slabs gels. As shown in Fig. 1, Precision Plus Protein™ SDS-PAGE Standards (Bio-Rad, USA) was loaded into the first and last lanes of each gel. Scanning the gel was done through G: Box chemi-xx8 GeneSys image analyser (SYNGEN, Japan). The image (GeneSys V1.3.9.0 -SYNGEN) was captured by the GeneSys program (Fig. 1).

In solution tryptic digestion: According to Ru *et al.*[25], 50 µg protein which was taken from each GELFREE combined fraction was exposed to In-Solution Tryptic Digestion method. This procedure was done to prepare the sample before being subjected to nanoflow UHPLC separation and high resolution MS and MS/MS analysis. At the beginning, the re-suspension of each of these combined fractions was done in 6 M guanidine-HCL/25 mM ammonium bicarbonate (NH_4HCO_3) pH 8.5. Next, it was reduced by 250 µL of 1.0 mg mL^{-1} Dithiothreitol (DTT) in 25 mM NH_4HCO_3. Then, it was incubated at 55°C for 30 min and alkylated with 500 µL of 1 mg mL^{-1} iodoacetamide (IAA) in 25 mM NH_4HCO_3 (1 mg mL^{-1} IAA/25 Mm NH_4HCO_3). Then, it was subjected to incubation for 15 min in the dark at 55°C. In order to make the protein concentrated and to desalt the reduced and alkylated protein samples, buffer exchange was applied with 25 mM ammonium bicarbonate using spin-column with molecular cut-off of 3 kDa 3 times at 3500 rpm for 45 min. After that, the addition of 1.0 µL of the reconstituted 1.0 µg $µL^{-1}$ trypsin in the protein solution (ratio 1:50) was done. This was followed by the incubation of the sample at 37°C for 18 h. This incubation was done overnight and followed by the addition of 0.1 µL of formic acid, which was done to stop the trypsin reaction. The digested peptides were used for examining the glycoproteomics.

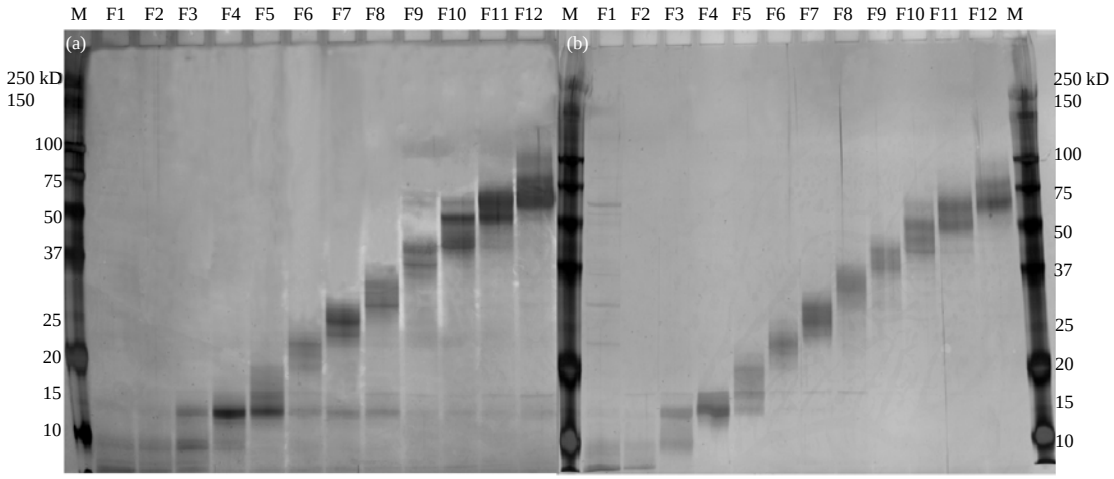

Fig. 1(a-b): Protein bands from GELFREE fractionated samples (sample no. 109) as seen on SDS-PAGE, (a) Breast normal tissue and (b) Breast cancer tissue

Detection and analysis of Glycopeptides: As described by Ru *et al.*[25] each (50 µL) GELFREE combined fraction was subjected to the tryptic digestion, which was followed by glycoproteomics analysis of peptides fractions. In consistent with glycopeptides/glycoprotein enrichment Top Tip carbon (Graphite), purification and enrichment were done during the detection of glycopeptides. For releasing the oligosaccharides (glycan) from proteins, enzymatic deglycosylation was applied to the enriched glycopeptides. This was done by the application of enzyme PNGase F (V483A, Promega, USA). The detection of the types of glycan structures was done by employing Enzyme PNGase F, which was further fragmented for structure analysis by mass spectrometry[26].

Glycoproteomic enrichment: Based on Glycopeptide /Glycoprotein Enrichment Top Tip carbon (graphite) product instructions (Glygen, Part No. TT2CAR), glycopeptides were purified from the peptide fractions and enrichment. Gentle tapping of Top Tip was done in order to prevent any packing material to get stuck to the top white caps which were taken out from top and bottom. For the purpose of wetting the packing material, 50 µL of the releasing solution was added through a pipette tip inserted in the top of the Top Tip. The Top Tip was fixed to a pipette or syringe and air pressure was employed to push the solution through the packed bed. The TopTip was removed from the pipette through washing for 2-3 times. This washing was done with the binding solution by applying the sample solution as described earlier. The Top Tip was connected to a pipette or syringe and the liquid was pushed strongly the packed bed. The washing of the sample was followed by washing the packed bed for 2-3 times with 50 µL of binding solution for eluting salts and other non-retained components. Finally, washing the packed bed with 25-50 µL of releasing solution (bed volume is 25 µL) released the sample fraction. The repetition of the washing and combination of the eluents was done to elute all adsorbed peptides. To prepare the solvent for the analysis, it was subjected to direct evaporation. Additionally, the concentration of the glycopeptides was done as a procedure to prepare them for the analysis. Based on the accurate mass and known biosynthetic pathways, structures were assigned by using GlycoWork Bench platform.

LC-MS/MS: Prior to LC-MS/MS analysis, the peptides samples were mixed with 100 µL of 0.1% formic acid in deionized water and filtered using a 0.45 µm syringe filter. All analyses of MS/LC and MS/MS were performed using a Thermo LTQ-Orbitrap fusion mass spectrometer coupled with the nano UHLPC system (Thermo Scientific, USA). The chromatographic separation of tryptic-digested peptides was carried out using easy-column C18 (10 cm, 0.75 mm i.d., 3 µm; Thermo Scientific, USA), which was used as the analytical column. On the other hand, easy-column C18 (2 cm, 0.1 mm i.d., 5 µm; Thermo Scientific, USA) was used as the pre-column. Approximately 1.0 µL of samples were injected and chromatographically separated at a flow rate of 0.3 µL min^{-1}. The mobile phase running buffers used were (A) 0.1% formic acid in deionized water and (B) acetonitrile with 0.1% formic acid. The samples were eluted using the gradient 5.0-100% of buffer B in 85 min.

The mass spectrometer and mass calibration coefficients were calibrated using a mixture of MRFA peptide, caffeine and Ultramark 1600 for normal-mass calibration and PPG 2700 with 7 mM sodium acetate in 70% methanol for high-mass calibration. The X calibur ver. 2.1 (Thermo Scientific, San Jose, CA, USA) with a mass tolerance threshold of 5 ppm was used to perform data acquisition. The eluent was sprayed into the mass spectrometer at 2.1 kV (source voltage) and a capillary temperature of 220°C was used. Peptides were detected using a full-scan mass analysis, from m/z 300-2,000 at a resolving power of 60,000 (at m/z 400, FWHM; 1-s acquisition), with data-dependent MS/MS analyses (ITMS) triggered by the eight most abundant ions from the parent mass list of predicted peptides with rejection of singly or unassigned charge state. The Collision Induced Dissociation (CID) was used as the fragmentation technique with collision energy of 35 eV.

Peaks Client software version 7.5 (Bioinformatics Solution, USA) was used to perform de novo sequencing and database matching. The MS/MS analysis was carried out using similar resolving power (60,000). The CID was applied with isolation width of 2 Da, normalized collision energy of 35 eV, activation q of 0.25, activation time of 50 msec and charge state of 2. Higher-energy Collisional Dissociation (HCD) was applied with isolation width of 2 Da, normalized collision energy of 35 eV, activation time of 0.1 msec and FT first mass value (m/z) of 100[27].

Statistical analysis: In this study, chi square test was used to analyze the statistical significance of the presence of the biomarker in the 40 samples of normal and tumor tissues. A t-test was used to analyze the statistical significance of the difference between the amount of the biomarker in the normal and tumor tissues. In addition, Person correlation was used to evaluate the correlation between glycoproteins of normal and tumor tissues[28].

RESULTS AND DISCUSSION

The extracted proteins were subjected to molecular weight fractionation using liquid phase (GELFREE) fraction (Fig. 1). Then, the fractions were reduced, alkylated and digested by trypsin using in-solution digestion protocol. This was followed by glycopeptides enhancement using glycopeptide enrichment techniques (Glycopeptide/ Glycoprotein Enrichment Top Tip carbon (graphite). In this research, enzymatic deglycosylation was applied to the first parts of the sample using enzyme PNGaseF in order to release the oligosaccharides (glycan) from proteins. Furthermore, it was used to detect types of glycan structures.

Detection of glycopeptides: Confirming with glycopeptides/ glycoprotein enrichment Top Tip carbon (Graphite), purification and enrichment were done during the detection of glycopeptides. Then, the enriched glycopeptides were subjected to enzymatic deglycosylation using enzyme PNGase F (Peptide N-glycanase F) for the purpose of getting rid of the oligosaccharides (glycan) from proteins. As a result of this, the LTQ-Orbitrap Velos pro (Thermo Scientific, USA) was implemented for analysing the results using ETD activation protocol.

Structures were distributed with the aid of GlycoWorkBench platform taking into account two important things: The accurate mass and known biosynthetic pathways (Fig. 2). A noticeable doubly charged N-glycan ion 1134.06 ($z = 2$) in both normal and tumour tissues was observed, while 1488.22 doubly charged N-glycan was detected in only normal tissue (Fig. 2a). On the other hand, there was a discovery of multiple charged N-glycan 903.43, 1089.63 and 1495.23 in tumour tissues (Fig. 2b).

An example of base peak chromatogram as typical full scan spectra depicted from group 3 combined tissue sample acquired on Orbitrap LTQ LC/MS-CID activation protocol are shown in Fig. 3. A number of peaks were observed in the Base Peak chromatogram at run time from 10-120 min. The HPLC is extremely useful for the isolation of peptides and proteins from a wide variety of synthetic or biological sources. Different peaks profiles were occurred in the chromatogram as a result of the elution of different peptides from the combined fraction.

The digestion process of peptides using trypsin leads to produce different sequences of peptides (i.e., formation of peptides in the preferred mass range for effective fragmentation by tandem mass spectrometry). Several peaks were shown in the chromatogram at a range between 10 to 68 min run time (Fig. 3a). Different charge states (z) ranging from 1 to 4 were produced as shown in Fig. 3. Masses which possess charge state of 2 or more has the higher potential to be peptides. A mass of 476.76, 519.93, 647.67 and 600.33 at charge state of 2, 3, 4 and 3, respectively detected in full scan spectra is considered a peptide, while 462.15 with single charge and 508.31 without charge state are not peptides. Therefore, doubly and triply charged masses were selected to be fragmented using CID.

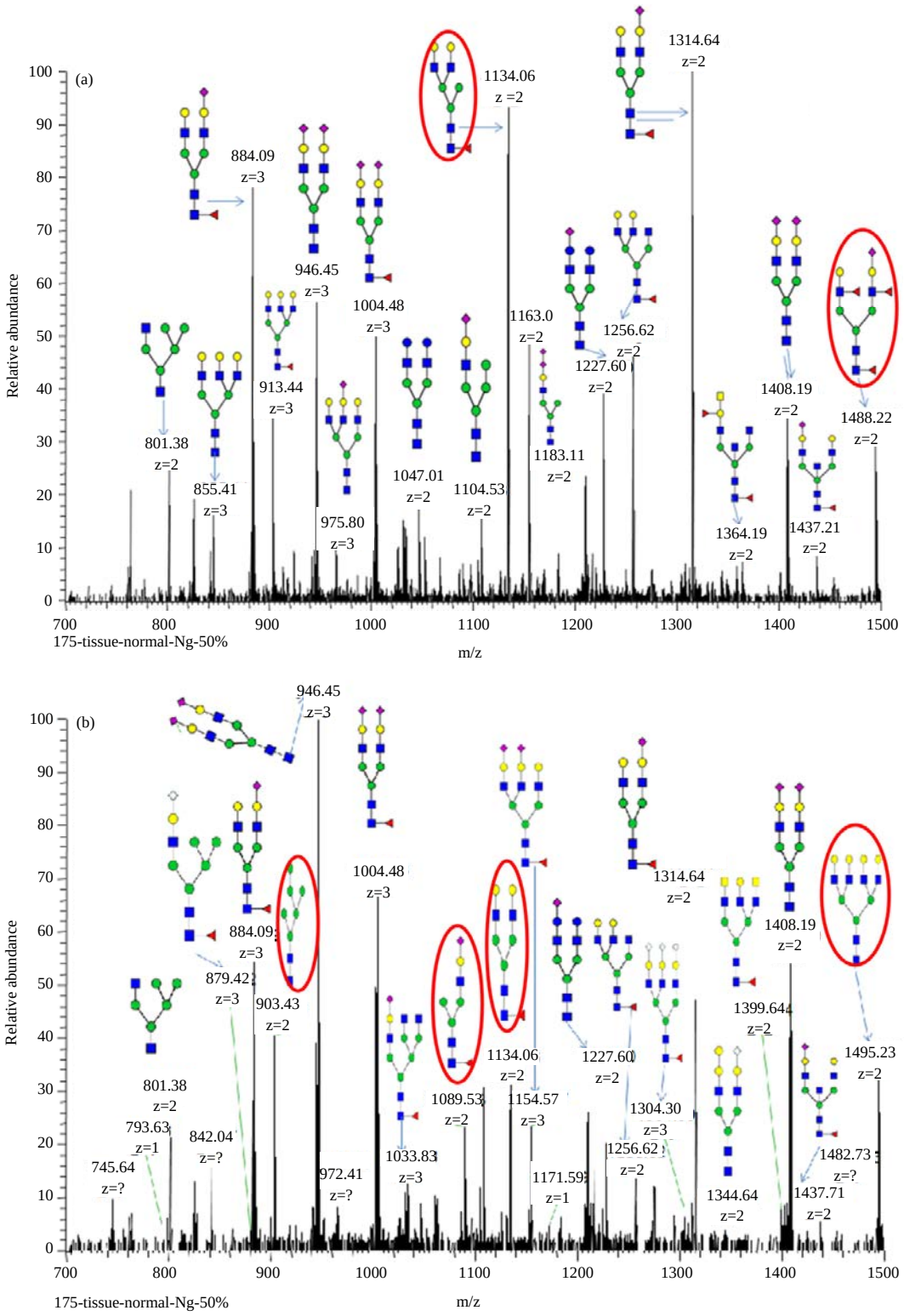

Fig. 2(a-b): Examples of the N-glycome of (a) Normal and (b) Tumour breast tissues

Peptide fragments which were produced due to CID fragmentation are shown in Fig. 3. It is an MS/MS spectrum of the doubly charged molecular mass 612.32 observed in group 3 at the retention time of 12.64 min. Rich sequence information was observed at this point. Subsequently, fragments were submitted for de novo

sequencing followed by PEAKS software database analysis. In fact, the peptide MS/MS spectrum produced by CID activation protocol represents valuable information that illustrates the peptide sequence and consequently match these sequences in the database for protein identification.

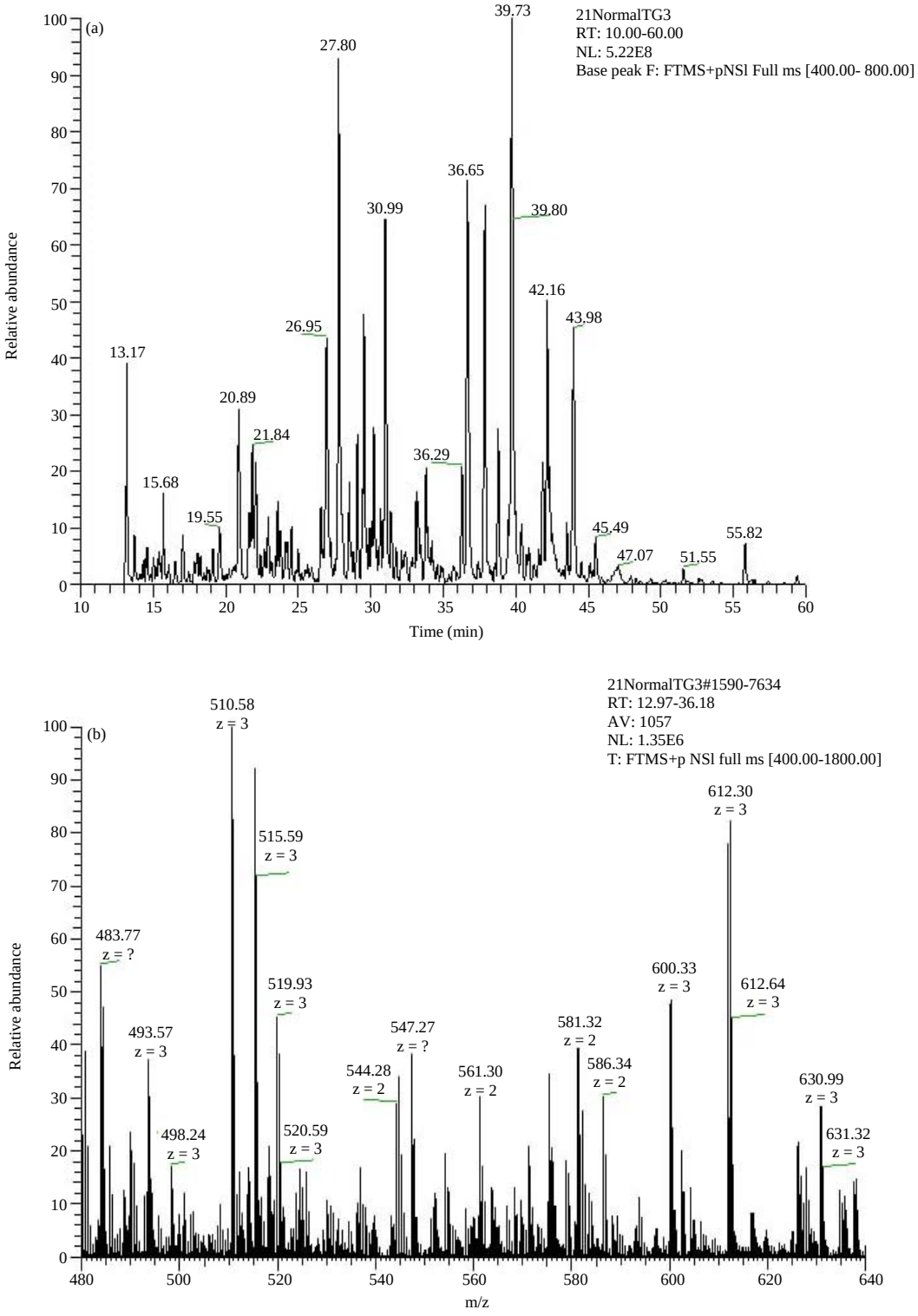

Fig. 3(a-c): Continue

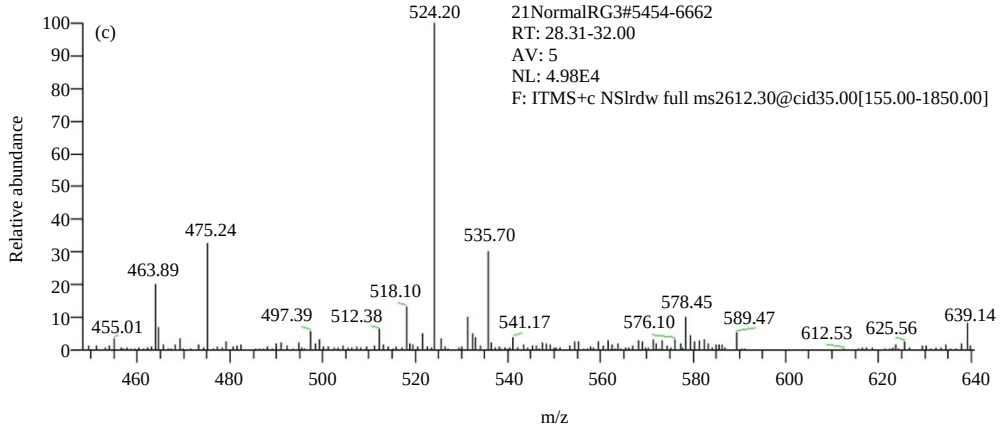

Fig. 3(a-c): (a) Base peak chromatogram of group 3 combined tumour tissue acquired on Orbitrap LTQ LC/MS-ETD activation protocol, (b) Expanded region of a single FTMS full scan information, mass range m/z 400-1600 and (c) MS/MS 655.81at ETD31.21, acquired MS/MS spectrum

The average of glycosylated peptides in tumour tissue samples was higher compared to normal samples (Fig. 4). The means of glycosylation were compared using t-test statistical analysis and the outcomes of this test are shown in Table 1. Compared to the number of glycosylation in normal sample, there was a significant increase in the numbers of glycosylation in tumour sample ($p<0.01$). The result reporting the higher numbers of glycosylation compared with normal tissue samples was in consistent with findings discussed earlier in this study, which reported that protein concentration and profile numbers were higher in tumour compared to normal tissues .

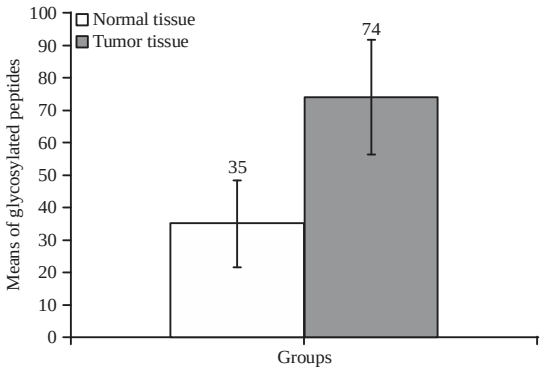

Fig. 4: Number of glycosylated peptides in normal and tumour samples. Values presented as Mean± SD

Glycosylation variable search and glycoprotein identification: The execution of *de novo* sequencing and database matching was carried out using PEAKS Client Version 7.5 (Bioinformatics Solution, Waterloo, Canada). Database matching was performed by employing Uniprot-Homo sapiens (Human). While, carbamidomethylation and methionine oxidation were put as fixed modifications and maximum missed cleavage was adjusted to 2. Furthermore, parent mass and precursor mass tolerance were put at 0.1. Protein acceptance was decided through the employment of False Detection Rate (FDR) <1% and significant score (-10lgP) for protein >20. While the minimum unique peptide was set at 1 and maximum variable post-translational modification was put at 4. In the second process, the search continued in the workflow, PEAKS PTM, where anticipation and addition of post-translation modifications in the alignment were carried out. The parameters were fixed by applying glycosylation variable; Biantennary (ms = 1622.58), Biantennary (-1 galactose) (ms = 1460.53), Biantennary (-2 galactose) (ms = 1298.48), Fucosylated biantennary (ms = 1768.64), Fucosylated biantennary (-1 galactose) (ms = 1606.59), Fucosylated biantennary (-2 galactose) (ms = 1444.53), Hex1HexNAC1 (ms = 365.13) HexNAC acylation (ST) (ms = 203.08) and N-linked glycan core (ms = 1216.42).

Eleven glycoproteins were detected in tumour tissue samples; while they were not detected in normal tissue (number and percentages of proteins in normal tissues were zero as mentioned in Table 2). Chi-square values are statistically significant as they lie between 57.14 and 66.67. Among the identified glycoproteins, the significance up-regulation of tropomyosin alpha-3 chain protein was

Table 1: Comparison between the means of glycosylation peptides in normal and tumour tissue samples using t-test

Type of PTMs	Mean	STD	SEM	t	df	Significance (2-tailed)
Glycosylation						
Normal	38	13.36	2.99	-10.566	19	0.01
Tumour	74	17.94	4.011			

PTMs: Post-translational modifications. STD: Standard deviation, SEM; Standard error mean (n = 40)

Table 2: Detection of glycoproteins in tissue samples

		Normal glycoproteins		Tumour glycoproteins			
ID	Proteins	No. of proteins	%	No. of proteins	%	Chi-square	p-value
1	Filamin-A	0	0	16	80	66.67	0.000
2	Fibrinogen beta chain	0	0	14	70	61.54	0.0001
3	Serotransferrin	0	0	14	70	61.54	0.0001
4	Collagen alpha-3 (VI) chain	0	0	14	70	61.54	0.0001
5	Immunoglobulin lambda-like polypeptide 5	0	0	14	70	61.54	0.0001
6	Ig gamma-2 chain C region	0	0	12	60	57.14	0.0007
7	Gelsolin	0	0	12	60	57.14	0.0007
8	Tropomyosin alpha-4 chain	0	0	12	60	57.14	0.0007
9	Tropomyosin alpha-3 chain	0	0	12	60	57.14	0.0007
10	Gelsolin	0	0	12	60	57.14	0.0007
11	Tropomyosin beta chain	0	0	12	60	57.14	0.0007

detected among Malaysian breast cancer (Table 2). The main significant roles played by this protein are the regulation of muscle contraction and the stabilization of cytoskeleton actin filaments. This finding is consistent with Lee *et al.*[29], who examined breast cancer tissue among Malaysian women. The family of microfilament-associated structural proteins include tropomyosins[29,30]. The expression modifications of tropomyosins were happened by several carcinogens such as chemical carcinogens, UV radiation, DNA and RNA tumour viruses during cancer cell transformation. Similar to the finding of Lee *et al.*[29], the present study reported tropomyosin alpha 3 chain, which is defined to be high molecular weight with 284 amino acids, was found to be up-regulated significantly among Malaysian breast cancer. Furthermore, the finding of the current study concerning the up-regulation of tropomyosin alpha 3 chain is in agreement with Helfman *et al.*[30] who reported this up-regulation in central nervous system tumours. However, the findings of this study regarding tropomyosin alpha 3 chain were not in agreement with those who reported down-regulation of tropomyosin alpha 3 chain in bladder cancer[30].

Serotransferrin (TF): The TF is a glycoprotein that has almost molecular weight of 76.5 kDa and its roles are to manage and to transform iron from the intestine, reticuloendothelial system and liver parenchymal cells to all proliferating cells in the body[31]. To compare the serum profiles between lung cancer patients and healthy persons, Chang *et al.*[32] employed 2-D liquid phase fractionation system (PF2D) and mass spectrometry approach. Their study revealed eight proteins from total eight fractions in the 2nd dimension, including the fibrinogen beta chain, HP, fibrinogen alpha chain, APOA1, transthyretin, serotransferrin, Ig alpha-1 chain and Ig alpha-2 chain. Chang *et al.*[32] had reported some of these proteins as secreted proteins or cancer markers such as fibrinogen, transthyretin, serotransferrin and Ig. Furthermore, in consistent with Chang *et al.*[32], this study revealed that fibrinogen beta chain and serotransferrin are glycoproteins in breast cancer tissue. At the same time, Jung *et al.*[33] reported the existence of serotransferrin as a glycoprotein in plasma and the up-regulation of glycoprotein which is similar to the finding of this study.

Filamin-A: Filamin-A (FLNA) was detected to be up-regulated glycoprotein in breast cancer tissue. However, compared with localized cancers, Bedolla *et al.*[34] revealed high levels of cytoplasmic FLNA and low levels of nuclear FLNA as expressed by metastatic prostate cancer tissues. Furthermore, they showed that the localization of this protein in *in vitro* studies was responsible for the regulation of cell invasion and motility. They also reported that inhibition of Protein Kinase A (PKA) is probably significant method of regulating the activity of FLNA. This is an important issue because if FLNA cytoplasmic localization is a cause of metastasis, the existence of FLNA in the cytoplasm can be considered as a predictive marker of future metastasis, whereas inhibition of PKA may be used as a therapeutic tool to strongly stop this effect[33]. Additionally, Zhao *et al.*[35] exhibited the function of the FLNA as a modulator of chemo sensitivity to docetaxel in Triple-Negative Breast Cancer (TNBC) cells through

regulating MAPK/ERK pathway in both *in vitro* and *in vivo*. The FLNA can function as a novel therapeutic target for improvement of chemotherapy efficacy in TNBC[35].

Fibrinogen beta chain protein as one of the identified glycoproteins, was reported to be up-regulated among Malaysian breast cancer. Fibrinogen is a 350 kDa glycoprotein synthesised essentially by liver epithelium. Fibrinogen is connected with cancer cell growth, progression and metastasis, where elevated levels existed in a number of cancer types[36].

The present research revealed novel nuclear membrane biomarkers for Hepatocellular carcinoma (HCC). Among 76 variously expressed proteins, fibrinogen beta chain protein, was measured across the specimens and outlined into liver fibrosis and carcinoma specific proteins[37]. Dowling *et al.*[38] revealed that fibrinogen beta chain and CA15-3 were demonstrated to have similar p-values and AUC-values for distinguishing stage IV breast cancer from the control, Inflammatory Bowel Disease (IBD) and benign breast disease groups. The IBD and benign breast disease groups were included as control for inflammation/immunological reactions that are connected with cancer at all stages[38].

Dowling *et al.*[38] found that fibrinogen alpha chain, beta chain and gamma chain were remarkably elevated in advanced breast cancer, stressing its important function in cancer growth and development. Discovery phase analysis identified fibrinogen chains as responsible for the transferrin identified, revealing that several of these high abundant proteins may form complex with each other in the blood. In spite of considering fibrinogen in the stroma as a hallmark of breast cancer, the ability of cancer cells to synthesise and secrete fibrinogen should be recognized as a contributor to the elevated levels that were demonstrated in the circulation, especially in patients with a large tumour burden[38].

Collagen alpha-3 (VI) chain: Collagen alpha-3 (VI) chain is commonly considered to be distributed extracellular matrix macromolecule that has a significant function in tissue development and is greatly discovered in cancers[39]. During inflammation in breast cancer, the two important things that produce collagen VI are adipocytes and macrophages. Furthermore, collagen VI increases chemotherapy resistance, marking it as a potential biomarker for cancer diagnosis[40]. Thus, the findings of this study strongly confirm the function of molecule in tumour progression and suggest that they may be employed as a prognostic factor for the treatment of breast cancer.

In this study, significant up-regulation of immunoglobulin lambda-like polypeptide 5 (IGLL5) was reported. This finding is consistent with what was reported by previous studies[41]. In previous study mass spectrometry-based quantitative proteomic approach was employed for determining proteins that were divergently shown in retinoblastoma tumour[41]. Moreover, this study reported that immunoglobulin lambda-like polypeptide 5 (IGLL5) can be considered as a one of novel proteins in retinoblastoma.

An important finding of the present study is the up-regulated proteins as detected in tropomyosin alpha 4 chain; this finding is consistent with what was reported by Harada *et al.*[42]. Furthermore, in this study the up-regulation of tropomyosin beta chain protein was also reported. It is significant to mention that the detection of down-regulated protein of tropomyosin beta chain protein was reported in esophageal squamous cell carcinomas tissue[42]. The functions of the tropomyosins include the regulation of actin–myosin interaction leading to muscle contraction and the regulation of cytoskeletal functions such as cell adhesion and motility. Li *et al.*[43], reported that in breast cancer, there is a correlation between increased expression of tropomyosins 4 and lymph node metastasis.

In consistent with previous studies that have reported the up-regulation of Ig gamma-2 chain[44] C region in cervical dysplasia and cervical carcinoma *in Situ*, this study reported similar findings. This protein may receive modulated relevant to precancerous and cancerous conditions. In addition, this protein may be beneficial for studies examining cervical cancer progression, treatment efficacy and the tailoring of individualized patient care[44]. Panis *et al.*[45] who studied breast cancer reported the down-regulation of a total of 80 proteins after chemotherapy, including Ig gamma-2 chain C region. The findings of Panis *et al.*[45] demonstrated that doxorubicin down regulated acute phase proteins after breast cancer chemotherapy a finding that may provide useful insights for future developments of markers for breast cancer treatment response. Caragata *et al.*[46] pointed out that Ig gamma-2 chain C region may be useful in screening for potential glycoprotein biomarkers in human saliva to detect glycosylation modifications that are specific for particular diseases.

In this study, gelsolin protein was found to be one of the up regulated glycoprotein. It is an actin-binding protein that is a core regulator of actin filament assembly and disassembly and it inhibits apoptosis by stabilizing the mitochondria[47]. The role of the gelsolin in metastasis is a disputed issue because previous studies have revealed connections between it and

both worse and better prognosis. High gelsolin levels are related to better prognosis in ER^+HER2^- breast cancer and a minimization in tumour cell migration[48].

CONCLUSION

The glycoprotein profiling using the high sensitivity LTQ-Orbitrap Fusion Mass Spectrometer has allowed detecting the glycoprotein biomarker in breast cancer tissue. This study identified 11 upregulated glycoproteins in tumour tissues of breast cancer but not in normal ones which suggested that the alterations in proteins expressions in breast cancer could provide important clues to develop novel biomarkers as well as strategy for efficient prevention and therapy for breast cancer. The findings of this study provided useful insights into the glycoproteins identification and could establish fundamental insights for further differential glycoproteomics profiling for breast cancer biomarker discovery and it could be further identified and validated for early breast cancer diagnosis.

SIGNIFICANCE STATEMENTS

- Glycosylation is considered to be one of the most important and wide forms of protein post-translational modification which is playing an important role in several physiological functions and biological activities
- Among tumour tissues, 11 glycoproteins were detected which are not found in the normal tissues

ACKNOWLEDGMENT

We would like to thank the Ministry of Heath, Malaysia for providing the human ethical clearance for conducting this research. Special thanks to Doping Control Centre Malaysia and Universiti Sains Malaysia for providing all support throughout the study. This work was partly funded by USM Short Term Grant (304/PDOPING/6312022). We would like to thank the Anti-Doping Lab Qatar management for providing us all kinds of technical and scientifically support, facilities, instrumentation and chemicals. Finally, we would like to thank the management of the SJH (Penang, Malaysia) and the KBH (Penang, Malaysia) for providing the samples.

REFERENCES

1. Draisma, G., R. Etzioni, A. Tsodikov, A. Mariotto and E. Wever *et al*., 2009. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J. Natl. Cancer Inst., 101: 374-383.
2. Calligaris, D., C. Villard and D. Lafitte, 2011. Advances in top-down proteomics for disease biomarker discovery. J. Proteomics, 74: 920-934.
3. Ross, J.S., C. Hatzis, W.F. Symmans, L. Pusztai and G.N. Hortobagyi, 2008. Commercialized multigene predictors of clinical outcome for breast cancer. Oncologist, 13: 477-493.
4. Solassol, J., N. Boulle, T. Maudelonde and A. Mange, 2005. Clinical proteomics: Towards early detection of cancers. Med. Sci. (Paris), 21: 722-729, (In French).
5. Wallin, E. and G. von Heijne, 1998. Genome wide analysis of integral membrane proteins from eubacterial, archaean and eukaryotic organisms. Prot. Sci., 7: 1029-1038.
6. Lundqvist, A., E. Andersson, I. Ahlberg, M. Nilbert and U. Gerdtham, 2016. Socioeconomic inequalities in breast cancer incidence and mortality in Europe-a systematic review and meta-analysis. Eur. J. Public Health, 26: 804-813.
7. Jemal, A., F. Bray, M.M. Center, J. Ferlay, E. Ward and D. Forman, 2011. Global cancer statistics. CA: Cancer J. Clin., 61: 69-90.
8. Tao, Z., A. Shi, C. Lu, T. Song, Z. Zhang and J. Zhao, 2015. Breast cancer: Epidemiology and etiology. Cell Biochem. Biophys., 72: 333-338.
9. Lam, S.W., C.R. Jimenez and E. Boven, 2014. Breast cancer classification by proteomic technologies: Current state of knowledge. Cancer Treat. Rev., 40: 129-138.
10. Yip, C.H., P.N. Bhoo and S.H. Teo, 2014. A review of breast cancer research in Malaysia. Med. J. Malaysia, 69: 8-22.
11. Pagel, O., S. Loroch, A. Sickmann and R.P. Zahedi, 2015. Current strategies and findings in clinically relevant post-translational modification-specific proteomics. Expert Rev. Proteomics, 12: 235-253.
12. Hardman, M. and A.A. Makarov, 2003. Interfacing the orbitrap mass analyzer to an electrospray ion source. Anal. Chem., 75: 1699-1705.
13. Macek, B., L.F. Waanders, J.V. Olsen and M. Mann, 2006. Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. Mol. Cell. Proteomics, 5: 949-958.
14. Zhang, Z. and B. Shah, 2007. Characterization of variable regions of monoclonal antibodies by top-down mass spectrometry. Anal. Chem., 79: 5723-5729.
15. Bondarenko, P.V., T.P. Second, V. Zabrouskov, A.A. Makarov and Z. Zhang, 2009. Mass measurement and top-down HPLC/MS analysis of intact monoclonal antibodies on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. J. Am. Soc. Mass Spectrom., 20: 1415-1424.
16. Roth, Z., G. Yehezkel and I. Khalaila, 2012. Identification and quantification of protein glycosylation. Int. J. Carbohydrate Chem., Vol. 2012. 10.1155/2012/640923.
17. Drake, R.R. and L. Ball, 2015. Glycosylation and Cancer. Vol. 126, Academic Press, London, UK., ISBN: 9780128016145, Pages: 418.

18. Clark, D.J., Y. Mei, S. Sun, H. Zhang, A.J. Yang and L. Mao, 2015. Glycoproteomic approach identifies KRAS as a positive regulator of CREG1 in non-small cell lung cancer cells. Theranostics, 6: 65-77.
19. Pinho, S.S. and C.A. Reis, 2015. Glycosylation in cancer: Mechanisms and clinical implications. Nat. Rev. Cancer, 15: 540-555.
20. Taniguchi, N. and Y. Kizuka, 2015. Chapter two-glycans and cancer: Role of N-glycans in cancer biomarker, progression and metastasis and therapeutics. Adv. Cancer Res., 126: 11-51.
21. Tran, J.C. and A.A. Doucette, 2008. Gel-eluted liquid fraction entrapment electrophoresis: An electrophoretic method for broad molecular weight range proteome separation. Anal. Chem., 80: 1568-1573.
22. Ismail, M.N., E.L. Stone, M. Panico, S.H. Lee, Y. Luu, K. Ramirez and A. Dell, 2011. High-sensitivity O-glycomic analysis of mice deficient in core 2 β1, 6-N-acetylglucosaminyltransferase. Glycobiology, 21: 82-98.
23. Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem., 72: 248-254.
24. Witkowski, C. and J. Harkins, 2009. Using the GELFREE 8100 fractionation system for molecular weight-based fractionation with liquid phase recovery. J. Visualized Exp., Vol. 34. 10.3791/1842.
25. Ru, Q.C., L.A. Zhu, R.A. Katenhusen, J. Silberman, H. Brzeski, M. Liebman and C.D. Shriver, 2006. Exploring human plasma proteome strategies: High efficiency in-solution digestion protocol for multi-dimensional protein identification technology. J. Chromatogr. A, 1111: 175-191.
26. North, S.J., J. Jang-Lee, R. Harrison, K. Canis, M.N. Ismail, A. Trollope and A.T. Etienne, 2010. Chapter two-mass spectrometric analysis of mutant mice. Meth. Enzymol., 478: 27-77.
27. Saba, J., S. Dutta, E. Hemenway and R. Viner, 2012. Increasing the productivity of glycopeptides analysis by using higher-energy collision dissociation-accurate mass-product-dependent electron transfer dissociation. Int. J. Proteomics, Vol. 2012. 10.1155/2012/560391.
28. Tian, H.M., X.H. Liu, W. Han, L.L. Zhao, B. Yuan and C.J. Yuan, 2013. Differential expression of filamin A and its clinical significance in breast cancer. Oncol. Lett., 6: 681-686.
29. Lee, H.H., C.A. Lim, Y.T. Cheong, M. Singh and L.H. Gam, 2012. Comparison of protein expression profiles of different stages of lymph nodes metastasis in breast cancer. Int. J. Biol. Sci., 8: 353-362.
30. Helfman, D.M., P. Flynn, P. Khan and A. Saeed, 2008. Tropomyosin as a Regulator of Cancer Cell Transformation. In: Tropomyosin, Gunning, P. (Ed.). Springer, New York, ISBN: 9780387857664, pp: 124-131.
31. Rajan, S., G. Jagatheesan, C.N. Karam, M.L. Alves and I. Bodi *et al*., 2010. Molecular and functional characterization of a novel cardiac-specific human tropomyosin isoform. Circulation, 121: 410-418.
32. Chang, Y.K., Y.H. Lai, Y. Chu, M.C. Lee, C.Y. Huang and S. Wu, 2016. Haptoglobin is a serological biomarker for adenocarcinoma lung cancer by using the Proteome Lab PF2D combined with mass spectrometry. Am. J. Cancer Res., 6: 1828-1836.
33. Jung, K., W. Cho and F.E. Regnier, 2008. Glycoproteomics of plasma based on narrow selectivity lectin affinity chromatography. J. Proteome Res., 8: 643-650.
34. Bedolla, R.G., Y. Wang, A. Asuncion, K. Chamie and S. Siddiqui *et al*., 2009. Nuclear versus cytoplasmic localization of filamin A in prostate cancer: Immunohistochemical correlation with metastases. Clin. Cancer Res., 15: 788-796.
35. Zhao, P., W. Ma, Z. Hu, L. Zang, Z. Tian and K. Zhang, 2016. Filamin A (FLNA) modulates chemosensitivity to docetaxel in triple-negative breast cancer through the MAPK/ERK pathway. Tumor Biol., 37: 5107-5115.
36. Polterauer, S., C. Grimm, V. Seebacher, N. Concin and C. Marth *et al*., 2009. Plasma fibrinogen levels and prognosis in patients with ovarian cancer: A multicenter study. Oncologist, 14: 979-985.
37. Kwon, M., E. Hanna, D. Lorang, M. He and J.S. Quick *et al*., 2008. Functional characterization of filamin a interacting protein 1-like, a novel candidate for antivascular cancer therapy. Cancer Res., 68: 7332-7341.
38. Dowling, P., V. Palmerini, M. Henry, P. Meleady and V. Lynch *et al*., 2014. Transferrin-bound proteins as potential biomarkers for advanced breast cancer patients. BBA Clin., 2: 24-30.
39. Bernardi, P. and P. Bonaldo, 2008. Dysfunction of mitochondria and sarcoplasmic reticulum in the pathogenesis of collagen VI muscular dystrophies. Ann. N. Y. Acad. Sci., 1147: 303-311.
40. Karousou, E., M.L. D'Angelo, K. Kouvidi, D. Vigetti and M. Viola *et al*., 2014. Collagen VI and hyaluronan: The common role in breast cancer. BioMed. Res. Int., Vol. 2014. 10.1155/2014/606458.
41. Danda, R., K. Ganapathy, G. Sathe, A.K. Madugundu and S. Ramachandran *et al*., 2016. Proteomic profiling of retinoblastoma by high resolution mass spectrometry. Clin. Proteomics, Vol. 13. 10.1186/s12014-016-9128-7.
42. Harada, T., Y. Kuramitsu, A. Makino, M. Fujimoto and N. Iizuka *et al*., 2007. Expression of tropomyosin alpha 4 chain is increased in esophageal squamous cell carcinoma as evidenced by proteomic profiling by two dimensional electrophoresis and liquid chromatography mass spectrometry/mass spectrometry. Proteomics Clin. Applic., 1: 215-223.

43. Li, D.Q., L. Wang, F. Fei, Y.F. Hou and J.M. Luo *et al*., 2006. Identification of breast cancer metastasis associated proteins in an isogenic tumor metastasis model using two dimensional gel electrophoresis and liquid chromatography ion trap mass spectrometry. Proteomics, 6: 3352-3368.
44. Streckfus, C., 2016. Altered salivary protein profiles among individuals diagnosed with cervical dysplasia and cervical carcinoma in situ. Cancer Clin. Oncol., 5: 63-72.
45. Panis, C., L. Pizzatti, A.C. Bufalo, A.C. Herrera, V.J. Victorino, R. Cecchini and E. Abdelhay, 2016. Early downregulation of acute phase proteins after doxorubicin exposition in patients with breast cancer. Tumor Biol., 37: 3775-3783.
46. Caragata, M., A.K. Shah, B.L. Schulz, M.M. Hill and C. Punyadeera, 2016. Enrichment and identification of glycoproteins in human saliva using lectin magnetic bead arrays. Anal. Biochem., 497: 76-82.
47. Koya, R.C., H. Fujita, S. Shimizu, M. Ohtsu, M. Takimoto, Y. Tsujimoto and N. Kuzumaki, 2000. Gelsolin inhibits apoptosis by blocking mitochondrial membrane potential loss and cytochrome C release. J. Biol. Chem., 275: 15343-15349.
48. Stock, A.M., F. Klee, K. Edlund, M. Grinberg and S. Hammad *et al*., 2015. Gelsolin is associated with longer metastasis-free survival and reduced cell migration in estrogen receptor-positive breast cancer. Anticancer Res., 35: 5277-5285.