RESEARCH ARTICLE

Plasma Peptidome Pattern of Breast Cancer Using Magnetic Beads-Based Plasma Fractionation and MALDI-TOF MS: A Case Control Study in Egypt

Ahmad Zaki[1], Ragaa Abdelkader Ramadan[1]*, Pacint Moez[2], Hala Ghareeb[1], Ahmad Elkarmouty[3]

Abstract

Objective: The present study aimed to determine peptidome patterns in breast cancer (BC). **Methods:** We analyzed the plasma proteomic profiling of 80 BC patients and 50 healthy controls, using hydrophobic interaction chromatography magnetic beads (MB-HIC8) separation followed by Matrix assisted laser desorption ionization/ time of flight mass spectrometry (MALDI-TOF MS). **Results:** ClinProTools software identified 92 peaks that differed among the analyzed groups, 33 peaks were significantly different (P < 0.05). Of those, 22 peaks were up-regulated while 11 peaks were down-regulated in BC patients compared with the healthy controls. Three peptide ion signatures (*m/z* 1,570.31, 1,897.4 and 2,568.17) were provided by the Quick Classifier model to discriminate BC patients from healthy control subjects with 96.4% accuracy. External validation was performed by an independent group and this achieved a sensitivity of 100% and a specificity of 76.9%. **Conclusion:** MALDI-TOF MS has good analytical performance in distinguishing BC patients from healthy controls.

Keywords: Matrix assisted laser desorption ionization/ time of flight mass spectrometry- magnetic beads- ClinProTools

Introduction

Breast cancer (BC) is considered a leading cause of cancer death in women worldwide (Tan et al., 2018). In Egypt, BC is the most common cancer in women accounting for approximately 32.04% of the reported malignancies among Egyptian women (Ibrahim et al., 2014). Currently available imaging techniques such as mammography, ultrasonography, magnetic resonance imaging and digital breast tomosynthesis have improved diagnosis and reduced BC related mortality. However there are limitations especially in small tumors, abscent calcifications and high breast density as well as age related false results (Whelan et al., 2012; Moseley, 2016).

On the other hand, circulating tumor markers like cancer antigen (CA15-3) and carcinoembryonic antigen (CEA) lack sensitivity and specificity in diagnosis, which have limited their use for screening and confined them for prognosis and follow up (Patani et al., 2013; Tan et al., 2018). This necessitates the identification of biomarkers not only to improve diagnosis and prognosis, but also to support individualized therapy and clinical decision making.

Cancer biomarkers have been studied in different specimens, but indeed blood is the most convenient of all because of its noninvasive attainable nature. In addition it usually contains biomarkers which are secreted not only by the tumor but by the surrounding stroma as well (McCuaig et al., 2017). Several analytical tools have been employed in cancer biomarker research. Mass spectrometry outperformed other techniques in its ability to detect hundreds of proteins in a sole experiment. These proteins are either up or down regulated when comparing the proteomic profiling between malignant and non malignant biological fluids (Mazur and Pyatchanina, 2016; Yang et al., 2016).

Proteome profiling reveals the features of the whole proteome through defined *m/z* values in the mass spectrum but does not report protein identities for the discriminatory ion peaks. Numerous mass spectrometry techniques are in use for proteome profiling especially Matrix-Assisted Laser Desorption-Ionization spectrometry (MALDI) coupled to a Time-of-Flight (ToF) analyzer. This enables in addition to the high throughput and low sample consumption, high sensitivity and accuracy of elucidation of *m/z* values of ions in reported mass spectra of complex protein mixtures in different biological specimens (Zhang et al., 2016).

[1]Department of Chemical Pathology, [3]Department of Experimental and Clinical Surgery, Medical Research Institute, [2]Department of Clinical Pathology, Faculty of Medicine, Alexandria University, Egypt. *For Correspondence: ragaa.abdelkader@gmail.com

In the current study, we used hydrophobic interaction chromatography magnetic beads (MB-HIC8) followed by MALDI-TOF MS for plasma proteomic profiling analysis in patients with BC and healthy controls. By comparing the generated proteomic profiles by ClinPro Tools 3.0, differentially expressed peaks could be identified as potential biomarkers.

Materials and Methods

Eighty BC patients were consecutively enrolled from 2015 to 2016 from Medical Research Institute Teaching Hospital, Alexandria University, Egypt. All patients were newly diagnosed, histopathologically confirmed by ultrasound guided core needle biopsy, and untreated. The control group included fifty apparently healthy female volunteers receiving routine mammography at the breast diagnostic center. Exclusion criteria for controls included abnormal mammography or physical breast examination or prior personal history of any cancer. All subjects gave informed consent, approved by local ethics committee. Demographic data were obtained, via interviews and standardized questionnaire.

Blood samples and biopsy

Fasting venous blood sample was collected , from subjects in a seated position, into a purple capped vacutainer (BD diagnostics, Plymouth, UK) containing 50 μL of 3.8% di-potassium ethylene diamine tetra acetic acid (K2-EDTA) , then it was inverted 10 times then centrifuged under refrigerated conditions (5°C) at 1,800 x g for 15 minutes. The plasma samples were distributed into 200 μl aliquots and stored at -80°C till use. Biopsy material was used to assess for histopathological tumor grade.

Reagents

The hydrophobic interaction chromatography magnetic beads (MB-HIC8), peptide calibration standard and α-cyano-hydroxycinnamic (HCCA) were purchased from Bruker Daltonics. We used acetonitrile (ACN) HPLC grade (Scharlau, Barcelona, Spain) and trifluoroacetic acid (TFA) HPLC grade (Loba Chemie, Mumbai, India). Elution buffer was freshly prepared daily by mixing 50% of ACN with MB-HIC stabilization buffer. HCCA matrix was prepared by dissolving 1.2 mg HCCA in 200 μl ACN, 40 μl 10% TFA and 160 μl distilled water, this yielded 3mg/ml HCCA in 50% ACN: 2% TFA.

Sample purification

The samples were processed according to manufacturer instructions. Briefly, the plasma and MB-HIC8 Kit were left to reach room temperature, 8 μl of MB-HIC binding buffer and 4 μL of each plasma sample were combined in a microfuge tube. Then, 4 μL of HIC8-beads was added to the tube (after thoroughly shaking) and mixed by pipetting up and down 5 times. The sample tube was allowed to stand for one minute and then placed in a magnetic bead separator for 20 seconds, where the beads were pulled to the side by magnetic force, allowing for the supernatant to be removed and discarded carefully with a pipette.

Subsequently 90 μL of the washing buffer were added to the tube, the tube was moved 20 times back and forth in two adjacent holes of the magnetic separator. 20 seconds later the beads were collected on the wall of the tubes in the magnetic separator, and the supernatant was removed carefully, using a pipette and the beads remained in place. The washing process was repeated twice. Following binding and washing, 9 μL of the elution buffer was added to disperse beads in tubes by pipetting up and down. One minute later the tube was placed in a magnetic bead separator for 30 seconds where the beads were pulled to the side and the clear elute was transferred to a fresh tube. Finally, 1 μL of the resulting elute was spotted on the MALDI-TOF MS polished steel target (MTP 384 polished steel target plate). After air drying, 1 μL of HCCA matrix was applied onto each spot, and the target was air dried again (co-crystallization). The ClinPro Standard (CPS) was applied for calibrating the machine.

Mass Spectrometry Analysis

The proteomic profiling analysis was performed using an ultrafleXtreme MALDI-TOF/TOF MS instrument (Bruker Daltonics, Germany), operating in a linear mode with the following setting: ion source 1, 25.00 kV; ion source 2, 23.65 kV; lens, 6.8 kV; pulsed ion extraction, 300 ns; Ionization was achieved by irradiation with a laser operating at 2000.0Hz. Matrix suppression effect was enabled with signal suppression up to 800 Da. Mass spectrum were detected using linear positive mode. A standard calibration mixture in the range of 4000 to 20000 Da was used for mass calibration.

Four MALDI preparations (MALDI spots) for each sample were measured. For each MALDI spot, 3,000 spectrum were acquired (500 laser shots at 6 different spot positions). The ClinPro Tools software version 3.0 was used to compile spectra and detect peaks. The *m/z* ratios between 900 and 20,000 with a signal to noise threshold of 8.0 were selected as the target mass range for analysis because this range contained the resolved protein and peptides with smaller molecular weight. The study was conducted in the Proteomic lab in Faculty of Medicine, Alexandria University which is funded by STDF capacity building project 2,897.

Statistical analysis

Statistical analyses were perfromed by SPSS.18. Descriptive measures were done for all variables and a p-value less than 0.05 was considered statistically significant. The Student's t-test, chi-squared (x^2) test, and Fisher exact test were used to assess the general characteristics between groups. Receiver operator characteristic (ROC) curve was generated and Youden's index was calculated. Sample size was estimated to promote the detection of a difference of at least 25% in a measured variable among the studied groups at the 5% significance level ($\alpha = 0.05$) and statistical power of at least 0.8, assuming group coefficients of variation of 50% ($\sigma = 0.5$).

The Bruker Compass ClinProTools 3.0 application (referred to as 'ClinProTools') was used for data analysis. ClinProTools combines visualization features and multiple

mathematical algorithms to generate pattern recognition models for classification and prediction of disease from MS based profiling data. Data analysis began with raw data pre-treatment, including baseline subtraction on spectra, normalization and recalibration of spectra, followed by internal peak alignment using prominent peaks, and a peak picking procedure. Statistically significant different quantity of peptides was determined by means of Wilcoxon test.

Results

The mean age of BC patients and control subjects was 53.44±10.88 and 40.90±13.62 respectively. The infiltrating ductal carcinoma was the predominant histopathological type in BC patients (91.3%) followed by lobular carcinoma (5%), while mixed ductal and lobular carcinoma (3.8%) (Table 1). The data set was randomly split into model generation group (including 45 BC patients and 30 control subjects) and external validation group (including 35 BC patients and 20 control subjects). The model generation group was used for identification of the differentially expressed peptides between BC and controls, while the external validation group was used for independent validation of the peptide signatures. There was no significance difference between both groups regarding age as well as the pathological tumor criteria (Table 2).

A pilot study was performed on 10 BC patients and 10 control subjects to determine the optimal type of MB for sample purification. Two profiling kits were used; MB with hydrophobic interaction chromatography (MB-HIC8) and weak cation exchange chromatography (MB-WCX). MB-HIC8 managed to capture larger number of peptide peaks compared with MB-WCX and hence were utilized for the proteome fractionation. MB-HIC8 also showed better results regarding recognition capability and cross validation.

Model generation groups were constructed to differentiate BC from controls. A total of 92 distinct peaks were identified and 33 of them were significantly expressed. Of those, 22 peaks were up-regulated while 11 peaks were down-regulated in BC patients compared with the control subjects (Table 3). The stack and simulated two dimensional gel electrophoresis views of samples are depicted in Figure 1 and 2.

The ClinProTools supports four kinds of algorithms for model generating classification; Genetic Algorithm, Supervised Neural Network, Quick Classifier (QC) and Support Vector Machine. In our study, the QC model achieved the best results with 100% recognition capability and 96.4% cross validation accuracy. The cross validation refers to the accuracy of the algorithm to correctly assign a random sample to the correct group. In the QC algorithm the peak areas are sorted per peak and a weighted average over all peaks is calculated. Three peptide ion signatures with m/z 1,570.31 (start mass: 1,566.09, end mass: 1,574.99), 1,897.4 (start mass: 1,892.07, end mass: 1,909.46) and 2,568.17 (start mass: 2,560.23, end mass: 2,577.52) were obtained as a proteomic profile for a cross validation set to discriminate the BC patients from

Table 1. Demographic and Pathological Criteria of the Studied Participants

Variable	Breast cancer (n=80)	Control (n=50)	p value
Age (years)			
Mean ± SD	53.4±10.9	49.9±13.6	0.124
Menopausal state, n (%)			
Premenopausal	27 (33.8)	22 (44)	0.241
Postmenopausal	53 (66.3)	28 (56)	
Parity, n (%)			
Nulliparous	5 (6.3)	7 (14)	0.331
1-2	31 (38.8)	18 (36)	
≥3	44 (55)	25 (50)	
Family history of cancer, n (%)			
Yes	6 (7.5)	2 (4)	0.71
No	74 (92.5)	48 (96)	
Histopathological type			
Ductal	73 (91.3)		
Lobular	4 (5)		
Others	3 (3.8)		
Tumor stage			
I	0		
II	0		
III	73 (91.3)		
IV	7 (8.8)		
Tumor size			
≤ 2 cm	5		
> 2 cm	75		
LN			
Yes	70		
No	10		
Metastasis			
Yes	6		
No	74		
ER			
Positive	49		
Negative	14		
Missing	5		
PR			
Positive	47		
Negative	16		
Missing	5		
HER2			
Overexpression	14		
Non overexpression	49		
Missing	5		

LN, axillary lymph node involvement; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; *, Statistically significant at p < 0.05

control subjects (Figure 3 and 4). The receiver operating characteristic curve (ROC) analysis was performed for each peak. Based on quantitative data of peak intensities
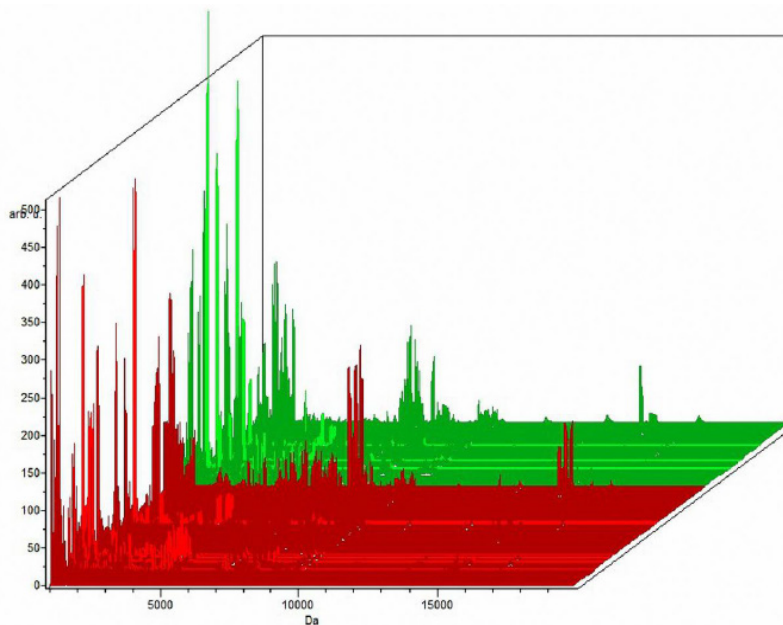
Figure 1. Stack View of the Plasma Protein Profile Mass Spectra of the Loaded Classes in a Three Dimensional Space (red represents 45 breast cancer patients and green represents 30 control subjects).

the area under curve (AUCs) of peaks m/z; 1,570.31, 1,897.4 and 2,568.17 were 0.992, 0.984 and 1 respectively with 95% confidence interval (Figure 5).

To verify the accuracy of the established QC classification model, we performed external validation study which consisted of 35 BC patients and 20 control subjects (not used in the model generation group). The QC model had higher external validation values than both SNN and GA models. The QC model correctly classified 100% of the breast cancer (sensitivity) and 76.9% of the control (specificity) samples.

Two samples were used to determine the within- and between-run precision. In each profile, three peaks with different molecular masses were selected to evaluate assay precision. Within-run imprecision was determined by evaluating the coefficients of variance (CV) for each sample, using 8 assays within a run, then between-run imprecision was established by carrying out 8 different assays for a sample over a period of 7 days. The peak CVs were all <4% and <10% in the within- and between-run assays, respectively (figure 6). These values were consistent with the reproducibility data for the Protein
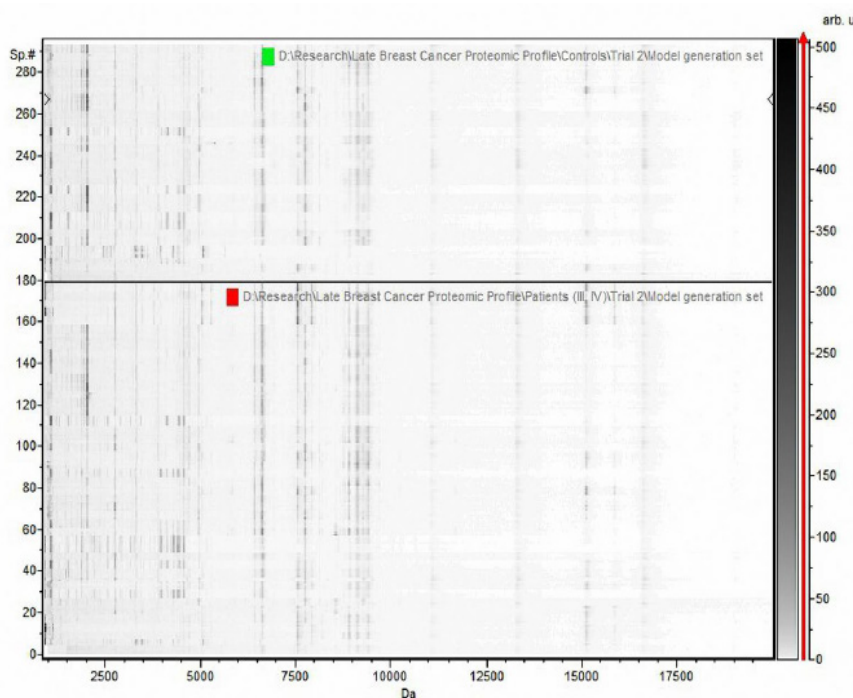

Figure 2. The Simulated Two-Dimensional Gel Electrophoresis Views of Plasma Protein Profile Mass Spectra of All Samples (the upload, benign group; the download, malignant group).
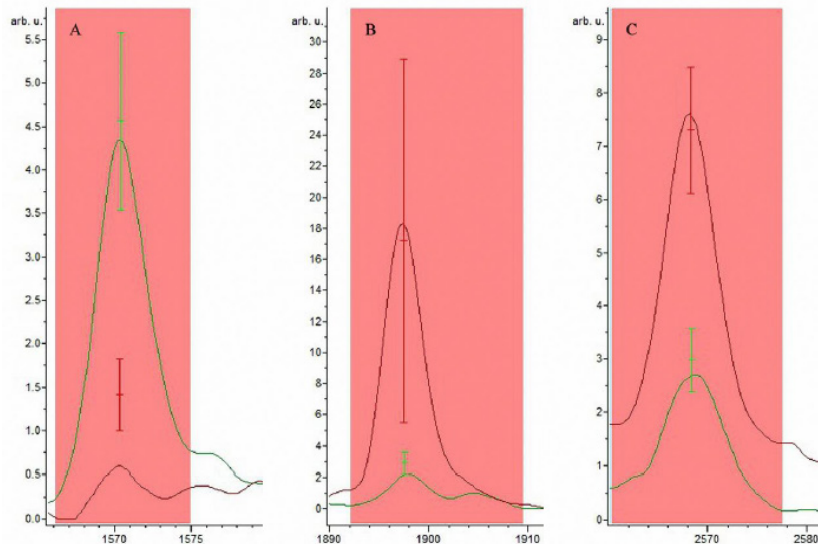
Figure 3. Zoom of the Mass Range of the Three Peaks (MALDI-TOF MS Linear Mode) Used in the QC Model to Differentiate Breast Cancer Patients (Red) from the Control Subjects (Green). The red line demarcates the peak as an integration region. A, peak at m/z 1570.31 (start mass 1566.09 and end mass 1574.99); B, peak at m/z 1897.4 (start mass 1892.07 and end mass 1909.46); C, peak at m/z 2568.17 (start mass 2560.23 and end mass 2577.52). The peak is up regulated in BC patients than in control subjects.

Biology System reported by Bruker Daltonics.

Discussion

Owing to the heterogenous nature of BC as well as the different molecular subtypes, no single molecular feature per say could be a decisive diagnostic tool. Instead, multi-component molecular classifiers of either genes or proteins are suggested (Tayyari et al., 2018). Although a number of genetic panels are currently in practice for diagnostic and prognostic applications, medical management guidelines do not explicitly advocate for or against the use these testing methods (NCCN guidelines, 2017; The American Society of Breast Surgeons Consensus Guideline, 2017). Meanwhile BC proteomic panels are

still under research and are not commercialized for clinical use. Despite, peptide profiling has been approved for routine use in clinical microbiological laboratories for identification of microorganisms (Patel, 2015).

In this work, a case control comparative analysis between BC and healthy controls was performed. Plasma proteomic profiles were identified with MB-HIC8 fractionation followed by MALDITOF MS. A total 33 peak were significantly expressed. The QC model provided three peptide ion signatures (*m/z* 1570.31, 1,897.4 and 2,568.17) as a proteomic profile to discriminate the BC patients from control subjects with recognition capability and cross validation accuracy of almost 100%. These peaks are worthy of further sequence determination and functional analysis. Blinded verification of the QC
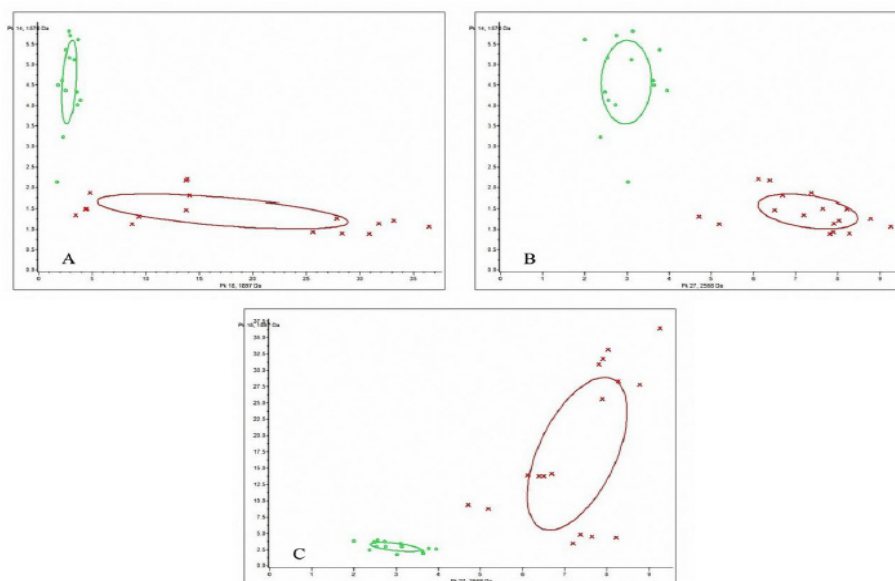

Figure 4. Two-Dimensional (2D) Peak Distribution View of Peptides with A; m/z 1570 (x-axis) and 1897 (y-axis), B; m/z 1570 (x-axis) and 2568 (y-axis), C; m/z 1897 (x-axis) and 2568 (y-axis)

Table 2. Characteristics of Breast Cancer Patients in Model Generation and External Validation Groups

Variable	Model Generation (n = 45)		External Validation (n = 35)		p
	n (%)		n(%)		
Age (years)					
Mean ± SD	53.78 ± 12.15		53.00 ± 9.17		0.753
≤ 50	19	42.2	12	34.3	0.47
> 50	26	57.8	23	65.7	
Tumor type					
Ductal Carcinoma	42	93.3	31	88.6	0.635
Lobular Carcinoma	2	4.4	2	5.7	
Mixed	1	2.2	2	5.7	
Tumor grade					
G II	37	82.2	22	62.9	0.721
G III	6	13.3	4	11.4	
G X	2	4.4	9	25.7	
Tumor stage					
III	43	95.6	30	85.7	0.782
IV	2	4.4	5	14.3	
Tumor size (cm)					
Mean ± SD	4.12 ± 1.19		4.13 ± 1.40		0.975
ER status					
Positive	31	68.9	18	51.4	0.944
Negative	9	20	5	14.3	
Missing	5	11.1	12	34.3	
PR status					
Positive	29	64.4	18	51.4	0.613
Negative	11	24.5	5	14.3	
Missing	5	11.1	12	34.3	
HER2 expression status					
Positive	11	24.5	3	8.6	0.184
Negative	29	64.5	20	57.1	
Missing	5	11.1	12	34.3	

ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.

classification model proved to correctly classify 100% of BC cases (sensitivity) and 76.9% of control subjects (specificity). These findings highlight the possibility of the use of our classification peptidome model as a sensitive and specific diagnostic tool.

During tumorigenesis proteins and peptides may be abnormally secreted, over/under-expressed, modified or degraded. These cancer-specific low-molecular-weight proteins are an indirect snapshot of the enzyme activity in tumor cells (Hajduk et al., 2016). They mostly result from specifically released proteases that process the acute phase proteins generated by the host response to the tumor (Qin and Ling, 2012).

Different proteomic technologies have identified unique proteome patterns that can diagnose BC in different clinical stages and recognize different outcome and response to therapy (Mazur and Pyatchanina, 2016). Furthermore, reproducibility of MS based protein profiles for diagnosis of BC across clinical studies has been demonstrated in systematic review (Callesen et al., 2008).

By searching literature for MS-based studies using MALDI- TOF MS, we retrieved two protein peaks by comparing the obtained molecular masses in our work with the molecular weights of identified proteins and their subunits. Villanueva et al., (2006) identified a peak with *m/z* 1895.99 as a fragment of complement component 4a (C4a) by using MB-HIC8. The same study identified the peak at *m/z* 2,568.17 as a fragment of apolipoprotein E (ApoE). Additionally, a similar peak with *m/z* 1,897.4 was identified as C4a by Tiss et al., (2010) using reversed phase pre-packed C18-ZipTips Van den Broek et al., (2010) have validated a quantitative assay for peptides, generated by BC specific exoproteases, by liquid chromatography coupled to tandem mass spectrometry and one of these peptides was C4a. Also ApoE has been reported to significantly increase in sera of BC patients compared to healthy control confirming its role in promoting tumor cell growth (Xu et al., 2016).

Although proteomics based research have been growing, limited studies are reported till date in Egypt
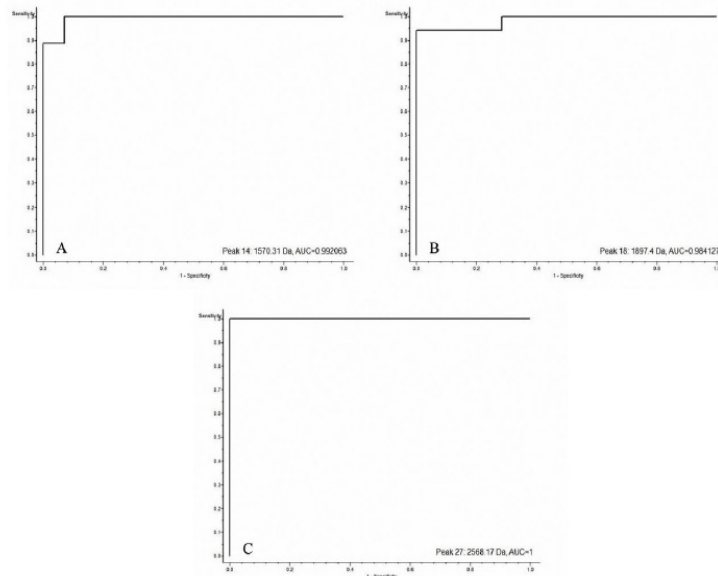
Figure 5. The Receiver Operating Characteristic (ROC) Curves of the Three Peaks Used in the QC Model to Differentiate Breast Cancer Patients from the Control Subjects.
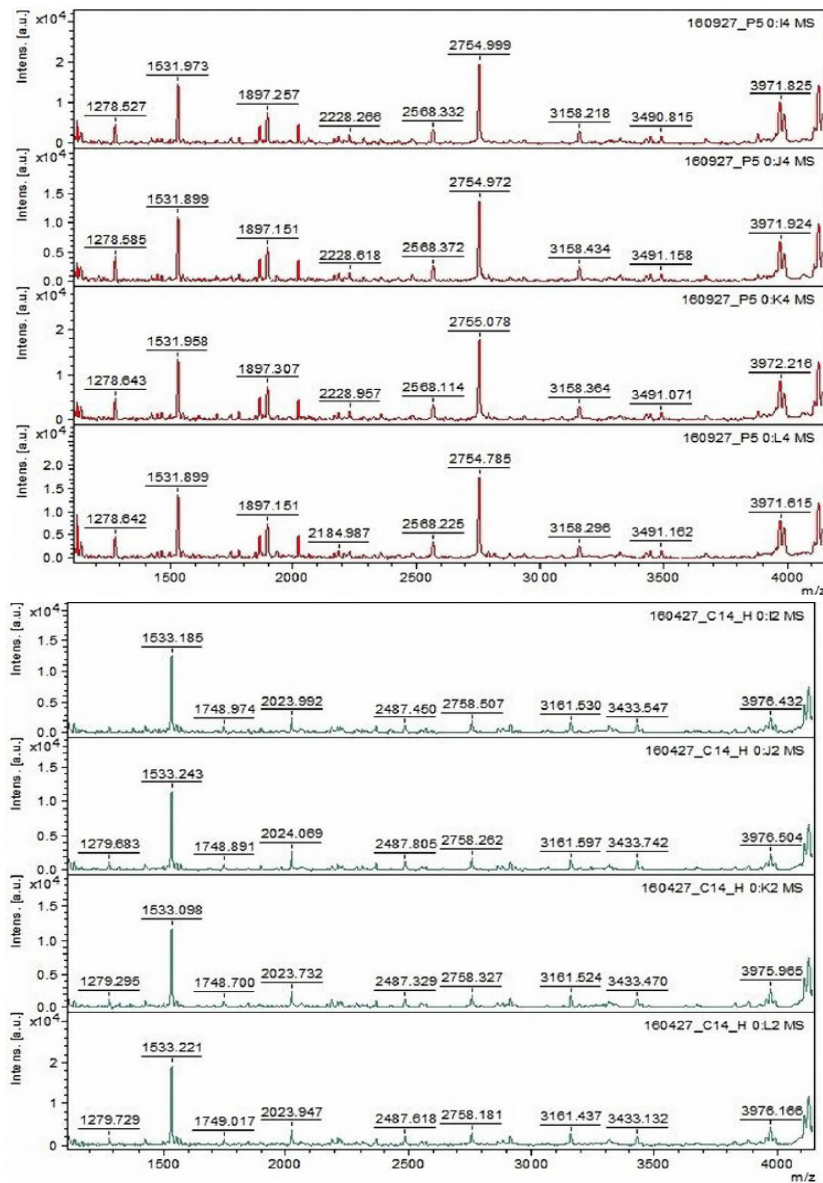

Figure 6. Reproducibility of Mass Spectra of the Plasma Protein Profile in one BC Patient (Red) and Control Subject (Green), Showing Low Variability between Replicates.

Table 3. Statistics of Differentially Expressed Peaks among Breast Cancer Patients and Control Subjects

m/z	Breast Cancer (Average ± SD)$^\Delta$	Control Group (Average ± SD)$^\Delta$	Regulation in Breast Cancer	p
1897.4*	19.12 ± 13.02	3.21 ± 0.78	Up	0.0000901
3158.86	13.31 ± 5.4	3.29 ± 1.61	Up	0.0000901
2568.17*	8.1 ± 1.31	3.32 ± 0.66	Down	0.0000901
1570.31*	1.58 ± 0.46	5.08 ± 1.14	Down	0.0000901
1532.25	15.86 ± 9.54	64.46 ± 20.65	Down	0.000251
2882	5.7 ± 2.45	2.23 ± 0.74	Up	0.00279
3883.79	3.52 ± 0.74	2.09 ± 0.8	Up	0.00279
1747.92	6.01 ± 2.44	3.22 ± 0.92	Up	0.00328
4283.43	13.86 ± 13.33	4.2 ± 9.24	Up	0.00338
2229.03	9.12 ± 6.18	2.56 ± 1.58	Up	0.0068
3446.38	3.88 ± 1.02	2.28 ± 1.05	Up	0.0068
5160.67	1.85 ± 0.57	1.19 ± 0.48	Up	0.0068
4298.48	21.95 ± 19.67	2.76 ± 3.05	Up	0.00814
3956.75	11.28 ± 9.23	4.05 ± 9	Up	0.00814
3290.25	4.32 ± 3.14	1.59 ± 1.29	Up	0.00814
2185.66	6.51 ± 3.36	3.23 ± 1.74	Up	0.0128
3430.87	7.24 ± 5.23	2.72 ± 1.59	Up	0.0137
1044.72	12.65 ± 2.94	16.75 ± 5.17	Down	0.0165
3972.4	23.24 ± 16.12	4.58 ± 2.83	Up	0.019
8602.7	0.82 ± 0.34	0.5 ± 0.23	Up	0.019
4419.15	26.32 ± 22.01	5.95 ± 2.23	Up	0.0216
1099.44	17.25 ± 8.39	27.71 ± 9.95	Down	0.0216
1278.71	7.73 ± 3.77	4.68 ± 1.24	Up	0.0216
2486.19	2.76 ± 1.73	4.86 ± 1	Down	0.0216
8567.34	0.67 ± 0.25	0.42 ± 0.2	Up	0.0216
1553.87	4.38 ± 2.63	6.85 ± 1.7	Down	0.0234
2037.85	2.08 ± 1.29	7.2 ± 8.07	Down	0.0252
2023.08	14.64 ± 16.97	62.61 ± 79.04	Down	0.034
7830.37	0.93 ± 0.45	0.58 ± 0.28	Up	0.0366
1061.42	289.6 ± 118.53	411.7 ± 130.26	Down	0.0394
7766.43*	9.16 ± 2.79	5.12 ± 3.5	Up	0.0424
4145.18	6.79 ± 4.13	18.39 ± 15.88	Down	0.0442
8618.72	0.85 ± 0.54	0.42 ± 0.18	Up	0.0442

m/z, mass-to-charge ratio; $^\Delta$, Average and standard deviation of the peak intensity, up; upregulated, down;downregulated, * differential peaks used for model generation group, p value calculated with the Wilcoxon test; p < 0.05 suggest statistical significance

and they focused on the use of tandem MS in screening for inborn errors of metabolism (Hassan et al., 2016). Our work, is the first of its type in Egypt to employ MALDI TOF in BC.

Definitely the aim of MS proteomic profiling is reproducible, rapid and inexpensive acquisition of discriminating peptide signature that can be suitable in population screening or clinical triage. Provided standardized sample pretreatment, MS measurement, data processing and analysis, peptide and protein profiles are highly reproducible. Standardization of preanalytical factors such as anticoagulants, temperature, freeze–thaw cycles and storage conditions is critical in MS studies because they have significant impact on dynamic alterations of the acquired proteome (Baumann et al., 2005; Periano et al., 2016; Tsuchida et al., 2018).

In the current work we used a standardized protocol which started by the use of plasma. This was in agreement with the Human Proteomics Organization who recommended ethylenediamine tetraacetic acid (EDTA) plasma as the preferred specimen from blood. This can be attributed to less degradation ex vivo in addition to much less variability than the protease-rich process of clotting (Rai et al., 2005). Few studies reported peak identification lists for serum (Villanueva et al., 2006; Tiss et al., 2010) or plasma (Koomen et al., 2005) and this showed the inherent difference between the two type of samples in the moderate percent of concordance in the identified peaks. Several proteomic based studies used plasma (Periyasamy et al., 2015; Baralla et al., 2018).

Magnetic beads (MB) have been developed to purify and fractionate the proteome. They consist of

magnetic core and chromatographic surface. For each sample, fresh disposable beads are used, thus avoiding carry-over that may occur with other techniques such as liquid chromatography. There are various types of MB: cation exchange, anion exchange, reverse phase and metal affinity. This different functionality allow protein and peptide enrichment based on different chemical–physical interactions, thereby broadening the range of components covered (Velstra et al., 2012).

MB-hydrophobic interaction chromatography (MB-HIC) kits are based on paramagnetic micro-particles with a highly porous surface with hydrophobic coatings (C3,C8,C18). They capture 1-10 KDa polypeptides. While MB-weak cation-excahnge chromatography (MB-WCX) depend on supramagnetic micro-particles with negatively charged functional groups at their surface enabling cation exchange chromatography (Qiu et al., 2008).

Choosing the right type of MB enables the highest acquisition of proteins and peptides for the furthur proteomic profiling analysis, therfore we compared the performance of two MB and we found that MB-HIC8 had better average peak numbers, higher peak intensitites, and better capturing ability than MB-WCX. HIC showed good performance in previous studies such as the work by (Villanueva et al., 2004; De Noo et al., 2006; Periyasamy et al., 2015) as well as good reproducability (Baumann et al., 2005).

Several compounds were employed as MS matrix such as 2,5-dihydroxyacetophenone, sinapinic acid and 2,5-dihydroxybenzoic acid (Hajduk et al., 2016). We used alpha-cyano-4-hydroxycinnamic acid (HCCA), the most common MALDI-TOF MS matrix, yielding satisfactory quality spectra in both low and high mass ranges. Once again standardization of the type of organic solvent, matrix and its concentration and pH is mandatory because of its effect on crystallization (Penno et al., 2009). HCCA was employed by numerous research groups (Calandra et al., 2016; Kim et al., 2016; Baralla et al., 2018).

In conclusion, our study showed that MB-HIC8 fractionation followed by MALDI-TOF MS combined with ClinProTools software shows high sensitivity and specificity for the identification of BC. The study was limited by the absence of sequence identification of the expressed peptides. Nevertheless, we followed standardized protocols in sample collection, pre-analytical conditions and MS analysis. Moreover biological variables were matched for both groups and data was validated by an independent group not used for model generation. Also the control group included age-matched females undergoing a mammography and showing no aberrations. Our results need to be further confirmed in larger patient cohorts, and we recommend the construction of the protein panel in an immunoassay format exemplified in a multiplexed technique to facilitate its further evaluation and validation. The next step of our study will be to correlate our findings to overall survival, recurrence free survival and quality of life in BC patients. So that, in a near future we can fulfill the role in the personalized medicine and achieve the ultimate aim, the decrease in mortality rate from BC.

Acknowledgements

References

Baralla A, Fois AG, Sotgiu E, et al (2018). Plasma proteomic signatures in early chronic obstructive pulmonary disease. *Proteomics Clin Appl*, **7**, e1700088.

Baumann S, Ceglarek U, Fiedler GM, et al (2005). Standardized approach to proteome profiling of human serum based on magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin Chem*, **51**, 973-80.

Calandra E, Posocco B, Crotti S, et al (2016). Cross-validation of a mass spectrometric-based method for the therapeutic drug monitoring of irinotecan: implementation of matrix-assisted laser desorption/ionization mass spectrometry in pharmacokinetic measurements. *Anal Bioanal Chem*, **408**, 5369-77.

Callesen AK, Vach W, Jørgensen PE, et al (2008). Reproducibility of mass spectrometry based protein profiles for diagnosis of breast cancer across clinical studies: a systematic review. *J Proteome Res*, **7**, 1395-402.

De Noo ME, Mertens BJ, Ozalp A, et al (2006). Detection of colorectal cancer using MALDI-TOF serum protein profiling. *Eur J Cancer*, **42**, 1068-76.

Hajduk J, Matysiak J, Kokot ZJ (2016). Challenges in biomarker discovery with MALDI-TOF MS. *Clinica Chimica Acta*, **458**, 84-98.

Hassan FA, El-Mougy F, Sharaf SA, et al (2016). Inborn errors of metabolism detectable by tandem mass spectrometry in Egypt: The first newborn screening pilot study. *J Med Screen*, **23**, 124-9.

Ibrahim AS, Khaled HM, Mikhail NN, Baraka H, Kamel H (2014). Cancer incidence in Egypt: results of the national population-based cancer registry program. *J Cancer Epidemiol*, **2014**, 437971.

Kim NA, Jin JH, Kim KH, et al (2016). Investigation of early and advanced stages in ovarian cancer using human plasma by differential scanning calorimetry and mass spectrometry. *Arch Pharm Res*, **39**, 668-76.

Koomen JM1, Li D, Xiao LC, et al (2005). Direct tandem mass spectrometry reveals limitations in protein profiling experiments for plasma biomarker discovery. *J Proteome Res*, **4**, 972-81.

Mazur MG, Pyatchanina TV (2016). The use of proteomic technologies in breast cancer research. *Exp Oncol*, **38**, 146-57.

McCuaig R, Wu F, Dunn J, Rao S, E.Dahlstrom J (2017). The biological and clinical significance of stromal-epithelial interactions in breast cancer. *Pathology*, **49**, 133-40.

Moseley TW (2016). Digital mammography and digital breast tomosynthesis. *Clin Obstet Gynecol*, **59**, 362-79.

NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines). Genetic/familial high-risk assessment: breast and ovarian. Version 2.2017. https://www.nccn.org/. Accessed 31 March 2017.

Patani N, Martin LA, Dowsett M (2013). Biomarkers for the clinical management of breast cancer: international perspective. *Int J Cancer*, **133**, 1-13.

Patel R (2015). MALDI-TOF MS for the diagnosis of infectious diseases. Clin Chem , 61, 100-11.

Penno MAS, Ernst M, Hoffmann P (2009). Optimal preparation methods for automated matrix-assisted laser desorption/ionization time-of-flight mass spectrometry profiling of low molecular weight proteins and peptides. *Rapid Commun Mass Spectrom*, **23**, 2656-62.

Periyasamy A, Gopisetty G, Veluswami S, Joyimallaya Subramanium M, Thangarajan R (2015). Identification of candidate biomarker mass (m/z) ranges in serous ovarian adenocarcinoma using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry profiling. *Biomarkers*, **20**, 292-8.

Preiano M, Maggisano G, Lombardo N, et al (2016). Influence of storage conditions on MALDI-TOF MS profiling of gingival crevicular fluid: Implications on the role of S100A8 and S100A9 for clinical and proteomic based diagnostic investigations. *Proteomics*, **16**, 1033-45.

Qin XJ, Ling BX (2012). Proteomic studies in breast cancer (Review). *Oncol lett*, **3**, 735-43.

Qiu F, Liu H, Zhang X, Tian Y (2008). Optimization of magnetic beads for maldi-TOF MS analysis. *Front biosci (Landmark Ed)*, **14**, 3712-23.

Rai A, Gelfand C, Haywood B, et al (2005). HUPO plasma proteome project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. *Proteomics*, **5**, 3262-77.

Tan G, Chu C, Gui X, Li J, Chen Q (2018). The prognostic value of circulating cell-free DNA in breast cancer: A meta-analysis. *Medicine (Baltimore)*, **97**, e0197.

Tayyari F, Gowda GAN, Olopade OF, et al (2018). Metabolic profiles of triple-negative and luminal A breast cancer subtypes in African-American identify key metabolic differences. *Oncotarget*, **9**, 11677-90.

The American Society of Breast Surgeons Consensus Guideline on Hereditary Genetic Testing for Patients With and Without Breast Cancer (revised March 14, 2017). https://www.breastsurgeons.org/new_layout/about/statements/PDF_Statements/ BRCA_Testing.pdf. Accessed 1 April 2017).

Tiss A, Smith C, Menon U, et al (2010). A well characterised peak identification list of MALDI MS profile peaks for human blood serum. *Proteaomics*, **10**, 3388-92.

Tsuchida S, Satoh M, Umemura H, et al (2018). Assessment by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of the effects of preanalytical variables on serum peptidome profiles following long-term sample storage. *Proteomics Clin Appl*, **12**, e1700047

Van den Broek I, Sparidans RW, Schellens JHM, Beijnen JH (2010). Quantitative assay for six potential breast cancer biomarker peptides in human serum by liquid chromatography coupled to tandem mass spectrometry. *J Chromatogr B*, **87**8, 590-602.

Velstra B, Van der burgt YE, Mertens BJ, et al (2012). Improved classification of breast cancer peptide and protein profiles by combining two serum workup procedures. *J Cancer Res Clin Oncol*, **138**, 1983-92.

Villanueva J, Philip J, Entenberg D, et al (2004). Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal Chem*, **76**, 1560-70.

Villanueva J, Shaffer DR, Philip J, et al (2006). Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J Clin Invest*, **116**, 271-84.

Whelan SA, He J, Lu M, et al (2012). Mass spectrometry (LC-MS/MS) identified proteomic biosignatures of breast cancer in proximal fluid. *J Proteome Res*, **11**, 5034-45.

Xu X, Wan J, Yuan L, et al (2016). Serum levels of apolipoprotein E correlates with disease progression and poor prognosis in breast cancer. *Tumor Biol*, **37**, 15959-66.

Yang J, Xiong X, Liu S, et al (2016). Identification of novel serum peptides biomarkers for female breast cancer patients in Western China. *Proteomics*, **16**, 925-34.

Zhang F, Xu C, Ning L, et al (2016). Exploration of serum proteomic profiling and diagnostic model that differentiate Crohn's disease and intestinal tuberculosis. *PLoS One*, **11**, e0167109.