Roshanak Aslebagh[1]
Devika Channaveerappa[1]
Kathleen F. Arcaro[2]
Costel C. Darie[1]

[1]Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, Potsdam, NY, USA
[2]Department of Veterinary & Animal Sciences, University of Massachusetts Amherst, Amherst, MA, USA

Received January 16, 2018
Revised March 11, 2018
Accepted April 16, 2018

Research Article

Comparative two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) of human milk to identify dysregulated proteins in breast cancer

Breast cancer (BC) remains a major cause of mortality, and early detection is considered important for reducing BC-associated deaths. Early detection of BC is challenging in young women, due to the limitations of mammography on the dense breast tissue of young women. We recently reported results of a pilot proteomics study, using one-dimensional polyacrylamide gel electrophoresis (1D-PAGE) and mass spectrometry (MS) to investigate differences in milk proteins from women with and without BC. Here, we applied two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and MS to compare the protein pattern in milk from the breasts of a single woman who was diagnosed with BC in one breast 24 months after donating her milk. Statistically different gel spots were picked for protein digestion followed by nanoliquid chromatography tandem MS (nanoLC-MS/MS) analysis. The upregulated proteins in BC *versus* control are alpha-amylase, gelsolin isoform a precursor, alpha-2-glycoprotein 1 zinc isoform CRA_b partial, apoptosis-inducing factor 2 and vitronectin. Several proteins were downregulated in the milk of the breast later diagnosed with cancer as compared to the milk from the healthy breast, including different isoforms of albumin, cholesterol esterase, different isoforms of lactoferrin, different proteins from the casein family and different isoforms of lysozyme. Results warrant further studies to determine the usefulness of these milk proteins for assessing risk and detecting occult disease. MS data is available via ProteomeXchange with identifier PXD009860.

Keywords:
2D-PAGE / Biomarkers / Breast cancer / Breast milk / Proteomics / Mass spectrometry
DOI 10.1002/elps.201800025

Additional supporting information may be found online in the Supporting Information section at the end of the article.

1 Introduction

Finding a noninvasive method for accurate risk assessment and early diagnosis of BC remains a major research concern since BC is one of the most common cancers both worldwide [1] and in the United States [2, 3]. Accurate risk assessment and early detection of BC (diagnosis in early stages) is important to increase survival rate [4]. Investigation of potential BC biomarkers is an emerging technology for risk assessment and early diagnosis of BC. Recently, different studies have been performed to investigate potential biomarkers of caner [5–24], including the investigations on breast milk [25–30]. As reviewed previously, a non-invasive tool for early detection of BC could be particularly beneficial for women of reproductive age, as current imaging tools are less effective in detecting cancer in the relatively dense breasts of younger women, and because the incidence of pregnancy associated BC is predicted to increase as the age of first pregnancy increase. Empirical studies support the use of human milk to investigate potential BC biomarkers [27–35]. Dysregulated proteins in breast milk could provide potential biomarkers of BC. To identify dysregulated

Correspondence: Dr. Costel C. Darie, Biochemistry & Proteomics Group, Department of Chemistry & Biomolecular Science, Clarkson University, 8 Clarkson Avenue, Potsdam 13699–5810, NY, USA
Fax: + 315-268–6610
E-mail: cdarie@clarkson.edu

Abbreviations: **2D-PAGE**, two-dimensional polyacrylamide gel electrophoresis; **BC**, breast cancer; **DCIS**, ductal carcinoma in situ; **gi**, GenInfo; **MS**, mass spectrometry; **NanoLC-MS/MS**, nano liquid chromatography-tandem mass spectrometry; **NCBI**, national center for biotechnology information; **pI**, isoelectric point; **pkl**, peak list; **PLGS**, ProteinLynx Global Server; **RT-PCR**, reverse transcription polymerase chain reaction; **UniProt**, universal protein resource

proteins in breast milk, we applied MS-based proteomics, as it is one of the most common and robust tools for protein identification, quantitation and characterization [36–39] as well as for analysis of post translational modifications of proteins [40, 41]. MS also provides the opportunity to identify protein species, including modified proteins and protein isoforms [42–44].

We recently performed a pilot study on 10 milk samples including both *within* woman (milk from a diseased breast versus a healthy breast of the same woman) and across women (milk from a woman with cancer versus a woman without cancer) comparisons, using 1D-PAGE protein separation coupled with MS-based proteomics [30]. Here, we conducted 2D-PAGE protein separation coupled with MS-based proteomics on one of the *within* woman comparisons previously studied, to examine in greater depth the protein differences of milk from a breast that developed BC *versus* milk from the contralateral breast, which did not develop BC. The protein separation in 2D-PAGE is based on the isoelectric point and molecular weight (MW) of the proteins. The dysregulated protein spots in BC versus control were excised, trypsin digested and analyzed by nanoLC-MS/MS and were identified by database search. The milk was obtained from a 35-year old mother, with family history of BC. The mother was diagnosed with ductal carcinoma in situ (DCIS) 24 months after milk donation. In depth analysis of these samples is important for two reasons. First, this within person comparison eliminates differences due to genetic background, environmental exposures, age, subtype of BC, time of milk donation, etc., leaving only protein differences that are likely to be related to the cancer. Second, since the milk sample was provided 24 months prior to the BC diagnosis, any protein differences could potentially be considered as biomarkers of risk.

2 Materials and methods

2.1 Milk samples

The milk samples were provided from an archived bank at the University of Massachusetts Amherst. University of Massachusetts Amherst Institutional Review Board approved the parent study and secondary analyses were performed at collaborating laboratories. The details of milk collection and subject enrollment are described elsewhere [27, 28]. BC and control milk samples were from the same woman. The milk sample with BC was from the breast diagnosed with DCIS 24 months after milk donation and the other sample (control) was from the other breast, which was not affected by the BC tumor, therefore was used as the control. Milk was obtained from a 35-year old mother with family history of BC. The milk was donated separately from each breast. The breast that was diagnosed later with DCIS (we refer to this sample as BC) and from the other breast (we refer to this sample as control).

2.2 Reagents

All the basic reagents and chemicals used in this study were from Sigma-Aldrich (St. Louis, MO) unless it is mentioned differently.

2.3 2D-PAGE

Intact human breast milk samples (without any centrifugation or fractionation prior to 2D-PAGE analysis) were used for this study. The concentrations of breast milk samples were measured performing Bradford assay and equal amount of proteins for each sample were loaded for both Coomassie stained (550 μg) and silver stained gels (150 μg). Carrier ampholine method of isoelectric focusing was applied by Kendrick Labs, Inc. (Madison, WI) to run 2D-PAGE which has been described elsewhere in details [45, 46]. The first part of 2D was isoelectric focusing which was performed in a 3.3 mm (inner diameter) glass tube with the pH range of 3 to 10 made by 2% Isodalt Servalytes (Serva, Heidelberg, Germany), running at 20 000 volt-hrs. Tropomyosin (protein spot with MW 33 000 and p*I* 5.2, marked with a red arrow in Fig. 1) was used as the internal standard. For Coomassie stained gels, 1 μg per sample and for silver stained gels 100 ng per sample of tropomyosin was added. The pH gradient plot for the tube gel, (not shown here) was obtained using surface pH electrode. The tube gel was equilibrated in the equilibration buffer (10% glycerol, 50 mM dithiothreitol, 2.3% SDS and 0.0625 M tris, pH 6.8) for 10 min, then sealed on top of the stacking gel which overlaid the slab gel (10% acrylamide slab gel, 1.0 mm thick). SDS-PAGE ran for approximately 5 hours (25 mA/gel). The MW standards (marked in Fig. 1**)** used for SDS-PAGE, were as follow: myosin (220 kDa), phosphorylase A (94 kDa), catalase (60 kDa), actin (43 kDa), carbonic anhydrase (29 kDa) and lysozyme (14 kDa) (Sigma-Aldrich, St. Louis, MO and EMD Millipore, Billerica, MA). The gels were put between cellophane paper sheets (lower pH to the left) and dried.

2.4 Computerized comparisons

For computerized analysis, we had two technical replicates of silver stained gels for each sample, which were scanned with laser densitometer (model PDSI, Molecular Dynamics Inc, Sunnyvale, CA). The linearity of the scanner was checked with a calibrated neutral density filter set (Melles Griot, Irvine, CA) before scanning the gels. Progenesis Same Spots software (version 4.5, 2011, TotalLab, UK) and Progenesis PG240 software (version 2006, TotalLab, UK) were used for analysis of the gel images. Briefly, the images were wrapped, and the protein spots were found, then the background was subtracted followed by matching and quantification (measuring the spot percentage). Accurate manual checking was performed all the time. Spot percentage was defined as the spot integrated density above background.

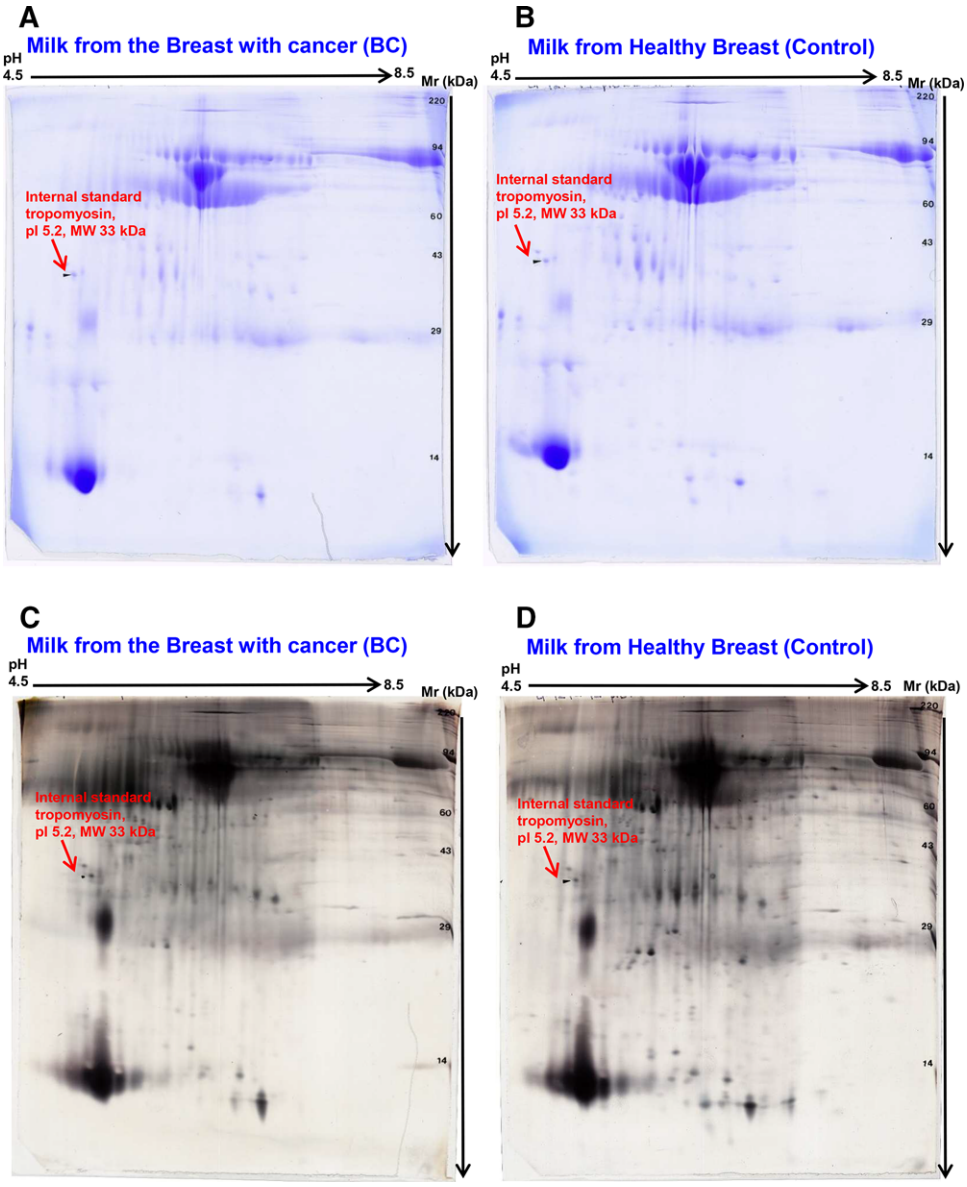

Figure 1. Gel images of Coomassie and silver stained gels for BC sample and control sample. (A) Coomassie stained gel of milk from the breast with cancer (BC) (B) Coomassie stained gel of milk from the healthy breast (control). (C) Silver stained gel of milk from the breast with cancer (BC) (D) Silver stained gel of milk from the healthy breast (control).

The dysregulation was measured as the fold change of the spot percentages. Fold change for upregulation (spot % of BC divided by spot % of control) is shown with positive numbers and downregulation (spot % of Control sample divided by spot % of BC sample) is shown with negative numbers.

2.5 In-gel trypsin digestion of the differentially regulated protein spots

The differentially regulated protein spots from the Coomassie blue stained gels were picked and digested performing in-gel trypsin digestion: the gel spots were cut into very small pieces, then washed and de-stained followed by overnight in-gel trypsin digestion. After digestion, the peptide extraction was done, and the peptide mixture was cleaned using Zip-Tip reversed phase chromatography (C18 Ziptip™; Millipore, Billerica, MA).

2.6 Nano LC-MS/MS and data processing

After ziptip was done and the sample were dried, the peptide mixtures were re-solubilized and analyzed by nanoliquid chromatography tandem mass spectrometry (nanoLC-MS/MS) using a nanoACQUITY® UPLC coupled with a QTOF Ultima API mass spectrometer (Waters, Milford, MA) as described elsewhere [47]. Using ProteinLynx Global Server (PLGS version 2.4, Waters) software, the raw data files from the mass spectrometer were converted to pkl (peak list) files as described elsewhere [48]. The processing parameters used for this conversion were as follows: Background polynomial of

order 5 with background subtraction threshold of 35%, with smoothing iterations of 2, smoothing window of 3 channels and smoothing type of Savitzky-Golay. Centroid top was considered as top 80% of peaks and minimum peak width was considered as four channels. The pkl files were undergone data base search using web based MASCOT search engine (www.matrixscience.com). The database search parameters used in this study were as follow: National Center for Biotechnology Information (NCBI), NCBInr 20150425 (65519838 sequences; 23472502492 residues) database, selected for Homo sapiens (human) (304503 sequences), cysteine propionamide as fixed modification, methionine oxidation as variable modifications, peptide mass tolerance of ±1.3 Da (one ^{13}C isotope), fragment mass tolerance of ±0.8 Da, maximum missed cleavages = 4. Apart from the proteins, which were identified directly from mascot database search, there were some spots that did not show any protein identification in mascot, but peptides from the corresponding protein to that spot were identified. For those peptides we performed *de novo* sequencing on MS/MS raw data files to identify more dysregulated proteins in our experiment. The outline of 2D-PAGE based nanoLC-MS/MS experiment is shown in Supporting Information Fig. 1.

2.7 Data sharing

MS and MS/MS raw data files and Mascot database search links from matrix science website are available upon request, according to Clarkson University' Material Transfer Agreement. The MS data was deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD009860.

3 Results and discussion

3.1 2D-PAGE of breast milk from breast with BC tumor versus healthy breast as a control

To compare the protein pattern of breast milk from the breast with cancer versus control, one Coomassie and two silver stained gels (technical replicate) were run for each breast milk sample (BC and Control). Figure 1 shows the image of the Coomassie stained gels (Fig. 1A and 1B), as well as one of the silver stained gels for each of breast milk sample (Fig. 1C and 1D). The difference image of the averaged gel of BC samples versus the averaged gel of control samples is shown in Supporting Information Fig. 2. Based on the 2D gels overlay, total of 688 spots were analyzed and the ones with a statistically significant difference were picked for the comparisons. The criteria for statistically significant differences was fold change ≥ 2.7 and p value ≤ 0.05 or fold change ≥ 3.0 for upregulated proteins in BC versus control and fold change ≤ –2.7 and p value ≤ 0.05 or fold change of ≤ –3.0 for downregulated proteins in BC versus control. Supporting Information Table 1 shows all the specifications for these spots such as spot number, the isoelectric point (pI), spot percentage (individual spot density divided by total density of all measured spots) in each replicate of silver stained gels as well as the average spot percentage, BC versus control difference (calculated based on average spot percentages for BC and control) and BC versus control T-test p value. These spots underwent in-gel trypsin digestion followed by nanoLC-MS/MS analysis. The resulted raw data were converted to pkl files using PLGS software and the data base search was done on these pkl files, using Mascot database and the corresponding proteins to dysregulated spots were identified. Several dysregulated proteins were identified directly from MASCOT database searches. In addition to those, *de novo* sequencing was done to identify more dysregulated proteins (data is shown in Supporting Information Fig. 3). All these proteins and their specifications (NCBI gi (GenInfo) identifier, protein score, fold change and *T*-test (p value)) are shown in Supporting Information Table 2. Examples of zoomed images of dysregulated protein spots in BC versus control (from the 2D gels), are shown in Fig. 2. The comparative plots of the ratio (%) of spot percentages for BC (in red) versus control (in blue) are also given for these spots. After identification of statistically significant dysregulated proteins, we reviewed the literature for function and the possible role of these proteins in cancer development.

3.2 Upregulated proteins in BC sample versus control

Several proteins were upregulated in BC versus control. The upregulated proteins with fold change ≥ 2.7 and p value ≤ 0.05 or fold change ≥ 3.0 were selected for comparison. As shown in Supporting Information Table 2, the upregulated proteins identified in BC versus control were: alpha-amylase, gelsolin isoform a precursor, alpha-2-glycoprotein 1 zinc isoform CRA_b partial, apoptosis-inducing factor 2 and vitronectin. Below we provide a brief and selected review of the role of these proteins in cancer.

3.2.1 Alpha-amylase

Alpha-amylase, an enzyme for hydrolysis of polysaccharides, is dysregulated in cancer patients. In one study of ovarian cancer patients, the level of amylase was investigated in serum samples of Japanese women with benign (158 patients) and malignant (70 patients) ovarian tumors. Hyperamylasemia (high level of amylase in blood) was observed in 15 of 70 (21%) patients with malignant tumors, while it was only observed in 10 of 158 (6.3%) of patients with benign tumors [8]. In another study of 33 patients with different stages of lung cancer, the amylase level was measured both in serum and tumor tissues. Serum amylase was measured in 27 of 33 patients and hyperamylasemia was observed in 3 of 20 of these cases Also, 18 out of 28 lung cancer tumor tissues had positive staining

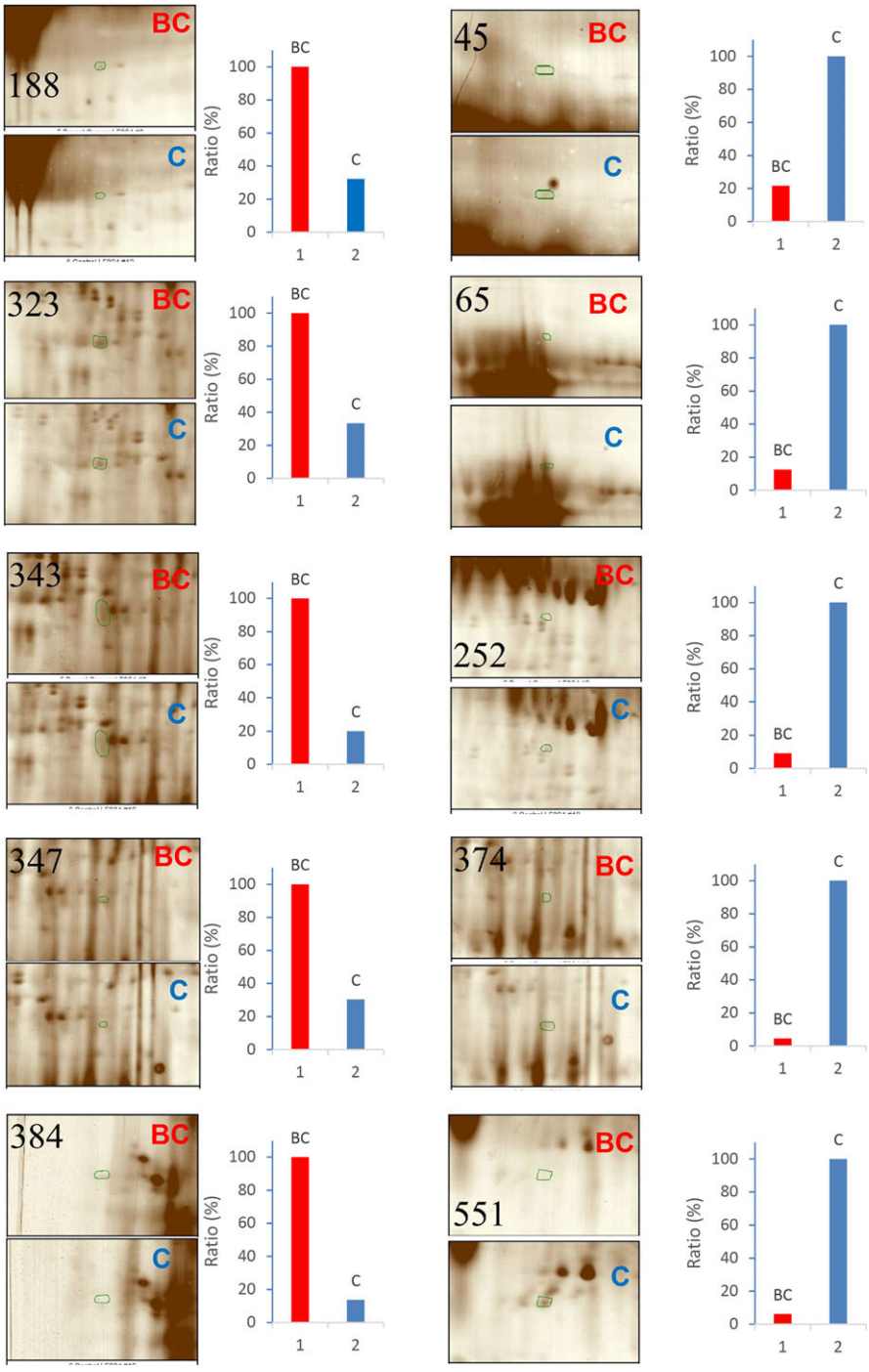

Figure 2. Examples of zoomed images of dysregulated protein spots in BC versus control and their corresponding comparative plots (C represents the control sample).

using immunohistochemistry and 15 out of 19 lung tumor tissues showed positive results for amylase using Northern blot analysis [9]. In a different study, also on lung tumor tissues, 65 of 74 samples showed positive staining results for amylase using immunohistochemistry. The amylase activity was also measured in 52 of these cases, of which 10 had high amylase activity (greater than 1 U/g). Additionally, serum amylase levels were measured in 70 of the cases of which 13 showed hyperamylasemia [10].

3.2.2 Gelsolin isoform a precursor

Gelsolin, a Ca2+ regulated protein, binds to actin filaments and blocks the ends to prevent the exchange of monomer. It plays a role in remodeling of cytoskeleton. The higher levels of gelsolin in BC cell lines can affect important events in cancer development such as proliferation, cell cycle phases and cell migration [11]. The role of gelsolin in cellular events has been investigated in oral cancer cells in which upregulation

of gelsolin resulted in increased cell migration and invasiveness [12]. In another study on colorectal cancer patients, tumor tissues of 50 patients and 15 controls were investigated using immunohistochemistry. Upregulation of gelsolin was observed on the borders of tumors as well as in colorectal cancer cell lines [13].

3.2.3 Alpha-2-glycoprotein 1 zinc isoform CRA_b partial

Zinc alpha 2 glycoprotein is involved in lipolysis (lipid degradation). Based on the Universal Protein Resource (UniProt), breakdown of lipids by this protein causes a considerable depletion in body fat. Interestingly, which happens in cancers, thus the high levels of Zinc alpha 2 glycoproteins that have been reported in cancers causes cancer related cachexia.

Zinc alpha 2 glycoprotein has been considered as a biomarker for different types of cancers including BC in several studies (reviewed in reference [15]). Levels of zinc alpha 2 glycoprotein have been measured in BC tumor tissues. Performing immunohistochemistry on 55 patients with breast tumor and 6 controls, positive staining results were observed for 81% of DCIS and 79% of invasive ductal carcinoma (IDC) cases and 94% of the normal tissues, taken from the surrounding tissues to the tumor tissue in BC patients. No staining results were observed for the control samples investigated in this study [14]. In a different study on 104 BC tumor tissues from different stages of differentiation using enzyme-linked immunoassay, levels of this protein reported to be significantly higher in advanced stages of BC tumor differentiation compared to the earlier stages [16]. High expression levels of zinc alpha 2 glycoprotein gene also has been reported previously based on cDNA clones from human breast library [17].

3.2.4 Apoptosis-inducing factor 2

Apoptosis-inducing factor is an enzyme involved in apoptosis through fragmentation of DNA. This protein plays a role in cancer development and has been considered a target for cancer chemotherapy [19]. Overexpression of the gene for apoptosis-inducing factor has been observed in prostate cancer previously based on the Oncomine database (13 out of 14 studies showed overexpression of the mRNA of apoptosis-inducing factor). In the same study, the prostate cancer tumor tissues of 51 patients were also analyzed and immunohistochemistry showed higher levels of apoptosis-inducing factor in cancer tissues compared to controls (normal prostate tissue) [18]. In a different study on tumor tissues of patients with gastric cancers, 70% of cases showed the positive staining results in immunohistochemical analysis of tissue microarrays, while the controls, except for the parietal cells, did not show any positive results [20].

3.2.5 Vitronectin

Vitronectin is an extracellular involved in cell growth and adhesion, which is important in cancer development. Vitronectin might be related to cancer due to its relationship with plasminogen activation system which is involved in tumor metastasis [22]. It also has been reported that vitronectin is involved in differentiation of stem cells in human serum [23]. In one study on 70 BC tumor tissues and 19 controls, using reverse transcription polymerase chain reaction (RT-PCR) and immunohistochemistry, high levels of vitronectin was observed in the surrounding extracellular matrix of cancer cells [22].

3.3 Downregulated proteins in BC sample versus control

We found several proteins to be downregulated in BC versus control. The downregulated proteins with fold change ≤ -2.7 and p value ≤ 0.05 or fold change ≤ -3.0 were picked for our comparison study. As shown in Supporting Information Table 2, the downregulated proteins we identified in BC versus control were as follows: different isoforms of albumin, cholesterol esterase, different isoforms of lactoferrin, chain A factor B serine protease domain, proline-rich phosphoprotein (gene PRH1 Db allele), serpin peptidase inhibitor clade A (alpha-1 antiproteinase, antitrypsin) member 3 isoform CRA_b, leucine-rich alpha-2-glycoprotein precursor, C2CD2L(C2 domain-containing protein 2-like) protein partial, leukocyte elastase inhibitor, different proteins from casein family, spen homolog transcriptional regulator (Drosophila), protein PAT1 homolog 1, different proteins from apolipoprotein A-I family, manganese-containing superoxide dismutase, DST protein partial (dystonin), galectin-3-binding protein precursor, neuropolypeptide h3, DMD (Dystrophin), different isoforms of lysozyme, chain A crystal structure of the Mrp14 complexed with chaps, chain A, crystal structure of heart fatty acid binding protein (Fabp3), protein S100-A8 (cystic fibrosis antigen) and CST6 (Cystatin M).

3.3.1 Different isoforms of albumin

Different isoforms of albumin were downregulated in our study. Albumin, which is the major protein in blood plasma, has several functions, the major one being binding to different molecules in blood to maintain osmotic pressure. Serum album levels have been reported to be downregulated compared to the normal level of albumin in blood (considered 35 g/L in this study) in 78 out of 183 (42.6%) of patients with carcinomas of unknown primary sites (the carcinoma that occurs in epithelial cells, but the origin is not detectable when the cancer is diagnosed) [24].

3.3.2 Cholesterol esterase

Cholesterol esterase, which is also called bile salt-stimulated lipase, is an enzyme in breast milk that digests triglycerides in milk [49]. The enzymes that are used for the metabolism of lipids as well as their corresponding pathways are dysregulated in different kinds of diseases, including cancer (reviewed in reference [50]). In a recent study done by our group on 5 paired BC versus control samples, including the same pair that is used for this study, on human breast milk, we found human bile salt also downregulated in 2 (including the same pair that is used in this study) out of 5 BC samples versus controls [30].

3.3.3 Different isoforms of lactoferrin

Different isoforms of lactoferrin were downregulated in our study. Several functions reported for lactoferrin are related to cancer development, such as activating transcription and RNase, and involvement in the immune system. Lactoferrin might be related to BC due to the fact that it is hormone responsive [51]. It has been reported that low levels of lactoferrin occur in BC, while higher levels correspond to the normal breast for either lactating or non-lactating women [51]. In a study on five different BC cell lines, downregulation of the mRNA for lactoferrin was observed compared to the control cell line using RT-PCR. The levels of lactoferrin also were measured in 99 primary tumor tissues and 75% of the tumors showed low levels of lactoferin [52]. In a different study, the mRNA levels of lactoferrin as well as the protein levels (using immunohistochemistry) were reported to be downregulated in 31 of 78 BC patients [53]. It also has been shown that the levels of lactoferrin depend on the BC subtype. In a study on 266 BC samples using immunohistochemistry, lower levels of lactoferrin were observed in ER-negative tumor samples compared to ER-positive and in total only 52% of samples were positive for lactoferrin [54].

3.3.4 Chain A factor B serine protease domain

Proteases are enzymes that hydrolyze peptide bonds and are named based on the amino acid involved in the active site of the enzyme. Proteases play a role in tumor growth in cancer [55]. Downregulation of serine protease HTRA has been reported previously in BC. In a study of 131 BC tissues, over 50% showed downregulation of this enzyme [56].

3.3.5 Proline-rich phosphoprotein (gene PRH1 Db allele)

Proline-rich phosphoprotein is a salivary protein that prevents the formation of calcium phosphate crystals to protect teeth. One of the proteins from the family of proline-rich phosphoproteins is Yes-associated protein (YAP), which according to NCBI, is involved in cancer development. YAP has been shown to be downregulated in BC. In one study, YAP positive staining was observed in 82.5% of 40 normal tissues, but only in 45.1% of 266 BC tumor tissue samples [57]. In a different study, YAP was downregulated in different stages of BC, including pre-invasive DCIS (21 out of 33 (63.6%) showed negative results), and invasive breast tumors (64 out of 101 (63.4%) showed negative results) as compared to 20 controls. Based on this study, the low levels of YAP are apparent in the early stages of BC [58]. It is interesting, since the milk samples used in our study are DCIS as well, and they correspond to the pre-cancer state. Based on the same study, lower levels of Yap correspond to more tumor growth and invasion [58]. Finally, in another study, YAP1 levels were investigated in two groups of primary BC samples using immunohistochemistry and Western blot analysis. In group one (screening cohort), the results showed 3, 44, 33 and 20% of non, weak, intermediate and strong protein levels of YAP, respectively. In group two (randomized cohort), the results showed 4, 47, 35 and 14% of non, weak, intermediate and strong protein levels of YAP, respectively. The decrease in YAP levels at early stages of BC was reported in this study [59].

3.3.6 Serpin peptidase inhibitor clade A (alpha-1 antiproteinase, antitrypsin) member 3 isoform CRA_b

Serpins are a family of protease inhibitors. The role of proteases and protease inhibitors in cancer development has been reported previously. Proteases play a role in tumor growth in cancer development, therefore, protease inhibitors could play a role in cancer therapy by inhibiting the activity of proteases [55]. SERPINB5 gene, from the same serpin peptidase inhibitor family, that suppresses the invasion in cancer cells [60], has been reported to be downregulated in BC cell lines and also downregulated as the cancer advances to higher levels of malignancy in different types of cancers such as BC [61]. In a study done on 26 BC primary tumors compared to matched lymph node metastases the levels of SERPINB5 gene (serine (or cysteine) proteinase inhibitor, clade B, member 5), a gene from the same serpin peptidase inhibitor family, was shown to be downregulated in 58% of the lymph node metastases compared to BC primary tumors 58% of the samples using. In the same study, SERPINB5 gene expression was also measured using *in situ* hybridization technique and 48% showed downregulation in lymph node metastases compared to BC primary tumors [62]. In another study on 42 lymph node primary BC tumor samples from Saudi Arabian cancer patients, serpin peptidase inhibitor clade A (SERPINA5) gene is reported to be among the 20 genes that found to be downregulated by fold change greater than 2 [63]. SERPINB1, a different member of SERPIN family, has shown to be downregulated in hepatocellular carcinoma by Western blot analysis of eight paired groups (cancer versus control) of hepatocellular carcinoma fresh samples, as well as immunohistochemical analysis on other 67

paraffin-embedded samples. Except for the tissue samples, three cell lines of hepatocellular carcinoma also showed downregulation of SERPINB1 compared to control normal cell line [64].

3.3.7 Leucine-rich alpha-2-glycoprotein precursor

Leucine-rich alpha-2-glycoprotein is from leucine-rich repeat family. This family is involved in cellular processes such as cell formation, cellular interaction and transmission of signals. Down regulation of leucine-rich alpha-2-glycoprotein has been reported in head and neck squamous cell carcinoma. This downregulation is reported based on the gene database (gene expression omnibus database), as well as in tumor tissues, using immunohistochemistry and Western blot analysis [65].

3.3.8 C2CD2L protein (C2 domain-containing protein 2-like) partial

C2CD2L protein binds to the phospholipids and plays a role in transportation of phospholipids. Lower levels or loss of this protein is reported for the malignant tumors for different types of cancers including BC, based on the human protein atlas [66].

3.3.9 Leukocyte elastase inhibitor

Leukocyte elastase inhibitor is a different name for serpin B1 (it is encoded by SERPINB1 gene). We discussed about serpins and their role in cancer development before (please refer to serpin peptidase inhibitor section).

3.3.10 Different proteins from casein family

We found different proteins of the casein family to be downregulated in BC. Caseins are multifunctional phosphoproteins and are among the most abundant proteins in milk. In our recent previous study several proteins from the casein family were downregulated in four of the five BC versus control comparisons, including kappa-casein isoform X2 [30]. It should be mentioned that among 11 dysregulated caseins identified in the present study, only one casein was found to be upregulated in spot number 347, the other ten caseins were downregulated.

3.3.11 Spen homolog transcriptional regulator (Drosophila)

This protein is from the family of spen (split ends) proteins and is a transcriptional regulator. Lower expression of SPEN gene was observed in invasive BC compared to control (normal breast tissue) and SPEN was shown to act as a tumor suppressor [67].

3.3.12 Protein PAT1 homolog 1

This protein, which is a RNA-binding protein, is involved in degradation of mRNAs and controlling of gene expression [68]. Further, PAT1 homolog 1 protein shows mostly weak to moderate levels in cancers, including BC [69].

3.3.13 Different proteins from apolipoprotein A-I family

Different proteins from apolipoprotein A-I family were downregulated in our study. These proteins play a role in lipids transportation. Proapolipoprotein was among downregulated proteins (fold change of –26.73) in a study of human adenoid cystic carcinoma cells, transplanted in nude mice, and later examined using 2-dimensional differential in-gel electrophoresis coupled with matrix- assisted laser desorption/ionization time-of-flight peptide mass fingerprinting [70].

3.3.14 Manganese-containing superoxide dismutase

Manganese superoxide dismutase is an antioxidant enzyme, which in the presence of reactive oxygen species, can protect cells. Importantly, BC is associated with high levels of reactive oxygen species [71]. Alterations in the levels of antioxidant enzymes have been observed in cancers and mostly downregulation has been reported [72]. In one study of 65 BC tumor tissues (DCIS and IDC compared to non-cancer tissue from the same patient), lower levels of manganese superoxide dismutase were observed in the BC samples (mean combined score of 3.78 for DCIS and 5.3 for IDC) compared to controls (mean combined score of 6.33) [73]. In another study, using RT-PCR, lower levels of SOD2 gene, the gene that encodes manganese superoxide dismutase, were reported in 4 BC cell lines compared to the control cell line (the gene expression in BC was reported as one third of the control cell line). In this study, the level of manganese superoxide dismutase protein, as well as the activity of this protein was shown to be lower in three BC cell lines using Western blot analysis and activity gel assay, respectively [74].

3.3.15 Dystonin (DST) protein partial

Dystonin functions in the cytoskeleton, where it is involved in linking and networking different cytoskeletal filaments. In a review of 13 translational studies (including 553 BC cases and 79 controls) dystonin (bullous pemphigoid antigen 1, isoforms 1/2/3/4/5/8 [fragment]) was among the down regulated genes in 3 studies [75].

3.3.16 Galectin-3-binding protein precursor

Galectin-3-binding protein and galectin 3 protein are involved in cell-cell interactions and signaling processes. These two proteins play a role in different diseases such as cancer [76]. In a study of 196 patients with colorectal carcinoma, lower levels of LGALS3BP gene, the gene that encodes Galectin-3-binding protein, was associated with shorter survival time and greater malignancy of colorectal tumors: 68.5 versus 91% of five-year cumulative survival rate for lower versus higher gene expression of LGALS3BP [77]. Downregulation of galectin-3 itself, also has been reported to be associated with invasiveness and shorter survival rate in BC in a study of 87 BC patients [78].

3.3.17 Neuropolypeptide h3

Neuropolypeptide h3 is also called phosphatidylethanolamine-binding protein 1 (PEBP1) or raf kinase inhibitor protein (RKIP). This protein plays a role in the regulation of multiple signaling pathways. Lower levels of RKIP gene have been reported in metastatic tumor tissues as well as in a metastatic prostate cancer cell line [79, 80]. Lower levels or loss of RKIP also were reported in a study of 103 BC patients (51 lymph node metastases and 52 no lymph node metastases) using immunohistochemistry. RKIP also is reported as a tumor suppressor gene in BC. 17.7% of lymph node metastases samples showed no results for RKIP, 58.8% showed weakly staining results and 23.5% were moderate. Therefore this study also confirms that the lower levels of RKIP are corresponded to the metastatic state [81]. In another study on prostate cancer patients also, lower expression of RKIP gene was associated with metastasis. This study was done on 758 tissue samples from 134 patients (including non-neoplastic prostate tissues, primary tumors and metastases) lower levels of RKIP gene were observed in 89% of metastases, while it was only observed in 48% and 5% of primary tumors and non-neoplastic prostate tissues [82].

3.3.18 Dystrophin (DMD)

Dystrophin is a part of dystrophin associated protein complex. In this complex, dystroglycan binds to dystrophin in cytoskeleton and acts as a bond between cytoskeleton and extracellular matrix. In a study done on 16 types of tumor tissues, all non-myogenic (27 paired comparison groups, including 1765 human cases) the gene expression of DMD was downregulated in 15 out of 27 (56%) of paired groups [83]. Except for downregulation of dystrophin, dystroglycan also has showed downregulation in cancer previously. The levels of dystroglycans have been investigated in BC and colon cancer cell lines and tumors. In this study, the levels of α- dystroglycan were lower in 2 BC cell lines, as well as prostate cancer and leukemia cells, compared to normal cell lines. Lower levels of 43 kD β-dystroglycan were also observed in four cell lines of BC, prostate cancer, colon cancer and leukemia. Using immunostaining, 67 out of 102 (66%) of tumor tissues, mostly the more advanced tumors, showed no results for α-dystroglycan staining. Also, in 30 out of 43 (70%) of colon cancer tissues, α and β dystroglycans were shown to be downregulated using Western blot analysis [84]. The lower levels of dystroglycan has been also reported in another study on six cases of ductal adenocarcinomas BC and 15 cases of prostate cancer tumor tissues using immunohistochemistry [85]. In a different study, levels of α and β dystroglycans were measured in 9 BC cell lines compared to two control cell lines using immunoblotting and six out of nine cell lines showed lower levels of α- dystroglycan [66].

3.3.19 Different isoforms of lysozyme

Different isoforms of lysozyme found to be downregulated in our study. Lysozyme is an enzyme involved in the immune system, performing antibacterial activity by hydrolysis of glycosidic bonds in the peptidoglycan in the cell wall of bacteria. In a previous study performing immunohistochemistry method on 177 BC tumor tissues, 69.4% (126 out of 177) showed positive results for lysozyme, but the level of this protein reported to be lower for lower stages of BC compared to the developed BC samples. Also, low levels of lysozyme corresponded to the patients with shorter survival [86].

3.3.20 Chain A crystal structure of the Mrp14 complexed with chaps

Migration inhibitory factor-related protein 14 (MRP14) is from the family of S100 proteins, named S100 A9. S100 proteins are calcium binding proteins and involved in cellular processes. These proteins are related to cancer development through their role in cancer metastasis and cell growth and have shown altered levels in different cancers [87]. This alteration could be either upregulation or downregulation depending on the type and stage of the cancer [87]. Low levels of a complex including two proteins from the S100 family, S100A8/A9 has shown to stimulate cell growth in a study done on 8 different cancer cell lines including BC cells [88]. Lower levels of mrp14 has been also reported in esophageal cancer in a study done on 66 tumor tissues and 3 cancer cell lines compared to controls (adjacent normal tissue). Performing Northern blot, RT-PCR, Western blot and immunochemistry, low levels of Mrp14 were observed in 60 out of 66 tumor samples and loss of this gene was observed in three out of three cell lines [89].

3.3.21 Chain A crystal structure of heart fatty acid binding protein (Fabp3)

Fatty acid binding proteins play a role in metabolism of fatty acids. Different proteins from the same family have shown

downregulation in cancer. In a study on BC cell lines compared to control cell lines, lower levels of epidermal and adipose fatty acid binding proteins have been reported in BC cell lines [90]. Downregulation (fold change of 3 to 20) of epidermal and adipose fatty acid binding proteins also has been observed in prostate cancer in a study done on both prostate cancer cell lines and prostate cancer tumor tissues using RT-PCR [69]. Regarding the alterations in fatty acids and their related proteins in BC, it must be mentioned that in our previous study done on five paired BC versus control samples, including the same pair used for this study, on human breast milk, we found fatty acid synthase also downregulated in all five out of five BC samples versus controls [30].

3.3.22 Protein S100-A8 (cystic fibrosis antigen)

We discussed about S100 protein and their role in cancer development before (please refer to Mrp14 section).

3.3.23 CST6 (Cystatin M)

Cystatin (cysteine protease inhibitor), same as serpin protease inhibitor, is from the protease inhibitor family, cystatins. Based on NCBI, and atlas of genetics and cytogenetics in oncology and haematology, the gene that encodes this protein is downregulated in BC as metastasis advances [54, 55]. Cystatin has been reported to be a BC tumor suppressor gene [55].

In addition to the proteins discussed above, several antibodies (shown in Supporting Information Table 3) were dysregulated (mostly downregulated, and sometimes upregulated) in the BC sample compared to the control. Since this study was in one individual, and there was no overall pattern for dysregulation of antibodies, we make no conclusions regarding the role of antibodies in the development of BC. However, we reported the data on antibodies because we primarily found them to be downregulated in BC versus control and there are other studies showing downregulation of antibodies in cancer. For instance, downregulation of anti-α-enolase autoantibodies has been investigated in serum samples of BC and lung cancer patients using enzyme linked immunosorbent assay. This study was done on 178 sera samples from BC and lung cancer patients versus 99 controls and 21 patients with other diseases including pneumonia, tuberculosis and heart diseases and statistically significant downregulation of anti-α-enolase autoantibodies was observed in cancer samples [6]. In another study, using affinity chromatography, significantly lower levels of IgG1 was observed in serum samples from 303 invasive and non-invasive BC patients as well as 324 benign breast diseases patients compared to 174 control samples. The mean level of IgG1 was lowest for BC patients compared to the other two groups [7]. The possible role of antibodies in BC development also has been studied previously. For instance, in one study the role of immunoglobulin superfamily has been investigated in BC and they have been considered as biomarkers for prediction of BC [5].

4 Concluding remarks

Expanding on our previous study of milk from 5 pairs of women (cancer versus control) performing 1D-PAGE coupled with MS [30], we now used 2D-PAGE coupled with MS-based proteomics to identify dysregulated proteins in human milk from a breast that developed DCIS 24 months after the milk was obtained. We identified several statistically significant dysregulated proteins (upregulated and downregulated proteins) in BC sample versus control. Some of the dysregulated proteins identified in the current study, were also identified in our previous study [30]. Most of the identified proteins have functions relevant to cancer development and metastasis, many of which have identified previously in cancer tissues, cell lines or other bodily fluids. This pilot study along with the previous pilot study, provide support for the use of proteomic analysis of breast milk to identify biomarkers of BC risk.

4.1 Limitations

The present study has limitations. First, although the within person comparison helps eliminate major differences in protein patterns that may not be associated with BC, differences between left and right breast could still be related to factors independent of disease. For example while one study, reported that the general composition of milk does not differ between right and left breasts, differences were observed in the case of infection [91]. Second, this study conducted with the milk of single participant must be considered as a proof-of-principle and used primarily to support future research.

The authors would like to thank the subject who provided breast milk for this pilot study and the team from Kendrick Labs, Inc. (Madison, WI) (https://kendricklabs.com/) for 2D-PAGE analysis. We also thank Avon foundation for women and the congressionally directed medical research program for supporting milk samples collection through grants to Prof. Kathleen F. Arcaro's group. We also thank the ProteomeXchange Consortium via the PRIDE for help to deposit the MS data [92].

The authors have declared no conflict of interest.

5 References

[1] http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics.

[2] https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html.

[3] https://seer.cancer.gov/statfacts/html/breast.html.

[4] https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html.

[5] Li, Y., et al., *Oncotarget* 2017, *8*, 2444.

[6] Shih, N.-Y., et al., *Jap.J. Clin. Oncol.* 2010, *40*, 663–669.

[7] Kronberger, L., et al., *Breast. Cancer Res. Treat.* 2000, *64*, 193–199.

[8] Kawakita, T., et al., *Gynecol. Oncol. Case Rep.* 2012, *2*, 112–114.

[9] Tomita, N., et al., *Cancer Res.* 1988, *48*, 3288–3291.

[10] Lenler-Petersen, P., et al., *Euro. Resp. J.* 1994, *7*, 941–945.

[11] Chen, Z.-Y., et al., *J. Biomed. Sci.* 2015, *22*, 90.

[12] Deng, R., et al., *Oncol. Lett.* 2015, *9*, 2129–2134.

[13] Zhuo, J., et al., *PloS One* 2012, *7*, e43594.

[14] Dubois, V., et al., *Anticancer. Res.* 2010, *30*, 2919–2925.

[15] Hassan, M. I., et al., *Mol. Cancer Res.* 2008, *6*, 892–906.

[16] Díez-Itza, I., et al., *Euro. J. Cancer* 1993, *29*, 1256–1260.

[17] Freije, J. P., et al., *FEBS Lett.* 1991, *290*, 247–249.

[18] Lewis, E. M., et al., *J. Biol. Chem.* 2012, *287*, 43862–43875.

[19] Millan, A., Huerta, S., *J. Sur. Res.* 2009, *151*, 163–170.

[20] Lee, J. W., et al., *Pathol.-Res. Prac.* 2006, *202*, 497–501.

[21] Delille, H. K., Bonekamp, N. A., Schrader, M., *Int. J. Biomed. Sci: IJBS* 2006, *2*, 308.

[22] Aaboe, M., et al., *Biochimica et Biophysica Acta (BBA)-Mol. Basis Dis.* 2003, *1638*, 72–82.

[23] Hurt, E. M., et al., *Stem Cells* 2010, *28*, 390–398.

[24] Seve, P., et al., *Cancer* 2006, *107*, 2698–2705.

[25] Thompson, P., et al., *Cancer Epidemiol. Prev. Biomark.* 1998, *7*, 37–42.

[26] Gu, Y.-Q., et al., *Int. J. Mol. Med.* 2014, *33*, 1243–1249.

[27] Browne, E. P., et al., *Epigenetics* 2011, *6*, 1425–1435.

[28] Wong, C. M., et al., *Epigenetics* 2010, *5*, 645–655.

[29] Schneider, S. S., et al., *Advancements of Mass Spectrometry in Biomedical Research*, 2014, Springer, 399–408.

[30] Aslebagh, R., et al., *Electrophoresis* 2017, *39*, 653–665.

[31] Faupel-Badger, J. M., et al., *J. National Cancer Institute* 2012, *105*, 166–174.

[32] Yang, H. P., et al., *Cancer Causes & Cont.* 2015, *26*, 345–354.

[33] Arcaro, K. F., et al., *J. Human Lact.* 2012, *28*, 543–546.

[34] Qin, W., et al., *BMC cancer* 2012, *12*, 100.

[35] Murphy, J., et al., *Breast Cancer Res. Treat.* 2016, *157*, 13–22.

[36] Ngounou Wetie, A. G., et al., *J. Lab. Automat.* 2013, *18*, 19–29.

[37] Darie, C. C., *Aust. J. Chem.* 2013, *66*, 719–720.

[38] Sokolowska, I., et al., *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 2013, *1834*, 1474–1483.

[39] Sokolowska, I., et al., *Aust. J. Chem.* 2013, *66* 721–733.

[40] Aslebagh, R., et al., *Mol. Life 1*, 29–42.

[41] Woods, A. G., et al., *Advancements of Mass Spectrometry in Biomedical Research*, 2014, Springer, 1–32,

[42] Jungblut, P., Thiede, B., Schlüter, H., *J. Proteomics* 2016, *134*, 1.

[43] Jungblut, P. R., et al., *Chem. Cent. J.* 2008, *2*, 16.

[44] Schlüter, H., et al., *Chem. Cent. J.* 2009, *3*, 11.

[45] O'Farrell, P. H., *J. Biol. Chem.* 1975, *250*, 4007–4021.

[46] Burgess-Cassler, A., et al., *Clin. Chem.* 1989, *35*, 2297–2304.

[47] Wetie, N., et al., *Autism Res.* 2015, *8*, 338–350.

[48] Sokolowska, I., et al., *Prote. Sci.* 2012, *10*, 47.

[49] Hernell, O., Bläckberg, L., *J. Pediatr.* 1994, *125*, S56-S61.

[50] Beloribi-Djefaflia, S., Vasseur, S., Guillaumond, F., *Oncogenesis* 2016, *5*, e189.

[51] Gopal, S. H., Das, S. K., *J. Cancer Clin. Trials* 2016, *1*.

[52] Benaïssa, M., et al., *Int. J. Cancer* 2005, *114*, 299–306.

[53] Penco, S., et al., *Cancer Biochem. Biophy.* 1999, *17*, 163–178.

[54] Naleskina, L., et al., *Experim. Oncol.* 2016, *38*, 181–186.

[55] Eatemadi, A., et al., *Biomed. Pharma.* 2017, *86*, 221–231,

[56] Lehner, A., et al., *PloS one* 2013, *8*, e60359.

[57] Tufail, R., et al., *Breast Cancer Res. Treat.* 2012, *131*, 743–750.

[58] Yuan, M., et al., *Cell Death Diff.* 2008, *15*, 1752.

[59] Lehn, S., et al., *BMC Cancer* 2014, *14*, 119.

[60] Sheng, S., et al., *Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells*. Proceedings of the National Academy of Sciences, 1996, *93*, 11669–11674.

[61] Yang, S.-F., et al., *Sci. Rep.* 2016, *6*.

[62] Vecchi, M., et al., *Oncogene* 2008, *27*, 2148–2158.

[63] Merdad, A., et al., *BMC Genomics* 2014, *15*, P55.

[64] Cui, X., et al., *J. Mol Histol.* 2014, *45*, 59–68.

[65] Wang, Y., et al., *Oncol. Reports* 1899, *37*, 1503–1510.

[66] Muschler, J., et al., *Cancer Res.* 2002, *62*, 7102–7109.

[67] Légaré, S., et al., *Cancer Res.* 2015, *75*, 4351–4363.

[68] Marnef, A., Standart, N., *Pat1 proteins: a life in translation, translation repression and mRNA decay*, Portland Press Limited, London 2010.

[69] Das, R., et al., *Clinic. Cancer Res.* 2001, *7*, 1706–1715.

[70] Nakashima, D., et al., *Int. J. Cancer* 2006, *118*, 704–713.

[71] Liou, G.-Y., Storz, P., *Free Radical Res.* 2010, *44*, 479–496.

[72] Oberley, T., Oberley, L., *Histol. Histopathol.* 1997, *12*, 525–535.

[73] Tsai, S.-M., et al., *Kaohsiung J. Med. Sci.* 2011, *27*, 167–172.

[74] Hitchler, M. J., Oberley, L. W., Domann, F. E., *Free Rad. Biol. Med.* 2008, *45*, 1573–1580.

[75] Srour, N., Reymond, M. A., Steinert, R., *Pathobiol.* 2008, *75*, 112–118.

[76] DeRoo, E. P., et al., *Blood* 2015, *125*, 1813–1821.

[77] Piccolo, E., et al., *J.Trans. Med.* 2015, *13*, 248.

[78] Ilmer, M., et al., *Breast Cancer Res.* 2016, *18*, 97.

[79] Escara-Wilke, J., Yeung, K., Keller, E. T., *Cancer Metastasis Rev.* 2012, *31*, 615–620.

[80] Keller, E. T., Fu, Z., Brennan, M., *Biochem. Pharmacol.* 2004, *68* 1049–1053.

[81] Hagan, S., et al., *Clinical Cancer Res.* 2005, *11*, 7392–7397.

[82] Fu, Z., et al., *The Prostate* 2006, *66*, 248–256.

[83] Luce, L. N., et al., *Oncotarget* 2017, *8*, 145.

[84] Sgambato, A., et al., *Am. J. Pathol.* 2003, *162*, 849–860.

[85] Henry, M. D., Cohen, M. B., Campbell, K. P., *Human Pathol.* 2001, *32*, 791–795.

[86] Vizoso, F., et al., *Annals Surg. Oncol.* 2001, *8*, 667–674.

[87] Bresnick, A. R., Weber, D. J., Zimmer, D. B., *Nat. Rev. Cancer* 2015, *15*, 96–109.

[88] Ghavami, S., et al., *J. Leukocyte Biol.* 2008, *83*, 1484–1492.

[89] Jie, W., et al., *Cell Res.* 2004, *14*, 46–53.

[90] Hammamieh, R., et al., *J. Exp. Ther. Oncol.* 2005, *5*, 133–143.

[91] Jenness, R., *Semin. Perinatol.*. 1979, *3*, 225–239.

[92] Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., et al., *Nucleic Acids Res.* 2016, *44*, D447–456.

