research articles Journal of proteome research

Secretomic and Proteomic Analysis of Potential Breast Cancer Markers by Two-Dimensional Differential Gel Electrophoresis

Tzu-Chia Lai,[†,#] Hsiu-Chuan Chou,[‡,#] Yi-Wen Chen,[†] Tian-Ren Lee,[†] Hsin-Tsu Chan,[†] Hsin-Hsin Shen,[‡] Wei-Ta Lee,[†] Szu-Ting Lin,[†] Ying-Chieh Lu,[†] Chieh-Lin Wu,[†] and Hong-Lin Chan*[,†]

Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan, and Industrial Technology Research Institute, Hsinchu, Taiwan

Received September 14, 2009

The transformation of a normal cell into a cancer cell has been correlated with alterations in gene regulation and protein expression. To identify altered proteins and link them to the tumorigenesis of breast cancer, we have distinguished normal breast cells (MCF-10A) from noninvasive breast cancer cells (MCF-7) and invasive breast cancer cells (MB-MDA-231) to identify potential breast cancer markers in transformed breast cells. Using the 2D-DIGE and MALDI-TOF MS techniques, we quantified and identified differentially expressed extracellular secreted proteins and total cellular proteins across MCF-7, MB-MDA-231 and MCF-10A. The proteomic analysis of the secreted proteins identified 50 unique differentially expressed proteins from three different media. In addition, 133 unique differentially expressed proteins from total cellular proteins were also identified. Note that 14 of the secreted proteins and 51 of the total cellular proteins have not been previously reported in breast cancer research. Some of these unreported proteins have been examined in other breast cancer cell lines and have shown positive correlations with 2D-DIGE data. In summary, this study identifies numerous putative breast cancer markers from various stages of breast cancer. The results of this study may aid in developing proteins identified as useful diagnostic and therapeutic candidates in research on cancer and proteomics.

Keywords: breast cancer • biomarker • proteomics • secretome • DIGE • MALDI-TOF • tumorigenesis • metastasis

Introduction

Breast cancer is one of the leading causes of death among women around the world. The 5-year survival rate for breast cancer is close to 97% when tumors are confined to breast tissue, but decrease dramatically to 23% when tumors have metastasized to other organs at the time of diagnosis.[1,2] Previous studies indicate that the transformation and metastasis of normal breast cells are correlated to altered expression in both transcription and translation levels.[3−7] To better understand the molecular mechanisms associated with tumorigenesis and metastasis, it is necessary to identify the gene expression signatures and protein expression markers among normal breast cells, noninvasive breast cancer cells, and invasive breast cancer cells. At the transcription level, microarray strategies have been used to classify breast tumors as highly invasive and noninvasive cancer.[8,9] At the translation level, proteomic strategies have been used to discern cancer markers from noninvasive and invasive breast cells.[9−12] Nagaraja et al. compared the proteomic profiling of cell lines corresponding to normal breast cells, noninvasive breast cancer cells, and invasive breast cancer cells using 2-DE. Although they identified 26 spots as potential cancer markers, no statistical analysis was included in their study. Pucci-Minafra et al. compared a ductal infiltrating carcinoma-derived cell line with a nontumoral mammary epithelial cell line using 2-DE, silver staining, immunodetection, and N-terminal sequencing and identified 58 differentially expressed proteins. In contrast to these cell line based studies, Pawlik et al. and Varnum et al. analyzed differentially expressed proteins among nipple aspirate fluid samples from tumor-bearing and disease-free breasts. Although these identified proteins are primarily abundant proteins, few of them have been validated as biomarkers.

During tumorigenesis and metastasis, secreted proteins in the extracellular space are major factors in cell invasion, migration, motility, growth control, angiogenesis, matrix-degradation and adhesion.[13] Consequently, the analysis of tumor secreted proteins is a promising strategy for identifying cancer biomarkers. In the past few years, researchers have used proteomic analysis to identify some secreted biomarker candidates for human cancer using 2D-DIGE and LC-tandem mass spectrometry. These markers have been found in lung cancer, liver cancer, pancreatic cancer and colorectal cancer.[14] In breast cancer research, Kulasingam and Diamandis used a LC-

* To whom correspondence should be addressed. Dr. Hong-Lin Chan, Institute of Bioinformatics and Structural Biology & Department of Life Sciences, National Tsing Hua University, No.101, Kuang-Fu Rd. Sec.2, Hsinchu, 30013, Taiwan. Tel, 886-3-5742476; fax, 886-3-5715934; e-mail, hlchan@life.nthu.edu.tw.

† National Tsing Hua University.

‡ Industrial Technology Research Institute.

# These authors contributed equally to this work.

10.1021/pr900825t

Published on Web 01/07/2010

MS/MS strategy to analyze and compare the expression of extracellular and membrane-bound proteins in conditioned media of three breast cell types corresponding to a normal control and cell lines derived from stage 2 and stage 4 patients. Their study identified several hundred proteins from conditioned media.[1]

2-DE is currently a key technique in profiling thousands of proteins within biological samples and plays a role complementary to LC/MS-based proteomic analysis.[15] However, reliable quantitative comparisons between gels and gel-to-gel variations remain the primary challenge in 2-DE analysis. A significant improvement in the gel-based analysis of protein quantitation and detection was achieved by the introduction of 2D-DIGE, which can codetect numerous samples in the same 2-DE. This approach minimizes gel-to-gel variations and compares the relative amount of protein features across different gels using an internal fluorescent standard. Moreover, the 2D-DIGE technique has the advantages of a broader dynamic range, higher sensitivity, and greater reproducibility than traditional 2-DE. This innovative technology relies on the prelabeling of protein samples with fluorescent dyes (Cy2, Cy3 and Cy5) before electrophoresis. Each dye has a distinct fluorescent wavelength, allowing multiple experimental samples with an internal standard to be simultaneously separated in the same gel. The internal standard, which is a pool of an equal amount of the experimental protein samples, helps provide accurate normalization data and increase statistical confidence in relative quantitation among gels.[16–18]

Whether comparisons of normal cell lines with cancer cell lines actually reflect common changes associated with cancer and can be successfully developed into clinically useful biomarkers or therapeutic targets remains debatable. Thus, a direct comparison of cancer tissue with normal tissue is the best theoretical method of obtaining protein expression signatures during tumor progression. However, a direct comparison of clinical samples increases the amount of false positives due to the heterogeneity of tumor specimens, which interferes with the identification of tumor-specific markers. For this reason, well-characterized model cell lines established from normal and tumor tissue are recognized as more informative in cancer proteomics research. In the field of breast cancer research, MCF-10A, MCF-7 and MDA-MB-231 are widely used to represent normal luminal epithelial cells, noninvasive breast cancer cells derived from the luminal duct and invasive breast cancer cells derived from the same tissue, respectively.[19,20] Accordingly, in this study, we are interested in early stage discovery of putative diagnostic markers and therapeutic targets from this cell model system. To achieve these goals, it is necessary to identify potential biomarkers that reflect the progression of tumorigenesis. Thus, we compared the proteomic profiles of total cellular proteins and secreted proteins of this cell model system using 2D-DIGE to quantitatively identify putative transformation markers in breast cancer.

Materials and Methods

Chemicals and Reagents. Generic chemicals were purchased from Sigma-Aldrich (St. Louis, MO), while reagents for 2D-DIGE were purchased from GE Healthcare (Uppsala, Sweden). All primary antibodies were purchased from Abcam (Cambridge, U.K.) and anti-mouse, anti-goat and anti-rabbit secondary antibodies were purchased from GE Healthcare (Uppsala, Sweden). All the chemicals and biochemicals used in this study were of analytical grade.

Cell Lines and Cell Cultures. The breast epithelial cell line MCF-10A was a gift from Dr. Wun-Shaing Wayne Chang, National Health Research Institute, Taiwan. The breast cancer cell lines MCF-7, MDA-MB-231, MDA-MB-453 and MDA-MB-361 were purchased from American Type Culture Collection (ATCC), Manassas, VA. MCF-10A was maintained in Dulbecco's Modified Eagle's medium and F-12 medium (DMEM/F-12) supplemented with 5% horse serum, L-glutamine (2 mM), streptomycin (100 μg/mL), penicillin (100 IU/mL), epidermal growth factor (20 ng/mL) (all from Gibco-Invitrogen Corp., U.K.), insulin (10 μg/mL) (Sigma) and hydrocortisone (0.5 μg/mL) (Sigma). MCF-7, MDA-MB-231, MDA-MB-453 and MDA-MB-361 were maintained in Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal calf serum (FCS), L-glutamine (2 mM), streptomycin (100 μg/mL), and penicillin (100 IU/mL) (all from Gibco-Invitrogen Corp., U.K.). All cells were incubated at 37 °C and 5% CO_2.

Sample Preparation for Proteomic Analysis. Cells in normal growth medium at ~80% confluence were used for proteomic analysis. For total cellular protein analysis, cells were washed in chilled 0.5× PBS and scraped in 2-DE lysis buffer containing 4% (w/v) CHAPS, 7 M urea, 2 M thiourea, 10 mM Tris-HCl, pH 8.3, and 1 mM EDTA. Lysates were homogenized by passage through a 25-gauge needle 10 times, insoluble material was removed by centrifugation at 13 000 rpm for 30 min at 4 °C, and protein concentrations were determined using Coomassie Protein Assay Reagent (BioRad). For secreted protein analysis, approximately 1.25×10^8 cells were seeded into 25 175-cm^2 cell culture plates for each cell line. After 2 days of incubation, the DMEM or DMEM/F-12 media were discarded, and the cells were rinsed three times with PBS. Subsequently, 375 mL of serum-free DMEM or DMEM/F-12 media was added for an additional 30 h. The media were collected and filtered with 0.45 μm microfilters to remove cell debris and then concentrated 1000-fold with 10-kDa molecular mass cutoff concentrators (Millipore). The concentrated media were then precipitated by adding 1 vol of 100% TCA (at −20 °C) to 4 vol of sample and incubated for 10 min at 4 °C. The precipitated protein was then recovered by centrifugation at 13 000 rpm for 10 min, and the resulting pellet was washed twice with ice-cold acetone. Air-dried pellets were resuspended in 2-DE lysis buffer for protein quantification.

2D-DIGE and Gel Image Analysis. Before performing 2D-DIGE, protein samples were labeled with *N*-hydroxy succinimidyl ester-derivatives of the cyanine dyes Cy2, Cy3 and Cy5 following the protocol described previously.[21,22] Briefly, 150 μg of protein sample was minimally labeled with 375 pmol of either Cy3 or Cy5 for comparison on the same 2-DE. To facilitate image matching and cross-gel statistical comparison, a pool of all samples was also prepared and labeled with Cy2 at a molar ratio of 2.5 pmol Cy2/μg of protein as an internal standard for all gels. Thus, the triplicate samples and the internal standard could be run and quantified on multiple 2-DE. The labeling reactions were performed in the dark on ice for 30 min and then quenched with a 20-fold molar ratio excess of free L-lysine to dye for 10 min. The differentially Cy3- and Cy5-labeled samples were then mixed with the Cy2-labeled internal standard and reduced with dithiothreitol for 10 min. IPG buffer, pH 3−10 nonlinear (2% (v/v), GE Healthcare) was added and the final volume was adjusted to 450 μL with 2D-lysis buffer for rehydration. The rehydration process was performed with immobilized nonlinear pH gradient (IPG) strips (pH 3−10, 24 cm) which were later rehydrated by CyDye-

labeled samples in the dark at room temperature overnight (at least 12 h). Isoelectric focusing was then performed using a Multiphor II apparatus (GE Healthcare) for a total of 62.5 kVh at 20 °C. Strips were equilibrated in 6 M urea, 30% (v/v) glycerol, 1% SDS (w/v), 100 mM Tris-HCl (pH 8.8), 65 mM dithiothreitol for 15 min and then in the same buffer containing 240 mM iodoacetamide for another 15 min. The equilibrated IPG strips were transferred onto 26 × 20-cm 12.5% polyacrylamide gels casted between low fluorescent glass plates. The strips were overlaid with 0.5% (w/v) low melting point agarose in a running buffer containing bromophenol blue. The gels were run in an Ettan Twelve gel tank (GE Healthcare) at 4 W/gel at 10 °C until the dye front had completely run off the bottom of the gels. Afterward, the fluorescence 2-DE were scanned directly between the low fluorescent glass plates using an Ettan DIGE Imager (GE Healthcare). This imager is a charge-coupled device-based instrument that enables scanning at different wavelengths for Cy2-, Cy3-, and Cy5-labeled samples. Gel analysis was performed using DeCyder 2-D Differential Analysis Software v7.0 (GE Healthcare) to codetect, normalize and quantify the protein features in the images. Features detected from nonprotein sources (e.g., dust particles and dirty backgrounds) were filtered out. Spots displaying $a \geq 1.5$ average-fold increase or decrease in abundance with a p-value <0.05 were selected for protein identification.

Protein Staining. Colloidal coomassie blue G-250 staining was used to visualize CyDye-labeled protein features in 2-DE. Bonded gels were fixed in 30% (v/v) ethanol, 2% (v/v) phosphoric acid overnight, washed three times (30 min each) with ddH_2O and then incubated in 34% (v/v) methanol, 17% (w/v) ammonium sulfate, 3% (v/v) phosphoric acid for 1 h., prior to adding 0.5 g/L coomassie blue G-250. The gels were then left to stain for 5−7 days. No destaining step was required. The stained gels were then imaged on an ImageScanner III densitometer (GE Healthcare), which processed the gel images as .tif files.

In-Gel Digestion. Excised poststained gel pieces were washed three times in 50% acetonitrile, dried in a SpeedVac for 20 min, reduced with 10 mM dithiothreitol in 5 mM ammonium bicarbonate pH 8. 0 (Ammonium bicarbonate) for 45 min at 50 °C and then alkylated with 50 mM iodoacetamide in 5 mM Ammonium bicarbonate for 1 h. at room temperature in the dark. The gel pieces were then washed three times in 50% acetonitrile and vacuum-dried before reswelling with 50 ng of modified trypsin (Promega) in 5 mM Ammonium bicarbonate. The pieces were then overlaid with 10 μL of 5 mM Ammonium bicarbonate and trypsinized for 16 h at 37 °C. Supernatants were collected, peptides were further extracted twice with 5% trifluoroacetic acid in 50% acetonitrile, and the supernatants were pooled. Peptide extracts were vacuum-dried, resuspended in 5 μL of ddH_2O, and stored at −20 °C prior to MS analysis.

Protein Identification by MALDI-TOF MS. Extracted proteins were cleaved with a proteolytic enzyme to generate peptides, then a peptide mass fingerprinting (PMF) database search following MALDI TOF mass analysis was employed for protein identification. Briefly, 0.5 μL of tryptic digested protein sample was first mixed with 0.5 μL of a matrix solution containing α-cyano-4-hydroxycinammic acid at a concentration of 1 mg in 1 mL of 50% acetonitrile (v/v)/0.1% trifluoroacetic acid (v/v), spotted onto an anchorchip target plate (Bruker Daltonics) and dried. The peptide mass fingerprints were acquired using an Autoflex III mass spectrometer (Bruker Daltonics) in reflector mode. The algorithm used for spectrum annotation was SNAP (**S**ophisticated **N**umerical **A**nnotation **P**rocedure). This process used the following detailed metrics: Peak detection algorithm, SNAP; Signal to noise threshold, 25; Relative intensity threshold, 0%; Minimum intensity threshold, 0; Maximal number of peaks, 50; Quality factor threshold, 1000; SNAP average composition, Averaging; Baseline subtraction, Median; Flatness, 0.8; Median Level, 0.5. The spectrometer was also calibrated with a peptide calibration standard (Bruker Daltonics) and internal calibration was performed using trypsin autolysis peaks at m/z 842.51 and m/z 2211.10. Peaks in the mass range of m/z 800−3000 were used to generate a peptide mass fingerprint that was searched against the Swiss-Prot/TrEMBL database (v57.12) with 513 877 entries using Mascot software v2.2.06 (Matrix Science, London, U.K.). The following parameters were used for the search: *Homo sapiens*; tryptic digest with a maximum of 1 missed cleavage; carbamidomethylation of cysteine, partial protein N-terminal acetylation, partial methionine oxidation and partial modification of glutamine to pyroglutamate and a mass tolerance of 50 ppm. Identification was accepted based on significant MASCOT Mowse scores ($p < 0.05$), spectrum annotation and observed versus expected molecular weight and pI on 2-DE.

Immunoblotting. Immunoblotting was used to validate the differential expression of mass spectrometry identified proteins. Cells were lysed with a lysis buffer containing 50 mM HEPES, pH 7.4, 150 mM NaCl, 1% NP40, 1 mM EDTA, 2 mM sodium orthovanadate, 100 μg/mL AEBSF, 17 μg/mL aprotinin, 1 μg/mL leupeptin, 1 μg/mL pepstatin, 5 μM fenvalerate, 5 μM BpVphen and 1 μM okadaic acid prior to protein quantification with Coomassie Protein Assay Reagent (BioRad). Thirty micrograms of protein samples was diluted in Laemmli sample buffer (final concentrations: 50 mM Tris, pH 6.8, 10% (v/v) glycerol, 2% SDS (w/v), 0.01% (w/v) bromophenol blue) and separated by 1D-SDS-PAGE following standard procedures. After electroblotting separated proteins onto 0.45 μm Immobilon P membranes (Millipore), the membranes were blocked with 5% (w/v) skim milk in TBST (50 mM Tris, pH 8.0, 150 mM NaCl and 0.1% Tween-20 (v/v)) for 1 h. Membranes were then incubated in primary antibody solution in TBS-T containing 0.02% (w/v) sodium azide for 2 h. Membranes were washed in TBS-T (3 × 10 min) and then probed with the appropriate horseradish peroxidase-coupled secondary antibody (GE Healthcare). After further washing in TBS-T, immunoprobed proteins were visualized using an enhanced chemiluminescence method (Visual Protein Co.).

Immunofluorescence. For immunofluorescence staining, cells were plated onto coverslips (VWR international) for overnight incubation. The cells were fixed with PBS containing 4% (v/v) paraformaldehyde for 25 min, washed three times with PBS, and followed by permeabilization in PBS containing 0.2% (v/v) Triton X-100 for 10 min. Coverslips were rinsed and blocked in PBS containing 5% (w/v) BSA for 10 min before incubation with primary antibodies diluted in 2.5% BSA/PBS for 1 h. After three washings with PBS, samples were incubated with the appropriate fluorescently labeled secondary antibodies diluted in 2.5% BSA/PBS for 1 h. Coverslips were then washed three times with PBS and at least twice with ddH_2O before mounting in Vectashield mounting medium (Vector Lab). Coverslip edges were sealed with nail polish onto glass slides (BDH) and then dried in the dark at 4 °C. For image analysis, cells were imaged using a Zeiss Axiovert 200 M fluorescent microscope (Carl Zeiss, Inc., Germany). The laser intensities used to detect the same immunostained markers from different

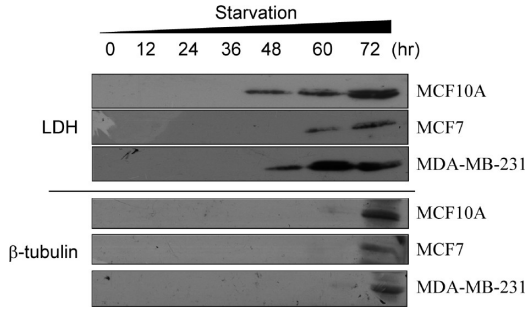

Figure 1. Optimization of starvation time for secretomic analysis. MCF-10A, MCF-7 and MDA-MB-231 on cell culture dishes were used to check starvation induced cell autolysis by detecting the release of cytoplasmic proteins, LDH and β-tubulin in serum-free media. The serum-free media were harvested and concentrated 1000-fold at indicative starvation periods prior to performing immunoblotting analysis.

cell lines were identical, and none of the laser intensities used to capture images were saturated. Images were exported as .tif files using the Zeiss Axioversion 4.0 and processed using Adobe Photoshop V.7.0 software.

Results

Optimization of Cell Conditions for Secreted Protein Analysis. For secretomic analysis, MCF-10A, MCF-7 and MDA-MB-231 were grown on cell culture dishes and the confluency of cells was checked prior to incubation in serum-free culture media to ensure that no other exogenous proteins were present. To minimize cell autolysis induced by starvation and to maximize secreted protein concentration in the media, the starvation time of each cell line was optimized (refer to Materials and Methods). Through immunoblotting, the LDH and β-tubulin levels were detected in the 1000-fold concentrated serum-free media starting at 48−60 h and at 60−72 h, respectively (Figure 1). LDH and β-tubulin are both cytoplasmic proteins and their levels in the media represent the amount of cell death taking place in cell culture. Accordingly, a starvation period of 30 h was chosen for further 2D-DIGE based secretomic analysis.

DIGE and MALDI-TOF Analysis of Secretomes among MCF-10A, MCF-7, and MDA-MB-231 Cells. Proteins secreted from each cell type were enriched from the serum-free medium followed by labeling with CyDyes for 2D-DIGE analysis. The secretomic profiling of MCF-10A, MCF-7 and MDA-MB-231 were visualized using a fluorescence scanner and the images were superimposed using ImageQuant software (Figure 2). To investigate the potential involvement of secreted proteins in tumorigenesis and metastasis for human breast cancer, biological variation analysis of spots showing greater than 1.5-fold change in expression with a *t* test score of less than 0.05 were visually checked before confirming the alterations for protein identification. MALDI-TOF MS identification revealed 50 unique differentially expressed proteins across MCF-10A, MCF-7 and MDA-MB-231 (Table 1). Of the proteins identified, 42 were differentially expressed between MCF-7/MCF-10A, 44 of them were differentially expressed between MDA-MB-231/MCF-10A, and 37 proteins were differentially expressed between MDA-MB-231 and MCF-7. In the three cell lines investigated, 39% of the total proteins identified were extracellular and plasma membrane-anchored proteins (Figure 3A). Most of the identified proteins were involved in signaling transduction, redox-

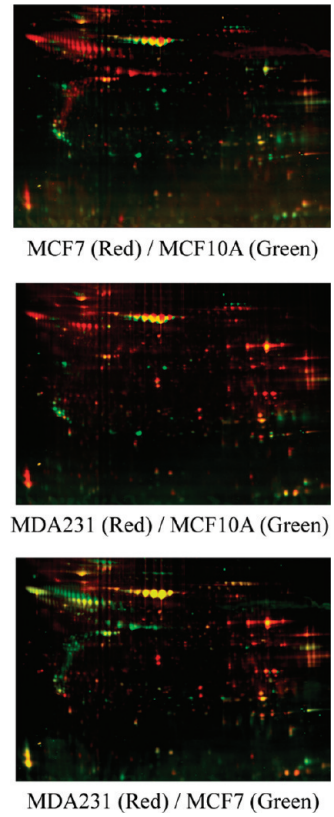

Figure 2. Secretomic comparisons across MCF-10A, MCF-7 and MDA-MB-231 cells using 2D-DIGE. Protein samples (50 μg each) enriched from serum-free media were labeled with Cy-dyes and separated using 24 cm, pH 3−10 nonlinear IPG strips. 2D-DIGE images of MCF-10A, MCF-7, and MDA-MB-231 at appropriate excitation and emission wavelengths were pseudocolored and overlaid with ImageQuant Tool (GE Healthcare).

regulation and metabolism (Figure 3B). To our knowledge, 14 out of these identified spots, including IFIT3, have not been reported in any breast cancer related studies. Consequently, these proteins might have the potential to be putative breast cancer markers. As expected, this 2D-DIGE experiment also identified a number of reported breast cancer markers, including Cathepsin D[23] and IGFBP4.[24]

DIGE and MALDI-TOF Analysis of the Total Cell Proteomes among MCF-10A, MCF-7, and MDA-MB-231 Cells. To identify the altered abundance of proteins and relate them to the tumorigenesis of breast cancer, the proteomic profiles of MCF-10A, MCF-7 and MDA-MB-231 were analyzed. Triplicates of the three different cell lysates were compared using 2D-DIGE to obtain an overview of breast cell tumorigenesis. Image analysis using DeCyder v7.0 clearly defined more than 2500 protein spots (Figure 4). To reduce the intrinsic variability derived from protein samples and gel-to-gel variation, only those protein spots that appeared in all of the triplicate gel images were used for statistical analysis. Furthermore, biological variation analysis of spots showing greater than 1.5-fold change in expression with a *t* test score of less than 0.05 were visually checked before confirming the alterations for protein identification. MALDI-TOF MS identification revealed 133 unique differentially expressed proteins across MCF-10A, MCF-7, and MDA-MB-231 (Table 2). Of the 133 proteins identified, 107 of them had differential expressions between MCF-7/MCF-10A, 63 were differentially expressed between MDA-MB-231/

Table 1. Alphabetical List of Identified Differentially Expressed Secreted Proteins across MCF-10A, MCF-7 and MDA-MB-231 Breast Cells Obtained after 2D-DIGE Coupled with MALDI-TOF Mass Spectrometry Analysis

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Kulasingam et al.[d]
P63104	14-3-3 protein zeta/delta	4.73	27899	17/43	52	137/56	Cytoplasm	Siganl transduction	−2.25	−4.28	−1.85		K
O95861	3′(2′),5′-bisphosphate nucleotidase 1	5.46	33713	8/12	27	76/56	Cytoplasm	Biosynthesis	1.16	2.89	2.56		
O95336	6-phosphogluconolactonase	5.7	27815	9/44	41	60/56	Cytoplasm	Metabolism	1.05	1.95	1.91		K
P60709	Actin	5.29	42052	7/12	23	65/56	Cytoplasm	Cytoskeleton	2.57	1.17	−2.14		K
P60709	Actin	5.29	42052	7/15	22	65/56	Cytoplasm	Cytoskeleton	2.21	−1.04	−2.25		K
P15121	Aldose reductase	6.51	36230	7/24	25	57/56	Cytoplasm	Metabolism	3.14	8.14	2.66		K
P06733	Alpha-enolase	7.01	47481	10/20	26	70/56	Cytoplasm	Metabolism	−2.57	1.86	4.94		K
P08758	Annexin A5	4.94	35971	10/17	26	77/56	Plasma membrane	Signal transduction/ Ca regulation	−1.03	3.28	3.47		K
Q8N1M1	Bestrophin-3/BEST3	6.13	76457	9/22	14	66/56	Plasma membrane	Transport/channel	3.6	2.62	−1.34	A	
P00918	Carbonic anhydrase 2	6.87	29285	7/31	29	62/56	Cytoplasm	Metabolism	−4.72	−6.11	−1.25	B	
P46527	Cyclin-dependent kinase inhibitor 1B/Cyclin-dependent kinase inhibitor p27/p27Kip1	6.54	22288	6/28	32	83/56	Nucleus	Cell cycle	1.65	3.41	2.07		
Q9C0G6	Dynein heavy chain 6	5.72	479761	19/31	5	66/56	Cytoplasm	Cytoskeleton	5.07	2.87	−1.71	A	
Q9NPC3	E3 ubiquitin-protein ligase CCNB1IP1	8.59	31923	6/23	24	60/56	Nucleus	Cell cycle	4.92	5.8	1.22		
Q93070	Ecto-ADP-ribosyltransferase 4/ CD297/ART4	9.31	36197	5/49	15	57/56	Plasma membrane	Biosynthesis	1.75	1.92	1.14	B	
P04075	Fructose-bisphosphate aldolase A	8.3	39851	15/35	47	103/56	Cytoplasm	Metabolism	1.89	1.15	−1.61		K
P21266	Glutathione S-transferase Mu 3	5.37	26998	7/11	25	58/56	Cytoplasm	Redox regulation	4.18	−1.09	−4.46		
P09211	Glutathione S-transferase P	5.43	23569	6/15	38	56/56	Cytoplasm	Redox regulation	−5.67	−4.47	1.3		K
P09211	Glutathione S-transferase P	5.43	23569	7/15	41	73/56	Cytoplasm	Redox regulation	−13.19	−13.28	1.02		K
P04406	Glyceraldehyde-3-phosphate dehydrogenase	8.57	36201	6/17	20	59/56	Mito	Metabolism	−1.71	−1.12	1.57		K
Q9HC38	Glyoxalase domain-containing protein 4	5.4	35170	6/15	20	58/56	Mito	Redox regulation	−1.19	1.71	2.09		
Q8IUY3	GRAM domain-containing protein 2/GRAMD2	8.73	40908	6/26	13	60/56	Plasma membrane	Unknown	2.33	2.85	1.25	A	
Q8IUY3	GRAM domain-containing protein 2/GRAMD2	8.73	40908	9/56	25	63/56	Plasma membrane	Unknown	6.8	6.97	1.06	A	
Q8IUY3	GRAM domain-containing protein 2/GRAMD2	8.73	40908	7/29	19	61/56	Plasma membrane	Unknown	3.81	3.47	−1.06	A	
P04792	Heat shock protein beta-1/HSP 27	5.98	22826	4/26	25	69/56	Cytoplasm	Protein folding	3.43	−1.5	−5.02		K
P04792	Heat shock protein beta-1/HSP 27	5.98	22826	6/14	30	72/56	Cytoplasm	Protein folding	2.39	−2.82	−6.57		K
P29218	Inositol monophosphatase	5.16	30568	6/32	23	64/56	Cytoplasm	Biosynthesis	−2.41	1.1	2.72		K
P22692	Insulin-like growth factor-binding protein 4/IGFBP4	6.81	29113	10/42	37	123/56	Secreted	Siganl transduction	6.64	29.51	4.56		K
P22692	Insulin-like growth factor-binding protein 4/IGFBP4	6.81	29113	8/23	28	105/56	Secreted	Siganl transduction	8.64	25.19	2.99		K
O14879	Interferon-induced protein with tetratricopeptide repeats 3/IFIT3/ ISG60	5.12	56691	6/16	16	66/56	Plasma membrane	Siganl transduction	−9.31	−7.91	1.22	A	
P30740	Leukocyte elastase inhibitor/Serpin B1	5.9	42829	14/34	33	126/56	Cytoplasm	Protease inhibitor	−4.73	−3.25	1.49		K

Table 1 Continued

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Kulasingam et al.[d]
Q14168	MAGUK p55 subfamily member 2/MPP2	6.28	64887	7/24	14	60/56	Plasma membrane	Siganl transduction	2.66	13.13	5.06		
Q14168	MAGUK p55 subfamily member 2/MPP2	6.28	64887	7/24	14	57/56	Plasma membrane	Siganl transduction	1.31	2.53	1.98		
P35240	Merlin/Neurofibromin-2	6.11	69874	10/29	15	70/56	Plasma membrane	Cell motility/signal transduction	4.77	1.54	−3.03		
P26038	Moesin	6.08	67892	14/34	19	96/56	Plasma membrane	Cell motility/ cytoskeleton	−3	1.95	6.04		K
O75380	NADH dehydrogenase [ubiquinone] iron−sulfur protein 6	8.59	14045	6/52	33	58/56	Mito	Electron transport	2.09	2.16	1.07		
P62937	Peptidyl-prolyl *cis*−*trans* isomerase A/cyclophilin A	7.68	18229	10/25	64	112/56	Cytoplasm	Protein folding	−1.88	−1.2	1.61		K
Q06830	Peroxiredoxin-1	8.27	22324	13/33	56	139/56	Cytoplasm	Redox Regulation	2.11	1.29	−1.6		K
P32119	Peroxiredoxin-2	5.66	22049	6/24	35	75/56	Cytoplasm	Redox Regulation	2.06	1.84	−1.09		K
P30086	Phosphatidylethanolamine-binding protein 1/PEBP1/Raf kinase inhibitor protein	7.01	21158	8/33	54	97/56	Cytoplasm	Siganl transduction	1.52	1.26	−1.17		K
P18669	Phosphoglycerate mutase 1	6.67	28900	8/27	35	70/56	Cytoplasm	Metabolism	−1.2	4.81	5.9	B	
Q01814	Plasma membrane calcium-transporting ATPase 2/PMCA2/ATP2B2/Plasma membrane calcium ATPase isoform 2	5.66	137987	7/16	7	71/56	Plasma membrane	Transport	−2.96	−6.05	−1.99		
P05121	Plasminogen activator inhibitor 1/PAI1/SERPINE1	6.68	45088	8/14	14	72/56	Secreted	Protease inhibitor	−10.93	−8.8	1.29		K
P25786	Proteasome subunit alpha type-1	6.15	29822	6/26	23	58/56	Proteasome	Proteins degradation	−1.05	1.56	1.68	A	
P25787	Proteasome subunit alpha type-2	6.92	255996	6/22	33	70/56	Proteasome	Proteins degradation	−1.49	1.05	1.61		K
P49720	Proteasome subunit beta type-3	6.14	23219	7/20	38	67/56	Proteasome	Proteins degradation	−1.45	1.03	1.52	A	
P31150	Rab GDP dissociation inhibitor alpha/GDI1	5	51177	8/32	25	65/56	Plasma membrane	Siganl transduction	12.45	7.78	−1.55		K
P50395	Rab GDP dissociation inhibitor beta/GDI2	6.11	51087	5/38	18	58/56	Plasma membrane	Siganl transduction	1.21	1.68	1.44		K
Q9H5N1	Rab GTPase-binding effector protein 2	4.76	63690	8/26	13	62/56	Plasma membrane	Membrane trafficking	−15.11	−8.81	1.78	A	
Q8WUD1	Ras-related protein Rab-2B	7.68	24427	5/20	25	68/56	Plasma membrane	Siganl transduction	1.2	−1.77	−2.07	A	
Q13228	Selenium-binding protein 1/ SELENBP1	5.93	52928	11/23	20	94/56	Plasma membrane	Transport	1.38	−1.52	−2.03	B	
P36952	Serpin B5	5.72	42568	6/22	18	61/56	Secreted	Protease inhibitor	−16.8	−21.59	−1.25		
P04179	Superoxide dismutase [Mn], mitochondrial	8.35	24878	6/27	30	65/56	Mito	Redox Regulation	−4.36	−1.72	2.6		
P55072	Transitional endoplasmic reticulum ATPase/Valosin-containing protein/VCP	5.14	899550	25/41	34	197/56	Cytoplasm	Vascular transport	−1.31	1.36	1.85		K

Table 1 Continued

accession code	protein name	pI	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Kulasingam et al.[d]
Q8N912	Transmembrane protein C14orf180	11.2	18382	4/38	25	58/56	Plasma membrane	Unknown	1.88	1.91	1.05	A	
Q9P1Q0	Vacuolar protein sorting-associated protein 54/Hepatocellular carcinoma protein 8/Tumor antigen SLP-8p/VPS54	6.1	111545	7/24	11	64/56	Endosome	Vascular transport	−18.84	−25.44	−1.31	A	
O75083	WD repeat-containing protein 1/Actin-interacting protein 1	6.17	66836	18/32	40	197/56	Cytoplasm	Cytoskeleton	−1.26	1.76	2.3		K

[a] The subcellular locations and functional classes of identified proteins were obtained from the Uniprot Web site (http://www.uniprot.org/). [b] The average ratio of differentially expressed ($p < 0.05$) proteins across MCF-7/MCF-10A, MDA-MB-231/MCF-10A, and MDA-MB-231/MCF-7 were calculated considering 3 replica gels. [c] Identified proteins which have not been reported in any cancer research are marked "A", while proteins which have been reported in cancer research but not in breast cancer studies are marked "B". [d] Proteins in this list have been reported in Kulasingam et al.'s study.[1]

MCF-10A, and 96 had differential expressions between MDA-MB-231 and MCF-7. Almost half of the total proteins identified in this breast cell model were cytosolic proteins (Figure 5A), and most of the identified proteins were involved in signaling transduction, metabolism, protein folding, and cell motility (Figure 5B). To our knowledge, 51 of these identified spots, including Calumenin, have not been reported in any breast cancer related studies. As such, these proteins might have the potential to be putative breast cancer markers. As expected, some well-known breast cancer markers, such as 14-3-3 proteins,[25] annexins,[26] calmodulin,[27] AGR-2,[28,29] Galectin-1[30] and ROCK2,[31] were also identified in this 2D-DIGE experiment, lending credence to the reliability of early phase biomarker detection using this experimental strategy.

Validation of Characterized Breast Cancer Related Proteins through Immunoblotting and Immunofluorescence. This secretomic study indentified some of the well-characterized breast cancer related cytosolic proteins such as Cyclophilin A, 14-3-3delta and peroxiredoxin 2 in culture media.[32–34] It is essential to validate the levels of these cytosolic proteins in the medium from independent experiments. To this end, the expression level of cyclophilin A, 14-3-3delta and peroxiredoxin 2 from the culture media of MDA-MB-231, MCF-7 and MCF-10A was validated with immunoblotting. The results indicate that both the proteomic and immunoblot analysis showed cyclophilin A and 14-3-3 delta down-regulated in MCF-7 in comparison to the levels in MCF-10A. In contrast, peroxiredoxin 2 showed up-regulation in MCF-7 in comparison to the levels in MCF-10A. Comparing the secreted protein levels between MCF-10A and MDA-MB-231 indicates that the peroxiredoxin 2 and 14-3-3 delta expression levels increased in MDA-MB-231 and MCF-10A, respectively; however, the cyclophilin A level showed no significant change (Figure 6A−C). This observation confirmed that cyclophilin A, 14-3-3delta and peroxiredoxin 2 were differentially secreted across the breast cells.

Immunoblot and immunofluorescence analysis were carried out to further confirm the differential protein levels observed in the total cellular proteins (profilin, cathepsin D, annexin 2, protein disulfide isomerase A1 and HDAC1) across MDA-MB-231, MCF-7 and MCF-10A (Figure 6D−H). These proteins have been reported to play important roles in cytoskeleton regulation, proteolysis, calcium regulation, protein disulfide bond rearrangement and chromatin assembly during tumorigenesis.[35–39] The results of the immunoblotting indicate that cathepsin D and PDI showed up-regulation in MCF-7 cells but down-regulation in MDA-MB-231 compared to the two protein expressions in MCF-10A. The expression levels of the profilin and annexin 2 proteins showed down-regulation in MCF-7 but no significant changes in MDA-MB-231 compared to the levels in MCF-10A. These immunoblotting results demonstrate a positive correlation with the 2D-DIGE results (Figure 6D−G). In addition to immunoblotting, validation was also performed with immunofluorescent analysis. Figure 6H shows that most of the HDAC1 signal was distributed within the nucleus, which is consistent with the subcellular location of HDAC1 in cells. As expected, the fluorescent intensity with the same exposure indicates that HDAC1 showed increased expressions in MCF-7 and MDA-MB-231 compared to its expression in MCF-10A.

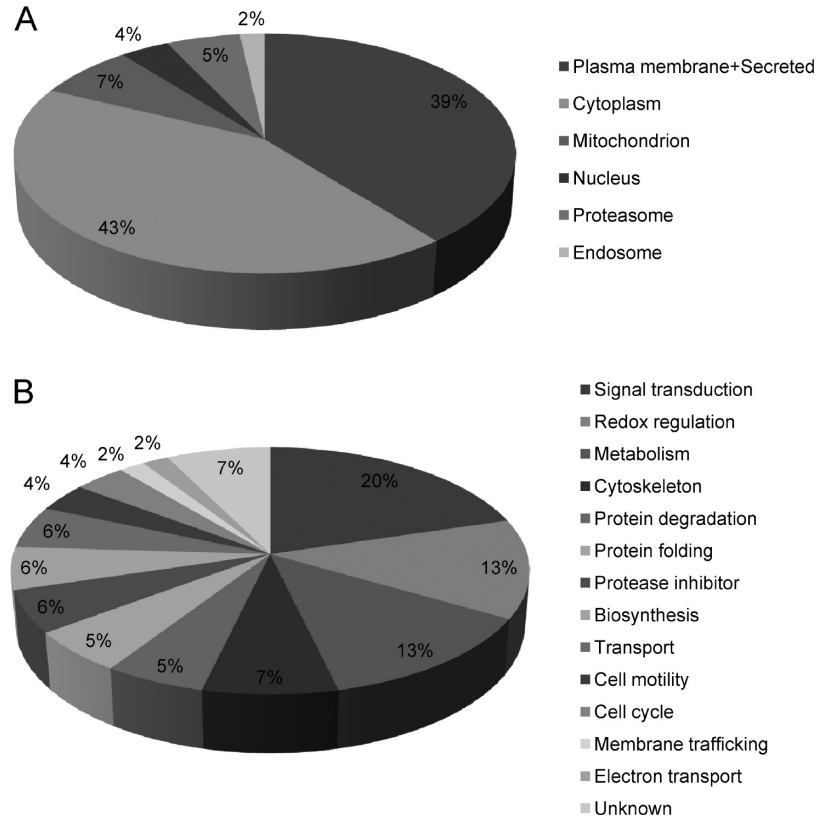

Figure 3. Percentage of secreted proteins identified from serum-free media by 2D-DIGE/MALDI-TOF MS for MCF-10A, MCF-7 and MDA-MB-231 cells according to their subcellular locations (A) and biological functions (B).

Altogether, the results from immunoblotting and immunofluorescence agreed with the results from 2D-DIGE data.

Validation of Unreported Identified Putative Tumorigenesis Markers through Immunoblotting and Immunofluorescence. The cellular proteomic and secretomic analyses above reveal that a number of identified proteins may be breast cancer markers (Tables 1 and 2). To verify this observation, immunoblotting and immunofluorescence were used to validate these differentially expressed proteins including bestrophin 3, MPP2, parvalbumin, PdLIM1, IFIT3 and BANF1 as these proteins showed relatively significant changes (>3-fold) in comparison with most of the unreported identified proteins across MCF-10A, MCF-7 and MDA-MB-231. The immunoblotting analysis of concentrated serum-free media shows that more bestrophin 3 was secreted in the cell lines of MCF-7 and MDA-MB-231 than MCF-10A, while MPP2 was only detected in MDA-MB-231. Notably, the bestrophin 3 blotting result did not completely agree with the 2D-DIGE data, where levels in MCF-7 were higher than MB-231 (Figure 7A). Using immunofluorescence staining, the robust increase of parvalbumin signal in both the MCF-7 and MDA-MB-231 cells was first confirmed after comparison with the signal in MCF-10A. Parvalbumin was primarily localized in the nucleus, which coincided with the DAPI stained nucleus. Further investigation of parvabumin expression in other breast cancer cell lines indicates that parvabumin was overexpressed in MDA-MB-453, a line of noninvasive breast cancer cells, and slightly up-regulated in MDA-MB-361, an adenocarcinoma with metastatic ability (Figure 7B). These results imply that parvabumin might have the potential to be a breast cancer marker. In contrast, PdLIM1, a cytosolic protein, was down-regulated in all breast cancer lines: MCF-7, MDA-MB-231, MDA-MB-453 and MDA-MB-361 (Figure 7B). In addition, IFIT3, a plasma membrane protein, was down-regulated in transformed cells, especially in MCF-7 and MDA-MB-231, and was consistent with the proteomic data from 2D-DIGE (Figure 7B). Interestingly, BANF1, a major nucleus-located protein, was distributed in the cytoplasm of the MCF-10A cells, but was confined within the nucleus in MCF-7, MDA-MB-231 and MDA-MB-453 cells; in addition, BANF1 was distributed within the cytoplasm and nucleus in MDA-MB-361 (Figure 7B). These results indicate that the BANF1 levels were different between normal breast cells and breast cancer cells, and that the subcellular locations of the protein may account for tumorigenesis.

Discussion

The results of this study show that the differentially expressed protein profiles across normal and transformed breast cell lines range from extracellular secreted proteins to intracellular proteins. The 2D-DIGE strategy is powerful enough to identify numerous breast cancer signatures and offers a complementary role to LC/MS-based proteomic analysis. Even though the global coverage of protein mixtures identified by LC-MS based analysis is generally higher than that of 2-DE based analysis, 2-DE based analysis offers some distinct advantages, such as direct protein quantification at protein isoform levels instead of peptide levels to reduce analytical variations.[15] Using the LC-MS/MS strategy, Kulasingam and Diamandis analyzed and compared the expressions of extracellular and membrane-bound proteins in conditioned media of three breast cells corresponding to the normal control cells and cell lines derived from stage 2 and stage 4 patients, respectively.[1] Their study identified 1062 differentially expressed proteins across these

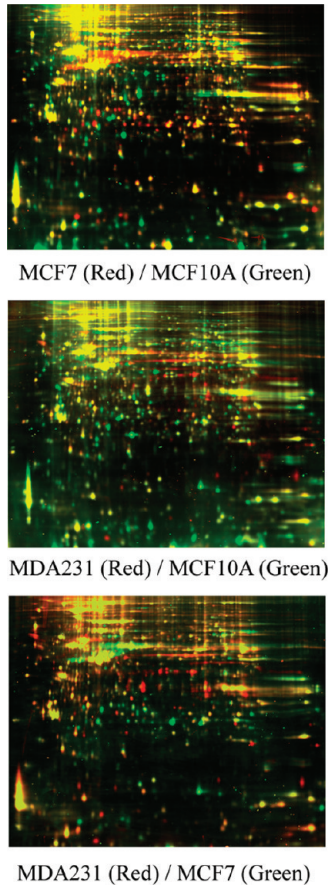

Figure 4. Proteomic comparisons among MCF-10A, MCF-7 and MDA-MB-231 cells using 2D-DIGE. Protein samples (150 μg each) purified from total cell lysates were labeled with Cy-dyes and separated using 24 cm, pH 3−10 nonlinear IPG strips. 2D-DIGE images of MCF-10A, MCF-7 and MDA-MB-231 at appropriate excitation and emission wavelengths were pseudocolored and overlaid with ImageQuant Tool (GE Healthcare).

three cell lines. A comparison between Kulasingam's study and our 2D-DIGE secretomic study shows that 25 out of 50 identified differentially expressed secreted proteins coincide with Kulasingam's study, indicating that both LC-MS/MS and 2D-DIGE are potential tools for discovering breast cancer markers with reasonable reproducibility. Importantly, 25 out of 50 identified proteins were not reported in Kulasingam's study or any other studies, demonstrating that 2D-DIGE plays a powerful complementary role in the assumed biomarker discovery (Table 1).

In a previous study, Nagaraja et al. used traditional 2-DE with poststains (silver stain and coomassie blue stain) to reveal 26 differentially expressed proteins among transformed breast cells with different levels of invasiveness and normal cells which were the same cell lines used in the present study.[9] Their study showed no evidence of visualizing protein spots with sensitive strategies, and protein expression changes were not quantifiable because no broader linear-ranged methods and statistical analysis were employed. Only six out of those 26 proteins coincide with our statistical 2D-DIGE data, which implies that differences might have derived from artificial variations or from results with no statistical analysis (Table 2).

Secreted proteins, plasma membrane bound proteins and extracellular proteins mediate cell adhesion, cell motility, cell−cell interactions and cell invasion. These proteins have the highest possibility of being found in the circulation system, including the blood, and thus serve as cancer markers or important markers involved in cancer formation.[40] To identify potential proteins that may be involved in tumor formation and metastasis, this study develops a strategy for preparing secreted proteins from normal and cancer cell lines with minimal cytosolic protein contamination. Although these cell lines are generally grown in serum-supplemented media, a serum-free conditioned medium is necessary to prevent serum protein contamination and to allow accurate detection of proteins secreted by cells. A serum-free medium is believed to affect the growth of cells and the production of secreted proteins; however, recent studies indicate that the serum-free condition does not significantly affect the composition of the secreted proteins.[41,42] In addition, it is impossible to prevent cell death, and the release of considerable amounts of cytosolic proteins into culture media in either the serum-free condition or the serum-supplemented medium. Accordingly, an intensive wash step was performed prior to incubating these cells in serum-free media to remove both cytosolic proteins and serum proteins. Meanwhile, the incubation time in serum-free media was optimized in advance, minimizing the serum-free induced autolysis of the cells, and enabling the recovery of an adequate amount of secreted proteins for 2D-DIGE analysis. The concentration of secreted proteins in this study was extremely low at approximately 1−2 μg/mL. For this reason, a concentration step was essential to enrich secreted proteins enough for analysis, and a desalting step was also required for the 2D-DIGE experiment. In this study, a spin-concentrator strategy and TCA−acetone precipitation strategy were used for secreted protein enrichment and sample desalting, respectively. After using 2D-DIGE and MALDI-TOF mass spectrometry for protein separation, quantification, and identification, 61% of the total identified proteins in the medium were neither secreted proteins nor membrane-bound proteins. Most of them were sublocated in the cytoplasm, implying that some level of cell necrosis or autolysis was taking place. Interestingly, 39% of the identified proteins in the medium were secreted proteins, plasma membrane bound proteins, or peripheral proteins, indicating that these membrane-associated proteins might be trimmed off the plasma membrane by proteases or might not be completely integrated into the plasma membrane. These results demonstrate that the proposed approach significantly enriches secreted proteins and membrane proteins in comparison with the previous report that only 2% of the entire mammary epithelial cell proteomes are classified as secreted and membrane proteins.[43]

Breast cancer is the leading cause of cancer-induced mortality in women. Early detection of breast cancer greatly improves its survival rates. Accordingly, the primary aim of this study is to discover putative biomarkers with the greatest potential to facilitate early detection of breast cancer and monitor the progress of breast tumorigenesis. Numerous proteins, including bestrophin-3 and parvalbumin, are highly expressed in both low-invasive and aggressive breast cancer cells and are verified as potential breast cancer markers in this study. Importantly, several of these identified proteins, including bestrophin-3, GRAMD2, and nuclear distribution protein nudE homologue 1, have not been reported in previous breast cancer studies, implying that these proteins need to be further investigated to confirm them as valuable breast cancer markers. The identification of cellular targets that play a role in highly invasive breast cancer may also contribute to a better understanding

Table 2. Alphabetical List of Identified Differentially Expressed Total Cellular Proteins across MCF-10A, MCF-7 and MDA-MB-231 Breast Cells Obtained after 2D-DIGE Coupled with MALDI-TOF Mass Spectrometry Analysis

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Nagaraja et al.[d]
P62258	14-3-3 protein epsilon	4.63	29326	8	20%	67/56	Cytoplasm	Sgnal Transduction	−3.37	1.3	4.55		
P63104	14-3-3 protein zeta/delta	4.73	27899	8	28%	74/56	Cytoplasm	Sgnal Transduction	−1.95	1.13	2.26		
P61254	60S ribosomal protein L26	10.55	17248	6	31%	56/56	Cytoplasm	Biosynthesis	9.72	−2.69	−22.09		
P11021	78 kDa glucose-regulated protein	5.07	72402	11	20%	98/56	ER	Biosynthesis	1.73	−1.63	−2.59		
Q6XD76	Achaete-scute homologue 4/ASCL4	9.23	19469	4	31%	59/56	Nucleus	Gene Regulation	1.18	1.74	1.47	B	
P21399	Aconitate hydratase	6.23	98850	10	15%	92/56	Mito	TCA cycle	1.24	1.62	1.31	A	
Q9Y6K8	Adenylate kinase isoenzyme 5	5.38	22358	5	35%	63/56	Cytoplasm	Sgnal Transduction	−2.07	−1.07	2.01		
O95994	AGR2/Anterior gradient protein 2 homologue	9.03	20024	5	34%	63/56	Secreted	Unknown	31.03	1.75	−15		
P15121	Aldose reductase	6.51	36230	7	33%	82/56	Cytoplasm	Metabolism	−1.18	2.3	2.73		
P06733	Alpha-enolase	7.01	47481	16	47%	151/56	Cytoplasm	Metabolism	1.3	1.62	1.24		
Q9H4A4	Aminopeptidase B	5.51	73234	10	20%	98/56	Secreted	Protein catabolism	−1.56	−2.43	−1.55	B	
P04083	Annexin A1	6.57	38918	9	33%	105/56	Cell Membrane	Signal transduction/ Ca regulation	−3.96	−1.03	3.97		
P07355	Annexin A2	7.57	38808	10	34%	92/56	Cell Membrane	Signal transduction/ Ca regulation	−5.67	−2.37	2.47		
P07355	Annexin A2	7.57	38808	14	41%	158/56	Cell Membrane	Signal transduction/ Ca regulation	−8.1	−3.56	2.35		
P12429	Annexin A3	5.63	36524	16	45%	151/56	Cytoplasm	Signal transduction/ Ca regulation	1.61	−3.86	−6.21	B	
P09525	Annexin A4	5.84	36088	6	26%	61/56	Cell Membrane	Signal transduction/ Ca regulation	−1.39	−3.91	−2.82		
P09525	Annexin A4	5.84	36088	10	34%	109/56	Cell Membrane	Signal transduction/ Ca regulation	−1.1	−4.59	−4.19		
P08758	Annexin A5	4.94	35971	9	35%	127/56	Cell Membrane	Signal transduction/ Ca regulation	−2.84	1.65	4.83		
P08758	Annexin A5	4.94	35971	11	42%	142/56	Cell Membrane	Signal transduction/ Ca regulation	−1.81	−1.11	1.63		
P08133	Annexin A6	5.42	76168	12	20%	104/56	Cell Membrane	Signal transduction/ Ca regulation	2.28	1.39	−1.64		
P14868	Aspartyl-tRNA synthetase	6.11	57499	7	21%	61/56	Cytoplasm	Biosynthesis	−3.18	−3.65	−1.15		
P06576	ATP synthase subunit beta	5.26	56525	9	23%	98/56	Mito	Sgnal Transduction	−1.26	−1.66	−1.32		
O75531	Barrier-to-autointegration factor/ Breakpoint cluster region protein 1	5.81	10280	6	55%	76/61	Nucleus	Gene Regulation	3.09	−3.18	−9.49	A	
P31939	Bifunctional purine biosynthesis protein PURH/IMP cyclohydrolase/ATIC	6.27	65089	11	23%	78/56	Cytoplasm	Biosynthesis	2.37	1.14	−2.08	A	
P62158	Calmodulin	4.09	16827	5	39%	81/56	Cytoplasm	Signal transduction/ Ca regulation	2.92	−2.15	−6.08		
P62158	Calmodulin	4.09	16827	5	39%	78/56	Cytoplasm	Signal transduction/ Ca regulation	3.85	−2.6	−9.67		
P04632	Calpain small subunit 1/Capn4/ Calcium-dependent protease small subunit/Calpain regulatory subunit	5.05	28469	5	28%	64/56	Cytoplasm Cell membrane	Signal transduction/ Cytoskeleton remodelling/Ca regulation	−1.59	−1.55	1.02		

Table 2 Continued

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Nagaraja et al.[d]
P27797	Calreticulin	4.29	48283	7	16%	71/56	ER	Signal transduction/ Ca regulation	2.16	−1.36	−2.84		N
O43852	Calumenin	4.47	37198	9	35%	129/56	ER	Signal transduction/ Ca regulation	2.03	−1.16	−2.26	B	
P00918	Carbonic anhydrase 2	6.87	29285	6	26%	69/56	Cytoplasm	Metabolism	−1.54	−9.13	−5.94	B	
P07339	Cathepsin D	6.1	45037	5	14%	63/56	Lysosome	Protein degradation	8.45	−2.1	−17.7		
P07339	Cathepsin D	6.1	45037	5	16%	59/56	Lysosome	Protein degradation	9.91	−9.22	−40.11		
P07339	Cathepsin D	6.1	45037	6	16%	74/56	Lysosome	Protein degradation	5.19	−1.48	−53.69		
P07339	Cathepsin D	6.1	45037	7	16%	57/56	Lysosome	Protein degradation	7.8	−5.32	−88.32		
P29373	Cellular retinoic acid-binding protein 2/CRABP2	5.42	15854	6	46%	66/56	Cytoplasm	Transport	2.37	−5.13	−12.14		
O00299	Chloride intracellular channel protein 1/CLIC1	5.09	27248	8	43%	114/56	Cell Membrane	Transport	−1.14	1.6	1.82		
O00299	Chloride intracellular channel protein 1/CLIC1	5.09	27248	9	44%	98/56	Cell Membrane	Transport	−1.13	1.37	1.55		
Q9Y696	Chloride intracellular channel protein 4/CLIC4	5.45	28982	6	32%	64/56	Cell Membrane	Transport	−5.72	−2.25	2.54		
Q9Y696	Chloride intracellular channel protein 4/CLIC4	5.45	28982	6	26%	69/56	Cell Membrane	Transport	−4.09	−1.87	2.19		
O75390	Citrate synthase	8.45	51908	4	10%	58/56	Mito	TCA cycle	−2.09	−2.68	−1.24		
O14579	Coatomer subunit epsilon (COPE)	4.97	34688	15	50%	181/56	Golgi	Vascular transport	2.17	1.13	−1.86		
P23528	Cofilin-1	8.22	18719	5	37%	56/56	Cytoplasm	Cell motility/Ca regulation	1.33	−3.8	−5.06		
A6NKD9	Coiled-coil domain-containing protein/CCDC85C	6.48	45467	7	16%	58/56	Cytoplasm	Unknown	5.12	6.54	1.28	A	
Q14993	Collagen alpha-1(XIX) chain/ Collagen alpha-1	8.57	115947	7	9%	61/56	Secreted	Cell-cell interaction	3.23	3.84	1.23		
Q15828	Cystatin-M	8.31	16785	4	34%	61/56	Secreted	Other	2.34	1.76	−1.33		
Q04695	Cytokeratin 1	4.97	48361	8	18%	68/56	Cytoplasm	Cytoskeleton	1.54	1.34	−1.15		
P19012	Cytokeratin 15	4.71	49365	9	25%	106/56	Cytoplasm	Cytoskeleton	−6.26	−11.21	−1.73		
P05787	Cytokeratin 8	5.52	53671	14	31%	108/56	Cytoplasm	Cytoskeleton	17.75	1.67	−10.66		
Q9UBS4	DnaJ homologue subfamily B member 11	5.81	40774	6	19%	69/56	ER	Protein folding	2.28	−1.06	−2.34		
P29692	Elongation factor 1-delta	4.9	31217	6	28%	84/56	Cytoplasm	Biosynthesis	1.29	1.56	1.21		
P26641	Elongation factor 1-gamma	6.25	50429	10	36%	124/56	Cytoplasm	Biosynthesis	1.74	1.85	1.06		
P13639	Elongation factor 2	6.41	96246	10	12%	67/56	Cytoplasm	Biosynthesis	1.62	1.11	−1.46		
P30040	Endoplasmic reticulum protein ERp29	6.77	29032	6	24%	69/56	ER	Protein folding	4.31	1.23	−2.92		
P30040	Endoplasmic reticulum protein ERp29	6.77	29032	7	24%	70/56	ER	Protein folding	4.89	1.62	−3.39		
P14625	Endoplasmin	4.76	92696	10	10%	89/56	ER	Protein folding	2.41	1.08	−1.89		
O94905	Erlin-2/SPFH2	5.47	38044	9	27%	111/56	Cell Membrane	Protein catabolism	1.94	−1.53	−2.87	A	
P63241	Eukaryotic translation initiation factor 5A-1	5.08	17049	8	59%	109/56	Nucleus	Biosynthesis	−1.69	1.28	2.53		
P63241	Eukaryotic translation initiation factor 5A-1	5.08	17049	4	50%	56/56	Nucleus	Biosynthesis	−2.54	−1.04	2.23		

Table 2 Continued

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Nagaraja et al.[d]
P47756	F-actin-capping protein subunit beta	5.36	31613	9	34%	97/56	Cytoskeleton	Cell motility	1.53	−1.38	−2.1	B	
P47756	F-actin-capping protein subunit beta	5.36	31613	7	27%	74/56	Cytoskeleton	Cell motility	1.2	−1.81	−2.18	B	
Q02790	FK506-binding protein 4/ Peptidyl-prolyl cis−trans isomerase/FKBP52	5.35	52057	12	26%	94/56	Cytoskeleton	Protein folding	2.83	−1.07	−3.03		
P30043	Flavin reductase	7.13	22219	6	44%	82/56	Cytoplasm	Redox regulation	11.67	5.46	−2.14	B	
P09467	Fructose-1,6-bisphosphatase 1	6.54	37190	12	40%	128/56	Cytoplasm	Metabolism	4.57	1.19	−3.83	A	
P09467	Fructose-1,6-bisphosphatase 1	6.54	37190	14	42%	123/56	Cytoplasm	Metabolism	12.62	−1.12	−14.19	A	
P04075	Fructose-bisphosphate aldolase A	8.3	39851	6	28%	60/56	Cytoplasm	Metabolism	1.66	−1.31	−2.18	B	
P04075	Fructose-bisphosphate aldolase A	8.3	39851	6	28%	80/56	Cytoplasm	Metabolism	1.8	−1.29	−2.32	B	
P09382	Galectin-1	5.34	15048	5	52%	82/56	Cytoplasm	Cell-cell interaction	5.99	−1.18	−6.83		N
P11413	Glucose-6-phosphate 1-dehydrogenase/G6PD	6.39	59675	13	20%	100/56	Cytoplasm	Metabolism	4.39	−1.05	−4.62		
P14314	Glucosidase 2 subunit beta/ PRKCSH/80K−H protein	4.33	60357	13	17%	87/56	ER	Metabolism	2.91	−1.05	−2.96		
P21266	Glutathione S-transferase Mu 3	5.37	26998	8	40%	100/56	Cytoplasm	Redox regulation	16.67	2.54	−6.57		
P21266	Glutathione S-transferase Mu3	5.37	26998	7	37%	88/56	Cytoplasm	Redox regulation	46.76	2.56	−18.29		
P09211	Glutathione S-transferase P	5.43	23569	10	57%	123/56	Cytoplasm	Redox regulation	−27.61	−11.61	2.38		
P09211	Glutathione S-transferase P	5.43	23569	5	31%	60/56	Cytoplasm	Redox regulation	−17.35	−8.71	2.06		
P62993	Growth factor receptor-bound protein 2 (Grb2)	5.89	25304	7	26%	91/56	Golgi	Sgnal Transduction	3.19	1.63	−1.96		
Q12931	Heat shock protein 75 kDa/ TNFR-associated protein 1/TRAP1	8.3	80345	10	18%	64/56	Mito	Protein folding/Cell survival	1.83	−1.09	−1.99	B	
P04792	Heat shock protein beta-1/HSP 27	5.98	22826	9	40%	92/56	Cytoplasm	Protein folding	4.22	−4.51	−19.01		N
P04792	Heat shock protein beta-1/HSP 27	5.98	22826	6	30%	68/56	Cytoplasm	Protein folding	7.97	−3.26	−25.98		N
P04792	Heat shock protein beta-1/HSP 27	5.98	22826	8	36%	78/56	Cytoplasm	Protein folding	7.62	−3.7	−28.21		N
P68871	Hemoglobin subunit beta	6.75	16102	6	53%	83/56	Cytoplasm	Oxygen transport	2.9	2.44	−1.15		
P61978	Heterogeneous nuclear ribonucleoprotein K	5.39	51230	5	13%	74/56	Nucleus	Gene Regulation	−2.66	−2.36	1.17		
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	8.97	37464	7	34%	85/56	Nucleus	Gene Regulation	−2.74	−2.52	1.13	B	
Q13547	Histone deacetylase 1	5.31	55638	7	22%	60/56	Nucleus	Gene Regulation	12.52	11.23	−1.08		
Q9Y4L1	Hypoxia up-regulated protein 1	5.16	111494	11	13%	77/56	ER	Redox regulation	2.05	−1.12	−1.94		
Q04760	Lactoylglutathione lyase/Glyoxalase 1/Aldoketomutase	5.12	20992	6	26%	67/56	Cytoplasm	Metabolism	1.52	1.64	1.08		
P09960	Leukotriene A-4 hydrolase/LTA4H	5.8	69868	9	19%	84/56	Cytoplasm	Metabolism	−1.6	−1.28	1.25	B	
P07195	L-lactate dehydrogenase B chain	5.71	36900	6	17%	58/56	Cytoplasm	Metabolism	−7.5	−1.09	6.89		
P07195	L-lactate dehydrogenase B chain	5.71	36900	6	21%	64/56	Cytoplasm	Metabolism	−3.25	−1.09	2.98		
P40926	Malate dehydrogenase	8.92	35965	7	26%	69/56	Mito	TCA cycle	1.32	−1.92	−2.53		
O00264	Membrane-associated progesterone receptor component 1/PGRMC1	4.56	21772	4	16%	58/56	Cell Membrane	Receptor	3.27	−1.43	−4.51		
O75121	Microfibrillar-associated protein 3 like/MFAP3L	5.13	45750	7	13%	72/56	Cell membrane	Unknown	−1.58	9.38	14.78	B	

Table 2 Continued

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Nagaraja et al.[d]
Q15691	Microtubule-associated protein RP/EB family member 1 (End-binding protein 1)	5.02	30151	9	48%	143/56	Cytoplasm	Cytoskeleton	−1.36	1.44	1.97	B	
P26038	Moesin	6.08	67892	7	12%	73/56	Cell Membrane	Cytoskeleton	−2.33	3.51	8.46		
P60660	Myosin light polypeptide 6	4.56	17090	5	30%	63/56	Cytoplasm	Cell motility	5.49	−1.79	−9.48		
O14950	Myosin regulatory light chain MRLC2	4.71	19824	5	23%	559/56	Cytoplasm	Cell motility	1.84	1.8	1.16		
P19105	Myosin regulatory light chain MRLC3	4.67	19839	4	30%	67/56	Cytoplasm	Cell motility	−3.9	−1.93	2.09		
P30084	Myosin-IXa	4.95	294989	9	5%	58/56	Cytoplasm	Cell motility	−1.03	−2.08	−1.95		
Q14697	Neutral alpha-glucosidase AB/ Glucosidase II subunit alpha	5.74	107263	26	32%	233/56	Golgi	Metabolism	2.34	−1.22	−2.32		
Q14697	Neutral alpha-glucosidase AB/ Glucosidase II subunit alpha	5.74	107263	16	20%	164/56	Golgi	Metabolism	1.87	−1.28	−2.75		
Q9NXR1	Nuclear distribution protein nudE homologue 1/NDE1	5.2	38842	7	21%	57/56	Cytoplasm	Cell cycle	7.78	6.44	−1.17	A	
P15531	Nucleoside diphosphate kinase A/NDP kinase A	5.83	17309	6	42%	89/56	Cytoplasm	Biosynthesis	−1.85	1.07	2.04		N
P20472	Parvalbumin alpha	4.98	12051	6	55%	62/56	Nucleus	Cell motility/Ca regulation	16.77	10.64	−1.52	B	
O00151	PDZ and LIM domain protein 1/LIM domain protein CLP-36/ Elfin/CLP36	6.56	36505	8	27%	97/56	Cytoplasm	Cell motility/Ca regulation	−3.96	−1.46	2.8	B	
P62937	Peptidyl-prolyl cis−trans isomerase A/cyclophilin A	7.68	18229	9	53%	91/56	Cytoplasm	Protein folding	−1.34	−1.65	−1.24		N
P62937	Peptidyl-prolyl *cis−trans* isomerase A/cyclophilin A	7.68	18229	6	36%	68/56	Cytoplasm	Protein folding	−1.41	−1.76	−1.25		N
Q96BP3	Peptidylprolyl isomerase domain and WD repeat-containing protein 1	6.7	74098	6	12%	68/56	Splicesome	Protein folding	−4.68	−3.06	1.58	A	
Q13162	Peroxiredoxin-4	5.86	30749	6	24%	58/56	Cytoplasm	Redox regulation	1.74	1.19	−1.46		
Q13162	Peroxiredoxin-4	5.86	30749	7	32%	80/56	Cytoplasm	Redox regulation	4.92	−1.11	−5.29		
P30041	Peroxiredoxin-6	6	25133	10	51%	116/56	Cytoplasm	Redox regulation	1.4	−1.23	−1.73		
P30041	Peroxiredoxin-6	6	25133	7	31%	98/56	Cytoplasm	Redox regulation	1.09	−1.67	−1.82		
Q9Y617	Phosphoserine aminotransferase/ PSAT	7.56	40796	6	16%	62/56	Cytoplasm	Metabolism	−5.37	−1.26	4.25	B	
P13797	Plastin-3/T plastin	5.52	70904	7	11%	62/56	Cytoplasm	Cytoskeleton	−3.88	−1.67	2.32	A	
P07737	Profilin-1	8.44	15216	5	47%	62/56	Cytoplasm	Cell motility	−1.35	1.2	1.62		
P07737	Profilin-1	8.44	15216	5	47%	63/56	Cytoplasm	Cell motility	−1.29	1.16	1.5		
Q8WUM4	Programmed cell death 6-interacting protein/ ALG-2-interacting protein 1/Hp95/ PDCD6IP	6.13	96590	9	14%	61/56	Cytoplasm	Vascular transport	1.88	−1.07	−2.02	B	
P35232	Prohibitin	5.57	29843	7	37%	82/56	Mito	Gene Regulation	1.08	1.63	1.51		
Q15185	Prostaglandin E synthase 3/PGES	4.35	18971	6	35%	71/56	Cytoplasm	Sgnal Transduction	1.63	1.18	−1.38		

Table 2 Continued

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Nagaraja et al.[d]
Q06323	Proteasome activator complex subunit 1	5.78	28876	12	39%	98/56	Proteasome	Protein degradation	1.73	−1.3	−2.24	A	
Q9UL46	Proteasome activator complex subunit 2	5.44	27515	7	30%	67/56	Proteasome	Protein degradation	1.36	−1.22	−1.66	B	
Q9UL46	Proteasome activator complex subunit 2 (PSME2)	5.44	27515	7	33%	83/56	Proteasome	Protein degradation	1.58	−1.25	−1.98	B	
P61289	Proteasome activator complex subunit 3 (PSME3)	5.69	29602	8	36%	72/56	Proteasome	Protein degradation	−1.11	1.48	1.64		
Q9Y2B0	Protein canopy homologue 2/ CNPY2/MIR-interacting saposin-like protein	4.81	20981	10	53%	161/56	ER	Gene Regulation	1.85	−3.25	−5.82	A	
Q6XLA1	Protein CASC2, isoform 3/Cancer susceptibility candidate gene 2 protein, isoform 3	8.53	12065	4	21%	64/56	Peroxisome	Gene Regulation	−2.97	1.09	3.34	B	
P07237	Protein disulfide-isomerase A1/PDI	4.76	57480	6	12%	66/56	ER	Protein folding	3.33	−1.75	−5.63		
P30101	Protein disulfide-isomerase A3/ ERp57	5.98	57146	10	23%	84/56	ER	Protein folding	−2	−1.41	1.42		
Q15084	Protein disulfide-isomerase A6/ PDIA6/Protein disulfide isomerase P5	4.95	48490	7	24%	77/56	ER	Protein folding	−1.12	−1.66	−1.49	A	
Q9HCY8	Protein S100A14	5.16	11826	5	63%	92/56	Cytoplasm	Signal transduction/ Ca regulation	−5.11	−6.92	−1.31		
Q6PI26	Protein SHQ1	4.7	65712	6	14%	56/56	Cytoplasm	Unknown	−1.27	−4.27	−3.36	A	
P14618	Pyruvate kinase isozymes M1/M2/ Pyruvate kinase 2/3/PKM2/THBP1	7.96	58470	10	23%	101/56	Cytoplasm	Metabolism	−1.77	−1.54	1.15		
P50395	Rab GDP dissociation inhibitor beta/GDI2	6.11	51087	15	39%	158/56	Cell Membrane	Sgnal Transduction	1.53	1.19	−1.29	B	
Q8WUD1	Ras-related protein Rab-2B	7.68	24427	5	29%	72/56	Cell Membrane	Sgnal Transduction	−5.45	−13.37	−2.07	A	
Q14257	Reticulocalbin-2/ERC-55/RCN2	4.26	36911	6	25%	76/56	ER	Signal transduction/ Ca regulation	1.5	−1.78	−2.59	B	
O94788	Retinal dehydrogenase 2/Aldehyde dehydrogenase family 1 member A2/ALDH1A2	5.79	57144	7	14%	61/56	Cytoplasm	Biosynthesis	6.02	−1.12	−6.51		
P42331	Rho GTPase-activating protein 25	5.83	72955	6	11%	62/56	Cell Membrane	Cell motility	1.54	1.11	−1.17	A	
Q13017	Rho GTPase-activating protein 5/p190-B/ARHGAP5/RhoGAP 5	6.18	173834	8	6%	63/56	Cell Membrane	Cell motility	1.93	1.31	−1.47	B	
O75116	Rho-associated protein kinase 2/ROCK2	5.75	161952	9	5%	59/56	Cytoplasm	Cell motility	5.16	−1.56	−8.07		
O75116	Rho-associated protein kinase 2/ROCK2	5.75	161952	14	11%	66/56	Cytoplasm	Cell motility	8.04	−2.56	−20.57		
P13489	Ribonuclease inhibitor/RNase inhibitor	4.71	51766	9	30%	115/56	Cytoplasm	Sgnal Transduction	−1.54	−1.68	−1.09	A	
Q13228	Selenium-binding protein 1/ SELENBP1	5.93	52928	8	15%	71/56	Cell Membrane	Transport	3.04	−1.1	−3.34	B	
Q13228	Selenium-binding protein 1/ SELENBP1	5.93	52928	9	15%	83/56	Cell Membrane	Transport	6.76	−1.12	−7.56	B	

Table 2 Continued

accession code	protein name	p*I*	MW	no. match peptides	seq cov (%)	score	subcellular location[a]	functional class[a]	MCF7/ MCF10A[b]	MDA231/ MCF10A[b]	MDA231/ MCF7[b]	putative breast markers[c]	reported by Nagaraja et al.[d]
Q9NVA2	Septin-11	6.36	49652	9	27%	110/56	Nucleus	Cell cycle	2.15	−1.9	−3.98	A	
Q92599	Septin-8	5.89	56234	9	22%	104/56	Nucleus	Cell cycle	2.36	1.05	−2.17	A	
Q8TDX7	Serine/threonine-protein kinase Nek7	8.49	34985	7	22%	66/56	Cytoplasm	Cell cycle	−2.14	−1.53	1.66	A	
Q00536	Serine/threonine-protein kinase PCTAIRE-1	7.23	55909	7	15%	68/56	Cytoplasm	Cell cycle	4.12	−1.28	−5.28	A	
P36952	Serpin B5	5.72	42568	7	28%	81/56	Secreted	Protease inhibitor	−10.78	−11.05	−1.02		
P55854	Small ubiquitin-related modifier 3/SUMO3	5.32	11687	5	37%	81/56	Cytoplasm	Protein degradation	−2.61	−1.83	1.48	B	
P31948	Stress-induced-phosphoprotein 1/STIP1	6.4	63227	9	20%	62/56	Cytoplasm	Protein folding	−2.2	1.07	2.42	B	
Q8NBS9	Thioredoxin domain-containing protein 5/ERp46/TXNDC5	5.63	48283	7	22%	68/56	ER	Redox regulation	5.71	−1.11	−6.13	B	
Q8NBS9	Thioredoxin domain-containing protein 5/ERp46/TXNDC5	5.63	48283	9	26%	124/56	ER	Redox regulation	5.8	−1.53	−8.61	B	
Q9UI15	Transgelin-3	6.84	22629	6	27%	63/56	Cytoplasm	Unknown	5.15	2.84	−1.53		
P60174	Triosephosphate isomerase/TPI	6.45	26938	12	66%	142/56	Cytoplasm	Metabolism	1.61	−1.05	−1.68		N
P60174	Triosephosphate isomerase/TPI	6.45	26938	7	34%	95/56	Cytoplasm	Metabolism	1.5	−1.45	−2.18		
P67936	Tropomyosin alpha-4 chain/ Tropomyosin-4	4.67	28619	7	22%	80/56	Cytoplasm	Cell motility/Ca regulation	−1.72	1.38	2.46		
P61088	Ubiquitin-conjugating enzyme E2/Ubc13	6.13	17184	6	41%	70/56	Nucleus	Gene Regulation	−2.28	1.29	3.04	A	
Q5 VYV7	UPF0492 protein C20orf94	9.5	46094	5	17%	58/56	Unknown	Unknown	−4.17	−2.6	1.66	A	
P08670	Vimentin	5.06	53676	6	15%	70/56	Cytoplasm	Cytoskeleton	−3.55	4.14	15.21		
P08670	Vimentin	5.06	53676	16	32%	93/56	Cytoplasm	Cytoskeleton	−3.77	2.65	10.32		
P21796	Voltage-dependent anion-selective channel protein 1/VDAC1	8.62	30868	8	40%	110/56	Mito	Transport	−1.65	−2.13	−1.29	B	
P45880	Voltage-dependent anion-selective channel protein 2/VDAC2	7.49	32060	6	31%	84/56	Mito	Transport	−1.55	−1.62	−1.05	B	
Q2 VY69	Zinc finger protein 284	8.77	71198	8	11%	59/56	Nucleus	Gene Regulation	−8.32	−2.7	3.19		
Q8N7K0	Zinc finger protein 433	9.38	79872	11	24%	75/56	Nucleus	Gene Regulation	7.92	−1.02	−8.04	B	
Q6ZNA1	Zinc finger protein 836	9.39	111011	6	8%	56/56	Nucleus	Gene Regulation	2.2	1.21	−1.53		

[a] The subcellular locations and functional classes of identified proteins were obtained from the Uniprot Web site (http://www.uniprot.org/). [b] The average ratio of differentially expressed ($p < 0.05$) proteins across MCF-7/MCF-10A, MDA-MB-231/MCF-10A, and MDA-MB-231/MCF-7 were calculated considering 3 replica gels. [c] Identified proteins which have not been reported in any cancer research are marked "A", while proteins which have been reported in cancer research but not in breast cancer studies are marked "B". [d] Identified proteins in this study have been reported in Nagaraja et al's study.[9]

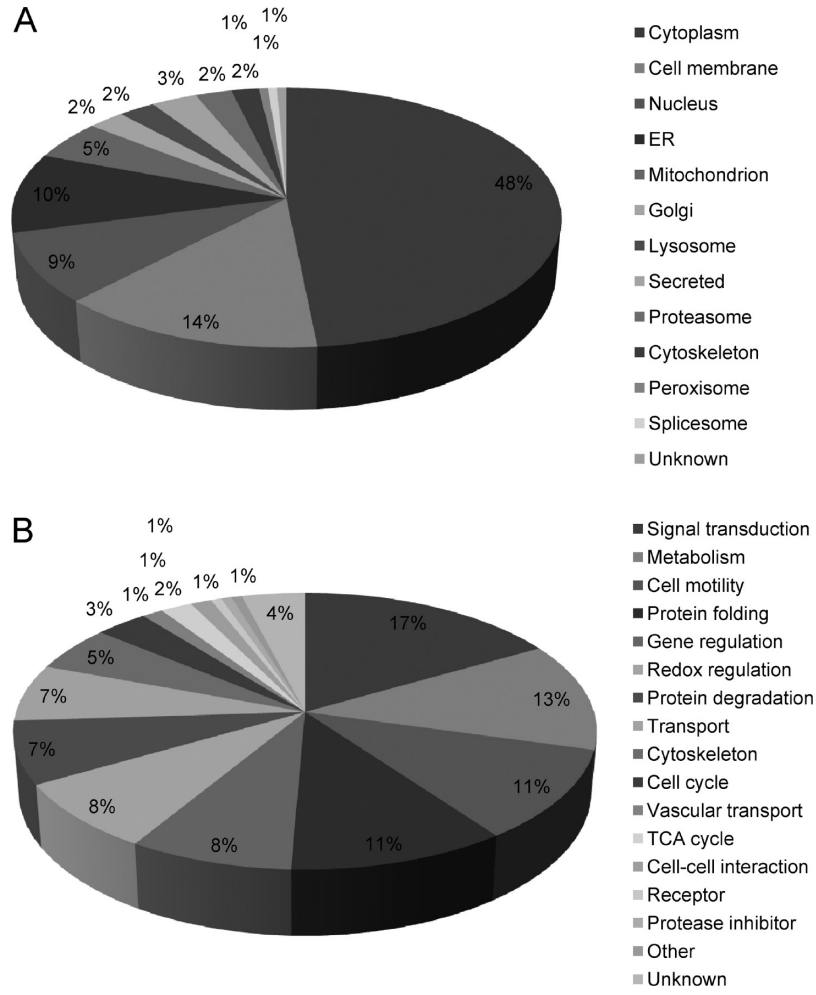

Figure 5. Percentage of total cellular proteins identified by 2D-DIGE/MALDI-TOF MS for MCF-10A, MCF-7 and MDA-MB-231 cells according to their subcellular locations (A) and biological functions (B).

of the biological mechanisms inherent in the aggressive progression of cancer and may be of use in the development of new diagnostic or therapeutic strategies for breast cancer. For this purpose, invasion associated markers were further investigated, including upregulated microfibrillar-associated protein 3 and downregulated annexin A4, carbonic anhydrase 2, plasma membrane calcium-transporting ATPase 2, protein SHQ1, Rab-2B and ROCK2.

With the basis of a Swiss-Prot search and KEGG pathway analysis, numerous potential biological functions of the identified proteins across MCF-10A, MCF-7 and MDA-MB-231 were determined. The information should be useful for studying the mechanisms of breast cancer tumorigenesis and metastasis. Figure 8 compares the expression profiles of the identified differentially expressed proteins in these 3 cell lines. Proteins known to regulate cell cycle are found to be upregulated in both MCF-7 and MDA-MB-231 (Figure 8A), and are associated with the promotion of tumorigenesis.[44] In addition, the expression of proteins linked to redox-regulation increased in the MCF-7 cells in comparison to the levels in MCF-10A (Figure 8B). Induced expression of these proteins may be able to account for cancer development and progression. For example, Noh et al. showed that peroxiredoxins are greatly overexpressed in most breast cancer tissues.[45] Proteomic analysis also reveals that proteins involved in carbohydrate metabolism are significantly overexpressed in MCF-7 cells (Figure 8C). This demonstrates that cancer cells rely heavily on glycolysis to obtain ATP for proliferation and tumorigenesis in the presence of adequate oxygen levels;[46] this mechanism has been implicated in numerous cancer therapies.[47,48] Figures 8D−F show the downregulated profiles of proteins in both MCF-7 and MDA-MB-231 cells. These proteins are involved in calcium regulation, vascular transport and protease inhibition. Calcium-binding proteins, such as annexin 1, whose function is modulated by an estrogen receptor, have been reported to show decreased expression in correlation with breast cancer development and progression.[49–52] The S100 protein family is a family of low molecular weight calcium-binding proteins that is responsible for the regulation of protein phosphorylation, intracellular calcium homostasis, the dynamics of cytoskeleton constituents and cell proliferation.[53] The S100 family has become a major interest because of its deregulated expression in human diseases, especially in cancer. According to Ji et al, S100 families exhibit significantly reduced expression in esophageal squamous cell carcinoma[54] and are hence recognized as a prognostic esophageal cancer marker. In this study, S100A14 was identified as downregulated in MCF-7 and MDA-MB-231, suggesting their potential roles in breast cancer. Interestingly, proteins involved in vascular transport, including Rab GTPase-binding effector protein and vacuolar protein sorting-associated protein 54, were decreased in expression in MCF-7 and MDA-MB-231 (Figure 8F). This may be explained by a previous report

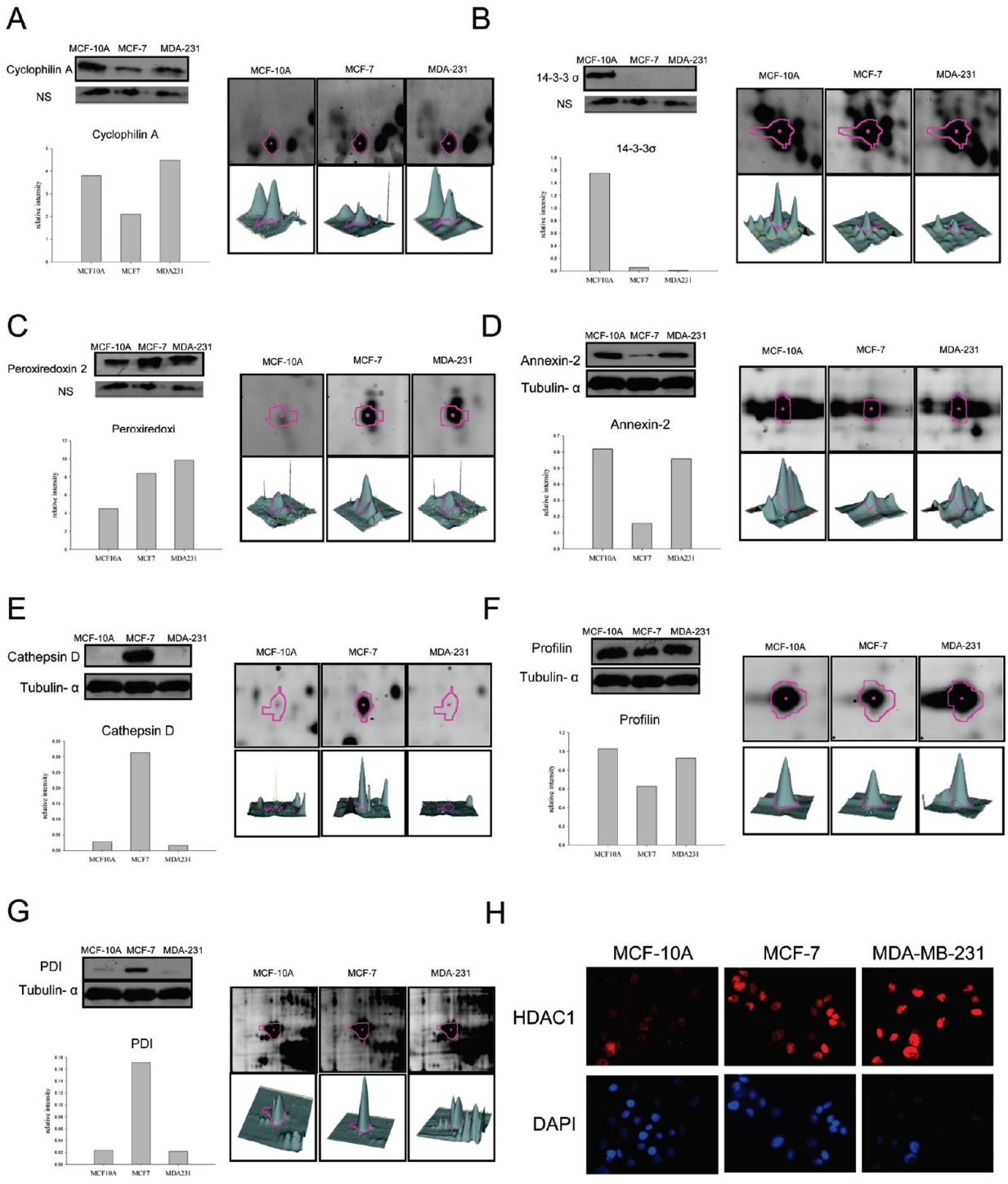

Figure 6. Representative immunoblotting and immunofluorescent analyses for selected differentially expressed proteins identified by proteomic analysis in MCF-10A, MCF-7 and MDA-MB-231 cells. The levels of identified proteins in serum-free media, (A) Cyclophilin A, (B) 14-3-3 delta and (C) Peroxiredoxin 2 and total cellular proteins, (D) Profilin, (E) Cathepsin D, (F) Annexin 2 and (G) Protein disulfide isomerase A1 in MDA-MB-231 and MCF-7 versus MCF-10A confirmed by immunoblot (left top panels), densitometry results with normalized values using nonspecific bands (NS) of secreted proteins and β-tubulin as loading controls (left bottom panels), protein expression map (right top panels) and three-dimensional spot image (right bottom panels). (H) MCF-10A, MCF-7 and MDA-MB-231 cells were fixed and incubated with anti-HDAC antibody and stained with a Texas Red-conjugated secondary antibody (Red). Nuclei were stained with DAPI (Blue). Each set of three fields was taken using the same exposure, and images are representative of three different fields. Scale bar = 20 μm.

indicating that the downregulation of Rab5 GDP/GTP exchange factor enhances receptor tyrosine kinase signaling and promotes the growth factor-directed migration of tumor cells.[55] However, there are few studies on tumorigenesis regarding the roles of the Rab GTPase-binding effector protein and the vacuolar protein sorting-associated protein 54. Serpin is a group of proteins able to inhibit protease and block the growth, invasion, and metastatic properties of breast tumors. Hence,

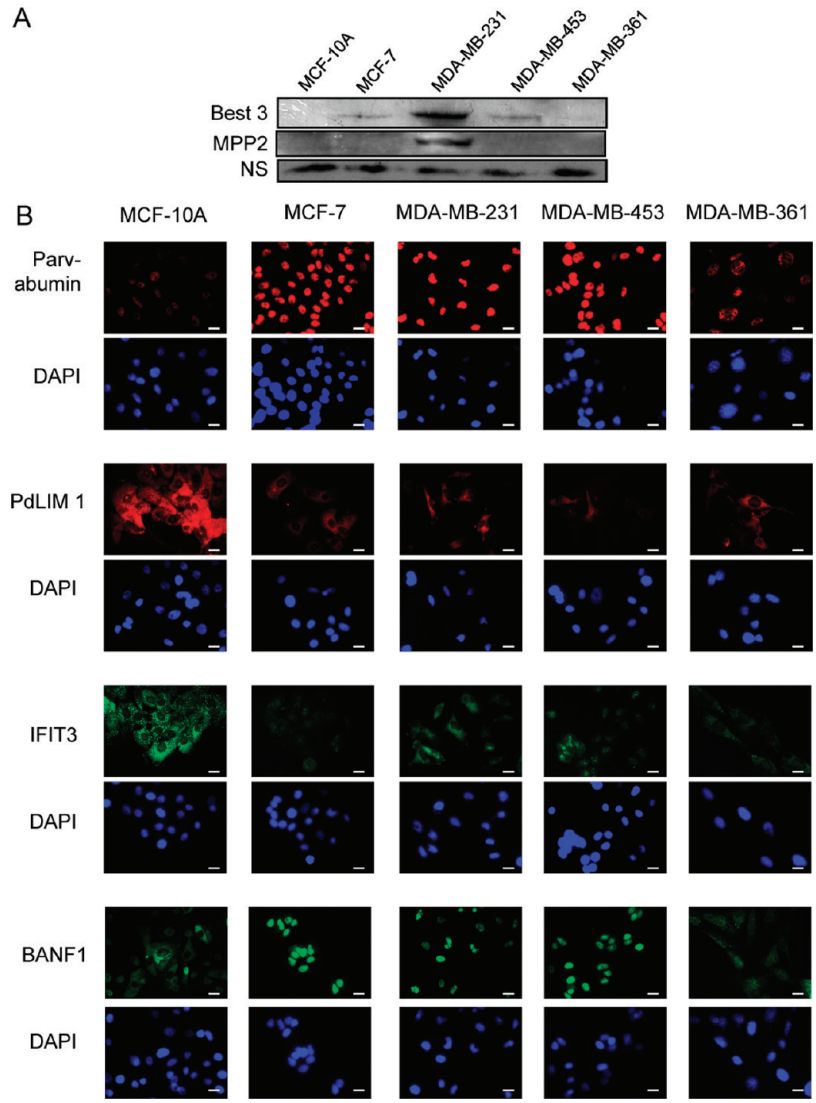

Figure 7. Immunoblotting and immunofluorescence analyses of the expression and protein localization changes of newly identified putative breast cancer markers across MCF-10A, MCF-7, MDA-MB-231, MDA-MB-453 and MDA-MB-361 cells. (A) The profile of the secreted proteome changes across MCF-10A, MCF-7, MDA-MB-231, MDA-MB-453 and MDA-MB-361 cells. The serum-free media from the cell lines was concentrated and 10 μg of the total protein was resolved using SDS-PAGE and immunoblotted for MPP2 and Bestrophin 3. NS represents a nonspecific band used to show equal loading of secreted proteins. (B) A total of 5×10^4 MCF-10A, MCF-7, MDA-MB-231, MDA-MB-453 and MDA-MB-361 cells were seeded on coverslips before fixation and staining for Parvabumin, BANF1, PdLIM1 and IFIT3. Each set of three fields was taken using the same exposure, and images are representative of three different fields. Scale bar = 20 μm.

serpin families function as tumor suppressors in cancer research.[56] The downregulation of serpin is well-correlated with the progression of breast cancer[57] and our own observations in MCF-7 and MDA-MB-231 cells (Figure 8F).

Other differentially expressed proteins of interest across MCF-10A, MCF-7 and MDA-MB-231 include cathepsin D, bestrophin-3 and interferon-induced protein with tetratricopeptide repeats 3 (IFIT3). Cathepsin D, a lysosomal aspartic protease, is overexpressed in estrogen receptor positive breast cancer cells[58] and is generally of good prognostic value in comparison with estrogen receptor negative breast cancer in clinical studies.[59] Our study indicates that cathepsin D is highly expressed in MCF-7, both in total cellular proteins or in secreted fraction. In contrast, cathepsin D is significantly downregulated in MDA-MB-231 cells compared with MCF-7. Thus, our proteomic results display good correlation with these reports. To our knowledge, bestrophin-3, a cGMP-dependent calcium-activated chloride channel, has not been reported to be associated with cancer and shows upregulation in MCF-7 and MDA-MB-231 in this study. Nevertheless, the related study in bestrophin-1 shows the protein improves intracellular Ca^{2+} signaling and increases cell growth rate in colonic carcinoma cells. The proliferation of the cells is significantly suppressed by bestrophin-1 RNA interference treatment.[60] This indicates bestrophin-3 may be a potential target for breast cancer therapy. IFIT3 plays a key role in the antiproliferative activity of the interferon-related signaling pathway through inducing expression of cell cycle inhibitors, p21 and p27 proteins.[61] The 2D-DIGE results in this study show that IFIT3 is downregulated in both MCF-7 and MDA-MB-231 cells, implying that breast cancer cells may maintain a high level of proliferative activity by downregulating the expression of IFIT3.

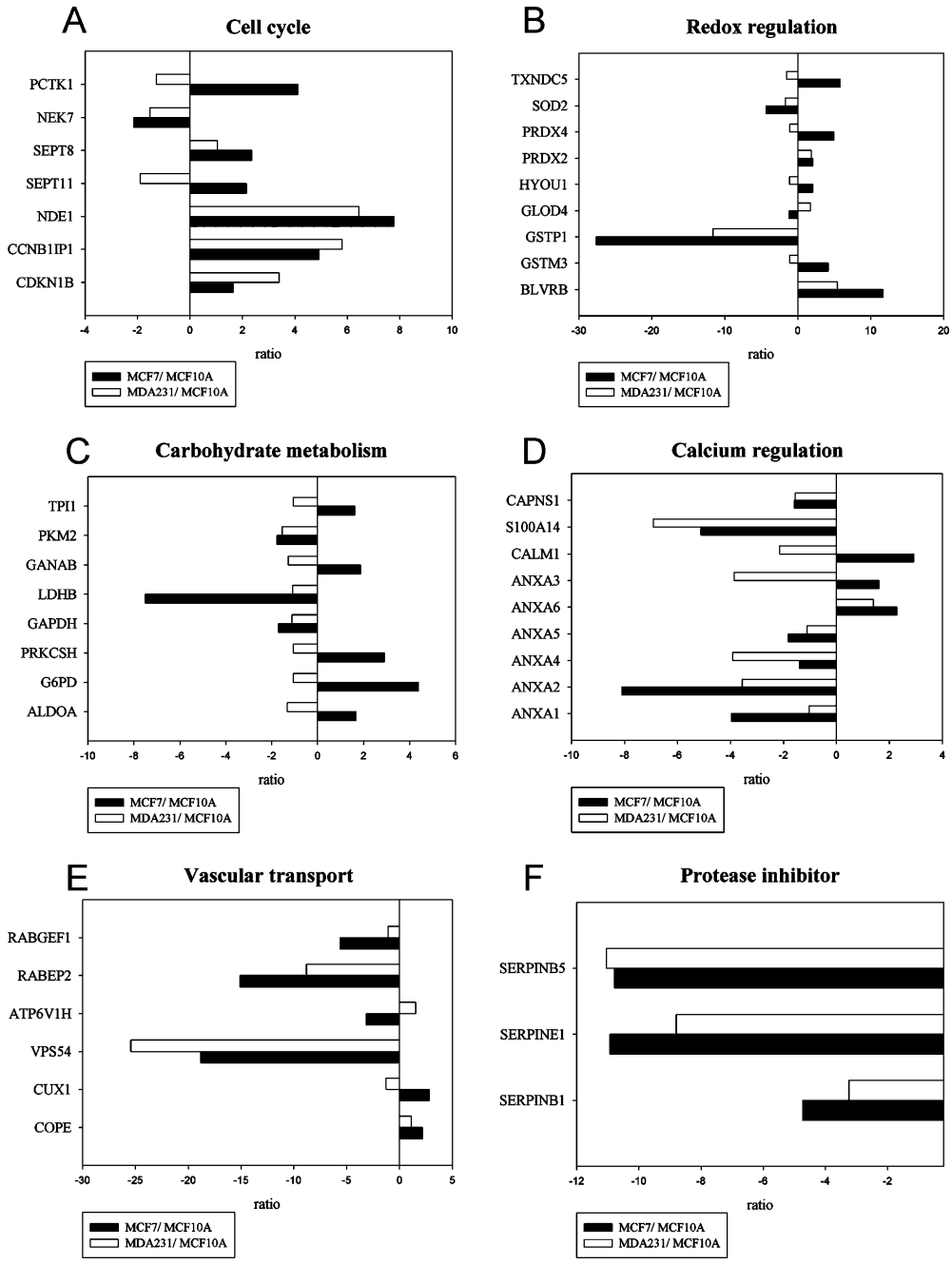

Figure 8. Expression profiles for proteins potentially contributing to (A) cell cycle, (B) redox regulation, (C) carbohydrate metabolism, (D) calcium regulation, (E) vascular transport, (F) protease inhibition in comparing MCF-7 and MDA-MB-231 with MCF-10A. White bars represent fold change in protein expression in MDA-MB-231 versus MCF-10A. Black bars represent fold change in protein expression in MCF-7 versus MCF-10A. The vertical axis indicates the identified proteins; the horizontal axis indicates the fold change in protein expression. Additional details for each protein can be found in Tables 1 and 2.

Recently, Neal et al. reported that the 14-3-3 zeta showed a significant up-regulation in both MCF-7 cells and MDA-MB-231 cells compared with its expression in MCF-10A cell.[62] However, our proteomics results showed that the expression level of 14-3-3 zeta in both MCF-7 cells and MDA-MB-231 cells was down-regulated by 2.25-fold and 4.28-fold in comparison to its expression in MCF-10A in the secretomic experiment, respectively. Another independent experiment to analyze the total cell proteins shows that the 14-3-3 zeta expression in MCF-7 cells and MDA-MB-231 cells showed a 1.95-fold down-regulation and 1.13-fold up-regulation, respectively, from expression in MCF-10A. We compared the cell growth conditions and found the MCF-7 cell culture medium was DMEM/F12 with 8% FBS in Neal's study, but the MCF-7 was cultured in DMEM with 10% FBS in our proteomics study. In addition, although this study and Neal's study use the same growth medium for MCF-10A cells, our culture medium did not contain 0.1 μg/mL cholera toxin. These different culture conditions might account for the different experimental results. Importantly, our secretomic experiments were performed after the cells were starved in a serum-free medium for 30 h, while Neal's work was performed under medium with 8% FBS.

Pauly's study demonstrates that the cell energy status/metabolism status might account for the differentially expressed and subcellular locations of 14-3-3 proteins in hydra.[63] This might offer an alternative clue to explain the different differential expressions of 14-3-3 zeta between the two laboratories. Also, another secretomic study that analyzed the starved conditioned media from MCF-10A, BT474 (noninvasive breast cancer line) and MDA-MB-468 (invasive breast cancer line) was conducted. In this experiment, 14-3-3 zeta was down-regulated in BT474 and up-regulated in MDA-MB-468 when compared with its protein expression level in MCF-10A.[1] Additionally, some cancer studies indicate that 14-3-3 protein shows down-regulation during cancer progression.[64,65] These studies show that the levels of 14-3-3 proteins during breast tumorigenesis were regulated, and further studies are essential before using the proteins as breast cancer biomarkers.

In summary, this study establishes a comprehensive proteomic strategy for putative cancer marker discovery. Using 2D-DIGE and MALDI-TOF MS, we identified a set of candidate protein markers that could serve as new diagnostic and prognostic biomarkers of breast cancer after the subsequent careful characterization of these candidates.

Acknowledgment. This work was supported by grant NSC 97-2311-B-007-005 from the National Science Council, Taiwan, grant CMU-NTHU Joint Research No.98N2443E1 and grant VGHUST98-P5-15 & VGHUST99-P5-22 Veteran General Hospitals University System of Taiwan Joint Research Program. The authors thank Dr. Wun-Shaing Wayne Chang (National Health Research Institute, Taiwan) for providing the MCF-10A cell line.

References

(1) Kulasingam, V.; Diamandis, E. P. *Mol.Cell. Proteomics* **2007**, *6*, 1997.
(2) Jemal, A.; Tiwari, R. C.; Murray, T.; Ghafoor, A.; Samuels, A.; Ward, E.; Feuer, E. J.; Thun, M. J. *CA Cancer J. Clin.* **2004**, *54*, 8.
(3) Nuyten, D. S.; van de Vijver, M. J. *Semin. Radiat. Oncol.* **2008**, *18*, 105.
(4) Morrow, T. *Manag. Care* **2007**, *16*, 51.
(5) Lee, W. Y.; Huang, S. C.; Tzeng, C. C.; Chang, T. L.; Hsu, K. F. *Nutr.Cancer* **2007**, *58*, 239.
(6) Kulasingam, V.; Diamandis, E. P. *Mol. Cell. Proteomics* **2007**, *6*, 1997.
(7) Hondermarck, H.; Dolle, L.; Yazidi-Belkoura, I.; Vercoutter-Edouart, A. S.; Adriaenssens, E.; Lemoine, J. *J. Mammary Gland Biol. Neoplasia* **2002**, *7*, 395.
(8) Sorlie, T.; Tibshirani, R.; Parker, J.; Hastie, T.; Marron, J. S.; Nobel, A.; Deng, S.; Johnsen, H.; Pesich, R.; Geisler, S.; Demeter, J.; Perou, C. M.; Lonning, P. E.; Brown, P. O.; Borresen-Dale, A. L.; Botstein, D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 8418.
(9) Nagaraja, G. M.; Othman, M.; Fox, B. P.; Alsaber, R.; Pellegrino, C. M.; Zeng, Y.; Khanna, R.; Tamburini, P.; Swaroop, A.; Kandpal, R. P. *Oncogene* **2006**, *25*, 2328.
(10) Pawlik, T. M.; Hawke, D. H.; Liu, Y.; Krishnamurthy, S.; Fritsche, H.; Hunt, K. K.; Kuerer, H. M. *BMC Cancer* **2006**, *6*, 68.
(11) Pucci-Minafra, I.; Fontana, S.; Cancemi, P.; Alaimo, G.; Minafra, S. *Ann. N.Y. Acad. Sci.* **2002**, *963*, 122.
(12) Varnum, S. M.; Covington, C. C.; Woodbury, R. L.; Petritis, K.; Kangas, L. J.; Abdullah, M. S.; Pounds, J. G.; Smith, R. D.; Zangar, R. C. *Breast Cancer Res. Treat.* **2003**, *80*, 87.
(13) Mbeunkui, F.; Fodstad, O.; Pannell, L. K. *J. Proteome Res.* **2006**, *5*, 899.
(14) Xue, H.; Lu, B.; Lai, M. *J. Transl. Med.* **2008**, *6*, 52.
(15) Timms, J. F.; Cramer, R. *Proteomics* **2008**, *8*, 4886.
(16) Minden, J. *Biotechniques* **2007**, *43*, 739.
(17) Westermeier, R.; Scheibe, B. *Methods Mol. Biol.* **2008**, *424*, 73.
(18) Marouga, R.; David, S.; Hawkins, E. *Anal. Bioanal. Chem.* **2005**, *382*, 669.
(19) Singh, B.; Berry, J. A.; Vincent, L. E.; Lucci, A. *J. Surg. Res.* **2006**, *134*, 44.
(20) Lu, T. J.; Lai, W. Y.; Huang, C. Y.; Hsieh, W. J.; Yu, J. S.; Hsieh, Y. J.; Chang, W. T.; Leu, T. H.; Chang, W. C.; Chuang, W. J.; Tang, M. J.; Chen, T. Y.; Lu, T. L.; Lai, M. D. *J. Biol. Chem.* **2006**, *281*, 38405.
(21) Chan, H. L.; Gharbi, S.; Gaffney, P. R.; Cramer, R.; Waterfield, M. D.; Timms, J. F. *Proteomics* **2005**, *5*, 2908.
(22) Gharbi, S.; Gaffney, P.; Yang, A.; Zvelebil, M. J.; Cramer, R.; Waterfield, M. D.; Timms, J. F. *Mol. Cell. Proteomics* **2002**, *1*, 91.
(23) Zhang, Y. G.; DU, J.; Tian, X. X.; Zhong, Y. F.; Fang, W. G. *Chin. Med. J.* **2007**, *120*, 1597.
(24) Mita, K.; Zhang, Z.; Ando, Y.; Toyama, T.; Hamaguchi, M.; Kobayashi, S.; Hayashi, S.; Fujii, Y.; Iwase, H.; Yamashita, H. *Jpn. J. Clin. Oncol.* **2007**, *37*, 575.
(25) Danes, C. G.; Wyszomierski, S. L.; Lu, J.; Neal, C. L.; Yang, W.; Yu, D. *Cancer Res.* **2008**, *68*, 1760.
(26) Cao, Y.; Li, Y.; Edelweiss, M.; Arun, B.; Rosen, D.; Resetkova, E.; Wu, Y.; Liu, J.; Sahin, A.; Albarracin, C. T. *Appl. Immunohistochem. Mol. Morphol.* **2008**, *16*, 530.
(27) Gallo, D.; Jacquot, Y.; Laurent, G.; Leclercq, G. *Mol. Cell. Endocrinol.* **2008**, *291*, 20.
(28) Zweitzig, D. R.; Smirnov, D. A.; Connelly, M. C.; Terstappen, L. W.; O'Hara, S. M.; Moran, E. *Mol. Cell. Biochem.* **2007**, *306*, 255.
(29) Fritzsche, F. R.; Dahl, E.; Pahl, S.; Burkhardt, M.; Luo, J.; Mayordomo, E.; Gansukh, T.; Dankof, A.; Knuechel, R.; Denkert, C.; Winzer, K. J.; Dietel, M.; Kristiansen, G. *Clin. Cancer Res.* **2006**, *12*, 1728.
(30) Jung, E. J.; Moon, H. G.; Cho, B. I.; Jeong, C. Y.; Joo, Y. T.; Lee, Y. J.; Hong, S. C.; Choi, S. K.; Ha, W. S.; Kim, J. W.; Lee, C. W.; Lee, J. S.; Park, S. T. *Int. J. Cancer* **2007**, *120*, 2331.
(31) Fu, X. D.; Giretti, M. S.; Baldacci, C.; Garibaldi, S.; Flamini, M.; Sanchez, A. M.; Gadducci, A.; Genazzani, A. R.; Simoncini, T. *PLoS One* **2008**, *3*, e2790.
(32) Harding, M. W.; Handschumacher, R. E. *Transplantation* **1988**, *46*, 29S.
(33) Aitken, A. *Semin. Cancer Biol.* **2006**, *16*, 162.
(34) Fujii, J.; Ikeda, Y. *Redox. Rep.* **2002**, *7*, 123.
(35) Feldner, J. C.; Brandt, B. H. *Exp. Cell Res.* **2002**, *272*, 93.
(36) Liaudet-Coopman, E.; Beaujouin, M.; Derocq, D.; Garcia, M.; Glondu-Lassis, M.; Laurent-Matha, V.; Prebois, C.; Rochefort, H.; Vignon, F. *Cancer Lett.* **2006**, *237*, 167.
(37) Sharma, M. C.; Sharma, M. *Curr. Pharm. Des.* **2007**, *13*, 3568.
(38) Fu, X.; Wang, P.; Zhu, B. T. *J. Steroid Biochem. Mol. Biol.* **2008**, *112*, 127.
(39) Kawai, H.; Li, H.; Avraham, S.; Jiang, S.; Avraham, H. K. *Int. J. Cancer* **2003**, *107*, 353.
(40) Xue, H.; Lu, B.; Lai, M. *J. Transl. Med.* **2008**, *6*, 52.
(41) Yamaguchi, N.; Yamamura, Y.; Koyama, K.; Ohtsuji, E.; Imanishi, J.; Ashihara, T. *Cancer Res.* **1990**, *50*, 7008.
(42) Inoue, Y.; Kawamoto, S.; Shoji, M.; Hashizume, S.; Teruya, K.; Katakura, Y.; Shirahata, S. *Cytotechnology* **2000**, *33*, 21.
(43) Jacobs, J. M.; Mottaz, H. M.; Yu, L. R.; Anderson, D. J.; Moore, R. J.; Chen, W. N.; Auberry, K. J.; Strittmatter, E. F.; Monroe, M. E.; Thrall, B. D.; Camp, D. G.; Smith, R. D. *J. Proteome Res.* **2004**, *3*, 68.
(44) Dictor, M.; Ehinger, M.; Mertens, F.; Akervall, J.; Wennerberg, J. *Am. J. Clin. Pathol.* **1999**, *112*, S40−S52.
(45) Noh, D. Y.; Ahn, S. J.; Lee, R. A.; Kim, S. W.; Park, I. A.; Chae, H. Z. *Anticancer Res.* **2001**, *21*, 2085.
(46) Lopez-Lazaro, M. *Anticancer Agents Med. Chem.* **2008**, *8*, 305.
(47) Gatenby, R. A.; Gillies, R. J. *Int. J. Biochem. Cell Biol.* **2007**, *39*, 1358.
(48) Rivenzon-Segal, D.; Boldin-Adamsky, S.; Seger, D.; Seger, R.; Degani, H. *Int. J. Cancer* **2003**, *107*, 177.
(49) Ang, E. Z.; Nguyen, H. T.; Sim, H. L.; Putti, T. C.; Lim, L. H. *Mol. Cancer Res.* **2009**, *7*, 266.
(50) Cao, Y.; Li, Y.; Edelweiss, M.; Arun, B.; Rosen, D.; Resetkova, E.; Wu, Y.; Liu, J.; Sahin, A.; Albarracin, C. T. *Appl. Immunohistochem. Mol. Morphol.* **2008**, *16*, 530.
(51) Shen, D.; Nooraie, F.; Elshimali, Y.; Lonsberry, V.; He, J.; Bose, S.; Chia, D.; Seligson, D.; Chang, H. R.; Goodglick, L. *Hum. Pathol.* **2006**, *37*, 1583.
(52) Shen, D.; Chang, H. R.; Chen, Z.; He, J.; Lonsberry, V.; Elshimali, Y.; Chia, D.; Seligson, D.; Goodglick, L.; Nelson, S. F.; Gornbein, J. A. *Biochem. Biophys. Res. Commun.* **2005**, *326*, 218.
(53) Donato, R. *Microsc. Res. Tech.* **2003**, *60*, 540.
(54) Ji, J.; Zhao, L.; Wang, X.; Zhou, C.; Ding, F.; Su, L.; Zhang, C.; Mao, X.; Wu, M.; Liu, Z. *J. Cancer Res. Clin. Oncol.* **2004**, *130*, 480.
(55) Hu, H.; Milstein, M.; Bliss, J. M.; Thai, M.; Malhotra, G.; Huynh, L. C.; Colicelli, J. *Mol. Cell. Biol.* **2008**, *28*, 1573.
(56) Sager, R.; Sheng, S.; Pemberton, P.; Hendrix, M. J. *Adv. Exp. Med. Biol.* **1997**, *425*, 77.
(57) Webber, B. A.; Lawson, D.; Cohen, C. *Appl. Immunohistochem. Mol. Morphol.* **2008**, *16*, 19.
(58) Rochefort, H. *Bull. Acad. Natl. Med.* **1999**, *183*, 955.

(59) Rochefort, H. C. R. *Seances Soc. Biol. Fil.* **1998**, *192*, 241.
(60) Spitzner, M.; Martins, J. R.; Soria, R. B.; Ousingsawat, J.; Scheidt, K.; Schreiber, R.; Kunzelmann, K. *J. Biol. Chem.* **2008**, *283*, 7421.
(61) Xiao, S.; Li, D.; Zhu, H. Q.; Song, M. G.; Pan, X. R.; Jia, P. M.; Peng, L. L.; Dou, A. X.; Chen, G. Q.; Chen, S. J.; Chen, Z.; Tong, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16448.
(62) Neal, C. L.; Yao, J.; Yang, W.; Zhou, X.; Nguyen, N. T.; Lu, J.; Danes, C. G.; Guo, H.; Lan, K. H.; Ensor, J.; Hittelman, W.; Hung, M. C.; Yu, D. *Cancer Res.* **2009**, *69*, 3425.
(63) Pauly, B.; Stiening, B.; Schade, M.; Alexandrova, O.; Zoubek, R.; David, C. N.; Bottger, A. *Exp. Cell Res.* **2003**, *285*, 15.
(64) Moreira, J. M.; Ohlsson, G.; Rank, F. E.; Celis, J. E. *Mol. Cell. Proteomics* **2005**, *4*, 555.
(65) Yang, H.; Zhang, Y.; Zhao, R.; Wen, Y. Y.; Fournier, K.; Wu, H. B.; Yang, H. Y.; Diaz, J.; Laronga, C.; Lee, M. H. *Oncogene* **2006**, *25*, 4585.

PR900825T