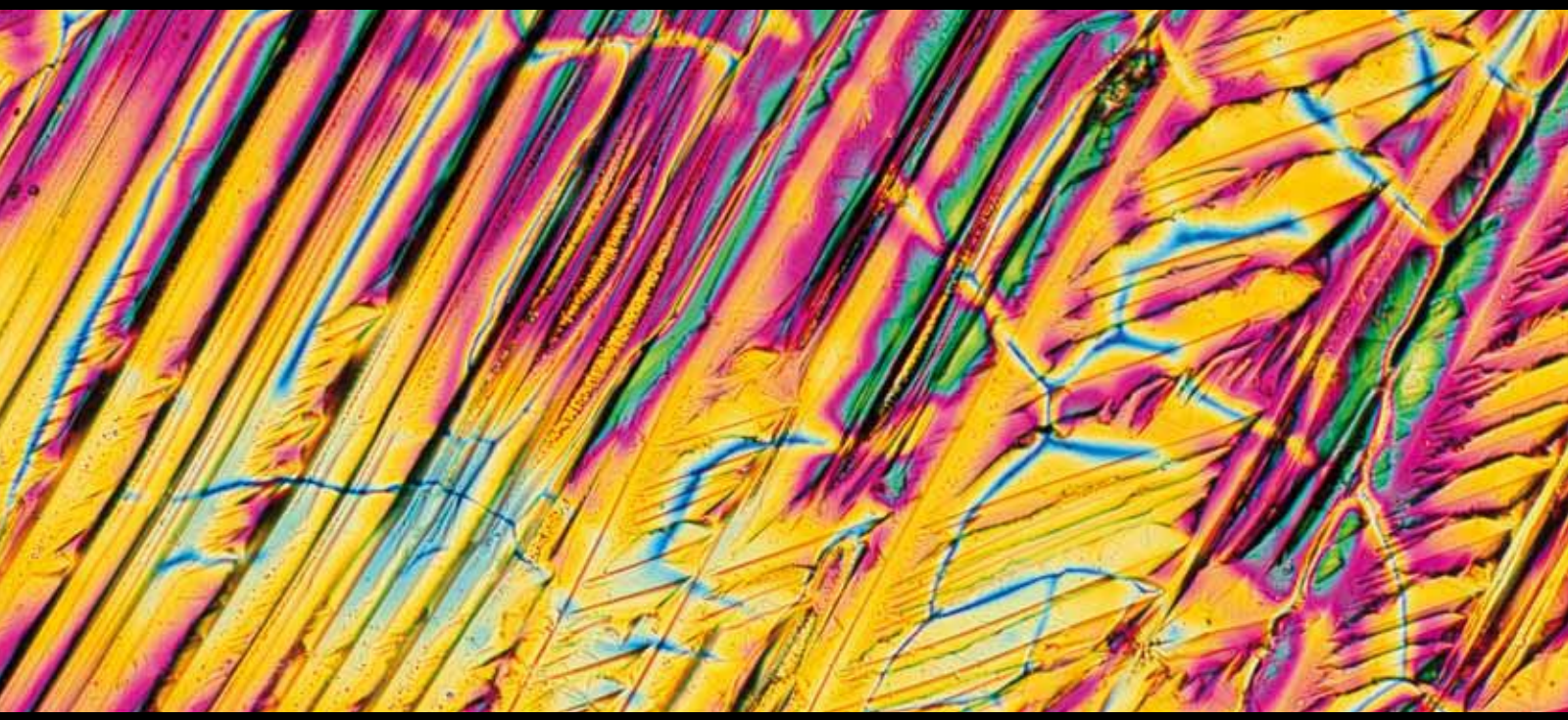
International Journal of Proteomics

Proteomics-Based Disease Biomarkers

Guest Editors: David E. Misek, Tadashi Kondo, and Mark W. Duncan



Proteomics-Based Disease Biomarkers

International Journal of Proteomics

Proteomics-Based Disease Biomarkers

Guest Editors: David E. Misek, Tadashi Kondo,
and Mark W. Duncan

Copyright © 2011 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in volume 2011 of "International Journal of Proteomics." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Ruth Hogue Angeletti, USA
A. I. Archakov, Russia
Christoph H. Borchers, Canada
Robert Bowser, USA
Mark Chance, USA
Yen-Chung Chang, Taiwan
Jen-Fu Chiu, Hong Kong
William C. Cho, Hong Kong
Richard Christopherson, Australia
Juliet A. Gerrard, New Zealand
Michael Hippler, Germany
Christian Huck, Austria
Djuro Josic, USA
S. Komatsu, Japan
Hana Kovarova, Czech Republic
Thallapuranam K. Kumar, USA
Ho Jeong Kwon, Republic of Korea
Andrew J. Link, USA
Kazuyuki Nakamura, Japan
Peter Nilsson, Sweden
Bernd Rehm, New Zealand
Peter Selby, UK
David Sheehan, Ireland
Louisa Braal Tabatabai, USA
Visith Thongboonkerd, Thailand
Vladimir Uversky, USA
Uwe Völker, Germany
Etienne Waelkens, Belgium
Shao Q. Yao, Singapore
Rong Zeng, China
Yaoqi Zhou, USA

Contents

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 894618, 2 pages
doi:10.1155/2011/894618

Editorial

Proteomics-Based Disease Biomarkers

David E. Misek,[1] Tadashi Kondo,[2] and Mark W. Duncan[3,4]

[1] *Department of Surgery, University of Michigan, Ann Arbor, MI 48109, USA*
[2] *Division of Pharmacoproteomics, National Cancer Center Research Institute, Tokyo 104-0045, Japan*
[3] *Division of Endocrinology, Metabolism and Diabetes, School of Medicine, University of Colorado, Aurora, CO 80045, USA*
[4] *Obesity Research Center, College of Medicine, King Saud University, Riyadh, Saudi Arabia*

Correspondence should be addressed to David E. Misek, dmisek@umich.edu

Received 4 September 2011; Accepted 4 September 2011

Copyright © 2011 David E. Misek et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sequencing of the human genome has greatly impacted the proteomics-based analysis of disease by providing a framework for understanding the proteome of diseased cells, tissues, and biological fluids. Consequently, there is a growing interest in applying proteomics technologies to define protein pathways involved in various diseases, to identify new biomarkers that correlate with diseases, ideally in their early stages, and to accelerate the development of new therapeutic targets. However, disease-related proteomics applications require that we improve our ability to separate and characterize the components of complex protein mixtures in such a way as to boost both throughput and sensitivity. In response to these demands, the proteomics technologies have been improved markedly over recent years. Today, proteomics, in all its various forms, is proving to be invaluable to our understanding of the biochemistry of health and disease and will likely play a central role in the evolution of personalized medicine. In this special issue, we include reports of novel research findings together with several reviews that highlight advances in key areas.

The first two papers of this special issue focus on lung cancer. The first paper, by H. C. Gong et al., addresses the profiling of receptor tyrosine kinase pathway activation and the role of key genetic mutations in human lung tumor cell lines and human lung tumors. The authors defined molecular pathways which may assist in development of targeted lung tumor therapies. Within the second paper, Q. Zhang et al. used proteomic profiling to delineate expression and subcellular localization of multiple forms of aldehyde dehydrogenase in lung adenocarcinoma cell lines. The next two papers focus on pancreatic cancer. The third paper, by R. S. Kwon and D. M. Simeone, reviews the use of protein-based biomarkers for the diagnosis of cystic tumors of the pancreas. The fourth paper, by M. Abulaizi et al., utilizes a three-step proteomic protocol (immunodepletion of abundant serum proteins, followed by fractionation by RP-HPLC and further separation by 2D-PAGE) to discover candidate early detection biomarkers of pancreatic cancer.

The next two papers focus on breast cancer, with the fifth paper, by D. E. Misek and E. H. Kim, reviewing the development of protein biomarkers for the early detection of breast cancer. The sixth paper, by J. He et al., addresses LC-MS/MS identification of protein biosignatures in breast tumors, as protein-based markers that correctly classify tumor subtypes and predict therapeutic response would be of great clinical utility in guiding patient treatment. The next two papers are both by M. S. Sabel et al., and focus on melanoma. The seventh paper reviews the use of proteomics for the discovery of new prognostic and predictive biomarkers. The eighth paper explores the clinical utility of serum autoantibodies that were detected in melanoma patients. The investigators profiled serum antibodies against melanoma-associated antigens to identify those that may predict nodal positivity, a widely accepted index of metastatic disease.

The ninth paper, by E. H. Kim and D. E. Misek, reviews the use of glycoproteomics to identify cancer biomarkers. The tenth paper, by A. Vivekanandan-Giri et al., utilized glycoproteomics to identify novel urinary glycoprotein biomarkers of chronic kidney disease. The issue concludes with two papers that report on novel approaches and related considerations. The eleventh paper, by I. Kiyokawa et al., describes the development of a new surface coating for urinary collection tubes that minimizes the amount of urine protein adsorption onto the walls of the collection tube.

Within the final paper of this special issue, T. Hagiwara et al. examine the utility of a solid-phase hexapeptide ligand library in combination with conventional plasma proteomics modalities for comprehensive profiling of intact plasma proteins.

David E. Misek
Tadashi Kondo
Mark W. Duncan

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 215496, 13 pages
doi:10.1155/2011/215496

Research Article

Signatures of Drug Sensitivity in Nonsmall Cell Lung Cancer

Hua C. Gong,[1] Sean Wang,[1] Gary Mayer,[1] Guoan Chen,[2] Glen Leesman,[1] Sharat Singh,[1] and David G. Beer[3]

[1] *Department of Research and Development, Prometheus Inc. 9410, Carroll Park Drive, San Diego, CA 92121, USA*
[2] *Department of Surgery, School of Medicine, University of Michigan, 6304 Cancer Center, Ann Arbor, MI 48109, USA*
[3] *Cancer Center, University of Michigan, Room 6304, Ann Arbor, MI 48109, USA*

Correspondence should be addressed to
Sharat Singh, ssingh@prometheuslabs.com and David G. Beer, dgbeer@umich.edu

Received 27 April 2011; Accepted 1 June 2011

Academic Editor: David E. Misek

Copyright © 2011 Hua C. Gong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We profiled receptor tyrosine kinase pathway activation and key gene mutations in eight human lung tumor cell lines and 50 human lung tumor tissue samples to define molecular pathways. A panel of eight kinase inhibitors was used to determine whether blocking pathway activation affected the tumor cell growth. The HER1 pathway in HER1 mutant cell lines HCC827 and H1975 were found to be highly activated and sensitive to HER1 inhibition. H1993 is a c-MET amplified cell line showing c-MET and HER1 pathway activation and responsiveness to c-MET inhibitor treatment. IGF-1R pathway activated H358 and A549 cells are sensitive to IGF-1R inhibition. The downstream PI3K inhibitor, BEZ-235, effectively inhibited tumor cell growth in most of the cell lines tested, except the H1993 and H1650 cells, while the MEK inhibitor PD-325901 was effective in blocking the growth of KRAS mutated cell line H1734 but not H358, A549 and H460. Hierarchical clustering of primary tumor samples with the corresponding tumor cell lines based on their pathway signatures revealed similar profiles for HER1, c-MET and IGF-1R pathway activation and predict potential treatment options for the primary tumors based on the tumor cell lines response to the panel of kinase inhibitors.

1. Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide, resulting in 1.61 million new cases and 1.38 million deaths per year according to the Global cancer statistics estimation in 2011 [1]. Lung cancer is generally classified histologically into two major types, small cell lung cancer (SCLC) and nonsmall cell lung cancer (NSCLC). Approximately 85–90% of lung cancers are NSCLC representing three major subtypes based on tumor cell size, shape, and composition, with adenocarcinoma accounting for 40%, squamous cell lung carcinoma 25–30%, and large-cell lung carcinoma accounting for 10–15% of all lung cancers [2, 3]. Although less than optimal, current conventional treatment for lung cancer consists of surgery for operable candidates and chemotherapy for disease-advanced patients with the mean survival for most advanced lung cancer patients less than one year [4]. During the last decade, considerable progress has been made in the treatment of NSCLC due to the emergence of new targeted therapies specific to the oncogenic tyrosine kinase pathways activated in tumor cells. For example, two epidermal growth factor receptor (HER1) tyrosine kinase inhibitors (TKI), Gefitinib (Iressa) and Erlotinib (Tarceva), have been FDA approved for the treatment of locally advanced or metastatic NSCLC that has failed at least one prior chemotherapy regimen [5, 6]. Other receptor tyrosine kinase (RTK) pathway inhibitors, such as Sunitinib (Sutent), which targets the platelet-derived growth factor receptors and vascular endothelial growth factor receptors, as well as Crizotinib, a hepatocyte growth factor RTK inhibitor, are in advanced clinical trials for NSCLC [7, 8].

The advances made in targeted therapy for NSCLC are based on understanding the mechanism by which mutated genes confer a neoplastic phenotype on tumor cells and how the targeted interruption of these oncogenic pathways leads to clinical response. Thus, analysis of a pathway-focused panel of biomarkers in fresh tumor tissue samples collected

from patients could pave the way for determining if the markers are associated with the optimal clinical therapy and may provide predictive value in identifying responsive patients. In addition, drug combinations targeted against the receptors affecting downstream signaling molecules may overcome pathway activation and drug resistance often seen in NSCLC therapy. Difficulties in predicting efficacy in targeted therapy is due to the limited knowledge of the activated oncogenic pathways in the patient's tumor so that the appropriate inhibitor(s) are not prescribed. Thus, preclinical cellular response profiling of tumor tissue samples has become a cornerstone in the development of novel cancer therapeutics. To this end, we have developed and trademarked a channel enzyme enhanced reaction (CEER) assay methodology to profile some of the major oncogenic pathways activated in cancer cells and have used this assay together with genotyping to characterize the activated oncogenic pathways in eight human NSCLC tumor cell lines as well as 50 fresh-frozen NSCLC samples collected from patients. The aim of this study was to assess the potential to prospectively classify lung cancer patients into different treatment groups based on correlation of pathway activation profiles, gene mutational status, and clinical features between the patient tumor samples and the tumor cell lines. In addition, we evaluated the efficacy of a panel of eight kinase pathway inhibitors to block the pathway activation and proliferation of these eight lung tumor cell lines and used the results to identify treatment options for the 50 lung cancer patients.

2. Materials and Methods

2.1. Human Lung Tumor Cell Lines, Lung Cancer Tissue Samples, and Kinase Inhibitors. Eight NSCLC cell lines, HCC827, H1975, H1734, H1993, H358, H1650, A549, and H460, were selected, and they represent the major NSCLC cancer subtypes, adenocarcinoma and large-cell lung carcinoma. The cell lines were purchased from ATCC (Table 1). Fifty lung adenocarcinomas samples were collected from patients operated on for lung cancer at the University of Michigan. Collection and use of all tissue samples were approved by the Human Subjects Institutional Review Boards of the University of Michigan. The demographic information of the patients is shown in Supplementary Table 1. The primary tumor samples were snap frozen and cryostat-sectioned to identify regions representing >70% tumor cellularity for subsequent pathway analysis. The samples (~2 cubic millimeters in size) were shipped to Prometheus Laboratories on dry ice for analysis. Eight kinase inhibitors representing a diverse panel of potential cancer therapeutics were purchased from Selleck Chemicals (Houston, TX). The collection included specific as well as multiple RTK inhibitors, that is, compounds targeting the cellular kinase pathways: HER1/2/4 (epidermal growth factor receptors) inhibitors (Erlotinib for HER1, Lapatinib for HER1/2, Gefitinib for HER1/2/4, and BIBW-2992 is an irreversible inhibitor for HER1/2); c-MET (hepatocyte growth factor receptor) inhibitor, PF-2341066; IGF-1R (insulin-like growth factor-1 receptor) inhibitor BMS-536924; MEK (mitogen-activated protein kinase kinase) inhibitor, PD-325901, and PI3K (phosphatidylinositol-3-kinase) and mTOR (mammalian target of rapamycin) inhibitor BEZ-235.

2.2. Preparation of Lysates from Cell Lines and Primary Tumor Samples. Tumor cells were cultured in their respective growth medium recommended by ATCC plus 10% fetal bovine serum (FBS). Cells were grown in 35 mm 6-well cell culture plates until reaching 80% confluence. After washing the cells with phosphate buffered saline (PBS) 3 times, the cell culture plate was placed on ice and then the plate was carefully tilted on its side for 10 sec to completely remove all residual media. Then, 150 μL of ice cold lysis buffer was added to each plate and the plate was then left on ice for 5 min. The lysed cells were scraped off and together with the crude lysate transferred to a 1.5 mL centrifuge tube. The mixture was vortexed in the tube, placed on ice for 15 min and then centrifuged at 14,000 rpm for 15 min at 4°C. The supernatant was transferred to another centrifuge tube and stored at −70°C until analysis. The frozen tumor samples were similarly processed by the addition of 4 volumes of ice-cold lysis buffer per tissue volume and homogenized in a Powergren High Throughput Homogenizer (Fisher Scientific) at a speed setting of 7 for 2 min. The homogenate was transferred to a 1.5 mL centrifuge tube and centrifuged at 14,000 rpm for 15 min, at 4°C. The supernatant from the tumor lysate was harvested and stored at −70°C until analysis.

2.3. Profiling of Signaling Pathways Using the CEER Assay in Tumor Cell Lines and Tissue Samples. The principle of the CEER assay is based on the capture of the target protein by a target-specific antibody printed in two dilutions on the surface of a microarray slide. Measurement of the activation status of the captured target protein is revealed by the formation of a unique immunocomplex, requiring the colocalization of two detecting enzyme-conjugated antibodies on the same target protein captured on the microarray surface as illustrated in Supplemental Figure 1. Formation of this complex is initiated by the binding of the first detecting antibody, which is coupled to glucose oxidase (GO), to an epitope on the captured target protein that is different from the epitope recognized by the capture antibody, followed by the binding of a second detecting antibody, which is coupled to horseradish peroxidase (HRP), to a phosphorylated tyrosine (p-Tyr) residue on the target protein. Upon the addition of glucose, the immobilized GO on the captured target protein produces H_2O_2 and due to the close proximity, the locally generated H_2O_2 is then utilized by the HRP coupled to the p-Tyr-specific second detecting antibody to generate a chemical signal that can be amplified with biotinyl-tyramide. The sensitivity and specificity for the detection of the phosphorylated target protein are greatly enhanced by this collaborative reaction and amplification process, which is mediated by the simultaneous binding of three different antibodies on the same target protein [9] (for details of the CEER assay method to profile the tumor cell lines and tissue samples see Supplemental Methods).

Table 1: Clinical features of the eight lung tumor cell lines.

Name	ATCC Cat. no.	Type	Source
HCC827	CRL-2868	Adenocarcinoma	Primary, lung epithelial
H1975	CRL-5908	Adenocarcinoma	Primary, lung epithelial
H1734	CRL-5891	Adenocarcinoma	Primary, lung epithelial
H1993	CRL-5909	Adenocarcinoma	Metastatic, lymph node
H358	CRL-5807	Bronchioalveolar carcinoma	Primary, lung epithelial
H1650	CRL-5883	Bronchioalveolar carcinoma	Metastatic, pleural effusion
A549	CCL-185	Adenocarcinoma	Primary, lung epithelial
H460	HTB-177	Large-cell carcinoma	Metastatic, pleural effusion

2.4. Inhibition of Activated Signaling Pathways in Tumor Cell Lines by Kinase Inhibitors. The tumor cells were cultured in their respective growth medium with 10% FBS in 35 mm 6-well cell culture plates until they reached ~80% confluence. The cells were then starved overnight in serum-free medium, followed by a 4-hour treatment with various concentrations of the kinase inhibitor. Afterwards, cell lysates were prepared from the treated cells as before and aliquots of the lysates subjected to the CEER assay.

2.5. Inhibition of Tumor Cell Line Growth by Kinase Inhibitors. The tumor cells were seeded into 96-well cell culture plates and maintained in culture for 24 hours. After washing, the cultured cells were incubated in their respective medium containing 5% FBS and various concentrations of the indicated inhibitor for 48 hr. Determination of tumor cell growth inhibition was performed by adding 100 μL of the combined Cell Titer-Glo Buffer and Cell Titer-Glo Substrate Labeling Reagent (Promega) to each well of the plates, followed by incubation at room temperature for 10 min to stabilize the luminescence. The luminescent signal from the cell samples was detected by using an M5 microtiter plate reader. For studies involving treatment with more than one inhibitor, the selected inhibitor that showed more than 25% inhibition of tumor cell growth at 10 μM concentration when treated individually was further tested in combination treatment with another inhibitor. A 5 μM concentration of each inhibitor was combined to make a 10 μM dose, and the same half log dilution was made as in the single drug treatment for adding to the cells. Tumor cells were treated for 48 hr and cell viability was measured as in the single inhibitor treatment.

2.6. Anchorage-Independent Inhibition of Tumor Cell Line Growth by Kinase Inhibitors. A single cell suspension of 3000 cells from each of the eight tumor cell lines in 1 mL mixture of 1.2% Agarose (Seaplaque; FMC, Rockland, ME) in DMEM (Life Technologies, Carlsbad, CA) plus 10% FBS was added on top of 1% soft agarose that had been allowed to gel previously in the wells of a 35 mm 6-well cell culture plate. The plates were kept at 4°C for 2 hr to solidify the cell-containing layer. The plates were then incubated at 37°C in a CO_2 incubator with 2 mL of medium containing various concentrations of the kinase inhibitor. The medium with the inhibitor was changed every 3 days. After 2 weeks, cell colonies larger than 10 cells were scored under a Nikon inverted-phase microscope [10].

2.7. Genotyping of Lung Tumor Tissue Samples. Genomic DNA was isolated from human tumor tissue samples using the DNeasy kit (Qiagen, Valencia, CA). Primers and probes for all of the measured SNPs were obtained from the ABI TaqMan SNP Genotyping Assay (Applied Biosystems), using the Assay-by-Design service for which we provided the sequences, or the Assay-on-Demand service when the assays were already designed by Applied Biosystems. Reactions were performed in 5 μL volume and contained 10 ng DNA, 1x TaqMan Universal Mastermix (Applied Biosystems), 200 nM of each probe and 900 nM of each primer. Cycling conditions on the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) were 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C and 1 min at 60°C. After cycling, the endpoint fluorescence was measured and the amplified sequences determined by DNA sequencing analysis. Alleles were assigned using the SDS 2.1 software (Applied Biosystems).

2.8. Lung Tumor Tissue Sample Clustering Analysis and Heat Map Generation. Hierarchical clustering analysis was performed on the 50 lung tumor tissue samples to explore whether the pathway activation profiles determined by the CEER assay and the gene mutational analysis done for these samples could segregate them into distinct subsets that are similar to the pathway activation and mutational signatures of the tumor cell lines. The general construction of a hierarchical agglomerative classification was achieved by using an algorithm to find the two closest objects and merge them into a cluster, and then find and merge the next two closest points, where a point is either an individual object or a cluster of objects. A heat map of one-dimensional hierarchical clustering result was generated in the analysis to demonstrate the sample clustering structure based on pathway activation signatures and mutational status.

3. Results

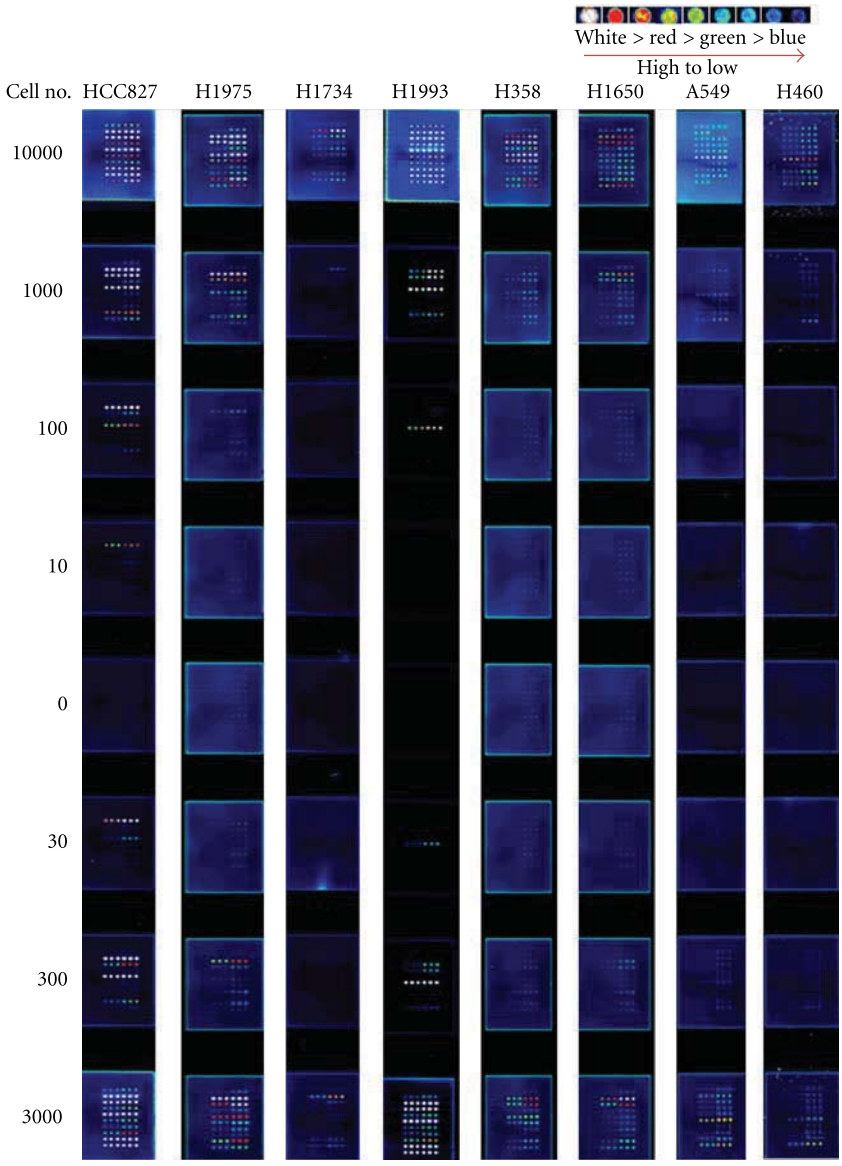
3.1. Profiling of the Activated Kinase Pathways in Tumor Cell Lines. The activated cell signaling pathways for eight lung tumor cell lines were profiled using the CEER assay. The assay measured the activated (p-Tyr) HER1, HER2, HER3,

c-MET, IGF-1R, c-Kit, PI3K, and SHC levels in the cells based on the number of cells being assayed. The tumor cells were cultured in presence of 10% FBS and harvested at about 80% confluence for preparation of cell lysates and signaling pathway profiling. Serial dilutions of tumor cell lysates equivalent to 10–10,000 cells were assayed, and the raw data captured on the slides are shown in Figure 1(a) and a graphic representative of the data is shown in Figure 1(b). As seen in Figure 1(b), the dose-response curves representing the level of activation in each RTK pathway is inversely proportional to the number of cells being assayed. The more activated the pathway, the less number of cells are needed to generate the maximal signal. Thus, the relative activation of the activated pathways can be determined based on the EC_{50} value of the number of cells being assayed in each pathway activation curve (see Supplemental Table 2).

As shown in Figure 1(b), each of the eight tumor cell lines exhibited a distinct RTK activation pattern. The lung adenocarcinoma cell line, HCC827, exhibited the greatest number of activated RTK pathways, with HER1, c-MET, and HER2 being highly activated, PI3K being moderately activated, and HER3 as well as IGF-1R being lowly activated. The other adenocarcinoma cell line H1975 showed only moderate activation of HER2, c-MET, and SHC pathways and a low activation of the IGF-1R pathway. The remaining adenocarcinoma cell line H1734 exhibited a moderate activation of the HER1 pathway and a very low activation of the HER2 and c-MET pathways. By contrast, the adenocarcinoma cell line H1993 from a metastatic tumor showed a very potent activation of the c-MET pathway, with also a high activation of the HER2 and SHC pathways, and a moderate activation of the HER3 and HER1 pathways. Both primary and metastatic bronchioalveolar carcinoma cell lines H358 and H1650 showed a moderate activation of the HER1 and HER2 pathways, with H358 also exhibited a moderate activation of the c-MET and IGF-1R pathways. The large-cell carcinoma cell line A549 exhibited a moderate activation of the HER1 and IGF-1R pathways with a low activation of the c-MET and HER2 pathways, whereas the metastatic large-cell carcinoma cell line H460 showed only a very low activation of the IGF-1R pathway.

3.2. Inhibition of Activated Signaling Pathways in Tumor Cell Lines by Kinase Inhibitors. Profiling of the eight tumor cell lines showed that the HER1 and HER2 pathways are highly activated in the HCC827 cells and thus treatment of these cells with an irreversible HER1/2 inhibitor BIBW-2992 should be able to block the activation of these pathways. Indeed, the data presented in Figure 2, showed that a potent dose-dependent inhibition of the HER1 and HER2 pathways in the HCC827 cells was observed by treatment with BIBW-2992. Other HER1 and/or HER2 pathway-activated cell lines, H1975 and H1650, were likewise had these pathway activations blocked by the treatment with BIBW-2992. What is remarkable is that the H1975 cell line harbors the T790M and L858R mutations in the HER1 gene (information obtained from the Sanger Institute website), which confer resistance to HER1 kinase inhibitors, Gefitinib and Erlotinib, responded to the irreversible HER1 kinase inhibitor BIBW-2992. Two other additional HER1 and/or HER2 pathway-activated cell lines, H358 and H1734, also had their activation blocked by the HER1/2 kinase inhibitors, Gefitinib and Lapatinib, respectively. In the c-MET amplified cell line H1993, activation of this pathway was blocked by the treatment with PF-2341066, a c-MET kinase inhibitor. The c-MET inhibitor also inhibited the HER1 signaling pathway in this cell line, most likely due to crosstalk between the HER1 and c-MET pathways. Treatment with the IGF-1R kinase inhibitor BMS-536924 was able to block the activated IGF-1R pathway exhibited by the A549 and H460 cell lines.

3.3. Inhibition of Tumor Cell Line Proliferation by Kinase Inhibitors. Since one purpose of this study was to correlate the activated RTK pathways and gene mutations found in the tumor cell lines with the appropriate kinase inhibitors to determine whether treatment with the inhibitors could inhibit the growth of these cells, the eight tumor cell lines were treated with the selected kinase inhibitors and the results shown in Figure 3. As expected, the lung adenocarcinoma cell line HCC827, which exhibited highly activated HER1 and HER2 pathways as well as HER1 mutation, responded exceedingly well to the HER1 inhibitors, Erlotinib and Gefitinib. In addition, two other HER1/2 inhibitors, Lapatinib and BIBW-2992, were able to inhibit the proliferation of this cell line. By contrast, the specific c-MET inhibitor PF-2341066 (Crizotinib) was not able to inhibit the proliferation of this cell line even though it exhibited the activated c-MET pathway. This is not surprising because it has been shown that Crizotinib is effective in treating patients carrying the ALK/EML4 fusion gene and this drug response is not correlated with c-MET amplification. Moreover, inhibition of proliferation by BMS-536924, an IGF-1R inhibitor, in the HCC827 cells could be due to crosstalk between c-MET and IGF-1R pathways. The proliferation of two other HER1 and HER2 pathway-activated adenocarcinoma cell lines, H1734 and H1975, was also inhibited by Erlotinib. Moreover, the irreversible HER1/2 kinase inhibitor, BIBW-2992, was also able to potently inhibit the growth of the H1975 cell line, which harbored the T790M and L858R mutations in the HER1 gene and thus rendering the cells resistant to Gefitinib and Lapatinib treatment. In addition, the IGF-1R inhibitor BMS-536924 was also able to inhibit the growth of this cell line. In regard to the metastatic adenocarcinoma cell line H1993, which exhibited a potently activated c-MET pathway, the c-MET inhibitor PF-2341066 was able to block its proliferation very effectively. What is more, the downstream MEK inhibitor PD-325901 was also able to block the proliferation of the H1993 cells whereas the irreversible HER1/2 inhibitor BIBW-2992 and the PI3K inhibitor BEZ-235 were able to block the proliferation of this cell line but only weakly. Proliferation of the carcinoma cell line H358, which exhibited a moderate activation of the HER1, HER2, and IGF-1R pathways, was weakly blocked by the HER1/2 inhibitors, Lapatinib and Gefitinib, but the IGF-1R inhibitor BMS-536924 was able to inhibit its proliferation more effectively. Surprisingly, growth of the H358 cell line



(a)

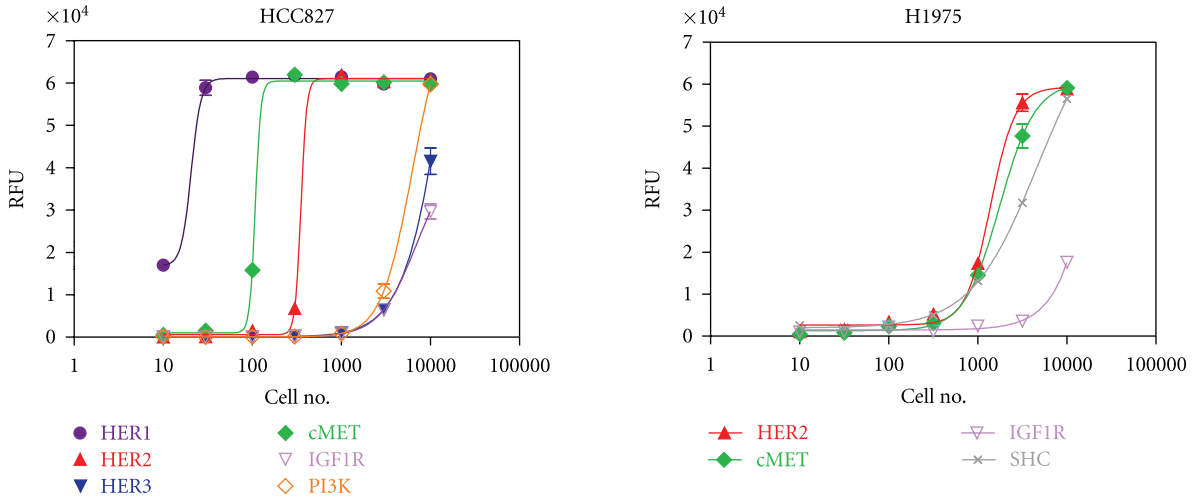
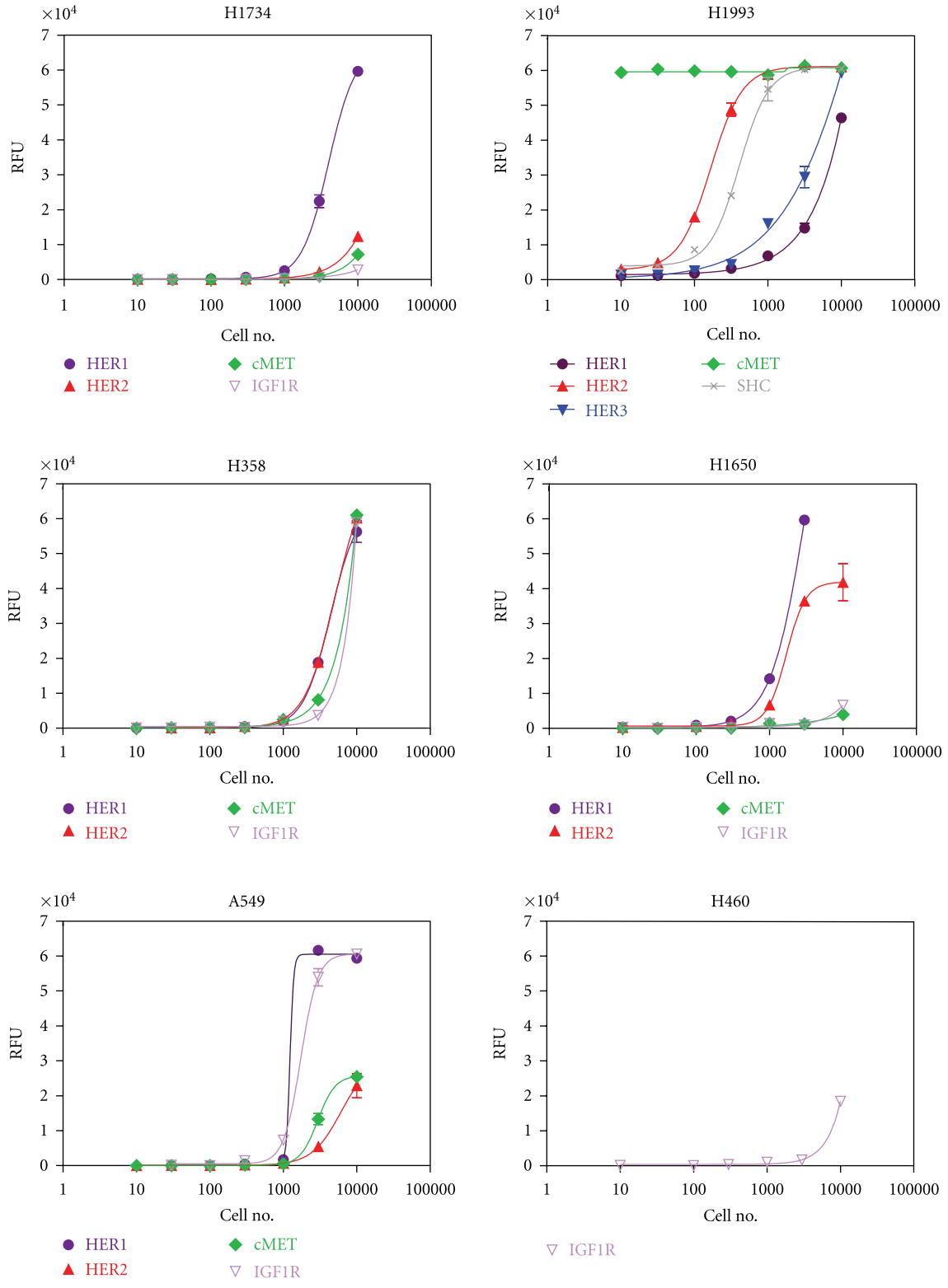


Figure 1: Continued.



(b)

Figure 1: Profiling of the activated (p-Tyr) RTK signaling pathways in the lung tumor cell lines. (a) Slide images of the CEER assay obtained from the eight tumor cell lines (please see the supplemental data and methods for experimental details available online at doi: 10.1155/2011/215496). (b) Graphical representation of the activated (p-Tyr) RTK pathways in the lung tumor cell lines profiled by the CEER assay. The cells were cultured in the presence of 10% FBS in their respective medium until 80% confluence and cell lysates were prepared in lysis buffer. The activation signal was determined from the harvested lysate. In each cell line the activated pathways are shown in RFU (relative fluorescent units) based on the lysate obtained from the number of cells being assayed.

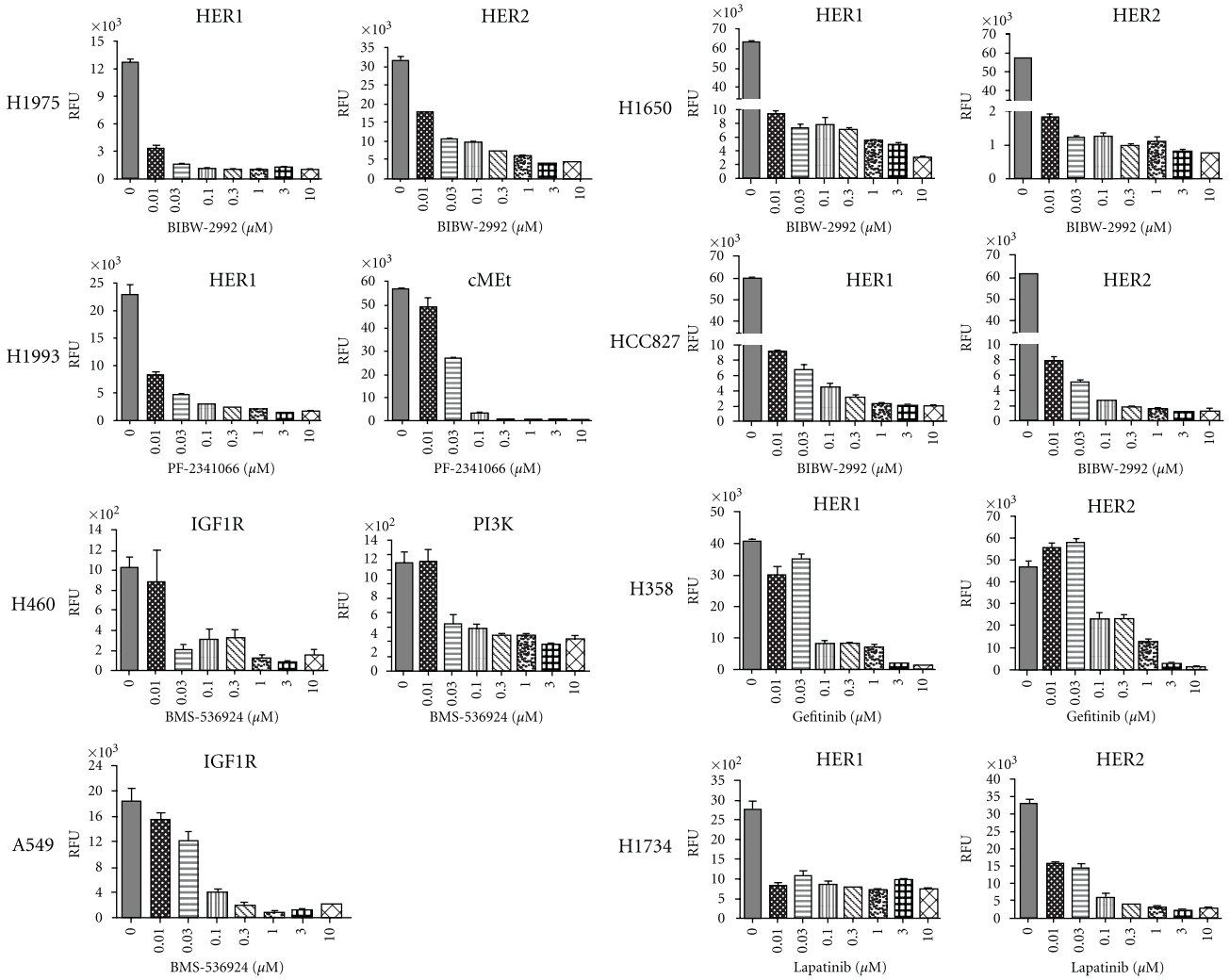


Figure 2: Inhibition of signaling pathway activation in lung tumor cell lines by kinase inhibitors. Lung tumor cells were cultured in 10% FBS until reaching ~80% confluence and then the cells were starved in serum-free medium for overnight, followed by 4-hour treatment with the inhibitors. Cell lysates were then prepared and used for determination of the pathway activation signals by the CEER assay.

was not inhibited by the c-MET inhibitor PF-2341066, even though it exhibited a moderately activated c-MET pathway. Growth of the metastatic carcinoma cell line H1650, which harbored moderately activated HER1 and HER2 pathways, was inhibited by the irreversible HER1/2 inhibitor BIBW-2292. Proliferation of the large-cell carcinoma cell line A549, which exhibited moderately activated HER1 and IGF-1R pathways, was inhibited by the PI3K inhibitor BEZ-235 and by the IGF-1R inhibitor BMS-536924 and weakly by the MEK inhibitor PD-325901. Growth of the remaining metastatic large-cell carcinoma cell line H460, which harbored the PIK3CA gene mutation and an activated IGF-1R pathway, was inhibited by the PI3K inhibitor BEZ-235 but weakly inhibited by the IGF-1R inhibitor BMS-536924.

3.4. Anchorage-Independent Growth Inhibition of Tumor Cell Lines by Kinase Inhibitors. Since anchorage-independent growth is a hallmark of transformed cells, we wanted to ensure that the kinase inhibitors which were able to strongly block tumor cell proliferation in tissue culture were also able to inhibit the same tumor cells' proliferation in an anchorage-independent manner. As seen in Figure 4, Erlotinib, which inhibited the growth of HCC827 cells cultured in anchorage-dependent cell culture plates, was also able to reduce the size of the cell colonies grown in agar plates in an anchorage-independent fashion. Similarly, reduction of cell colony size formation was also observed by the treatment with the irreversible HER1/2 inhibitor BIBW-2992 in H1975 and H1650 cells as seen when these cells were grown in an anchorage-dependent manner. Large colony formation of H1993 cells, whose proliferation in tissue culture was potently blocked by treatment with the c-MET kinase inhibitor PF-2341066, was also inhibited by treatment with the same inhibitor. Proliferation of cell colony size in the H1734 cells was potently blocked by treatment with the downstream MEK inhibitor PD-325901 as seen when these cells were grown in cell culture. Likewise, the IGF-1R kinase inhibitor BMS-536924 was able to reduced

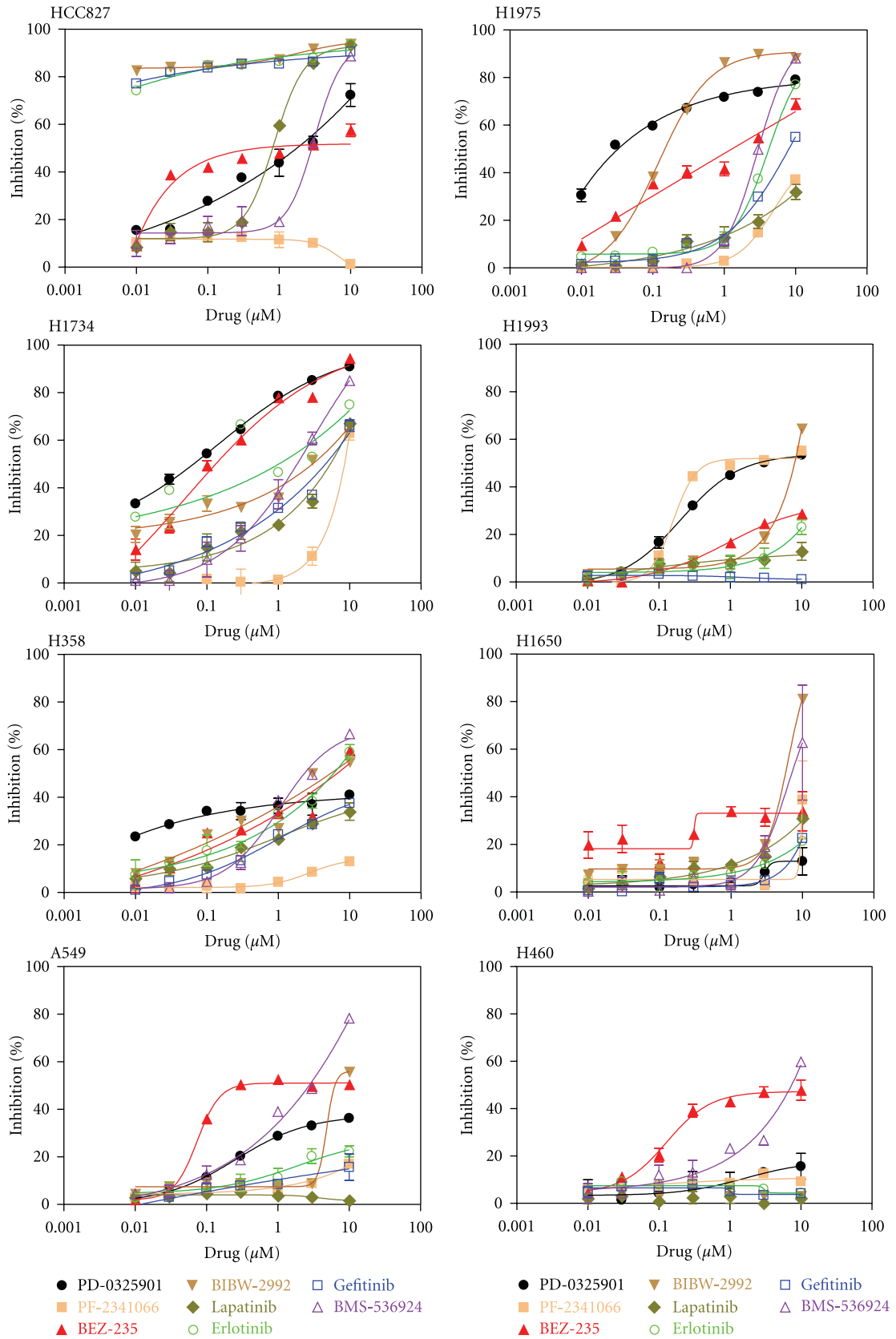


FIGURE 3: Inhibition of lung tumor cell growth by kinase inhibitors. Lung tumor cells were cultured in 5% FBS plus increasing concentrations of the indicated inhibitors, ranging from 0.01–10 μM, for 48 hours. Determination of cell proliferation was performed with the CellTiter-Glo Luminescent Cell Viability Assay.

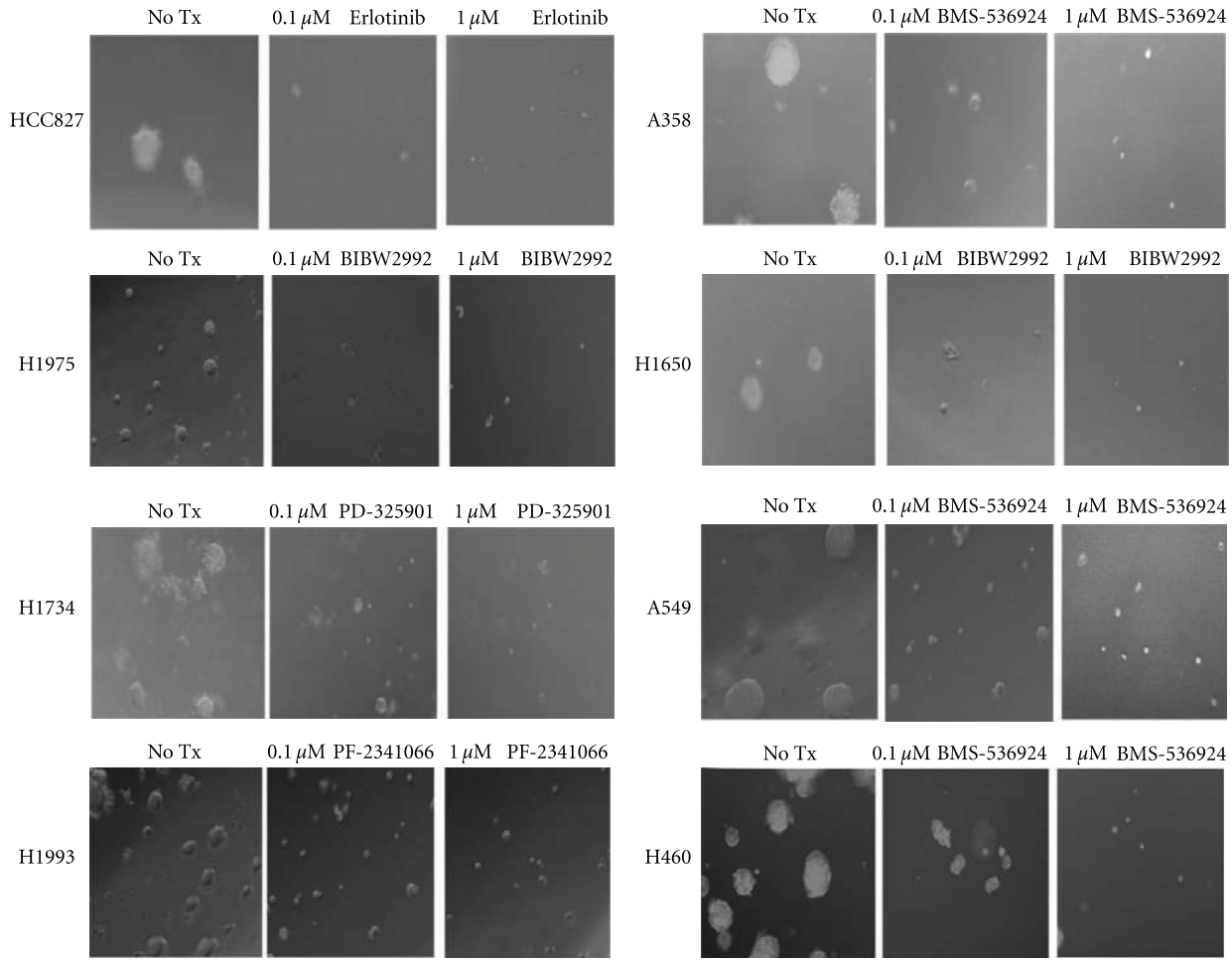


Figure 4: Inhibition of anchorage-independent growth of lung tumor cell lines by selected inhibitors. Each selected cell line was treated with the indicated inhibitor at 0.1 μM and 1 μM concentrations for two weeks and cell colony size formation was scored under the Nikon inverted-phase microscope.

cell colony size formation, dose dependently in the H358, A549 and H460 cells.

3.5. Inhibition of Tumor Cell Line Growth by a Combination of Two Kinase Inhibitors. Tumor cells often rely on signaling from multiple pathways and, hence, treating patients with a single agent can seldom eradicate tumor growth [11]. A common clinical practice is to treat patients with combination therapies. In order to identify therapeutics with synergistic effects, we selected four tumor cell lines out of the eight and treated them with a combination of two kinase inhibitors, one inhibiting the appropriate RTK and the other inhibiting a downstream signaling pathway. As seen in Figure 5, the H1975 cells, whose growth was moderately inhibited by the HER1/2 RTK inhibitor BIBW-2992, but only weakly inhibited by the downstream MEK inhibitor PD-325901 and the downstream PI3K inhibitor BEZ-235, responded more effectively to a combination of BIBW-2992 with either PD-325901 or BEZ-235 with almost 100% growth inhibition of this cell line at 10 μM concentration whereas a combination of the two downstream inhibitors, PD-325901 and BEZ-235, was much less effective. Likewise, combination of PD-325901 or BEZ-235 with the c-MET inhibitor PF-2341066 was more effective in blocking the proliferation of the H1993 cells, which exhibited a potently activated c-MET pathway, then by treating this cell line with the c-MET inhibitor PF-2341066 alone. Interestingly, a combination of both downstream kinase inhibitors, PD-325901 and BEZ-235, was the most effective in blocking the proliferation of this cell line. In the bronchioalveolar cell line H358, which exhibited HER1 and HER2 activation, treatment with the HER1 inhibitor Erlotinib in combination with either one of the two downstream inhibitors, PD-325901 or BEZ-235, showed synergistic inhibition of cell proliferation. A combination of the two downstream inhibitors, PD-0325901 and BEZ-235, was also highly effective. The same phenomenon was observed in the metastatic bronchioalveolar cell line H1650 when it was treated with the irreversible HER1/2 inhibitor BIBW-2992 in combination with one of the downstream inhibitor BEZ-235 or a combination of the two downstream inhibitors, PD-325901 and BEZ-235.

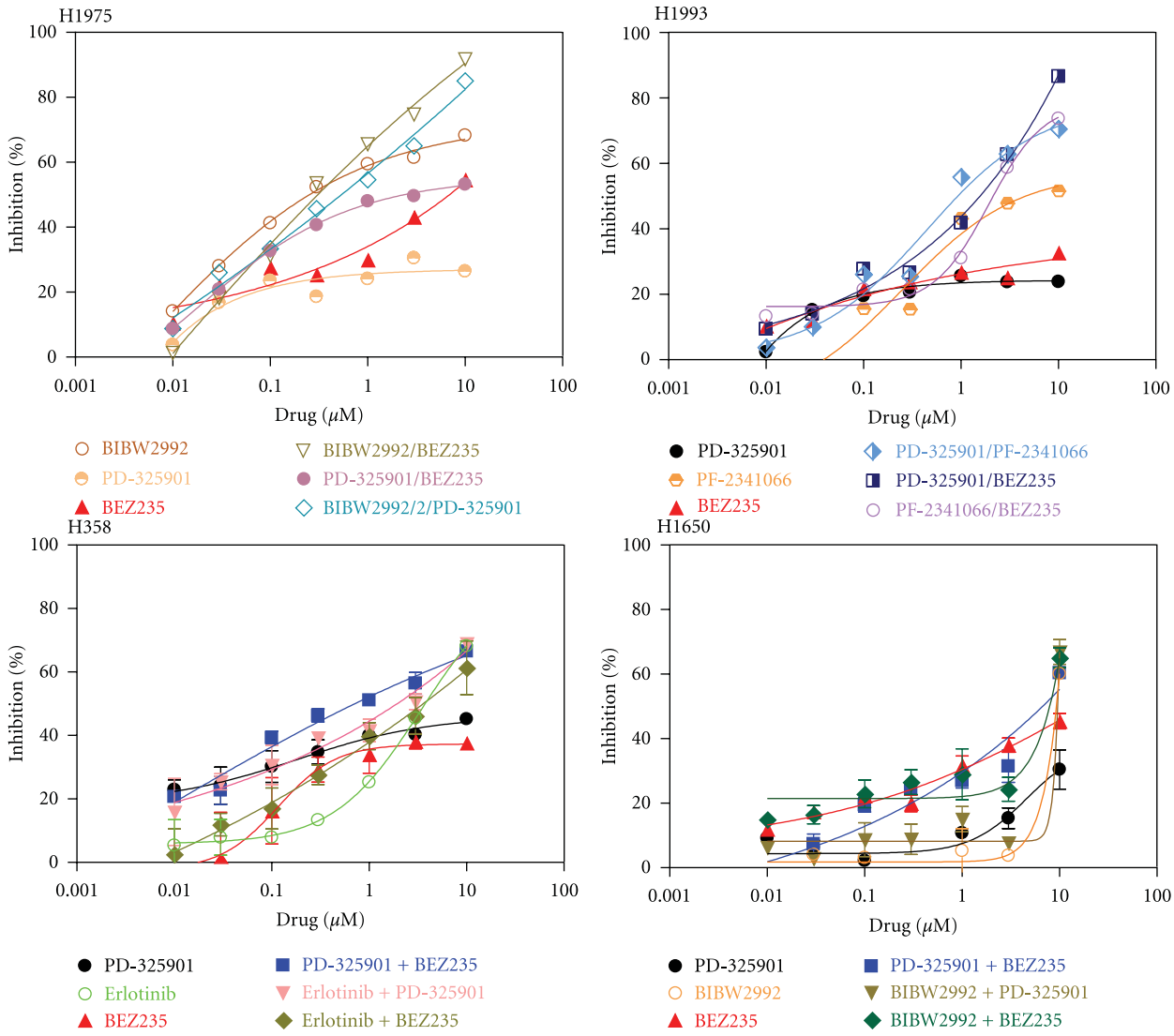


Figure 5: Inhibition of lung tumor cell growth by a combination of two kinase inhibitors. Lung tumor cell lines were cultured in 5% FBS plus increasing concentrations of the indicated single kinase inhibitor or a combination of the two indicated kinase inhibitors. Determination of cell proliferation was performed with the CellTiter-Glo Luminescent Cell Viability Assay.

3.6. Clustering of Tumor Tissue Samples with Tumor Cell Lines. Profiling of the 50 human lung tissue samples using the CEER assay revealed distinct activated signaling pathways in each of the tissue samples as shown in the heat map in Figure 6(a) (see also Supplemental Table 3). Based on these results, unsupervised one-dimensional clustering of the 50 lung tumor tissue samples with the corresponding cell lines H1993, H1975, HCC827, H1734, H1650, H358, H460, and A549, could be performed as shown in Figure 6(b) because a significant degree of shared pathway activation exists between the primary tumor tissue samples and the tumor cell lines. Clustering of the tumor tissue samples with their corresponding cell lines provides potential insight into targeted therapy based on the drug treatment results obtained from the tumor cell lines.

3.7. Gene Mutation Analysis of Tumor Tissue Samples. Mutation analysis of three frequently mutated genes (KRAS, P53, and STK11) in 50 human lung tissue samples further supported the clustering of human tumor samples into the cell line groups (see Supplemental Table 3). For example, patient samples LC16, LC44, and LC23 were originally aligned with tumor cell line H358 based on CEER assay profiling. This alignment was supported by the finding that these three tissue samples also harbored the G34A and G37T KRAS mutations as was found in the H358 cell line. Similarly, the original alignment of patient samples LC45 with cell line H1734 and LC21 with cell line A549 was substantiated by finding the same G34A and G37T KRAS mutations in both patient samples and cell lines. In addition, alignment of the patient sample LC15 with cell line H460 is supported by the finding of the A183T KRAS mutation in both patient sample and cell line. Likewise, the alignment of patient samples LC6, LC31, and LC13 with cell lines H1975 and H1734 was substantiated by the finding of the C726G and G818T P53, mutations in both patient samples and cell lines.

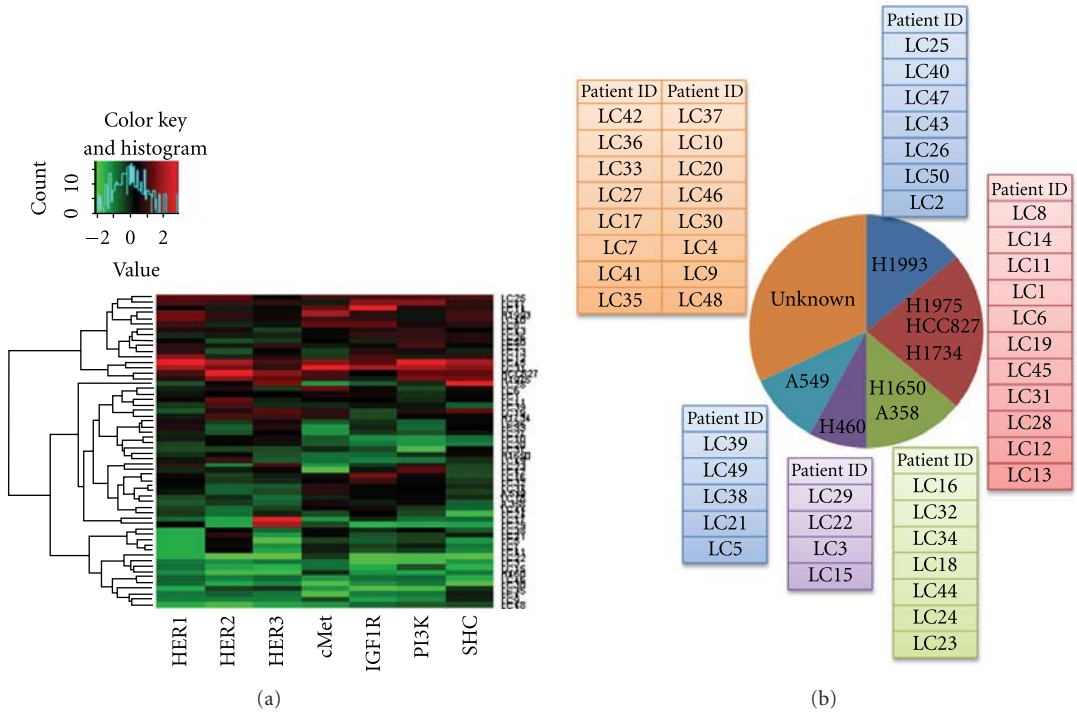


Figure 6: (a) Heat Map representing the activated signaling pathways found in the 50 lung tumor tissue samples and eight lung tumor cell lines. Each row constitutes all the pathway markers determined from an individual tumor sample organized in color columns. Green and red denote markers that are present at lower and higher levels, respectively. (b) Clustering of the 50 lung tumor samples with the corresponding lung tumor cell lines based on the similarities in the markers between the tissue samples and the cell lines.

Also, the alignment of patient samples LC32 and LC23 with cell line H1650 was supported by the finding of the same P53 mutations in both patient samples and cell line. Lastly, the alignment of patient sample LC25 with the H1993 cell line was supported by the finding of the C109T and G595T STK11 mutations in both patient sample and cell line. Thus, clustering of the patient tissue samples with the tumor cell lines based on mutational status of the KRAS, P53 and STK11 genes was also consistent with the clustering based on the RTK pathway signatures.

4. Discussion

Current lung cancer treatments are less than optimal, with a mean survival of less than one year for advanced lung cancer patients, regardless of treatment regimen [2]. Emerging new treatment modalities are generally targeted to inhibit specific tyrosine kinases activated in the tumor cells through basically two independent approaches [3]. The first approach is to use a highly specific monoclonal antibody to target the membrane growth factor receptor kinase that is responsible for tumor cell growth, and the resulting antibody/antigen complex invokes the host immune system to kill the tumor cells. This approach is exemplified by the treatment of HER2 receptor-positive breast cancer with Herceptin, a humanized monoclonal antibody against this receptor [4]. However, the high cost of monoclonal antibody drugs could be a disadvantage for this approach. The second approach is to develop cell-penetrating small organic molecules that target the specific tyrosine kinases in the signaling pathways of the tumor cells. This approach is best exemplified by the use of Gleevec to block the activation of the BCR-ABL fusion kinase in chronic myelogenous leukemia [5]. The design and synthesis of small molecule tyrosine kinase inhibitors have been greatly facilitated by the availability of crystal structures for the tyrosine kinases in the past decade and effective kinase inhibitors have thus been produced by many large and small pharmaceutical companies. Nevertheless, without prior knowledge of the activated kinase signaling pathways responsible for propagating and metastasis of the tumor cells, it is not possible to apply the targeted therapy approach with the available kinase inhibitors. Therefore, we have selected a panel of eight lung tumor cell lines that harbored the most frequently detected gene mutations: P53, KRAS, STK11, and HER1 as representative examples that cover the major human lung cancer subtypes.

Mutation analysis has provided valuable information in guiding targeted therapy for cancer patients. A good example is found in cancer patients who carry a KRAS gene mutation because these patients have been shown to be nonresponsive to anti-HER1 therapeutics. Therefore, KRAS mutation testing is becoming routinely performed in patients who are being considered for anti-HER1 therapy with either Cetuximab or Panitumumab in Europe and the United States [12, 13]. Our current study further confirmed an important role of gene mutation analysis in guiding the use of drugs

to treat lung cancers. For example, it has been reported that cancers from patients with lung adenocarcinoma that harbored mutations within the tyrosine kinase domain of the HER1 gene often responded initially to TKI drugs such as Gefitinib and Erlotinib [14, 15] but usually developed drug-resistance later [16–18]. Indeed, the HCC827, H1975, and H1650 lung tumor cell lines employed in this study harbored the HER1 gene mutation and they were found to be sensitive to HER1 inhibitor treatment. By contrast, KRAS gene mutation is associated with resistance to HER1 tyrosine kinase inhibitors [19, 20]. This phenomenon is substantiated by our finding that the A549 and H460 cells, which harbored the KRAS mutated gene, did not respond to treatment with HER1 inhibitors but they are sensitive to the downstream PI3K inhibitor BEZ-235.

Although targeted therapy based on association of somatic mutation analysis and drug sensitivity has greatly facilitated lung cancer treatment, the profiling of oncogenic signal transduction pathways in human tumor cells and tissues offers yet another complementary approach to guide therapeutic treatment [21–25]. Traditional signaling pathway profiling in tumor tissue samples by immunohistochemistry (IHC) staining or Western blotting methods are neither quantitative nor sensitive enough to have utility when only small amount of tumor tissue samples is available. The CEER assay is an assay that we have developed to overcome the low sensitivity and specificity issues associated with these traditional methods. The CEER assay uses a multiplexed, proximity-based, collaborative immunoassay platform that can provide clinical information on a limited amount of tissue samples with high sensitivity and specificity. The principle of the assay is based on the formation of a unique immunocomplex that requires the colocalization of two detecting antibodies against a target protein once the protein is captured on the microarray surface. It is the formation of this complex that enables the generation of a highly specific and sensitive signal to reveal the activation status of the target protein. We have compared CEER with the conventional IHC/FISH and Western blotting and found that the CEER assay provided more quantitative information in regard to the oncogenic kinases [26]. Using this assay, the activated HER1, HER2, HER3, c-MET, IGF-1R, PI3K, and SHC pathways present in the eight lung tumor cell lines as well as the 50 human lung tumor tissue samples were profiled. Cell lines that exhibited one or more of these activated pathways were treated with the corresponding kinase inhibitors and the results demonstrated that this matching approach is effective in inhibiting the pathway activation and growth of these cell lines under both anchorage-dependent and anchorage-independent culture conditions. Moreover, in those cell lines in which significant growth inhibition could not be achieved with a single kinase inhibitor treatment (H1993, H358, and H1650), a combination treatment with two kinase inhibitors, one targeting the RTK and the other targeting a downstream kinase, was effective in blocking their proliferation. However, we noted that there was no direct correlation between the IC_{50} value of the drug with the level of the corresponding pathway activation because we only measured a subset of pathway biomarkers. There might be other important pathways which may contribute to tumor growth. Another aspect is that the mutational status of tumor cells also plays an important role in driving the tumor growth. Nevertheless, characterization of the activated signaling pathways and mutational status as well as their cell growth inhibition by kinase inhibitors in the lung tumor cell lines could facilitate the target-focused treatment of lung cancers.

Pathway profiling using the CEER assay and gene mutational analysis of the 50 tumor tissues collected from lung cancer patients clearly showed some similarities in the biomarkers between the lung tumor tissue samples and tumor cell lines. Therefore, the panel of pathway biomarkers and mutated genes can discriminate and cluster different tumor tissue samples with the corresponding tumor cell lines in both supervised and unsupervised clustering analysis and this information could be used to guide treatment options based on the drug sensitivity of the tumor cell lines. However, it is important to interpret the data with caution due to the small sample size of the tumor tissue samples. Prospective validation studies with sufficient sample size and enough analytical power are essential to substantiate the role of these pathway biomarkers in cancer diagnosis and treatment. The ultimate validation of this approach is the patient's response to drug treatment based on the biomarker prediction. Irrespective of the slow progress made towards curing cancer, we have gained much knowledge through translational research by using new molecular and biological technology. We believe that the continuing gain in knowledge of lung cancer biology will provide the foundation for improvement in lung cancer treatment.

Abbreviations

SCLC: Small cell lung cancer
NSCLC: Nonsmall cell lung cancer
TKI: Tyrosine kinase inhibitor
RTK: Receptor tyrosine kinase
CEER: Channel enzyme enhanced reaction
FBS: Fetal bovine serum
PBS: Phosphate buffered saline
BSA: Bovine serum albumin.

Acknowledgments

The authors are indebted to Dr. Nicholas Ling and Dr. Belen Ybarrondo for their valuable inputs in editing and proofreading the paper.

References

[1] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA Cancer Journal for Clinicians*, vol. 61, pp. 69–90, 2011.

[2] A. A. Vaporciyan, J. C. Nesbitt, J. S. Lee et al., *Holland-Frei Cancer Medicine*, BC Decker, 5th edition, 2000.

[3] V. L. Roggli, R. T. Vollmer, S. D. Greenberg, M. H. McGavran, H. J. Spjut, and R. Yesner, "Lung cancer heterogeneity: a blinded and randomized study of 100 consecutive cases," *Human Pathology*, vol. 16, no. 6, pp. 569–579, 1985.

[4] J. H. Schiller, D. Harrington, C. P. Belani et al., "Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer," *The New England Journal of Medicine*, vol. 346, no. 2, pp. 92–98, 2002.

[5] R. L. Comis, "The current situation: erlotinib (Tarceva) and gefitinib (Iressa) in non-small cell lung cancer," *The Oncologist*, vol. 10, no. 7, pp. 467–470, 2005.

[6] S. Ramalingam and A. B. Sandler, "Salvage therapy for advanced non-small cell lung cancer: factors influencing treatment selection," *The Oncologist*, vol. 11, no. 6, pp. 655–665, 2006.

[7] G. S. Papaetis and K. N. Syrigos, "Sunitinib: a multitargeted receptor tyrosine kinase inhibitor in the era of molecular cancer therapies," *BioDrugs*, vol. 23, no. 6, pp. 377–389, 2009.

[8] J. W. Neal and L. V. Sequist, "Exciting new targets in lung cancer therapy: ALK, IGF-1R, HDAC, and Hh," *Current Treatment Options in Oncology*, vol. 11, no. 1-2, pp. 36–44, 2010.

[9] V. Serra, M. Scaltriti, L. Prudkin et al., "PI3K inhibition results in enhanced HER signaling and acquired ERK dependency in HER2-overexpressing breast cancer," *Oncogene*, vol. 30, no. 22, pp. 2547–2557, 2011.

[10] H. C. Gong, Y. Honjo, P. Nangia-Makker et al., "The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells," *Cancer Research*, vol. 59, no. 24, pp. 6239–6245, 1999.

[11] W. Pao, V. A. Miller, K. A. Politi et al., "Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the HER1 kinase domain," *PLoS Medicine*, vol. 2, no. 3, pp. 225–235, 2005.

[12] F. Le Calvez, A. Mukeria, J. D. Hunt et al., "TP53 and KRAS mutation load and types in lung cancers in relation to tobacco smoke: distinct patterns in never, former, and current smokers," *Cancer Research*, vol. 65, no. 12, pp. 5076–5083, 2005.

[13] W. Pao, T. Y. Wang, G. J. Riely et al., "KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib," *PLoS Medicine*, vol. 2, no. 1, pp. 57–61, 2005.

[14] T. Sjöblom, S. Jones, L. D. Wood et al., "The consensus coding sequences of human breast and colorectal cancers," *Science*, vol. 314, no. 5797, pp. 268–274, 2006.

[15] A. Potti, H. K. Dressman, A. Bild et al., "Genomic signatures to guide the use of chemotherapeutics," *Nature Medicine*, vol. 12, no. 11, pp. 1294–1300, 2006.

[16] W. Pao, V. Miller, M. Zakowski et al., "EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 36, pp. 13306–13311, 2004.

[17] T. J. Lynch, D. W. Bell, R. Sordella et al., "Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib," *The New England Journal of Medicine*, vol. 350, no. 21, pp. 2129–2139, 2004.

[18] E. L. Kwak, R. Sordella, D. W. Bell et al., "Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 21, pp. 7665–7670, 2005.

[19] J. G. Paez, P. A. Jänne, J. C. Lee et al., "HER1 mutations in lung cancer: correlation with clinical response to gefitinib therapy," *Science*, vol. 304, no. 5676, pp. 1497–1500, 2004.

[20] S. Kobayashi, T. J. Boggon, T. Dayaram et al., "HER1 mutation and resistance of non-small-cell lung cancer to gefitinib," *The New England Journal of Medicine*, vol. 352, no. 8, pp. 786–792, 2005.

[21] J. S. Sebolt-Leopold, "Advances in the development of cancer therapeutics directed against the RAS-mitogen-activated protein kinase pathway," *Clinical Cancer Research*, vol. 14, no. 12, pp. 3651–3656, 2008.

[22] D. W. Rusnak, K. J. Alligood, R. J. Mullin et al., "Assessment of epidermal growth factor receptor (HER1, ErbB1) and HER2 (ErbB2) protein expression levels and response to lapatinib (Tykerb, GW572016) in an expanded panel of human normal and tumour cell lines," *Cell Proliferation*, vol. 40, no. 4, pp. 580–594, 2007.

[23] K. D. Paull, R. H. Shoemaker, L. Hodes et al., "Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm," *Journal of the National Cancer Institute*, vol. 81, no. 14, pp. 1088–1092, 1989.

[24] U. McDermott, S. V. Sharma, L. Dowell et al., "Identification of genotype-correlated sensitivity to selective kinase inhibitors by using high-throughput tumor cell line profiling," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 50, pp. 19936–19941, 2007.

[25] N. Nakatsu, T. Nakamura, K. Yamazaki et al., "Evaluation of action mechanisms of toxic chemicals using JFCR39, a panel of human cancer cell lines," *Molecular Pharmacology*, vol. 72, no. 5, pp. 1171–1180, 2007.

[26] S. Singh, X. Liu, T. Lee et al., "Analysis of truncated HER2 expression and activation in breast cancer," *Journal of Clinical Oncology, ASCO Annual Meeting*, vol. 28, no. 15, p. 1099, 2010.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 145010, 8 pages
doi:10.1155/2011/145010

Research Article

Comprehensive Proteomic Profiling of Aldehyde Dehydrogenases in Lung Adenocarcinoma Cell Lines

Qing Zhang,[1] Ayumu Taguchi,[1] Mark Schliekelman,[1] Chee-Hong Wong,[1] Alice Chin,[1] Rork Kuick,[2] David E. Misek,[3] and Samir Hanash[1]

[1] *Molecular Diagnostics Program, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M5-C800, Seattle, WA 98109, USA*
[2] *University of Michigan Comprehensive Cancer Center, 8D17 300 NIB, University of Michigan, Ann Arbor, MI 48109, USA*
[3] *Department of Surgery, University of Michigan, 1150 West Medical Center Drive, Room A510A MSRB-1, Ann Arbor, MI 48109, USA*

Correspondence should be addressed to David E. Misek, dmisek@umich.edu

Received 1 July 2011; Accepted 14 August 2011

Academic Editor: Tadashi Kondo

Copyright © 2011 Qing Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We have explored the potential of proteomic profiling to contribute to the delineation of the range of expression and subcellular localization of aldehyde dehydrogenases (ALDHs) in lung adenocarcinoma. In-depth quantitative proteomics was applied to 40 lung adenocarcinoma cell lines resulting in the identification of the known members of the ALDH family. Substantial heterogeneity in the level and occurrence of ALDHs in total lysates and on the cell surface and in their release into the culture media was observed based on mass spectrometry counts. A distinct pattern of expression of ALDHs was observed in cells exhibiting epithelial features relative to cells exhibiting mesenchymal features. Strikingly elevated levels of ALDH1A1 were observed in two cell lines. We also report on the occurrence of an immune response to ALDH1A1 in lung cancer.

1. Introduction

Aldehyde dehydrogenases (ALDHs) constitute a large family of enzymes that catalyze the oxidation of endogenous and exogenous aldehydes to carboxylic acids. The different family members exhibit heterogeneous tissue distribution and have been localized predominantly in the cytosol and mitochondria. ALDHs have been investigated primarily based on their gene expression, their immunohistochemical localization and based on their activity. Cell sorting techniques have been utilized to enrich for cells expressing these enzymes based on activity. The relevance of ALDHs to cancer stems in part from the role they may play during carcinogenesis, their association with therapeutic resistance, and more recently from a distinct pattern of expression of ALDH1A1 and ALDH3A1 in cancer stem cells, which has been exploited as a means to define this cell population in tumors [1–3]. Several studies have pointed to evidence for epithelial to mesenchymal transitions in cancer stem cells, defined, in part, based on their ALDH expression or activity [4, 5]. While ALDH1 and 3 dominate the cancer stem cell literature, other ALDHs have also been explored. High-level expression of ALDH1B1 was observed in colon cancer by immunohistochemistry [6]. Moderate to strong staining for ALDH4A1, ALDH5A1, and ALDH6A1 was observed in most cancer tissue in Protein Atlas (http://www.proteinatlas.org/). Strong expression of ALDH7A1 has been found in human prostate cancer cell lines, primary tumors, and matched bone metastases, with evidence of its functional involvement in the formation of bone metastases [7]. Comparative analysis of hepatocellular carcinoma tissue and adjacent nontumor tissue identified changes in ALDHs1-3 proteins in tumor tissue [8].

Several studies have explored the biological significance of ALDHs specifically in lung cancer and have provided supportive evidence for the association between ALDH activity and lung cancer stem cells [9]. Flow cytometric analysis of a panel of lung cancer cell lines and patient-derived

tumors revealed the occurrence of a subpopulation of cells with elevated ALDH activity in most non-small cell lung cancers, which correlated with ALDH1A1 expression [10]. Immunohistochemical staining of a large panel of primary tumors revealed a significant correlation between ALDH1A1 expression and poor prognosis in patients, including those with early stage disease [10]. Likewise, in another study, expression of ALDH1 was found to be positively correlated with the stage and grade of lung tumors and related to a poor prognosis for patients with early-stage lung cancer [11]. Expression analysis of sorted cells revealed elevated Notch pathway transcript expression in ALDH-positive cells. Suppression of the Notch pathway resulted in a significant decrease in ALDH-positive lung cancer cells with concordant reduction in tumor cell proliferation and clonogenicity [10]. Downregulation of ALDH isozymes affects cell growth, cell motility, and gene expression in lung cancer cells [12]. Other ALDHs have also been explored in lung cancer. ALDH3B1 is expressed in a tissue-specific manner and in a limited number of cell types. ALDH3B1 expression was found to be upregulated in a high percentage of human tumors, particularly lung tumors [13].

In general, most studies of ALDHs in lung cancer have focused on particular members of the family. In-depth proteomic profiling allows delineation of proteins expressed in tumor cells and in subcellular compartments. In this study we applied quantitative in-depth proteomic profiling to assess the occurrence of ALDHs in whole lysates of 40 lung adenocarcinoma cell lines and to examine their association with the cell surface and the extent of their release into culture media. ALDH1A1 was further explored as a tumor antigen that induces an autoantibody response in lung cancer.

2. Methods

2.1. Lung Adenocarcinoma Cell Line Culture. Cells were grown in DMEM media (Invitrogen) containing 0.1% of dialyzed fetal bovine serum (FBS) (Invitrogen) and ^{13}C-lysine instead of regular lysine, for 7 passages (1 : 2) according to the standard SILAC protocol [14]. Incorporation of ^{13}C Lys isotope exceeded 90% of the total protein lysine content. The same batch of cells was used for extracting cell surface proteins and for analysis of conditioned media and whole-cell lysate proteins. The secreted proteins were obtained directly from the cell conditioned media after 48 h of culture. Cells and debris were removed by centrifugation at 5000 ×g and filtration through a 0.22 μM filter. Total extracts of cells were obtained by sonication of $\sim 2 \times 10^7$ cells in 1 mL of PBS containing the detergent octyl glucoside (OG) (1% w/v) and protease inhibitors (complete protease inhibitor cocktail, Roche Diagnostics, Germany) followed by centrifugation at 20,000 ×g.

2.2. Capture of Cell Surface Proteins. To isolate cell surface proteins, $\sim 2 \times 10^8$ cells were biotinylated in the culture plate after extensive PBS rinsing, with 10 mL of 0.25 mg/mL of Sulfo-NHS-SS-BIOTIN in PBS at room temperature (23-24°C) for 10 min. The residual biotinylation reagent was quenched with 10 mM lysine. Protein extraction was performed in a solution containing NP 40 detergent 2% (v/v) with cell disruption by sonication followed by centrifugation at 20,000 ×g. Biotinylated proteins were chromatographically isolated by affinity chromatography using 1 mL of UltraLink Immobilized NeutrAvidin (Pierce) according to manufacturers' instruction. Proteins bound to the column were recovered by reduction of the biotinylation reagent with 5 mL of a solution containing 65 μMol of DTT and 1% octyl glucoside (OG) detergent for 1 h at 37°C. Eluted proteins were subsequently alkylated with 200 μMol of iodoacetamide at room temperature.

2.3. Fractionation of Cell Extracts. Cell surface, conditioned media, and total extract were fractionated by reversed-phase chromatography, using, respectively 500 μg, 1 mg, and 1 mg of total protein. All the cell extracts were reduced and alkylated with iodoacetamide prior to chromatography. Separation was performed in a POROS R1/10 column (Applied Biosystems—4.6 × 50 mm) at 2.7 mL/min using a linear gradient of 10 to 80% of organic solvent over 30-minute run. Solvent system used was aqueous solvent, 5% acetonitrile/95% water/0.1% of trifluoroacetic acid; organic solvent, 75% acetonitrile/15% isopropanol/10% water/0.095% trifluoroacetic acid. Fractions were collected at a rate of 3 fractions/minute.

2.4. Protein Identification by LC-MS/MS. Protein digestion and identification by LC-MS/MS was performed as described previously [15]. Briefly, each one of the reversed-phase fractions was individually digested in-solution digestion with trypsin (400 ng/fraction) and grouped into 24 to 27 pools for each cell line and each compartment (i.e., cell surface, conditioned media, and soluble whole-cell lysate) based on chromatographic features. Pools were individually analyzed by LC-MS/MS in a LTQ-Orbitrap mass spectrometer (Thermo-Finnigan) coupled to a nanoflow chromatography system (Eksigent) using a 25 cm column (PicoFrit 75 μM ID, New Objectives, packed in-house with Magic C18 resin) over a 90-minute linear gradient. Acquired data was automatically processed by the Computational Proteomics Analysis System (CPAS) [16]. The tandem mass spectra were searched against version 3.57 of the human IPI database. A variable modification of 6.020129 mass units was added to lysine residues for database searching to account for incorporation of the heavy lysine isotope. We applied the tools PeptideProphet [17] and ProteinProphet [18] to estimate the significance of peptide and protein matches. Identifications with a PeptideProphet probability of greater than 0.2 were selected and submitted to ProteinProphet. The latter infers a minimal set of proteins that explain the peptide evidence, assigning a probability to each protein based on the combined peptide probabilities. The derived protein identifications were filtered at a <5% error rate based on the probability that the best match obtained would fall in the distribution of random database matches [19]. A spectral counting method [20] was used to estimate protein enrichment for each compartment.

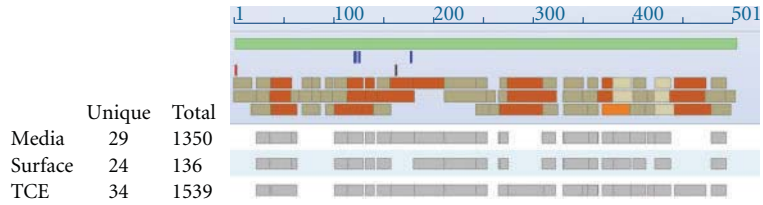


Figure 1: Extent of peptide coverage of ALDH1A1 in cell line H522 in the TCE, media, and cell surface, across the ALDH1A1 protein sequence. Blue bar: dbSNP; red bar: N-acetylserine; black bar: conflict, V->I at 162; tan bar: non-Cys-containing tryptic peptide; orange bar: Cys-containing tryptic peptide; light tan bar: N-linked non-Cys-containing tryptic peptide; light orange bar: N-linked Cys-containing peptide.

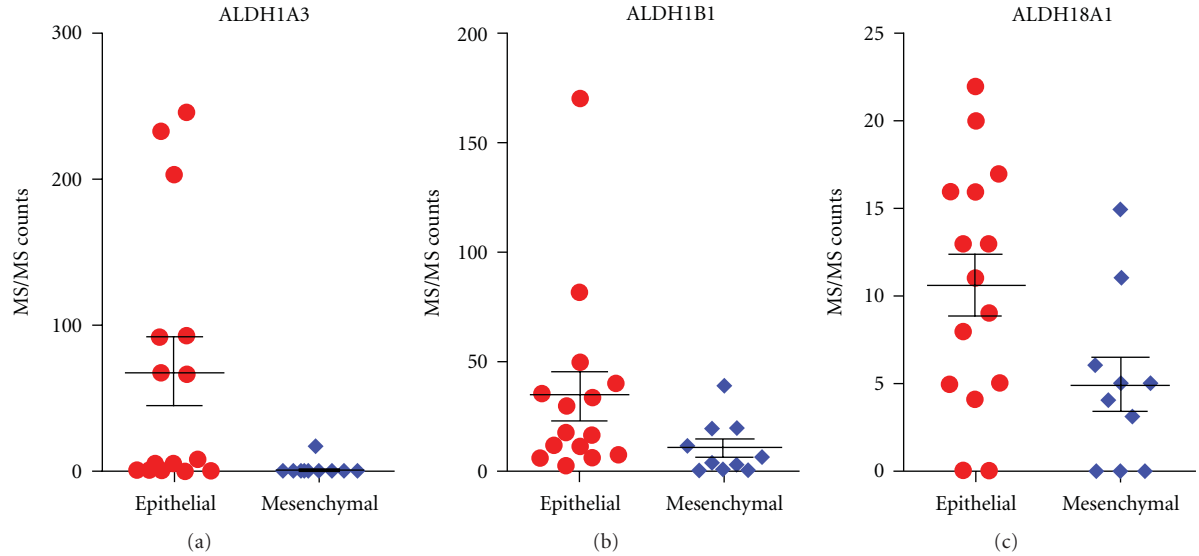


Figure 2: Differential expression of ALDH1A3, ALDH1B1, and ALDH18A1 between cell lines with epithelial and cell lines with mesenchymal features based on MS/MS counts. A Mann-Whitney test was performed using MS/MS counts of epithelial and mesenchymal cell lines. ALDH1A3 and ALDH1B1 have significantly higher expression from TCE in epithelial cells as compared to mesenchymal ones, with a P value of 0.007 and 0.030, respectively. ALDH18A1 has significant higher expression from the cell surface in epithelial as compared to mesenchymal with a P value of 0.042.

2.5. Two-Dimensional Polyacrylamide Gel Electrophoresis (2D PAGE). Proteins derived from the extracts of the cultured H522 lung adenocarcinoma cell line were separated into two dimensions as described previously [21]. Briefly, cultured NCI-H522 cells were lysed in solubilization buffer (8 M urea (Bio-Rad), 2% Nonidet P-40, 2% carrier ampholytes, pH 4–8 (Gallard/Schlesinger, Carle Place, NY), 2% ß-mercaptoethanol, and 10 mM PMSF). 200 μg of solubilized proteins were applied onto isoelectric focusing gels. Isoelectric focusing was performed using pH 4–8 carrier ampholytes at 700 V for 16 h, followed by 1000 V for an additional 2 h. The first-dimension gel was loaded onto the second-dimension gel, after equilibration in 125 mM Tris (pH 6.8), 10% glycerol, 2% SDS, 1% DTT, and bromophenol blue. For the second-dimension separation, a gradient of 11–14% acrylamide (Crescent Chemical, Hauppauge, NY) was used. The resolved proteins were transferred to an Immobilon-P polyvinylidene difluoride (PVDF) membrane (Millipore, Bedford, Mass). Protein patterns in some gels were visualized directly by silver staining or Sypro Ruby staining.

2.6. Western Blotting and Image Analysis. After transfer, membranes were incubated with a blocking buffer consisting of PBS and 0.1% Tween-20 containing 1.8% nonfat dry milk for 2 h. The membranes were incubated for 1 h at room temperature with serum obtained from either patients or control individuals as a source of primary antibody at a 1 : 100 dilution. Some additional membranes were incubated with an antibody to ALDH1A1 to visualize the protein. After three washes with washing buffer (PBS containing 0.1% Tween 20), the membranes were incubated with horseradish peroxidase-conjugated sheep anti-human IgG (Amersham Biosciences, Piscataway, NJ) at a dilution of 1 : 1000 for 1 h at room temp. Immunodetection was accomplished by enhanced chemiluminescence (Amersham Biosciences) followed by exposure on Hyperfilm MP (Amersham Biosciences). Films were digitized with a Kodak CCD camera. The spots on each image were detected and quantified as previously described [21]. After immunoblotting, all membranes were Coomassie Blue stained, and the patterns obtained were compared to those of the films in order to determine the locations of reactive spots. Spot integrated intensities were normalized

Cell line	Unique	Total
DFC1032	24	219
HCC4017	24	194
HCC2935	18	173
PC9	16	116
H820	17	76
HCC4011	18	72
HCC4019	11	71
HCC827	15	58
HCC4006	14	43
H1792	15	39
H1703	2	17
DFC1024	9	15
H522	2	15
H1373	7	8
H1819	5	8
H1944	2	6
H1385	1	6
H1355	1	5
H1437	1	4
H1650	3	3
H1693	1	3
H1395	1	2
H969	1	2
H2291	2	2

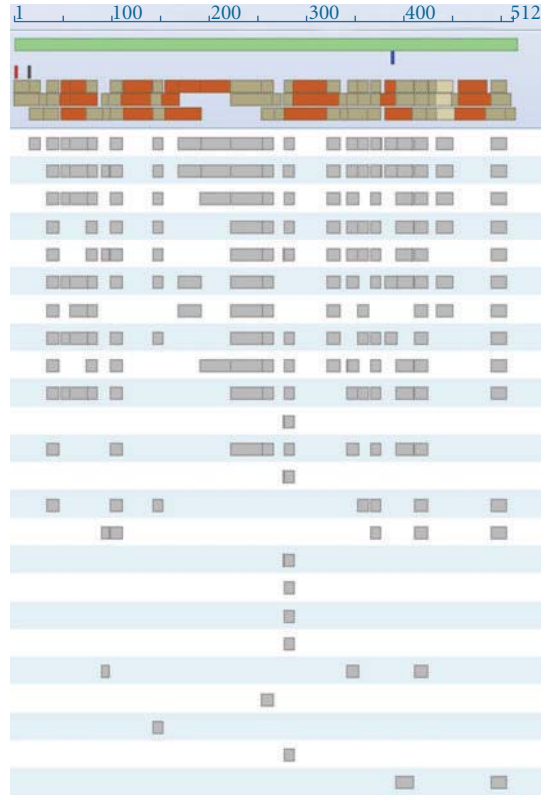


Figure 3: Extent of peptide coverage of ALDH1A3 across cell lines based on spectral counts from MS data.

by dividing by the intensity of a reference HSP27 protein spot. HRP-conjugated goat anti-mouse secondary antibody was spiked into the sheep anti-human cocktail at a 1 : 1000 dilution to facilitate detection.

2.7. Statistical Analysis. The data was analyzed using one-sided Rank-Sum tests.

3. Results

3.1. Occurrence of ALDHs among Lung Adenocarcinoma Cell Lines and in Subcompartments. Comprehensive proteomic profiling of 40 lung cancer cell lines resulted in the identification of known members of the ALDHs family in one or more cell lines. There was substantial heterogeneity in the abundance of individual family members among cell lines and in their occurrence and distribution between whole-cell lysates, on the cell surface and their release into the media of cultured cells (Table 1). ALDH1A1, ALDH1A3, ALDH2, ALDH3A1, ALDH7A1, and ALDH9A1 had more than 1,000 total MS counts each in total cell extracts across cell lines. ALDH1A1 and ALDH3A1 exhibited discordant expression with an overall total of MS counts for ALDH1A1 exceeding that of ALDH3A1 (Table 1), with some cell lines yielding high MS counts for ALDH1A1 with low counts or undetectable for ALDH3A1. The highest MS counts were observed for ALDH2 and ALDH1A1. However, whereas ALDH2 was relatively uniformly distributed across cell lines, expression of ALDH1A1 was more limited to just a few cell lines. Notably, cell line H522 yielded extremely high MS counts for ALDH1A1 relative to most other cell lines, and relative to other ALDHs (Table 2). Several ALDH proteins for which we provide evidence of expression in lung cancer cell lines (Table 1) were previously not characterized in lung cancer. ALDH9A1 was identified in most cell lines and was among the most abundant ALDH proteins identified based on spectral counts.

3.2. Occurrence of ALDHs in the Extracellular Compartment. We sought to determine the occurrence of ALDHs on the cell surface and their release into culture media. Whereas all ALDH proteins were identified in total cell extracts (TCE), for most proteins a significant fraction was observed in both the cell surface fraction (Surf) and the media (Med) (Table 1). For example, ALDH1A1 in H522 was predominant in both the TCE and Med (Table 2) whereas other proteins (e.g., ALDH3A2 and ALDH5A1) displayed a substantial proportion of protein abundance on the cell surface (Table 1). We examined whether MS data was suggestive of the occurrence of cleavage forms of ALDHs in the media or on the cell surface. However, the data obtained pertaining to peptide representation of the proteins was not suggestive of cleavage, or alternatively spliced forms of proteins associated with the cell surface or, for that matter, with their release into the media as exemplified for ALDH1A1 in H522 (Figure 1).

3.3. Differential Expression of ALDHs Based on Epithelial/Mesenchymal Features. Twenty-five of the 40 cell lines

Table 1: Identified Aldehyde Dehydrogenases in the total cell extracts (TCE), Media (Med) and the cell surface (Surf) across 40 lung adenocarcinoma cell lines based on mass spectrometry (MS/MS counts for corresponding peptides).

Gene symbol	Protein description	MS/MS counts (TCE)	MS/MS counts (Med)	MS/MS counts (Surf)	Unique peptides	Chromosome	Length	MW
ALDH16A1	Aldehyde dehydrogenase family 16 member A1	31	79	81	19	19	802	85127
ALDH18A1	Delta-1-pyrroline-5- carboxylate synthase	354	5	344	20	10	795	87302
ALDH1A1	Retinal dehydrogenase 1	4961	2640	1463	37	9	500	54731
ALDH1A2	Retinal dehydrogenase 2	10	1	1	5	15	480	53060
ALDH1A3	Aldehyde dehydrogenase family 1 member A3	1267	256	299	31	15	512	56108
ALDH1B1	Aldehyde dehydrogenase X, mitochondrial	843	22	384	29	9	517	57217
ALDH1L1	Aldehyde dehydrogenase family 1 member L1	5	3	3	8	3	505	55394
ALDH1L2	Aldehyde dehydrogenase family 1 member L2, mitochondrial	11	0	51	15	12	923	101776
ALDH2	Aldehyde dehydrogenase, mitochondrial	4964	889	1665	32	12	517	56381
ALDH3A1	Aldehyde dehydrogenase, dimeric NADP-preferring	1149	44	116	27	17	453	50379
ALDH3A2	Fatty aldehyde dehydrogenase	772	44	457	23	17	485	54848
ALDH3B1	Aldehyde dehydrogenase family 3 member B1	3	0	20	6	11	468	51840
ALDH3B2	Aldehyde dehydrogenase family 3 member B2	3	0	0	2	11	385	42670
ALDH4A1	Delta-1-pyrroline-5- carboxylate dehydrogenase, mitochondrial	256	0	41	22	1	563	61719
ALDH5A1	Succinate-semialdehyde dehydrogenase, mitochondrial	437	8	192	19	6	535	57215
ALDH6A1	Methylmalonate-Semialdehyde dehydrogenase [acylating], mitochondrial	131	0	13	17	14	535	5784
ALDH7A1	Alpha-aminoadipic semialdehyde dehydrogenase	2268	95	563	29	5	510	55235
ALDH8A1	Aldehyde dehydrogenase family 8 member A1	2	0	0	2	6	487	53401
ALDH9A1	4-trimethylaminobutyral dehyde dehydrogenase	1096	98	365	22	1	491	53374

investigated could be classified as either epithelial or mesenchymal based on their morphology and their expression of vimentin and E-cadherin. ALDH protein expression was examined in relation to cell line epithelial/mesenchymal characteristics. ALDH1A3, ALDH1B1, and ALDH18A1 exhibited statistically significant differences in their MS counts in epithelial versus mesenchymal cell lines (Table 3, Figure 2). Whereas the stem cell markers ALDH1A1 and ALDH3A1, did not yield statistically significant differences between epithelial and mesenchymal cell lines, a very high number of MS spectral counts were observed for ALDH1A1 in the two mesenchymal cell lines H522 and H1703 (Table 3). The extent of peptide coverage for ALDH1A3, ALDH1B1, and ALDH18A1 between epithelial and mesenchymal cell lines appears to be related to the total number of MS counts rather than to epithelial/mesenchymal characteristics, as shown in Figure 3 for ALDH1A3 in TCE.

3.4. Lung Cancer Sera Exhibit IgG-Based Reactivity to ALDH1A1. H522 tumor cell line proteins were separated by 2D PAGE and transferred onto Immobilon-P PVDF membranes, and the membranes were used to screen individual lung cancer and control sera for autoantibodies directed against H522 proteins. Sera from 25 lung cancer patients (9 adenocarcinomas, 6 small cell, and 10 squamous lung cancers) and 25 age- and sex-matched healthy controls were investigated. Two neighboring spots, both with similar apparent molecular weight (approximately 55 kDa) but slightly different isoelectric points (6.6 and 6.45, resp.), were observed to exhibit frequent reactivity in lung cancer patients

Table 2: Identified Aldehyde Dehydrogenases in cell line H522.

Gene symbol	MS/MS counts (TCE)	MS/MS counts (Med)	MS/MS counts (Surf)
ALDH16A1	4	1	1
ALDH18A1	31	0	6
ALDH1A1	2053	1706	173
ALDH1A3	0	0	1
ALDH1B1	39	0	24
ALDH1L2	0	0	3
ALDH2	29	3	47
ALDH3A1	0	0	8
ALDH3A2	3	0	2
ALDH4A1	15	0	0
ALDH5A1	15	0	5
ALDH7A1	164	2	26
ALDH9A1	21	0	8

Table 3: Occurrence of ALDH1A1, ALDH3A1, ALDH1A3, ALDH1B1 and ALDH18A1 in cell lines with epithelial or mesenchymal features.

Cell Lines	EMT	ALDH1A1			ALDH3A1			ALDH1A3			ALDH1B1			ALDH18A1		
		TCE	Med	Surf	TCE	Med	Surf	TCE	Med	Surf	TCE	Med	Surf	TCE	Med	Surf
H1437	Epithelial	82	2	0	0	6	2	0	0	0	33	0	9	18	0	16
H1650	Epithelial	0	0	0	0	0	0	3	0	0	11	0	5	4	0	0
H3255	Epithelial	0	0	0	0	0	0	0	0	0	2	0	0	8	0	5
HCC4019	Epithelial	0	0	0	0	0	0	93	0	19	29	0	6	0	0	0
DFC1032	Epithelial	0	0	0	0	0	3	247	112	77	17	0	7	0	0	16
HCC827	Epithelial	0	0	0	0	0	2	65	0	0	81	1	14	14	0	5
H1573	Epithelial	0	0	0	129	0	0	0	0	0	6	0	0	1	0	13
HCC4011	Epithelial	0	0	0	0	0	0	91	10	5	40	0	12	12	0	13
H1819	Epithelial	0	0	0	0	0	0	8	0	0	6	0	3	6	0	9
H1395	Epithelial	8	0	0	4	3	8	0	0	0	49	0	12	9	0	4
H969	Epithelial	134	108	33	7	1	0	0	0	0	11	0	6	62	0	20
H820	Epithelial	0	0	0	0	0	4	67	2	14	171	14	66	27	3	17
H2291	Epithelial	0	0	0	0	0	0	3	0	2	7	0	19	8	2	22
HCC4017	Epithelial	0	0	0	0	0	0	234	93	5	35	1	3	9	0	8
HCC2935	Epithelial	5	0	0	0	6	0	205	8	73	16	0	3	2	0	11
H1299	Mesenchymal	0	0	0	4	0	0	0	0	2	3	0	5	0	0	3
H23	Mesenchymal	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0
H838	Mesenchymal	0	0	0	0	0	1	0	0	2	0	0	2	0	0	0
H2030	Mesenchymal	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
H650	Mesenchymal	0	0	0	0	0	0	0	0	0	6	0	4	0	0	0
DFC1024	Mesenchymal	0	0	0	0	0	0	17	0	42	20	0	6	2	0	5
H1355	Mesenchymal	34	2	0	0	0	13	0	0	1	11	0	7	17	0	11
H2405	Mesenchymal	2	0	0	0	0	41	0	0	0	2	0	9	5	0	4
H522	Mesenchymal	2053	1706	173	0	0	8	0	0	1	39	0	24	31	0	6
H1703	Mesenchymal	1665	354	78	0	0	7	0	0	18	20	1	30	10	0	15

(Figure 4). The data for the more acidic spot was highly correlated to the more basic spot values, giving a Spearman's rank correlation of 0.62 with a P value of 2.3×10^{-6}. The more basic spot was excised from gels on two occasions and digested with trypsin, with the resulting tryptic digests subjected to mass spectrometric analysis. We obtained 46 and 45 spectra matching the aldehyde dehydrogenase 1A1 protein (NP_000680, gene symbol ALDH1A1) on the two runs. These matches amounted to 26 and 22 distinct peptides, with MASCOT search engine scores of 1503 and 1464, respectively. The more acidic spot was also identified as ALDH1A1, with 21 matching peptides and MASCOT score of 1329.

Spot integrated intensities on the films were measured and analyzed by one-sided Rank-Sum tests. We obtained

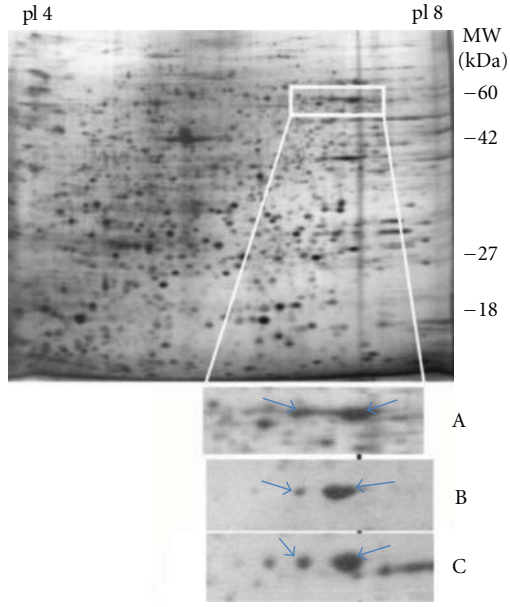


Figure 4: Occurrence of autoantibodies to ALDH1A1 in lung-cancer. A rectangle marks the location of ALDH1A1 in a silver-stained 2D gel of H522 lysate, also shown in a close-up below (A). The acidic and basic ALDH1A1 reactive spots in Western blots hybridized with subject sera (B and C) are indicated with arrows.

estimated P values of 0.0018 (basic) and 0.0026 (acidic) from 10000 permuted data sets, indicating greater reactivity in lung cancer patients compared to healthy controls. Two receiver operating characteristic (ROC) curves were constructed, each using the data for a single protein spot (either the more basic or more acidic spot). For the more basic spot, an AUC of 0.69 was obtained. An AUC of 0.68 was obtained for the more acidic spot.

4. Discussion

Comprehensive proteomic profiling delineated the range of expression and localization of ALDHs in total cell lysates, the cell surface, and their release into the media of cultured lung adenocarcinoma cell lines. A striking finding is the extent to which ALDHs localize to the cell surface and/or are released into the media. The evidence obtained based on the peptides identified for various ALDHs does not suggest the occurrence of distinct forms of ALDHs in different compartments, thus the mechanisms involved and the role of ALDHs localized to the cell surface or release into the media largely remain to be determined.

We observed a distinct pattern of expression of ALDHs in cell lines exhibiting epithelial versus mesenchymal features. In other studies, ALDH expression was found to mark pancreatic cancer cells that have stem cell and mesenchymal features [22]. ALDH expression was analyzed by immunohistochemistry in 269 primary surgical specimens of pancreatic adenocarcinoma and examined for association with clinical outcomes and in paired primary tumors and metastatic lesions from eight pancreatic cancer patients who had participated in a rapid autopsy program. The clonogenic growth potential of ALDH-positive pancreatic adenocarcinoma cells was assessed in vitro by a colony formation assay and by tumor growth in immunodeficient mice. ALDH-positive tumor cells were detected in 90 of the 269 primary surgical specimens, and their presence was associated with worse survival. Six of the eight patients with matched primary and metastatic tumor samples had ALDH-negative primary tumors, and in four of these six patients, the matched metastatic lesions contained ALDH-positive cells which expressed genes consistent with a mesenchymal state [22].

The extent to which mesenchymal features correlate with a stem cell phenotype reflected in a common ALDH expression pattern remains to be determined. In our studies, ALDH1A1 and ALDH3A1, which have been associated with stem cell features, exhibited distinct expression patterns among the cell lines analyzed. ALDH3A1 did not yield a significant association with a mesenchymal phenotype whereas ALDH1A1 exhibited strikingly high expression based on MS counts in two cells lines with a mesenchymal phenotype. Moreover, ALDH1A3, ALDH1B1, and ALDH18A1 exhibited a statistically significant difference in their protein expression between cell lines with epithelial and mesenchymal features.

We utilized lysates from H522, a lung adenocarcinoma cell line that expresses high levels of ALDH1A1 to characterize the humoral immune response in lung cancer. We obtained evidence for autoreactivity against two forms of ALDH1A in sera from subjects with lung cancer relative to healthy controls. The reactivity observed was not limited to lung adenocarcinoma, as sera from subjects with squamous cell lung cancer and small cell lung cancer exhibited similar reactivity. An immune response to ALDH1A1 has been reported in other studies. ALDH1A1 was identified as a novel CD8+ T-cell-defined tumor antigen in squamous cell carcinoma of the head and neck [23]. Mass spectral analysis of peptides in tumor-derived lysates was used to determine that the CTL line recognized the HLA-A2 binding ALDH1A1 (88–96) peptide. ALDH1A1(88–96) peptide-specific CD8(+) T cells recognized only HLA-A2(+) cell lines which over expressed ALDH1A1 and cells transfected with ALDH1A1 cDNA. In another study, sera from three of five patients with lung adenocarcinoma and none of ten controls with lung tuberculosis were found to exhibit IgG-based seroreactivity against aldehyde dehydrogenase identified in A549 lung adenocarcinoma cell line lysate using an approach similar to the approach for IgG reactivity against ALDH1A1 utilized in this study [24].

References

[1] G. K. Rekha, L. Sreerama, and N. E. Sladek, "Intrinsic cellular resistance to oxazaphosphorines exhibited by a human colon carcinoma cell line expressing relatively large amounts of a class-3 aldehyde dehydrogenase," *Biochemical Pharmacology*, vol. 48, no. 10, pp. 1943–1952, 1994.

[2] N. Tsukamoto, J. Chen, and A. Yoshida, "Enhanced expressions of glucose-6-phosphate dehydrogenase and cytosolic aldehyde dehydrogenase and elevation of reduced glutathione level in cyclophosphamide-resistant human leukemia cells,"

Blood Cells, Molecules and Diseases, vol. 24, no. 2, pp. 231–238, 1998.

[3] J. S. Moreb, "Aldehyde dehydrogenase as a marker for stem cells," *Current Stem Cell Research & Therapy*, vol. 3, no. 4, pp. 237–246, 2008.

[4] C. Chen, Y. Wei, M. Hummel et al., "Evidence for epithelial-mesenchymal transition in cancer stem cells of head and neck squamous cell carcinoma," *PLoS One*, vol. 6, no. 1, Article ID e16466, 2011.

[5] J. Li and B. P. Zhou, "Activation of β-catenin and Akt pathways by Twist are critical for the maintenance of EMT associated cancer stem cell-like characters," *BMC Cancer*, vol. 11, p. 49, 2011.

[6] Y. Chen, D. J. Orlicky, A. Matsumoto, S. Singh, D. C. Thompson, and V. Vasiliou, "Aldehyde dehydrogenase 1B1 (ALDH1B1) is a potential biomarker for human colon cancer," *Biochemical and Biophysical Research Communications*, vol. 405, no. 2, pp. 173–179, 2011.

[7] C. van den Hoogen, G. van der Horst, H. Cheung et al., "The aldehyde dehydrogenaseenzyme 7A1 is functionally involved in prostate cancer bone metastasis," *Clinical and Experimental Metastasis*. In press.

[8] K. S. Park, S. Y. Cho, H. Kim, and Y. K. Paik, "Proteomic alterations of the variants of human aldehyde dehydrogenase isozymes correlate with hepatocellular carcinoma," *International Journal of Cancer*, vol. 97, no. 2, pp. 261–265, 2002.

[9] D. Liang and Y. Shi, "Aldehyde dehydrogenase-1 is a specific marker for stem cells in human lungadenocarcinoma," *Medical Oncology*. In press.

[10] J. P. Sullivan, M. Spinola, M. Dodge et al., "Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on notch signaling," *Cancer Research*, vol. 70, no. 23, pp. 9937–9948, 2010.

[11] J. Feng, Q. Qi, A. Khanna et al., "Aldehyde dehydrogenase 1 is a tumor stem cell-Associated marker in lung cancer," *Molecular Cancer Research*, vol. 7, no. 3, pp. 330–338, 2009.

[12] J. S. Moreb, H. V. Baker, L. J. Chang et al., "ALDH isozymes downregulation affects cell growth, cell motility and gene expression in lung cancer cells," *Molecular Cancer*, vol. 7, article 87, 2008.

[13] S. A. Marchitti, D. J. Orlicky, C. Brocker, and V. Vasiliou, "Aldehyde dehydrogenase 3B1 (ALDH3B1): immunohistochemical tissue distribution and cellular-specific localization in normal and cancerous human tissues," *Journal of Histochemistry and Cytochemistry*, vol. 58, no. 9, pp. 765–783, 2010.

[14] S. E. Ong and M. Mann, "A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC)," *Nature Protocols*, vol. 1, no. 6, pp. 2650–2660, 2006.

[15] V. Faca, S. J. Pitteri, L. Newcomb et al., "Contribution of protein fractionation to depth of analysis of the serum and plasma proteomes," *Journal of Proteome Research*, vol. 6, no. 9, pp. 3558–3565, 2007.

[16] A. Rauch, M. Bellew, J. Eng et al., "Computational proteomics analysis system (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments," *Journal of Proteome Research*, vol. 5, no. 1, pp. 112–121, 2006.

[17] A. Keller, A. I. Nesvizhskii, E. Kolker, and R. Aebersold, "Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search," *Analytical Chemistry*, vol. 74, no. 20, pp. 5383–5392, 2002.

[18] A. I. Nesvizhskii, A. Keller, E. Kolker, and R. Aebersold, "A statistical model for identifying proteins by tandem mass spectrometry," *Analytical Chemistry*, vol. 75, no. 17, pp. 4646–4658, 2003.

[19] H. Choi and A. I. Nesvizhskii, "False discovery rates and related statistical concepts in mass spectrometry-based proteomics," *Journal of Proteome Research*, vol. 7, no. 1, pp. 47–50, 2008.

[20] H. Liu, R. G. Sadygov, and J. R. Yates III, "A model for random sampling and estimation of relative protein abundance in shotgun proteomics," *Analytical Chemistry*, vol. 76, no. 14, pp. 4193–4201, 2004.

[21] F. M. Brichory, D. E. Misek, A. M. Yim et al., "An immune response manifested by the common occurrence of annexins I and II autoantibodies and high circulating levels of IL-6 in lung cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 17, pp. 9824–9829, 2001.

[22] Z. A. Rasheed, J. Yang, Q. Wang et al., "Prognostic significance of tumorigenic cells with mesenchymal features in pancreatic adenocarcinoma," *Journal of the National Cancer Institute*, vol. 102, no. 5, pp. 340–351, 2010.

[23] C. Visus, D. Ito, A. Amoscato et al., "Identification of human aldehyde dehydrogenase 1 family member a1 as a novel CD8+ T-cell-defined tumor antigen in squamous cell carcinoma of the head and neck," *Cancer Research*, vol. 67, no. 21, pp. 10538–10545, 2007.

[24] T. Nakanishi, T. Takeuchi, K. Ueda, H. Murao, and A. Shimizu, "Detection of eight antibodies in cancer patients' sera against proteins derived from the adenocarcinoma A549 cell line using proteomics-based analysis," *Journal of Chromatography B*, vol. 838, no. 1, pp. 15–20, 2006.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 413646, 9 pages
doi:10.1155/2011/413646

Review Article

The Use of Protein-Based Biomarkers for the Diagnosis of Cystic Tumors of the Pancreas

Richard S. Kwon[1] and Diane M. Simeone[2]

[1] *Department of Internal Medicine, University of Michigan, 1500 E. Medical Center Drive, Taubman 3912, Ann Arbor, MI 48109-5362, USA*
[2] *Departments of Surgery and Molecular and Integrative Physiology, University of Michigan, 1500 E. Medical Center Drive, Taubman 2210B, Ann Arbor, MI 48109-5343, USA*

Correspondence should be addressed to Diane M. Simeone, simeone@umich.edu

Received 10 June 2011; Accepted 15 August 2011

Academic Editor: David E. Misek

Copyright © 2011 R. S. Kwon and D. M. Simeone. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Proteomics is a powerful method used to identify, characterize, and quantify proteins within biologic samples. Pancreatic cystic neoplasms are a common clinical entity and represent a diagnostic and management challenge due to difficulties in accurately diagnosing cystic lesions with malignant potential and assessing the risk of malignant degeneration. Currently, cytology and other biomarkers in cyst fluid have had limited success in accurately distinguishing both the type of cystic neoplasm and the presence of malignancy. Emerging data suggests that the use of protein-based biomarkers may have greater utility in helping clinicians correctly diagnose the type of cyst and to identify which cystic neoplasms are malignant. Several candidate proteins have been identified within pancreatic cystic neoplasms as potential biomarkers. Future studies will be needed to validate these findings and move these biomarkers into the clinical setting.

1. Background

Pancreatic cysts are increasing in prevalence as cross-sectional imaging has become widely utilized. In recent population-based studies using magnetic resonance imaging (MRI) [1–4] and computerized tomography (CT) scans [5, 6], the estimated prevalence of cystic lesions ranges from 2.6% to as high as 44.7%. An autopsy study of 300 patients from Japan reported a prevalence of 24.3% [7]. Not surprisingly, the number of evaluations for these lesions is increasing [8, 9]. Management of these increasingly prevalent lesions can utilize a significant amount of health care resources in the form of diagnostic studies and surgical resections. Therefore, developing an accurate and cost-effective diagnostic test to assist in patient management is a clear priority.

This paper will focus on cystic neoplasms which are distinguished by the presence of mucinous or nonmucinous epithelium. Ninety percent of all cystic neoplasms are comprised of serous cystadenomas (SCAs), a non-mucinous lesion, as well as mucin-producing cystic tumors comprised of mucinous cystic neoplasms (MCNs) and intraductal papillary mucinous neoplasms (IPMNs) [10]. Rare cystic neoplasms include solid pseudopapillary lesions, lymphoepithelial cysts, and cystic degeneration of pancreatic ductal adenocarcinoma or neuroendocrine tumors (see Table 1). The most common nonneoplastic cyst is a pancreatic pseudocyst which is associated with acute pancreatitis and has no epithelium.

SCAs are characterized by bland cuboidal glycogen-rich epithelium. They tend to occur predominantly in women (87%) with a median age in the early 50s [11, 12]. SCAs are usually comprised of microcystic components, with a classic honeycomb appearance, though they can be macrocystic in appearance [12]. Up to 30% of these lesions will have a characteristic central scar. Their malignant potential is considered so low that they are generally not resected unless symptomatic.

MCNs have columnar mucinous epithelium with surrounding ovarian stroma (defined as hypercellular spindle

Table 1: Types of pancreatic cystic neoplasms.

Mucinous
Mucinous cystic neoplasm
Intraductal papillary mucinous neoplasm (IPMN)
Non-mucinous
Serous cystadenoma
Solid pseudopapillary neoplasm
Lymphoepithelial cysts
Cystic degeneration of ductal adenocarcinoma
Cystic neuroendocrine tumor
Cystic acinar cell carcinoma

cell bundles that lay just beneath the epithelium and usually show positive staining for estrogen and progesterone receptors) [13, 14] and typically present as large solitary macrocystic lesions in the body or tail of the pancreas [15]. They are thought to be separate from the main pancreatic duct but can be connected in up to 20% of cases [14]. They occur almost exclusively in women (95–98%) [14] during the fourth or fifth decades of life [15]. The rate of malignancy ranges from 6–30% at the time of resection [15–17]. Risk factors for malignancy include older age, the presence of a mural nodule with the cyst, and cyst size >4 cm [14, 15, 17]. The five-year survival is 100% in patients with benign disease and 60% in those patients that develop invasive carcinoma [17, 18]. Recurrence appears to occur only with invasive disease [15, 19].

IPMNs are characterized by columnar papillary mucinous epithelium that involves the main pancreatic duct, the side branch ducts (SB-IPMN) or both (mixed IPMN). IPMNs tend to occur more frequently in the head of the pancreas than the body and tail. Males have a slightly higher predilection for IPMNs than females do. The risk of malignancy is much higher in main duct disease (mean 70%) than side-branch disease (mean 25%) [20]. The reported overall 5-year survival rate for resected noninvasive IPMN ranges from 77 to 100%, whereas 5-year survival rate for invasive IPMN ranges from 30% to 75% [21].

Recently, IPMNs have been categorized into four epithelial subtypes—gastric, intestinal, pancreaticobiliary, and oncocytic, based upon cell morphology and expression patterns of glycoproteins containing mucin (MUC) [22]. Combinations of epithelium subtypes may be present within an individual lesion and therefore each IPMN is classified by the dominant component [22]. Based upon recent studies, these categories may explain the clinical behavior of the different IPMN subtypes [23, 24]. Gastric-type IPMNs primarily are located in the side branches and express MUC5AC but not MUC1 or MUC2 [23]. These rarely undergo malignant transformation [23], but if they do, they develop into tubular adenocarcinomas, which have a survival that is almost as poor as ductal adenocarcinoma [24]. Intestinal-type IPMNs are located mainly in the main pancreatic duct and express both MUC2 and MUC5AC [22]. These have a high frequency of malignant transformation to colloid carcinoma, which has a better prognosis than ductal adenocarcinoma [23, 24]. When compared to the other subtypes of IPMN, pancreaticobiliary-type IPMN is noted to occur more frequently in women than men and at a later age (mean 69.2 y versus 60.3–65.6 y) [23]. By immunohistochemistry, they express MUC1 and MUC5AC [22, 23] and may progress to tubular adenocarcinoma [23, 24]. The oncocytic-type IPMNs also express MUC5AC and MUC1 [22]. These tend to develop in people at a younger age than the other IPMN subtypes [23]. While these can progress to malignancy (oncocytic adenocarcinoma), they tend to be noninvasive [23] and have better survival than ductal adenocarcinoma [23, 24].

To date, the subtypes of mucinous tumors (MCN and IPMN) can only reliably be distinguished by surgical pathology. There are no presurgical tests that distinguish these cyst types with a high level of accuracy. Moreover, the triggers or markers for malignant transformation are unknown and the timeline to transformation remains unclear. As such, our knowledge of the natural history of these lesions is still limited.

In 2006, International Consensus Guidelines were developed by a team of experts to define management of cystic mucinous neoplasms [20]. For cystic neoplasms, the decision to undergo surgical resection versus surveillance should be tempered by patient's wishes, comorbidities, life expectancy, and the risk of malignancy versus the risk of surgery. If the patient is an appropriate surgical candidate, the guidelines suggest resection of all MCNs and any IPMN which involve the main duct, or side branch IPMN lesions that are symptomatic, have a solid component, or are >3 cm in size [20]. They recommend yearly surveillance for lesions <10 mm, and surveillance every 6–12 months for lesions 10–20 mm and every 3–6 months for lesions >20 mm. The surveillance interval can be lengthened after two years of no change [20]. A retrospective study of 147 patients demonstrated that the algorithm proposed by the guidelines had a sensitivity and negative predictive value of 100% but a specificity of 23% [25]. Other studies have shown similar results [26–30]. The algorithm therefore seems reasonably sensitive to identify those who do not need surgery, but given the low specificity, there remains a fairly high resection rate for patients with benign disease [21].

The clinical challenges of managing patients with pancreatic cystic neoplasms have several layers of complexity. First, one must differentiate between mucinous and non-mucinous cysts. This differentiation is important because their clinical management is different. Non-mucinous lesions generally do not require follow-up, whereas because of their premalignant potential, mucinous neoplasms are either resected or monitored in a surveillance program. Second, once a mucinous lesion is identified, one should distinguish between MCN and IPMN (in particular focal SB-IPMN) since the former should be resected whereas the latter can be monitored. This can be a difficult task in part because MCNs occasionally have communication with the main duct which is sometimes difficult to accurately identify by cross-sectional imaging or EUS [10, 20]. Moreover, there is no preoperative test that can identify the characteristic ovarian

stroma of MCNs [10], and even on surgical pathologic analysis, the stroma may not be uniformly present, particularly with malignant transformation [31]. Third, amongst the mucinous lesions, one must differentiate between those that have high-grade dysplasia and cancer and those that are benign in order to appropriately refer those patients to surgery who would most benefit from resection. This differentiation would allow more selective recommendation of surgical resection for those who truly need it. Finally, given that the natural history of these lesions remains to be clarified, one must be able to identify those lesions that will go onto malignant transformation. To date, no single test or tests adequately addresses these challenges and as such, a biomarker or set of biomarkers are needed in order to address all four of these challenges.

Current tests have limited ability to distinguish between mucinous and non-mucinous lesions or to identify malignant cysts (Figure 1). Cross-sectional radiologic imaging is limited in its ability to distinguish between the different types of cysts. The accuracy of CT and MRI to determine the correct histology ranges from 40–60% [32, 33]. New advances in CT and MRI technology that provide more detailed images may account for a modest increase in accuracy up to 84% [34]. Morphology by endoscopic ultrasound is also limited in its ability to distinguish between types of cystic tumors, with a sensitivity and specificity of 56% and 45%, respectively [35]. Furthermore, the accuracy of EUS morphology is limited by a lack of interobserver agreement [36].

Endoscopic ultrasound allows for fine needle aspiration of cyst fluid analysis. The tests of choice for diagnostic evaluation include cytology and carcinoembryonic antigen (CEA). Fluid cytology can be limited due to luminal contamination, highly variable amounts of extracellular mucin, and scant cellularity within the cyst [37]. The overall accuracy of cytology in identifying mucinous lesions is around 50% [35, 38]. There are no cytological findings which currently distinguish MCN from IPMN [39, 40]. Cytology has a specificity that approaches 100% but lower sensitivity in identifying the presence of malignancy [38, 41, 42]. To date, cyst fluid CEA remains the most accurate test to distinguish mucinous from non-mucinous cysts [35]. A prospective, multicenter study determined that a CEA >192 ng/mL had a 75% sensitivity and 85% specificity in distinguishing between mucinous and non-mucinous cysts. Its overall accuracy of 79% was higher than morphology, cytology, and other tumor markers previously identified in pancreatic cysts such as CA72-4 [43], CA19-9 [44], and CA 15-3 [35, 45]. However, fluid CEA is limited by the fact that there is broad overlap between types of lesions. In addition, this test is unable to distinguish between types of mucinous cysts, nor is there any correlation between elevated concentrations and risk of malignancy [46].

2. Biomarkers for Cystic Neoplasms

As a result of these current limitations, there has been considerable interest in finding other biomarkers that can better distinguish mucinous lesions and identify patients with tumors of higher malignant potential (i.e., with high-grade dysplasia or carcinoma) who would benefit most from surgical resection. The general approach that has been taken is to aspirate cyst fluid and use a variety of techniques to try to create highly sensitive and specific assays to identify subjects with high grade dysplasia or frank malignancy [47]. Pancreatic cyst fluid would appear to be an ideal source for a biomarker development due to its relative ease at obtainment by endoscopic ultrasound, which has a low complication rate when performed by an experienced endoscopist, and the presumed localization of relevant biological material from cyst epithelium [48].

There is a considerable interest in genetic material within cyst fluid and its potential as to serve as biomarkers. DNA mutations, such as K-ras, and allelic loss amplitude of a proprietary list of specific pancreatic cancer-related genes within cyst fluid have been studied as surrogate markers for mucinous and malignant cysts [49]. In a multicenter study of 113 patients, the authors reported that the presence of a cyst fluid K-ras mutation had a high specificity of 95% but low sensitivity of 45% for diagnosing mucinous cysts. The combination of a K-ras mutation followed by allelic loss showed a high specificity of 96% but a low sensitivity of 37% in diagnosing malignant cysts [49]. Subsequent studies have reported mixed correlation between these DNA mutations and final surgical pathology [50–52]. The added benefit over existing tests remains unclear, and as such, the role for DNA analysis will need to be clarified [46, 53].

MicroRNAs (miRNAs) are small (22 nucleotides) non-coding RNAs that regulate the stability and translation of mRNA transcripts. Deregulation of miRNA expression has been identified in several human cancers, including pancreatic adenocarcinoma [54–59]. Using a panel of 12 miRNAs that are upregulated in pancreatic cancer, Habbe et al. described the identification of abnormal miRNA expression in surgical histology from 15 noninvasive IPMNs compared to normal pancreatic tissue [60]. Moreover, they established the feasibility of identifying miRNA in pancreatic juice. The two miRNAs with the highest expression, miR-21 and miR-155, both of which have been shown in laboratory studies to inhibit apoptosis, had higher expression in the IPMNs (6 of 10, 60%) than normal controls (0 of 5), though this did not reach statistical significance due to a small number of samples. There was also an increased frequency of miR-155 expression in IPMN lesions with pancreaticobiliary and intestinal epithelium. Further studies will be required to validate these findings and to define the true utility of using miRNAs as biomarkers in pancreatic cyst fluid.

3. Protein-Based Biomarker Strategies

Another strategy for biomarker development is to identify specific proteins already known to be involved in pancreatic cancer. One group used multiplex assays with 54 proteins associated with pancreatic cancer to demonstrate differential protein expression between noninvasive IPMN and SCA (34 out of 51 proteins, 67%) and noninvasive MCN and SCA

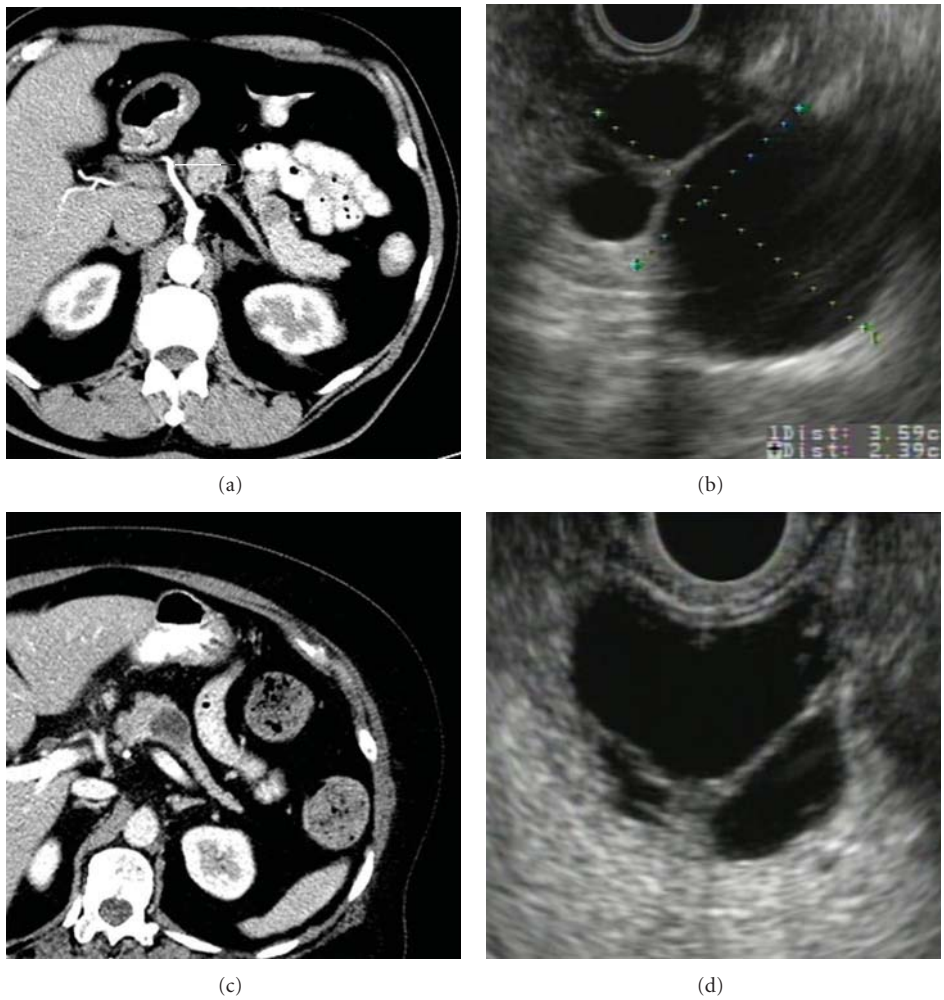


Figure 1: These images highlight the limitations of cross-sectional imaging and endoscopic ultrasound (EUS) in differentiating cyst types. By CT ((a)+(c)) and by EUS ((b)+(d)), the two cysts look very similar. The cyst in (a) and (b) was a macrocystic serous cystadenoma and the cyst in (c) and (d) was a mucinous cystadenoma. The histology of both was confirmed by surgical resection.

(13 out of 51 proteins, 25%) [61]. When using a panel of 14 proteins, the accuracy of distinguishing IPMNs from SCAs was 92% [61]. Another group took a more specific approach and examined the role of Prostaglandin E(2), which they had shown to have increased expression in pancreatic cancer tissue over normal pancreatic tissue [62], in distinguishing between types of mucinous cysts. From fluid obtained from 58 resected cystic lesions, they demonstrated that cyst fluid PGE(2) concentrations were greater in IPMNs versus MCNs (2.2 ± 0.6 versus 0.2 ± 0.1 pg/mol, $P < 0.05$) and that the mean level of PGE(2) increased with the degree of dysplasia in IPMN lesions [63]. However, there was noted to be an overlap in PGE(2) concentrations in benign MCNs ($n = 11$) and SCAs ($9n = 5$), thereby limiting the utility of this biomarker in the clinical setting. These studies demonstrate that targeting proteins associated with pancreatic cancer show potential in identifying appropriate biomarkers for cystic lesions and will require further investigation and validation.

As described above, immunohistochemistry of surgical pathology has demonstrated differential mucin profiles in IPMN [64, 65] and it would seem logical to expect mucin profiles in cyst fluid to identify lesions at risk for malignancy. A recent study has demonstrated differential mucin expression in cyst fluid from 40 surgically resected IPMNs using enzyme-linked immunosorbent assays [66]. Patients with high grade dysplasia or carcinoma ($n = 19$) were categorized as "high risk." Cyst fluids MUC2 and MUC4 were elevated in high-risk patients as compared to low risk patients (10 ± 3.0 ng/mL versus 4.4 ± 1.2 ng/mL, $P = 0.03$; 20.06 ± 10.6 ng/mL versus 4.5 ± 1.4 ng/mL, $P = 0.03$, resp.). There was no difference in MUC1 or MUC5AC concentrations between the two groups. Cysts with gastric epithelium ($n = 23$) had statistically significant lower expressions of MUC2, MUC4, and MUC5AC compared to pancreatic cystic tumors without gastric epithelium. Cysts with intestinal epithelium ($n = 8$) had statistically significant higher elevations of MUC2 compared to nonintestinal cysts and a trend towards higher MUC4 concentrations. There was no discernible difference in MUC concentrations in pancreaticobiliary epithelium cysts compared to those without pancreaticobiliary epithelium. These findings have

not yet been validated; however, this study highlights the potential for risk stratification based upon MUC expression in cyst fluid.

Other proteins previously identified in pancreatic cancer have also been studied to more accurately identify IPMN harboring malignancy. Mutant K-*ras* protein has been identified in cyst fluid using mass spectrometry [67]. The expression of Plectin-1, a marker found to be increased in ductal adenocarcinoma, has also been identified in fluid from malignant mucinous cysts [68]. Cytokine IL-1β is markedly elevated in high-risk patients compared to low risk patients in a small number of IPMNs [69]. These findings have yet to be validated and are thus far experimental in nature.

4. Proteomics of Cyst Fluid

Proteomics is an attractive method for identifying proteins within the cyst fluid which can differentiate mucinous cysts or identify malignancy with greater accuracy. Traditionally, proteins are separated by two-dimensional gel electrophoresis with subsequent mass spectrometric identification of protein spots or by protein digestion and mass spectrometric identification of peptide sequences. Proteomics overcomes the shortcoming of using DNA or mRNA analysis, whose changes may not reflect actual protein expression [70, 71] or include posttranslational modifications. Furthermore, proteomic analysis may provide information on the pathogenesis of these lesions.

The challenge of using proteomics is the complexity of the proteome. The method can identify a large number of proteins but interpretation of the results may be clouded by the signal of the most abundant proteins, and thus proteins present in very small concentrations may not be easy to identify. Within the pancreatic cyst fluid itself, the vast variety of proteins are a reflection of cyst epithelium, luminal contamination (if fluid is obtained by transgastric or transduodenal aspiration), plasma proteins, mucus, and possibly pancreatic enzymes, if there is a connection of the cyst to the pancreatic ductal system. Protein concentration yield may be subject to degradation by endogenous peptidases [72] or post-translational modifications [73]. Proteomics has been used to successfully identify potential biomarkers in the tissue [74, 75], serum [76], and pancreatic juice [77, 78] of patients with pancreatic ductal adenocarcinoma. A method to perform proteomic analysis using paraffin-embedded archival slides of a noninvasive IPMN carcinoma-in-situ has also been described [79].

Interest in using proteomics in pancreatic cyst fluid analysis is growing. The feasibility of proteomic analysis of pancreatic cyst fluid was established by Scarlett et al. [80]. In this proof of concept study, cyst fluids from 10 patients (including 3 ductal adenocarcinomas, 2 mucinous cystadenoma, and 1 IPMN) were analyzed using SELDI-TOF mass spectrometry. Reproducible protein profiles were demonstrated amongst the adenocarcinoma patients with differential expression in twelve protein peaks identified. These findings suggest that proteomics is a viable method for identifying potential biomarkers within cyst fluid.

Two recent studies advanced the use of proteomic analysis to identify biomarkers in cyst fluid. Ke et al. used small volumes (<40 μL) from EUS fine needle aspirates and grouped patients according to their cytology results ((a) benign: no evidence of benign mucinous epithelium, atypical cells or carcinoma; (b) benign mucinous epithelium; (c) atypical/suspicious; (d) malignant) [81]. Fluid was analyzed using MALDI-TOF mass spectrometry with LC/MS/MS protein identification, 2D gel electrophoresis, or GeLC/MS/MS (tryptic digestion of proteins fractionated by SDS-PAGE and identified by LC/MS/MS). The first two techniques proved to be unsatisfactory, presumably from endogenous peptidases which splintered native proteins in numerous locations [81]. Mass spectrometry yielded homologs within three families of proteins associated with pancreatic cancer, including mucins, CEACAMs, and S100s. The authors conclude that LC/MS/MS mass spectrometry provides useful information on biomarkers within cyst fluid using small volumes of fluid.

The same technique was used by Cuoghi et al. in a study of 8 patients who underwent surgical resection for symptomatic pancreatic neoplasms. Fluid was aspirated directly from the surgical specimens, thereby avoiding potential gastrointestinal luminal contamination. Proteins were separated by SDS-PAGE and then analyzed by LC/MS/MS. The total number of proteins identified in the cyst samples ranged from 220 to 727. They identified 38 proteins unique to neuroendocrine tumors, 18 unique to serous cystadenomas, 92 unique to MCNs, and 29 unique to IPMNs. Analysis of known proteins revealed that several proteins identified in the mucinous lesions (MCNs and IPMNs) were previously reported to be upregulated pancreatic cancer-associated proteins. The findings were confirmed by immunohistochemistry for two of the identified proteins, olfactomedin-4 (OLFM4) and the cell surface glycoprotein MUC18. Clearly, proteomics shows great promise in identifying potential biomarkers (see Table 2). Further studies and refinement of technique will hopefully yield reliable candidate biomarkers that can be validated in clinical studies.

5. Glycoproteomics

Glycoproteomics specifically examines carbohydrate modification or glycosylation of proteins. Aberrant glycosylation is a hallmark for tumorigenesis and tumor progression and not surprisingly, many previously identified biomarkers are glycoproteins. The advantage of glycoproteomics is the focused isolation of glycoproteins by specific binding of glycosylation sites. This specificity reduces the complexity of sample protein populations. As such, this method significantly increases the detection sensitivity for low abundance proteins [84]. Analytic approaches have been broadly categorized as glycoprotein-based analysis or glycopeptide-based analysis [84, 85]. The former begins with enrichment of glycoproteins using lectin and separation techniques to enrich the protein fractions. The latter uses glycopeptides that are digested and then deglycosylated. Peptide identification is then performed by mass spectrometry.

Table 2: Potential biomarkers identified to date in pancreatic cyst fluid.

Genetic biomarkers	References
DNA-based	
K-ras	[49]
Allelic loss amplitude	[49]
RNA-based	
miR-21	[60]
miR-155	[60]
Protein-based biomarkers	
Prostaglandin E(2)	[63]
Interleukin-1β	[69]
MUC1	[81, 82]
MUC2	[66]
MUC4	[66]
MUC5AC	[81–83]
MUC5B	[81]
MUC6	[81, 83]
MUC16	[81]
MUC18	[83]
CA 19-9	[82]
Plectin-1	[68]
S100-A6, 8, 9, 11	[81]
CEACAM 1, 5, 6, 7	[81]
BGP-1	[83]
Tspan-8, 27, 28	[83]
CD55	[83]
E-cad	[83]
Glutathione S-transferase P	[83]
Olfactomedin-4	[83]
Prostate stem cell antigen	[83]
Pyruvate kinase isozymes M1/M2	[83]
Ras-related protein Rab-8A	[83]
Rho-related GTP-binding protein RhoC	[83]
Trefoil factor 1,2	[83]
VE-cadherin	[83]
Protein-Z-dependent protease inhibitor	[83]
von Willebrand antigen 2	[83]

Glycoproteomics has already shown promise as a biomarker development tool in pancreatic cancer. A technique using lectin affinity chromatography, liquid separation, and characterization by mass spectrometry was demonstrated in serum of patients with pancreatic cancer [85]. Sialylated plasma protease C1 inhibitor was shown to be downregulated in cancer serum. Downregulation of the N83 glycosylation sites was also observed. Ninety-two individual glycosylation sites with 41 glycoproteins were identified and 202 glycan peaks with 104 unique carbohydrate structures were detected during glycan profiling using different separation techniques. Forty-five oligosaccharides were found altered in pancreatic cancer serum of which 44 were distinct in the cancer sample [85]. Based on these promising results, glycoprotein microarrays have been created as a high throughput tool to differentiate serum samples from patients with pancreatic cancer, from chronic pancreatitis and normal subjects [86, 87].

This approach has been expanded to the use of glycoproteomics in cystic neoplasms. Using a novel antibody-lectin sandwich array (ALSA) that targets glycan moieties on proteins [88], Haab et al. measured protein expression and glycosylation of MUC1, MUC5AC, MUC16, CEA, and other proteins associated with pancreatic cancer in 53 cyst fluid samples from surgically resected lesions (17 MCN, 15 IPMN, 15 SCA, and 9 pseudocysts) [82]. Wheat germ agglutination of MUC5AC was markedly elevated in MCN and IPMN but not SCAs or pseudocysts. CA19-9 could distinguish between MCN and IPMN with a sensitivity and specificity of 82% and 93%, respectively. MUC1 was elevated in serous lesions compared to pseudocysts and mucinous cysts. MUC5AC in combination with CA19-9 (sensitivity 87%, specificity 86%) outperformed fluid CEA (37% sensitivity, 80% specificity) in distinguishing mucinous from nonmucinous cysts. This study shows that glycan variants of proteins within cyst fluid may prove to be useful biomarkers and highlights an area warranting further evaluation. Validation studies are currently in progress. In addition, it remains to be determined if this approach will be useful in separating malignant and non-malignant lesions.

6. Conclusions

The clinical management of pancreatic cystic neoplasms is difficult due to the lack of sufficiently sensitive and specific diagnostic tests to differentiate cyst types and the presence of malignancy. Pancreatic cyst fluid provides an appealing source for improved biomarker development, particularly by proteomic analysis. Preliminary work with cyst fluid glycosylated mucins show promise in distinguishing mucinous from non-mucinous cysts and differentiating types of mucinous cysts. Cyst fluid homologs of mucin, CEACAMs and S100s, and other proteins associated with pancreatic tumorigenesis have been identified as potential biomarkers for malignancy within cyst fluid. These results will all need to be studied and validated in larger more adequately sized test and training sets of pancreatic cyst fluid for full biomarker development. Given that the field is currently limited by the lack of adequate numbers of pancreatic cyst fluid samples for analysis, it will be important that resources for fluid samples are further developed. As we close our gaps in knowledge regarding natural history of mucinous cysts and the relationship between epithelial subtypes and prognosis, biomarkers will likely play a prominent role in the management of cystic neoplasms.

References

[1] K. S. Lee, A. Sekhar, N. M. Rofsky, and I. Pedrosa, "Prevalence of incidental pancreatic cysts in the adult population on MR imaging," *The American Journal of Gastroenterology*, vol. 105, no. 9, pp. 2079–2084, 2010.

[2] K. de Jong, C. Y. Nio, J. J. Hermans et al., "High prevalence of pancreatic cysts detected by screening magnetic resonance

imaging examinations," *Clinical Gastroenterology and Hepatology*, vol. 8, no. 9, pp. 806–811, 2010.

[3] R. Girometti, S. Intini, G. Brondani et al., "Incidental pancreatic cysts on 3D turbo spin echo magnetic resonance cholangiopancreatography: prevalence and relation with clinical and imaging features," *Abdominal Imaging*, vol. 32, no. 2, pp. 196–205, 2010.

[4] X. M. Zhang, D. G. Mitchell, M. Dohke, G. A. Holland, and L. Parker, "Pancreatic cysts: depiction on single-shot fast spin-echo MR images," *Radiology*, vol. 223, no. 2, pp. 547–553, 2002.

[5] T. A. Laffan, K. M. Horton, A. P. Klein et al., "Prevalence of unsuspected pancreatic cysts on MDCT," *The American Journal of Roentgenology*, vol. 191, no. 3, pp. 802–807, 2008.

[6] K. S. Spinelli, T. E. Fromwiller, R. A. Daniel et al., "Cystic pancreatic neoplasms: observe or operate," *Annals of Surgery*, vol. 239, no. 5, pp. 651–657, 2004.

[7] W. Kimura, H. Nagai, A. Kuroda, T. Muto, and Y. Esaki, "Analysis of small cystic lesions of the pancreas," *International Journal of Pancreatology*, vol. 18, no. 3, pp. 197–206, 1995.

[8] P. J. Allen, M. D'Angelica, M. Gonen et al., "A selective approach to the resection of cystic lesions of the pancreas: results from 539 consecutive patients," *Annals of Surgery*, vol. 244, no. 4, pp. 572–579, 2006.

[9] C. Fernández-del Castillo and A. L. Warshaw, "Current management of cystic neoplasms of the pancreas," *Advances in Surgery*, vol. 34, pp. 237–248, 2000.

[10] S. M. Jeurnink, F. P. Vleggaar, and P. D. Siersema, "Overview of the clinical problem: facts and current issues of mucinous cystic neoplasms of the pancreas," *Digestive and Liver Disease*, vol. 40, no. 11, pp. 837–846, 2008.

[11] C. Bassi, R. Salvia, E. Molinari, C. Biasutti, M. Falconi, and P. Pederzoli, "Management of 100 consecutive cases of pancreatic serous cystadenoma: wait for symptoms and see at imaging or vice versa?" *World Journal of Surgery*, vol. 27, no. 3, pp. 319–323, 2003.

[12] B. Khurana, K. J. Mortelé, J. Glickman, S. G. Silverman, and P. R. Ros, "Macrocystic serous adenoma of the pancreas: radiologic-pathologic correlation," *The American Journal of Roentgenology*, vol. 181, no. 1, pp. 119–123, 2003.

[13] A. Izumo, K. Yamaguchi, T. Eguchi et al., "Mucinous cystic tumor of the pancreas: immunohistochemical assessment of 'ovarian-type stroma'," *Oncology Reports*, vol. 10, no. 3, pp. 515–525, 2003.

[14] K. Yamao, A. Yanagisawa, K. Takahashi et al., "Clinicopathological features and prognosis of mucinous cystic neoplasm with ovarian-type stroma: a multi-institutional study of the Japan pancreas society," *Pancreas*, vol. 40, no. 1, pp. 67–71, 2010.

[15] B. K. P. Goh, Y. M. Tan, Y. F. A. Chung et al., "A review of mucinous cystic neoplasms of the pancreas defined by ovarian-type stroma: clinicopathological features of 344 patients," *World Journal of Surgery*, vol. 30, no. 12, pp. 2236–2245, 2006.

[16] R. P. Reddy, T. C. Smyrk, M. Zapiach et al., "Pancreatic mucinous cystic neoplasm defined by ovarian stroma: demographics, clinical features, and prevalence of cancer," *Clinical Gastroenterology and Hepatology*, vol. 2, no. 11, pp. 1026–1031, 2004.

[17] C. Fernández-del Castillo, "Mucinous cystic neoplasms," *Journal of Gastrointestinal Surgery*, vol. 12, no. 3, pp. 411–413, 2008.

[18] S. Crippa, R. Salvia, A. L. Warshaw et al., "Mucinous cystic neoplasm of the pancreas is not an aggressive entity: lessons from 163 resected patients," *Annals of Surgery*, vol. 247, no. 4, pp. 571–579, 2008.

[19] L. D. R. Thompson, R. C. Becker, R. M. Przygodzki, C. F. Adair, and C. S. Heffess, "Mucinous cystic neoplasm (mucinous cystadenocarcinoma of low-grade malignant potential) of the pancreas: a clinicopathologic study of 130 cases," *The American Journal of Surgical Pathology*, vol. 23, no. 1, pp. 1–16, 1999.

[20] M. Tanaka, S. Chari, V. Adsay et al., "International consensus guidelines for management of intraductal papillary mucinous neoplasms and mucinous cystic neoplasms of the pancreas," *Pancreatology*, vol. 6, no. 1-2, pp. 17–32, 2006.

[21] T. Augustin and T. J. VanderMeer, "Intraductal papillary mucinous neoplasm: a clinicopathologic review," *Surgical Clinics of North America*, vol. 90, no. 2, pp. 377–398, 2010.

[22] T. Furukawa, G. Klöppel, N. Volkan Adsay et al., "Classification of types of intraductal papillary-mucinous neoplasm of the pancreas: a consensus study," *Virchows Archiv*, vol. 447, no. 5, pp. 794–799, 2005.

[23] T. Furukawa, T. Hatori, I. Fujita et al., "Prognostic relevance of morphological types of intraductal papillary mucinous neoplasms of the pancreas," *Gut*, vol. 60, no. 4, pp. 509–516, 2010.

[24] M. Mino-Kenudson, C. F. del Castillo, and Y. Baba, "Prognosis of invasive intraductal papillary mucinous neoplasm depends on histological and precursor epithelial subtypes," *Gut*. In press.

[25] M. Pelaez-Luna, S. T. Chari, T. C. Smyrk et al., "Do consensus indications for resection in branch duct intraductal papillary mucinous neoplasm predict malignancy? A study of 147 patients," *The American Journal of Gastroenterology*, vol. 102, no. 8, pp. 1759–1764, 2007.

[26] R. S. Tang, B. Weinberg, D. W. Dawson et al., "Evaluation of the guidelines for management of pancreatic branch-duct intraductal papillary mucinous neoplasm," *Clinical Gastroenterology and Hepatology*, vol. 6, no. 7, pp. 815–819, 2008.

[27] J. R. Rodriguez, R. Salvia, S. Crippa et al., "Branch-duct intraductal papillary mucinous neoplasms: observations in 145 patients who underwent resection," *Gastroenterology*, vol. 133, no. 1, pp. 72–79, quiz 309-10, 2007.

[28] R. Salvia, S. Crippa, M. Falconi et al., "Branch-duct intraductal papillary mucinous neoplasms of the pancreas: to operate or not to operate?" *Gut*, vol. 56, no. 8, pp. 1086–1090, 2007.

[29] S. Tanno, Y. Nakano, T. Nishikawa et al., "Natural history of branch duct intraductal papillary-mucinous neoplasms of the pancreas without mural nodules: Long-term follow-up results," *Gut*, vol. 57, no. 3, pp. 339–343, 2008.

[30] K. Nagai, R. Doi, T. Ito et al., "Single-institution validation of the international consensus guidelines for treatment of branch duct intraductal papillary mucinous neoplasms of the pancreas," *Journal of Hepatobiliary Pancreatic Surgery*, vol. 16, no. 3, pp. 353–358, 2009.

[31] M. Sugiyama and Y. Atomi, "Recent topics in mucinous cystic tumor and intraductal papillary mucinous tumor of the pancreas," *Journal of Hepatobiliary Pancreatic Surgery*, vol. 10, no. 2, pp. 123–124, 2003.

[32] B. C. Visser, B. M. Yeh, A. Qayyum, L. W. Way, C. E. McCulloch, and F. V. Coakley, "Characterization of cystic pancreatic masses: relative accuracy of CT and MRI," *The American Journal of Roentgenology*, vol. 189, no. 3, pp. 648–656, 2007.

[33] W. E. Fisher, S. E. Hodges, V. Yagnik et al., "Accuracy of CT in predicting malignant potential of cystic pancreatic neoplasms," *HPB*, vol. 10, no. 6, pp. 483–490, 2008.

[34] N. I. Sainani, A. Saokar, V. Deshpande, C. Fernández-Del Castillo, P. Hahn, and D. V. Sahani, "Comparative performance of MDCT and MRI with MR cholangiopancreatography in characterizing small pancreatic cysts," *The American Journal of Roentgenology*, vol. 193, no. 3, pp. 722–731, 2009.

[35] W. R. Brugge, K. Lewandrowski, E. Lee-Lewandrowski et al., "Diagnosis of pancreatic cystic neoplasms: a report of the cooperative pancreatic cyst study," *Gastroenterology*, vol. 126, no. 5, pp. 1330–1336, 2004.

[36] N. A. Ahmad, M. L. Kochman, C. Brensinger et al., "Interobserver agreement among endosonographers for the diagnosis of neoplastic versus non-neoplastic pancreatic cystic lesions," *Gastrointestinal Endoscopy*, vol. 58, no. 1, pp. 59–64, 2003.

[37] M. B. Pitman and V. Deshpande, "Endoscopic ultrasound-guided fine needle aspiration cytology of the pancreas: a morphological and multimodal approach to the diagnosis of solid and cystic mass lesions," *Cytopathology*, vol. 18, no. 6, pp. 331–347, 2007.

[38] J. L. Frossard, P. Amouyal, G. Amouyal et al., "Performance of endosonography-guided fine needle aspiration and biopsy in the diagnosis of pancreatic cystic lesions," *The American Journal of Gastroenterology*, vol. 98, no. 7, pp. 1516–1524, 2003.

[39] L. J. Layfield and H. Cramer, "Fine-needle aspiration cytology of intraductal papillary-mucinous tumors: a retrospective analysis," *Diagnostic Cytopathology*, vol. 32, no. 1, pp. 16–20, 2005.

[40] E. B. Stelow, M. W. Stanley, R. H. Bardales et al., "Intraductal papillary-mucinous neoplasm of the pancreas. The findings and limitations of cytologic samples obtained by endoscopic ultrasound-guided fine-needle aspiration," *The American Journal of Clinical Pathology*, vol. 120, no. 3, pp. 398–404, 2003.

[41] F. Maire, A. Couvelard, P. Hammel et al., "Intraductal papillary mucinous tumors of the pancreas: the preoperative value of cytologic and histopathologic diagnosis," *Gastrointestinal Endoscopy*, vol. 58, no. 5, pp. 701–706, 2003.

[42] D. B. Williams, A. V. Sahai, L. Aabakken et al., "Endoscopic ultrasound guided fine needle aspiration biopsy: a large single centre experience," *Gut*, vol. 44, no. 5, pp. 720–726, 1999.

[43] A. J. Alles, A. L. Warshaw, J. F. Southern, C. C. Compton, and K. B. Lewandrowski, "Expression of CA 72-4 (TAG-72) in the fluid contents of pancreatic cysts. A new marker to distinguish malignant pancreatic cystic tumors from benign neoplasms and pseudocysts," *Annals of Surgery*, vol. 219, no. 2, pp. 131–134, 1994.

[44] C. L. H. Snozek, R. C. Mascarenhas, and D. J. O'Kane, "Use of cyst fluid CEA, CA19-9, and amylase for evaluation of pancreatic lesions," *Clinical Biochemistry*, vol. 42, no. 15, pp. 1585–1588, 2009.

[45] D. Rubin, A. L. Warshaw, J. F. Southern, M. Pins, C. C. Compton, and K. B. Lewandrowski, "Expression of CA 15.3 protein in the cyst contents distinguishes benign from malignant pancreatic mucinous cystic neoplasms," *Surgery*, vol. 115, no. 1, pp. 52–55, 1994.

[46] A. Khalid and W. Brugge, "ACG practice guidelines for the diagnosis and management of neoplastic pancreatic cysts," *The American Journal of Gastroenterology*, vol. 102, no. 10, pp. 2339–2349, 2007.

[47] N. B. Kiviat and C. W. Critchlow, "Novel approaches to identification of biomarkers for detection of early stage cancer," *Disease Markers*, vol. 18, no. 2, pp. 73–81, 2002.

[48] W. G. Park, "Screening for pancreatic cancer: what can cyst fluid analysis tell us?" *F1000 Medicine Reports*, vol. 3, p. 3, 2011.

[49] A. Khalid, M. Zahid, S. D. Finkelstein et al., "Pancreatic cyst fluid DNA analysis in evaluating pancreatic cysts: a report of the PANDA study," *Gastrointestinal Endoscopy*, vol. 69, no. 6, pp. 1095–1102, 2009.

[50] M. S. Sawhney, S. Devarajan, P. O'Farrel et al., "Comparison of carcinoembryonic antigen and molecular analysis in pancreatic cyst fluid," *Gastrointestinal Endoscopy*, vol. 69, no. 6, pp. 1106–1110, 2009.

[51] J. Shen, W. R. Brugge, C. J. Dimaio, and M. B. Pitman, "Molecular analysis of pancreatic cyst fluid: a comparative analysis with current practice of diagnosis," *Cancer Cytopathology*, vol. 117, no. 3, pp. 217–227, 2009.

[52] J. Sreenarasimhaiah, L. F. Lara, S. F. Jazrawi, C. C. Barnett, and S. J. Tang, "A comparative analysis of pancreas cyst fluid CEA and histology with DNA mutational analysis in the detection of mucin producing or malignant cysts," *Journal of the Pancreas*, vol. 10, no. 2, pp. 163–168, 2009.

[53] M. A. Anderson, R. S. Kwon, and J. M. Scheiman, "PANDA cyst-fluid analysis: eats, shoots and leaves?" *Gastrointestinal Endoscopy*, vol. 69, no. 6, pp. 1103–1105, 2009.

[54] S. Ali, K. Almhanna, W. Chen, P. A. Philip, and F. H. Sarkar, "Differentially expressed miRNAs in the plasma may provide a molecular signature for aggressive pancreatic cancer," *The American Journal of Translational Research*, vol. 3, no. 1, pp. 28–47, 2011.

[55] J. Hao, S. Zhang, Y. Zhou, X. Hu, and C. Shao, "MicroRNA 483-3p suppresses the expression of DPC4/Smad4 in pancreatic cancer," *FEBS Letters*, vol. 585, no. 1, pp. 207–213, 2011.

[56] J. Hao, S. Zhang, Y. Zhou, C. Liu, X. Hu, and C. Shao, "MicroRNA 421 suppresses DPC4/Smad4 in pancreatic cancer," *Biochemical and Biophysical Research Communications*, vol. 406, no. 4, pp. 552–557, 2011.

[57] N. Ikenaga, K. Ohuchida, K. Mizumoto et al., "MicroRNA-203 expression as a new prognostic marker of pancreatic adenocarcinoma," *Annals of Surgical Oncology*, vol. 17, no. 12, pp. 3120–3128, 2010.

[58] J. K. Park, J. C. Henry, J. Jiang et al., "miR-132 and miR-212 are increased in pancreatic cancer and target the retinoblastoma tumor suppressor," *Biochemical and Biophysical Research Communications*, vol. 406, no. 4, pp. 518–523, 2011.

[59] X. J. Zhang, H. Ye, C. W. Zeng, B. He, H. Zhang, and Y. Q. Chen, "Dysregulation of miR-15a and miR-214 in human pancreatic cancer," *Journal of Hematology and Oncology*, vol. 3, article 46, 2010.

[60] N. Habbe, J. B. M. Koorstra, J. T. Mendell et al., "MicroRNA miR-155 is a biomarker of early pancreatic neoplasia," *Cancer Biology and Therapy*, vol. 8, no. 4, pp. 340–346, 2009.

[61] P. J. Allen, L. X. Qin, L. Tang, D. Klimstra, M. F. Brennan, and A. Lokshin, "Pancreatic cyst fluid protein expression profiling for discriminating between serous cystadenoma and intraductal papillary mucinous neoplasm," *Annals of Surgery*, vol. 250, no. 5, pp. 754–760, 2009.

[62] P. L. Crowell, C. M. Schmidt, M. T. Yip-Schneider, J. J. Savage, D. A. Hertzler, and W. O. Cummings, "Cyclooxygenase-2 expression in hamster and human pancreatic neoplasia," *Neoplasia*, vol. 8, no. 6, pp. 437–445, 2006.

[63] C. M. Schmidt, M. T. Yip-Schneider, M. C. Ralstin et al., "PGE(2) in pancreatic cyst fluid helps differentiate IPMN from MCN and predict IPMN dysplasia," *Journal of Gastrointestinal Surgery*, vol. 12, no. 2, pp. 243–249, 2008.

[64] S. Yonezawa, A. Nakamura, M. Horinouchi, and E. Sato, "The expression of several types of mucin is related to the biological behavior of pancreatic neoplasms," *Journal of Hepatobiliary Pancreatic Surgery*, vol. 9, no. 3, pp. 328–341, 2002.

[65] N. V. Adsay, K. Merati, O. Basturk et al., "Pathologically and biologically distinct types of epithelium in intraductal papillary mucinous neoplasms: delineation of an "intestinal" pathway of carcinogenesis in the pancreas," *The American Journal of Surgical Pathology*, vol. 28, no. 7, pp. 839–848, 2004.

[66] A. V. Maker, N. Katabi, M. Gonen et al., "Pancreatic cyst fluid and serum mucin levels predict dysplasia in intraductal papillary mucinous neoplasms of the pancreas," *Annals of Surgical Oncology*, vol. 18, no. 1, pp. 199–206, 2010.

[67] Q. Wang, R. Chaerkady, J. Wu et al., "Mutant proteins as cancer-specific biomarkers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 6, pp. 2444–2449, 2011.

[68] D. Bausch, M. Mino-Kenudson, C. Fernández-Del Castillo, A. L. Warshaw, K. A. Kelly, and S. P. Thayer, "Plectin-1 is a biomarker of malignant pancreatic intraductal papillary mucinous neoplasms," *Journal of Gastrointestinal Surgery*, vol. 13, no. 11, pp. 1948–1954, 2009.

[69] A. V. Maker, N. Katabi, L. X. Qin et al., "Cyst fluid interleukin-1beta (IL1beta) levels predict the risk of carcinoma in intraductal papillary mucinous neoplasms of the pancreas," *Clinical Cancer Research*, vol. 17, no. 6, pp. 1502–1508, 2011.

[70] S. P. Gygi, Y. Rochon, B. R. Franza, and R. Aebersold, "Correlation between protein and mRNA abundance in yeast," *Molecular and Cellular Biology*, vol. 19, no. 3, pp. 1720–1730, 1999.

[71] Q. Tian, S. B. Stepaniants, M. Mao et al., "Integrated genomic and proteomic analyses of gene expression in Mammalian cells," *Molecular and Cellular Proteomics*, vol. 3, no. 10, pp. 960–969, 2004.

[72] M. E. Forgue-Lafitte, R. Arambam, and J. Bara, "Proteases present in some pancreatic cyst fluids may affect mucin immunoassay by degrading antibodies and antigens," *Pancreas*, vol. 39, no. 7, pp. 1070–1076, 2010.

[73] M. Tyers and M. Mann, "From genomics to proteomics," *Nature*, vol. 422, no. 6928, pp. 193–197, 2003.

[74] A. R. Shekouh, C. C. Thompson, W. Prime et al., "Application of laser capture microdissection combined with two-dimensional electrophoresis for the discovery of differentially regulated proteins in pancreatic ductal adenocarcinoma," *Proteomics*, vol. 3, no. 10, pp. 1988–2001, 2003.

[75] J. Shen, M. D. Person, J. Zhu, J. L. Abbruzzese, and D. Li, "Protein expression profiles in pancreatic adenocarcinoma compared with normal pancreatic tissue and tissue affected by pancreatitis as detected by two-dimensional gel electrophoresis and mass spectrometry," *Cancer Research*, vol. 64, no. 24, pp. 9018–9026, 2004.

[76] J. Koopmann, Z. Zhang, N. White et al., "Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry," *Clinical Cancer Research*, vol. 10, no. 3, pp. 860–868, 2004.

[77] M. Gronborg, J. Bunkenborg, T. Z. Kristiansen et al., "Comprehensive proteomic analysis of human pancreatic juice," *Journal of Proteome Research*, vol. 3, no. 5, pp. 1042–1055, 2004.

[78] C. Rosty, L. Christa, S. Kuzdzal et al., "Identification of hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein I as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology," *Cancer Research*, vol. 62, no. 6, pp. 1868–1875, 2002.

[79] W. Cheung, M. M. Darfler, H. Alvarez et al., "Application of a global proteomic approach to archival precursor lesions: deleted in malignant brain tumors 1 and tissue transglutaminase 2 are upregulated in pancreatic cancer precursors," *Pancreatology*, vol. 8, no. 6, pp. 608–616, 2008.

[80] C. J. Scarlett, J. S. Samra, A. Xue, R. C. Baxter, and R. C. Smith, "Classification of pancreatic cystic lesions using SELDI-TOF mass spectrometry," *ANZ Journal of Surgery*, vol. 77, no. 8, pp. 648–653, 2007.

[81] E. Ke, B. B. Patel, T. Liu et al., "Proteomic analyses of pancreatic cyst fluids," *Pancreas*, vol. 38, no. 2, pp. e33–e42, 2009.

[82] B. B. Haab, A. Porter, T. Yue et al., "Glycosylation variants of mucins and CEACAMs as candidate biomarkers for the diagnosis of pancreatic cystic neoplasms," *Annals of Surgery*, vol. 251, no. 5, pp. 937–945, 2010.

[83] A. Cuoghi, A. Farina, K. Z'graggen et al., "Role of proteomics to differentiate between benign and potentially malignant pancreatic cysts," *Journal of Proteome Research*, vol. 10, no. 5, pp. 2664–2670, 2011.

[84] Y. Tian and H. Zhang, "Glycoproteomics and clinical applications," *Proteomics*, vol. 4, no. 2, pp. 124–132, 2010.

[85] J. Zhao, D. M. Simeone, D. Heidt, M. A. Anderson, and D. M. Lubman, "Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum," *Journal of Proteome Research*, vol. 5, no. 7, pp. 1792–1802, 2006.

[86] J. Zhao, T. H. Patwa, W. Qiu et al., "Glycoprotein microarrays with multi-lectin detection: unique lectin binding patterns as a tool for classifying normal, chronic pancreatitis and pancreatic cancer sera," *Journal of Proteome Research*, vol. 6, no. 5, pp. 1864–1874, 2007.

[87] C. Li, D. M. Simeone, D. E. Brenner et al., "Pancreatic cancer serum detection using a lectin/glyco-antibody array method," *Journal of Proteome Research*, vol. 8, no. 2, pp. 483–492, 2009.

[88] B. B. Haab, "Antibody-lectin sandwich arrays for biomarker and glycobiology studies," *Expert Review of Proteomics*, vol. 7, no. 1, pp. 9–11, 2010.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 628787, 13 pages
doi:10.1155/2011/628787

Research Article

The Application of a Three-Step Proteome Analysis for Identification of New Biomarkers of Pancreatic Cancer

Mayinuer Abulaizi,[1,2] Takeshi Tomonaga,[1,3] Mamoru Satoh,[1,2] Kazuyuki Sogawa,[1,2] Kazuyuki Matsushita,[1,2] Yoshio Kodera,[2,4] Jurat Obul,[1] Shigetsugu Takano,[5] Hideyuki Yoshitomi,[5] Masaru Miyazaki,[5] and Fumio Nomura[1,2]

[1] *Department of Molecular Diagnosis, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan*
[2] *Clinical Proteomics Research Center, Chiba University Hospital, Chiba 260-8677, Japan*
[3] *Laboratory of Proteome Research, National Institute of Biomedical Innovation, Osaka 567-0085, Japan*
[4] *Laboratory of Biomolecular Dynamics, Department of Physics, Kitasato University School of Science, Sagamihara 252-0373, Japan*
[5] *Department of General Surgery, Graduate School of Medicine, Chiba University, 260-8670 Chiba, Japan*

Correspondence should be addressed to Mayinuer Abulaizi, ablizmaynur@gmail.com

Received 31 May 2011; Accepted 2 August 2011

Academic Editor: Tadashi Kondo

Copyright © 2011 Mayinuer Abulaizi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We searched for novel tumor markers of pancreatic cancer by three-step serum proteome analysis. Twelve serum abundant proteins were depleted using immunoaffinity columns followed by fractionation by reverse-phase high-performance liquid chromatography. Proteins in each fraction were separated by two-dimensional gel electrophoresis. Then the gel was stained by Coomassie Brilliant Blue. Protein spots in which the expression levels were significantly different between cancer and normal control were identified by LC-MS/MS. One hundred and two spots were upregulated, and 84 spots were downregulated in serum samples obtained from patients with pancreatic cancers, and 58 proteins were identified by mass spectrometry. These candidate proteins were validated using western blot analysis and enzyme-linked immunosorbent assay (ELISA). As a result of these validation process, we could confirm that the serum levels of apolipoprotein A-IV, vitamin D-binding protein, plasma retinol-binding protein 4, and tetranectin were significantly decreased in patients with pancreatic cancer.

1. Introduction

Pancreatic cancer is one of the most lethal malignancies, with a 5-year survival rate of only 4-5% [1]. The major reasons for the poor prognosis may be late diagnosis and limited therapeutic options; early diagnosis of pancreatic cancer is a pressing clinical problem.

Serum levels of the conventional tumor markers including carcinoembryonic antigen (CEA) and the Lewis blood group carbohydrate antigen (CA19-9) often remain in normal range at early stages of this malignancy [2]. Therefore, search for novel biomarkers of pancreatic cancer is needed.

Recent advances in proteomic technologies have provided promising ways to discover and identify novel biomarkers in various fields of clinical medicine. Although there has been long and uncertain path from marker discovery to clinical utility [3], sophisticated technologies have facilitated the discovery of potential tumor markers with improved sensitivities and specificities for the diagnosis of cancer patients [4]. Also, proteome analysis can lead to biomarkers that may be useful in the prediction of clinical response to anticancer therapy [5].

Surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) is a representative example of a proteomics technique for the high-throughput fingerprinting of serum proteins and peptides and biomarker discovery [6]. Using this technology, we could detect and identify novel diagnostic markers for alcohol abuse [7] and also a new prognostic marker for pancreatic cancer [8].

One of the technical challenges in serum proteome analysis is that serum contains thousands of proteins and peptides

that are present in a large dynamic concentration [9]. Indeed, 22 abundant proteins such as albumin, immunoglobulins, and transferring constitute up to 99% of the protein content of plasma [10]. In proteomic studies searching for low-abundance serum proteins or peptides, depletion of those abundant proteins and further fractionation of samples will be necessary.

We recently conducted a three-step proteome analysis involving removal of 12 abundant proteins and subsequent reversed-phase high-performance liquid chromatography fractionation and one-dimensional electrophoresis: we successfully identified three proteins including YKL-50 as a promising biomarker of sepsis [11].

Proteomics in pancreatic cancer research including serum or plasma biomarker search has been reviewed [12]. A three-step approach as-we used in this study has not been tried in biomarker search for pancreatic cancers before.

In this study, we applied this three-step proteome analysis to find novel biomarkers of pancreatic cancer.

2. Method

2.1. Patients Studied. Serum samples were obtained preoperatively from a total of 32 patients diagnosed with primary invasive pancreatic ductal carcinoma who had surgery at the Department of General Surgery, Chiba University Hospital. Clinical data of 32 patients are summarized in Table 1(a). Serum samples were also obtained from apparently healthy and age-matched subjects who had medical checkup at the Port-square Kashiwado Clinic, Kashiwado Memorial Foundation (Table 1(b)). All samples were frozen by liquid nitrogen and were stored at −80°C until analysis. Written informed consent was obtained from all the subjects. The ethics committee of our institute approved the protocol.

2.2. A Three-Step Serum Proteome Analysis

2.2.1. Immunoaffinity Subtraction of Highly Abundant Proteins from Human Serum. Serum samples obtained from 4 patients with pancreatic cancer (Nos 1∼4 in Table 1(a)) were pooled. Sera obtained from 4 age-matched healthy volunteers were also pooled (Nos 1∼4 in Table 1(b)). As the first step of proteome analysis, the twelve most abundant proteins (albumin, Immunoglobulin G, transferrin, fibrinogen, Immunoglobulin A, Immunoglobulin M, apolipoprotein A-I, apolipoprotein A-II, haptoglobin, α1-acid glycoprotein and α2-macroglobulin) were removed from serum by passage through a commercially available immunoaffinity column, the ProteomeLab IgY12HC LC10 (Beckman coulter, Inc. Fullerton, CA. USA.) Ninety microliters of each pooled sample was subjected to the immunoaffinity subtraction as we previously described [11]. The combined flow-through fractions were concentrated by Vivaspin2 spin concentrators (molecular weight cutoff, 10 kDa, Vivascience, Hannover, Germany) and were stored at −80°C until use.

In addition, sera from 32 patients with pancreatic cancer and 32 healthy volunteers were used for validation. Eight healthy controls and 8 of relatively advanced cases (Nos 1∼8 in Tables 1(a) and 1(b)) were chosen for first validation and 24 of them (Nos 9∼32 in Tables 1(a) and 1(b)) were used for the second validation experiment.

2.3. HPLC Separation of Immunoaffinity-Subtracted Serum Samples. Immunoaffinity-subtracted serum samples prepared as described above were separated by reverse-phase HPLC in an automated HPLC system, the SHISEIDO Nanospace SI-2 (Shiseido Fine Chemicals, Tokyo, Japan) essentially as we described before [11]. A total of 40 fractions were collected at 0.5 min intervals from 19.6 to 39.6 min. Each fraction was immediately lyophilized by a centrifugal vacuum concentrator and stored at −80°C until further analysis.

2.4. Two-Dimensional Gel Electrophoresis. The IEF gels (70 mm length, Inner 2.5 mm diameter and pH ranges from 3 to 10) were prepared as previously described [13, 14]. The lyophilized samples (from fraction 6 to fraction 25) were dissolved with 15 μL sample preparation buffer and proteins were separated by two-dimensional gel electrophoresis with agarose gels in the first dimension as described by Oh-Ishi et al. [13].

2.5. In-Gel Digestion and LC-MS/MS. CBB stained 2-DE images of pooled serum samples obtained from patients with pancreatic cancer were compared with those obtained from healthy volunteers. Differentially expressed protein bands were excised from the gel and were subjected to in-gel tryptic digestion as previously reported [14]. Digested peptides were injected into a trap column: 0.3 × 5 mm L-trap column (Chemicals Evaluation and Research Institute, Saitama, Japan) and an analytical column: 0.1 × 50 mm Monolith column (AMR, Tokyo, Japan), which was attached to a HPLC system (Nanospace SI-2; Shiseido Fine Chemicals, Tokyo, Japan). The flow rate of a mobile phase was 1 μL/min. The solvent composition of the mobile phase was programmed to change in 35 min cycles with varying mixing ratios of solvent A (2% v/v CH_3CN and 0.1% v/v HCOOH) to solvent B (90% v/v CH_3CN and 0.1% v/v HCOOH): 5–50% B 20 min, 50–95% B 1 min, 95% B 3 min, 95–5% B 1 min, 5% B 10 min. Purified peptides were introduced from HPLC to an LTQ-XL (Thermo Scientific, CA, USA), an ion trap mass spectrometer (ITMS), via an attached Pico Tip (New Objective, MA, USA). The MS and MS/MS peptide spectra were measured in a data-dependent manner according to the manufacturer's operating specifications. The Mascot search engine (Matrixscience, London, UK) was used to identify proteins from the mass and tandem mass spectra of peptides. Peptide mass data were matched by searching the Human International Protein Index database (IPI, July 2008, 72079 entries, European Bioinformatics Institute) using the MASCOT engine. The minimum criterion of the probability-based MASCOT/MOWSE score was set with 5% as the significant threshold level. When the candidates had SEQUEST scores lower than 100 or when the SEQEUST score was computed by using fewer than one peptides fragment, we inspected the raw MS and MS/MS spectra of peptides to judge their qualities (see Figures a–f in Supplementary Material available online at doi: 10.1155/2011/628787).

Table 1

(a) Clinical features of pancreatic cancer patients.

No	Gender	Age (years)	UICC-stage	Tumor size (mm)	TP (g/dL)	ALB (g/dL)	Che (U/L)	T-Cho (mg/dL)
1	M	66	III	35	6.7	4.3	270	220
2	M	66	III	39	6.6	4.4	359	198
3	M	79	IV	10	6.2	3.8	179	155
4	M	71	III	18	6.8	4.0	207	210
5	M	65	III	35	7.0	3.7	189	131
6	M	78	III	50	6.6	4.4	359	198
7	M	66	IIA	30	6.3	4.1	267	176
8	M	62	IV	10	5.9	3.6	171	169
9	M	38	IA	10	6.7	4.0	289	179
10	M	50	IB	30	7.3	4.6	408	203
11	M	63	IIA	18	6.6	3.9	260	138
12	M	62	IIA	38	6.0	3.4	124	149
13	M	54	IIA	24	7.0	4.5	198	162
14	M	73	IIA	25	6.7	4.1	242	176
15	F	76	IIA	26	5.0	3.3	162	95
16	M	63	IIB	15	7.1	4.2	175	163
17	M	65	IIB	32	5.8	3.5	227	113
18	M	68	IIB	80	6.8	4.0	316	217
19	M	71	IIB	24	7.5	4.1	221	144
20	M	74	IIB	27	7.0	4.3	319	178
21	F	68	IIB	27	6.5	4.0	223	177
22	M	63	IIB	26	6.9	4.3	262	219
23	F	68	IIB	27	6.1	3.5	339	163
24	M	61	IIB	30	6.2	4.0	200	176
25	F	74	IIB	40	6.6	3.8	221	170
26	M	62	IIB	60	8.3	3.5	140	155
27	F	73	IIB	35	5.9	3.2	130	169
28	F	59	IIB	18	7.2	4.4	356	130
29	M	73	IIB	50	5.7	3.3	127	135
30	F	62	IIB	25	6.8	4.1	255	293
31	F	71	III	25	6.8	4.2	294	159
32	F	78	III	50	5.9	3.3	197	182
Ave ± SD		66 ± 8.6		30.9 ± 15.0	6.6 ± 0.6	3.9 ± 0.4	240.2 ± 75.6	171.9 ± 6.6

UICC: international union against cancer, M: male, F: female. TP: total protein. ALB: albumin. Che: cholinesterase. T-Cho: total cholesterol. Ave: average. SD: standard deviation.
From number 1 to 4 were used for 2-DE, from number 1 to 8 were for first western blot, from number 9 to 32 were for second western blot.

(b) Clinical features of healthy controls.

No	Gender	Age (years)	TP (g/dL)	ALB (g/dL)	Che (U/L)	T-Cho (mg/dL)
1	M	62	7.5	4.7	401	203
2	M	61	7.7	5.1	396	269
3	M	64	7.2	4.4	284	229
4	M	73	7.6	4.7	216	203
5	M	57	8.3	5.0	430	301
6	M	57	6.7	4.8	328	253
7	M	65	7.2	4.8	375	230
8	M	64	7.2	4.6	327	296

(b) Continued.

No	Gender	Age (years)	TP (g/dL)	ALB (g/dL)	Che (U/L)	T-Cho (mg/dL)
9	M	55	6.9	4.6	300	211
10	M	71	7.1	4.7	290	267
11	F	64	8.1	5.4	293	192
12	M	71	7.2	4.5	233	213
13	M	55	7.2	4.5	365	227
14	M	68	7.4	4.4	300	255
15	F	67	7.0	4.5	279	172
16	F	71	7.0	4.5	297	220
17	M	60	6.7	4.2	284	183
18	M	61	7.0	4.7	260	198
19	M	55	6.4	3.9	304	176
20	F	70	7.6	4.7	398	225
21	M	70	7.3	4.6	257	250
22	F	67	7.4	4.5	304	194
23	M	60	7.0	4.4	338	234
24	F	77	8.1	5.0	416	274
25	F	62	7.5	4.6	284	169
26	F	64	7.2	4.5	280	208
27	M	65	7.4	4.6	264	231
28	M	61	7.8	4.3	319	239
29	M	61	7.3	4.6	309	271
30	M	66	7.5	5.0	375	278
31	F	65	7.6	4.5	339	288
32	M	73	7.1	4.4	306	211
Ave ± SD		64.4 ± 5.7	7.3 ± 0.4	4.6 ± 0.3	317 ± 53.4	230.3 ± 36.9

M: male, F: female. TP: total protein. ALB: albumin. Che: cholinesterase. T-Cho: total cholesterol. Ave: average. SD: standard deviation.
From number 1 to 4 were used for 2-0 E, from number 1 to 8 were for first western blot, from number 9 to 32 were for second western blot.

2.6. Western Blot Analysis. Western blotting was performed as we previously described [15].

Briefly, immunoaffinity-subtracted serum samples were separated on SDS-PAGE in 10–20% polyacrylamide gradient gel (DRC, Tokyo, Japan) and were transferred to polyvinylidene fluoride membranes (0.45 μm thickness, Millipore, Bedford, MA) at 10 V for overnight. The following antibodies commercially available were used as primary antibodies; mouse anti-human ApoA-IV antibody (BML Inc., Tokyo, Japan), mouse anti-human GC antibody (LifeSpan, Inc., UK), mouse monoclonal anti-human RBP4 antibody (Abnova.Com., Taipei, Taiwan) and mouse anti-human CLEC3B antibody (BioPorto, Grusbakken 8, DK-2820 Gentofte, Denmark). Antigens on the membrane were detected with enhanced chemiluminescence detection reagents (GE Healthcare). Band intensities of the western blot images were quantified by TotalLab TL12 imaging analysis software (Shimadzu Co., Ltd. Kyoto, Japan) and were presented by arbitrary units.

2.7. Other Procedures. In addition to western blotting, ELISA was conducted in some marker candidates using human ApoA-IV ELISA kit (Millipore, Missouri, USA), GC ELISA kit (immundiagnostik AG, Bensheim), and RBP4 ELISA kit; (R & D systems). Their optical density was measured at 450 nm using a microplate reader (iMark Microplate Reader S/N 10288). Serum levels of CEA and CA19-9 were determined by established commercially available kits.

2.8. Statistical Analysis. Statistical analysis was conducted using KaleidaGraph 4.0 J (Synergy Software, Reading, PA) and IBM SPSS Statistics 18 (SPSS Inc., IL, USA). Significance was defined as $P < 0.05$.

3. Results

3.1. Discovery and Identification of Differentially Expressed Proteins by a Three-Step Proteome Analysis. To discover and identify novel serum markers for pancreatic cancer, we employed a comparative three-step proteome analysis of the pooled serum samples obtained from patients with pancreatic cancer and healthy volunteers. As the first step, 12 abundant proteins were removed by immunosubtractions. The immunoaffinity-subtracted samples were separated by

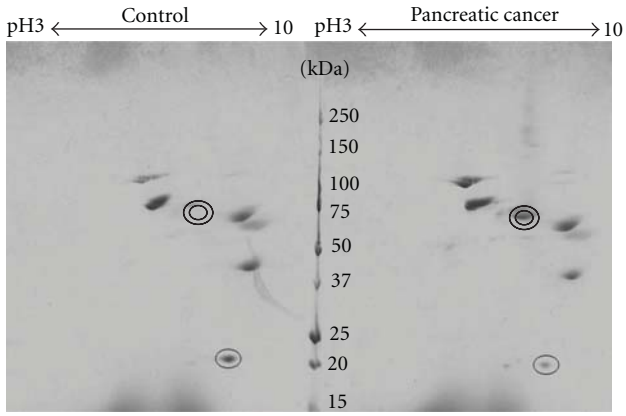


Figure 1: Comparison of 2-DE images of the same fraction of healthy volunteer control and pancreatic cancer patient sample. Electrophoresis was performed on the same gel and at the same condition. Figure 1 is an example of Coomassie blue-stained gel displaying spots from depleted and fractionated serum from control (left, $n = 4$, pooled) and pancreatic cancer patients (right, $n = 4$, pooled). (Fraction number is 10th). Double circles indicate increased spot in pancreatic cancer. Single circle indicates decreased spot in pancreatic cancer.

RP-HPLC, and 20 fractions (fractions Nos 6–25) were subjected to 2-DE. A representative example is shown in Figure 1. By comparing the 2-DE images of the proteins included in the 20 fractions, a total 186 spots were found to be differently expressed. Subsequent LC-MS/MS could identify 100 proteins. Excluding keratins, complements and trypsin, 58 proteins were selected; 37 of them were upregulated and 21 were downregulated (Tables 2(a) and 2(b)).

3.2. Validation of Marker Candidates by Western Blotting. Out of the 58 proteins listed in Tables 2(a) and 2(b), we focused on 19 proteins the alterations of which at serum level have not been studied in detail before, and also antibodies to be used for validation studies are available. Initial validation was conducted using 8 serum samples (nos. 1–8 in Table 1) obtained from relatively advanced cases with pancreatic cancer including the four cases used for the three-step analysis. Western blotting of the 19 proteins indicated in Tables 2(a) and 2(b) revealed that expression of 7 proteins were found to be significantly decreased in patients with pancreatic cancers compared with controls: they were inter-alpha trypsin inhibitor heavy chain H1 (ITIH1), hemopexin precursor (HPX), alpha-1B-glycoprotein precursor (A1BG), apolipoprotein A-IV precursor (ApoA-IV), vitamin D-binding protein precursor (GC), plasma retinol-binding protein precursor (RBP4), and tetranectin (CLEC3B).

We then conducted the second validation study to test whether differential expression of the 7 protein candidates described above is reproducible using another set of serum samples obtained from 24 patients with pancreatic cancers including cases with relatively early stages (nos. 9–32 in Table 1(a)). As shown in Figure 2(a) the expression levels of the four proteins ApoA-IV, GC, RBP4, and CLEC3B were greater in cancer patients than in controls. The differences were statistically significant assessed by densitometry Figure 2(b).

3.3. Validation of Marker Candidates by ELISA. ELISA kits were commercially available for GC, ApoA-IV, and RBP4. Their serum levels were determined in the 15 pairs of serum samples obtained from relatively early stages of patients used for the second validation by western blotting. Serum ApoA-IV levels of patients with pancreatic cancer (107.8 ± 99.9 AU) were significantly lower than those in healthy volunteers (195.2 ± 66.9 AU, $P = 0.008$) Figure 3(a). GC levels were significantly lower in the patient group with pancreatic cancer (25.4 ± 10 UA) when compared with healthy group (34.3 ± 10.3 AU, $P = 0.03$) Figure 3(b). Also, serum RBP levels in the patients (43.0 ± 5.9 AU) were significantly lower than in the controls (50.2 ± 4.2, $P = 0.0004$) Figure 3(c).

3.4. Comparison of the Marker Candidates with CEA and CA19-9. Figure 4 shows the receiver-operating characteristic curve (ROC) analysis for the three marker candidates determined by ELISA and those for CEA and CA19-9. The areas under the curves for ApoA-IV, GC, RBP4, CA19-9, and CEA were 0.79, 0.72, 0.85, 0.88, 0.58, 0.88, 0.89, and 0.89, respectively. Also, AUCs of the combination assay of GC/CA19-9, ApoA-IV/CA19-9, and RBP/CA19-9 were 0.88, 0.89, and 0.89, respectively.

In Tables 3(a) and 3(b), serum levels of ApoA-IV, GC and RBP4 determined by ELISA are listed together with CEA and CA19-9. There were 7 cases in which serum CA19-9 level was not elevated. Out of these 7 cases, ApoA-IV levels were below the lower reference interval value (mean SD) in 2 cases. Also, GC levels were below the lower reference interval value in one case.

4. Discussion

The sequencing of the human genome has opened the door for comprehensive analysis of all the messenger RNA (transcriptome) and proteins (proteome). Messenger RNA concentrations, however, are not necessarily predictive of corresponding protein concentrations. Indeed, a recent report indicates that the sharing rate between cDNA microarray and proteome-based profilings is limited for the identification of candidate biomarkers in renal cell carcinoma [16]. Therefore, proteome analysis is one of the prerequisite for development of novel biomarkers. Proteomic studies in pancreatic cancers have been conducted by many research groups as reviewed [12, 17]. Hwang et al. found by using 2-DE/MS that phosphoglycerate kinase (PGK) 1, a secretable glycolytic enzyme involved in angiogenesis, is overexpressed in serum samples of pancreatic cancer patients, as compared to controls [18]. More recently, using the two-dimensional image-converted analysis of liquid chromatography and mass spectrometry (2DICAL) and a "glyco capturing" through concanavalin A-agarose, Ono et al. identified a novel prolyl-hydroxylation of fibrinogen alpha chain in plasma samples obtained from patients with pancreatic cancers [19].

Table 2

(a) Proteins upregulated in pancreatic cancer.

Protein's name	Experimental mass (Da)	Theoretical mass (Da)	Score[1]	Queries matched[2]	Validation
Histidine-rich glycoprotein precursor	80000	59541	150	3	WB[3]
Plasminogen precursor	100000	90510	331	8	
IGHM protein	50000	52754	105	2	
TF Serotransferrin precursor	75000	77000	601	14	WB
Isoform LMW of Kininogen-1 precursor	70000	47853	242	6	
F2 Prothrombin precursor (Fragment)	90000	69992	486	8	
Alpha-1B-glycoprotein precursor	43000	54239	590	8	WB
Vitronectin precursor	62000	54271	581	8	WB
Hepatocyte growth factor-like protein precursor	85000	80268	284	6	WB
Plasma kallikrein precursor	90000	71323	246	8	
Ceruloplasmin precursor	115000	122128	1452	49	
Isoform 1 of Ficolin-3 precursor	34400	32883	58	1	
Transthyretin precursor	95000	15877	4489	13	
Serum amyloid P-component precursor	24000	25371	2293	12	
Antithrombin III variant	61000	52658	1984	20	
SERPINC1 protein	61000	29074	1147	11	
Carbonic anhydrase 1	30800	28852	679	10	WB
Isoform 1 of C-reactive protein precursor	26000	25023	169	6	
Apolipoprotein A-1	31600	30759	225	4	
Isoform 1 of Inter-alpha-trypsin inhibitor heavy chain H3	138000	75031	529	12	
Leucine-rich alpha-2-glycoprotein precursor	47000	38154	549	8	
Isoform 1 of N-acetylmuramoyl-L-alanine amidase precursor	68900	67957	193	4	
Xaa-Pro dipeptidase	53700	54513	144	3	
Inter-alpha (globulin) inhibitor H4	130500	103261	1242	26	WB
Vitamin K-dependent protein S precursor	84000	75074	136	4	WB
Serpin peptidase inhibitor, clade D (Heparin cofactor),	72000	57034	374	9	
Isoform 1 of Fibronectin precursor	200000	262442	553	14	WB
Alpha-1-antichymotrypsin precursor	62800	50566	3315	33	WB
Kallistatin precursor	59600	48511	123	2	
Plastin-2	72000	70245	215	5	
Corticosteroid-binding globulin precursor	62500	45112	52	1	
Myosin-1	31600	222976	96	2	
Isoform 1 of Serum albumin precursor	73000	69321	261	9	
Cholinesterase precursor	83000	72836	50	1	
AMBP protein precursor	200000	38974	68	1	
Plasma protease C1 inhibitor precursor	87000	55119	1168	15	
Apolipoprotein B-100 precursor	300000	515241	3982	117	

[1] MOWSE score of candidate proteins.
[2] Number of peptide fragments yielding informative MS/MS.
[3] WB: western blot.

(b) Proteins downregulated in pancreatic cancer.

Protein's name	Experimental mass (Da)	Theoretical mass (Da)	Score[1]	Queries matched[2]	Validation
Plasma retinol-binding protein precursor	19000	22995	373	8	WB[3]
Coagulation factor XII precursor	75000	67774	140	5	
Tetranectin precursor	19000	22552	61	1	WB
Hyaluronan-binding protein 2 precursor	68000	62630	195	6	

(b) Continued.

Protein's name	Experimental mass (Da)	Theoretical mass (Da)	Score[1]	Queries matched[2]	Validation
Vitamin D-binding protein precursor	55000	52883	284	14	WB
Hemopexin precursor	75000	51643	635	8	WB
Lumican precursor	100000	38405	120	6	WB
Isoform 1 of Gelsolin precursor	80000	85644	1360	23	WB
Afamin precursor	80000	69024	307	7	
Carboxypeptidase N catalytic chain precursor	49000	52253	244	12	WB
Inter-alpha-trypsin inhibitor heavy chain H1 precursor	200000	101326	812	13	WB
Histone H4	25700	11360	59	1	
JUP JUP protein	100000	81675	90	3	
apolipoprotein A-IV precursor	42000	45371	2188	35	WB
Inter-alpha-trypsin inhibitor heavy chain H2 precursor	200000	106370	1636	26	WB
Pigment epithelium-derived factor precursor	50000	46313	529	11	
Angiotensinogen precursor	56300	53121	986	13	
SERPINF2 protein	58000	55029	75	2	
Actin, cytoplasmic 1	100000	41710	208	5	
Thrombospondin-1 precursor	175000	129300	152	4	
Alpha-2-macroglobulin precursor	180000	163175	559	18	

[1] MOWSE score of candidate proteins.
[2] Number of peptide fragments yielding informative MS/MS data. The minimum significant threshold level of the probability-based MASCOT/MOWSE score was set at 5%.
[3] WB: western blot.

In this study, the three-step procedure was carried out to discover novel markers of pancreatic cancer. The outline of the three-step procedures is shown in Figure 5.

As a first step, serum samples were subjected to antibody-based immunoaffinity column that simultaneously removes 12 abundant serum proteins. The concentrated flow-through was then fractionated using reversed-phase HPLC. Proteins obtained in each HPLC fraction were further separated by 2-DE. A total of 58 differentially expressed proteins were identified. As results of initial validation by western blotting in relatively advanced cases and further validation including the less advanced cases by western blotting, the expression levels of the four proteins ApoA-IV, GC, RBP4, and CLEC3B were greater in cancer patients than in controls. Out of these four proteins, ELISA were available in apolipoprotein A-IV, retinol-binding protein precursor (RBP4), and vitamin D binding protein (GC). Serum levels of these 3 proteins were significantly lower in patients with pancreatic cancer than in healthy volunteers. In ROC analyses, the area under the curves for these three proteins was not significantly greater than that for CA19-9, but it is noteworthy that among the 4 cases of pancreatic cancers in which serum levels of both CEA and CA19-9 were within the reference intervals, at least one of ApoA-IV, RBP4, and GC was found to be decreased in 2 cases, suggesting that these candidate markers could be complementary to the conventional markers in diagnosis of pancreatic cancer.

ApoA-IV is present in human intestinal epithelial cells and is secreted as a chylomicron and VLDL apoprotein [20].

Retinol binding protein 4 (RBP4) is a 21-kDa protein synthesized in the liver and adipose tissue; its major function is to deliver retinol to tissue [21]. Fabris et al. determined serum RBP levels in patients with pancreatic cancer and found that the levels decreased concomitant with zinc and prealbumin levels [22]. Serum zinc levels were not significantly correlated with RBP4 levels in the present study (data not shown).

Vitamin D-binding protein is a plasma protein involved in vitamin D transport and other function. Although diagnostic role of this protein in pancreas cancer has not been reported yet, inhibitory role of vitamin D binding protein-macrophage activating factor (DBP-maf) in pancreatic carcinogenesis has been pointed out [23].

Tetranectin binds to kringle 4 of plasminogen, enhancing the plasminogen activation by tissue-type plasminogen activator in the presence of poly-D-lysine [24] Low serum levels of tetranectin (CLEC3B) are associated with increased risk of second-line chemoresistance in patients with ovarian cancer [25]. Also, in colorectal cancer, significantly shorter survival was found for patients with CLEC3B levels below a cut-off point of compared to patients with levels above [26].

Thus, the results of this study show that four serum proteins, apolipoprotein A-IV, vitamin D binding protein, retinol-binding protein 4, and tetranectin are significantly

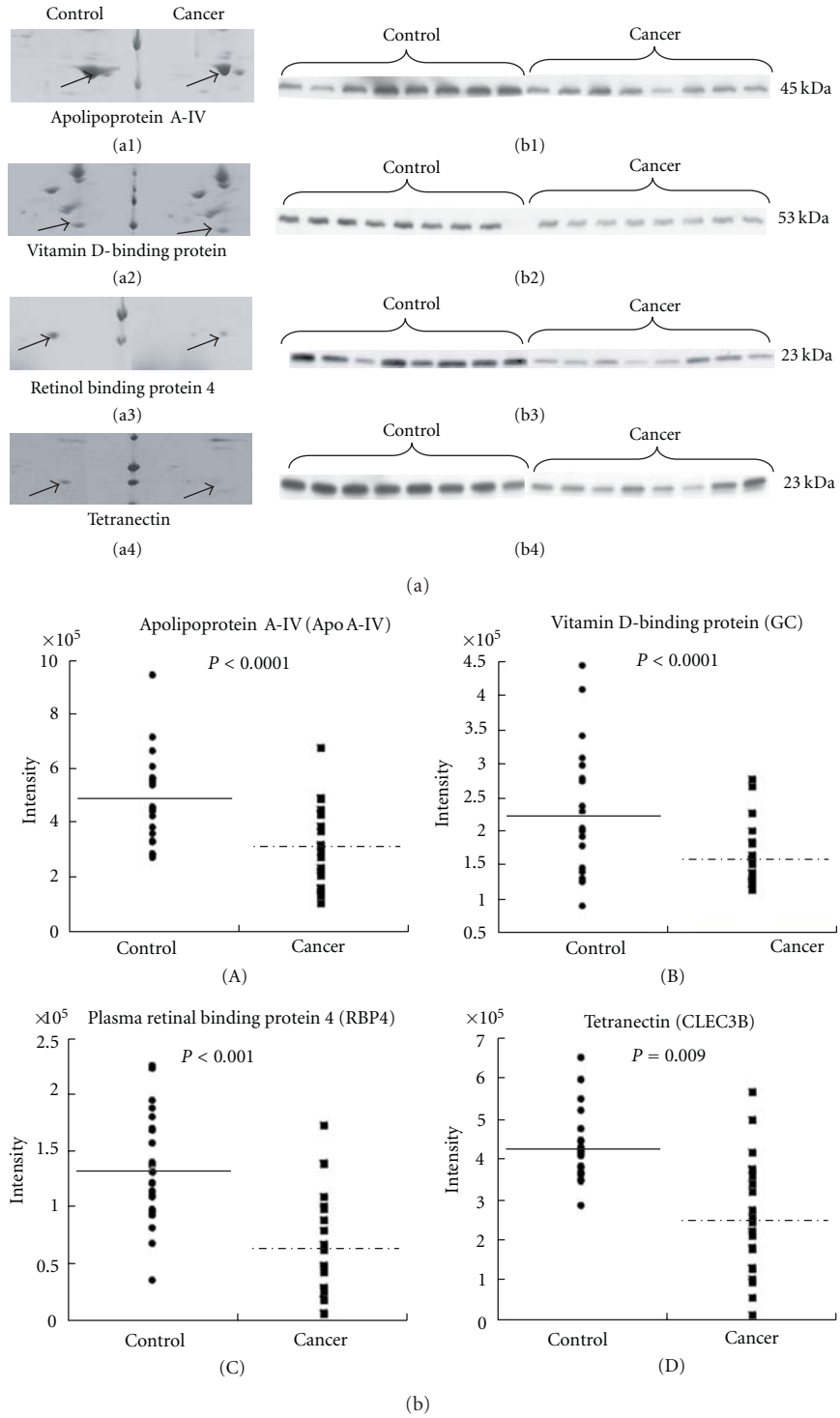


Figure 2: (a) Magnified views of 2-D gel images and western blotting analysis of ApoA-IV, GC, plasma retinol binding protein 4 RBP4, and CLEC3B in serum samples. Coomassie blue-stained 2-D gel images from pooled control and pancreatic cancer displaying the protein spots for ApoA-IV, GC, RBP4, and CLEC3B are shown in left panels (a1), (a2), (a3) and (a4). Western blotting of these four proteins are shown in the right panel (b1), (b2), (b3) and (b4). (b) Quantitation of differentially expressed serum proteins in pancreatic cancer and healthy volunteers by Western blot analysis. Intensities of each band were calculated by TotalLab TL 120 software. Closed circles indicate healthy volunteers and closed squares indicate patients with pancreatic cancer. Significance of the differences were calculated by using Wilcoxon Mann-Whitney test. Panel A: ApoA-IV levels of serum were significantly lower in the depleted sera of pancreatic cancer when compared with the depleted sera of healthy volunteers ($P < 0.0001$). Panel (B, C and D) are for proteins GC, RBP4, and CLEC3B and their serum levels were likewise lower in the pancreatic cancer patients. Their P values are lower than 0.0001, 0.001, and 0.009, respectively.

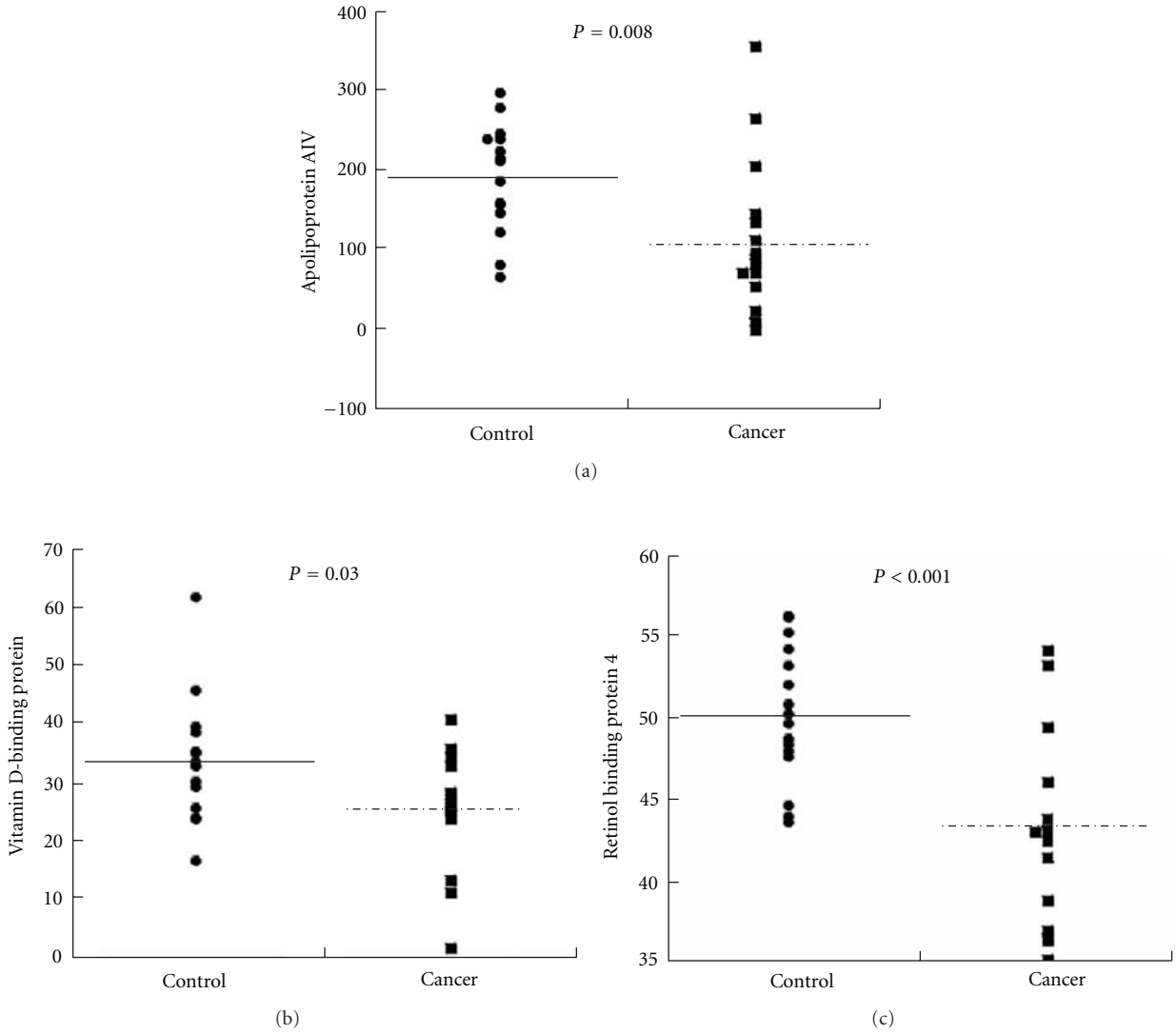


Figure 3: Quantitation of differentially expressed proteins in pancreatic cancer by ELISA. ELISA was performed using human ApoA-IV ELISA kit vitamin D binding protein ELISA kit and RBP4 ELISA kit, in serum samples obtained from 15 patients with pancreatic cancers and 15 control subjects. Analysis was performed by using Wilcoxon Mann-Whitney test. Closed circles indicate control and closed squares indicate cancer. (a): ApoA-IV levels of serum of patient with pancreatic cancer group (107.76 ± 25.8 AU) were lower than those in healthy group (185.27 ± 16.0 AU, $P = 0.01$); (b): GC levels were significantly lower in the patient group with pancreatic cancer (25.35 ± 9.8 UA) when compared with healthy group (34.40 ± 10.2 AU, $P = 0.03$); (c): RBP4 levels were lower in the pancreatic cancer group (42.99 ± 1.5 AU) than healthy group (50.7 ± 1.00 AU, $P < 0.001$).

decreased in patients with pancreatic cancer. It was notable that these changes were observed in some patients in whom conventional tumor markers for this malignancy were not altered.

The reasons why serum levels of these proteins were decreased in pancreatic cancer patients are not clear at the moment. It is unlikely that the alterations were entirely due to malnutrition because serum levels of the 4 proteins were not significantly correlated with their serum albumin levels. It is possible that some negative mediators originated from tumor and/or the cancer-tissue microenvironments were regulating their production. It is unlikely that the alterations were due to biliary obstruction because the extent of the alterations of the four markers were not related to the extent of biliary obstruction (data not shown). Alterations of these four proteins in chronic pancreatitis as well as biliary tract diseases remain to be studied. Also, it remains to be determined whether serum levels of these four proteins are changed in other gastroenterological cancers.

Although exact mechanisms responsible for the reduction remain to be investigated, alterations of serum levels of apolipoprotein A-IV, vitamin D binding protein, tetranectin, and retinol binding protein may have complementary role in diagnosis of pancreas cancer.

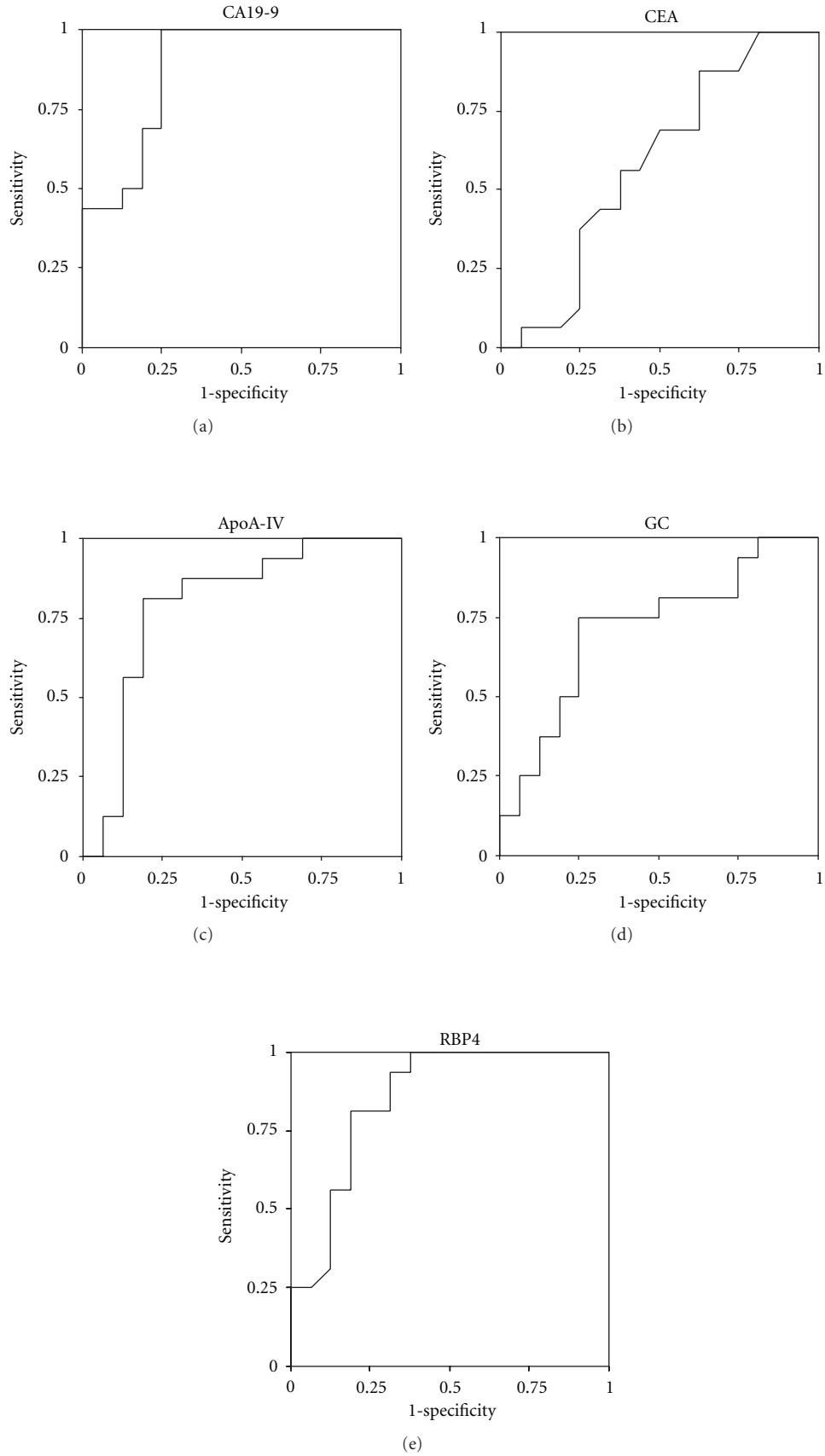


Figure 4: Receiver operating characteristic (ROC) curves for CA19-9, CEA, ApoA-IV, GC, RBP4 Their AUCs are described in the text.

Table 3

(a) Serum tumor marker levels in patients with pancreatic cancer.

Gender	Age (years)	UICC-stage	Tumor size (mm)	CA19-9 (U/mL)	CEA (ng/mL)	ApoA-IV (AU)	GC (AU)	RBP4 (AU)
M	38	IA	10	26.2	2.3	262.1	35.9	54.2
M	50	IB	30	46.5	2.3	8.5	34.3	43.9
M	63	IIA	18	157	1.1	10.2	26.7	37.0
M*	62	IIA	38	9	5	70.7	25.1	41.5
M*	54	IIA	24	11	1.4	87.5	11.3	42.5
M	73	IIA	25	15.5	3.5	203.5	13.6	49.5
F*	76	IIA	26	13.3	1.5	70.7	33.1	43.2
M	65	IIB	32	1579	3.4	352.1	24.3	36.7
M*	74	IIB	27	10.9	3.4	0	28.0	35.3
F	74	IIB	40	43	3.2	54	25.8	38.9
F	68	IIB	27	302	2.1	82.8	40.8	36.4
M	61	IIB	30	10	1.4	134.7	27.3	46.2
M	63	IIB	15	1080	—	144.6	23.9	53.3
M*	73	IIB	50	13.9	2.1	23.2	1.8	43.1
F	62	IIB	25	11.5	1.1	111.7	28.4	43.1
Ave ± SD	63.7 ± 1 0.4		27.8 ± 9.9	221.9 ± 466.2	63.7 ± 10.4	107.8 ± 99.9	25.4 ± 10	50.2 ± 4.2

CA19-9: carbohydrate antigen, CEA: carcinoembryonic antigen, ApoA-IV: apolipoprotein A-IV, GC: vitamin D-binding protein, RBP4: plasma retinol binding protein 4.
The stars indicate the patient who had a normal CA19-9 level and a low ApoA-IV.
Ave: average. SD: standard deviation. AU: arbitrary unit.

(b) Serum tumor marker levels in healthy controls.

Gender	Age (years)	CA19-9 (u/mL)	CEA (ng/mL)	ApoA-IV (AU)	GC (AU)	RBP4 (AU)
M	71	33.9	2.2	157.4	29.5	47.7
M	55	0.1	4.7	243.8	35.0	48.5
M	55	6.4	1	223.0	35.3	54.3
M	55	0.1	3.5	237.1	33.0	48.1
M	60	7.7	0.8	237.1	23.9	44.1
M	60	37.4	7	214.0	39.6	48.8
M	61	7.6	4.8	277.0	16.8	49.7
F	77	7.1	1.1	294.5	24.2	43.7
M	61	7.8	1.3	80.6	33.4	55.3
F	61	6	3	210.1	45.8	56.3
F	65	2.9	2.1	65.7	61.7	53.3
M	66	5.1	0.7	156.7	33.8	44.7
F	62	11.7	1.7	122.9	33.0	50.9
F	64	8.4	1.2	222.9	38.7	50.3
F	71	24.4	2.6	184.7	30.5	56.3
Ave ± SD	62.9 ± 6.3	11.1 ± 11.5	2.5 ± 1.8	195.2 ± 66.9	34.3 ± 10.3	50.2 ± 4.2

CA19-9: carbohydrate antigen, CEA: carcinoembryonic antigen, ApoA-IV: apolipoprotein A-IV, GC: vitamin D-binding protien, RBP4: plasma retinol binding protein
Ave: average. SD: standard deviation. AU: arbitrary unit.

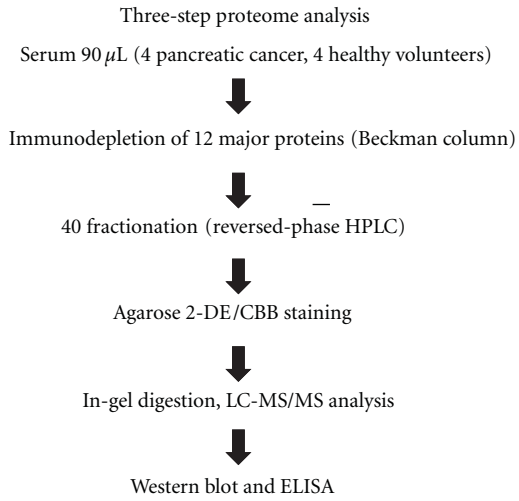


Figure 5: Outline of the procedure for the three-step serum proteome analysis.

Acknowledgments

The authors would like to thank Masanori Semiya for giving them helpful discussions and suggestions. Also, they thank Fumie Iida and Manami Miura for their technical supports.

References

[1] A. Jemal, R. Siegel, E. Ward et al., "Cancer statistics, 2006," *Ca-A Cancer Journal for Clinicians*, vol. 56, no. 2, pp. 106–130, 2006.

[2] G. Barugola, L. Frulloni, R. Salvia, and M. Falconi, "Is CA 19-9 a screening marker?" *Digestive and Liver Disease*, vol. 41, no. 5, pp. 325–327, 2009.

[3] N. Rifai, M. A. Gillette, and S. A. Carr, "Protein biomarker discovery and validation: the long and uncertain path to clinical utility," *Nature Biotechnology*, vol. 24, no. 8, pp. 971–983, 2006.

[4] S. C. C. Wong, C. M. L. Chan, B. B. Y. Ma et al., "Advanced proteomic technologies for cancer biomarker discovery," *Expert Review of Proteomics*, vol. 6, no. 2, pp. 123–134, 2009.

[5] L. Smith, M. J. Lind, K. J. Welham, and L. Cawkwell, "Cancer proteomics and its application to discovery of therapy response markers in human cancer," *Cancer*, vol. 107, no. 2, pp. 232–241, 2006.

[6] L. C. Whelan, K. A. R. Power, D. T. McDowell, J. Kennedy, and W. M. Gallagher, "Applications of SELDI-MS technology in oncology," *Journal of Cellular and Molecular Medicine*, vol. 12, no. 5A, pp. 1535–1547, 2008.

[7] F. Nomura, T. Tomonaga, K. Sogawa et al., "Identification of novel and downregulated biomarkers for alcoholism by surface enhanced laser desorption/ionization-mass spectrometry," *Proteomics*, vol. 4, no. 4, pp. 1187–1194, 2004.

[8] Y. Shirai, K. Sagawa, T. Yamaguchi et al., "Protein profiling in pancreatic juice for detection of intraductal papillary mucinous neoplasm of the pancreas," *Hepato-Gastroenterology*, vol. 55, no. 86-87, pp. 1824–1829, 2008.

[9] N. L. Anderson and N. G. Anderson, "The human plasma proteome: history, character, and diagnostic prospects," *Molecular & Cellular Proteomics*, vol. 1, no. 11, pp. 845–867, 2002.

[10] R. S. Tirumalai, K. C. Chan, D. A. Prieto, H. J. Issaq, T. P. Conrads, and T. D. Veenstra, "Characterization of the low molecular weight human serum proteome," *Molecular & Cellular Proteomics*, vol. 2, no. 10, pp. 1096–1103, 2003.

[11] N. Hattori, S. Oda, T. Sadahiro et al., "YKL-40 identified by proteomic analysis as a biomarker of sepsis," *Shock*, vol. 32, no. 4, pp. 393–400, 2009.

[12] D. Cecconi, M. Palmieri, and M. Donadelli, "Proteomics in pancreatic cancer research," *Proteomics*, vol. 11, no. 4, pp. 816–828, 2011.

[13] M. Oh-Ishi, M. Satoh, and T. Maeda, "Preparative two-dimensional gel electrophoresis with agarose gels in the first dimension for high molecular mass proteins," *Electrophoresis*, vol. 21, no. 9, pp. 1653–1669, 2000.

[14] M. Satoh, E. Haruta-Satoh, A. Omori et al., "Effect of thyroxine on abnormal pancreatic proteomes of the hypothyroid rdw rat," *Proteomics*, vol. 5, no. 4, pp. 1113–1124, 2005.

[15] T. Nishimori, T. Tomonaga, K. Matsushita et al., "Proteomic analysis of primary esophageal squamous cell carcinoma reveals downregulation of a cell adhesion protein, periplakin," *Proteomics*, vol. 6, no. 3, pp. 1011–1018, 2006.

[16] B. Seliger, S. P. Dressler, E. Wang et al., "Combined analysis of transcriptome and proteome data as a tool for the identification of candidate biomarkers in renal cell carcinoma," *Proteomics*, vol. 9, no. 6, pp. 1567–1581, 2009.

[17] I. Gräntzdörffer, S. Carl-McGrath, M. P. Ebert, and C. Röcken, "Proteomics of pancreatic cancer," *Pancreas*, vol. 36, no. 4, pp. 329–336, 2008.

[18] T. L. Hwang, Y. Liang, K. Y. Chien, and J. S. Yu, "Overexpression and elevated serum levels of phosphoglycerate kinase 1 in pancreatic ductal adenocarcinoma," *Proteomics*, vol. 6, no. 7, pp. 2259–2272, 2006.

[19] M. Ono, J. Matsubara, K. Honda et al., "Prolyl 4-hydroxylation of α-fibrinogen. A novel protein modification revealed by plasma proteomics," *Journal of Biological Chemistry*, vol. 284, no. 42, pp. 29041–29049, 2009.

[20] P. H. R. Green, R. M. Glickman, J. W. Riley, and E. Quinet, "Human apolipoprotein A-IV. Intestinal origin and distribution in plasma," *Journal of Clinical Investigation*, vol. 65, no. 4, pp. 911–919, 1980.

[21] L. Quadro, W. S. Blaner, D. J. Salchow et al., "Impaired retinal function and vitamin A availability in mice lacking retinol-binding protein," *EMBO Journal*, vol. 18, no. 17, pp. 4633–4644, 1999.

[22] C. Fabris, A. Piccoli, and A. Meani, "Study of retinol-binding protein in pancreatic cancer," *Journal of Cancer Research and Clinical Oncology*, vol. 108, no. 2, pp. 227–229, 1984.

[23] O. Kisker, S. Onizuka, C. M. Becker et al., "Vitamin D binding protein-macrophage activating factor (DBP-maf) inhibits angiogenesis and tumor growth in mice," *Neoplasia*, vol. 5, no. 1, pp. 32–40, 2003.

[24] I. Clemmensen, L. Petersen Chr. L., and C. Kluft, "Purification and characterization of a novel, oligomeric, plasminogen kringle 4 binding protein from human plasma: tetranectin," *European Journal of Biochemistry*, vol. 156, no. 2, pp. 327–333, 1986.

[25] B. Gronlund, E. V. S. Høgdall, I. J. Christensen et al., "Pre-treatment prediction of chemoresistance in second-line chemotherapy of ovarian carcinoma: value of serological tumor marker determination (tetranectin, YKL-40, CASA, CA 125)," *International Journal of Biological Markers*, vol. 21, no. 3, pp. 141–148, 2006.

[26] C. K. Høgdall, I. J. Christensen, R. W. Stephens, S. Sørensen, B. Nørgaard-Pedersen, and H. J. Nielsen, "Serum tetranectin is an independent prognostic marker in colorectal cancer and weakly correlated with plasma suPAR, plasma PAI-1 and serum CEA," *Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, vol. 110, no. 9, pp. 630–638, 2002.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 343582, 9 pages
doi:10.1155/2011/343582

Review Article

Protein Biomarkers for the Early Detection of Breast Cancer

David E. Misek[1,2] and Evelyn H. Kim[1]

[1] *Department of Surgery, University of Michigan Medical School, Ann Arbor, MI 48109-5656, USA*
[2] *The Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109-5656, USA*

Correspondence should be addressed to David E. Misek, dmisek@umich.edu

Received 24 May 2011; Accepted 23 June 2011

Academic Editor: Tadashi Kondo

Copyright © 2011 D. E. Misek and E. H. Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Advances in breast cancer control will be greatly aided by early detection so as to diagnose and treat breast cancer in its preinvasive state prior to metastasis. For breast cancer, the second leading cause of cancer-related death among women in the United States, early detection does allow for increased treatment options, including surgical resection, with a corresponding better patient response. Unfortunately, however, many patients' tumors are diagnosed following metastasis, thus making it more difficult to successfully treat the malignancy. There are, at present, no existing validated plasma/serum biomarkers for breast cancer. Only a few biomarkers (such as HER-2/neu, estrogen receptor, and progesterone receptor) have utility for diagnosis and prognosis. Thus, there is a great need for new biomarkers for breast cancer. This paper will focus on the identification of new serum protein biomarkers with utility for the early detection of breast cancer.

1. Introduction

Advances in breast cancer control will be greatly aided by early detection, thereby facilitating diagnosis and treatment of breast cancer in its preinvasive state prior to metastasis. Breast cancer is the most frequently occurring malignancy and the second leading cause of cancer-related death for women in the United States [1]. The most efficacious screening modality utilized in the clinic is mammography though lesions less than 0.5 cm in size remain undetectable by present technology. Importantly, however, even though a breast lesion may be detected, given the low sensitivity/specificity of mammography, approximately 4-fold more women (than those with breast malignancies) have resultant biopsies. Five-year survival of women with breast cancer is highly correlated with tumor stage, with tumor detection at very early stages (stages 0 and I) having an approximate 98% 5-year survival. Five-year survival for stage II tumors is approximately 85%, stage III approximately 60%, and stage IV approximately 20%. Overall, breast cancer has an approximate 80% 5-year survival, with 207,090 new cases and 39,840 deaths expected in women in the United States in 2010 [1].

Early detection of breast cancer does allow for increased treatment options, including surgical resection, with a corresponding better patient response. Surgical resection may involve lumpectomy or mastectomy with removal of some of the axillary lymph nodes. Following early detection, radiation therapy, chemotherapy (before or after surgery), and hormone therapy (tamoxifen [2] and aromatase inhibitors [3–5]) also have utility for therapeutic intervention. Targeted biologic therapy with trastuzumab (Herceptin) [6] or lapatinib (Tykerb) [7, 8] also has utility to treat HER2/neu-positive breast tumors. Unfortunately, however, in the absence of good serum/plasma biomarkers many breast cancer patients are diagnosed too late in the disease process (i.e., after the tumors metastasize) for surgical resection to be an effective option. Thus, these patients are typically offered various therapeutic treatment modalities dependent upon tumor subtype (ER^+ or ER^-; $HER2^+$ or $HER2^-$). The available treatment modalities may include hormonal (antiestrogen), taxane (docetaxel or paclitaxel) or nontaxane chemotherapy. In general, women with metastatic breast cancer are provided one therapeutic modality until treatment failure and are then switched to another therapeutic modality.

The origin of most breast cancer cases is not known. However, many risk factors have been identified, including female gender, increasing patient age, family history of breast

cancer at an early age, early menarche, late menopause, older maternal age at first live childbirth, prolonged hormone replacement therapy, exposure to therapeutic chest irradiation, benign proliferative breast disease, and genetic mutations in genes such as BRCA1/2 [9]. The overwhelming majority of breast masses detected by palpation and/or by mammography are epithelial lesions, which include benign fibrocystic change, hyperplasia, carcinoma in situ, and infiltrating mammary carcinoma. Although several histologic types and subtypes of mammary carcinomas exist, >95% are either ductal or lobular carcinomas [10], with the majority (75%–80%) of mammary carcinomas being ductal carcinomas [11, 12].

A number of genetic alterations have been identified in breast tumors. The most frequent genomic aberrations identified are gains along chromosomes 1q, 8q, 17q, 20q, and 11q and losses along 8p, 13q, 16q, 18q, and 11q [13–18]. Interestingly, many of these chromosomal segments harbor known proto-oncogenes and/or tumor suppressor genes such as BRCA1, BRCA2, HER2-neu, C-MYC, and Cyclin D-1. Low-grade (grade 1) infiltrating ductal carcinomas have relatively few numbers of chromosomal alterations with the highest frequency of aberrations occurring as losses on 16q and gains on 1q.

It is generally accepted that estrogen receptor-positive (ER^+) and ER-negative (ER^-) breast cancers are two different disease entities. ER^- tumors tend to be of high grade, have more frequent p53 mutations, and have worse prognosis compared with ER^+ disease. Both ER^+ and ER^- tumors can be either HER2 positive or negative. Low-grade tumors are typically ER positive, almost always HER2 nonamplified, and frequently overexpress cyclin D-1 [10]. In contrast, high-grade (grade 3) tumors tend to be ER negative, have frequent loss of p53 function, usually overexpress C-MYC and commonly overexpress HER2 [13–15, 19]. In the high-grade tumors, loss of p53 function is usually due to 17p13 deletion, mutation or inactivation, while overexpression of HER2 is usually because of 17q12 amplification [20–28].

Although early detection of cancer has improved survival for a number of cancers, including breast cancer [29], colon cancer [30–32], prostate cancer [33, 34], and cervical cancer [35], existing serum biomarkers for breast cancer are not adequate for early detection. The possibility of early detection of breast cancer may be realized through both noninvasive (i.e., imaging technologies) and invasive means (patient serum profiling). To date, gains in the early detection of breast cancer have been largely made due to routine mammography and/or by palpation (either self-examination or by physician or nurse practitioner). Imaging technologies (mammography, digital mammography, and magnetic resonance imaging (MRI)) have been adopted clinically for mass screening purposes, but there is resistance for seeking such services on a yearly basis, given the relative complexity and high cost-to-benefit ratio of these imaging methodologies. As a result, there has been much interest in development and validation of serum-based biomarkers for the early detection, risk stratification, prediction, and disease prognosis of breast cancer. This paper will focus on recent developments in identification of new serum protein biomarkers with potential utility for the early detection of breast cancer (Table 1).

2. Autoantibodies and Breast Cancer

The humoral immune response to cancer in humans has been well demonstrated by identification of autoantibodies to a number of different intracellular and surface antigens in patients with various tumor types [36–39]. A tumor-specific humoral immune response directed against oncoproteins [40, 41], mutated proteins such as p53 [42, 43], or other aberrantly expressed proteins have all been described. While it is currently unknown whether the occurrence of such antibodies is beneficial, knowledge of potential tumor antigens that may evoke tumor-specific immune responses may have utility in early cancer diagnosis, in establishing prognosis and in immunotherapy against the disease.

Several approaches are currently available for the identification of tumor antigens. In contrast to identification of tumor antigens based on analysis of recombinant proteins (which do not contain posttranslational modifications as found in tumors or tumor cell lines), it may be preferable to utilize a proteomics-based approach for the identification of tumor antigens. This may facilitate the identification of autoantibodies to naturally occurring proteins, such as in lysates prepared from tumors and tumor cell lines, and may uncover antigenicity associated with aberrant posttranslational modification of tumor cell proteins. Such a proteomics approach was implemented for the identification of breast tumor antigens that elicit a humoral response against proteins that are expressed in the SUM-44 breast cancer cell line. 2D PAGE was used to simultaneously separate individual cellular proteins from the SUM-44 cell line. The separated proteins were transferred onto PVDF membranes. Sera from breast cancer patients were screened individually for antibodies that reacted against the separated proteins by Western blot analysis. Proteins specifically reacting with sera from the breast cancer patients were identified by mass spectrometry. Le Naour and colleagues [36] have shown that a humoral response directed against RS/DJ-1 occurred in 13.3% of newly diagnosed breast cancer patients. None of the 25 healthy controls (0%) or 46 patients (0%) with hepatocellular carcinoma exhibited autoantibodies to RS/DJ-1. Only 2/54 (3.7%) samples of sera from lung adenocarcinoma patients demonstrated autoantibodies to RS/DJ-1.

In breast cancer, besides RS/DJ-1 [36], autoimmunity has also been shown against a number of other cellular proteins. These proteins include p53 [44–47], heat shock protein 60 [48, 49], heat shock protein 90 [50, 51], and mucin-related antigens [49, 52–54]. The presence of p53 autoantibodies have been observed in 15% of patients with breast cancer and were shown to be associated with a poor prognosis [44, 45, 47]. However, p53 autoantibodies have also been found in patients with other malignancies and inflammatory conditions [42, 43], thus the humoral response to p53 is not specific to breast cancer. A humoral response to the 90 kDa heat shock protein has also been associated with poor survival in breast cancer [51]. In contrast, the presence of MUC1

Table 1: Current promising biomarkers for the detection of breast cancer.

Name of biomarker	Technology used for discovery	Type	Reference
RS/DJ-1	Humoral response	autoantibody	[36]
p53	Humoral response	autoantibody	[44–47]
HSP60	Humoral response	autoantibody	[48, 49]
HSP90	Humoral response	autoantibody	[50, 51]
Mucin-related	Humoral response	autoantibody	[49, 52–54]
CA 15-3	Serum profiling	serum protein	[55, 56]
RS/DJ-1	Serum profiling	serum protein	[36]
HER-2/neu	Serum profiling	serum protein	[72]
α-2-HS-glycoprotein	Nipple aspirate fluid profiling	Ductal protein	[90]
Lipophilin B	Nipple aspirate fluid profiling	Ductal protein	[90]
beta-globin	Nipple aspirate fluid profiling	Ductal protein	[90]
Hemopexin	Nipple aspirate fluid profiling	Ductal protein	[90]
Vitamin D-binding protein	Nipple Aspirate Fluid Profiling	Ductal protein	[90]

autoantibodies has been associated with a reduced risk for disease progression in patients with breast cancer [53, 54]. While the antigenic epitope on MUC1 (or, for that matter, any of the other breast tumor antigens discussed above) is unknown, MUC1 has been shown to be aberrantly glycosylated frequently in breast cancer [54]. At present, CA 15-3 (a soluble or secreted form of MUC1) has utility as a circulating marker for breast cancer [55, 56]. Serial measurements of CA 15-3 have utility to detect recurrences and to monitor the treatment of metastatic breast cancer [55–57]. Additionally, the CA 15-3 concentration at initial presentation does have prognostic significance [58–62].

In order to circumvent many of the difficulties associated with 2D-PAGE (namely, inadequate resolution, slow throughput, and limited dynamic range), protein microarrays were developed that have the capability to screen patient's sera for autoantibodies directed against tumor antigens [63–66]. In comparison to traditional ELISAs that use single purified recombinant proteins, the protein microarrays are capable of presenting and analyzing >1000 tumor antigens simultaneously. In addition, as these tumor antigens are typically derived from diseased tissues or disease-related cells, they possess disease-related, potentially antigenic, posttranslational modifications not normally expressed by the particular cells or tissue. In this technology, proteins from diseased tissues or disease-related cell lines are separated by 2-dimensional liquid chromatography (chromatofocusing or ion exchange HPLC in the first dimension, followed by reverse phase HPLC in the second dimension). Following separation, all fractions ($\geq$1700 fractions) from each separation are printed onto nitrocellulose-coated microscope slides and are subsequently probed with sera from patients or control subjects [63–66]. As each reactive fraction may contain a number of different proteins, each reactive fraction would need to be further assessed to determine the tumor antigen of interest.

More recently, Ramachandran et al. [67, 68] developed a novel protein microarray technology, termed nucleic acid protein programmable array (NAPPA). NAPPA arrays are generated by printing full-length cDNA encoding the target proteins at each feature of the array. The proteins are then transcribed and translated by a cell-free system and immobilized *in situ* using epitope tags fused to the proteins. Although this technology circumvents many of the difficulties of traditional protein microarrays (i.e., the need to resolve complex protein lysates), the printed proteins on the array lack all normal posttranslational modifications. Thus, any antigenicity resulting from aberrant modification of tumor proteins is not assessed. Anderson and colleagues [69] utilized the NAPPA arrays to screen 4988 candidate tumor antigens with sera from patients with early stage breast cancer for autoantibodies. Twenty-eight of these antigens were confirmed using an independent serum cohort ($n = 51$ cases/38 controls, $P < 0.05$). Using all 28 antigens, a classifier was identified with a sensitivity of 80.8% and a specificity of 61.6% (AUC = 0.756). Although the sensitivity and specificity are not high, these 28 recombinant protein antigens may be considered as potential biomarkers for the early detection of breast cancer.

It is not clear why only a subset of patients with a particular tumor type develop a humoral response to particular tumor antigens. Immunogenicity may depend on the level of expression, posttranslational modification, or other types of protein processing, the extent of which may be variable among tumors of a similar histological type. Other factors that may influence the immune response include variability among tumors and individuals in major histocompatibility complex molecules and in antigen presentation. Although a number of autoantibodies have been identified in breast cancer, in most cases, they occur in less than 50% of patient's sera. Therefore, they are not likely to be effective individually for the early detection of breast cancer but may show efficacy if utilized as a panel of biomarkers.

3. Detection of Altered Plasma Protein Expression for Identification of Breast Cancer-Specific Biomarkers

There has been great interest in the hypothesis that tumor-specific proteins may be found in patient's circulation, and

they may have utility for the early detection of cancer. For example, proteins such as CA125 in ovarian cancer and prostate-specific antigen (PSA) in prostate cancer have been used clinically as diagnostic markers of cancer. CA125 is a mucin commonly employed as a diagnostic marker for epithelial ovarian cancer. PSA is secreted primarily by prostate epithelial cells into the seminal plasma and is one of the best characterized examples of a secreted glycoprotein used in cancer diagnostics.

There are a number of reports that have described aberrantly expressed proteins in the serum of breast cancer patients. The most widely used serum marker in breast cancer diagnostics is CA 15-3, which detects soluble forms of the mucin MUC1. MUC1 is normally found in the apical membrane of normal secretory epithelium. Following malignant transformation, however, MUC1 may be localized throughout the external surface of the entire plasma membrane. In addition, changes in MUC1 glycosylation have been reported during neoplastic transformation [70, 71]. Although MUC1 is expressed in normal and neoplastic breast epithelium, the clinical utility of MUC1 measurements is confined to measurements of shed or soluble forms (termed CA 15-3), released from the cell surface by proteolytic cleavage. Unfortunately, CA 15-3 is not suitable for early detection, as serum levels are rarely increased in patients with early or localized breast cancer. The main utility for CA 15-3 is for monitoring therapy in patients with metastatic breast cancer.

Le Naour and coworkers [36] have evaluated RS/DJ-1 as a serum biomarker of breast cancer. In normal tissue, expression of RS/DJ-1 was observed in epithelium, smooth muscle, blood vessels, and nerves. All 15 (100%) invasive ductal carcinomas and 3 (100%) invasive lobular carcinomas showed some level of cytoplasmic and nuclear reactivity in the neoplastic cells. Significantly elevated levels of serum RS/DJ-1 was observed in the sera of 11/30 patients with newly diagnosed breast cancer, as compared to serum from 25 healthy subjects. However, these authors did not evaluate serum RS/DJ-1 levels in patients with other types of breast lesions. Thus, it is unknown whether the increased serum RS/DJ-1 levels are cancer-specific.

In another study [72], significantly higher serum HER-2/neu levels were found in patients with tissue overexpression of HER-2/neu. Univariate analysis showed that HER-2/neu serum levels were prognostic factors in disease-free survival and overall survival only in patients with tissue overexpression. When only patients with HER-2/neu overexpression in tissue were studied, tumor size, nodal involvement, and tumor markers (at least one positive) were found to be independent prognostic factors for both disease-free survival and overall survival.

4. Use of Mass Spectrometric Methodologies for Identification of Breast Cancer-Specific Biomarkers

Methodologies have been developed to directly analyze the proteins contained within complex protein mixtures, such as that found within human biofluids (plasma or serum, nipple aspirate fluid, ductal lavage fluid, saliva, etc.). Among these technologies, some, like SELDI (Surface-Enhanced Laser Desorption and Ionization) are mass spectrometry-based. A number of investigators have used SELDI-TOF mass spectrometry to interrogate serum [73–81] and nipple aspirate/ductal lavage fluid [82–89] from patients with breast cancer. In one study, serum samples from women with or without breast cancer were analyzed using SELDI protein chip mass spectrometry [77]. Using a case-control study design, serum samples from 48 female patients with primary invasive breast cancer were compared with samples from 48 age- and sex-matched healthy controls. To increase the number of identifiable proteins, patient's serum was profiled on IMAC30 (activated with nickel) ProteinChip surfaces. Differences in protein intensity between breast cancer cases and controls were measured by the Mann-Whitney U test and adjusted for confounding variables in a multivariate logistic regression model. Three peaks, with mass-to-charge ratio (m/z) 4276, 4292 and 8941 were found that showed significant decreased expression in cancer sera, as compared to control sera ($P < 0.001$). One drawback of the SELDI technology, however, is that given the limited dynamic range of SELDI, it is likely that distinctive features observed in serum with this approach represent relatively abundant proteins that are not necessarily specific to breast cancer. Further, SELDI has difficulties in providing the identification of the distinctive proteins when used to directly profile complex protein mixtures.

5. Mass-Spectrometric Profiling of Nipple Aspirate Fluid or Ductal Lavage Fluid

Other mass spectrometric profiling methods have been utilized to profile proteins found in nipple aspirate fluid [90] and ductal lavage fluid in order to identify breast cancer-specific biomarkers. These investigators [90] analyzed paired nipple aspirate fluid samples from 18 women with stage I or stage II unilateral invasive breast cancer and 4 healthy volunteers using ICAT (isotope-coded affinity tag) labeling, followed by SDS-PAGE. Gel slices were cut from each sample, with subsequent analysis by liquid chromatography tandem mass spectrometry (LC-MS/MS). They identified 353 peptides from the tandem mass spectra. Alpha-2-HS-glycoprotein was found to be underexpressed in nipple aspirate fluid from tumor-bearing breasts, while lipophilin B, beta-globin, hemopexin and vitamin D-binding protein were all overexpressed. Unfortunately, these authors only identified abundant proteins whose over- or underexpression was somewhat modest. Moreover, these authors did not analyze nipple aspirate fluid from patients with inflammatory breast disease. Thus, conclusions cannot be drawn regarding breast cancerspecificity of protein expression.

6. N-linked Glycan Profiling for Biomarker Identification in Breast Cancer Serum

Glycoproteins are the most heterogeneous group of post-translational modifications known in proteins. Glycans show a high structural diversity reflecting inherent functional

diversity. N- and O-oligosaccharide variants on glycoproteins (glycoforms) can lead to alterations in protein activity or function that may manifest itself as overt disease [91, 92]. Many clinical biomarkers and therapeutic targets in cancer are glycoproteins [93–95], such as CA125 in ovarian cancer, HER2/neu in breast cancer, and prostate-specific antigen (PSA) in prostate cancer. The human epidermal growth factor receptor 2 (HER2/neu) is a transmembrane glycoprotein, where the presence of HER2 overexpression appears to be a key factor in malignant transformation and is predictive of a poor prognosis in breast cancer. CA125 is a mucin commonly employed as a diagnostic marker for epithelial ovarian cancer. Although CA125 has been used as an ovarian cancer marker for a long time, many of its O- and N-glycan structures have only recently been characterized [96]. PSA is secreted primarily by prostate epithelial cells into the seminal plasma. It is one of the best characterized examples of a secreted glycoprotein used in cancer diagnostics, and its glycoforms have been described [97]. The alteration in protein glycosylation that occurs through varying the heterogeneity of glycosylation sites or changing glycan structure of proteins on the cell surface and in body fluids has been shown to correlate with the development or progression of cancer and other disease states [98]. It has been reported that the glycosylation of PSA secreted by the tumor prostate cell line LNCaP differs significantly from that of PSA from seminal plasma (normal control). These carbohydrate differences allow a distinction to be made between PSA from normal and tumor origins and provide a valuable biochemical tool for diagnosis of prostate cancer [99].

There is growing evidence that glycan structures on glycoproteins are modified in breast cancer [100–109]. Breast cancer-associated alterations have been demonstrated for fucosylation groups and for sialylations on the plasma protein α-1-proteinase inhibitor [106]. Increased GlcNAc β1-6Man α1-6Man β-branching in asparagine-linked oligosaccharides has been observed in human tumor cells. The levels of the β1-6 branched oligosaccharides were evaluated in a series of benign and malignant human breast biopsies. Normal human breast tissue and benign lesions showed low expression but 50% of the primary malignancies examined showed significantly elevated β1-6 branching [107]. Subsequently, L-PHA (a lectin that binds specifically to the β1-6 branched oligosaccharides) lectin histochemistry was performed on paraffin sections of human breast tissues. All breast carcinomas and epithelial hyperplasia with atypia demonstrated significantly increased L-PHA staining as compared to fibroadenomas and hyperplasia without atypia [108]. More recently, L-PHA reactive glycoproteins were identified from matched normal (nondiseased) and malignant tissue isolated from patients with invasive ductal breast carcinoma [109]. Comparison analysis of the data identified 34 proteins that were enriched by L-PHA fractionation in tumor relative to normal tissue for at least 2 cases of ductal invasive breast carcinoma. Of these 34 L-PHA tumor enriched proteins, 12 were common to all 4 matched cases analyzed.

Abd Hamid and coworkers [110] analyzed fluorescently tagged serum *N*-glycans of advanced breast cancer patients using exoglycosidases and LC-MS/MS. They found that the expression of a trisialylated triantennary glycan containing an α-1,3-linked fucose was increased in the presence of breast cancer. Kyselova and coworkers. profiled the permethylated *N*-glycans in sera of breast cancer patients at different stages (stages I to IV) using MALDI TOF/TOF MS in one study [111]. In a second study, they profiled reduced and methylated serum *N*-glycans of late-stage breast cancer patients using nanoliquid chromatography (LC) chip/time-of-flight (TOF) MS [112]. In both studies, they found an increase in fucosylation in both core and branched segments of *N*-glycans in the presence of breast cancer. In the latter study, they found a decrease in expression of a biantennary-monosialylated *N*-linked glycan and an increase in expression of a fucosylated triantennary-trisialylated *N*-linked glycan in the presence of Stage IV breast cancer. These glycosylation changes in a tumor-secreted protein may reflect fundamental activity changes in the enzymes involved in the glycosylation pathway, either through altered levels of enzymes or altered enzymatic activity. Importantly, the changes in glycan structure may serve as early detection biomarkers of breast cancer.

7. Summary

Early detection of breast cancer, so as to diagnose and treat cancer in its preinvasive state prior to metastasis, may greatly impact the treatment and prognosis of patients with this common, but deadly, malignancy. Unfortunately, at present, suitable biomarkers have not been identified for the early detection of breast cancer. Biomarker discovery for this disease is still very much in its discovery phase. Multiple approaches have been developed, as described above, that hold promise for the identification of serum biomarkers. The protein biomarkers that have been identified to date do not possess the requisite sensitivity/specificity to have utility individually as a biomarker for the early detection of breast cancer but ultimately may have utility within a panel of protein biomarkers. Additionally, other emerging technologies, such as genetically engineered mouse models of breast cancer may have utility to identify panels of serum biomarkers that can be further explored in human sera. In order to determine the utility of any promising protein biomarkers, the candidates will need to be tested and validated by multiple independent studies using an adequately sized test and training set of sera samples from very early-stage breast cancer. Development of such resources, including serum from patients with nonmalignant breast lesions and prospective serum collection from individuals at high risk of being diagnosed with breast cancer as well as serum from patients with other breast lesions and other types (nonbreast) of malignancies is of critical need for the identification of biomarkers with utility for the early detection of breast cancer. Up until now, serum/plasma collection has been primarily performed in individual laboratories, using heterogeneous sample collection methods. The Human Proteome Organization (HUPO) has conducted a study to assess efficacious serum collection methods. These findings have

lead to efforts presently being made by the National Cancer Institute, through the Early Detection Research Network, to develop suitable serum resources for both the discovery phase and the subsequent validation phase of biomarkers for the early detection of cancer. With the ultimate development of these standardized resources, it is expected that suitable biomarkers would be validated and have utility for the early clinical detection of breast cancer within the next five-to-ten years.

Acknowledgment

This work was supported by Award no. W81XWH-09-1-0043 from the DoD (Army). The information presented reflects the opinions of the authors only and does not reflect the opinion of the DoD.

References

[1] A. Jemal, R. Siegel, J. Xu, and E. Ward, "Cancer statistics, 2010," *CA Cancer Journal for Clinicians*, vol. 60, no. 5, pp. 277–300, 2010.

[2] Early Breast Cancer Trialists' Collaborative Group (EBCTCG), "Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials," *The Lancet*, vol. 365, pp. 1687–1717, 2005.

[3] J. M. Nabholtz, A. Buzdar, M. Pollak et al., "Anastrozole is superior to tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: results of a North American multicenter randomized trial," *Journal of Clinical Oncology*, vol. 18, no. 22, pp. 3758–3767, 2000.

[4] J. Bonneterre, B. Thurlimann, J. F. Robertson et al., "Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the tamoxifen or arimidex randomized group efficacy and tolerability study," *Journal of Clinical Oncology*, vol. 18, no. 22, pp. 3748–3757, 2000.

[5] J. Bonneterre, A. Buzdar, J. M. Nabholtz et al., "Anastrozole is superior to tamoxifen as first-line therapy in hormone receptor positive advanced breast carcinoma," *Cancer*, vol. 92, no. 9, pp. 2247–2258, 2001.

[6] C. A. Hudis, "Trastuzumab—Mechanism of action and use in clinical practice," *New England Journal of Medicine*, vol. 357, no. 1, pp. 39–51, 2007.

[7] D. Bilancia, G. Rosati, A. Dinota, D. Germano, R. Romano, and L. Manzione, "Lapatinib in breast cancer," *Annals of Oncology*, vol. 18, supplement 6, pp. vi26–vi30, 2007.

[8] G. M. Higa and J. Abraham, "Lapatinib in the treatment of breast cancer," *Expert Review of Anticancer Therapy*, vol. 7, no. 9, pp. 1183–1192, 2007.

[9] R. W. Carlson, B. O. Anderson, H. J. Burstein et al., "Invasive breast cancer: clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 5, no. 3, pp. 246–312, 2007.

[10] B. J. Yoder, E. J. Wilkinson, and N. A. Massoll, "Molecular and morphologic distinctions between infiltrating ductal and lobular carcinoma of the breast," *Breast Journal*, vol. 13, no. 2, pp. 172–179, 2007.

[11] F. A. Tavassoli, "Invasive lobular carcinoma," in *Pathology of the Breast*, F. A. Tavassoli, Ed., pp. 426–36, Appleton & Lange, Stamford, Conn, USA, 2nd edition, 1999.

[12] "Invasive lobular carcinoma," in *Rosen's Breast Pathology*, P. P. Rosen, Ed., pp. 627–52, Lippincott Williams & Wilkins, Philadelphia, Pa, USA, 2nd edition, 2001.

[13] J. S. Reis-Filho, P. T. Simpson, T. Gale, and S. R. Lakhani, "The molecular genetics of breast cancer: the contribution of comparative genomic hybridization," *Pathology Research and Practice*, vol. 201, no. 11, pp. 713–725, 2005.

[14] S. E. Shackney and J. F. Silverman, "Molecular evolutionary patterns in breast cancer," *Advances in Anatomic Pathology*, vol. 10, no. 5, pp. 278–290, 2003.

[15] P. T. Simpson, J. S. Reis-Filho, T. Gale, and S. R. Lakhani, "Molecular evolution of breat cancer," *Journal of Pathology*, vol. 205, no. 2, pp. 248–254, 2005.

[16] D. E. Stange, B. Radlwimmer, F. Schubert et al., "High-resolution genomic profiling reveals association of chromosomal aberrations on 1q and 16p with histologic and genetic subgroups of invasive breast cancer," *Clinical Cancer Research*, vol. 12, no. 2, pp. 345–352, 2006.

[17] D. G. Albertson, C. Collins, F. McCormick, and J. W. Gray, "Chromosome aberrations in solid tumors," *Nature Genetics*, vol. 34, no. 4, pp. 369–376, 2003.

[18] D. G. Albertson and D. Pinkel, "Genomic microarrays in human genetic disease and cancer," *Human Molecular Genetics*, vol. 12, no. 2, pp. R145–R152, 2003.

[19] H. Buerger, E. C. Mommers, R. Littmann et al., "Ductal invasive G2 and G3 carcinomas of the breast are the end stages of at least two different lines of genetic evolution," *Journal of Pathology*, vol. 194, no. 2, pp. 165–170, 2001.

[20] H. Tsuda and S. Hirohashi, "Multiple developmental pathways of highly aggressive breast cancers disclosed by comparison of histological grades and c-erbB-2 expression patterns in both the non-invasive and invasive portions," *Pathology International*, vol. 48, no. 7, pp. 518–525, 1998.

[21] J. E. Somerville, L. A. Clarke, and J. D. Biggart, "c-erbB-2 overexpression and histological type of in situ and invasive breast carcinoma," *Journal of Clinical Pathology*, vol. 45, no. 1, pp. 16–20, 1992.

[22] A. N. Jain, K. Chin, A. L. Børresen-Dale et al., "Quantitative analysis of chromosomal CGH in human breast tumors associates copy number abnormalities with p53 status and patient survival," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 14, pp. 7952–7957, 2001.

[23] D. N. Poller, C. E. Hutchings, M. Galea et al., "p53 protein expression in human breast carcinoma: relationship to expression of epidermal growth factor receptor, c-erbB-2 protein overexpression, and oestrogen receptor," *British Journal of Cancer*, vol. 66, no. 3, pp. 583–588, 1992.

[24] M. Barbareschi, E. Leonardi, F. A. Mauri, G. Serio, and P. Dalla Palma, "p53 and c-erbB-2 protein expression in breast carcinomas: an immunohistochemical study including correlations with receptor status, proliferation markers, and clinical stage in human breast cancer," *American Journal of Clinical Pathology*, vol. 98, no. 4, pp. 408–418, 1992.

[25] M. Rudas, R. Neumayer, M. F. X. Gnant, M. Mittelböck, R. Jakesz, and A. Reiner, "p53 Protein expression, cell proliferation and steroid hormone receptors in ductal and lobular in situ carcinomas of the breast," *European Journal of Cancer A*, vol. 33, no. 1, pp. 39–44, 1997.

[26] J. E. Eyfjord, S. Thorlacius, M. Steinarsdottir, R. Valgardsdottir, H. M. Ogmundsdottir, and K. Anamthawat-Jonsson, "p53 Abnormalities and genomic instability in primary human breast carcinomas," *Cancer Research*, vol. 55, no. 3, pp. 646–651, 1995.

[27] C. Wiltschke, I. Kindas-Muegge, A. Steininger, A. Reiner, G. Reinger, and P. N. Preis, "Coexpression of HER-2/neu and p53 is associated with a shorter disease-free survival in node-positive breast cancer patients," *Journal of Cancer Research and Clinical Oncology*, vol. 120, no. 12, pp. 737–742, 1994.

[28] R. Seshadri, A. S. Y. Leong, K. Mccaul, F. A. Firgaira, V. Setlur, and D. J. Horsfall, "Relationship between p53 gene abnormalities and other tumour characteristics in breast-cancer prognosis," *International Journal of Cancer*, vol. 69, no. 2, pp. 135–141, 1996.

[29] A. Goldhirsch, M. Colleoni, G. Domenighetti, and R. D. Gelber, "Systemic treatments for women with breast cancer: outcome with relation to screening for the disease," *Annals of Oncology*, vol. 14, no. 8, pp. 1212–1214, 2003.

[30] D. Lieberman, "How to screen for colon cancer," *Annual Review of Medicine*, vol. 49, pp. 163–172, 1998.

[31] J. A. Ferguson, "Early detection of unsuspected colon cancers in asymptomatic people," *Diseases of the Colon and Rectum*, vol. 36, no. 4, p. 411, 1993.

[32] J. S. Mandel, J. H. Bond, T. R. Church et al., "Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study," *New England Journal of Medicine*, vol. 328, no. 19, pp. 1365–1371, 1993.

[33] W. J. Catalona, "Management of cancer of the prostate," *New England Journal of Medicine*, vol. 331, no. 15, pp. 996–1004, 1995.

[34] S. J. Jacobsen, S. K. Katusic, E. J. Bergstralh et al., "Incidence of prostate cancer diagnosis in the eras before and after serum prostate-specific antigen testing," *Journal of the American Medical Association*, vol. 274, no. 18, pp. 1445–1449, 1995.

[35] E. Lynge, "Screening for cancer of the cervix uteri," *World Journal of Surgery*, vol. 13, no. 1, pp. 71–78, 1989.

[36] F. Le Naour, D. E. Misek, M. C. Krause et al., "Proteomics-based identification of RS/DJ-1 as a novel circulating tumor antigen in breast cancer," *Clinical Cancer Research*, vol. 7, no. 11, pp. 3328–3335, 2001.

[37] F. M. Brichory, D. E. Misek, A. M. Yim et al., "An immune response manifested by the common occurrence of annexins I and II autoantibodies and high circulating levels of IL-6 in lung cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 17, pp. 9824–9829, 2001.

[38] A. O. Güre, N. K. Altorki, E. Stockert, M. J. Scanlan, L. J. Old, and Y. T. Chen, "Human lung cancer antigens recognized by autologous antibodies: definition of a novel cDNA derived from the tumor suppressor gene locus on chromosome 3p21.3," *Cancer Research*, vol. 58, no. 5, pp. 1034–1041, 1998.

[39] E. Stockert, E. Jäger, Y. T. Chen et al., "A survey of the humoral immune response of cancer patients to a panel of human tumor antigens," *Journal of Experimental Medicine*, vol. 187, no. 8, pp. 1349–1354, 1998.

[40] K. Ben-Mahrez, I. Sorokine, D. Thierry et al., "Circulating antibodies against c-myc oncogene product in sera of colorectal cancer patients," *International Journal of Cancer*, vol. 46, no. 1, pp. 35–38, 1990.

[41] S. M. Pupa, S. Menard, S. Andreola, and M. I. Colnaghi, "Antibody response against the c-erbB-2 oncoprotein in breast carcinoma patients," *Cancer Research*, vol. 53, no. 24, pp. 5864–5866, 1993.

[42] S. F. Winter, J. D. Minna, B. E. Johnson, T. Takahashi, A. F. Gazdar, and D. P. Carbone, "Development of antibodies against p53 in lung cancer patients appears to be dependent on the type of p53 mutation," *Cancer Research*, vol. 52, no. 15, pp. 4168–4174, 1992.

[43] J. Raedle, G. Oremek, M. Welker, W. K. Roth, W. F. Caspary, and S. Zeuzem, "p53 autoantibodies in patients with pancreatitis and pancreatic carcinoma," *Pancreas*, vol. 13, no. 3, pp. 241–246, 1996.

[44] P. Lenner, F. Wiklund, S. O. Emdin et al., "Serum antibodies against p53 in relation to cancer risk and prognosis in breast cancer: a population-based epidemiological study," *British Journal of Cancer*, vol. 79, no. 5-6, pp. 927–932, 1999.

[45] T. Soussi, "p53 Antibodies in the sera of patients with various types of cancer: a review," *Cancer Research*, vol. 60, no. 7, pp. 1777–1788, 2000.

[46] K. Angelopoulou, H. Yu, B. Bharaj, M. Giai, and E. P. Diamandis, "p53 Gene mutation, tumor p53 protein overexpression, and serum p53 autoantibody generation in patients with breast cancer," *Clinical Biochemistry*, vol. 33, no. 1, pp. 53–62, 2000.

[47] A. Kulić, M. Sirotković-Skerlev, S. Jelisavac-Ćosić, D. Herceg, Z. Kovač, and D. Vrbanec, "Anti-p53 antibodies in serum: relationship to tumor biology and prognosis of breast cancer patients," *Medical Oncology*, vol. 27, no. 3, pp. 887–893, 2010.

[48] C. Desmetz, F. Bibeau, F. Boissière et al., "Proteomics-based identification of HSP60 as a tumor-associated antigen in early stage breast cancer and ductal carcinoma in situ," *Journal of Proteome Research*, vol. 7, no. 9, pp. 3830–3837, 2008.

[49] C. Desmetz, C. Bascoul-Mollevi, P. Rochaix et al., "Identification of a new panel of serum autoantibodies associated with the presence of in situ carcinoma of the breast in younger women," *Clinical Cancer Research*, vol. 15, no. 14, pp. 4733–4741, 2009.

[50] S. E. Conroy, S. L. Gibson, G. Brunstrom, D. Isenberg, Y. Luqmani, and D. S. Latchman, "Autoantibodies to 90 kD heat-shock protein in sera of breast cancer patients," *The Lancet*, vol. 345, no. 8942, p. 126, 1995.

[51] S. E. Conroy, P. D. Sasieni, I. Fentiman, and D. S. Latchman, "Autoantibodies to the 90 kDa heat shock protein and poor survival in breast cancer patients," *European Journal of Cancer*, vol. 34, no. 6, pp. 942–943, 1998.

[52] S. Von Mensdorff-Pouilly, M. M. Gourevitch, P. Kenemans et al., "Humoral immune response to polymorphic epithelial mucin (MUC-1) in patients with benign and malignant breast tumours," *European Journal of Cancer A*, vol. 32, no. 8, pp. 1325–1331, 1996.

[53] S. Von Mensdorff-Pouilly, A. A. Verstraeten, P. Kenemans et al., "Survival in early breast cancer patients is favorably influenced by a natural humoral immune response to polymorphic epithelial mucin," *Journal of Clinical Oncology*, vol. 18, no. 3, pp. 574–583, 2000.

[54] O. Blixt, D. Bueti, B. Burford et al., "Autoantibodies to aberrantly glycosylated MUC1 in early stage breast cancer are associated with a better prognosis," *Breast Cancer Research*, vol. 13, no. 2, article R25, 2011.

[55] M. J. Duffy, "CA 15-3 and related mucins as circulating markers in breast cancer," *Annals of Clinical Biochemistry*, vol. 36, no. 5, pp. 579–586, 1999.

[56] M. J. Duffy, D. Evoy, and E. W. McDermott, "CA 15-3: uses and limitation as a biomarker for breast cancer," *Clinica Chimica Acta*, vol. 411, no. 23-24, pp. 1869–1874, 2010.

[57] B. De La Lande, K. Hacene, J. L. Floiras, N. Alatrakchi, and M. F. Pichon, "Prognostic value of CA 15.3 kinetics for metastatic breast cancer," *International Journal of Biological Markers*, vol. 17, no. 4, pp. 231–238, 2002.

[58] F. G. Ebeling, P. Stieber, M. Untch et al., "Serum CEA and CA 15-3 as prognostic factors in primary breast cancer," *British Journal of Cancer*, vol. 86, no. 8, pp. 1217–1222, 2002.

[59] M. Gion, P. Boracchi, R. Dittadi et al., "Prognostic role of serum CA15.3 in 362 node-negative breast cancers. An old player for a new game," *European Journal of Cancer*, vol. 38, no. 9, pp. 1181–1188, 2002.

[60] E. J. Kumpulainen, R. J. Keskikuru, and R. T. Johansson, "Serum tumor marker CA 15.3 and stage are the two most powerful predictors of survival in primary breast cancer," *Breast Cancer Research and Treatment*, vol. 76, no. 2, pp. 95–102, 2002.

[61] A. Martín, M. D. Corte, A. M. Álvarez et al., "Prognostic value of pre-operative serum CA 15.3 levels in breast cancer," *Anticancer Research*, vol. 26, no. 5B, pp. 3965–3971, 2006.

[62] R. Molina, X. Filella, J. Alicarte et al., "Prospective evaluation of CEA and CA 15.3 in Patients with locoregional breast cancer," *Anticancer Research*, vol. 23, no. 2A, pp. 1035–1041, 2003.

[63] J. Madoz-Gúrpide, H. Wang, D. E. Misek, F. Brichory, and S. M. Hanash, "Protein based microarrays: a tool for probing the proteome of cancer cells and tissues," *Proteomics*, vol. 1, no. 10, pp. 1279–1287, 2001.

[64] M. J. Nam, J. Madoz-Gurpide, H. Wang et al., "Molecular profiling of the immune response in colon cancer using protein microarrays: occurrence of autoantibodies to ubiquitin C-terminal hydrolase L3," *Proteomics*, vol. 3, no. 11, pp. 2108–2115, 2003.

[65] K. Bouwman, J. Qiu, H. Zhou et al., "Microarrays of tumor cell derived proteins uncover a distinct pattern of prostate cancer serum immunoreactivity," *Proteomics*, vol. 3, no. 11, pp. 2200–2207, 2003.

[66] J. Qiu, J. Madoz-Gurpide, D. E. Misek et al., "Development of natural protein microarrays for diagnosing cancer based on an antibody response to tumor antigens," *Journal of Proteome Research*, vol. 3, no. 2, pp. 261–267, 2004.

[67] N. Ramachandran, E. Hainsworth, B. Bhullar et al., "Self-assembling protein microarrays," *Science*, vol. 305, no. 5680, pp. 86–90, 2004.

[68] N. Ramachandran, J. V. Raphael, E. Hainsworth et al., "Next-generation high-density self-assembling functional protein arrays," *Nature Methods*, vol. 5, no. 6, pp. 535–538, 2008.

[69] K. S. Anderson, S. Sibani, G. Wallstrom et al., "Protein microarray signature of autoantibody biomarkers for the early detection of breast cancer," *Journal of Proteome Research*, vol. 10, no. 1, pp. 85–96, 2011.

[70] C. L. Hattrup and S. J. Gendler, "Structure and function of the cell surface (tethered) mucins," *Annual Review of Physiology*, vol. 70, pp. 431–457, 2008.

[71] D. W. Kufe, "Mucins in cancer: function, prognosis and therapy," *Nature Reviews Cancer*, vol. 9, no. 12, pp. 874–885, 2009.

[72] R. Molina, J. M. Augé, J. M. Escudero et al., "Evaluation of tumor markers (HER-2/neu oncoprotein, CEA, and CA 15.3) in patients with locoregional breast cancer: prognostic value," *Tumor Biology*, vol. 31, no. 3, pp. 171–180, 2010.

[73] Y. Fan, J. Wang, Y. Yang et al., "Detection and identification of potential biomarkers of breast cancer," *Journal of Cancer Research and Clinical Oncology*, vol. 136, no. 8, pp. 1243–1254, 2010.

[74] M. C. Gast, E. J. van Dulken, T. K. van Loenen et al., "Detection of breast cancer by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry tissue and serum protein profiling," *International Journal of Biological Markers*, vol. 24, no. 3, pp. 130–141, 2009.

[75] M. C. Gast, C. H. van Gils, L. F. Wessels et al., "Serum protein profiling for diagnosis of breast cancer using SELDI-TOF MS," *Oncology Reports*, vol. 22, no. 1, pp. 205–213, 2009.

[76] A. Lebrecht, D. Boehm, M. Schmidt, H. Koelbl, and F. H. Grus, "Surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry to detect breast cancer markers in tears and serum," *Cancer Genomics and Proteomics*, vol. 6, no. 2, pp. 75–84, 2009.

[77] A. W. Van Winden, M. C. W. Gast, J. H. Beijnen et al., "Validation of previously identified serum biomarkers for breast cancer with SELDI-TOF MS: a case control study," *BMC Medical Genomics*, vol. 2, article 4, 2009.

[78] C. Belluco, E. F. Petricoin, E. Mammano et al., "Serum proteomic analysis identifies a highly sensitive and specific discriminatory pattern in stage 1 breast cancer," *Annals of Surgical Oncology*, vol. 14, no. 9, pp. 2470–2476, 2007.

[79] G. Ricolleau, C. Charbonnel, L. Lodé et al., "Surface-enhanced laser desorption/ionization time of flight mass spectrometry protein profiling identifies ubiquitin and ferritin light chain as prognostic biomarkers in node-negative breast cancer tumors," *Proteomics*, vol. 6, no. 6, pp. 1963–1975, 2006.

[80] C. Laronga, S. Becker, P. Watson et al., "SELDI-TOF serum profiling for prognostic and diagnostic classification of breast cancers," *Disease Markers*, vol. 19, no. 4-5, pp. 229–238, 2003.

[81] J. Li, Z. Zhang, J. Rosenzweig, Y. Y. Wang, and D. W. Chan, "Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer," *Clinical Chemistry*, vol. 48, no. 8, pp. 1296–1304, 2002.

[82] E. R. Sauter, W. Davis, W. Qin et al., "Identification of a β-casein-like peptide in breast nipple aspirate fluid that is associated with breast cancer," *Biomarkers in Medicine*, vol. 3, no. 5, pp. 577–588, 2009.

[83] J. Zhou, B. Trock, T. N. Tsangaris et al., "A unique proteolytic fragment of alpha1-antitrypsin is elevated in ductal fluid of breast cancer patient," *Breast Cancer Research and Treatment*, vol. 123, no. 1, pp. 73–86, 2010.

[84] J. L. Noble, R. S. Dua, G. R. Coulton, C. M. Isacke, and G. P. H. Gui, "A comparative proteinomic analysis of nipple aspiration fluid from healthy women and women with breast cancer," *European Journal of Cancer*, vol. 43, no. 16, pp. 2315–2320, 2007.

[85] J. He, J. Gornbein, D. Shen et al., "Detection of breast cancer biomarkers in nipple aspirate fluid by SELDI-TOF and their identification by combined liquid chromatography-tandem mass spectrometry," *International Journal of Oncology*, vol. 30, no. 1, pp. 145–154, 2007.

[86] J. Li, J. Zhao, X. Yu et al., "Identification of biomarkers for breast cancer in nipple aspiration and ductal lavage fluid," *Clinical Cancer Research*, vol. 11, no. 23, pp. 8312–8320, 2005.

[87] T. M. Pawlik, H. Fritsche, K. R. Coombes et al., "Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry," *Breast Cancer Research and Treatment*, vol. 89, no. 2, pp. 149–157, 2005.

[88] E. R. Sauter, S. Shan, J. E. Hewett, P. Speckman, and G. C. Du Bois, "Proteomic analysis of nipple aspirate fluid using SELDl-TOF-MS," *International Journal of Cancer*, vol. 114, no. 5, pp. 791–796, 2005.

[89] C. P. Paweletz, B. Trock, M. Pennanen et al., "Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF:

potential for new biomarkers to aid in the diagnosis of breast cancer," *Disease Markers*, vol. 17, no. 4, pp. 301–307, 2001.

[90] T. M. Pawlik, D. H. Hawke, Y. Liu et al., "Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein," *BMC Cancer*, vol. 6, article no. 68, 2006.

[91] P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, and R. A. Dwek, "Glycosylation and the immune system," *Science*, vol. 291, no. 5512, pp. 2370–2376, 2001.

[92] A. Kobata and J. Amano, "Altered glycosylation of proteins produced by malignant cells, and application for the diagnosis and immunotherapy of tumours," *Immunology and Cell Biology*, vol. 83, no. 4, pp. 429–439, 2005.

[93] D. H. Dube and C. R. Bertozzi, "Glycans in cancer and inflammation—Potential for therapeutics and diagnostics," *Nature Reviews Drug Discovery*, vol. 4, no. 6, pp. 477–488, 2005.

[94] T. F. Ørntoft and E. M. Vestergaard, "Clinical aspects of altered glycosylation of glycoproteins in cancer," *Electrophoresis*, vol. 20, no. 2, pp. 362–371, 1999.

[95] O. J. Semmes, G. Malik, and M. Ward, "Application of mass spectrometry to the discovery of biomarkers for detection of prostate cancer," *Journal of Cellular Biochemistry*, vol. 98, no. 3, pp. 496–503, 2006.

[96] N. K. Wong, R. L. Easton, M. Panico et al., "Characterization of the oligosaccharides associated with the human ovarian tumor marker CA125," *Journal of Biological Chemistry*, vol. 278, no. 31, pp. 28619–28634, 2003.

[97] S. Prakash and P. W. Robbins, "Glycotyping of prostate specific antigen," *Glycobiology*, vol. 10, no. 2, pp. 173–176, 2000.

[98] T. M. Block, M. A. Comunale, M. Lowman et al., "Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 3, pp. 779–784, 2005.

[99] R. Peracaula, G. Tabarés, L. Royle et al., "Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins," *Glycobiology*, vol. 13, no. 6, pp. 457–470, 2003.

[100] D. F. Hayes, M. Abe, J. Siddiqui, C. Tondini, and D. W. Kufe, "Clinical and molecular investigations of the DF3 breast cancer-associated antigen," in *Immunological Approaches to the Diagnosis and Therapy of Breast Cancer II*, R. L. Ceriani, Ed., pp. 45–53, Plenum, New York, NY, USA, 1989.

[101] S. R. Hull, A. Bright, K. L. Carraway, M. Abe, D. F. Hayes, and D. W. Kufe, "Oligosaccharide differences in the DF3 sialomucin antigen from normal human milk and the BT-20 human breast carcinoma cell line," *Cancer Communications*, vol. 1, no. 4, pp. 261–267, 1989.

[102] L. Perey, D. F. Hayes, and D. Kufe, "Effects of differentiating agents on cell surface expression of the breast carcinoma-associated DF3-P epitope," *Cancer Research*, vol. 52, no. 22, pp. 6365–6370, 1992.

[103] L. Perey, D. F. Hayes, P. Maimonis, M. Abe, C. O'Hara, and D. W. Kufe, "Tumor selective reactivity of a monoclonal antibody prepared against a recombinant peptide derived from the DF3 human breast carcinoma-associated antigen," *Cancer Research*, vol. 52, no. 9, pp. 2563–2568, 1992.

[104] R. Sewell, M. Bäckström, M. Dalziel et al., "The ST6GalNAc-I sialyltransferase localizes throughout the golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer," *Journal of Biological Chemistry*, vol. 281, no. 6, pp. 3586–3594, 2006.

[105] J. M. Burchell, A. Mungul, and J. Taylor-Papadimitriou, "O-linked glycosylation in the mammary gland: changes that occur during malignancy," *Journal of Mammary Gland Biology and Neoplasia*, vol. 6, no. 3, pp. 355–364, 2001.

[106] M. T. Goodarzi and G. A. Turner, "Decreased branching, increased fucosylation and changed sialylation of alpha-1-proteinase inhibitor in breast and ovarian cancer," *Clinica Chimica Acta*, vol. 236, no. 2, pp. 161–171, 1995.

[107] J. W. Dennis and S. Laferte, "Oncodevelopmental expression of -GlcNAcβ1-6Manα1-6Manβ1- branched asparagine-linked oligosaccharides in murine tissues and human breast carcinomas," *Cancer Research*, vol. 49, no. 4, pp. 945–950, 1989.

[108] B. Fernandes, U. Sagman, M. Auger, M. Demetrio, and J. W. Dennis, "β1-6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia," *Cancer Research*, vol. 51, no. 2, pp. 718–723, 1991.

[109] K. L. Abbott, K. Aoki, J. M. Lim et al., "Targeted glycoproteomic identification of biomarkers for human breast carcinoma," *Journal of Proteome Research*, vol. 7, no. 4, pp. 1470–1480, 2008.

[110] U. M. Abd Hamid, L. Royle, R. Saldova et al., "A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression," *Glycobiology*, vol. 18, no. 12, pp. 1105–1118, 2008.

[111] Z. Kyselova, Y. Mechref, P. Kang et al., "Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles," *Clinical Chemistry*, vol. 54, no. 7, pp. 1166–1175, 2008.

[112] W. R. Alley, M. Madera, Y. Mechref, and M. V. Novotny, "Chip-based reversed-phase liquid chromatography-mass spectrometry of permethylated N-linked glycans: a potential methodology for cancer-biomarker discovery," *Analytical Chemistry*, vol. 82, no. 12, pp. 5095–5106, 2010.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 896476, 16 pages
doi:10.1155/2011/896476

Research Article

Proteomic-Based Biosignatures in Breast Cancer Classification and Prediction of Therapeutic Response

Jianbo He,[1,2] Stephen A. Whelan,[3] Ming Lu,[4] Dejun Shen,[5] Debra U. Chung,[2] Romaine E. Saxton,[1,2] Kym F. Faull,[6] Julian P. Whitelegge,[6] and Helena R. Chang[1,2,7]

[1] *Gonda/UCLA Breast Cancer Research Laboratory, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA 90095, USA*
[2] *Department of Surgery, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA 90095, USA*
[3] *Cardiovascular Proteomics Center, Center for Biomedical Mass Spectrometry, Boston University School of Medicine, Boston, MA 02118, USA*
[4] *Department of Medicine, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA 90095, USA*
[5] *Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA*
[6] *Pasarow Mass Spectrometry Laboratory, Semel Institute and Department of Psychiatry and Biobehavioral Science, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA 90095, USA*
[7] *Revlon/UCLA Breast Center, David Geffen School of Medicine at UCLA, 200 UCLA Medical Plaza, B265, Los Angeles, CA 90095, USA*

Correspondence should be addressed to Helena R. Chang, hchang@mednet.ucla.edu

Received 22 June 2011; Accepted 12 August 2011

Academic Editor: David E. Misek

Copyright © 2011 Jianbo He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Protein-based markers that classify tumor subtypes and predict therapeutic response would be clinically useful in guiding patient treatment. We investigated the LC-MS/MS-identified protein biosignatures in 39 baseline breast cancer specimens including 28 HER2-positive and 11 triple-negative (TNBC) tumors. Twenty proteins were found to correctly classify all HER2 positive and 7 of the 11 TNBC tumors. Among them, galectin-3-binding protein and ALDH1A1 were found preferentially elevated in TNBC, whereas CK19, transferrin, transketolase, and thymosin β4 and β10 were elevated in HER2-positive cancers. In addition, several proteins such as enolase, vimentin, peroxiredoxin 5, Hsp 70, periostin precursor, RhoA, cathepsin D preproprotein, and annexin 1 were found to be associated with the tumor responses to treatment within each subtype. The MS-based proteomic findings appear promising in guiding tumor classification and predicting response. When sufficiently validated, some of these candidate protein markers could have great potential in improving breast cancer treatment.

1. Introduction

Chemotherapy has long been used to treat all types of cancer. Although survival benefits from adjuvant systemic chemotherapy in breast cancer have been thoroughly documented [1], success is not uniform with many still dying after the initial chemotherapy. The unpredictable tumor response to chemotherapy in any given patient and the significant toxicity manifested in all demand a better strategy for delivering cancer therapy.

In selective subtypes of breast cancer, therapies targeting specific signal transduction and/or metabolic pathways have been successful. For example, Herceptin for HER2/neu positive breast cancer [2, 3] and poly(ADP ribose) polymerase (PARP) inhibitors for triple-negative breast cancer with defective DNA-repair [4, 5] are among the recent successes of targeted therapy. The success of target therapy has led to an explosion of interest in developing tailored systemic therapy.

Breast cancer is a heterogeneous disease molecularly, histologically, and clinically. Clinical outcomes from the same treatment vary widely even among patients with tumors of identical stage and histology. Breast cancers developed from an accumulation of genetic alterations may partially explain the differences observed including tumor responses

to anticancer agents [6]. Recently gene expression analysis has identified five subtypes of breast cancer which overlaps with clinical tumor classification according to the expression of three biomarkers, estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor 2 (HER2). Clinically these three markers are prognostically and therapeutically important in guiding treatment selection [7–10]; however, they do not fully reflect the complexity and heterogeneity of breast cancer and do not always predict the outcome of the treatment. For example, Herceptin as a single agent or in combination with chemotherapy has been shown to reduce recurrent disease and to save lives in patients with HER2-positive breast cancer, yet a significant number of HER2 overexpression tumors do not respond to the treatment [11]. Additional molecular targets are expected to improve tailored treatment in the future.

Proteomics has been employed in recent years to identify new disease-related biomarkers for cancer diagnosis and implementation of tailored treatment [12–15]. The tumor proteomes representing a global protein expression of cancer may provide new insights into the molecules that govern the dynamic cellular activities of tumor cells. Therefore, we choose to study breast tumor protein signatures in breast cancer classification and in predicting tumor response to treatment.

Previously we used SELDI mass spectrometry to profile tumor response to neoadjuvant treatment and found that significant m/z profile differences existed between cancers of nonresponders (tumor regression rate ≤25%) and others (tumor regression rate >25%) [16]. In this current study we have applied the LC-MS/MS technology to study the breast cancer proteomes in human tissues and identify unique proteins that may have the potential to separate two subtypes of breast cancer (TNBC versus HER2+) and to predict drug responses within each subtype.

2. Materials and Methods

2.1. Collection of Breast Tumor Tissues and Classification of Response. Breast tumors were collected, processed, and banked as previously described [16]. This study was approved by the UCLA institutional review board (IRB). Tumors from 39 consented patients with locally advanced breast cancer were collected from a neoadjuvant clinical trial [17]. Eleven were triple-negative breast tumors (TNBC, ER-/PR-/HER2-), and 28 were HER2-positive tumors (HER2+). The tumor specimens were uniformly collected according to a standard operating procedure established in our laboratory. Baseline tumor specimens were obtained by either core needle biopsy or surgical biopsy before starting the neoadjuvant Taxotere/Carboplatin/±Herceptin treatment (TC ± H). Evaluation of tumor response to the treatment was measured both by pathologic examination of surgically removed tissue and by clinical assessment including physical examination and/or imaging studies. The pathological response of the tumor was reported as either pathologically complete response (pCR) or having residual tumor. Because a baseline tumor size by pathologic evaluation was not possible in patients receiving neoadjuvant treatment, the clinically or imaging-measured tumor size prior to chemotherapy was used as the baseline tumor size. Pathological assessment after chemotherapy including tumor size, lymph node staging, and tumor biomarkers was performed on the specimen obtained from the definitive breast cancer surgery [16]. The tumor regression rate (TRR) was used to evaluate tumor response induced by neoadjuvant therapy, and it was calculated as follows: (baseline tumor size − residual tumor size)/baseline tumor size × 100%, where the baseline tumor size was measured clinically, and the postchemotherapy residual invasive tumor size was measured pathologically. The tumor response was categorized into three groups: responders (TRR > 75%, R), intermediate responders (25% < TRR ≤ 75%, IR), and nonresponders (TRR ≤ 25%, NR).

2.2. Protein Extraction and Abundant Protein Depletion. Protein extraction from tumors and depletion of abundant proteins from tumor lysates were performed as previously described [16]. Briefly, frozen tumors were homogenized in liquid nitrogen and suspended in 1% Triton X-100. The samples were refrozen at −80°C and thawed on ice twice. Following centrifugation (10,000 g, 10 min, 4°C), the supernatants were subjected to albumin and immunoglobulin depletion using an albumin and IgG removal kit (Amersham) as well as hemoglobin depletion using Ni-NTA magnetic agarose beads (Qiagen). Protein concentrations of each preparation were determined by the BioRad protein assays.

Because the blood proteins in the breast cancer tissue can cause significant ion suppression of lower abundance cancer-related proteins/peptides which may mask ion signals of less abundant peptides with similar M/Z ratios and retention times. In addition, the over presentation of serum proteins in the specimen may lower the amount of the cancer-related proteins available for LC-MS/MS analysis [18]. As a result, selected abundant serum proteins were depleted from the tissue extracts. Our preliminary test has shown more than 95% albumin, IgG and hemoglobin were removed by the described method, and more meaningful proteins have been detected.

2.3. Trypsin Digestion. The dried protein samples were dissolved in 6 M guanidine HCl, reduced with DTT (5 mM–15 mM), and alkylated using 10 mM iodoacetamide. Samples were then diluted with NH_4HCO_3 to lower guanidine HCl concentration (1 M), mixed with trypsin (1 : 50 w/w ratio, sequencing grade, Promega) containing 50 mM ammonium bicarbonate, and incubated at 37°C overnight. Samples were desalted by C18 Microspin columns (The Nest Group), and the eluates were dried in a vacuum centrifuge.

2.4. LC-MS/MS Analysis. Each digested and dried sample was prepared for LC-MS/MS analysis as previously reported [19]. Briefly the samples were redissolved in Buffer A (H_2O/acetonitrile/formic acid, 98.9/1/0.1, typically at 0.7 μg protein/uL), and aliquots were injected (5 μL) onto an

in-house-prepared C18 trap. The retained materials were placed onto a reverse phase column (New Objective C18, 15 cm, 75 μM diameter, 5 μm particle size equilibrated in Buffer A) and eluted (300 nL/min, Eksigent Nanolc-2D) with an increasing concentration of Buffer B (acetonitrile/water/formic acid, 98.9/1/0.1; min/%B: 0/5, 10/10, 112/40, 130/60, 135/90, 140/90). Eluted peptides were analyzed by MS and data-dependent MS/MS (collision-induced dissociation) using online data-dependent tandem mass spectrometry (LTQ Orbitrap, Thermo Fisher Scientific) in which the seven most abundant precursor ions were selected for MS/MS. Before testing the experimental specimens, the reproducibility of LC-MS/MS analysis was confirmed by examining the triplicates of two different tissue samples, and similar proteins were identified from the triplicates of each sample with more than 90% overlapping.

2.5. Database Searching and Analysis. BioWorks software (version 3.3.1, Thermo Fisher Sceintific), based on the SEQUEST algorithm (SRF v.5, Thermo Fisher Scientific), was used to search the mass spectra against a human trypsin indexed database (human.fasta.hdr database, Version 12.2, 227246 entries) as described by Whelan et al. [19]. SEQUEST was searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 50 PPM. The search tolerated up to two missed trypsin cleavages with variable modifications for carboxyamidomethylation (57.02146 Da) and methionine oxidation (15.99492 Da). Scaffold (version 3.0.3, Proteome Software, Inc.) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet Algorithm [20]. Protein identifications were accepted if they could be established with a greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet Algorithm [21].

From the resulting MS/MS protein identifications, a list of proteins was generated for each sample. A list of semi-quantitative protein abundances in the different samples was developed using the normalized spectrum counts of the identified tryptic peptides from each protein, as compiled by the Scaffold program. The protein lists and their relative abundances were then compared to find differentially expressed proteins between two groups.

2.6. Statistic Analysis. The data files exported from Scaffold were further processed as Excel files. The top 60% (180) abundant proteins of the 315 identified proteins were further selected for hierarchical clustering and supervised classification studies. Those proteins with at least a 2-fold difference in mean spectral counts between any two groups were selected for analysis in the web-based Gene Expression Profile Analysis Suite (GEPAS, version 4.0, http://www.gepas.org). Five different classification algorithms were tested to select candidate markers in the GEPAS software: Support Vector Machines (SVM), K-Nearest Neighbor Clustering (KNN), Diagonal Linear Discriminant Analysis (DLDA), Prediction Analysis with Microarrays (PAM), and Self-Organizing Map (SOM).

2.7. Immunohistochemistry Staining. A small portion of each of the baseline tumors was embedded in OCT and stored at -80°C. Endogenous peroxidase activity was quenched with 0.6% hydrogen peroxide in methanol for 10 minutes, and endogenous biotin was eliminated by Biotin Blocking System (DAKO, x0590). After blocking with 1 : 5 diluted normal goat serum or fetal bovine serum, slides were incubated for 1 hour with primary antibody (CK19, mouse IgG, ready to use, DAKO; galectin-3-binding protein, Goat IgG, 1 : 200 dilution, R&D) and 30 minutes with biotinylated secondary antibody (biotinylated anti-mouse Ig, 1 : 800 dilution, DAKO; biotinylated anti-goat Ig, 1 : 200 dilution, Vector Labs). Antigen-antibody complexes were then detected by the StreptABComplex/HRP method (DAKO) using diaminobenzidine as a chromogenic substrate (DAKO). Immunostained slides were lightly counterstained with hematoxylin. For negative controls, primary antibodies were replaced by mouse IgG or goat IgG.

3. Results

3.1. Patient Characteristics. The reported thirty-nine baseline tumor specimens included 28 HER2-positive breast cancers and 11 TNBC with HER2 status determined by fluorescence in situ hybridization (FISH) assay. Fifteen of the 28 HER2+ patients were randomized to receive TC, and the remaining 13 received TC and Herceptin (TCH) before surgery. All eleven patients with TNBC received neoadjuvant TC. Following the neoadjuvant treatment, 28 patients with HER2+ tumors showed 12 responders (R), including 7 with pathological complete response (pCR) and 5 with a tumor regression rate >75%, 12 intermediate responders (IR), and 4 nonresponders (NR). In the TNBC group, there were 7 responders including 6 pCR and 1 with tumor regression rate >75%, 3 IR, and 1 NR. The clinical characteristics and pathologic features of the 11 TNBC and 28 HER2+ cases are summarized in Tables 1(a) and 1(b).

3.2. Protein Comparison between HER2+ and TNBC Groups. Proteins identified by MS/MS from the 39 tumors showed that 48 proteins were only found in HER2+ tumors, 24 were only seen in TNBC, and 243 proteins were shared by both, but the quantity of the shared proteins differed widely in the two tumor types. In this study, we focused the analysis on the top 60% abundant proteins (180/315) detected in the 39 tumors.

The 20 most abundant shared proteins by both subtypes of cancer were summarized in Table 2. Among them apolipoprotein A-I and D, enolase 1, tumor rejection antigen (gp96) 1, transgelin 2, cofilin 1, profilin, heat shock proteins 70, and annexins 5 were found to be present in significant quantity in both types of breast cancer. Some of these shared proteins found in sufficient amount may be useful for breast cancer detection.

Table 1

(a) Clinical characteristics of 11 TNBC tumors

LTQ Orbitrap sample ID	Patient age	Ethnicity	TR %	Response	T stage	Histological type	ER	PR	FISH R/G ratio	Neoadjuvant
#1	61	White	80	R	T3	IDC	−	−	1.10	TC
#2	29	Hispanic	−60	NR	T3	IDC	−	−	0.92	TC
#5	55	Hispanic	100	R (pCR)	T3	IDC	−	−	0.92	TC
#6	54	Hispanic	100	R (pCR)	T3	IDC	−	−	1.01	TC
#7	40	Asian	45	IR	T3	IDC	−	−	1.17	TC
#8	44	White	100[a]	R (pCR)	T3	IDC	−	−	1.00	TC
#9	49	Hispanic	48	IR	T4	IDC	−	−	1.10	TC
#10	53	White	100	R (pCR)	T3	IDC	−	−	1.20	TC
#11	84	Asian	100	R (pCR)	T4	IDC	−	−	1.03	TC
#36	45	Hispanic	30	IR	T3	IDC	−	−	1.27	TC
#37	38	White	100	R (pCR)	T2	IDC	−	−	1.10	TC

[a] LN positive without residual primary cancer.

(b) Clinical characteristics of 28 HER2+ tumors

LTQ Orbitrap sample ID	Patient age	Ethnicity	TRR %	Response	T stage	Histological type	ER	PR	FISH R/G ratio	Neoadjuvant
#17	38	White	40	IR	T3	IDC	+	−	12.4	TC
#18	63	Asian	100	R (pCR)	T3	IDC	+	−	12.7	TCH
#19	57	White	100	R (pCR)	T3	IDC	−	−	4.6	TC
#20	56	Asian	78.2	R	T4	IDC	+	−	10.71	TC
#21	51	Black	56	IR	T3	IDC	−	−	19.97	TCH
#22	31	White	45.5	IR	T3	IDC	+	+	2.2	TC
#23	55	White	80	R	T4	IDC	+	+	3.8	TC
#24	45	Asian	75	IR	T4	IDC	+	+	2.7	TCH
#25	42	Hispanic	63.5	IR	T4	IDC	+	−	2.5	TC
#26	50	White	67.1	IR	T3	IDC	−	−	2.41	TCH
#27	33	White	82.9	R	T3	IDC	+	+	3.03	TCH
#28	40	White	66.7	IR	T3	IDC	−	−	8.1	TC
#29	35	Hispanic	100[a]	R (pCR)	T3	IDC	−	−	42.2	TCH
#30	44	White	97.3	R	T4	IDC	−	−	4.2	TC
#31	30	White	−7.7	NR	T3	IDC	+	−	5	TC
#32	57	White	25%	NR	T4	IDC	+	−	>4	TCH
#33	37	White	33.3	IR	T2	IDC	+	+	9.49	TC
#34	36	Black	25	NR	T3	IDC	−	−	5.1	TC
#35	42	White	60	IR	T2	IDC	+	+	4.5	TCH
#38	55	White	42.3	IR	T4	IDC	−	−	3.9	TCH
#39	47	White	100[b]	R (pCR)	T3	IDC	+	+	>20	TCH
#40	50	Asian	100[b]	R (pCR)	T3	IDC	+	+	4.19	TCH
#41	58	White	50	IR	T4	IDC	+	+	2.1	TCH
#42	40	Asian	60	IR	T2	IDC	−	−	16	TC
#43	37	White	−85.6	NR	T3	IDC	+	+	3.1	TC
#44	49	White	100[b]	R (pCR)	T2	IDC	+	−	7.7	TCH
#45	55	Asian	92.6	R	T3	IDC	−	−	9.9	TC
#46	41	Hispanic	100	R (pCR)	T2	IDC	−	−	9.2	TC

[a] LN positive and residual DCIS; [b] residual DCIS only.

Table 2: The 20 most abundant proteins shared by both HER2-positive and TNBC tumors.

Identified proteins	Accession no.	MW
Apolipoprotein A-I	gi|90108664	28 kDa
Vimentin	gi|62414289	54 kDa
Enolase 1	gi|4503571	47 kDa
Alpha-1 antitrypsin	gi|157086955	45 kDa
Triosephosphate isomerase 1	gi|4507645 (+2)	27 kDa
Cyclophilin A	gi|1633054	18 kDa
Apolipoprotein D	gi|619383	28 kDa
Cofilin 1	gi|5031635	19 kDa
Chaperonin	gi|31542947	61 kDa
Transgelin 2	gi|4507357	22 kDa
Heat shock 70 kDa protein 5	gi|16507237	72 kDa
Tumor rejection antigen (gp96) 1	gi|4507677	92 kDa
S100 calcium-binding protein A11	gi|5032057	12 kDa
Lumican precursor	gi|4505047	38 kDa
Tropomyosin 4	gi|4507651	29 kDa
ATP synthase, H+ transporting, mitochondrial F1 complex	gi|32189394	57 kDa
Prosaposin isoform a preproprotein	gi|11386147	58 kDa
Profilin	gi|157838211 (+4)	15 kDa
Heat shock 70 kDa protein 8 isoform 1	gi|5729877	71 kDa
Annexin 5	gi|4502107	36 kDa

Of the 180 top abundant proteins observed in the 39 breast cancer specimens, 61 were found to have a ≥2-fold difference of spectrum counts between the two subtypes of breast cancer (HER2+ versus TNBC). Because some of these proteins were not detected in every sample, we further refined the list of differential proteins by selecting only those detected in ≥50% of the cases in either group. The selected 44 differentially expressed proteins were tested by hierarchical clustering to classify HER2+ breast cancer versus TNBC. These differentially expressed proteins correctly classified all 28 HER2+ tumors and 8 of the 11 TNBC by unweighted pair-group method using arithmetic average (UPGMA) (Figure 1).

Self-validation of selected proteins in tumor classification was tested using a supervised classification. The 44 differentially expressed proteins were used to build a model separating subtypes of the tumors by Prophet, a web interface from the Gene Expression Profile Analysis Suite. Error rates were calculated as the number of misclassified tumors divided by total tumor cases tested. The error rates using various numbers of proteins by different models (SVM, KNN, DLDA, PAM, and SOM) were estimated by leaving-one-out tests (see File A in Supplementary Material available online at doi:10.1155/2011/896476). SVM had the lowest error rate (10%, 4/39) with 90% accuracy in tumor classification. The top 20 protein candidates (Table 3) selected by SVM model successfully classified all 28 HER2+ tumors and 7 of the 11 TNBC. Among the 20 differentially expressed proteins, G3BP, ALDH1A1, and complement component 1 inhibitor overexpression were found to be associated with TNBC subtype, whereas overexpression of CK19, transferrin, transketolase, and thymosin β4 and β10 were associated with HER2+ tumors (Figure 2).

3.3. Proteins Correlated with Different Tumor Response to Neoadjuvant Treatment among HER2+ Tumors. Of the 28 HER2+ tumors, there were 12 R (including 7 with pCR), 12 IR, and 4 NR. We compared proteomic differences between the two groups with extreme tumor response (pCR and NR) and found that 48 of the 180 proteins had an expressional difference ≥2-fold between 7 pCR versus 4 NR tumors. Self-validation of these potential marker proteins by five supervised classification methods suggested that the KNN had the lowest error rate (9%, 1/11) in predicting tumor response (Files B and C). By using KNN = 1 method, 100% (4/4) NR and 85.7% (6/7) pCR were correctly grouped by 20 selected proteins (Table 4). Of the 20 proteins, overexpressions of enolase1, vimentin, and L-plastin in HER2-positive tumors were associated with pCR, whereas high level of heat shock proteins 70 (Hsp70) and peroxiredoxin 5 (Prx V) were found only in the NR cases.

3.4. Proteins Predicting TNBC Tumor Response to Neoadjuvant Treatment. Among the 11 TNBC cases, there were 7 R (including 6 pCR), 3 IR, and 1 NR. Due to the small sample size, the proteins of responders' tumor (R) were compared to all the remaining tumors with less response (IR + NR). Sixty-three of 180 proteins had a ≥2-fold mean differences between the two groups of TNBC with different response to the same treatment. Self-validation of these proteins by five supervised classification methods was used to compare the accuracy in predicting a tumor response. Using DLDA method, 6 of 7 tumors in the R group and 3 of 4 tumors in IR/NR group were correctly classified by the 30 selected proteins (error rate 18%) (Files D and E). Of these 30 proteins, the increased heat shock 70 kDa protein 8, periostin, Ras homolog gene family member A (RhoA), actinin alpha 4, cathepsin D preproprotein, annexin 1, and several other proteins were associated with drug resistance in TNBC (Table 5).

3.5. Evaluation of CK19 and G3BP Expression in TNBC and HER2+ Frozen Tumors. CK19 and G3BP protein expressions were tested in breast cancer tumors by immunohistochemistry. The CK19 and G3BP staining showed cytoplasmic/membrane staining pattern in breast cancer cells. The overexpression of CK19 was found in HER2+ breast tumors (Figure 3) while the expression of G3BP was found to be upregulated in most TNBC (Figure 4). The concordance findings between the mass spectrometry analysis and immunohistochemical staining of the same tumor suggested that high-throughput mass spectrometry may be used as a screening tool to discover disease-related biomarkers.

Table 3: Top 20 differentially expressed proteins selected by supervised classification methods for classifying two tumor subtypes.

Rank	Accession no.	Protein name	MW	HER2+/TNBC mean	Subcellular location	Function
1	gi|10946578	Thymosin β4	5 kDa	2.99	Cytoplasm, cytoskeleton	For cytoskeletal binding, involved in cell growth and maintenance
2	gi|4507521	Transketolase	68 kDa	4.20	Cytosol	Involved in metabolism. Associated with cell proliferation of uterine and cervical cancer.
3	gi|1633054	Cyclophilin A	18 kDa	2.45	Cytoplasma	Involved in accelerate the folding of proteins
4	gi|73858568	Complement component 1 inhibitor	55 kDa	0.33	Secreted	Regulating the complement cascade
5	gi|4557871	Transferrin	77 kDa	16.38	Secreted	Essential for cell growth and iron-dependent metabolic processes
6	gi|90111766	Keratin type I cytoskeletal 19	44 kDa	11.29	Cytoskeleton	Involved in metastatic progression of breast cancer
7	gi|10863895	Thymosin β10	5 kDa	2.25	Cytoplasm, cytoskeleton	For cytoskeletal binding, involved in cell growth and maintenance
8	gi|5031863	Galectin-3-binding protein	65 kDa	0.41	Secreted	Modulating cell-cell and cell-matrix interactions
9	gi|4505753 (+1)	Phosphoglycerate mutase 1	29 kDa	2.51	Cytosol	Involved in glycolysis
10	gi|5174391	Aldo-keto reductase family 1, member A1	37 kDa	0.30	Cytosol	Involved in the reduction of biogenic and xenobiotic aldehydes
11	gi|21361176	Aldehyde dehydrogenase 1A1	55 kDa	0.39	Cytoplasm	Detoxifying enzyme responsible for oxidating of intracellular aldehydes. A marker for cancer stem cells
12	gi|4505185	Macrophage migration inhibitory factor	12 kDa	0.36	Secreted, cytoplasm	Involved in integrin signaling pathways
13	gi|4507645 (+2)	Triosephosphate isomerase 1	27 kDa	2.49	Cytosol, nucleus	Fatty acid biosynthesis, gluconeogenesis, glycolysis, lipid synthesis
14	gi|4930167	Aldolase A	39 kDa	6.41	Extracellular, cytoskeleton	Involved in glycolysis
15	gi|116241280	Adenylyl cyclase-associated protein 1 (CAP 1)	52 kDa	3.03	Membrane	Regulating filament dynamics, cell polarity and signal transduction,
16	gi|21624607 (+5)	Coactosin-like 1	16 kDa	0.42	Cytoplasm, cytoskeleton	Regulating the actin cytoskeleton
17	gi|160420317	Filamin A, alpha isoform 2	281 kDa	3.10	Cytoplasm	Anchoring transmembrane proteins to the actin cytoskeleton, scaffold for cytoplasmic signaling proteins
18	gi|6005942	Valosin-containing protein	89 kDa	3.26	Cytosol, nucleus	Fragmentation of Golgi stacks during mitosis and reassembly
19	gi|5174539	Cytosolic malate dehydrogenase	36 kDa	2.52	Cytoplasm	Involved glycolysis, oxidation reduction, and tricarboxylic acid cycle
20	gi|33286418 (+2)	Pyruvate kinase 3	58 kDa	6	Cytoplasm, nucleus	Involved in glycolysis

4. Discussion

In this discovery study, the MS-detected proteomic differences between two subtypes of breast cancer (HER2+ versus TNBC tumors) were explored, and proteomic prediction of tumor response to neoadjuvant chemotherapy was investigated. LC-MS/MS data sets of proteins from the 39 tumors analyzed allowed us to identify several candidate proteins that could classify tumor subtypes and predict tumor response to neoadjuvant chemotherapy.

Two clinical subtypes of breast cancer, HER2-positive and triple-negative breast cancers, defined by immunohistochemical staining and fluorescence in situ hybridization of three biomarkers of breast cancer have also been confirmed by gene analysis as two distinctive types of breast cancer. In this study, we reported that proteomic analysis could also separate the two subtypes by the unique biosignature associated with each type of breast cancer (Table 3). We also reported the potential of proteomic analysis in classifying drug-resistant TNBC and HER2+ breast cancer (Tables 4 and 5).

Table 4: Top 20 proteins predicting tumor response to neoadjuvant treatment in HER2-positive tumors.

Rank	Protein name	Accession no.	pCR/NR mean	Subcellular location	Function
1	Enolase 1	gi|4503571	2.59	Cytoplasm, cell membrane	Multifunctional enzyme
2	Heterogeneous nuclear ribonucleoprotein A2/B1 isoform B1	gi|14043072	3.51	Nucleus, cytoplasm	Pre-mRNA processing
3	Heat shock 70 kDa protein 1	gi|75061728	0.24	Cytoplasm	Stress response
4	Vimentin	gi|62414289	9.94	Cytosol	Class III intermediate filaments
5	Vesicle amine transport protein 1	gi|18379349	0.50	Cytoplasmic vesicle membrane	Neurotransmitter transport
6	Coronin, actin-binding protein, 1A	gi|5902134	2.00	Cytoplasm	Component of the cytoskeleton of highly motile cells
7	Fatty acid-binding protein 4	gi|4557579 (+1)	0.23	Cytoplasm, nucleus	Lipid transport protein
8	Peroxiredoxin 5	gi|15826629	0.37	Mitochondrion, cytoplasm, peroxisome	Antioxidant, oxidoreductase peroxidase
9	Heat shock 70 kDa protein 9	gi|24234688	0.15	Mitochondrion	Control of cell proliferation and cellular aging
10	Leucine aminopeptidase 3	gi|41393561	2.94	Cell membrane, secreted	Cell-cell signaling
11	Apolipoprotein D	gi|619383	2.90	Secreted	Lipid metabolic process
12	L-plastin	gi|4504965	3.14	Cytoplasm, cell membrane	Activation of T cells, intracellular protein transport
13	Anterior gradient protein 2 homolog precursor	gi|5453541	0.11	Secreted, endoplasmic reticulum	Mucus secretion
14	Heat shock 10 kDa protein 1	gi|4504523	0.37	Mitochondrion	Stress response
15	ATP synthase, H+ transporting, mitochondrial F1 complex	gi|4757810	0.41	Mitochondrion	Proton-transporting ATP synthase complex assembly
16	Glutathione transferase	gi|20664358 (+5)	3.29	Cytoplasm	Glutathione metabolic process
17	Chaperonin	gi|31542947	0.33	Mitochondrion	Stress response
18	Complement component 3 precursor	gi|115298678	3.00	Secreted	Activation of the complement system
19	Heterogeneous nuclear ribonucleoprotein D isoform a	gi|14110420	2.19	Nucleus, cytoplasm	Transcription regulation
20	Malate dehydrogenase	gi|6648067 (+1)	0.22	Cytoplasm	Tricarboxylic acid cycle

Through an extensive literature review, some of the identified proteins have reported roles that are relevant to cancer biology and treatment. In TNBC tumors, we observed that the levels of G3BP, ALDH1A1, and complement component 1 inhibitor protein were preferentially elevated. All of them have been reported to have important biological properties in cancer progression. G3BP, also known as 90-kDa Mac-2-binding protein, is a member of the beta-galactoside-binding protein family and has a role in modulating cell-cell and cell-matrix interactions. It has been shown that G3BP is overexpressed in a variety of cancer cells such as colon, gastric, and breast cancer, and its overexpression appears to correlate with tumor progression, and metastasis [22–26]. Our report is the first to describe G3BP overexpression in human TNBC by both mass spectrometry analysis and immunohistochemical staining method.

One protein correlated with triple-negative breast cancer meriting a discussion is ALDH1A1, a detoxifying enzyme responsible for oxidizing intracellular aldehydes. This process is important in early differentiation of stem cells

Table 5: Top 30 proteins predicting tumor response to neoadjuvant chemotherapy in TNBC tumors.

Rank	Protein name	Accession no.	R/IR + NR mean	Subcellular location	Function
1	Heat shock 70 kDa protein 8 isoform 1	gi|5729877	0.32		Stress response
2	Periostin precursor (PN) (osteoblast-specific factor 2)	gi|93138709	0.31	Nucleus	Transcription regulation
3	Cyclophilin A	gi|1633054	0.41	Secreted	Cell attachment adhesion and spreading
4	Tyrosine 3/tryptophan 5-monooxygenase activation protein	gi|5803225 (+1)	3.71	Nucleus	Protein binding
5	Profilin	gi|157838211 (+4)	0.32	Cytoplasm, cytoskeleton	Actin cytoskeleton organization
6	Cardiac muscle alpha actin 1 proprotein	gi|4885049	0.08	Cytoplasm, cytokeleton	actin filament-based movement, apoptosis
7	Beta actin	gi|4501885	0.22	Cytoplasm, cytokeleton	Cell motility
8	Caldesmon (CDM)	gi|2498204	0.42	Cytoplasm, cytokeleton	Actin- and myosin-binding protein
9	Tubulin β5	gi|7106439	0.19	Cytosol	Major constituent of microtubules
10	Tropomyosin 2 (beta) isoform 1	gi|42476296	0.11	Cytoplasm, cytokeleton	Binding to actin filaments
11	Actinin, α4	gi|12025678	0.11	Nucleus, cytoplasm	Protein transport
12	Ras homolog gene family, member A (RhoA)	gi|10835049 (+4)	0.33	Cytoplasm, cell membrane	Regulating a signal transduction pathway
13	Heterogeneous nuclear ribonucleoprotein K	gi|13384620	0.33	Cytoplasm, nucleus	Pre-mRNA-binding proteins
14	Tubulin α1	gi|6755901	0.36	Cytosol	Major constituent of microtubules
15	Tropomyosin 4	gi|4507651	0.35	Cytoplasm, cytokeleton	Binds to actin filaments
16	Complement component 1 inhibitor precursor	gi|73858568	4.57	Secreted	Complement pathway
17	ATP synthase, H+ transporting, mitochondrial F1 complex	gi|4757810	0.43	Mitochondrion	Proton-transporting ATP synthase complex assembly
18	Calnexin precursor	gi|10716563	0.42	Endoplasmic reticulum membrane, cell membrane	Calcium-binding protein
19	Eukaryotic translation elongation factor 1 alpha 1	gi|4503471	0.29	Cytoplasm	Protein biosynthesis
20	Annexin I	gi|4502101	0.25	Nucleus, cytoplasm, membrane	Calcium/phospholipid-binding protein
21	Triosephosphate isomerase 1	gi|4507645 (+2)	0.35	Cytosol, nucleus	Fatty acid biosynthesis, gluconeogenesis, glycolysis, lipid synthesis
22	Cathepsin D preproprotein	gi|4503143	0.35	Lysosome	proteolysis
23	Alpha glucosidase II alpha subunit isoform 2	gi|38202257	0.19	Cytosol	Glycan metabolism, N-glycan metabolism
24	Tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein	gi|4507949 (+1)	0.42	Nucleus	Protein binding
25	Thymosin β10	gi|10863895	0.42	Cytoplasm, cytokeleton	cytoskeleton organization

Table 5: Continued.

Rank	Protein name	Accession no.	R/IR + NR mean	Subcellular location	Function
26	Aconitase 2 precursor	gi|4501867	0.44	Mitochondrion	Carbohydrate metabolism, tricarboxylic acid cycle
27	Heterogeneous nuclear ribonucleoprotein D isoform a	gi|14110420	0.48	Nucleus, cytoplasm	Transcription regulation
28	Serine (or cysteine) proteinase inhibitor	gi|32454741	0.25	Secreted	Inhibits activated protein C, plasminogen activator
29	Lumican precursor	gi|4505047	0.28	Secreted	Binds to laminin
30	Apolipoprotein D	gi|619383	2.86	Secreted	Transport

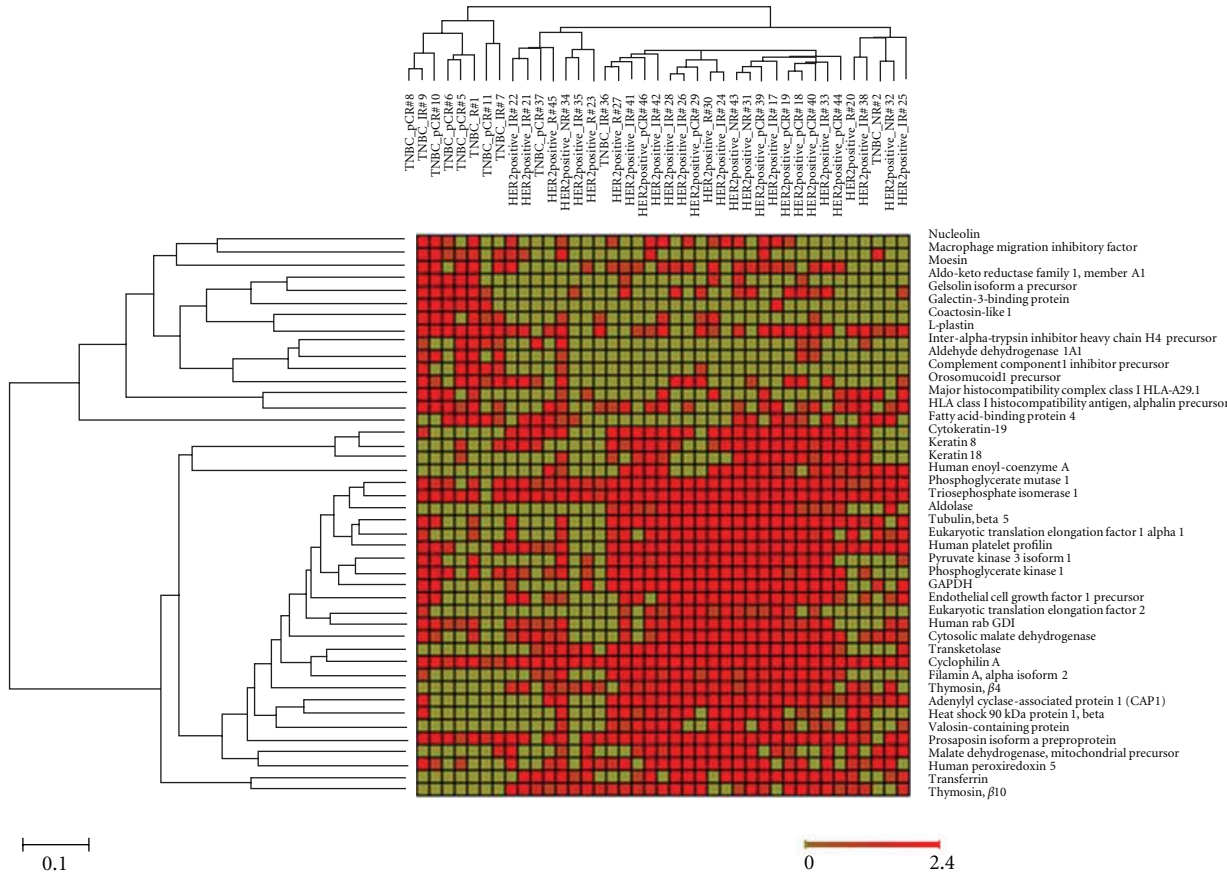
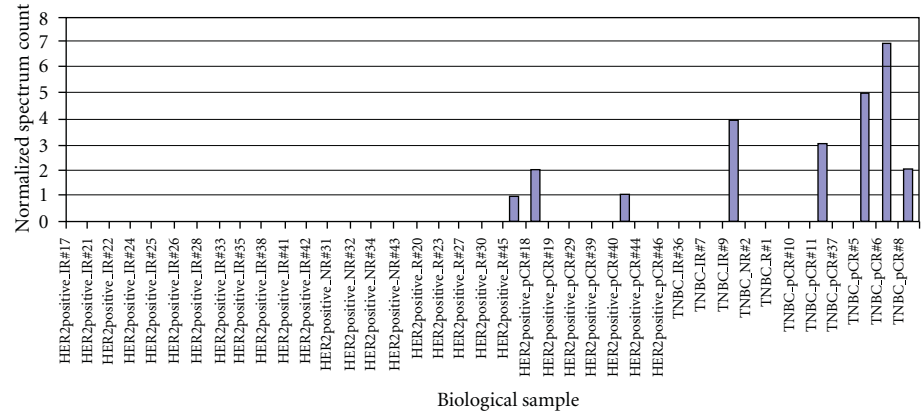
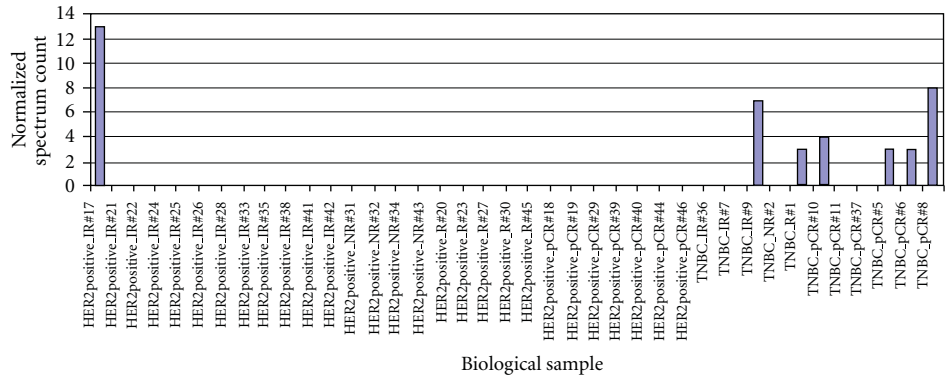


Figure 1: Heat map displaying the expression of 44 proteins in 28 HER2-positive and 11 TNBC tumors. Classification of 39 breast cancer cases into 2 groups based on tumor subtypes (HER-positive tumors and TNBC tumors) by the hierarchical clustering using GEPAS software. Each column represents a case as labeled on top, the short labeling cases are "TNBC" with sample ID, and long labeling cases are "HER2-positive" with sample ID. Each row represents a protein ID as indicated at the right. 44 proteins were expressed by ≥2-fold differences and detected in ≥50% of the cases in either group.

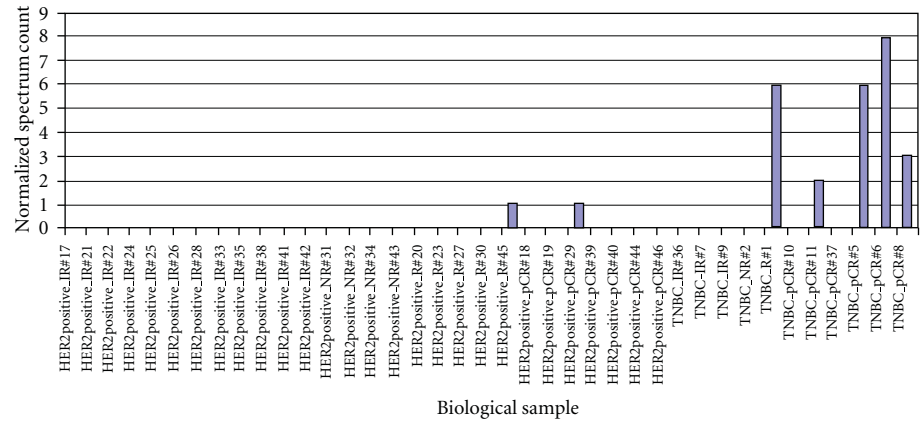
through conversion of retinol to retinoic acid [27]. ALDH1 is considered to be a breast cancer stem cell marker and also a predictor for poor prognosis [27]. Because breast cancer stem cells have been implicated in radiation and chemotherapy resistance, as well as increasing the potential for metastasis, our finding of ALDH1A1 in TNBC may explain the more frequent relapse in TNBC patients. Previously, we have observed an overexpression of ALDH1 in TNBC when compared with hormone-receptor-positive and HER2-negative breast cancer [28]. In this paper, we also found a preferential overexpression of ALDH1 in TNBC over HER2-positive tumors. The unique overexpression of ALDH1 in TNBC tumors may point out an important population as the origin of some TNBC whereby notch signaling-dependent stem cell targets may be leveraged for target therapy development [29].



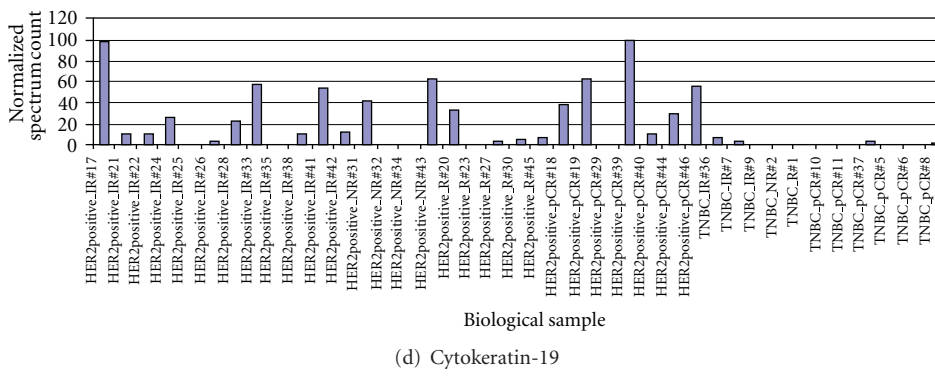
(a) ALDH1A1



(b) Galectin-3-binding protein

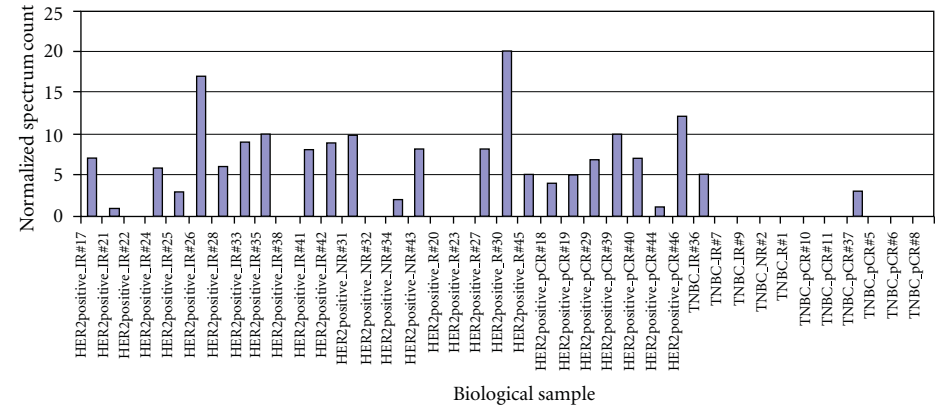


(c) Complement component 1 inhibitor

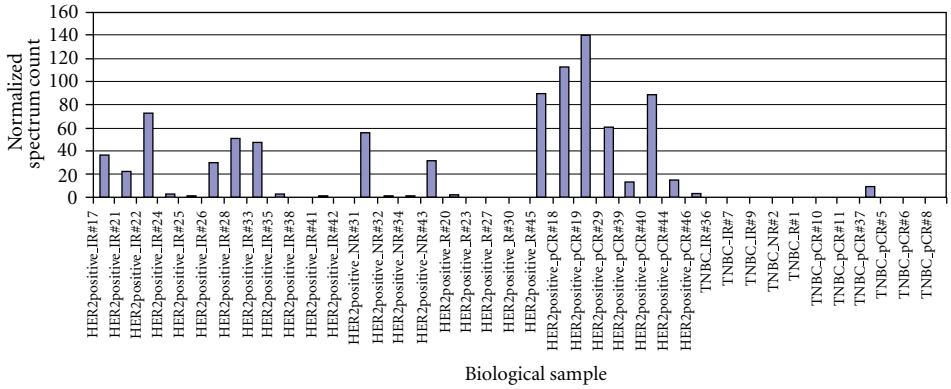


(d) Cytokeratin-19

Figure 2: Continued.



(e) Transketolase



(f) Transferrin

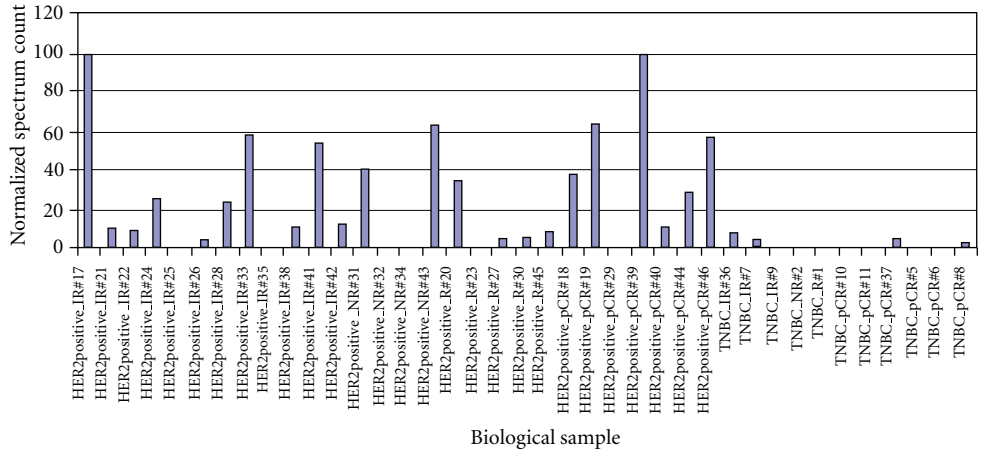
Figure 2: Representative proteins differentially expressed by HER2+ and TNBC tumors. (a)–(c): proteins preferentially expressed in TNBC. (d)-(e): proteins preferentially expressed in HER2+ tumors.

A different set of proteins was found preferentially elevated in HER2+ tumors. This list included CK19, transferrin, transketolase, and thymosin β4 and β10, and the biological significance of some of them will be discussed.

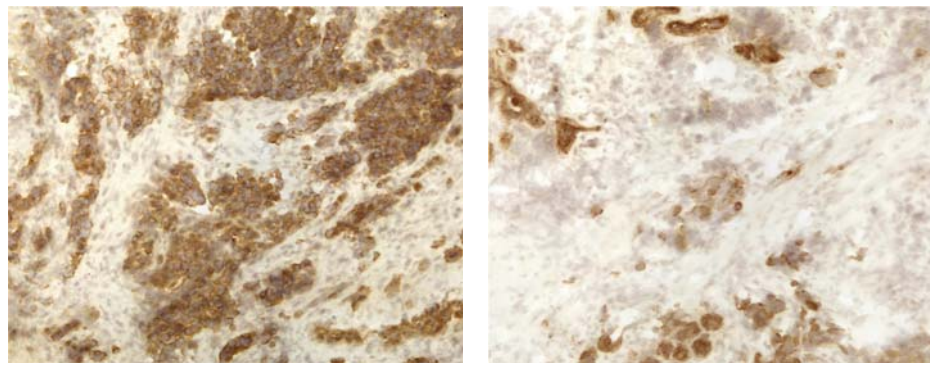
Cytokeratins are known to be important in cellular motility, signaling, and division. While CK8/CK18 were similarly detected in both HER2+ and TNBC tumors, elevated CK19 was more commonly found in HER2+ tumors. Our observation coincides with the finding reported by Schulz et al. using a combination of 2D-DIGE/mass spectrometry and western blot [30]. Both our current and previous papers suggest that CK19 is low in TNBC when compared with either HER2-positive or HER2-negative but hormone-receptor-positive breast cancer [31]. Although the biological significance and the mechanism of reduced CK19 in TNBC are not clear, it could be related to the frequent recurrence and poor overall survival rate seen in TNBC patients [32].

Transferrin, another protein associated with HER2-positive cancer, is essential for cell growth and iron-dependent metabolic activities including DNA synthesis, electron transport, and mitogenic signaling pathways [33]. The elevation of transferrin receptor (CD71) was reported to be a marker of poor outcome [33]. Vyhlidal et al. reported that transferrin is regulated by estrogen hormone [34], and tamoxifen was shown to be ineffective in ER-positive breast cancer with transferrin overexpression which coincides with the tamoxifen resistance observed in HER2-positive hormone-receptor-positive breast cancer. Taken together, the ineffectiveness of tamoxifen in women with HER2-positive and hormone-receptor-positive cancer may be related to the prevalent expression of transferrin in these tumors.

Thymosin β4 and β10 are members of a family of highly conserved small acid peptides that control the growth and differentiation of many cell types. They act as major actin-sequestering factor and play a role in cancer cell motility, invasion and metastasis [35, 36]. Thymosin β4 stimulates tumor metastasis by activating cell migration and angiogenesis in lung cancer and is associated with poor prognosis [37–39]. Elevations of thymosin β4 and β10 have also been reported in a number of other cancers including melanoma and breast cancer [40]. In the same tumor, the level of thymosin β10 was significantly higher in the cancer cells than in the normal breast parenchymal cells of the uninvolved area [41]. Its association with high-grade and poorly differentiated cancer cells is consistent with our findings of thymosin β10 overexpression in HER2-positive breast cancer. Further studies are required to confirm its



(a) Cytokeratin-19



(b) (c)

Figure 3: CK19 expressions detected by LC-MS/MS and immunohistochemical (IHC) staining. Elevated CK19 expressions found in HER+ tumor group by LC-MS/MS and confirmed by IHC in most of the frozen HER2+ tumors. (a) Normalized spectrum count of CK19 detected in 39 breast cancer tissues. (b) Immunohistochemical staining of CK19 in a HER2+ frozen tumor (power 200x). (c) Immunohistochemical staining of CK19 in a TNBC frozen tumor (power 200x).

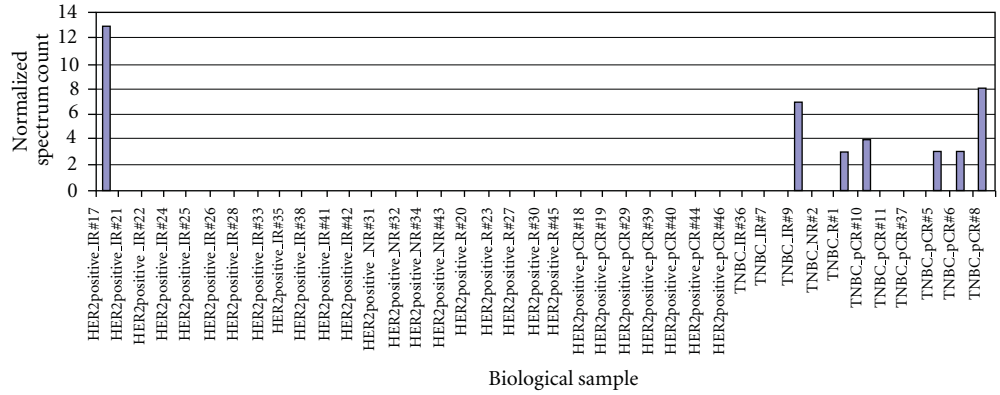
overexpression and to determine its role in HER2-positive breast cancer.

In this study, we also reported the MS-identified protein signature predicting drug-induced tumor response in HER2-positive tumors. We found that enolase 1, vimentin, L-plastin, and ApoD predicted a favorable response of HER2-positive tumors. In contrast, elevated peroxiredoxin 5 and heat shock proteins 70 were found in nonresponding HER2-positive tumors.

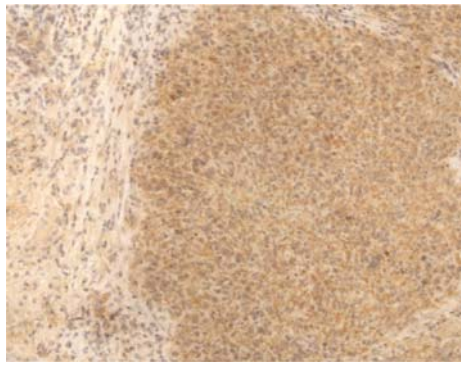
Enolase 1(ENO1), a phosphopyruvate dehydratase, is a key glycolytic enzyme involved in anaerobic metabolism under hypoxic conditions of cancer growth, and a cell surface plasminogen receptor for tumor invasion. Overexpression of ENO1 in breast, and lung cancers is associated with tumor progression and rapid tumor growth [42]. While our study did not specifically studying the prognostic value of ENO1, the observation of ENO1 in HER2-positive tumors supports the prognostic importance of this molecule. While it was seen in the more aggressive subtype of breast cancer, our study showed ENO1 elevation in HER2-positive tumors seemed to indicate a better tumor response to chemotherapy.

Vimentin is a member of the intermediate filament family. Along with microtubules and actin microfilaments, vimentin is an integral component of the cell cytoskeleton. In cancer, altered vimentin level is associated with a dedifferentiated phenotype, increased motility, invasiveness, and poor clinical prognosis [43, 44]. Vimentin overexpression was found in 90.5% of grade III breast carcinomas [45] which may explain its presence in both HER2-positive breast cancer and in TNBC. Our study found that vimentin, although an aggressive marker for breast cancer growth, is another indicator for a favorable tumor response to chemotherapy. L-plastin is an actin-binding protein involved in cancer cell migration, invasion, and metastasis, and its expression in breast cancer cell lines correlates with the degree of invasiveness [30, 46]. In this paper, L-plastin overexpression in HER2-positive breast cancer was associated with a likelihood of pCR.

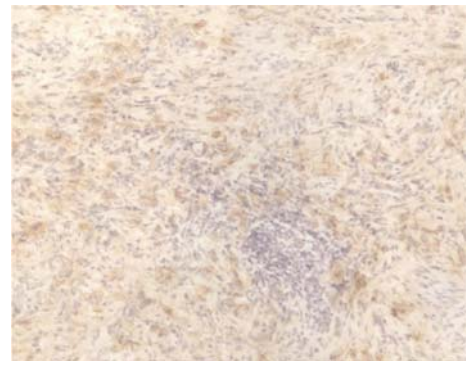
In contrast to those molecules associated with favorable tumor response to neoadjuvant therapy, high levels of Prx V in HER2-positive breast cancers were found to be associated with poor response to the same chemotherapy regimen. Peroxiredoxins (Prxs) represent a novel group of peroxidases



(a) Galectin-3-binding protein



(b)



(c)

Figure 4: G3BP expressions detected by LC-MS/MS and immunohistochemical (IHC) staining. Elevated G3BP expressions found in TNBC group by LC-MS/MS and confirmed by IHC in most of the frozen TNBC tumors. (a) Normalized spectrum count of G3BP detected from 39 breast cancer tissues. (b) Immunohistochemical staining of G3BP in a TNBC frozen tumor (power 200x). (c) Immunohistochemical staining of G3BP in a HER2+ frozen tumor (power 200x).

containing high antioxidant activity involved in cell differentiation and apoptosis [47], and Prx V is particularly effective in reducing reactive oxygen species (ROS). Moreover, Prx V is found in peroxisomes and mitochondria where protection against ROS is mostly needed. The antioxidant activity of Prx V may be associated with drug resistance of the tumor cells.

While some molecules are unique to the characteristics of individual subtype of breast cancer, Hsp70 overexpression was found by us to be associated with drug resistance in both HER2-positive and TNBC tumors. Heat shock proteins are overexpressed in a wide range of human cancers and are implicated in tumor cell proliferation, differentiation, death, invasion, metastasis, and immune recognition [48]. Consistent with the cellular functions of Hsp70, clinically it has been correlated with poor prognosis in breast, endometrial, cervical, and bladder cancers. Others have also reported that Hsp70 mediated drug resistance through its inhibitory effect on chemotherapy-induced tumor cell apoptosis [48–50].

In TNBC tumors, a list of different proteins was found to be overexpressed in tumors resistant to neoadjuvant chemotherapy. In addition to Hsp70, proteins such as periostin precursor (OSF-2), RhoA, actinin α4, cathepsin D preproprotein, and annexin 1 predicted a poor response of TNBC to treatment. Although all of them were known to have important cancer biological properties, they have not been linked to chemotherapy susceptibility until now.

Periostin was originally identified in a mouse osteoblastic cell line as an extracellular matrix adhesion protein for preosteoblasts. In addition to forming bones, teeth, and heart, periostin was recently found to be overexpressed in various types of human cancer. Periostin interacts with multiple cell-surface receptors (most notable integrins) and signals via the PI3-K/Akt and other pathways to promote cancer cell survival, epithelial-mesenchymal transition, invasion, and metastasis [51]. In breast cancer, periostin was found upregulated at both the mRNA and protein levels [51–55]. Activation of the Akt/PKB cellular survival pathway with consequential protection of tumor cells and endothelial cells from stress-induced cell death [51, 56] may contribute to the periostin-mediated drug resistance in cancer. To our knowledge, this is the first paper to link periostin to drug resistance in TNBC.

RhoA is a member of the Ras superfamily. It is involved in the regulation and timing of cell division. It is a small GTPase protein known to regulate the actin cytoskeleton in the formation of stress fibers. RhoA protein levels were significantly increased in breast cancer compared with the matched normal tissue. It has been reported by Fritz et al.

that an elevated RhoA protein level correlated with increasing breast tumor grade and poor prognosis [57].

Actinin $\alpha 4$ is another interesting protein that we found to indicate a poor tumor response to neoadjuvant therapy. It is thought that the actinin $\alpha 4$ cross-links actin filaments and connects the actin cytoskeleton to the cell membrane. The accumulation of actinin $\alpha 4$ in the cytoplasm is related to tumor invasiveness and metastasis, probably by enhancing cell motility, and was suggested to be a novel prognosticator in patients with ovarian and breast cancer [58].

Cathepsin D, an acid protease, is active in intracellular protein breakdown and is involved in the pathogenesis of several diseases. Its preproprotein secreted by cancer cells, acting as a mitogen on both cancer and stromal cells, stimulates both proinvasive and prometastatic properties of cancer cells. Many studies found that cathepsin D preproprotein/cathepsin D level represents an independent prognostic factor in a variety of cancers and is, therefore, considered to be a potential target for anticancer therapy [59]. Others have also shown that overexpression of cathepsin D in human breast cancers is associated with a higher risk of relapse and metastasis [59–61]. In our study, cathepsin D preproprotein appeared to be a drug-resistant marker in TNBC.

Although many proteins identified in this pilot study are interesting with promising potential, this study has several limitations. First, the tumors used in this study were collected from a clinical trial which provided many controlled clinical data; however, the sample size available for proteomic analysis was small. As a result, the findings derived from a small sample size always warrant a cautious interpretation. Second, the HER2-positive group consisted of tumors with different ER and PR status which might interfere with the conclusion. The potential false associations with HER2 might be solved by stratifying the HER2-positive tumors according to hormonal receptor status in a larger study. Lastly, the HER2-positive patients in this study were randomized to receive either chemotherapy alone or chemotherapy with Herceptin. The selected drug-resistant markers may represent the resistance not only to the chemotherapy but also to Herceptin.

In summary, our study has led to the identification of a list of important breast cancer proteins. The study also suggests that MS-based protein profiling may be an important tool in discovery of cancer biosignatures for tumor subtyping and prediction of treatment outcome. When sufficiently validated, some of these candidate protein markers could be used to improve breast cancer care. In addition, due to the heterogeneous and complex nature of the breast cancer tissue specimens, more refined methods need to be developed to maximize the protein identification to allow the capture of the best protein candidate markers for clinical use.

Abbreviations

DLDA: Diagonal linear discriminant analysis
G3BP: Galectin-3-binding protein
GEPAS: Gene expression pattern analysis suite
IR: Intermediate responders (>25% but ≤75% tumor regression)
KNN: K nearest neighbor
MS: Mass spectrometry
NR: Nonresponders, chemoresistant tumors (≤25% tumor regression)
PAM: Prediction analysis with microarrays
pCR: Pathological complete response, no residual cancer found at primary tumor site
R: Responders (>75% of tumor regression)
SOM: Self-organizing map
SVM: Support vector machines
TC ± H: Taxotere/Carboplatin/±Herceptin treatment
TNBC: Triple negative breast tumors
TRR: Tumor regression rate.

Conflict of Interests

The authors declare that they have no conflict of interests.

Funding

Grant support: The present study was supported by grant funds from NIH (NCI no. 1RO1 CA 093736-01A1), the Gonda Foundation, the Entertainment Industry Foundation (EIF)/Women's Cancer Research Fund, and the Friends of the Breast Program at UCLA.

Acknowledgments

The authors thank Jeffrey A. Gornbein (Department of Biomathematics, UCLA) for providing critical advice on statistical analysis.

References

[1] R. W. Carlson, E. Brown, H. J. Burstein et al., "NCCN task force report: adjuvant therapy for breast cancer," *Journal of the National Comprehensive Cancer Network*, vol. 4, supplement 1, pp. S1–S26, 2006.

[2] J. Baselga, L. Gianni, C. Geyer, E. A. Perez, A. Riva, and C. Jackisch, "Future options with trastuzumab for primary systemic and adjuvant therapy," *Seminars in Oncology*, vol. 31, no. 5, supplement 10, pp. 51–57, 2004.

[3] J. S. Ross, D. P. Schenkein, R. Pietrusko et al., "Targeted therapies for cancer 2004," *American Journal of Clinical Pathology*, vol. 122, no. 4, pp. 598–609, 2004.

[4] P. C. Fong, D. S. Boss, T. A. Yap et al., "Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers," *The New England Journal of Medicine*, vol. 361, no. 2, pp. 123–134, 2009.

[5] C. K. Anders, E. P. Winer, J. M. Ford et al., "Poly(ADP-ribose) polymerase inhibition: "targeted" therapy for triple-negative breast cancer," *Clinical Cancer Research*, vol. 16, no. 19, pp. 4702–4710, 2010.

[6] R. Radpour, Z. Barekati, C. Kohler, W. Holzgreve, and X. Y. Zhong, "New trends in molecular biomarker discovery for breast cancer," *Genetic Testing and Molecular Biomarkers*, vol. 13, no. 5, pp. 565–571, 2009.

[7] J. S. Reis-Filho and A. N. J. Tutt, "Triple negative tumours: a critical review," *Histopathology*, vol. 52, no. 1, pp. 108–118, 2008.

[8] C. M. Perou, T. Sørile, M. B. Eisen et al., "Molecular portraits of human breast tumours," *Nature*, vol. 406, no. 6797, pp. 747–752, 2000.

[9] T. Sørlie, C. M. Perou, R. Tibshirani et al., "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 19, pp. 10869–10874, 2001.

[10] T. Sørlie, R. Tibshirani, J. Parker et al., "Repeated observation of breast tumor subtypes in independent gene expression data sets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 14, pp. 8418–8423, 2003.

[11] H. R. Chang, "Trastuzumab-based neoadjuvant therapy in patients with HER2-positive breast cancer," *Cancer*, vol. 116, no. 12, pp. 2856–2867, 2010.

[12] V. Kulasingam and E. P. Diamandis, "Strategies for discovering novel cancer biomarkers through utilization of emerging technologies," *Nature Clinical Practice Oncology*, vol. 5, no. 10, pp. 588–599, 2008.

[13] N. L. Anderson and N. G. Anderson, "The human plasma proteome: history, character, and diagnostic prospects," *Molecular & Cellular Proteomics*, vol. 1, no. 11, pp. 845–867, 2002.

[14] H. Hondermarck, C. Tastet, I. E. Yazidi-Belkoura, R. A. Toillon, and X. L. Bourhis, "Proteomics of breast cancer: the quest for markers and therapeutic targets," *Journal of Proteome Research*, vol. 7, no. 4, pp. 1403–1411, 2008.

[15] C. Röwer, J. P. C. Vissers, C. Koy et al., "Towards a proteome signature for invasive ductal breast carcinoma derived from label-free nanoscale LC-MS protein expression profiling of tumorous and glandular tissue," *Analytical and Bioanalytical Chemistry*, vol. 395, no. 8, pp. 2443–2456, 2009.

[16] J. He, D. Shen, D. U. Chung et al., "Tumor proteomic profiling predicts the susceptibility of breast cancer to chemotherapy," *International Journal of Oncology*, vol. 35, no. 4, pp. 683–692, 2009.

[17] H. R. Chang, J. Glaspy, M. A. Allison et al., "Differential response of triple-negative breast cancer to a docetaxel and carboplatin-based neoadjuvant treatment," *Cancer*, vol. 116, no. 18, pp. 4227–4237, 2010.

[18] N. Rifai, M. A. Gillette, and S. A. Carr, "Protein biomarker discovery and validation: the long and uncertain path to clinical utility," *Nature Biotechnology*, vol. 24, no. 8, pp. 971–983, 2006.

[19] S. A. Whelan, M. Lu, J. He et al., "Mass spectrometry (LC-MS/MS) site-mapping of N-glycosylated membrane proteins for breast cancer biomarkers," *Journal of Proteome Research*, vol. 8, no. 8, pp. 4151–4160, 2009.

[20] A. Keller, A. I. Nesvizhskii, E. Kolker, and R. Aebersold, "Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search," *Analytical Chemistry*, vol. 74, no. 20, pp. 5383–5392, 2002.

[21] A. I. Nesvizhskii, A. Keller, E. Kolker, and R. Aebersold, "A statistical model for identifying proteins by tandem mass spectrometry," *Analytical Chemistry*, vol. 75, no. 17, pp. 4646–4658, 2003.

[22] T. A. Ulmer, V. Keeler, L. Loh et al., "Tumor-associated antigen 90K/Mac-2-binding protein: possible role in colon cancer," *Journal of Cellular Biochemistry*, vol. 98, no. 5, pp. 1351–1366, 2006.

[23] Y. P. Park, S. C. Choi, J. H. Kim et al., "Up-regulation of Mac-2 binding protein by hTERT in gastric cancer," *International Journal of Cancer*, vol. 120, no. 4, pp. 813–820, 2007.

[24] F. Mbeunkui, B. J. Metge, L. A. Shevde, and L. K. Pannell, "Identification of differentially secreted biomarkers using LC-MS/MS in isogenic cell lines representing a progression of breast cancer," *Journal of Proteome Research*, vol. 6, no. 8, pp. 2993–3002, 2007.

[25] Y. Wang, X. Ao, H. Vuong et al., "Membrane glycoproteins associated with breast tumor cell progression identified by a lectin affinity approach," *Journal of Proteome Research*, vol. 7, no. 10, pp. 4313–4325, 2008.

[26] Q. Zhao, X. Guo, G. B. Nash et al., "Circulating galectin-3 promotes metastasis by modifying MUC1 localization on cancer cell surface," *Cancer Research*, vol. 69, no. 17, pp. 6799–6806, 2009.

[27] C. Ginestier, M. H. Hur, E. Charafe-Jauffret et al., "ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome," *Cell Stem Cell*, vol. 1, no. 5, pp. 555–567, 2007.

[28] M. Lu, J. P. Whitelegge, S. A. Whelan et al., "Hydrophobic fractionation enhances novel protein detection by mass spectrometry in triple negative breast cancer," *Journal of Proteomics and Bioinformatics*, vol. 3, no. 2, pp. 1–10, 2010.

[29] J. P. Sullivan, M. Spinola, M. Dodge et al., "Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on notch signaling," *Cancer Research*, vol. 70, no. 23, pp. 9937–9948, 2010.

[30] D. M. Schulz, C. Böllner, G. Thomas et al., "Identification of differentially expressed proteins in triple-negative breast carcinomas using DIGE and mass spectrometry," *Journal of Proteome Research*, vol. 8, no. 7, pp. 3430–3438, 2009.

[31] M. Lu, S. A. Whelan, J. He et al., "Hydrophobic proteome analysis of triple negative and hormone-receptor- positive-her2-negative breast cancer by mass spectrometer," *Clinical Proteomics*, vol. 6, no. 3, pp. 93–103, 2010.

[32] R. R. Parikh, Q. Yang, S. A. Higgins, and B. G. Haffty, "Outcomes in young women with breast cancer of triple-negative phenotype: the prognostic significance of CK19 expression," *International Journal of Radiation Oncology Biology Physics*, vol. 70, no. 1, pp. 35–42, 2008.

[33] H. O. Habashy, D. G. Powe, C. M. Staka et al., "Transferrin receptor (CD71) is a marker of poor prognosis in breast cancer and can predict response to tamoxifen," *Breast Cancer Research and Treatment*, vol. 119, no. 2, pp. 283–293, 2010.

[34] C. Vyhlidal, X. Li, and S. Safe, "Estrogen regulation of transferrin gene expression in MCF-7 human breast cancer cells," *Journal of Molecular Endocrinology*, vol. 29, no. 3, pp. 305–317, 2002.

[35] T. Nakagawa, S. K. Huang, S. R. Martinez et al., "Proteomic profiling of primary breast cancer predicts axillary lymph node metastasis," *Cancer Research*, vol. 66, no. 24, pp. 11825–11830, 2006.

[36] A. L. Goldstein, "Thymosin $\beta 4$: a new molecular target for antitumor strategies," *Journal of the National Cancer Institute*, vol. 95, no. 22, pp. 1646–1647, 2003.

[37] H. J. Cha, M. J. Jeong, and H. K. Kleinman, "Role of thymosin $\beta 4$ in tumor metastasis and angiogenesis," *Journal of the National Cancer Institute*, vol. 95, no. 22, pp. 1674–1680, 2003.

[38] S. M. J. Rahman, A. L. Gonzalez, M. Li et al., "Lung cancer diagnosis from proteomic analysis of preinvasive lesions," *Cancer Research*, vol. 71, no. 8, pp. 3009–3017, 2011.

[39] B. J. Xu, A. L. Gonzalez, T. Kikuchi et al., "MALDI-MS derived prognostic protein markers for resected non-small cell lung cancer," *Proteomics—Clinical Applications*, vol. 2, no. 10-11, pp. 1508–1517, 2008.

[40] C. W. Sutton, N. Rustogi, C. Gurkan et al., "Quantitative proteomic profiling of matched normal and tumor breast tissues," *Journal of Proteome Research*, vol. 9, no. 8, pp. 3891–3902, 2010.

[41] S. Verghese-Nikolakaki, N. Apostolikas, E. Livaniou, D. S. Ithakissios, and G. P. Evangelatos, "Preliminary findings on the expression of thymosin beta-10 in human breast cancer," *British Journal of Cancer*, vol. 74, no. 9, pp. 1441–1444, 1996.

[42] A. Hennipman, B. A. van Oirschot, J. Smits, G. Rijksen, and G. E. J. Staal, "Glycolytic enzyme activities in breast cancer metastases," *Tumor Biology*, vol. 9, no. 5, pp. 241–248, 1988.

[43] M. I. Kokkinos, R. Wafai, M. K. Wong, D. F. Newgreen, E. W. Thompson, and M. Waltham, "Vimentin and epithelial-mesenchymal transition in human breast cancer—observations in vitro and in vivo," *Cells Tissues Organs*, vol. 185, no. 1–3, pp. 191–203, 2007.

[44] M. J. C. Hendrix, E. A. Seftor, R. E. B. Seftor, and K. T. Trevor, "Experimental co-expression of vimentin and keratin intermediate filaments in human breast cancer cells results in phenotypic interconversion and increased invasive behavior," *American Journal of Pathology*, vol. 150, no. 2, pp. 483–495, 1997.

[45] E. Korsching, J. Packeisen, C. Liedtke et al., "The origin of vimentin expression in invasive breast cancer: epithelial-mesenchymal transition, myoepithelial histogenesis or histogenesis from progenitor cells with bilinear differentiation potential?" *Journal of Pathology*, vol. 206, no. 4, pp. 451–457, 2005.

[46] C. S. Lin, Z. P. Chen, T. Park, K. Ghosh, and J. Leavitt, "Characterization of the human L-plastin gene promoter in normal and neoplastic cells," *Journal of Biological Chemistry*, vol. 268, no. 4, pp. 2793–2801, 1993.

[47] P. Karihtala, A. Mäntyniemi, S. W. Kang, V. L. Kinnula, and Y. Soini, "Peroxiredoxins in breast carcinoma," *Clinical Cancer Research*, vol. 9, no. 9, pp. 3418–3424, 2003.

[48] D. R. Ciocca and S. K. Calderwood, "Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications," *Cell Stress and Chaperones*, vol. 10, no. 2, pp. 86–103, 2005.

[49] M. Jäättelä, "Escaping cell death: survival proteins in cancer," *Experimental Cell Research*, vol. 248, no. 1, pp. 30–43, 1999.

[50] D. D. Mosser, A. W. Caron, L. Bourget et al., "The chaperone function of hsp70 is required for protection against stress-induced apoptosis," *Molecular and Cellular Biology*, vol. 20, no. 19, pp. 7146–7159, 2000.

[51] K. Ruan, S. Bao, and G. Ouyang, "The multifaceted role of periostin in tumorigenesis," *Cellular and Molecular Life Sciences*, vol. 66, no. 14, pp. 2219–2230, 2009.

[52] R. Shao, S. Bao, X. Bai et al., "Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression," *Molecular and Cellular Biology*, vol. 24, no. 9, pp. 3992–4003, 2004.

[53] X. Liang, J. Zhao, M. Hajivandi et al., "Quantification of membrane and membrane-bound proteins in normal and malignant breast cancer cells isolated from the same patient with primary breast carcinoma," *Journal of Proteome Research*, vol. 5, no. 10, pp. 2632–2641, 2006.

[54] X. Liang, J. Huuskonen, M. Hajivandi et al., "Identification and quantification of proteins differentially secreted by a pair of normal and malignant breast-cancer cell lines," *Proteomics*, vol. 9, no. 1, pp. 182–193, 2009.

[55] Y. Zhang, G. Zhang, J. Li, Q. Tao, and W. Tang, "The expression analysis of periostin in human breast cancer," *Journal of Surgical Research*, vol. 160, no. 1, pp. 102–106, 2010.

[56] L. Gillan, D. Matei, D. A. Fishman, C. S. Gerbin, B. Y. Karlan, and D. D. Chang, "Periostin secreted by epithelial ovarian carcinoma is a ligand for $\alpha V\beta 3$ and $\alpha V\beta 5$ integrins and promotes cell motility," *Cancer Research*, vol. 62, no. 18, pp. 5358–5364, 2002.

[57] G. Fritz, C. Brachetti, F. Bahlmann, M. Schmidt, and B. Kaina, "Rho GTPases in human breast tumours: expression and mutation analyses and correlation with clinical parameters," *British Journal of Cancer*, vol. 87, no. 6, pp. 635–644, 2002.

[58] K. Honda, T. Yamada, R. Endo et al., "Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion," *Journal of Cell Biology*, vol. 140, no. 6, pp. 1383–1393, 1998.

[59] P. Benes, V. Vetvicka, and M. Fusek, "Cathepsin D-many functions of one aspartic protease," *Critical Reviews in Oncology/Hematology*, vol. 68, no. 1, pp. 12–28, 2008.

[60] S. M. Thorpe, H. Rochefort, M. Garcia et al., "Association between high concentrations of M(r) 52,000 cathepsin D and poor prognosis in primary human breast cancer," *Cancer Research*, vol. 49, no. 21, pp. 6008–6014, 1989.

[61] F. Spyratos, T. Maudelonde, J. P. Brouillet et al., "Cathepsin D: an independent prognostic factor for metastasis breast cancer," *The Lancet*, vol. 2, no. 8672, pp. 1115–1118, 1989.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 181890, 8 pages
doi:10.1155/2011/181890

Review Article

Proteomics in Melanoma Biomarker Discovery: Great Potential, Many Obstacles

Michael S. Sabel, Yashu Liu, and David M. Lubman

Department of Surgery, University of Michigan Health Systems, 1500 East Medical Center Drive, Ann Arbor, MI 48109, USA

Correspondence should be addressed to Michael S. Sabel, msabel@med.umich.edu

Received 12 July 2011; Accepted 2 August 2011

Academic Editor: David E. Misek

Copyright © 2011 Michael S. Sabel et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The present clinical staging of melanoma stratifies patients into heterogeneous groups, resulting in the application of aggressive therapies to large populations, diluting impact and increasing toxicity. To move to a new era of therapeutic decisions based on highly specific tumor profiling, the discovery and validation of new prognostic and predictive biomarkers in melanoma is critical. Genomic profiling, which is showing promise in other solid tumors, requires fresh tissue from a large number of primary tumors, and thus faces a unique challenge in melanoma. For this and other reasons, proteomics appears to be an ideal choice for the discovery of new melanoma biomarkers. Several approaches to proteomics have been utilized in the search for clinically relevant biomarkers, but to date the results have been relatively limited. This article will review the present work using both tissue and serum proteomics in the search for melanoma biomarkers, highlighting both the relative advantages and disadvantages of each approach. In addition, we review several of the major obstacles that need to be overcome in order to advance the field.

1. Introduction

The field of oncology is rapidly attempting to move to a new era of personalized therapy, where individualized therapeutic decisions are based on highly specific tumor profiling. For this to become a reality, the discovery and validation of new prognostic and predictive biomarkers are necessary. This goal seems increasingly realistic thanks to high-throughput screening methods, particularly genomic profiling. The large-scale analysis of gene expression has improved our knowledge of tumorigenesis, invasion, and metastasis. More recently, gene expression assays have helped guide therapeutic decisions, such as the use of multigene assays for decisions regarding systemic therapy in breast cancer [1, 2]. However, despite continued success in the preclinical setting, clinical translation has been slow and several tumor types present unique challenges to the use of genomic profiling.

One example is melanoma. The present staging system for melanoma, using Breslow thickness, ulceration, mitotic rate, and the presence of regional and distant metastases, stratifies patients into heterogeneous groups, with wide variability in outcome or response to therapy. Clinically, this results in applying more aggressive surgical and adjuvant therapies to large populations, diluting the impact of therapy while exposing more patients to toxicity. For those treating melanoma, better prognostic and predictive markers in melanoma are sorely needed, but to date have been elusive. The high-throughput analysis of genomic data requires fresh tissue from a large number of primary tumors. This presents a unique challenge in melanoma where the primary is often only a few millimeters in size, with no residual tissue after the diagnosis has been made. For this reason, proteomics appears to be an ideal choice for the discovery of new prognostic and predictive biomarkers in melanoma.

There are other potential advantages to proteomics over genomics. Genes are transcribed into mRNA, but because cells can use alternative splicing, there is no one-to-one relationship between the genome and the transcript. The transcripts are further translated into proteins, which often undergo posttranslational modifications (PTMs), or can be aberrant in cancer cells. Therefore, one gene can result in several different protein isoforms. Protein structure can also be influenced by environmental factors, including interaction

with other proteins, degradation, or compartmentalization of proteins within protein complexes. As the structure and availability of the final versions of the proteins ultimately determine the behavior of the cell, high-throughput screening methods for changes in protein expression may be better suited to identify biomarkers with prognostic or predictive value.

2. Biomarkers in Melanoma

The clinical potential for melanoma biomarkers covers the spectrum of disease. Protein changes associated with the transition from melanocyte to atypia or dysplasia and ultimately to melanoma could be used to aid in diagnosis or to screen high-risk patients. Proteins associated with pathophysiology and malignant properties could be used to further classify melanoma, stratifying patients by risk of recurrence in order to better select surgical and adjuvant treatments. Likewise, protein expression (baseline or changes in expression) may predict response to specific therapies, so that selection of systemic therapies can be tailored to the individual patient. The detection of low levels of melanoma-associated proteins in the serum may also lead to early recognition of recurrent disease or monitoring the response to therapy for metastatic disease.

Despite the significant potential for tissue-based or serological biomarkers to help diagnose early stage disease, tailor therapy, or detect recurrence, very few biomarkers are in clinical use. Several tissue markers are used to help distinguish melanoma from other types of cancers, including S100, MART-1, and gp100/HMB45. To date, however, there are no tissue-based biomarkers that are utilized clinically for prognostic classification. This is despite the identification of multiple biomarkers, through genomic or immunohistochemical analysis, whose abnormal expression has been linked to poor outcome. These include tumor suppressors/oncogenes/signal transducers (p16, PTEN, EGFR, c-KIT, c-myc, bcl-6, HER3), cell-cycle associated proteins (Ki67, Cyclins A, B, D, E, p21, Geminin, PCNA), regulators of apoptosis (bcl-2, bax, Bak, ING3, ING4), proteins involved with cell adhesion and motility (P, E and N-cadherin, β-catenin, β1 and β3 integrins, matrix metalloproteinases (MMPs)), and others (Hsp90, RGS1, NCOA3, MCM4, MCM6) [3, 4].

There have also been several serologic markers that have been associated with poor prognosis. These include differentiation antigens (S100β, MIA, tyrosinase), proangiogenic factors (vascular endothelial growth factor (VEGF), basic fibroblast growth factor (BFGF), IL-8), cell adhesion and motility molecules (soluble intracellular adhesion molecule 1 (sICAM-1), soluble vascular cell adhesion molecule 1 (sVCAM), MMP-1, MMP-9), cytokines (IL-6, IL-10, soluble IL-2 receptor (sIL-2R)), and others (TA90 immune complex, YKL-40) [3–5]. However, only two of these are utilized clinically.

The strongest prognostic serum biomarker is lactate dehydrogenase (LDH), which correlates with tumor load in advanced disease and is the strongest independent prognostic factor in stage IV melanoma [6]. It is the only biomarker included in the AJCC staging system [7] and is often used to stratify patients for randomized trials in stage IV disease. There is limited benefit, however, to the measurement of LDH among patients with earlier stage disease, particularly in the followup of patients who appear tumor-free after surgical resection.

The S100 protein consists of two subunits, alpha and beta. The beta subunit is expressed in cells of melanocytic lineage and is often used as an immunohistochemical marker for histological diagnosis of melanoma. Serum S100β has also been studied as a prognostic biomarker. Several studies have demonstrated an association between serum S100β levels and outcome, independent of stage. S100β serum concentrations can be a useful marker for monitoring therapy response in patients with advanced disease. While some European guidelines recommend routine S100β measurements as part of the surveillance of melanoma patients, [8, 9], there is limited evidence that this impacts outcome [10, 11].

3. Tissue Proteomics in Melanoma

Proteomic approaches can be divided into two categories: those that characterize the entire protein complement of the cells or tissue of interest and those that examine only those proteins found in specific specimens (primarily blood, but this could also include other fluids such as saliva or urine). Several groups have utilized functional proteomics to identify alterations in protein expression and posttranslational modifications to identify markers of melanoma progression as well as predictive markers, such as identifying proteins that may be associated with response to therapy (Table 1). While this approach has been utilized across a spectrum of primary tumors, it is more difficult in melanoma secondary to the limited accessibility of primary melanoma tissues. Therefore most of this work has been carried out in melanoma cell lines.

Two-dimensional electrophoresis (2DE) has been the mainstay tool for separating proteins for many years. Proteins in a 2-dimensional gel are separated in the first dimension based on isoelectric points and then in a second dimension based on molecular masses. Differences between the samples can be compared and relative quantities determined by quantifying the ratios of spot intensities in the 2D gels. Matrix-assisted desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) can then be used to analyze small amounts of protein isolated from the gel. The mass information obtained can then be used for protein identification using an appropriate protein sequence database and search program.

Bernard et al. [12] used 2DE and mass spectrometry to identify proteins that differentiated melanocytes from melanoma cell lines and therefore may be important in the early progression to melanoma. Two proteins, nuclophosmin/B23 and hepatoma-derived growth factor (HDGF), were strongly upregulated in melanoma, while cathepsin D was downregulated in melanoma cell lines. Carta et al. [13] also used 2DE and mass spectrometry to examine the

Table 1: Proteomics for biomarker discovery in melanoma.

Author	Year	Specimens	Comparison	Methodology	Proteins of interest
Bernard et al. [12]	2003	Cell lines	Melanocyte versus melanoma (primary and metastatic)	2DE and mass spectrometry	Nucleophosmin/B23, HDGF, CTSD
Sinha et al. [14]	2003	Cell lines	Responsive to chemotherapy versus nonresponsive	2DE and MALDI-TOF	Multiple (25)
Wilson et al. [28]	2004	Serum	Stage I or II melanoma patients, recurrence versus none	SELDI-TOF	n/a (expression profiles)
Carta et al. [13]	2005	Cell lines	Primary versus metastatic melanoma	2DE and mass spectrometry	HSP27, HSP60, HSPA8, PRDX2
Mian et al. [29, 30]	2005	Serum	Stage I melanoma versus Stage IV melanoma	SELDI-TOF	n/a (expression profiles), further work identified SAA
Mian et al. [29, 30]	2005	Serum	Stage III melanoma patients, recurrence versus none	SELDI-TOF	
Takikawa et al. [24]	2009	Serum	Volunteers versus melanoma patients	Nano LC and MALDI-TOF	PPBP
Caron et al. [27]	2009	Serum	Volunteers versus melanoma patients	SELDI-TOF	n/a (expression profiles)
Greco et al. [31]	2009	Serum	Patients undergoing biopsy, melanoma versus not	2DE and MALDI-TOF	TTR, AGT, DBP
Paulitschke et al. [46]	2009	Secreted proteins from and cell lines and skin samples	Normal skin versus Melanoma	Nano LC and mass spectrometry	GPX5, periostin, stanniocalcini-1
Suzuki et al. [36]	2010	Serum	Volunteers versus melanoma patients	Autoantibody detection	EEF2, ENO1, ALDOA, GAPDH, HNRNP-A2B1
Hood et al. [47]	2010	Cell lines	Melanoma versus normal skin	Nano LC and mass spectrometry	Tenascin-C, fibronectin, ACN4, TSP-1
Liu et al. [37]	2010	Serum	Node negative versus node positive melanoma patients	Autoantibody detection	GRP94, ASAH1, CTSD, LDHB
Hardesty et al. [16]	2011	Surgically resected lymph nodes	Recurrence versus no recurrence	MALDI-IMS	Cytochrome C calmodulin

HDGF: hepatoma-derived growth factor, CTSD: cathespin D, PPBP: prop-platelet basic protein precursor, SAA: serum amyloid A, TTR: transthyretin, AGT: angiotensinogen, DBP: vitamin D binding protein, EEF2: eukaryotic elongation factor 2, ENO1: enolase 1, ALDOA: aldolase A, GAPDH: glyceraldehyde-3-phosphate dehydrogenase, HNRNP-A2B1: heterogeneous nuclear ribonucleoprotein A2B1, ACN4: alpha-actinin-4, TSP-1: thrombospondin-1, GRP94: 94 kD glucose-regulated protein, ASAH1: acid ceramidase, LDHB: lactate dehydrogenase B.

proteomes of cultured melanocytes and melanoma cell lines from primary and metastatic lesions. They identified several candidate proteins, many of which were stress proteins. They then used RT-PCR to evaluate mRNA expression of these proteins and found that overexpression of HSP27, HSP60, and HSPA8 and downregulation of PRDX2 were observed more commonly in metastatic melanoma versus primary melanoma.

As opposed to identifying proteins associated with melanoma development and progression, Sinha et al. [14] compared melanoma cell lines with varying degrees of resistance to commonly used anticancer drugs to identify proteins that may be responsible for resistance to therapy. Starting with a single melanoma cell line, they created a panel of sublines that exhibited different levels of drug sensitivity [15]. Using 2DE and matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry for protein identification, they identified a variety of proteins that were differentially expressed in chemoresistant melanoma cell lines, many of which were chaperones, including heat-shock proteins (HSPs).

These studies utilized electrophoresis to physically separate the proteins, requiring the use of melanoma cell lines. Hardesty et al. [16] used MALDI-imaging mass spectrometry (MALDI-IMS) analysis, which acquires information from intact proteins directly from thin sections of the tissue [17]. This allows for the analysis of specific cellular regions and precludes the need to generate a cell line from which

the proteins are isolated. Using lymph nodes involved with metastatic melanoma from 69 stage III patients, they identified two proteins that were associated with recurrence, cytochrome C and calmodulin, with a better prognosis as the intensity of both proteins increase.

While this approach is applicable to patients with larger metastatic deposits, such as clinically involved lymph nodes, it is not useful for discovery of primary tumor biomarkers. For approximately 90% of melanoma patients, the entire primary tumor is excised during the initial biopsy, so proteomic investigations requiring fresh or frozen primary tumor tissue are not applicable. Many academic centers have collections of formalin-fixed paraffin-embedded (FFPE) melanoma specimens, with accompanying clinicopathologic and outcome data. If these specimens could be used to scrutinize the proteome, the use of proteomics for melanoma biomarker discovery would take a giant leap forward. Unfortunately, FFPE tissues are typically refractory to proteomic investigations using today's methodologies, largely due to the high level of covalently linked proteins arising from formalin fixation [18]. Shotgun proteomics involves direct digestion of protein mixtures to complex peptide mixtures which are then separated and analyzed. These approaches can be used to extract proteins or peptides from fixed tissues for analysis. Several proteomics studies using FFPE tissue have been reported [19–22]. A newer approach, a modified shotgun proteomic strategy, termed direct tissue proteomics (DTP), can extract and identify peptides and proteins directly from tissues using micro-reverse-phase (μRP) LC-MS/MS and has been proposed for use in melanoma [18, 23]. While there are still several obstacles to overcome, DTP with efficient extraction of proteins from FFPE tissue could open up a new avenue of retrospective proteomics-based biomarker investigation in melanoma.

4. Serum Proteomics

As discussed, examining the entire proteome of the cell of interest requires either the generation of a cell line or adequate harvestable tissue, which inherently limits and biases the study population. This presents several obstacles to both validating the findings and ultimately using them clinically. Blood carries not only plasma-specific proteins but also multiple proteins derived from other tissues, including tumors. Many proteins are secreted, shed, or lost into the circulation, either directly by tumor cells or indirectly after destruction of the tumor cells. The development, validation, and use of serum tests hold several potential advantages over tests that require primary tumor tissue, particularly for melanoma. As an example, Takikawa et al. [24] compared the serum proteome between healthy volunteers and melanoma patients using NanoLC and MALDI-TOF-MS, and identified 9 proteins detectable in plasma from the melanoma patients but not healthy plasma. Ultimately they identified pro-platelet basic protein precursor (PPBP) as a protein whose level corresponded with outcome and may serve as a serological prognostic biomarker.

Several investigators have utilized surface-enhanced laser desorption/ionization (SELDI) and protein chip technology to find serum protein patterns that may be associated with the presence or the stage of melanoma. While gene chips have allowed for the detection of thousands of genes from very small samples, the creation of protein chips has faced several obstacles. Compared to DNA, proteins are not as robust and tend to denature. Proteins are more difficult to attach to chip surfaces, and while PCR can be used to amplify DNA, there is no method of amplifying minute amounts of protein. The ProteinChip Biology System uses SELDI-TOF MS to retain proteins on a solid-phase chromatographic surface that are subsequently ionized and detected by TOF MS [25, 26]. The surface of the ProteinChip is designed to select proteins from extracts due to either chemical (anionic, cationic, hydrophobic, hydrophilic) or biochemical (antibody, receptor, DNA, enzyme) properties. This is a more user-friendly approach to proteomics; SELDI has several advantages over other technologies for high-throughput screening as it is rapid, of relatively low cost, and reproducible. It requires smaller amounts of sample than 2DE as protein profiles can be made from fewer cells. It is also readily adaptable to a diagnostic format.

This technology can be used to detect protein expression patterns and then compare these patterns between different population sets. These patterns could be used as a clinical diagnostic test. Caron et al. [27] used this technology to try and discriminate between serum samples from melanoma patients and healthy volunteers and demonstrated a good diagnostic accuracy of 98.1% (sensitivity 96.7%, specificity 100%). Wilson et al. [28] examined the serum from patients with AJCC stage I or II melanoma who recurred ($n = 25$) or did not ($n = 24$) using SELDI ProteinChip mass spectrometry (MS) and identified three protein expression pattern differences that could discriminate between the two. Mian et al. demonstrated the potential of this technology with artificial neural networks (ANNs) to discriminate between serum samples from 101 stage I melanoma patients and 104 stage IV melanoma patients, as well as from 28 stage III patients who recurred from 27 stage III patients who did not [29]. Further research, focusing on the highest peak, ultimately identified serum amyloid A (SAA) as a potential prognostic biomarker in melanoma [30].

There are some drawbacks to serum-based proteomics. The expression and release of proteins into the serum can be variable. As an example of this, Greco et al. [31] obtained serum from 50 patients undergoing biopsy for suspected melanoma. Using 2DE and MALDI-TOF-MS, they identified increases of transthyretin (TTR) and angiotensinogen (AGT) and decreased expression of vitamin D binding protein (DBP). The investigators also examined serum samples after surgical removal of the melanoma and found that these were no longer elevated 1 month after surgery. This is an important consideration in melanoma patients, particularly if one is seeking to develop a prognostic serum test. This test would most likely be ordered after the diagnosis of melanoma and therefore might be greatly impacted by whether or not the entire tumor was removed with the diagnostic biopsy.

Another drawback to serum proteomics is that most (97%) of the proteins found in plasma belong to one of 7

major groups of high-abundance plasma proteins: albumin, immunoglobulins, fibrinogen, alpha-1 antitrypsin, alpha-2 macroglobulin, transferring, and lipoproteins. As these are primarily proinflammatory proteins, they are unlikely to represent prognostic or predictive biomarkers. SELDI pattern recognition studies do not depend on detecting low-abundance proteins, but depend instead on fluctuations of protein expression patterns. Therefore this technology, as well as others, is limited in specifically identifying low-abundance proteins that may be useful biomarkers.

5. Antibody-Based Proteomics

An alternate approach to examining the proteome directly is to screen for antibodies in the serum of patients with melanoma. As with serum-based proteomics, primary tumor tissue from each patient is not required, a significant advantage in melanoma patients. Another advantage is that antibodies are more sensitive and stable than proteins, which is clinically important in creating clinically useful diagnostic or prognostic tests. Antibody-based proteomics is particularly well suited to melanoma as the presence of an immune response to melanoma-associated antigens has been well documented [32–35]. Monitoring the presence or absence of antibodies in the serum that recognize specific tumor antigens can provide insights into the propensity of melanoma to metastasize, serving as biomarkers of melanoma biology and perhaps identifying ideal targets for therapeutic intervention.

There are several high-throughput methods for the discovery of autoantibodies including serological screening of cDNA expression libraries (SEREX), phage display libraries and proteomics-based techniques. Serological protein analysis (SERPA) uses 2DE to separate proteins from tumor tissues or cell lines which are then transferred onto membranes by electroblotting and subsequently probed with sera from different populations of interest. Suzuki et al. [36] used this approach to identify 5 proteins that differentiated serum samples from melanoma patients and healthy volunteers (eukaryotic elongation factor 2 (EEF2), enolase 1 (ENO1), aldolase A (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and heterogeneous nuclear ribonucleoprotein (HNRNP-A2B1).

Another proteomics-based approach to detecting serum autoantibodies is the use of protein microarrays. Protein microarrays spot proteins (purified, recombinant, or extracted from tumor cell lysates) onto microarrays which can be two-dimensional (glass slides, nitrocellulose membranes and microtitre plates) or three-dimensional (beads, nanoparticles). These are then incubated with sera from different populations. Advantages include the fact that less sample and reagents are needed, and autoantibodies to proteins with posttranslational modifications, such as glycosylated proteins, can be detected. There are, however, challenges to identifying the specific immune-reactive proteins in the respective protein fractions. Rather than screening the entire proteome, Liu et al. [37] focused on the subproteome of glycoproteins. Dual-lectin (ConA and WGA) affinity chromatography was applied to extract both glycoproteins from the lysate of a cell line generated from an intra-abdominal melanoma metastasis. Liquid-based reverse phase separation and natural protein microarray were then applied to separate the enriched proteins and spot the separated fractions on nitrocellulose slides. These were used to probe the sera from patients with newly diagnosed melanoma for antibodies that correlated with the presence of regional metastases. After validation, antibodies to 4 proteins including 94 kD glucose-regulated protein (GRP94), acid ceramidase (ASAH1), cathepsin D (CTSD), and lactate dehydrogenase B (LDHB) were identified that differentiated node-negative from node-positive patients.

6. Clinical Translation and Obstacles

Although the clinical potential of both tissue and serum-based proteomics in melanoma biomarker discovery seem strong, progress has been relatively limited to date. The reasons for this are multifactorial, both related to the inherent obstacles specific to melanoma and problems with study design. For the field to advance, several issues need to be tackled.

A key obstacle to biomarker discovery is reproducibility. This has been a significant problem with proteomics. As an example, in a meta-analysis of prostate cancer proteomic data obtained with SELDI-TOF, published results seem to differ greatly between different groups, and even from within the same groups [38, 39]. As evidenced in Table 1, there is little overlap between the proteins identified by different groups using different techniques. There are multiple approaches to proteomics, each with significant advantages and disadvantages depending on what the question is. Compared with serum proteomics, tissue proteomics has a higher likelihood of identifying marker candidates based on the higher concentration of protein within the tissue than after dilution in peripheral blood. For melanoma, the use of cell lines is most feasible however results are immediately biased by (1) selecting patients with harvestable tumor, (2) selecting melanomas that grow well *in vitro,* and (3) by changes in protein expression induced by *in vitro* culturing. Variations in cell culture technique could yield different results from the same cell line. Using tissue obtained directly from patients avoids some of these issues, but limits the patient population and hence the questions that can be asked. In addition, the results obtained can be greatly impacted by stroma, necrotic tissue, serum proteins, and blood cells within the specimen. Pure cancer cell populations can be created using fine needle aspiration, calcium starvation, immunomagnetic separation, or laser capture microdissection (LCM), [40–43] but this is extremely difficult to do when there is a limited supply of primary tumor tissue. Until newer technologies for examining FFPE tissues are more fully developed, this approach will be limited to select patient populations. Serum proteomics may identify fewer candidate proteins, but given the drawbacks of tissue proteomics in melanoma, this approach may be more translatable to clinical use. Examining the serum proteome is impossible without reducing the complexity of the protein mixture by removing highly abundant serum proteins, for

which there are also several methods. When the blood is drawn, how it is stored, and whether serum or plasma is used can potentially impact the results. Therefore one can see that before any analysis is even performed, variations in the methods used to prepare the samples (cell culturing, tissue procurement, extraction of highly abundant serum proteins) can significantly impact the results. The methods used to do this must be highly reproducible, as even small variations in buffers or agents can alter the results. It becomes easy to see how radically different results can be reached even when the same experimental technique is utilized [44].

Quite often, initial experiments are designed based on the available samples and technologies, without as much forethought into the clinical question that the findings hope to address. In melanoma, for example, tissue proteomics are often carried out on samples from resected lymph nodes and metastatic deposits, as there is ample tissue available. Clinically, however, this is a population with a very poor prognosis, and outside of very specific questions, it is less likely that biomarkers identified from these highly dedifferentiated samples would impact clinical decision making or serve as useful biomarkers in patients with early stage disease. It is imperative that translational oncology be hypothesis driven, whereby even discovery studies are designed with a specific clinical question in mind. Even something as simple as when blood samples are obtained, if they do not reflect when they would be drawn in the clinical setting, could negatively impact the results. Close collaboration between clinical experts in melanoma and basic scientists is imperative to results that have a high likelihood of clinical impact.

Patient selection is also critical to results that are both reproducible and clinically relevant. In oncology, it is quite common to discover a biomarker among a highly heterogenous group of patients that on univariate analysis is significantly associated with outcome, but in reality correlates so strongly with known staging factors (tumor size, grade) that on multivariate analysis it provides no independent prognostic information. While these may be of scientific interest in unraveling the genes/proteins associated with dedifferentiation and metastases, they are of limited clinical benefit in stratifying patients beyond our current staging systems. It is therefore imperative that samples be obtained from as homogenous a population as is feasible and in adequate numbers so that newly discovered biomarkers are analyzed in the context of the known prognostic and predictive factors used in clinical decision making. Often smaller and more heterogeneous sample sets are chosen for practical reasons, which unfortunately contribute to the large number of reported biomarkers that are never validated or demonstrate clinical utility.

Equally important is the appropriate selection of controls. For prognostic markers meant to differentiate between melanoma patients with different outcomes, it is imperative that the "good" cohort have adequate followup to be sure that this is truly a group at low likelihood of recurrence. Frequently these controls only have a median followup of 2-3 years. With longer followup, several of the "good" players may recur, and with small sample sets the conversion of only 1 or 2 patients from "good" to "bad" can dramatically change the results. Likewise, the search for biomarkers associated with melanoma development, with potential use as a screening tool, often compares melanoma patients (of varying stages) to healthy volunteers. Neither of these populations is appropriate—the melanoma group should only include patients with early stage, recently diagnosed disease, and the control group should not be healthy volunteers but rather high-risk individuals with similar characteristics regarding age and sun exposure.

While the field of proteomics and other "-omics" fields are replete with articles describing discovery, there are dramatically fewer articles validating previously published results. Several reports include internal validations, using a fraction of their samples for discovery and then validating the results in the complete set. However, the bioinformatics tools used in discovery sets often seek to overfit the data, erring on the side of not missing a potential biomarker, but resulting in sensitivities and specificities that may not be reproducible. Even if an institution is able to validate their own findings, given the impact that sample preparation and patient selection can have, validation from other institutions is absolutely critical. Beyond interlaboratory variations, many biomarkers demonstrate varying expression based on patient characteristics (age, race and ethnicity, genetic lineage, environmental exposures). Therefore, biomarkers discovered and validated on a population in one geographic area may not be validated on another, even though the techniques are the same and the populations appear matched by known prognostic factors (Breslow thickness, nodal status, etc.). Unfortunately, there is limited enthusiasm for one institution to attempt and validate a published result. Validation requires larger numbers and can be costly and labor intensive. As many candidate biomarkers will not be validated, few researchers are interested in devoting time and money to a project that will likely not result in a publication, as there is limited interest on the part of prominent journals in publishing negative results. This leads to significant publication bias, which all of us as editors and reviewers are in part responsible for.

As new proteomic-based biomarkers are discovered, it is increasingly important that a mechanism exists by which the most promising biomarkers can be validated using external samples. This will require a collaborative effort on the part of the leading melanoma research centers. As an example of this, the National Cancer Institute has created the Early Detection Research, which hopes to promote biomarker discovery, validation, and translation into clinical practice for biomarkers associated with screening and risk [45]. Similar disease-specific collaborative efforts, centered on prospectively collecting data, blood and tumor tissue from multiple centers *not* for biomarker discovery but rather for validation, will be necessary to validate prognostic and predictive biomarkers.

References

[1] S. Paik, S. Shak, G. Tang et al., "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer,"

The New England Journal of Medicine, vol. 351, no. 27, pp. 2817–2826, 2004.

[2] K. S. Albain, W. Barlow, S. Shak et al., "Prognostic and predictive value of the 21-gene recurrence score assay in post-menopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial," *The Lancet Oncology*, vol. 11, no. 1, pp. 55–65, 2010.

[3] J. Utikal, D. Schadendorf, and S. Ugurel, "Serologic and immunohistochemical prognostic biomarkers of cutaneous malignancies," *Archives of Dermatological Research*, vol. 298, no. 10, pp. 469–477, 2007.

[4] H. Gogas, A. M. M. Eggermont, A. Hauschild et al., "Biomarkers in melanoma," *Annals of Oncology*, vol. 20, pp. vi8–vi13, 2009.

[5] S. Liu, P. Kirschmeier, J. Simon, C. Seidel-Dugan, and M. Puhlmann, "Prognostic and predictive molecular markers in cutaneous malignant melanoma: the first step toward personalized medicine," *Current Pharmacogenomics and Personalized Medicine*, vol. 6, no. 4, pp. 272–294, 2008.

[6] M. Deichmann, A. Benner, M. Bock et al., "S100-beta, melanoma-inhibiting activity, and lactate dehydrogenase discriminate progressive from nonprogressive American joint committee on cancer stage IV melanoma," *Journal of Clinical Oncology*, vol. 17, no. 6, pp. 1891–1896, 1999.

[7] C. M. Balch, J. E. Gershenwald, S. Seng-Jaw et al., "Final version of 2009 AJCC melanoma staging and classification," *Journal of Clinical Oncology*, vol. 27, no. 36, pp. 6199–6206, 2009.

[8] R. Dummer, R. Panizzon, P. H. Bloch, and G. Burg, "Updated Swill guidelines for the treatment and follow-up of cutaneous melanoma," *Dermatology*, vol. 210, no. 1, pp. 39–44, 2005.

[9] C. Garbe, D. Schadendorf, W. Stolz et al., "Short German guidelines: malignant melanoma," *Journal der Deutschen Dermatologischen Gesellschaft*, vol. 6, no. 1, pp. S9–S15, 2008.

[10] R. Dummer, A. Hauschild, M. M. Guggenheim, L. Jost, and G. Pentheroudakis, "Melanoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up," *Annals of Oncology*, vol. 21, no. 5, pp. v194–v197, 2010.

[11] S. Ugurel, J. Utikal, and J. C. Becker, "Tumor biomarkers in melanoma," *Cancer Control*, vol. 16, no. 3, pp. 219–224, 2009.

[12] K. Bernard, E. Litman, J. L. Fitzpatrick et al., "Functional proteomic analysis of melanoma progression," *Cancer Research*, vol. 63, no. 20, pp. 6716–6725, 2003.

[13] F. Carta, P. P. Demuro, C. Zanini et al., "Analysis of candidate genes through a proteomics-based approach in primary cell lines from malignant melanomas and their metastases," *Melanoma Research*, vol. 15, no. 4, pp. 235–244, 2005.

[14] P. Sinha, J. Poland, S. Kohl et al., "Study of the development of chemoresistance in melanoma cell lines using proteome analysis," *Electrophoresis*, vol. 24, no. 14, pp. 2386–2404, 2003.

[15] M. A. Kern, H. Helmbach, M. Artuc, D. Karmann, K. Jurgovsky, and D. Schadendorf, "Human melanoma cell lines selected in vitro displaying various levels of drug resistance against cisplatin, fotemustine, vindesine or etoposide: modulation of proto-oncogene expression," *Anticancer Research*, vol. 17, pp. 4359–4370, 1997.

[16] W. M. Hardesty, M. C. Kelley, M. Deming, R. L. Low, and R. M. Caprioli, "Protein signatures for survival and recurrence in metastatic melanoma," *Journal of Proteomics*, vol. 74, no. 7, pp. 1002–1014, 2011.

[17] M. Stoeckli, P. Chaurand, D. E. Hallahan, and R. M. Caprioli, "Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues," *Nature Medicine*, vol. 7, no. 4, pp. 493–496, 2001.

[18] K. Rezaul, L. L. Wilson, and D. K. Han, "Direct tissue proteomics in human diseases: potential applications to melanoma research," *Expert Review of Proteomics*, vol. 5, no. 3, pp. 405–412, 2008.

[19] B. L. Hood, M. M. Darfler, T. G. Guiel et al., "Proteomic analysis of formalin-fixed prostate cancer tissue," *Molecular and Cellular Proteomics*, vol. 4, no. 11, pp. 1741–1753, 2005.

[20] S. R. Shi, C. Liu, B. M. Balgley et al., "Protein extraction from formalin-fixed paraffin-embedded tissue sections: quality evaluation by mass spectometry," *Journal of Histochemistry & Cytochemistry*, vol. 54, no. 6, pp. 739–743, 2006.

[21] X. Jiang, S. Feng, R. Tian, M. Ye, and H. Zou, "Development of efficient protein extraction methods for shotgun proteome analysis of formalin-fixed tissues," *Journal of Proteome Research*, vol. 6, no. 3, pp. 1038–1047, 2007.

[22] D. E. Palmer-Toy, B. Krastins, D. A. Sarracino, J. B. Nadol, and S. N. Merchant, "Efficient method for the proteomic analysis of fixed and embedded tissues," *Journal of Proteome Research*, vol. 4, no. 6, pp. 2404–2411, 2005.

[23] S. I. Hwang, J. Thumar, D. H. Lundgren et al., "Direct cancer tissue proteomics: a method to identify candidate cancer biomarkers from formalin-fixed paraffin-embedded archival tissues," *Oncogene*, vol. 26, no. 1, pp. 65–76, 2007.

[24] M. Takikawa, Y. Akiyama, T. Ashizawa et al., "Identification of melanoma-specific serological markers using proteomic analyses," *Proteomics—Clinical Applications*, vol. 3, no. 5, pp. 552–562, 2009.

[25] H. J. Issaq, T. D. Veenstra, T. P. Conrads, and D. Felschow, "The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification," *Biochemical and Biophysical Research Communications*, vol. 292, no. 3, pp. 587–592, 2002.

[26] T. W. Hutchens and T. T. Yip, "New desorption strategies for the mass spectrometric analysis of macromolecules," *Rapid Communications in Mass Spectrometry*, vol. 7, pp. 576–580, 1993.

[27] J. Caron, A. Mange, B. Guillot, and J. Solassol, "Highly sensitive detection of melanoma based on serum proteomic profiling," *Journal of Cancer Research and Clinical Oncology*, vol. 135, no. 9, pp. 1257–1264, 2009.

[28] L. L. Wilson, L. Tran, D. L. Morton, and D. S. B. Hoon, "Detection of differentially expressed proteins in early-stage melanoma patients using SELDI-TOF mass spectrometry," *Annals of the New York Academy of Sciences*, vol. 1022, pp. 317–322, 2004.

[29] S. Mian, S. Ugurel, E. Parkinson et al., "Serum proteomic fingerprinting discriminates between clinical stages and predicts disease progression in melanoma patients," *Journal of Clinical Oncology*, vol. 23, no. 22, pp. 5088–5093, 2005.

[30] P. Findeisen, M. Zapatka, T. Peccerella et al., "Serum amyloid A as a prognostic marker in melanoma identified by proteomic profiling," *Journal of Clinical Oncology*, vol. 27, no. 13, pp. 2199–2208, 2009.

[31] M. Greco, M. De Metri, F. Chiriaco, G. Leo, E. Brienza, and M. Maffia, "Serum proteomic profile of cutaneous malignant melanoma and relation to cancer progression: association to tumor derived alpha-N-acetylgalactosaminidase activity," *Cancer Letters*, vol. 283, no. 2, pp. 222–229, 2009.

[32] M. Neagu, C. Constantin, and C. Tanase, "Immune-related biomarkers for diagnosis/prognosis and therapy monitoring of cutaneous melanoma," *Expert Review of Molecular Diagnostics*, vol. 10, pp. 897–919, 2010.

[33] A. Wankowicz-Kalinska, C. Le Poole, R. Van Den Wijngaard, W. J. Storkus, and P. K. Das, "Melanocyte-specific immune response in melanoma and vitiligo: two faces of the same coin?" *Pigment Cell Research*, vol. 16, no. 3, pp. 254–260, 2003.

[34] H. Uchi, R. Stan, M. J. Turk et al., "Unraveling the complex relationship between cancer immunity and autoimmunity: lessons from melanoma and vitiligo," *Advances in Immunology*, vol. 90, pp. 215–241, 2006.

[35] T. Ramirez-Montagut, M. J. Turk, J. D. Wolchok, J. A. Guevara-Patino, and A. N. Houghton, "Immunity to melanoma: unraveling the relation of tumor immunity and autoimmunity," *Oncogene*, vol. 22, no. 20, pp. 3180–3187, 2003.

[36] A. Suzuki, A. Iizuka, M. Komiyama et al., "Identification of melanoma antigens using a serological proteome approach (SERPA)," *Cancer Genomics and Proteomics*, vol. 7, no. 1, pp. 17–23, 2010.

[37] Y. Liu, J. He, X. Xie et al., "Serum autoantibody profiling using a natural glycoprotein microarray for the prognosis of early melanoma," *Journal of Proteome Research*, vol. 9, no. 11, pp. 6044–6051, 2010.

[38] E. P. Diamandis, "Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems," *Journal of the National Cancer Institute*, vol. 96, no. 5, pp. 353–356, 2004.

[39] E. P. Diamandis, "Proteomic patterns in biological fluids: do they represent the future of cancer diagnostics?" *Clinical Chemistry*, vol. 49, no. 8, pp. 1272–1275, 2003.

[40] P. R. Srinivas, M. Verma, Y. Zhao, and S. Srivastava, "Proteomics for cancer biomarker discovery," *Clinical Chemistry*, vol. 48, no. 8, pp. 1160–1169, 2002.

[41] R. J. Simpson and D. S. Dorow, "Cancer proteomics: from signaling networks to tumor markers," *Trends in Biotechnology*, vol. 19, supplement 10, pp. S40–S48, 2001.

[42] B. Franzen, T. Hirano, K. Okuzawa et al., "Sample preparation of human tumors prior to two-dimensional electrophoresis of proteins," *Electrophoresis*, vol. 16, no. 7, pp. 1087–1089, 1995.

[43] R. E. Banks, M. J. Dunn, M. A. Forbes et al., "The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis—preliminary findings," *Electrophoresis*, vol. 18, pp. 622–624, 1999.

[44] H. T. Tan, J. Low, S. G. Lim, and M. C. M. Chung, "Serum autoantibodies as biomarkers for early cancer detection," *FEBS Journal*, vol. 276, no. 23, pp. 6880–6904, 2009.

[45] M. S. Pepe, R. Etzioni, Z. Feng et al., "Phases of biomarker development for early detection of cancer," *Journal of the National Cancer Institute*, vol. 93, no. 14, pp. 1054–1056, 2001.

[46] V. Paulitschke, R. Kunstfeld, T. Mohr et al., "Entering a new era of rational biomarker discovery for early detection of melanoma metastases: secretome analysis of associated stroma cells," *Journal of Proteome Research*, vol. 8, no. 5, pp. 2501–2510, 2009.

[47] B. L. Hood, J. Grahovac, M. S. Flint et al., "Proteomic analysis of laser microdissected melanoma cells from skin organ cultures," *Journal of Proteome Research*, vol. 9, no. 7, pp. 3656–3663, 2010.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 413742, 9 pages
doi:10.1155/2011/413742

Research Article

Clinical Utility of Serum Autoantibodies Detected by Protein Microarray in Melanoma

Michael S. Sabel, Yashu Liu, Kent A. Griffith, Jintang He, Xaiolei Xie, and David M. Lubman

Department of Surgery, University of Michigan Health Systems Biostatistics Core, University of Michigan Comprehensive Cancer Center, Ann Arbor, MI 48109, USA

Correspondence should be addressed to Michael S. Sabel, msabel@med.umich.edu

Received 12 July 2011; Accepted 11 August 2011

Academic Editor: David E. Misek

Copyright © 2011 Michael S. Sabel et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Better prognostic and predictive markers in melanoma are needed to select patients for therapy. We utilized a dual-lectin affinity chromatography and a natural protein microarray-based analysis to select a subproteome of target glycoproteins to profile serum antibodies against melanoma associated antigens that may predict nodal positivity. We identified 5 melanoma-associated antigens using this microarray coupled to mass spectrometry; GRP75, GRP94, ASAH1, CTSD and LDHB. We evaluated their predictive value for nodal status adjusting for age, gender, Breslow thickness, mitotic rate and ulceration using standard logistic regression. After adjustment, ASAH1, CTSD and LDHB were significantly negatively associated with nodal status ($P = 0.0008$) and GRP94 was significantly positively associated ($P = 0.014$). Our best multivariate model for nodal positivity included Breslow thickness, presence of serum anti-ASAH1, anti-LDHB or anti-CTSD, and presence of serum anti-GRP94, with an area under the ROC curve of 0.869. If validated, these results show promise for selecting clinically node negative patients for SLN biopsy. In addition, there is strong potential for glycoprotein microarray to screen serum autoantibodies that may identify patients at high risk of distant metastases or those likely or unlikely to respond to treatment, and these proteins may serve as targets for intervention.

1. Introduction

The present staging system for melanoma, using Breslow thickness, ulceration, mitotic rate, and the presence of regional and distant metastases, stratifies patients into heterogenous groups, with wide variability in outcome or response to therapy. This results in applying more aggressive surgical and adjuvant therapies to large populations, diluting the impact of therapy while exposing more patients to toxicity. Better biomarkers in melanoma are needed to target both surgical and adjuvant therapies, but to date have been elusive. For many solid tumors, the large-scale analysis of gene expression at the RNA level can provide patterns of gene expression that may stratify patients better than TNM staging and help guide therapy. However, this approach requires fresh tissue from a large number of primary tumors, a unique challenge in melanoma where the primary is often only a few millimeters in size, with no residual tissue after the diagnosis has been made.

For this reason, we chose to examine serum protein markers, hypothesizing that antibody discovery was ideal for the patient with malignant melanoma, as primary tumor tissue is not required and the presence of an immune response to melanoma-associated antigens has been well documented [1–4]. The investigation of humoral response provides new perspective to focus on melanoma-associated antibodies, which are more sensitive and stable to become diagnostic biomarkers for early-stage melanoma. We focused on glycoproteins, as most of the tumor-associated antigens are cell surface proteins or released to the extracellular matrix, where glycosylation is the major type of posttranslational modifications [5, 6]. Moreover, glycoproteins are considered to be the linkage between T cells and antigen-presenting cells to help the orientation of binding, and play important roles in the generation and loading of antigenic peptides into MHC class I and MHC class II [5–7].

Using this approach we sought biomarkers that correlated with the presence of regional metastases among

melanoma patients. Using dual-lectin affinity chromatography and a natural protein microarray-based analysis to select a subproteome of target glycoproteins which were then used as baits to profile the antibodies against melanoma-associated antigens [8]. This significantly improved technology using lectin affinity chromatography allows us to concentrate low abundant glycoproteins which are typically undetectable in whole cell lysate. This approach led us to the discovery of antibodies to 5 interesting melanoma-associated antigens (75 kD glucose-regulated protein (GRP75), 94 kD glucose-regulated protein (GRP94), acid ceramidase (ASAH1), cathepsin D (CTSD), and lactate dehydrogenase B (LDHB)) that correlated with the presence of melanoma within the regional lymph nodes [8]. GRP75, also known as mortalin, is a transport protein. A member of the heat shock protein-70 family, it also inactivates the tumor suppressor p53. GRP94, also known as heat shock protein-90, is a chaperone protein that is involved in the function and stability of many cell-signaling molecules. ASAH1 is a catabolic lysosomal enzyme that deacylates ceramide, which when phosphorylated forms the poten mitogen S1P. CTSD is a lysosomal acid proteinase which is involved in regulation of programmed cell death. Lactate dehydrogenase (LDH) is an enzyme that catalyzes the conversion of lactate to pyruvate, and serum levels are associated with outcome in stage IV melanoma. We proposed that these autoantibodies may form the basis of a serum test that could select patients for sentinel lymph node biopsy. However, many prognostic factors show limited utility when used clinically in the context of known prognostic factors. We therefore sought to examine the potential clinical utility of these novel serum markers for predicting regional involvement among patients with melanoma.

2. Materials and Methods

2.1. Patients. In our previous work, we identified serum autoantibodies that recognized glycoproteins from a melanoma cell line and distinguished between 27 node-negative patients and 16 node-positive patients. In that work, we subsequently validated these results using recombinant proteins among a larger sample set of 79 patients. For this University of Michigan Institutional Review Board approved project, we used this latter sample set to examine the clinical utility of these serum autoantibodies as a predictor of regional node involvement. Serum samples were obtained from patients being evaluated at our melanoma multidisciplinary clinic, a few weeks after the diagnostic biopsy, but 2 to 3 weeks prior to undergoing wide local excision and lymph node surgery (SLN biopsy for clinically node-negative patients ($n = 71$) or lymph node dissection for clinically node-positive patients ($n = 8$)). Blood was allowed to clot at room temperature, after which the tubes were centrifugated at 2500 g for 10 minutes. The serum phase was then harvested and frozen in 1 mL aliquots.

2.2. Measurement of Serum Autoantibody Levels. Our initial discovery (using extracted glycoproteins) and validation (using recombinant proteins) of these serum autoantibodies have been previously described [8]. For this study we used the results obtained with the recombinant proteins. 75 kD glucose-regulated protein (GRP75), 94 kD glucose-regulated protein (GRP94), cathepsin D (CTSD), and lactate dehydrogenase B (LDHB) were purchased from Abcam (Cambridge, MA, USA). Recombinant acid ceramidase (ASAH1) was purchased from Abnova (Taiwan). These 5 recombinant proteins were chosen because the amino acid sequences described in the manufacturers' instructions are perfectly matched with the sequences acquired from Swiss-Prot database. The sequences and the purity of purchased recombinant proteins were reconfirmed by MALDI-QIT (Shimadzu, CA, USA).

To summarize our previous work, the recombinant proteins were dissolved in the printing buffer (62.5 mM Tris-HCl (pH 6.8), 1% w/v sodium dodecyl sulfate (SDS), 5% w/v dithiothreitol (DTT), and 1% glycerol in 1x PBS) to reach a final concentration of 100 μg/mL, respectively. Each protein solution was then transferred to a well in a 200 μL 96-well clear printing plate (Bio-rad). The recombinant proteins from the printing plate were spotted onto nitrocellulose (Whatman, USA) slides using a noncontact piezoelectric printer (nanoplotter 2 GeSiM). Each spot contains five spotting events of 500 pL each so that the total volume of each protein solution was 2.5 nL. Each spot was found to be $\sim$450 μm in diameter, with the distance between spots maintained at 600 μm. Printed slides were dried on the printer deck overnight and stored in a refrigerator desiccated at 4°C if the slides were not used immediately. Each recombinant protein was printed in triplicate, and 14 identical blocks were printed on each slide.

The slides were washed three times with 0.1% Tween-20 in PBS buffer (PBST) and then blocked with 1% bovine serum albumin (Roche) in PBST for 1 hr. The blocked slides were dried by centrifugation and inserted into a SIMplex (Gentel Bioscience) multiarray device which divides each slide by 16 wells. The wells separate the neighboring blocks and prevent cross-contamination. Each serum sample was diluted 1 : 200 in probe buffer which consisted of 1% BSA, 0.05% Triton X-100, 0.1% brij-30 (Sigma-Aldrich, USA) in 1x PBS. The sample hybridization was totally randomized on each slide in no specific order to prevent bias. Each block was hybridized in 100 microliter of diluted serum for 2 hrs at 4°C. Then goat-anti-human IgG (H+L) conjugated with Alexa Fluor 647 (1 μg/mL, Invitrogen, Carlsbad, CA) was applied to each block to bind with the antibodies attached on the protein array. Anti-human IgG was printed on the array as positive control and printing buffer served as the negative control. All processed slides were immediately scanned using an Axon 4000B microarray scanner (Axon Instruments, Foster City, USA). GenePix Pro 6.0 was used to extract the numerical data from each spot on the slides. The background subtracted median intensity of each spot was taken as a single data point. Then the mean intensity of each protein from the triplicate was used for the further analysis.

2.3. Statistical Analysis. For this study, patient and tumor characteristics were collected for our sample of 79 melanoma cases, it included patient's age at surgery, gender, tumor

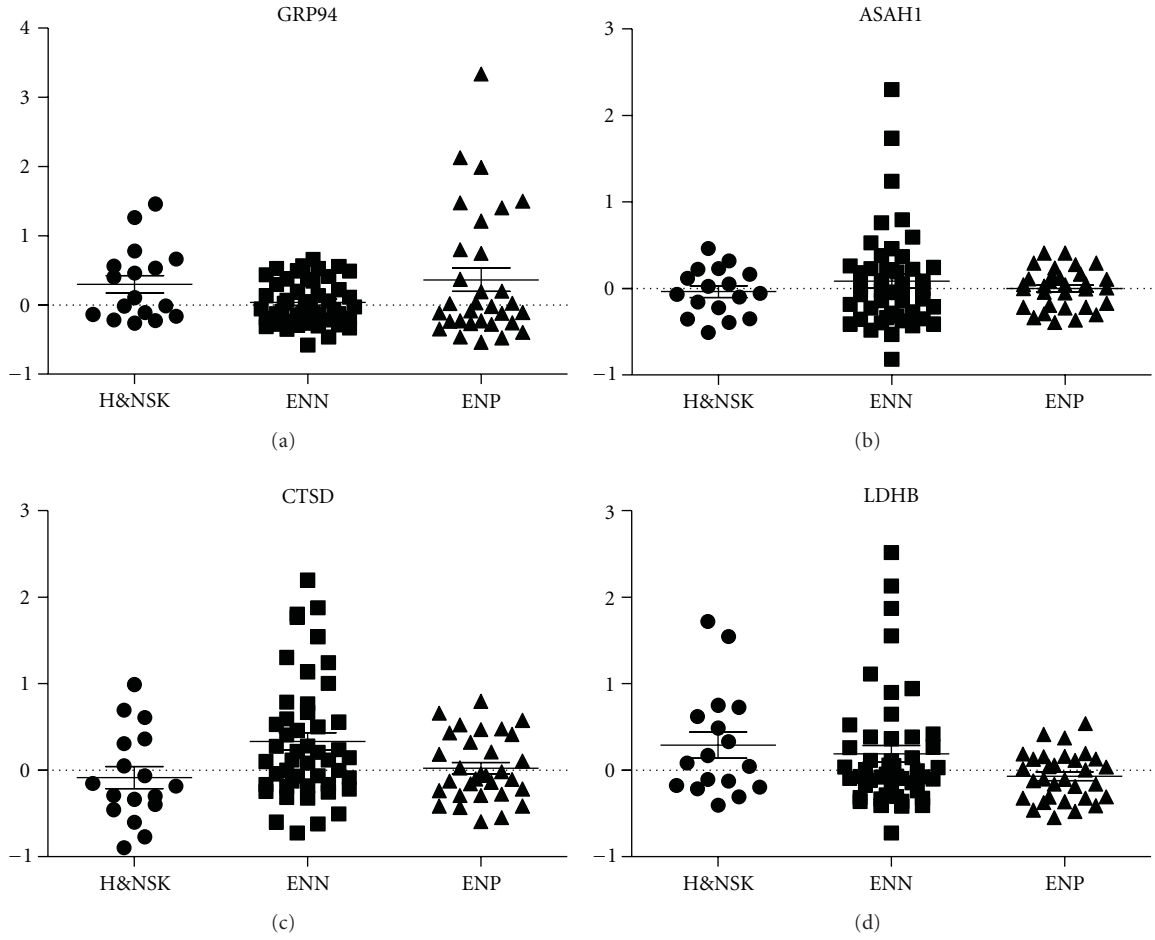


Figure 1: Humoral response to recombinant 75 kD glucose-regulated protein (GRP75), 94 kD glucose-regulated protein (GRP94), acid ceramidase (ASAH1), cathepsin D (CTSD), and lactate dehydrogenase B (LDHB) among healthy volunteers and nonmelanoma skin cancer patients (H&NSK), node-negative melanoma patients (ENN) and node-positive melanoma patients (ENP).

thickness (Breslow), mitotic rate, presence of ulceration, and nodal status. For use as a potential clinical test, the sample distribution of each auto-antibody classified as over- or underexpressed, defined as a serum antibody level one standard deviation increment above the sample mean value. Association between the levels of the autoantibodies is summarized by the Spearman rank correlation coefficient with P values testing for significant correlations reported. The associations between patient, tumor, and antibody covariates with nodal disease was compared using the two-sample t-test for continuous covariates and the chi-square or Fisher's exact test for categorical covariates. The magnitude of the association between each serum antibody level and nodal disease is reported categorically as the odds ratio and 95% confidence interval for cases with overexpression versus cases without. Odds ratios and confidence intervals were reported separately for the univariate associations and after adjustment for the patient and tumor characteristics using standard logistic regression. All statistical analyses were conducted using SAS Version 9.2 (SAS Institute, Inc., Cary, NC, USA) with P values less than 5% considered statistically significant.

3. Results

In the previous work, we used the native proteins extracted by a dual-lectin column from the melanoma cell line as bait to detect the presence of autoantibodies in the sera of melanoma patients, identifying 5 antigens including 75 kD glucose-regulated protein (GRP75), 94 kD glucose-regulated protein (GRP94), acid ceramidase (ASAH1), cathepsin D (CTSD), and lactate dehydrogenase B (LDHB), and we investigated the humoral response against the recombinant proteins using a larger sample set of 79 melanoma patients (Figure 1). The clinical characteristics of the patient population are shown in Table 1 for the total population and stratified by nodal disease status. Of note, one patient with a negative SLN subsequently recurred in a regional basin, changing the population from 48 node-negative and 31 node-positive to 47 node-negative and 32 node-positive patients. Among the 32 node-positive patients, 8 were clinically node-positive, 23 were SLN positive, and 1 represented a regional recurrence after false-negative SLN. Each glycoprotein was summarized by over- (>1SD), standard (±1SD), and underexpression (<1SD), relative to the sample mean. ASAH1, CTSD, and

Table 1: Patient and tumor characteristics of the study cohort.

Characteristic	Total	Node negative	Node positive	*P* value[1]
Frequency	79	47	32	
Age				
Mean (SD)	51.9 (12.6)	51.8 (13.2)	52.0 (12.0)	0.944
Gender				
Males: *N* (%)	38 (48.1)	22 (46.8)	16 (50.0)	0.781
Females: *N* (%)	41 (51.9)	25 (53.2)	16 (50.0)	
Breslow depth				
Mean (SD)	2.39 (1.65)	1.94 (1.15)	3.11 (2.04)	0.007
Mitotic rate				
Mean (SD)	5.51 (5.72)	4.45 (4.93)	7.04 (6.50)	0.075
Ulceration				
Absent: *N* (%)	24 (30.4)	35 (74.5)	16 (50.0)	0.059[2]
Present: *N* (%)	51 (64.6)	11 (23.4)	13 (40.6)	
Unknown: *N* (%)	4 (5.0)	1 (2.1)	3 (9.4)	
ASAH1				
Mean (SD)	803 (353)	837 (432)	753 (179)	0.238
±1 SD: *N* (%)	68 (86.1)	36 (76.6)	32 (100)	0.008
>1 SD: *N* (%)	8 (10.1)	8 (17.0)	0	
<1 SD: *N* (%)	3 (3.8)	3 (6.4)	0	
CTSD				
Mean (SD)	8664 (4355)	9607 (5014)	7278 (2660)	0.009
±1 SD: *N* (%)	62 (78.5)	33 (70.2)	29 (90.6)	0.039
>1 SD: *N* (%)	12 (15.2)	11 (23.4)	1 (3.1)	
<1 SD: *N* (%)	5 (6.3)	3 (6.4)	2 (6.3)	
GRP75				
Mean (SD)	4087 (2107)	4221 (2367)	3891 (1671)	0.471
±1 SD: *N* (%)	62 (78.5)	36 (76.6)	26 (81.3)	0.913
>1 SD: *N* (%)	13 (16.5)	8 (17.0)	5 (15.6)	
<1 SD: *N* (%)	4 (5.1)	3 (6.4)	1 (3.1)	
GRP94				
Mean (SD)	6798 (3782)	6032 (1840)	7924 (5364)	0.063
±1 SD: *N* (%)	70 (88.6)	47 (100)	23 (71.9)	<0.001
>1 SD: *N* (%)	9 (11.4)	0	9 (28.1)	
<1 SD: *N* (%)	0	0	0	
LDH				
Mean (SD)	7863 (4093)	8587 (4933)	6798 (2021)	0.029
±1 SD: *N* (%)	68 (86.1)	37 (78.7)	31 (96.9)	0.056
>1 SD: *N* (%)	10 (12.7)	9 (19.2)	1 (3.1)	
<1 SD: *N* (%)	1 (1.3)	1 (2.1)	0	
Combination: ASAH1, CTSD, and LDH				
Overexpressed[†]	22 (27.9)	20 (42.5)	2 (6.2)	<0.001
Normal	57 (72.1)	27 (57.5)	30 (93.8)	

[1] Comparing between node-negative and node-positive groups.
[2] Unknown group omitted for statistical test.
[†] Overexpressed defined as >1 SD for ASAH1, CTSD, or LDH.

Table 2: Correlation of antibodies and continuous patient and tumor characteristics: Spearman *r*, *P* value.

	ASAH1	CTSD	GRP75	GRP94	LDH
Age	**−0.2317**	**0.0811**	**−0.0024**	**0.0214**	**0.0539**
	0.0400	0.4774	0.9836	0.8514	0.6371
Breslow	**0.1294**	**0.0822**	**0.1975**	**0.0878**	**0.1444**
	0.2686	0.4834	0.0894	0.4539	0.2165
Mitotic rate	**0.0630**	**−0.0253**	**0.0516**	**0.0705**	**0.0485**
	0.6019	0.8341	0.6693	0.5593	0.6881

Table 3: Univariate associations of glycoproteins with positive nodal status.

Characteristic	Odds ratio	95% CI	*P* value
ASAH1			
500 unit increase	0.685	0.332–1.415	0.306
>1 SD versus not	0.123†		0.019
CTSD			
500 unit increase	0.928	0.869–0.991	0.026
>1 SD versus not	0.106	0.013–0.864	0.022
GRP75			
500 unit increase	0.962	0.860–1.076	0.495
>1 SD versus not	0.903	0.266–3.059	0.999
GRP94			
500 unit increase	1.076	1.002–1.156	0.0440
>1 SD versus not	38.4†		<0.0001
LDH			
500 unit increase	0.931	0.862–1.007	0.0731
>1 SD versus not	0.193†		0.0427
Combination: ASAH1, CTSD, and LDH			
Any >1 SD versus not	0.173	0.045–0.663	<0.0001

† Continuity correction applied when calculating the estimate of odds ratios due to cell sample sizes ≤1. Confidence interval not reportable in these cases.

LDHB all had significant negative associations with the presence of nodal disease, with overexpression associated with a lower risk. Higher GRP94 levels were associated with a higher risk of nodal disease; however, the level of GRP75 was not significantly associated with nodal status. Correlation between antibody levels and age, gender, Breslow thickness, mitotic rate, or ulceration are shown in Table 2. Correlations included ASAH1 which negatively correlated with patient age (although the magnitude of the correlation was mild, $r < 0.3$) and GRP75 which was higher in females than males ($P = 0.03$).

The magnitude of the association of the glycoproteins with nodal disease status is reported in Table 3. Even after adjustment for clinical parameters (age, gender, ulceration, and Breslow thickness) ASAH1, CTSD, and LDHB remained significantly negatively associated with nodal disease, and GRP94 positively associated (Table 4). A composite measure for the overexpression of any of the three proteins shown to have a negative association with nodal disease (ASAH1, CTSD, and LDHB) was constructed with 22 (27.9%) of the total population with composite overexpression. Only 2 (6.2%) of the node-positive patients had composite overexpression in contrast to 20 (42.5%) node-negative patients. For the 9 of 32 node-positive patients (28%) who had overexpression of GRP94, 5 (16%) patients had clinically involved nodes while the remaining 4 (14%) had microscopic disease (<2% surface area) in clinically negative nodes. Table 5 reports the best multivariable model for nodal positivity. Overexpression of anti-ASAH1, anti-LDHB, or anti-CTSD (decreased risk), the overexpression of anti-GRP94 (increased risk) and Breslow thickness (increased risk) significantly correlated with the likelihood of regional metastases. The ROC curve for this model is presented in Figure 2, with an area under the curve of 0.8690.

4. Discussion

Using dual-lectin affinity chromatography to generate a subproteome of glycoproteins from a melanoma cell line generated from a metastatic deposit, we discovered 4 antibodies in the serum of melanoma patients recently diagnosed with melanoma that were strongly correlated with the presence or absence of nodal metastases. In this analysis, we demonstrate that overexpression of these antibodies were independent of

Table 4: Adjusted[†] associations of glycoproteins with positive nodal status.

Characteristic	Odds ratio	95% CI	*P* value
ASAH1			
500 unit increase	0.272	0.076–0.979	0.0464
>1 SD versus not		Model not estimable	
CTSD			
500 unit increase	0.867	0.786–0.956	0.0042
>1 SD versus not	0.067	0.006–0.740	0.0275
GRP75			
500 unit increase	0.880	0.749–1.034	0.1216
>1 SD versus not	0.644	0.153–2.712	0.5488
GRP94			
500 unit increase	1.052	0.976–1.133	0.1871
>1 SD versus not		Model not estimable	
LDH			
500 unit increase	0.890	0.797–0.993	0.0363
>1 SD versus not	0.135	0.015–1.210	0.0736
Combination: ASAH1, CTSD, and LDH			
Any >1 SD versus none	0.045	0.007–0.309	0.0016

[†] Adjusted for age, gender, ulceration, and Breslow depth. Due to the very high correlation between mitotic rate and Breslow depth, mitotic rate was omitted from the model.

Table 5: Best multivariable model explaining positive nodal status.

Characteristic	Odds ratio	95% CI	*P* value
Age			
1 year increase	0.987	0.935–1.043	0.6468
Gender			
Male	1.663	0.426–6.495	0.4645
Female	1.000		
Ulceration			
Present	2.046	0.359–11.662	0.4201
Absent	1.000		
Breslow			
1 mm increase	2.178	1.104–4.298	0.0248
Combination: ASAH1, CTSD, and LDH			
Overexpressed	0.006	< 0.001–0.117	0.0008
Not overexpressed	1.000		
GRP94			
Overexpressed	1.223	1.041–1.436	0.0141
Not overexpressed	1.000		

other known prognostic factors in melanoma, and on multivariate analysis maintained highly significant, independent prognostic value. These results demonstrate the potential of these 4 autoantibodies as a serum test for the purpose of selecting clinically node-negative patients for SLN biopsy. Elevation (defined as >1 SD above the mean) of anti-ASAH1, anti-CTSD, or anti-LDHB was highly significantly associated with being SLN negative, with an odds ratio of 0.05 (0.01–0.31, $P = 0.002$) after adjusting for age, gender, ulceration, mitotic rate, and Breslow thickness. Elevation of one of these three antibodies was not uncommon; among the 70 clinically node-negative patients in this study, 22 patients (33%) had elevation of one of these antibodies. If it were validated that patients with elevation of one of these antibodies had a very low risk of SLN metastases, then potentially these patients could be treated by wide excision alone, reducing the cost and morbidity of melanoma treatment. In contrast, elevation of anti-GRP94 was associated with an increased risk of regional metastases, and was elevated in both patients with clinically evident disease (5 of 8 patients) and microscopic disease

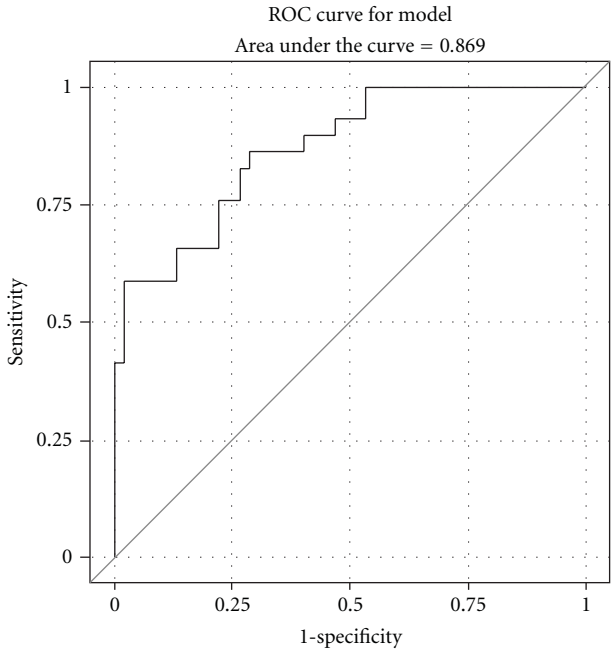


Figure 2: Receiver operating characteristic (ROC) curve for our best multivariable model for nodal positivity incorporating increasing Breslow thickness (increased risk), autoantibody response to GRP94 (increased risk), and autoantibody response to ASAH1, LDHB, or CTSD (decreased risk).

(4 of 23 patients). While detection of serum anti-GRP94 levels would be less useful clinically, it could potentially identify some patients with thin melanoma for whom SLN might otherwise be omitted.

A bigger question is the role these proteins may play in the development and progrsion of melanoma. Beyond the development of serum-based diagnostic tests, proteomics may identify targets for therapeutic intervention. Indeed, 3 of the 4 proteins have strong associations with melanoma progression and prognosis. Lactate dehydrogenase (LDH) is an enzyme that catalyzes the conversion of lactate to pyruvate, and serum levels of LDH are strongly associated with melanoma prognosis. Serum LDH levels strongly correlate with outcome among stage IV patients and serum LDH measurements are part of the American Joint Cancer Commission (AJCC) staging system for melanoma [9–13]. However, serum LDH levels are rarely elevated and of no clinical utility in nonmetastatic melanoma [14–16]. Cathepsin D (CTSD) is a lysosomal acid proteinase which degrades proteins, peptides, and peptide precursors. In addition, it appears to be involved in other biological processes including regulation of programmed cell death, tissue remodeling and renewal, activation of proteolytic enzymes, and fibrinolysis [17]. Many tumors have altered processing, secretion, and activity levels of CTSD, and they are often associated with aggressive behavior, stimulating tumor cell proliferation, invasion, and metastases [17, 18]. Immunohistochemical studies have shown that CTSD is markedly expressed in melanoma cell lines and tissue biopsies from primary and metastatic melanoma, and these correlate with poor outcome [19–24]. As with LDH, measuring plasma levels of CTSD was not of clinical value for identifying patients at risk of recurrence [17, 25]. As we can detect very low levels of antibodies in the serum, measuring antibody levels may be more sensitive than measuring protein levels, allowing transition of known serum markers from utility in stage IV disease only to the setting of early-stage disease. Prior to this publication, the third protein, acid ceramidase (ASAH1), had not been strongly associated with melanoma progression, but has been associated with cancers of the breast, prostate, and thyroid [26–29]. ASAH1 is a catabolic lysosomal enzyme that deacylates ceramide and yields sphingosine, which when phosphorylated, forms the potent mitogen S1P. The cellular levels of ceramide, sphingosine, and S1P are integral in determining cell survival and growth [30–32]. Targeting this pathway holds promise for anticancer therapies [32–34].

In the case of these three proteins, expression of the proteins is associated with advanced stage, but the presence of antibodies to these proteins is associated with lower stage. This highlights one drawback to antibody-array-based proteomics—the presence of the antibody may be due to increased exposure of the proteins (overexpression), immune recognition of protein alteration, or the antibodies may be functionally blocking critical pathways. In the case of these three proteins/antibodies, it remains unclear whether this represents specific functional inhibition by the antibodies, or increased progression in the face of decreased immune recognition of overexpressed proteins (disease advancement in the face of decreased immune surveillance). As these proteins are identified, further analysis of their role in melanoma progression, and their posttranslational structure is necessary.

In contrast to these three proteins, for which the detection of autoantibodies was a favorable prognostic sign, the presence of autoantibodies to GRP94, or heat shock protein-90 (HSP90), was associated with an increased risk of regional metastases. Although this protein is highly conserved (and should not trigger a significant immune response), our primary data and validation studies using the recombinant proteins demonstrate the presence of anti-GRP94 antibodies in the serum of nearly between 1/4 and 1/3 of node-positive patients. HSP90 is a chaperone protein that is crucially involved in the function and stability of many oncogene products and cell-signaling molecules, including CRAF, ERB-B2, BCR-ABL, CDK4, CDK6, AKT, mutated p53, MEK, VEGFR, and importantly to melanoma mutated (but not wildtype) BRAF. HSP90 protects these proteins from deterioration caused by environmental stress, which includes cancer therapy. Expression of HSP90 is elevated in melanoma, correlates with increasing Breslow thickness, and is associated with advanced disease [35]. Because HSP90 chaperones so many proteins implicated in carcinogenesis, inhibiting HSP90 could inhibit several pathways at once, HSP90 inhibitors are presently in clinical trial in metastatic melanoma. While our data suggest that autoantibodies do little to inhibit HSP90 functionally, their presence as a response to overexpression is clearly related to melanoma progression.

In conclusion, the creation of a glycoprotein microarray to screen melanoma patient serum samples for autoantibodies yielded four autoantibodies that show promise in predicting regional metastases, and could potentially form the basis of a blood test to select clinically node-negative patients for SLN biopsy. If validated, this test could greatly minimize the cost and morbidity associated with the surgical treatment of melanoma, as well as identify patients with thin melanomas who should undergo the procedure. In addition, glycoproteins recognized by these antibodies may have important roles in the development and progression of melanoma and may serve as targets for intervention. On a broader scale, this approach could be used to identify additional serum autoantibodies that can identify patients at high risk of distant metastases and those unlikely to respond to treatment, allowing a more tailored use of adjuvant therapies.

Acknowledgment

This paper was supported in part by the National Cancer Institute under Grant R01CA106402 (D. M. Lubman) and R21CA124441 (D. M. Lubman), Schering Corporation, as subsidiary of Merck & Co., Inc., Whitehouse Station, NJ, and an MICHR grant from University of Michigan.

References

[1] M. Neagu, C. Constantin, and C. Tanase, "Immune-related biomarkers for diagnosis/prognosis and therapy monitoring of cutaneous melanoma," *Expert review of molecular diagnostics*, vol. 10, no. 7, pp. 897–919, 2010.

[2] A. Wankowicz-Kalinska, C. Le Poole, R. Van Den Wijngaard, W. J. Storkus, and P. K. Das, "Melanocyte-specific immune response in melanoma and vitiligo: two faces of the same coin?" *Pigment Cell Research*, vol. 16, no. 3, pp. 254–260, 2003.

[3] H. Uchi, R. Stan, M. J. Turk et al., "Unraveling the complex relationship between cancer immunity and autoimmunity: lessons from melanoma and vitiligo," *Advances in Immunology*, vol. 90, pp. 215–241, 2006.

[4] T. Ramirez-Montagut, M. J. Turk, J. D. Wolchok, J. A. Guevara-Patino, and A. N. Houghton, "Immunity to melanoma: unraveling the relation of tumor immunity and autoimmunity," *Oncogene*, vol. 22, no. 20, pp. 3180–3187, 2003.

[5] P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, and R. A. Dwek, "Glycosylation and the immune system," *Science*, vol. 291, no. 5512, pp. 2370–2376, 2001.

[6] P. M. Rudd, M. R. Wormald, R. L. Stanfield et al., "Roles for glycosylation of cell surface receptors involved in cellular immune recognition," *Journal of Molecular Biology*, vol. 293, no. 2, pp. 351–366, 1999.

[7] H. A. Chapman, "Endosomal proteolysis and MHC class II function," *Current Opinion in Immunology*, vol. 10, no. 1, pp. 93–102, 1998.

[8] Y. Liu, J. He, X. Xie et al., "Serum autoantibody profiling using a natural glycoprotein microarray for the prognosis of early melanoma," *Journal of Proteome Research*, vol. 9, no. 11, pp. 6044–6051, 2010.

[9] C. M. Balch, J. E. Gershenwald, S. J. Soong et al., "Final version of 2009 AJCC melanoma staging and classification," *Journal of Clinical Oncology*, vol. 27, no. 36, pp. 6199–6206, 2009.

[10] A. Y. Bedikian, M. M. Johnson, C. L. Warneke et al., "Prognostic factors that determine the long-term survival of patients with unresectable metastatic melanoma," *Cancer Investigation*, vol. 26, no. 6, pp. 624–633, 2008.

[11] U. Keilholz, P. Martus, C. J. A. Punt et al., "Prognostic factors for survival and factors associated with long-term remission in patients with advanced melanoma receiving cytokine-based treatments: second analysis of a randomised EORTC Melanoma Group trial comparing interferon-α2a (IFNα) and interleukin 2 (IL-2) with or without cisplatin," *European Journal of Cancer*, vol. 38, no. 11, pp. 1501–1511, 2002.

[12] J. Manola, M. Atkins, J. Ibrahim, and J. Kirkwood, "Prognostic factors in metastatic melanoma: a pooled analysis of Eastern Cooperative Oncology Group trials," *Journal of Clinical Oncology*, vol. 18, no. 22, pp. 3782–3793, 2000.

[13] A. Barth, L. A. Wanek, and D. L. Morton, "Prognostic factors in 1,521 melanoma patients with distant metastases," *Journal of the American College of Surgeons*, vol. 181, no. 3, pp. 193–201, 1995.

[14] F. Egberts, W. N. Hitschler, M. Weichenthal, and A. Hauschild, "Prospective monitoring of adjuvant treatment in high-risk melanoma patients: lactate dehydrogenase and protein S-100B as indicators of relapse," *Melanoma Research*, vol. 19, no. 1, pp. 31–35, 2009.

[15] T. M. Johnson, C. R. Bradford, S. B. Gruber, V. K. Sondak, and J. L. Schwartz, "Staging workup, sentinel node biopsy, and follow-up tests for melanoma: update of current concepts," *Archives of Dermatology*, vol. 140, no. 1, pp. 107–113, 2004.

[16] L. Lugovic, M. Situm, M. Buljan et al., "Results of the determination of serum markers in patietns wtih malignant melanoma," *Collegium Antropologicum*, vol. 31, supplement 1, pp. 7–11, 2007.

[17] G. Leto, F. M. Tumminello, M. Crescimanno, C. Flandina, and N. Gebbia, "Cathepsin D expression levels in nongynecological solid tumors: clinical and therapeutic implications," *Clinical and Experimental Metastasis*, vol. 21, no. 2, pp. 91–106, 2004.

[18] H. Rochefort and E. Liaudet-Coopman, "Cathepsin D in cancer metastasis: a protease and a ligand," *Acta Pathologica, Microbiologica. et Immunologica Scandinavica*, vol. 107, no. 1, pp. 86–95, 1999.

[19] E. Frohlich, B. Schalagenhauff, M. Mohrle et al., "Activity, expression, and transcription rate of the cathepsins B, D, H, and L in cutaneous malignant melanoma," *Cancer*, vol. 91, no. 5, pp. 972–982, 2001.

[20] I. Bartenjev, Z. Rudolf, B. Štabuc, I. Vrhovec, T. Perkovič, and A. Kansky, "Cathepsin D expression in early cutaneous malignant melanoma," *International Journal of Dermatology*, vol. 39, no. 8, pp. 599–602, 2000.

[21] F. J. Otto, T. Goldmann, B. Biess, A. Lippold, L. Suter, and U. Westhoff, "Prognostic classification of malignant melanomas by combining clinical, histological, and immunohistochemical parameters," *Oncology*, vol. 56, no. 3, pp. 208–214, 1999.

[22] T. Goldmann, L. Suter, D. Ribbert, and F. Otto, "The expression of proteolytic enzymes at the dermal invading front of primary cutaneous melanoma predicts metastasis," *Pathology Research and Practice*, vol. 195, no. 3, pp. 171–175, 1999.

[23] T. Kageshita, A. Yoshii, T. Kimura et al., "Biochemical and immunohistochemical analysis of cathepsins B, H, L and D in human melanocytic tumours," *Archives of Dermatological Research*, vol. 287, no. 3-4, pp. 266–272, 1995.

[24] O. L. Podhajcer, L. Bover, A. I. Bravo et al., "Expression of cathepsin D in primary and metastatic human melanoma and dysplastic nevi," *Journal of Investigative Dermatology*, vol. 104, no. 3, pp. 340–344, 1995.

[25] U. Westhoff, C. Fox, and F. J. Otto, "Quantification of cathepsin D in plasma of patients with malignant melanoma," *Anticancer Research*, vol. 18, no. 5 B, pp. 3785–3788, 1998.

[26] A. E. M. Mahdy, J. C. Cheng, J. Li et al., "Acid ceramidase upregulation in prostate cancer cells confers resistance to radiation: AC inhibition, a potential radiosensitizer," *Molecular Therapy*, vol. 17, no. 3, pp. 430–438, 2009.

[27] X. Liu, J. C. Cheng, L. S. Turner et al., "Acid ceramidase upregulation in prostate cancer: role in tumor development and implications for therapy," *Expert Opinion on Therapeutic Targets*, vol. 13, no. 12, pp. 1449–1458, 2009.

[28] E. Ruckhberle, U. Holtrich, K. Engels et al., "Acid ceramidase 1 expression correlates with a better prognosis in ER-positive breast cancer," *Climacteric*, vol. 12, no. 6, pp. 502–513, 2009.

[29] I. Maeda, T. Takano, F. Matsuzuka et al., "Rapid screening of specific changes in mRNA in thyroid carcinomas by sequence specific-differential display: decreased expression of acid ceramidase mRNA in malignant and benign thyroid tumors," *International Journal of Cancer*, vol. 81, no. 5, pp. 700–704, 1999.

[30] A. Haimovitz-Friedman, C. C. Kan, D. Ehleiter et al., "Ionizing radiation acts on cellular membranes to generate ceramide and initiate apoptosis," *Journal of Experimental Medicine*, vol. 180, no. 2, pp. 525–535, 1994.

[31] S. J. Chmura, E. Nodzenski, M. A. Beckett, D. W. Kufe, J. Quintans, and R. R. Weichselbaum, "Loss of ceramide production confers resistance to radiation-induced apoptosis," *Cancer Research*, vol. 57, no. 7, pp. 1270–1275, 1997.

[32] R. Kolesnick, "The therapeutic potential of modulating the ceramide/sphingomyelin pathway," *Journal of Clinical Investigation*, vol. 110, no. 1, pp. 3–8, 2002.

[33] A. Huwiler and J. Pfeilschifter, "Altering the sphingosine-1-phosphate/ceramide balance: a promising approach for tumor therapy," *Current Pharmaceutical Design*, vol. 12, no. 35, pp. 4625–4635, 2006.

[34] Y. H. Zeidan, R. W. Jenkins, J. B. Korman et al., "Molecular targeting of acid ceramidase: implications to cancer therapy," *Current Drug Targets*, vol. 9, no. 8, pp. 653–661, 2008.

[35] M. M. McCarthy, E. Pick, Y. Kluger et al., "HSP90 as a marker of progression in melanoma," *Annals of Oncology*, vol. 19, no. 3, pp. 590–594, 2008.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 601937, 10 pages
doi:10.1155/2011/601937

Review Article

Glycoproteomics-Based Identification of Cancer Biomarkers

Evelyn H. Kim[1] and David E. Misek[1, 2]

[1] *Department of Surgery, University of Michigan Medical School, Ann Arbor, MI 48109, USA*
[2] *Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109, USA*

Correspondence should be addressed to Evelyn H. Kim, hyeyk@umich.edu

Received 1 July 2011; Accepted 16 July 2011

Academic Editor: Tadashi Kondo

Copyright © 2011 E. H. Kim and D. E. Misek. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Protein glycosylation is one of the most common posttranslational modifications in mammalian cells. It is involved in many biological pathways and molecular functions and is well suited for proteomics-based disease investigations. Aberrant protein glycosylation may be associated with disease processes. Specific glycoforms of glycoproteins may serve as potential biomarkers for the early detection of disease or as biomarkers for the evaluation of therapeutic efficacy for treatment of cancer, diabetes, and other diseases. Recent technological developments, including lectin affinity chromatography and mass spectrometry, have provided researchers the ability to obtain detailed information concerning protein glycosylation. These in-depth investigations, including profiling and quantifying glycoprotein expression, as well as comprehensive glycan structural analyses may provide important information leading to the development of disease-related biomarkers. This paper describes methodologies for the detection of cancer-related glycoprotein and glycan structural alterations and briefly summarizes several current cancer-related findings.

1. Introduction

Within the past decade, proteomics has become an intensive field of research; one which may help to define biomarkers that could facilitate the early detection of disease or to provide important information for risk stratification, prediction of therapeutic efficacy, and disease prognosis. Proteins are known to be involved in biological activity and physiological changes in organisms [1]. Large-scale profiling of cellular proteins, using comparative expression levels between disease and normal homeostatic conditions, may reveal the basic underpinnings of disease processes. It may also facilitate the identification of proteins that are modified, either in structure or in levels of expression. Along with proteomic analysis of proteins, the analysis of protein posttranslational modifications (PTMs) also plays an important role in the study of disease. There are many types of PTMs, including acetylation, ubiquitination, phosphorylation, and glycosylation [2]. Each type of PTM may play a significant role in protein functionality. It is estimated that PTMs can be found on up to 80% of mammalian proteins [3]. Glycosylation is one of the most common PTMs, estimated to be found on over 50% of human proteins [4, 5].

Carbohydrate modifications are important in host-pathogen interactions, inflammation, development, and malignancy. Aberrant glycosylation may result in abnormal changes in biological function/activity, protein folding, and molecular recognition in disease. As such, analysis of altered cancer-related glycoprotein expression may facilitate discovery of potential biomarkers, as well as discovery of novel targets of therapeutics. Glycoproteins from various biological samples that are known to be cancer biomarkers are shown in Table 1. There are several different types of protein glycosylation, including (1) N-linked glycosylation, (2) O-linked glycosylation, (3) C-glycosylation [6], and (4) S-linked glycosylation (only found in bacteria) [7, 8]. N-glycosylation occurs on the asparagine in the sequence of Asn-X-Ser/Thr (and occasionally Cys) with X being any amino acid with the exception of proline. It is initiated on the cytoplasmic side of the rough endoplasmic reticulum (ER), where the oligosaccharide $Man_5GlcNAc_2$ is delivered to the precursor, dolichol pyrophosphate. The best known core glycan precursor is $Glc_3Man_9GlcNAc_2$-PP-dol [9–11]. N-glycans can be further categorized by the type and position of monosaccharide residues added to the core, being either

Table 1: List of some of the US Food and Drug Administration (FDA) approved glycoprotein cancer biomarkers. CA: cancer antigen, FDP: fibrin degradation protein, sPIgR: secreted chain of the polymeric immunoglobulin receptor.

Biomarker(a)	Glycosylation	Source	Disease
CA15.3	Yes	Serum	Breast
CA27-29	Yes	Serum	Breast cancer
HER2/NEU	Yes	Serum	Breast cancer
Fibrin/FDP	Yes	Urine	Bladder
CEA (carcinoembryonic antigen)	Yes	Serum	Colon cancer
Carcinoembryonic antigen (CEA)	Yes	Serum	Colon, breast, lung, pancreatic
Epidermal growth factor receptor	Yes	Tissue	Colon cancer
CA19-9	Yes	Serum	Gastrointestinal
KIT	Yes	Tissue	Gastrointestinal tumor
α-fetoprotein(AFP)	Yes	Serum	Hepatoma, testicular cancer
Human chorionic gonadotropin-β	Yes	Serum	Testicular cancer
Thyroglobulin	Yes	Serum	Thyroid cancer
CA125	Yes	Serum	Ovarian
PSA (prostate-specific antigen)	Yes	Serum	Prostate

a high-mannose type, an antennary complex type, or a hybrid type (Figure 1(a) [10]). The high-mannose type of N-glycan consists of mostly mannose in the core structure. The complex type of N-glycan contains N-acetylgalactosamine (Galβ1-3/4GlcNAc) in the N-glycan antennal region. The antennae can be further extended by adding Gal and GlcNAc residues. The hybrid type of N-glycan has both high mannose and N-acetylglucosamine. High-mannose and hybrid types share some similar features, such as two-mannose attachment on the trimannosyl core.

Serine or threonine residues can be O-glycosylated by addition of N-acetylglucosamine, mannose, fucose, glucose, N-acetylgalactosamine, or xylose (Figure 1(b)) [10, 11]. The most common O-linked glycosylation is initiated by N-acetylgalactosamine, bound through α-glycosidic linkages to Ser/Thr residues. In mucin-type O-glycosylation (mucins may be a cancer biomarker due to involvement in cancer development and influence in cell adhesion, invasion, and immune response [12]), the carbohydrate is linked to a hydroxyl group on Ser/Thr residues. This linkage often occurs while the protein is transiting through the Golgi apparatus as it is being secreted through the classical secretory pathway. The O-glycan core structure is formed by adding galactose/N-acetylglucosamine and may also contain sialic acid and/or fucose. C-glycosylation involves α-mannose C-linked to tryptophan. Still yet, many glycans have further modifications, such as sulfation and phosphorylation. The degree of complexity of glycan structures/composition changes, and unknown modification of glycans are beyond what can be adequately described in this paper. However, it is important to note that this complexity does contribute to cancer glycoproteomics.

2. Glycoproteomics Methodology

In general, glycoproteomics methodology consists of glycoprotein isolation, enrichment of the glycoproteins/glycopeptides, proteolytic digestion, and detection/identification of peptides or glycan structures using mass spectrometry-based techniques. Since many biological samples, such as plasma or serum, are very complex mixtures of proteins, extensive chromatographic separation techniques have been utilized (including ion exchange, size exclusion, hydrophobic interaction, and affinity chromatography) in order to reduce sample complexity and enhance dynamic ranges [13, 14]. Although glycoproteins can be separated by 2D-PAGE, their hydrophobic nature and tendency to precipitate at their isoelectric point, inadequate resolution, and the limited dynamic range of the gel system tend to greatly limit recovery and sequence coverage rates (by mass spectrometry). Following separation, the glycoproteins can be identified by MALDI-TOF-MS or LC-MS/MS. Recent technological improvements with LC-MS/MS have allowed mass spectrometry to play a major role in glycoproteomics analysis of disease. For glycoprotein analysis, the most commonly used methodology is a bottom-up approach [15] in which proteins are digested, after which the peptides go through an enrichment process or a deglycosylation process using Peptide N-Glycosidase F (PNGase F). PNGase F cleaves between the innermost GlcNAc and asparagine residues from N-linked glycoproteins/peptides, with the exception of ones carrying $\alpha 1 \rightarrow 3$ linked core fucose [16]. Following digestion, the peptides are subjected to MS analysis. With MS data, bioinformatics with various algorithms and glycol-related database are heavily relied upon to analyze glycoproteins and glycans [17, 18]. UniProt and PeptideAtlas libraries [19] also can provide information of glycoproteins and glycopeptide mass spectra. Using this methodology, it is necessary to integrate obtained data about the glycan and the glycoprotein [20]. Ultimately, however, it is challenging to find disease-related glycosylation changes due to the relatively low abundance of the altered glycan/glycoprotein structures.

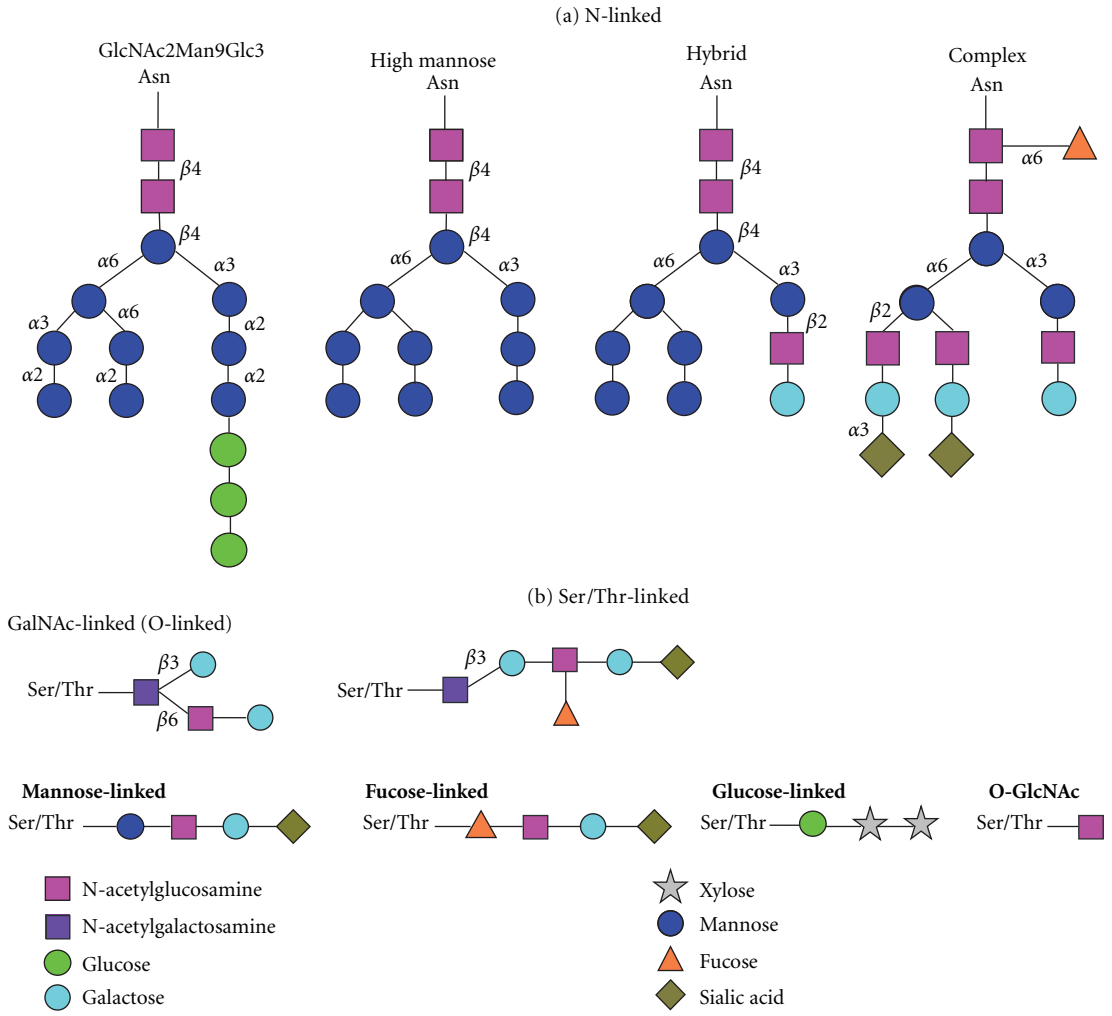


Figure 1: Common N- and O-linked glycans [10]. (a) Asparagine (N)-linked glycans, (b) serine/threonine (O)-linked glycans.

3. Glycoprotein Enrichment by Lectin Affinity Chromatography

Lectin affinity chromatographic enrichment is a routinely utilized methodology designed to concentrate glycoproteins/peptides that contain specific glycan structures, while eliminating nonspecific binding [21–24]. Various lectins (Table 2) can be used to isolate glycoproteins selectively based on glycan structure. As Concanavalin A (ConA) recognizes α-linked mannose residues, it will bind to high mannosyl, glycosyl, and hybrid-type glycans [25–27]. Wheat germ agglutinin (WGA) will bind to N-acetylglucosamine and possibly sialic acid residues on glycoproteins [28]. *Aleuria aurantia lectin* (AAL) recognizes specific binding to L-fucose-containing oligosaccharides [29]. Jacalin lectin (JAC) preferentially binds to galactosyl (β-1,3) N-acetylgalactosamine but will also bind to other O-glycosidically linked oligosaccharides (O-glycan) [30, 31]. In addition, there are many other lectins, each with their own carbohydrate binding specificity. Multiple agarose bound lectins can be used simultaneously/sequentially to purify/enrich different types of glycoproteins from various complex protein mixtures [32–34]. Utilizing multiple lectins provides the advantage of increasing detection coverage and providing global analysis. To date, however, lectin affinity chromatography has been more focused on the study of N-linked glycosylation, in part due to the fact that lectin affinity chromatography targets specific oligosaccharide structures and isolation/purification of O-linked glycan structures still in need of technological improvement.

For the enrichment of O-linked glycosylation glycoprotein/glycopeptides, it is more common to use β-elimination followed by Michael addition of DTT (BEMAD) or biotin-pentylamine, to label the O-glycosylation site (*O*-GlcNAc) [35]. Another method to isolate O-linked glycopeptides utilizes hydrophilic binding followed by multiple-staged MS/MS analysis [36, 37]. The detection of *O*-GlcNAc is challenging, since GlcNAcylation and phosphorylation are confined to a similar residue. One can also utilize galactosyltransferase to tag *O*-GlcNAc with ketone biotin [38]. Another extraction method, boronic acid-based beads, can provide fast, efficient, and specific enrichment of glycoproteins by binding to *cis*-diol groups on sugar residues [39, 40]. Hydrazide chemistry,

Table 2: A partial list of lectins commonly used for enrichment of glycoproteins/glycopeptides.

Lectin	Specificity
Aleuria Aurantia Lectin (AAL)	Fucα1-6 GlcNAc, Fucα1-3(Galβ1-4)GlcNAc
Concanavalin A (Con A)	High-Mannose, Manα1-6(Manα1-3)Man
Erythrina Cristagalli Lectin (ECA)	Galβ1-4GlcNAc
Hippeastrum Hybrid Lectin (HHL, AL)	High-Mannose, Manα1-3Man, Manα1-6Man
Jacalin	Galβ1-3GalNAc, GalNAc
Lens Culinaris Agglutinin (LCA)	Fucα1-6 GlcNAc, α-D-Glc, α-D-Man
Maackia Amurensis Lectin (MAL)	Siaα2-3Galβ1-4GlcNAc
Peanut Agglutinin (PNA)	Galβ1-3GalNAc
Phaseolus vulgaris Leucoagglutinin (PHA-L)	Tri/tetra-antennary complex-type N-glycan
Sambucus Nigra Lectin (SNA, EBL)	Siaα2-6Gal/GalNAc
Ulex Europaeus Agglutinin-I (UEA-I)	Fucα1-2Galβ1-4GlcNAc
Wheat Germ Agglutinin (WGA)	Chitin oligomers, Sia

interacting with glycoprotein carbonyl groups [41, 42], has also proved to be useful and is often combined with lectin affinity chromatography. Following the enrichment process, further purification with ethanol [43] or acetone [44] and a C18 stationary phase [45] or graphitized carbon column [46, 47] for glycans could be introduced prior to mass spectrometric analysis. For glycan analysis, lyophilization [48] or drying under nitrogen [49] is also useful, as increasing the temperature of the sample could cause destruction of the glycan. Further, solvent removal is very critical due to the low quantity of glycan for detection. In order to increase efficiency of isolation and detection of glycoproteins from complex protein lysates, the work flow of an online lectin microcolumn (ConA and SNA lectins) connected to LC-MS has been introduced [50]. Using silica-based columns instead of agarose-based columns, online selective concentration and detection of glycoproteins/glycopeptides gave shorter analysis time, reduced sample loss, and provided greater coverage uniformity.

4. Mass Spectrometry for Glycopeptides

The most widely used methods for glycomics involve characterization of glycopeptides generated by digestion and/or deglycosylation. Directly analyzing glycoproteins with attached glycans is complicated due to many fragment ions from backbone peptides, carbohydrates, and also ions from the mass spectrometry ion source. For mass analysis, there are several instruments including MALDI-TOF/TOF (time-of-flight)-MS [51–53], electrospray-based quadrupole ion trap (QIT) [54], quadrupole/TOF [55], Fourier transform ion cyclotron resonance (FTICR) [56], Orbitrap [48, 57] with CID (collision-induced dissociation), electron-capture dissociation (ECD) [58, 59], electron-transfer disassociation (ETD) [60, 61], and infrared multiphoton dissociation (IRMPD) [62, 63]. MALDI ionization generates stable singly charged precursor ions from the glycan. These precursor ions can be later characterized, using the MS/MS mode, by cleavage of glycosidic bonds and peptide with loss of glycan, leaving information regarding the glycan moiety. TOF/TOF fragmentation spectra could give additional attachment site information [64, 65] and structural analysis [66, 67]. QTOF mass spectrometers provide spectra with less chemical noise than spectra obtained by triple quadrupole or MALDI mass spectrometers. The advantages of using QTOF are higher mass accuracy, sensitivity, and resolution, thereby being able to detect ions with low intensity. Also, QTOF mass spectrometry could determine sensitive glycosylation site(s) and of the type of attached carbohydrate moiety [68]. FT-ICR-MS with ECD or IRMPD is very powerful tool for study of glycomics, since not only does FT-ICR-MS have high mass accuracy and high mass resolution, but also it has the ability to sequence peptides with no loss of glycans when it is equipped with ECD. It can also produce abundant fragment ions with IRMPD, resulting from dissociation at glycosidic linkages [57, 64].

5. Mass Spectrometry for Glycans

N-glycan release, resulting from cleavage with PNGase F, and O-glycan release by chemical methods [69] can be detected by mass spectrometry, although it is often necessary to adapt another step to improve ionization of the glycans, such as by permethylation [70, 71] or methylation [72, 73] derivatization [74, 75]. Figure 2 shows the nomenclature for tandem mass spectrometric product ions of glycans and glycoconjugated forms [76]. Neutral glycan produces gave singly and doubly charged ions and strong signals with $[M+Na]^+$, often along with $[M+K]^+$ ions. Also, cross-ring and c-type fragmentation could be generated with CID. In order to increase the glycan signal, stable anionic adducts can be generated by having unstable adducts react with chloride, bromide, iodide, nitrate, and phosphate, with the products analyzed in negative mode. Occasionally, glycan analysis in negative mode provides great advantage for a strong signal, and yet still contains structural information. Acidic glycan gives ions in higher charge states due to anionic groups. It will also fragment differently due to charge localization [77]. Glycans with different isoforms from glycoconjugates can be also detected with ion mobility, since ion mobility can differ

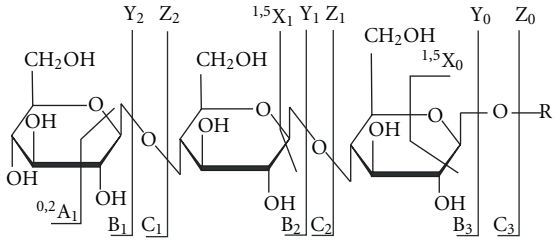


Figure 2: Nomenclature for tandem mass spectrometric product ions of glycans and glycoconjugated forms. Ions retaining the charge at the reducing terminus are named X, Y, and Z. Complementary ions are labeled A, B, and C [76].

based on molecular size and shape. Fragmentation of the glycan can have several factors, such as energy level applied to ion and the charge state of the ion. There are two major ions produced, the ion from the glycosidic cleavage between sugar ring and the ion from cross-ring cleavages. Both major ions can provide linkage information of glycan.

6. Cancer Glycoproteomics

Aberrant protein glycosylation may result due to genetic defects, cancer, and inflammation. These changes in protein glycosylation may result in abnormal changes in biological function/activity, protein folding, and molecular recognition in cancer. The site of protein glycosylation and the structure of the oligosaccharide could be altered during initiation or progression of disease. For example, oligosaccharides such as polysialic acid, sialic acid α2,6-linked to galactose, β1-6 branching, and extended lactoseries chains (antigens) have been known to be altered in cancer [78]. There have been glycoprotein biomarker studies with identification and profiling of glycoprotein/glycan to help early diagnosis and development of new therapeutics. Current advanced technologies provide enhanced ability to detect glycoproteins/glycans with increased dynamic range and lower detection limit of analytes from complex protein lysates such as plasma [79, 80], serum [81, 82], tissues [83, 84], and bodily fluids [85, 86]. These advanced technologies also facilitate analyte quantification using labeling [87] or label-free [88] methods. There are several studies for N-linked glycoproteins associated with cancer including prostate cancer [89], ovarian cancer [90], and breast cancer [91], as well as for O-linked glycoproteins associated with disease, including prostate cancer [92] and colon cancer [93]. Also, N- and O-linked glycans have been studied for cancer biomarkers [94, 95] with glycan mass profiling, since there is alteration in the branching and differential expression of glycoforms [96]. Analysis of individual glycans may be more suitable for biomarker studies, in spite of difficulty of study due to complexity of target itself, by using specific isoform and linkage information in the future.

7. Glycoproteomic Analysis of Prostate Cancer

Prostate cancer is the most common cancer in men in the United States, with an estimated 186,320 newly diagnosed cases and 28,660 deaths in 2008 [97]. Currently, the serum glycoprotein PSA is used clinically for mass population screening for prostate cancer. Unfortunately, however, assessment of PSA levels does not have the required specificity for a definitive prostate cancer diagnosis. This is due, in part, to the observed increase in expression of this protein in other prostatic pathologies such as benign prostatic hyperplasia (BPH) or prostatitis (prostate gland infection or inflammation). PSA also increases with age and infections of the prostate. Oligosaccharide profiling by mass spectrometry showed that PSA from prostate cancer sera has a higher content of α2,3-linked sialic acid than that from seminal fluid [98]. Lectin affinity column chromatography followed by the determination of total and free PSA by immunoassay has shown lower α2,6-linked sialic acid in serum free PSA from prostate cancer than that from BPH and higher α2,3-linked sialic acid in serum free PSA from cancer compared to BPH [99]. More recently, Meany and colleagues [100] demonstrated in a pooled sera study that α2,3-linked and α2,6-linked sialylation of PSA are more heterogeneous in cancer than in noncancer. Li and colleagues [101] monitored glycosylated and sialylated prostate-specific antigen (PSA) in prostate cancer and noncancerous tissues. They coupled a glycopeptide extraction strategy for specific glycosylation with selected reaction monitoring (SRM). Results of this study demonstrated that the relative abundance of glycosylated PSA isoforms were not correlated with total PSA protein levels measured in the same prostate cancer tissue samples by clinical immunoassay. Furthermore, the sialylated PSA was differentially distributed in cancer and noncancer tissues. These data suggest that differently glycosylated isoforms of glycoproteins can be quantitatively analyzed and may provide useful information for clinically relevant studies.

8. Glycoproteomics of Ovarian Cancer

Ovarian carcinoma is the leading cause of death from gynecological cancers in many Western countries. Machado et al. [90] analyzed N-linked glycans in the SKOV3 ovarian cancer cell line and on recombinant-secreted glycoprotein erythropoietin (EPO) expressed by transfected SKOV3 cells. The N-glycans were released using PNGase F and were then desalted and analyzed by high-performance anion exchange chromatography with pulsed amperometric detection and MALDI-TOF/TOF mass spectrometry. High-mannose type and fucosylated, complex type of glycans were found in the SKOV3 cancer cell line; predominant core-fucosylated structure glycans and partial LacdiNAc motif glycans were found on secreted recombinant human EPO. A large amount of N-acetylneuraminic acid in α-2,3-linkage was detected as were endogenous glycoproteins containing the LacdiNAc motif in N-linked glycans. The study suggests that high-mannose type glycans and LacdiNAc motif glycans may have a role as potential biomarkers for ovarian cancer. Abbott and colleagues [102] performed comparative glycotranscriptome analysis of ovarian cancer and normal ovarian tissues. Multiple lectins were utilized followed by nano-LC-linear ion trap

mass spectrometry. The identified proteins were verified by immunoprecipitation and lectin blot detection. The study showed 47 potential tumor-specific lectin reactive biomarkers for ovarian cancer; periostin and thrombospondin were presented as tumor-specific glycan changes that can be used to distinguish ovarian cancer patient serum from normal serum.

N- and O-linked glycans have also been analyzed as potential ovarian cancer biomarkers [94, 95] by glycan mass profiling, since there is alteration in the branching and differential expression [103]. Bereman and colleagues [48] reported plasma glycan profiling in 10 ovarian tumors, 10 controls with a differential diagnosis of benign gynecologic tumors, and 10 healthy controls using nano-HPLC-MS using reverse phase and amide-based stationary phase column under hydrophilic interaction. The experimental data suggests that amide-based stationary phase columns may have more robustness with high mass-measurement accuracy. Results of a comparison of glycan profiling between tumor, benign tumor and normal, two fucosylated glycans showed overexpression in healthy controls. Analysis of individual glycans may be more suitable for biomarker studies, using specific isoform and linkage information in the future.

9. Glycoproteomics of Breast Cancer

There is growing evidence that glycan structures on glycoproteins are modified in breast cancer [104–113]. Breast-cancer-associated alterations have been demonstrated for fucosylation groups and for sialylations on the plasma protein α-1-proteinase inhibitor [110]. Increased GlcNAc β1-6Man α1-6Man β-branching in asparagine-linked oligosaccharides has been observed in human tumor cells. The levels of the β1-6 branched oligosaccharides were evaluated in a series of benign and malignant human breast biopsies. Normal human breast tissue and benign lesions showed low expression, but 50% of the primary malignancies examined showed significantly elevated β1-6 branching [111]. Subsequently, L-PHA (a lectin that binds specifically to the β1-6 branched oligosaccharides) lectin histochemistry was performed on paraffin sections of human breast tissues. All breast carcinomas and epithelial hyperplasia with atypia demonstrated significantly increased L-PHA staining as compared to fibroadenomas and hyperplasia without atypia [112]. More recently, L-PHA reactive glycoproteins were identified from matched normal (nondiseased) and malignant tissue isolated from patients with invasive ductal breast carcinoma [113]. Comparison analysis of the data identified 34 proteins that were enriched by L-PHA fractionation in tumor relative to normal tissue for at least 2 cases of ductal invasive breast carcinoma. Of these 34 L-PHA tumor enriched proteins, 12 were common to all 4 matched cases analyzed.

Rudd and coworkers [114] analyzed fluorescently tagged serum *N*-glycans of advanced breast cancer patients using exoglycosidases and LC-MS/MS. They found that the expression of a trisialylated triantennary glycan containing an α-1,3-linked fucose was increased in the presence of breast cancer. Novotny and coworkers profiled the permethylated *N*-glycans in sera of breast cancer patients at different stages (stages I to IV) using MALDI TOF/TOF MS in one study [115]. In a second study, they profiled reduced and methylated serum *N*-glycans of late-stage breast cancer patients using nanoliquid chromatography (LC) Chip/time-of-flight (TOF) MS [116]. In both studies, they found an increase in fucosylation in both core and branched segments of *N*-glycans in the presence of breast cancer. In the latter study, they found a decrease in expression of a biantennary-monosialylated *N*-linked glycan and an increase in expression of a fucosylated triantennary-trisialylated *N*-linked glycan in the presence of stage IV breast cancer. These glycosylation changes in a tumor-secreted protein may reflect fundamental activity changes in the enzymes involved in the glycosylation pathway, either through altered levels of enzymes or altered enzymatic activity. Importantly, the changes in glycan structure may serve as early detection biomarkers of breast cancer.

10. Summary

Glycoproteomic analysis has become an important part of proteomics because not only does glycosylation reveal biological changes in terms of disease, but also it is possible to study glycoprotein/glycan in depth with the recent development of new mass spectrometry technology. For isolation and detection of glycoprotein/glycopeptides and glycans, an enrichment process using affinity chromatography followed by mass spectrometry with MS/MS can be used with quantitation. There is still massive, intense, manual data processing for studying glycan structure analysis, however, bioinformatics with various algorithms and glycol-related database has been developed to help analyze glycoprotein and glycan. The study of glycoproteomics is technically challenging yet has begun to produce promising results to identify biomarkers for early diagnosis and disease therapeutics.

Acknowledgments

This work was supported by award number W81XWH-09-1-0043 from the DoD (Army) and by award number CA140211 from the NCI/NIH. The information presented reflects the opinions of the authors only and does not reflect the opinion of the DoD or the NCI/NIH.

References

[1] M. R. Wilkins, J. C. Sanchez, A. A. Gooley et al., "Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it," *Biotechnology and Genetic Engineering Reviews*, vol. 13, pp. 41–50, 1996.

[2] R. Gudepu and F. Wold, in *Proteins: Analysis and Design*, R. H. Angeletti, Ed., pp. 121–207, Academics, San Diego, Calif, USA, 1998.

[3] R. Kornfeld and S. Kornfeld, "Assembly of asparagine-linked oligosaccharides," *Annual Review of Biochemistry*, vol. 54, pp. 631–664, 1985.

[4] R. Apweiler, H. Hermjakob, and N. Sharon, "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database," *Biochimica et Biophysica Acta*, vol. 1473, no. 1, pp. 4–8, 1999.

[5] C. H. Wong, "Protein glycosylation: new challenges and opportunities," *Journal of Organic Chemistry*, vol. 70, no. 11, pp. 4219–4225, 2005.

[6] X. Wei and L. Li, "Comparative glycoproteomics: approaches and applications," *Briefings in Functional Genomics and Proteomics*, vol. 8, no. 2, pp. 104–113, 2009.

[7] N. Floyd, B. Vijayakrishnan, J. R. Koeppe, and B. G. Davis, "Thiyl glycosylate of olefinic proteins: S-linked glycoconjugate synthesis," *Angewandte Chemie - International Edition*, vol. 48, no. 42, pp. 7798–7802, 2009.

[8] C. J. Lote and J. B. Weiss, "Identification in urine of a low-molecular-weight highly polar glycopeptide containing cysteinyl-galactose," *Biochemical Journal*, vol. 123, no. 4, p. 25, 1971.

[9] W. Morelle, K. Canis, F. Chirat, V. Faid, and J. C. Michalski, "The use of mass spectrometry for the proteomic analysis of glycosylation," *Proteomics*, vol. 6, no. 14, pp. 3993–4015, 2006.

[10] J. F. Rakus and L. K. Mahal, "New technologies for glycomic analysis: toward a systematic understanding of the glycome," *Annual Review of Analytical Chemistry*, vol. 4, pp. 367–392, 2011.

[11] K. Ohtsubo and J. D. Marth, "Glycosylation in cellular mechanisms of health and disease," *Cell*, vol. 126, no. 5, pp. 855–867, 2006.

[12] M. A. Hollingsworth and B. J. Swanson, "Mucins in cancer: protection and control of the cell surface," *Nature Reviews Cancer*, vol. 4, no. 1, pp. 45–60, 2004.

[13] A. K. Ottens, F. H. Kobeissy, R. A. Wolper et al., "A multidimensional differential proteomic platform using dual-phase ion-exchange chromatography-polyacrylamide gel electrophoresis/reversed-phase liquid chromatography tandem mass spectrometry," *Analytical Chemistry*, vol. 77, no. 15, pp. 4836–4845, 2005.

[14] H.-L. Huang, T. Stasyk, S. Morandell et al., "Enrichment of low-abundant serum proteins by albumin/immunoglobulin G immunoaffinity depletion under partly denaturing conditions," *Electrophoresis*, vol. 26, no. 14, pp. 2843–2849, 2005.

[15] R. Aebersold and M. Mann, "Mass spectrometry-based proteomics," *Nature*, vol. 422, no. 6928, pp. 198–207, 2003.

[16] T. Liu, W. J. Qian, M. A. Gritsenko et al., "Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry," *Journal of Proteome Research*, vol. 4, no. 6, pp. 2070–2080, 2005.

[17] K. F. Aoki-Kinoshita, "An introduction to bioinformatics for glycomics research," *PLoS Computational Biology*, vol. 4, no. 5, Article ID e1000075, 2008.

[18] H. Tang, Y. Mechref, and M. V. Novotny, "Automated interpretation of MS/MS spectra of oligosaccharides," *Bioinformatics*, vol. 21, no. 1, pp. i431–i439, 2005.

[19] E. W. Deutsch, H. Lam, and R. Aebersold, "PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows," *EMBO Reports*, vol. 9, no. 5, pp. 429–434, 2008.

[20] R. Uematsu, J. I. Furukawa, H. Nakagawa et al., "High throughput quantitative glycomics and glycoform-focused proteomics of murine dermis and epidermis," *Molecular and Cellular Proteomics*, vol. 4, no. 12, pp. 1977–1989, 2005.

[21] R. D. Cummings, "Use of lectins in analysis of glycoconjugates," *Methods in Enzymology*, vol. 230, pp. 66–86, 1994.

[22] K. Taketa, "Characterization of sugar chain structures of human α-fetoprotein by lectin affinity electrophoresis," *Electrophoresis*, vol. 19, no. 15, pp. 2595–2602, 1998.

[23] T. Endo, "Fractionation of glycoprotein-derived oligosaccharides by affinity chromatography using immobilized lectin columns," *Journal of Chromatography A*, vol. 720, no. 1-2, pp. 251–261, 1996.

[24] M. Saleemuddin and Q. Husain, "Concanavalin A: a useful ligand for glycoenzyme immobilization—a review," *Enzyme and Microbial Technology*, vol. 13, no. 4, pp. 290–295, 1991.

[25] Y. Ohyama, K. I. Kasai, H. Nomoto, and Y. Inoue, "Frontal affinity chromatography of ovalbumin glycoasparagines on a concanavalin A-Sepharose column. A quantitative study of the binding specificity of the lectin," *Journal of Biological Chemistry*, vol. 260, no. 11, pp. 6882–6887, 1985.

[26] K. Kornfeld, M. L. Reitman, and R. Kornfeld, "The carbohydrate-binding specificity of pea and lentil lectins. Fucose is an important determinant," *Journal of Biological Chemistry*, vol. 256, no. 13, pp. 6633–6640, 1981.

[27] J. W. Becker, G. N. Reeke, and J. L. Wang, "The covalent and three dimensional structure of concanavalin A. III. Structure of the monomer and its interactions with metals and saccharides," *Journal of Biological Chemistry*, vol. 250, no. 4, pp. 1513–1524, 1975.

[28] N. Bakry, Y. Kamata, and L. L. Simpson, "Lectins from triticum vulgaris and limax flavus are universal antagonists of botulinum neurotoxin and tetanus toxin," *Journal of Pharmacology and Experimental Therapeutics*, vol. 258, no. 3, pp. 830–836, 1991.

[29] N. Kochibe and K. Furukawa, "Purification and properties of a novel fucose-specific hemagglutinin of Aleuria aurantia," *Biochemistry*, vol. 19, no. 13, pp. 2841–2846, 1980.

[30] M. C. Roque-Barreira and A. Campos-Neto, "Jacalin: an IgA-binding lectin," *Journal of Immunology*, vol. 134, no. 3, pp. 1740–1743, 1985.

[31] R. L. Gregory, J. Rundegren, and R. R. Arnold, "Separation of human IgA1 and IgA2 using jacalin-agarose chromatography," *Journal of Immunological Methods*, vol. 99, no. 1, pp. 101–106, 1987.

[32] R. D. Cummings and S. Kornfeld, "Fractionation of asparagine-linked oligosaccharides by serial lectin-agarose affinity chromatography. A rapid, sensitive, and specific technique," *Journal of Biological Chemistry*, vol. 257, no. 19, pp. 11235–11240, 1982.

[33] M. Madera, Y. Mechref, I. Klouckova, and M. V. Novotny, "High-sensitivity profiling of glycoproteins from human blood serum through multiple-lectin affinity chromatography and liquid chromatography/tandem mass spectrometry," *Journal of Chromatography B*, vol. 845, no. 1, pp. 121–137, 2007.

[34] T. Plavina, E. Wakshull, W. S. Hancock, and M. Hincapie, "Combination of abundant protein depletion and Multi-Lectin Affinity Chromatography (M-LAC) for plasma protein biomarker discovery," *Journal of Proteome Research*, vol. 6, no. 2, pp. 662–671, 2007.

[35] L. Wells, K. Vosseller, R. N. Cole, J. M. Cronshaw, M. J. Matunis, and G. W. Hart, "Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications," *Molecular & Cellular Proteomics*, vol. 1, no. 10, pp. 791–804, 2002.

[36] Y. Wada, M. Tajiri, and S. Yoshida, "Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics," *Analytical Chemistry*, vol. 76, no. 22, pp. 6560–6565, 2004.

[37] C. D. Calvano, C. G. Zambonin, and O. N. Jensen, "Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry," *Journal of Proteomics*, vol. 71, no. 3, pp. 304–317, 2008.

[38] N. Khidekel, S. Arndt, N. Lamarre-Vincent et al., "A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications," *Journal of the American Chemical Society*, vol. 125, no. 52, pp. 16162–16163, 2003.

[39] K. Sparbier, S. Koch, I. Kessler, T. Wenzel, and M. Kostrzewa, "Selective isolation of glycoproteins and glycopeptides for MALDI-TOF MS detection supported by magnetic particles," *Journal of Biomolecular Techniques*, vol. 16, no. 4, pp. 407–413, 2005.

[40] L. Liang and Z. Liu, "A self-assembled molecular team of boronic acids at the gold surface for specific capture of cis-diol biomolecules at neutral pH," *Chemical Communications*, vol. 47, no. 8, pp. 2255–2257, 2011.

[41] D. J. O'Shannessy and W. L. Hoffman, "Site-directed immobilization of glycoproteins on hydrazide-containing solid supports," *Biotechnology and Applied Biochemistry*, vol. 9, no. 6, pp. 488–496, 1987.

[42] H. Zhang, X. J. Li, D. B. Martin, and R. Aebersold, "Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry," *Nature Biotechnology*, vol. 21, no. 6, pp. 660–666, 2003.

[43] C. Kirmiz, B. Li, H. J. An et al., "A serum glycomics approach to breast cancer biomarkers," *Molecular and Cellular Proteomics*, vol. 6, no. 1, pp. 43–55, 2007.

[44] C. E. Costello, J. M. Contado-Miller, and J. F. Cipollo, "A glycomics platform for the analysis of permethylated oligosaccharide alditols," *Journal of the American Society for Mass Spectrometry*, vol. 18, no. 10, pp. 1799–1812, 2007.

[45] S. Kamoda, M. Nakano, R. Ishikawa, S. Suzuki, and K. Kakehi, "Rapid and sensitive screening of N-glycans as 9-fluorenylmethyl derivatives by high-performance liquid chromatography: a method which can recover free oligosaccharides after analysis," *Journal of Proteome Research*, vol. 4, no. 1, pp. 146–152, 2005.

[46] C. Bleckmann, H. Geyer, V. Reinhold et al., "Glycomic analysis of N-linked carbohydrate epitopes from CD24 of mouse brain," *Journal of Proteome Research*, vol. 8, no. 2, pp. 567–582, 2009.

[47] Y. G. Kim, K. S. Jang, H. S. Joo, H. K. Kim, C. S. Lee, and B. G. Kim, "Simultaneous profiling of N-glycans and proteins from human serum using a parallel-column system directly coupled to mass spectrometry," *Journal of Chromatography B*, vol. 850, no. 1-2, pp. 109–119, 2007.

[48] M. S. Bereman, T. I. Williams, and D. C. Muddiman, "Development of a nanolc ltq orbitrap mass spectrometric method for profiling glycans derived from plasma from healthy, benign tumor control, and epithelial ovarian cancer patients," *Analytical Chemistry*, vol. 81, no. 3, pp. 1130–1136, 2009.

[49] Y. Huang, Y. Mechref, and M. V. Novotny, "Microscale nonreductive release of O-linked glycans for subsequent analysis through MALDI mass spectrometry and capillary electrophoresis," *Analytical Chemistry*, vol. 73, no. 24, pp. 6063–6069, 2001.

[50] M. Madera, Y. Mechref, and M. V. Novotny, "Combining lectin microcolumns with high-resolution separation techniques for enrichment of glycoproteins and glycopeptides," *Analytical Chemistry*, vol. 77, no. 13, pp. 4081–4090, 2005.

[51] M. Wuhrer, C. H. Hokke, and A. M. Deelder, "Glycopeptide analysis by matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry reveals novel features of horseradish peroxidase glycosylation," *Rapid Communications in Mass Spectrometry*, vol. 18, no. 15, pp. 1741–1748, 2004.

[52] K. Kubota, Y. Sato, Y. Suzuki et al., "Analysis of glycopeptides using lectin affinity chromatography with MALDI-TOF mass spectrometry," *Analytical Chemistry*, vol. 80, no. 10, pp. 3693–3698, 2008.

[53] K. Sparbier, A. Asperger, A. Resemann et al., "Analysis of glycoproteins in human serum by means of glycospecific magnetic bead separation and LC-MALDI-TOF/TOF analysis with automated glycopeptide detection," *Journal of Biomolecular Techniques*, vol. 18, no. 4, pp. 252–258, 2007.

[54] K. Sandra, B. Devreese, J. Van Beeumen, I. Stals, and M. Claeyssens, "The Q-Trap mass spectrometer, a novel tool in the study of protein glycosylation," *Journal of the American Society for Mass Spectrometry*, vol. 15, no. 3, pp. 413–423, 2004.

[55] A. Harazono, N. Kawasaki, S. Itoh et al., "Site-specific N-glycosylation analysis of human plasma ceruloplasmin using liquid chromatography with electrospray ionization tandem mass spectrometry," *Analytical Biochemistry*, vol. 348, no. 2, pp. 259–268, 2006.

[56] K. Håkansson, M. J. Chalmers, J. P. Quinn, M. A. McFarland, C. L. Hendrickson, and A. G. Marshall, "Combined electron capture and infrared multiphoton dissociation for multistage MS/MS in a Fourier transform ion cyclotron resonance mass spectrometer," *Analytical Chemistry*, vol. 75, no. 13, pp. 3256–3262, 2003.

[57] Z. M. Segu and Y. Mechref, "Characterizing protein glycosylation sites through higher-energy C-trap dissociation," *Rapid Communications in Mass Spectrometry*, vol. 24, no. 9, pp. 1217–1225, 2010.

[58] F. Kjeldsen, K. F. Haselmann, B. A. Budnik, E. S. Sørensen, and R. A. Zubarev, "Complete characterization of post-translational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage," *Analytical Chemistry*, vol. 75, no. 10, pp. 2355–2361, 2003.

[59] M. Mormann, H. Paulsen, and J. Peter-Katalinić, "Electron capture dissociation of O-glycosylated peptides: radical site-induced fragmentation of glycosidic bonds," *European Journal of Mass Spectrometry*, vol. 11, no. 5, pp. 497–511, 2005.

[60] Z. Darula and K. F. Medzihradszky, "Affinity enrichment and characterization of mucin core-1 type glycopeptides from bovine serum," *Molecular and Cellular Proteomics*, vol. 8, no. 11, pp. 2515–2526, 2009.

[61] M. I. Catalina, C. A. M. Koeleman, A. M. Deelder, and M. Wuhrer, "Electron transfer dissociation of N-glycopeptides: loss of the entire N-glycosylated asparagine side chain," *Rapid Communications in Mass Spectrometry*, vol. 21, no. 6, pp. 1053–1061, 2007.

[62] K. Håkansson, H. J. Cooper, M. R. Emmett, C. E. Costello, A. G. Marshall, and C. L. Nilsson, "Electron capture dissociation and infrared multiphoton dissociation MS/MS of an N-glycosylated tryptic peptide to yield complementary sequence information," *Analytical Chemistry*, vol. 73, no. 18, pp. 4530–4536, 2001.

[63] L. Bindila, K. Steiner, C. Schäffer, P. Messner, M. Mormann, and J. Peter-Katalinić, "Sequencing of O-glycopeptides derived from an S-layer glycoprotein of Geobacillus stearothermophilus NRS 2004/3a containing up to 51 monosaccharide

residues at a single glycosylation site by fourier transform ion cyclotron resonance infrared multiphoton dissociation mass spectrometry," *Analytical Chemistry*, vol. 79, no. 9, pp. 3271–3279, 2007.

[64] M. Kurogochin and S. I. Nishimura, "Structural characterization of N-glycopeptides by matrix-dependent selective fragmentation of MALDI-TOF/TOF tandem mass spectrometry," *Analytical Chemistry*, vol. 76, no. 20, pp. 6097–6101, 2004.

[65] N. Takemori, N. Komori, and H. Matsumoto, "Highly sensitive multistage mass spectrometry enables small-scale analysis of protein glycosylation from two-dimensional polyacrylamide gels," *Electrophoresis*, vol. 27, no. 7, pp. 1394–1406, 2006.

[66] E. Stephens, S. L. Maslen, L. G. Green, and D. H. Williams, "Fragmentation characteristics of neutral N-linked glycans using a MALDI-TOF/TOF tandem mass spectrometer," *Analytical Chemistry*, vol. 76, no. 8, pp. 2343–2354, 2004.

[67] W. Morelle, M. C. Slomianny, H. Diemer, C. Schaeffer, A. Van Dorsselaer, and J. C. Michalski, "Structural characterization of 2-aminobenzamide-derivatized oligosaccharides using a matrix-assisted laser desorption/ionization two-stage time-of-flight tandem mass spectrometer," *Rapid Communications in Mass Spectrometry*, vol. 19, no. 14, pp. 2075–2084, 2005.

[68] B. Maček, J. Hofsteenge, and J. Peter-Katalinić, "Direct determination of glycosylation sites in O-fucosylated glycopeptides using nano-electrospray quadrupole time-of-flight mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 15, no. 10, pp. 771–777, 2001.

[69] D. M. Carlson, "Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins," *Journal of Biological Chemistry*, vol. 243, no. 3, pp. 616–626, 1968.

[70] C. E. Costello, J. M. Contado-Miller, and J. F. Cipollo, "A glycomics platform for the analysis of permethylated oligosaccharide alditols," *Journal of the American Society for Mass Spectrometry*, vol. 18, no. 10, pp. 1799–1812, 2007.

[71] J. C. Botelho, J. A. Atwood, L. Cheng, G. Alvarez-Manilla, W. S. York, and R. Orlando, "Quantification by isobaric labeling (QUIBL) for the comparative glycomic study of O-linked glycans," *International Journal of Mass Spectrometry*, vol. 278, no. 2-3, pp. 137–142, 2008.

[72] I. Schabussova, H. Amer, I. van Die, P. Kosma, and R. M. Maizels, "O-Methylated glycans from Toxocara are specific targets for antibody binding in human and animal infections," *International Journal for Parasitology*, vol. 37, no. 1, pp. 97–109, 2007.

[73] I. Ciucanu and F. Kerek, "A simple and rapid method for the permethylation of carbohydrates," *Carbohydrate Research*, vol. 131, no. 2, pp. 209–217, 1984.

[74] S. H. Walker, L. M. Lilley, M. F. Enamorado, D. L. Comins, and D. C. Muddiman, "Hydrophobic derivatization of N-linked glycans for increased ion abundance in electrospray ionization mass spectrometry," *Journal of The American Society for Mass Spectrometry*, vol. 22, no. 8, pp. 1309–1317, 2011.

[75] M. Nakano, D. Higo, E. Arai et al., "Capillary electrophoresis-electrospray ionization mass spectrometry for rapid and sensitive N-glycan analysis of glycoproteins as 9-fluorenylmethyl derivatives," *Glycobiology*, vol. 19, no. 2, pp. 135–143, 2009.

[76] B. Domon and C. E. Costello, "A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates," *Glycoconjugate Journal*, vol. 5, no. 4, pp. 397–409, 1988.

[77] D. J. Harvey, "Fragmentation of negative ions from carbohydrates: Part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates," *Journal of the American Society for Mass Spectrometry*, vol. 16, no. 5, pp. 622–630, 2005.

[78] F. DairOlio, "Protein glycosylation in cancer biology: an overview," *Journal of Clinical Pathology: Clinical Molecular Pathology*, vol. 49, no. 3, pp. M126–M135, 1996.

[79] J. Rehman and M. A. Rahman, "Studies on glycosylated plasma proteins in diabetic patients," *Journal of the Pakistan Medical Association*, vol. 41, no. 1, pp. 16–18, 1991.

[80] Y. Zhou, R. Aebersold, and H. Zhang, "Isolation of N-linked glycopeptides from plasma," *Analytical Chemistry*, vol. 79, no. 15, pp. 5826–5837, 2007.

[81] R. Saldova, M. R. Wormald, R. A. Dwek, and P. M. Rudd, "Glycosylation changes on serum glycoproteins in ovarian cancer may contribute to disease pathogenesis," *Disease Markers*, vol. 25, no. 4-5, pp. 219–232, 2008.

[82] A. Klein, Y. Carre, A. Louvet, J. C. Michalski, and W. Morelle, "Immunoglobulins are the major glycoproteins involved in the modifications of total serum N-glycome in cirrhotic patients," *Proteomics: Clinical Applications*, vol. 4, no. 4, pp. 379–393, 2010.

[83] R. B. Parekh, A. G. Tse, R. A. Dwek, A. F. Williams, and T. W. Rademacher, "Tissue-specific N-glycosylation, site-specific oligosaccharide patterns and lentil lectin recognition of rat Thy-1," *EMBO Journal*, vol. 6, no. 5, pp. 1233–1244, 1987.

[84] R. Chen, X. Jiang, D. Sun et al., "Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry," *Journal of Proteome Research*, vol. 8, no. 2, pp. 651–661, 2009.

[85] S. Pan, Y. Wang, J. F. Quinn et al., "Identification of glycoproteins in human cerebrospinal fluid with a complementary proteomic approach," *Journal of Proteome Research*, vol. 5, no. 10, pp. 2769–2779, 2006.

[86] S. Y. Vakhrushev, M. Mormann, and J. Peter-Katalinić, "Identification of glycoconjugates in the urine of a patient with congenital disorder of glycosylation by high-resolution mass spectrometry," *Proteomics*, vol. 6, no. 3, pp. 983–992, 2006.

[87] P. L. Ross, Y. N. Huang, J. N. Marchese et al., "Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents," *Molecular and Cellular Proteomics*, vol. 3, no. 12, pp. 1154–1169, 2004.

[88] K. R. Rebecchi, J. L. Wenke, E. P. Go, and H. Desaire, "Label-free quantitation: a new glycoproteomics approach," *Journal of the American Society for Mass Spectrometry*, vol. 20, no. 6, pp. 1048–1059, 2009.

[89] R. E. Reiter, Z. Gu, T. Watabe et al., "Prostate stem cell antigen: a cell surface marker overexpressed in prostate cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 4, pp. 1735–1740, 1998.

[90] E. MacHado, S. Kandzia, R. Carilho, P. Altevogt, H. S. Conradt, and J. Costa, "N-Glycosylation of total cellular glycoproteins from the human ovarian carcinoma SKOV3 cell line and of recombinantly expressed human erythropoietin," *Glycobiology*, vol. 21, no. 3, pp. 376–386, 2011.

[91] C. I. Orazine, M. Hincapie, W. S. Hancock, M. Hattersley, and J. H. Hanke, "A proteomic analysis of the plasma glycoproteins of a MCF-7 mouse xenograft: a model system for the detection of tumor markers," *Journal of Proteome Research*, vol. 7, no. 4, pp. 1542–1554, 2008.

[92] P. Premaratne, K. Welén, J. E. Damber, G. C. Hansson, and M. Bäckström, "O-glycosylation of MUC1 mucin in prostate

cancer and the effects of its expression on tumor growth in a prostate cancer xenograft model," *Tumor Biology*, vol. 32, pp. 203–213, 2010.

[93] I. Brockhausen, J. Yang, M. Lehotay, S. Ogata, and S. Itzkowitz, "Pathways of mucin O-glycosylation in normal and malignant rat colonic epithelial cells reveal a mechanism for cancer-associated Sialyl-Tn antigen expression," *Biological Chemistry*, vol. 382, no. 2, pp. 219–232, 2001.

[94] D. E. Misek, T. H. Patwa, D. M. Lubman, and D. M. Simeone, "Early detection and biomarkers in pancreatic cancer," *Journal of the National Comprehensive Cancer Network*, vol. 5, no. 10, pp. 1034–1041, 2007.

[95] H. J. An, A. H. Franz, and C. B. Lebrilla, "Improved capillary electrophoretic separation and mass spectrometric detection of oligosaccharides," *Journal of Chromatography A*, vol. 1004, no. 1-2, pp. 121–129, 2003.

[96] G. Mondal, U. Chatterjee, Y. K. Chawla, and B. P. Chatterjee, "Alterations of glycan branching and differential expression of sialic acid on alpha fetoprotein among hepatitis patients," *Glycoconjugate Journal*, vol. 28, no. 1, pp. 1–9, 2011.

[97] A. Jemal, R. Siegel, E. Ward et al., "Cancer statistics, 2008," *CA Cancer Journal for Clinicians*, vol. 58, no. 2, pp. 71–96, 2008.

[98] R. Peracaula, G. Tabarés, L. Royle et al., "Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins," *Glycobiology*, vol. 13, no. 6, pp. 457–470, 2003.

[99] C. Ohyama, M. Hosono, K. Nitta et al., "Carbohydrate structure and differential binding of prostate specific antigen to Maackia amurensis lectin between prostate cancer and benign prostate hypertrophy," *Glycobiology*, vol. 14, no. 8, pp. 671–679, 2004.

[100] D. L. Meany, Z. Zhang, L. J. Sokoll, H. Zhang, and D. W. Chan, "Glycoproteomics for prostate cancer detection: changes in serum PSA glycosylation patterns," *Journal of Proteome Research*, vol. 8, no. 2, pp. 613–619, 2009.

[101] Y. Li, Y. Tian, T. Rezai et al., "Simultaneous analysis of glycosylated and sialylated prostate-specific antigen revealing differential distribution of glycosylated prostate-specific antigen isoforms in prostate cancer tissues," *Analytical Chemistry*, vol. 83, no. 1, pp. 240–245, 2011.

[102] K. L. Abbott, J. M. Lim, L. Wells, B. B. Benigno, J. F. McDonald, and M. Pierce, "Identification of candidate biomarkers with cancerspecific glycosylation in the tissue and serum of endometrioid ovarian cancer patients by glycoproteomic analysis," *Proteomics*, vol. 10, no. 3, pp. 470–481, 2010.

[103] G. Mondal, U. Chatterjee, Y. K. Chawla, and B. P. Chatterjee, "Alterations of glycan branching and differential expression of sialic acid on alpha fetoprotein among hepatitis patients," *Glycoconjugate Journal*, vol. 28, no. 1, pp. 1–9, 2011.

[104] D. F. Hayes, M. Abe, J. Siddiqui, C. Tondini, and D. W. Kufe, "Clinical and molecular investigations of the DF3 breast cancer-associated antigen," in *Immunological Approaches to the Diagnosis and Therapy of Breast Cancer II*, R. L. Ceriani, Ed., pp. 45–53, Plenum Press, New York, NY, USA, 1989.

[105] S. R. Hull, A. Bright, K. L. Carraway, M. Abe, D. F. Hayes, and D. W. Kufe, "Oligosaccharide differences in the DF3 sialomucin antigen from normal human milk and the BT-20 human breast carcinoma cell line," *Cancer Communications*, vol. 1, no. 4, pp. 261–267, 1989.

[106] L. Perey, D. F. Hayes, and D. Kufe, "Effects of differentiating agents on cell surface expression of the breast carcinoma-associated DF3-P epitope," *Cancer Research*, vol. 52, no. 22, pp. 6365–6370, 1992.

[107] L. Perey, D. F. Hayes, P. Maimonis, M. Abe, C. O'Hara, and D. W. Kufe, "Tumor selective reactivity of a monoclonal antibody prepared against a recombinant peptide derived from the DF3 human breast carcinoma-associated antigen," *Cancer Research*, vol. 52, no. 9, pp. 2563–2568, 1992.

[108] R. Sewell, M. Bäckström, M. Dalziel et al., "The ST6GalNAc-I sialyltransferase localizes throughout the golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer," *Journal of Biological Chemistry*, vol. 281, no. 6, pp. 3586–3594, 2006.

[109] J. M. Burchell, A. Mungul, and J. Taylor-Papadimitriou, "O-linked glycosylation in the mammary gland: changes that occur during malignancy," *Journal of Mammary Gland Biology and Neoplasia*, vol. 6, no. 3, pp. 355–364, 2001.

[110] M. T. Goodarzi and G. A. Turner, "Decreased branching, increased fucosylation and changed sialylation of alpha-1-proteinase inhibitor in breast and ovarian cancer," *Clinica Chimica Acta*, vol. 236, no. 2, pp. 161–171, 1995.

[111] J. W. Dennis and S. Laferte, "Oncodevelopmental expression of -GlcNAcβ1-6Manα1-6Manβ1-branched asparagine-linked oligosaccharides in murine tissues and human breast carcinomas," *Cancer Research*, vol. 49, no. 4, pp. 945–950, 1989.

[112] B. Fernandes, U. Sagman, M. Auger, M. Demetrio, and J. W. Dennis, "β1-6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia," *Cancer Research*, vol. 51, no. 2, pp. 718–723, 1991.

[113] K. L. Abbott, K. Aoki, J. M. Lim et al., "Targeted glycoproteomic identification of biomarkers for human breast carcinoma," *Journal of Proteome Research*, vol. 7, no. 4, pp. 1470–1480, 2008.

[114] U. M. Abd Hamid, L. Royle, R. Saldova et al., "A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression," *Glycobiology*, vol. 18, no. 12, pp. 1105–1118, 2008.

[115] Z. Kyselova, Y. Mechref, P. Kang et al., "Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles," *Clinical Chemistry*, vol. 54, no. 7, pp. 1166–1175, 2008.

[116] W. R. Alley, M. Madera, Y. Mechref, and M. V. Novotny, "Chip-based reversed-phase liquid chromatography-mass spectrometry of permethylated N-linked glycans: a potential methodology for cancer-biomarker discovery," *Analytical Chemistry*, vol. 82, no. 12, pp. 5095–5106, 2010.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 214715, 18 pages
doi:10.1155/2011/214715

Research Article

Urine Glycoprotein Profile Reveals Novel Markers for Chronic Kidney Disease

Anuradha Vivekanandan-Giri,[1] Jessica L. Slocum,[1] Carolyn L. Buller,[1] Venkatesha Basrur,[2] Wenjun Ju,[1] Rodica Pop-Busui,[3] David M. Lubman,[2,4,5] Matthias Kretzler,[1,4] and Subramaniam Pennathur[1,4]

[1] *Division of Nephrology, University of Michigan, Ann Arbor, MI 48105, USA*
[2] *Department of Pathology, University of Michigan, Ann Arbor, MI 48105, USA*
[3] *Division of Metabolism, Endocrinology and Diabetes, Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48105, USA*
[4] *Department of Computational Medicine and Biology, University of Michigan, Ann Arbor, MI 48105, USA*
[5] *Department of Surgery, University of Michigan, Ann Arbor, MI 48105, USA*

Correspondence should be addressed to Subramaniam Pennathur, spennath@umich.edu

Received 30 June 2011; Accepted 30 July 2011

Academic Editor: David E. Misek

Copyright © 2011 Anuradha Vivekanandan-Giri et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chronic kidney disease (CKD) is a significant public health problem, and progression to end-stage renal disease leads to dramatic increases in morbidity and mortality. The mechanisms underlying progression of disease are poorly defined, and current noninvasive markers incompletely correlate with disease progression. Therefore, there is a great need for discovering novel markers for CKD. We utilized a glycoproteomic profiling approach to test the hypothesis that the urinary glycoproteome profile from subjects with CKD would be distinct from healthy controls. N-linked glycoproteins were isolated and enriched from the urine of healthy controls and subjects with CKD. This strategy identified several differentially expressed proteins in CKD, including a diverse array of proteins with endopeptidase inhibitor activity, protein binding functions, and acute-phase/immune-stress response activity supporting the proposal that inflammation may play a central role in CKD. Additionally, several of these proteins have been previously linked to kidney disease implicating a mechanistic role in disease pathogenesis. Collectively, our observations suggest that the human urinary glycoproteome may serve as a discovery source for novel mechanism-based biomarkers of CKD.

1. Introduction

Chronic kidney disease (CKD) affects approximately 11% of the US population with over 100,000 individuals progressing to end-stage renal disease (ESRD) annually [1, 2]. Despite this significant and growing public health problem, it remains difficult to predict which individuals will progress to ESRD. As ESRD carries a substantial increase in morbidity and mortality, it is critical to identify this high-risk patient population that would most benefit from early and aggressive therapy.

Current strategies for predicting CKD progression are limited. Pathologic examination of renal tissue provides valuable information on degree of interstitial fibrosis and predilection for ESRD. However, renal biopsy is invasive with a limited role for longitudinal followup. Quantitative measures of proteinuria have long been used as noninvasive markers of CKD progression [3], yet these largely albumin-based methods detect nonselective proteinuria and incompletely correlate with disease. With recent advances in high through-put technology and mass spectrometry techniques, urine proteomic investigation is an attractive tool in the pursuit for noninvasive and specific markers of CKD progression [4, 5].

Numerous investigators have successfully applied broad-scale urine proteomic strategies to kidney disease. The urine

Table 1: Patient characteristics of study subjects.

Variable	Healthy control ($n = 6$)	CKD ($n = 6$)	P
Age (years)	46.3 (13.5)	47.2 (14.2)	0.92
Sex (male/female)	2/4	2/4	1.00
Body mass index (kg/m^2)	24.3 (3.0)	30.5 (4.8)	0.02
Serum creatinine (mg/dL)	0.85 (0.16)	1.75 (1.09)	0.07
eGFR (mL/min)*	83.0 (15.0)	52.0 (27.4)	0.05
Protein/creatinine ratio	0.03 (0.02)	2.15 (1.44)	0.01

All data expressed as mean ± SD.
eGFR estimated glomerular filtration rate.
*eGFR calculated from Modification of Diet in Renal Disease formula.

proteome predicts nephropathy and decline in renal function in diabetic subjects [6, 7]. It also correlates with early changes of focal segmental nephrosclerosis [8], can identify IgA nephropathy and renal allograft rejection [9, 10], and predicts treatment response and disease activity in nephrotic syndrome and lupus nephritis [11, 12]. Despite these advances, analysis of the entire urine proteome is particularly difficult in CKD. With disruption of the glomerular filtration barrier and leakage of abundant plasma proteins into the urine, a nonselective, largely albumin predominant, pattern often results [13]. To overcome this, methods to increase the detection of low-abundance proteins have been developed to provide disease specificity and clinical relevance of urine profiling and to mechanistically understand factors influencing disease progression.

Glycoprotein enrichment techniques allow depletion of albumin and other abundant plasma proteins while providing a more thorough analysis of a subfraction of the urine proteome. As glycosylated proteins are critical for cellular interactions and signaling cascades, disease states are likely to cause early and specific alterations in urinary glycoprotein excretion. Indeed, glycoproteins are now important markers of autoimmunity and malignancy [14, 15]. More recently, the plasma glycoproteome has been used to predict nephropathy in diabetic subjects [16]. Despite this promising role as a noninvasive and specific biomarker of disease, little is known about the urinary glycoproteome in CKD.

We hypothesized that the urinary glycoproteome would be altered in CKD compared to healthy controls and that specific glycoprotein alterations might be useful in predicting CKD progression. The overall goal of this study was to perform an initial exploratory analysis of the urine glycoproteins in patients with CKD compared to healthy controls. We present a comprehensive profiling of the urinary glycoproteome in control and CKD subjects utilizing a hydrazide enrichment technique combined with tandem mass spectrometry identification of the glycoproteins.

2. Methods

2.1. Sample Collection and Processing. Clean catch urine samples were obtained from six CKD subjects and six age-matched healthy controls following written informed consent approved by the University of Michigan Institutional Review Board. Samples were stored at −80°C and thawed immediately prior to proteomic analysis. An initial 5000 g centrifugation was performed at 4°C for 10 minutes to remove cellular debris. Approximately, 30–50 mL healthy control samples and 1-2 mL CKD samples were concentrated using a 3 kDa filter cut-off membrane (Vivaspin 3 kDa MWCO, GE healthcare, Buckinghamshire, UK and Amicon ultra 0.5 mL, Millipore, Ireland resp.). As CKD subjects had higher urinary protein content (Table 1), the processed volumes were lower.

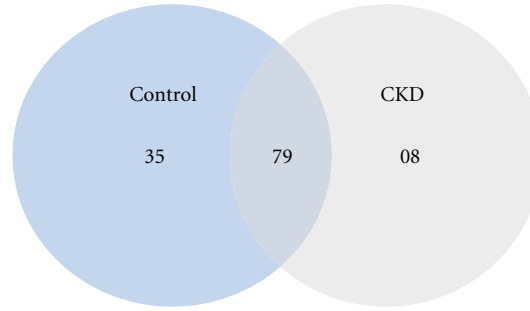


Figure 1: Venn diagram of the total urinary glycoproteins detected in healthy controls and CKD subjects. Tryptic digests of urine glycoproteins were subjected to LC-ESI-MS/MS analysis, and the proteins were identified as described in Section 2. 35 proteins were unique to healthy control subjects while 8 proteins were unique to subjects with CKD. 79 proteins were present in both groups.

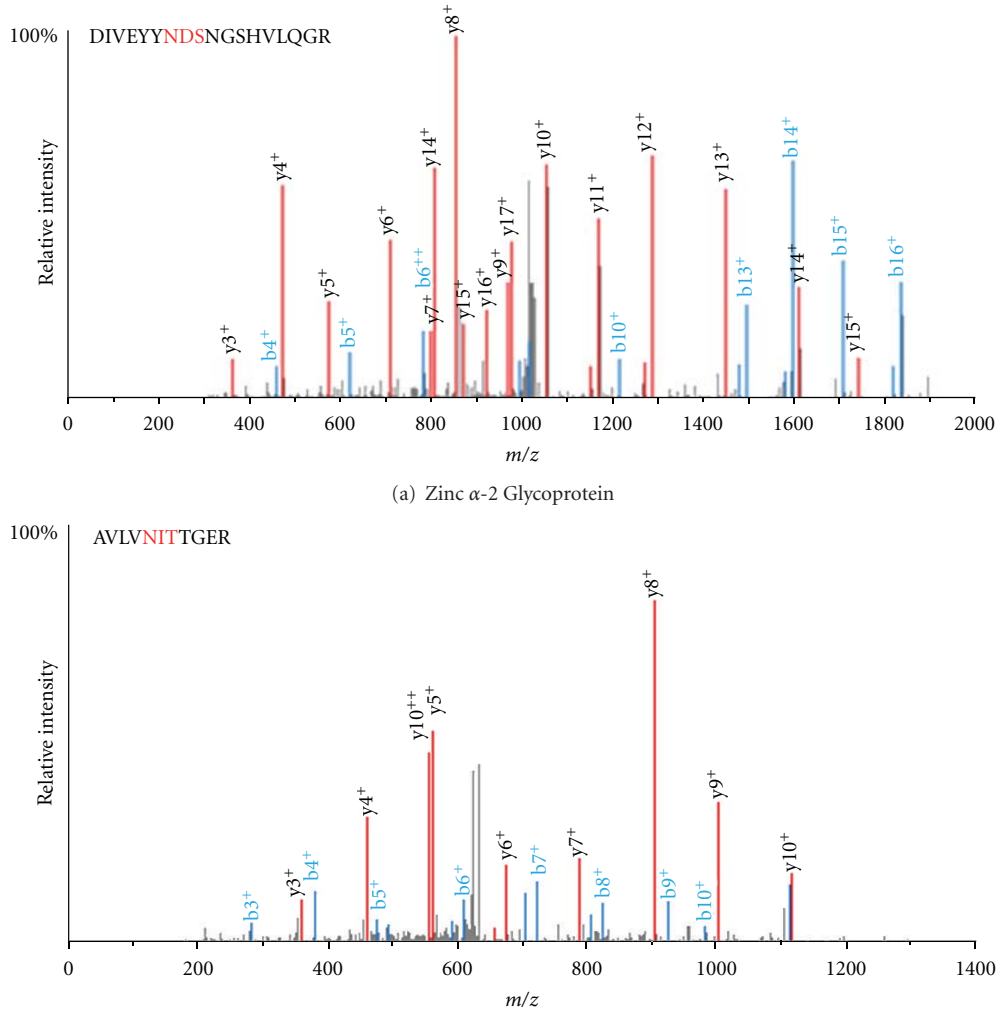
Urine protein concentration was determined using Coomassie Protein Assay Reagent with BSA standard (Thermo Scientific, Rockford, Illinois). 200 μg of concentrated protein were utilized for downstream processing. Protein samples were exchanged into 50 mM ammonium bicarbonate buffer (pH 7.4). Urine creatinine concentration was determined by tandem mass spectrometry (MS/MS) as described previously by our group [17]. To determine the level of creatinine, a known amount of [$^2H_3$]creatinine was spiked into each sample. A full-scan mass spectrum revealed molecular ions of m/z 114 and 117 for authentic creatinine and [$^2H_3$]creatinine, respectively. The transitions of the m/z 114 to 44 and m/z 117 to 47 were monitored in multiple-reaction monitoring mode for authentic and [$^2H_3$]creatinine, respectively, utilizing an Agilent Technologies (New Castle, DE) 6410 Triple Quadrupole mass spectrometer system, equipped with an Agilent 1200 series HPLC

Table 2: Urinary glycoproteins unique to CKD or healthy control subjects.

Unique proteins in healthy controls	Unique proteins in CKD
70 kDa lysosomal alpha-glucosidase (GAA)	Antithrombin-III (SERPINC1)
Alpha-1B-glycoprotein (A1BG)	Complement factor H-related 1 (CFHR1)
Basigin (BSG)	Desmoglein-2 (DSG2)
Beta-galactosidase (GLB1)	Lumican (LUM)
Beta-sarcoglycan (SGCB)	Lymphatic vessel endothelial hyaluronic acid receptor 1 (LYVE1)
Butyrophilin (BTN2A1)	Pigment epithelium-derived factor (SERPINF1)
Carboxypeptidase M (CPM)	Thyroxine-binding globulin (SERPINA7)
CD276 antigen (CD276)	Zinc-alpha-2-glycoprotein (AZP1)
Complement component C4B (C4B)	
Cubilin (CUBN)	
Colony stimulating factor 1 (macrophage) (CSF1)	
Delta and notch-like epidermal growth factor-related receptor (DNER)	
Desmocollin-2 (DSC2)	
Desmoglein-1 (DSG1)	
Epidermal growth factor (EGFR)	
Secreted frizzled-related protein-4 (SFRP4)	
Fibronectin 1 (FN1)	
Folate receptor alpha (FOLR1)	
Golgi phosphoprotein 2 (GOLPH2)	
Glutamyl aminopeptidase (ENPEP)	
Hepatitis B virus receptor binding protein (Q6PYX1)	
Hepatic asialoglycoprotein receptor 1 transcript variant b (ASGR1)	
Heparan sulfate proteoglycan 2 (HSPG2)	
Intercellular adhesion molecule 1 (ICAM1)	
Kallikrein-1 (KLK1)	
Kallikrein 3 (APS)	
Lysosomal alpha-glucosidase (GAA)	
Lysosomal-associated membrane protein 2 (LAMP2)	
Maltase-glucoamylase (MGAM)	
Microfibril-associated glycoprotein 4 (MFAP4)	
Mucin-6 (MUC6)	
Neuronal pentraxin receptor (NPTXR)	
Neuropilin and tolloid-like protein 1 (NETO1)	
Probable serine carboxypeptidase (CPVL)	
Sex hormone binding globulin (SHBG)	

system. The creatinine concentration in the urine sample was determined by comparing the peak areas for authentic and $[^2H_3]$creatinine for the above transitions.

2.2. Glycoprotein Separation and Enrichment. In order to assess recovery following the enrichment procedure, 5 μg of invertase from *Saccharomyces cerevisiae* (Sigma, St. Louis, MO) was spiked into 200 μg of protein in every sample. Glycoproteins were enriched from urinary proteins utilizing the hydrazide resin capture protocol as described previously by Zhang et al. [18]. Briefly, samples were oxidized with 10 mM sodium metaperiodate then incubated with hydrazide resin overnight at room temperature. Samples were then centrifuged at 3000 g for 2 minutes and the resin was washed successively with equal volumes 50 mM ammonium bicarbonate buffer (pH 7.4; Buffer A) supplemented with 8 M urea, followed by Buffer A alone and then water. The beads were resuspended in water, and the protein was reduced with 5 mm DTT followed by alkylation with 15 mM iodoacetamide. Trypsin (sequencing grade modified trypsin, Promega Corporation, Madison, WI) at 1 : 20 μg ratio was added to the samples and incubated overnight at 37°C for digestion. Following digestion, the beads were centrifuged at 3000 g for 2 minutes and the resin was then washed successively with 1.5 M NaCl, 80% acetonitrile, 100% methanol, and Buffer A. The resin was then resuspended in Buffer A and incubated with 5 units of PNGaseF (New England Biolabs, Ipswich, MA) overnight at 37°C for glycopeptide release. The glycopeptides were cleaned using a reverse phase column and



(a) Zinc α-2 Glycoprotein

(b) Golgi Phosphoprotein 2

Figure 2: Mass Spectra of glycopeptides derived from Zinc alpha 2 Glycoprotein (a) and Golgi phosphoprotein (b) in CKD urine. Tryptic digests of urine glycoproteins were subjected to LC-ESI-MS/MS analysis as described in Section 2. The mass spectra of peptides DIVEYYNDSNGSHVLQG from zinc alpha 2 glycoprotein which is upregulated (a) and those of peptide AVLVNNITTGER from Golgi phosphoprotein which is significantly downregulated in CKD subjects (b) are shown. The N-linked glycosylation site of each peptide is depicted in red.

eluted with 50% acetonitrile/0.1% TFA followed by elution with 80% acetonitrile/0.1% TFA. The peptides were then dried at 60°C in a vacuum centrifuge and stored for mass spectrometric analysis.

2.3. Liquid Chromatography Electrospray Ionization (ESI/LC) MS/MS Analysis. Peptide samples were resuspended in 0.1% formic acid and loaded onto an in-house packed reverse phase separation column (0.075 × 100 mm, MAGIC C18 AQ particles, 5 μm, Michrom Bioresources). The peptides were separated on a 1% acetic acid/acetonitrile gradient system (5–50% acetonitrile for 75 min, followed by a 10 min 95% acetonitrile wash) at a flow rate of ~300 nl/min. Peptides were directly sprayed onto the MS using a nanospray source. An LTQ Orbitrap XL (Thermo Fisher Scientific, Waltham, MA) was run in automatic mode collecting a high resolution MS scan (FWHM 30,000) followed by data-dependent acquisition of MS/MS scans on the 9 most intense ions (relative collision energy ~35%). Dynamic exclusion was set to collect 2 MS/MS scans on each ion and exclude it for an additional 2 min. Charge state screening was enabled to exclude +1 and undetermined charge states.

2.4. Data Processing and Statistical Analysis. The Human UniProt database (Release 2011-5) was appended with a reverse database, a common contaminant list, and yeast invertase. Raw files were converted to mzXML format and searched against the database using X!Tandem with a *k-score* plug-in, an open-source search engine developed by the Global Proteome Machine (http://www.thegpm.org/). The search parameters were as follows: (1) precursor mass tolerance window of 100 ppm and fragment mass

Table 3: Glycoproteins identified with peptides carrying NxS/T motif.

No.	Protein	Charge state	Peptide sequence with NxS/T motif
1	155 kDa platelet multimerin (MMRN1)	2^+ 2^+	LQ**N[115]LT**LPT**N[115]AS**IK FNPGAESVVLS**N[115]S**TLK
2	70 kDa lysosomal alpha-glucosidase (GAA)	2^+ 2^+	GVFIT**N[115]ET**GQPLIGK LE**N[115]LS**SSEM[147]GYTATLTR
3	Afamin (AFAM)	2^+ $2^+/3^+$	DIENF**N[115]ST**QK YAEDKF**N[115]ET**TEK
4	Aminopeptidase N (AMPN)	3^+ 2^+	KL**N[115]YT**LSQGHR **N[115]AT**LVNEADKLR
5	Attractin (ATRN)	2^+	IDSTG**N[115]VT**NELR
6	Apolipoprotein D (APO D)	$2^+/3^+$ 2^+	ADGTVNQIEGEATPV**N[115]LT**EPAK C[160]IQA**N[115]YS**LM[147]EN[115]GK
7	Apolipoprotein F (APO F)	2^+	Q[111]GGV**N[115]AT**QVLIQHLR
8	Apolipoprotein J (APO J)	2^+ $2^+/3^+$ 3^+ 3^+	LA**N[115]LT**QGEDQYYLR EDAL**N[115]ET**RESETK M[147]L**N[115]TS**SLLEQLNEQFNWVSR EIRH**N[115]ST**GC160LR
9	Alpha-1-antichymotrypsin (AACT)	$3^+/4^+$ $3^+/4^+$ $2^+/3^+$ $2^+/3^+$	GLKF**N[115]LT**ETSEAEIHQSFQHLLR YTG**N[115]AS**ALFILPDQDKM[147]EEVEAM[147]LLPETLKR TL**N[115]QS**SDELQLSM[147]GNAM[147]FVK KLIN[115]DYVK**N[115]GT**R
10	Alpha-2-HS-glycoprotein (FETUA)	$2^+/3^+$ $2^+/3^+$	AALAAFNAQN**N[115]GS**NFQLEEISR KVC[160]QDC[160]PLLAPL**N[115]DT**R
11	Alpha-1-acid glycoprotein 1 (ORM1)	$2^+/3^+$ 3^+ $3^+/4^+$	QDQC[160]IY**N[115]TT**YLNVQR SVQEIQATFFYFTP**N[115]KT**EDTIFLR N[115]EEY**N[115]KS**VQEIQATFFYFTP**N[115]KT**EDTIFLR
12	Alpha-1-acid glycoprotein 2 (ORM2)	$2^+/3^+$ 3^+	QNQC[160]FY**N[115]SS**YLNVQR SVQEIQATFFYFTP**N[115]KT**EDTIFLR
13	Alpha-1B-glycoprotein (A1BG)	$3^+/4^+$	EGDHEFLEVPEAQEDVEATFPVHQPG**N[115]YS**C[160]SYR
14	Antithrombin-III (SERPINC1)	$2^+/3^+$ 2^+	LGAC[160]**N[115]DT**LQQLM[147]EVFKFDTISEK SLTF**N[115]ET**YQDISELVYGAK
15	Basigin (BSG)	3^+ 3^+	ITDSEDKALM[147]**N[115]GS**ESR ILLTC[160]SL**N[115]DS**ATEVTGHR
16	Beta-galactosidase (GLB1)	2^+	NNVITL**N[115]IT**GK
17	Beta-sarcoglycan (SGCB)	2^+	ITS**N[115]AT**SDLNIK
18	Biotinidase (BTD)	$2^+/3^+$ $2^+/3^+$	NPVGLIGAE**N[115]AT**GETDPSHSK DVQIIVFPEDGIHGF**N[115]FT**R
19	Butyrophilin, subfamily 2, member A1 (BTN2A1)	2^+	GSVALVIH**N[115]IT**AQEN[115]GTYR
20	Cathepsin D heavy chain (CTSD)	2^+	GSLSYL**N[115]VT**R
21	Cathepsin L (CTSL)	3^+	YSVA**N[115]DT**GFVDIPKQEK
22	Carboxypeptidase B2 (CBPB2)	$2^+/3^+$	QVHFFV**N[115]AS**DVDNVK
23	Carboxypeptidase M (CBPM)	2^+ 4+	NFPDAFEYN**N[115]VS**R TVAQ**N[115]YS**SVTHLHSIGK
24	Calcium binding protein 39 (CAB39)	2^+	H**N[115]FT**IM[147]TK
25	CD276 antigen (CD276)	2^+	VVLGA**N[115]GT**YSC[160]LVR
26	CD163 antigen (CD163)	2^+	APGWA**N[115]SS**AGSGR
27	CD44 protein (CD44)	2^+	AF**N[115]ST**LPTM[147]AQM[147]EK
28	CD7 antigen (CD7)	3^+	GRIDFSGSQD**N[115]LT**ITM[147]HR

Table 3: Continued.

No.	Protein	Charge state	Peptide sequence with NxS/T motif
29	Cell adhesion molecule 1 (CADM1)	2^+ 2^+	VSLT**N[115]VS**ISDEGR FQLL**N[115]FS**SSELK
30	Ceruloplasmin (CP)	$2^+/3^+$ $3^+/4^+$ 2^+	EHEGAIYPD**N[115]TT**DFQR ELHHLQEQ**N[115]VS**NAFLDKGEFYIGSK E**N[115]LT**APGSDSAVFFEQGTTR
31	Complement component C4B (C4B)	2^+	GL**N[115]VT**LSSTGR
32	Complement factor H (CFH)	3^+	IPC[160]SQPPQIEHGTI**N[115]SS**R
33	Complement factor H-related 1 (CFHR1)	2^+	LQNNEN**N[115]IS**C[160]VER
34	Complement factor I (CFI)	2^+	FLN**N[115]GT**C[160]TAEGK
35	Cubilin (CUBN)	2^+ 2^+	LC[160]SSV**N[115]VS**NEIK AGF**N[115]AS**FHK
36	Corticosteroid-binding globulin (SERPINA6)	2^+ 3^+	AQLLQGLGF**N[115]LT**ER AVLQLNEEGVDTAGSTGVTL**N[115]LT**SKPIILR
37	Colony stimulating factor 1 (macrophage) (CSF1)	2^+	VKNVF**N[115]ET**K
38	Delta and notch-like epidermal growth factor-related receptor (DNER)	2^+	LVSFEVPQ**N[115]TS**VK
39	Desmocollin-2 (DSC2)	2^+ 2^+	LKAI**N[115]DT**AAR A**N[115]YT**ILK
40	Desmoglein-1 (DSC1)	2^+	DYNTK**N[115]GT**IK
41	Desmoglein-2 (DSG2)	2^+ 2^+	I**N[115]AT**DADEPNTLNSK YVQ**N[115]GT**YTVK
42	DNA ligase 4 (LIG4)	2^+	AP**N[115]LT**NVNK
43	Dual specificity protein phosphatase CDC14B (CDC14B)	2^+	NH**N[115]VT**TIIR
44	Epidermal growth factor (EGF)	$\mathbf{2^+}$	G**N[115]NS**HILLSALK
45	Epididymis secretory sperm binding protein Li 44a (SERPINA1)	$2^+/3^+/4^+$ $3^+/4^+$	YLG**N[115]AT**AIFFLPDEGKLQHLENELTHDIITK ADTHDEILEGLNF**N[115]LT**EIPEAQIHEGFQELLR
46	Extracellular link domain containing 1 (XLKD1)	$2^+/3^+$	KANQQL**N[115]FT**EAK
47	Fibrillin 1 (FBN1)	$\mathbf{2^+}$	TAIFAF**N[115]IS**HVSNK
48	Fibrinopeptide A (FGA)	$\mathbf{2^+}$	M[147]DGSLNF**N[115]RT**
49	Fibronectin type III domain-containing protein 5 (FNDC5)	$\mathbf{2^+}$	FIQEV**N[115]TT**TR
50	Frizzled protein 4 (FRP4)	$\mathbf{2^+}$	ISM[147]C[160]QNLGY**N[115]VT**K
51	Fibronectin 1 (FN1)	3^+	DQC[160]IVDDITYNV**N[115]DT**FHK
52	Folate receptor alpha (FOLR1)	$\mathbf{2^+}$	GW**N[115]WT**SGFNK
53	Galectin-3-binding protein (LGALS3BP)	2^+ 2^+ 2^+ 2^+	ALGFE**N[115]AT**QALGR AAIPSALDT**N[115]SS**K GL**N[115]LT**EDTYKPR TVIRPFYLT**N[115]SS**GVD
54	Glutaminyl-peptide cyclotransferase (QPCT)	$2^+/3^+$ $3^+/4^+$	NYHQPAIL**N[115]SS**ALR YFQ**N[115]YS**YGGVIQDDHIPFLR
55	Golgi phosphoprotein 2 (GOLPH2)	3^+ $1^+/2^+$	LQQDVLQFQKN[115]QTNLER AVLVN[115]**N[115]IT**TGER

Table 3: Continued.

No.	Protein	Charge state	Peptide sequence with NxS/T motif
56	Glutamyl aminopeptidase (ENPEP)	2^+	HTAEYAA**N[115]IT**K
57	Haptoglobin beta chain (HP)	$2^+/3^+$ $3^+/4^+$ $2^+/3^+/4^+$ 3^+	VVLHP**N[115]YS**QVDIGLIK MVSHH**N[115]LT**TGATLINEQWLLTTAK NLFL**N[115]HS**E**N[115]AT**AKDIAPTLTLYVGKK Q[111]LVEIEKVVLHP**N[115]YS**QVDIGLIK
58	HEG homolog 1 (HEG1)	2^+	SYSESSSTSSSESL**N[115]SS**APR
59	Hemopexin (HPX)	$3^+/4^+$ 2^+	GHGHR**N[115]GT**GHG**N[115]ST**HHGPEYM[147]R ALPQPQ**N[115]VT**SLLGC[160]TH
60	Hepatitis B virus receptor binding protein (Q6YPX1)	2^+	EEQY**N[115]ST**YR
61	Hepatic asialoglycoprotein receptor 1 transcript variant b (ASGR1)	2^+	ETFS**N[115]FT**ASTEAQVK
62	Heparan sulfate proteoglycan 2 (HSPG2)	2^+	ALV**N[115]FT**R
63	Ig alpha-1 chain C region (IGHA1)	3^+	LAGKPTHV**N[115]VS**VVM[147]AEVDGTC[160]Y
64	Ig gamma-1 chain C region (IGHG1)	2^+ $2^+/3^+$	EEQY**N[115]ST**YR TKPREEQY**N[115]ST**YR
65	Ig gamma-2 chain C region (IGHG2)	2^+ $2^+/3^+$	EEQF**N[115]ST**FR TKPREEQF**N[115]ST**FR
66	Ig gamma-4 chain C region (IGHG4)	2^+ $2^+/3^+$	EEQF**N[115]ST**FR TKPREEQF**N[115]ST**FR
67	Ig mu chain C region (IGHM)	2^+	YK**N[115]NS**DISSTR
68	Inducible T-cell co-stimulator ligand (ICOSLG)	2^+	TVVTYHIPQ**N[115]SS**LENVDSR
69	Insulin-like growth factor-binding protein 3 (IGFBP3)	2^+	GLC[160]V**N[115]AS**AVSR
70	Intercellular adhesion molecule 1 (ICAM1)	2 2^+	LNPTVTYG**N[115]DS**FSAK A**N[115]LT**VVLLR
71	Intercellular adhesion molecule 2 (ICAM2)	2^+	G**N[115]ET**LHYETFGK
72	Kallikrein-1 (KLK1)	4^+	HNLFDDEN[115]TAQFVHVSESFPHPGF**N[115]M[147]S**LLE**N[115]HT**R
73	KALLIKREIN-2 (KLK2)	2^+	**N[115]KS**VILLGR
74	Kininogen 1 (KNG1)	2^+ 2^+ $3^+/4^+$ 2^+	LNAEN**N[115]AT**FYFK ITYSIVQT**N[115]C[160]S**K HGIQYFN**N[115]NT**QHSSLFTLNEVKR YNSQ**N[115]QS**NNQFVLYR
75	Leucine-rich alpha-2-glycoprotein (LRG1)	2^+ $2^+/3^+$	MFSQ**N[115]DT**R KLPPGLLA**N[115]FT**LLR
76	Leukocyte-associated immunoglobulin-like receptor 1 (LAIR1)	$2^+/3^+$	STY**N[115]DT**EDVSQASPSESEAR
77	Lumican (LUM)	3^+	KLHINHN**N[115]LT**ESVGPLPK
78	Lysosomal acid phosphatase (ACP2)	2^+	YEQLQ**N[115]ET**R
79	Lysosomal alpha-glucosidase (GAA)	2^+ 2^+	GVFIT**N[115]ET**GQPLIGK LE**N[115]LS**SSEM[147]GYTATLTR
80	Lysosome-associated membrane glycoprotein 1 (LAMP1)	2^+	GHTLTL**N[115]FT**R

Table 3: Continued.

No.	Protein	Charge state	Peptide sequence with NxS/T motif
81	Lysosomal-associated membrane protein 2, (LAMP2)	2^+	VASVININP**N[115]TT**HSTGSC[160]R
82	Lymphatic vessel endothelial hyaluronic acid receptor 1 (XLKD1)	2^+	ANQQL**N[115]FT**EAK
83	Lysyl oxidase (LOX)	3^+ 3^+	AE**N[115]QT**APGEVPALSNLRPPSR RDPGAAVPGAA**N[115]AS**AQQPR
84	Major prion protein (PRNP)	2^+	Q[111]HTVTTTTKGE**N[115]FT**ETDVK
85	Membrane protein FAM174A (FAM174A)	2^+	GSEGG**N[115]GS**NPVAGLETDDHGGK
86	Maltase-glucoamylase (MGAM)	2^+ 2^+	ILGM[147]EEPS**N[115]VT**VK VILILDPAISG**N[115]ET**QPYPAFTR
87	Microfibril-associated glycoprotein 4 (MFAP4)	2^+	VDLEDFE**N[115]NT**AYAK
88	Monocyte differentiation antigen CD14	2^+	LR**N[115]VS**WATGR
89	Mucin-6 (MUC6)	2^+	GC[160]M[147]A**N[115]VT**VTR
90	N-acetylglucosamine-6-sulfatase (GNS)	2^+ 2^+	YY**N[115]YT**LSIN[115]GK TPMT**N[115]SS**IQFLDNAFR
91	N-acylsphingosine amidohydrolase (ASAH1)	2^+	TVLE**N[115]ST**SYEEAK
92	Neuronal pentraxin receptor (NPTXR)	2^+	ALPGGAD**N[115]AS**VASGAAASPGPQR
93	Neuropilin and tolloid-like protein 1 (NETO1)	2^+	HESEY**N[115]TT**R
94	Peptidase inhibitor 16 (PI16)	2^+	SLPNFP**N[115]TS**ATA**N[115]AT**GGR
95	Pigment epithelium-derived factor (SERPINF1)	3^+	VTQ**N[115]LT**LIEESLTSEFIHDIDR
96	Plasma protease C1 inhibitor (SERPING1)	$2^+/3^+$ 2^+ 3^+	GVTSVSQIFHSPDLAIRDTFV**N[115]AS**R VLS**N[115]NS**DANLELINTWVAK VGQLQLSH**N[115]LS**LVILVPQNLK
97	Plasma serine protease inhibitor (SERPINA5)	2^+	VVGVPYQG**N[115]AT**ALFILPSEGK
98	Platelet-derived growth factor subunit B (PDGFB)	3^+	LLHGDPGEEDGAELDL**N[115]M[147]T**R
99	Polytrophin (TROPH)	2^+	N[115]**N[115]VT**EDIK
100	Probable G-protein coupled receptor 116 (GPR116)	2^+ 2^+	ANEQVVQSL**N[115]QT**YK YEEQQLEIQ**N[115]SS**R
101	Probable serine carboxypeptidase (CPVL)	2^+	Q[111]AIHVG**N[115]QT**FNDGTIVEK
102	Prosaposin (PSAP)	$2^+/3^+$	NLEK**N[115]ST**KQEILAALEK
103	Prostaglandin D2 synthase 21 kDa (PTGDS)	$2^+/3^+$ 2^+	SVVAPATDGGL**N[115]LT**STFLR WFSAGLAS**N[115]SS**WLR
104	Prostatic acid phosphatase (ACPP)	3^+	FL**N[115]ES**YKHEQVYIR
105	Proteinase-activated receptor 1 (F2R)	2^+	AT**N[115]AT**LDPR
106	Protein shisa-7 (SHISA7)	2^+	LTGALTGGGGAASPGA**N[115]GT**R
107	RING finger protein 10 (RNF10)	2^+	**N[115]ES**FN[115]**N[115]QS**R

Table 3: Continued.

No.	Protein	Charge state	Peptide sequence with NxS/T motif
108	Secretory component (Polymeric IG Receptor) (PIGR)	3^+ 2^+ 2^+ 2^+	A**N[115]LT**NFPE**N[115]GT**FVVNIAQLSQDDSGR Q[111]IGLYPVLVIDSSGYVNP**N[115]YT**GR VPG**N[115]VT**AVLGETLK YKCGLGI**N[115]S**R
109	Slit homolog 1 (SLIT1)	2^+	LELN[115]GN[115]**N[115]IT**R
110	Sushi domain-containing protein 2 (SUSD2)	2^+	SELV**N[115]ET**R
111	Sex hormone binding globulin (SHBG)	2^+	LDVDQAL**N[115]RT**
112	Transferrin (TF)	$2^+/3^+$ $2^+/3^+$	Q[111]QQHLFGS**N[115]VT**DC[160]SGNFC[160]LFR C[160]GLVPVLAENY**N[115]KS**DN[115]C[160]EDTPEAGYFAVAVVK
113	Thrombin heavy chain (F2)	4^+	YPHKPEI**N[115]ST**THPGADLQENFC[160]R
114	Tripeptidyl-peptidase I variant (TPP1)	3^+	FLSSSPHLPPSSYFN[115]ASGR
115	Tyrosine-protein kinase receptor UFO (AXL)	2^+ 3^+	SLHVPGL**N[115]KT** **N[115]GS**QAFVHWQEPR
116	TIMP metallopeptidase inhibitor 1 (TIMP1)	2^+ 3^+	FVGTPEV**N[115]QT**TLYQR SH**N[115]RS**EEFLIAGK
117	Thyroxine-binding globulin (SERPINA7)	2^+	TLYETEVFSTDFS**N[115]IS**AAK
118	Trypstatin (AMBP)	$2^+/3^+$	SKW**N[115]IT**M[147]ESYVVHTNYDEYAIFLTK
119	Transmembrane protein 108 (TMEM108)	4^+	KGAG**N[115]SS**RPVPPAPGGHSR
120	Uromodulin (UMOD)	$2^+/3^+$ 2^+ $2^+/3^+$	Q[111]DF**N[115]IT**DISLLEHR FALLMTNCYATPSS**N[115]AT**DPLK CNTAAPMWL**N[115]GT**HPSSDEGIVSR
121	Vasorin (VASN)	2^+	LHEIT**N[115]ET**FR
122	Zinc-alpha-2-glycoprotein (AZGP1)	$2^+/3^+$ $3^+/4^+$ 2^+	DIVEYY**N[115]DSN[115]GS**HVLQGR AREDIFM[147]ETLKDIVEYY**N[115]DSN[115]GS**HVLQGR FGCEIEN**N[115]RS**

tolerance of 0.8 Da; (2) allowing two missed cleavages; (3) variable modification: oxidation of methionine (+15.9949 Da), carbamidomethyl cysteine (57.0214 Da), and +0.9840 Da, reflecting the conversion of asparagine in the NxS/T motif to aspartate due to the release of the N-linked glycopeptides from their oligosaccharides. All proteins with a ProteinProphet probability of greater than 0.9 were considered as positive identifications [19]. Only proteins containing peptides with the NxS/T sequence motif were included for statistical analysis.

Baseline characteristics of the control and CKD subjects were compared using Fisher's exact test for categorical variables and Student's *t*-test for continuous variables. Data is presented as means (±SD). Spectral counts for individual proteins were normalized to *Saccharomyces cerevisiae* invertase and to urine creatinine content. Spectral counts were compared across the two subject groups using the nonparametric Mann-Whitney test, and *P* values were adjusted for multiple comparisons using the False Discovery Rate (FDR) with reported *q*-values. All statistical analyses were performed with the use of SAS software, version 9.2.

2.5. Gene Ontology Analysis. Significant proteins of interest were analyzed using the Gene Ontology Database (Gene Ontology Consortium, http://www.geneontology.org, Princeton University, New Jersey, US; [20]). For a given Gene Ontology (GO) category, the relative enrichment of genes encoding the proteins detected in CKD relative to all reference genes in that category were calculated as previously described using GO Tools made available by the Bioinformatics Group at the Lewis-Sigler Institute (Princeton University, New Jersey, US; [21]). A cutoff value of $P < 0.01$ was used to report a functional category as significantly overrepresented. To address the multiple comparisons problem that arises when many processes are evaluated simultaneously, the analysis included calculation of the FDR [21]. To improve statistical confidence in our results, all enriched functional categories were required to be significant using both methods ($P < 0.01$ and FDR < 0.05).

3. Results

3.1. Study Subject Characteristics. Urine was isolated from six subjects with CKD and six age-matched healthy controls.

Baseline subject characteristics are provided in Table 1. Two important issues were considered with patient selection. First, the etiology of CKD was chosen to be diverse. This would ensure robustness of the putative markers as a CKD marker rather than a disease-specific marker. Second, we specifically targeted early Stage 3 CKD subjects to identify early disease markers that would potentially indicate pathways dysregulated early in the course of disease. This might offer mechanistic insights into disease pathogenesis and progression and have implications in therapeutic strategies. The six subjects had biopsy-proven diabetic nephropathy, lupus nephritis ($n = 2$), postacute tubular necrosis damage, NSAID nephropathy, and membranoproliferative glomerulonephritis, respectively. The mean estimated glomerular filtration rate (eGFR) was 83 mL/min in control subjects and 52 mL/min in CKD subjects.

3.2. Glycoprotein Spectral Count Normalization. Glycoproteins were extracted and enriched from the twelve urinary samples. To account for variations in the glycoprotein extraction efficiency, 5 μg of the yeast protein invertase from *Saccharomyces cerevisiae* was added to each sample prior to extraction. After addition to the database, invertase spectral count served as a surrogate marker for extraction efficiency in each individual sample. Invertase spectral counts ranged from 31 to 122 in the twelve samples with an average spectral count of 86 ($\pm$31). Each sample was normalized independently to the invertase spectral counts.

To account for intersubject urine concentration variability, spectral counts were then normalized to urine creatinine content. This provides standardization for urinary creatinine excretion and concentration differences which can vary with volume status, stress, diet, activity level, age, gender, and overall health status [22]. Indeed, this normalization is commonly followed in clinical practice where degree of urinary protein is normalized to creatinine to obtain protein excretion rates [23]. Final spectral counts were expressed per mmol creatinine.

3.3. Urine Glycoproteome Is Altered in CKD. Urinary glycoproteins were isolated from six subjects with CKD and six healthy controls using a hydrazide technique as described in Section 2. A total of 122 glycoproteins were identified, of which 35 proteins were unique to healthy control patients, 8 were unique to CKD subjects, and 79 were common proteins in both groups (Figure 1, Table 2). Unique proteins to the CKD group were Antithrombin-III (SERPINC1), Complement factor H-related 1 (CFHR1), Desmoglein-2 (DSG2), Lumican (LUM), Lymphatic vessel endothelial hyaluronic acid receptor 1 (LYVE1), Pigment epithelium-derived factor (SERPINF1), Thyroxine-binding globulin (SERPINA7), and Zinc-alpha-2-glycoprotein (AZP1).

Figure 2 displays MS spectra of two individual glycopeptides with glycosylation motifs which were altered in CKD subjects. Zinc-alpha-2-glycoprotein is significantly upregulated in CKD (Figure 2(a)), while Golgi phosphoprotein is significantly downregulated in CKD (Figure 2(b)). Table 3 displays motifs and specific peptide modifications for all unique 122 proteins. Proteins were only included if the peptides contained the NxS/T motif.

To test if proteins were significantly up- or downregulated in CKD, normalized spectral counts from the 6 CKD subjects were compared with those from the healthy controls. As sample size was small and spectral counts were not normally distributed, comparisons were made with the nonparametric Mann-Whitney test. As 122 proteins were being simultaneously tested, the FDR and corresponding q-values were determined to account for false positive results. Table 4 displays 23 proteins which are differentially expressed in CKD utilizing an uncorrected P value threshold of less than 0.05. These proteins include 70 kDa lysosomal alpha-glucosidase (GAA), Apolipoprotein D (APOD), Alpha-2-HS-glycoprotein chain B (FETUA), Alpha-1-acid glycoprotein 1 (ORM1), Antithrombin-III (SERPINC1), Beta-galactosidase (GLB1), Ceruloplasmin (CP), Cubilin (CUBN), Epidermal growth factor (EGF), Epididymis secretory sperm binding protein Li 44a (E9KL23), Galectin-3-binding protein (LGALS3BP), Golgi phosphoprotein 2 (GOLPH2), Haptoglobin beta chain (HP), Ig gamma-1 chain C region (IGHG1), Ig gamma-2 chain C region (IGHG2), Kininogen 1 (KNG1), Leucine-rich alpha-2-glycoprotein (LRG), Plasma protease C1 inhibitor (SERPING1), Prostaglandin D2 synthase 21 kDa (PTGDS), Transferrin (TF), Trypstatin (AMBP), Uromodulin (UMOD), and Zinc-alpha-2-glycoprotein (AZGP1). Following correction for multiple comparisons, differential expression remained significant in 12 proteins (APOD, ORM1, FETUA, E9KL23, LGALS3BP, GOLPH2, HP, KNG1, LRG, SERPING1, PTGDS, AZGP1). Incidentally, not all unique proteins to CKD or healthy control groups had statistically significant up- or down-regulation. For example, lumican was not isolated in any healthy control subjects and was found in only three of the six CKD subjects. Thus, lumican is unique to CKD; however, as it was only seen in three CKD subjects, it was not significantly upregulated in CKD via nonparametric testing.

3.4. Gene Ontology Analysis Reveals Enrichment for Distinct Biological Functions of Differentially Expressed Urinary Glycoproteins. The 23 proteins with differential expression in CKD were subjected to a GO Database search and further analyzed with GO Tools [20, 21]. GO Term Finder (http://go.princeton.edu/cgi-bin/GOTermFinder) allowed for clustered identification of proteins annotated to specific GO biological process, location, and function classifications. A subsequent GO Term Mapper (http://go.princeton.edu/cgi-bin/GOTermMapper) analysis of significantly altered proteins was performed to bin the proteins to GO parent terms or GO Slim terms (http://www.geneontology.org/GO.slims.shtml).

GO analysis (Figure 3) for biological processes demonstrated that 16 of the 23 proteins were linked to immune/stress response and biological process regulation ($P < 1 \times 10^{-4}$). 9 of the 23 were acute-phase and inflammatory response proteins ($P < 1 \times 10^{-3}$). Six proteins were regulators of hemostasis, platelet degranulation and coagulation ($P < 1 \times 10^{-4}$), and 10 were involved in

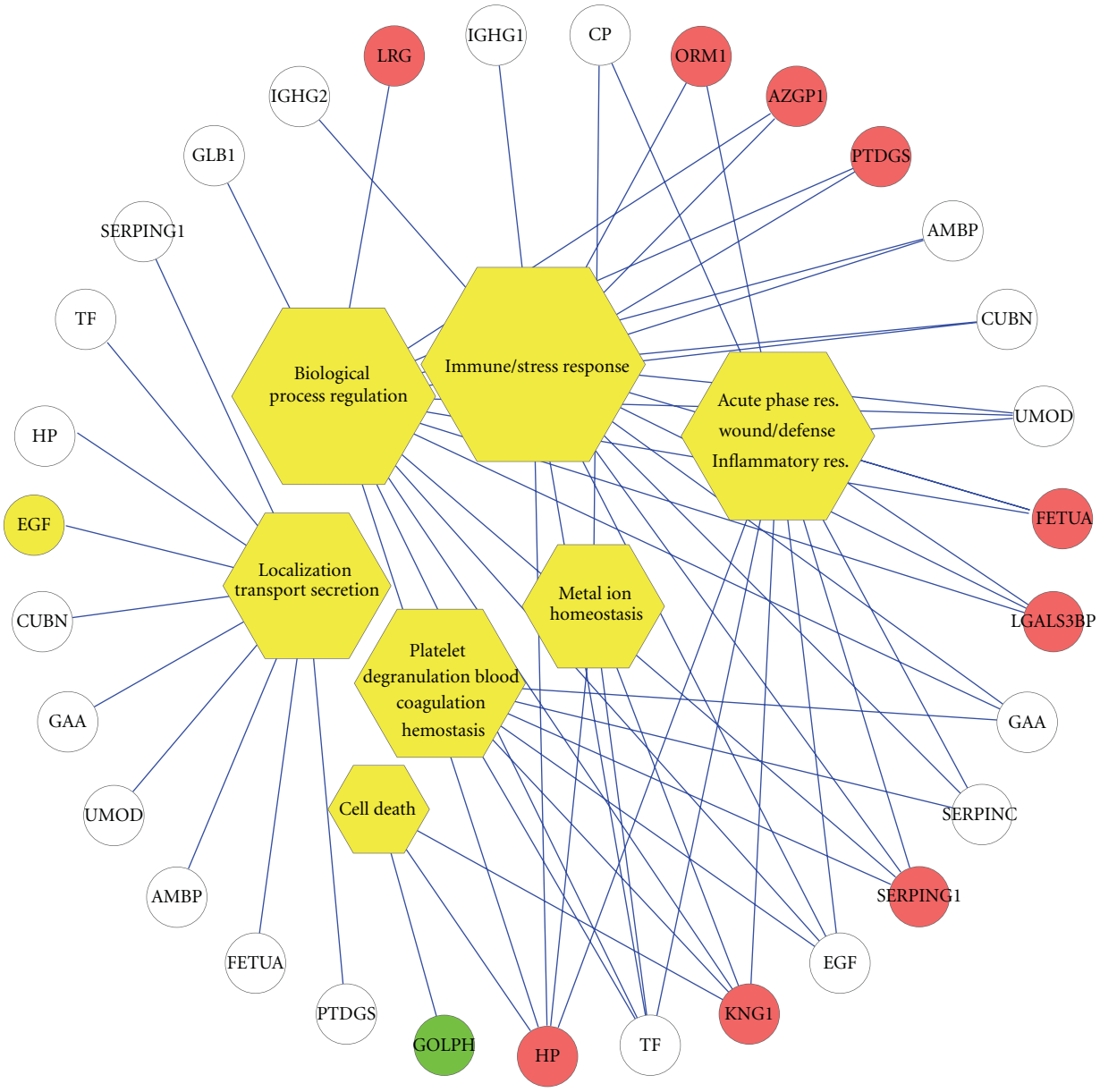


Figure 3: Global view of biological processes of differentially expressed urinary glycoproteins in CKD. Urinary glycoproteins that were differentially detected in CKD subjects were associated with biological functions using GO process annotations. This approach demonstrated significant overrepresentation of proteins involved in several categories, including regulation of response to stress, platelet activation/hemostasis/coagulation, acute-phase response, regulation of biological processes, localization, secretion, transport, and cell death. A bipartite network (generated using Cytoscape [24]) showing the relationship between GO process annotations (yellow hexagon nodes) and differentially regulated proteins in CKD subjects (white/red/green circular nodes). The size of the GO nodes is proportional to the number of edges (lines) that connect them to proteins. The 10 proteins that are altered with q-value of < 0.05 are depicted in red (up-regulation) and green (down-regulation) in CKD subjects. GAA, 70 kDa lysosomal alpha-glucosidase; APOD, Apolipoprotein D; FETUA: Alpha-2-HS-glycoprotein chain B; ORM1, Alpha-1-acid glycoprotein 1; SERPINC1, Antithrombin-III; GLB1, Beta-galactosidase; CP, Ceruloplasmin; CUBN, Cubilin; EGF, Epidermal growth factor; E9KL23, Epididymis secretory sperm binding protein Li 44a; LGALS3BP, Galectin-3-binding protein; GOLPH2, Golgi phosphoprotein 2; HP, Haptoglobin beta chain; IGHG1, Ig gamma-1 chain C region; IGHG2, Ig gamma-2 chain C region; KNG1, Kininogen 1; LRG, Leucine-rich alpha-2-glycoprotein; SERPING1,Plasma protease C1 inhibitor; PTGDS, Prostaglandin D2 synthase 21 kDa; TF, Transferrin; AMBP, Trypstatin; UMOD, Uromodulin; AZGP1, Zinc-alpha-2-glycoprotein.

Table 4: Differentially regulated proteins identified in CKD subjects.

Protein Code	Name of the protein identified	P value	q-value	Direction of change in CKD subjects
APOD	Apolipoprotein D	0.0022	0.0224	Up
FETUA	Alpha-2-HS-glycoprotein chain B	0.0022	0.0224	Up
ORM1	Alpha-1-acid glycoprotein 1	0.0022	0.0224	Up
E9KL23	Epididymis secretory sperm binding protein Li 44a	0.0022	0.0224	Up
LGALS3BP	Galectin-3-binding protein	0.0022	0.022	Up
GOLPH2	Golgi phosphoprotein 2	0.0022	0.0224	Down
HP	Haptoglobin beta chain	0.0022	0.0224	Up
KNG1	Kininogen 1	0.0022	0.0224	Up
LRG	Leucine-rich alpha-2-glycoprotein	0.0022	0.0224	Up
SERPING1	Plasma protease C1 inhibitor	0.0022	0.0224	Up
PTGDS	Prostaglandin D2 synthase 21kDa	0.0022	0.0224	Up
AZGP1	Zinc-alpha-2-glycoprotein	0.0022	0.0224	Up
GAA	70 kDa lysosomal alpha-glucosidase	0.0152	0.13	Down
SERPINC1	Antithrombin-III	0.0152	0.103	Up
GLB1	Beta-galactosidase	0.0152	0.103	Down
CUBN	Cubilin	0.0152	0.103	Down
EGF	Epidermal growth factor	0.0152	0.13	Down
UMOD	Uromodulin	0.0152	0.103	Up
TF	Transferrin	0.0216	0.1387	Up
AMBP	Trypstatin	0.0411	0.18	Up
CP	Ceruloplasmin	0.0433	0.18	Up
IGHG1	Ig gamma-1 chain C region	0.0433	0.18	Up
IGHG2	Ig gamma-2 chain C region	0.0433	0.18	Up

localization, transport, and secretion ($P < 1 \times 10^{-4}$). Other processes involved include metal ion homeostasis (4 proteins) and cell death (3 proteins).

Table 5 displays function and location for the 23 proteins which were differentially expressed in CKD. 18 out of the 23 proteins localized to the extracellular region consistent with possible extracellular matrix remodeling that typifies renal disease. The analysis also revealed 2 major clusters of molecular function: 20 out of the 23 proteins were involved in binding and protein-protein interactions ($P = 5 \times 10^{-4}$). 5 proteins were endopeptidase inhibitors ($P < 1 \times 10^{-6}$). Collectively, these observations implicate the inflammatory/acute-phase response and extracellular matrix remodeling in CKD. They also strongly support the proposal that glycoproteomic analysis of urine might reveal mechanisms underpinning CKD.

4. Discussion

CKD is a growing public health problem with dramatic increases in morbidity and mortality following progression to ESRD. Given this, there is a tremendous need for the development of biomarkers to predict CKD progression and allow for early therapeutic intervention. Urine proteomic strategies are now at the forefront of this search due to the sensitivity of MS/MS analysis and the ability to develop noninvasive biomarkers from a readily available biofluid. Significant progress has been made, particularly in diabetes, where urine proteomic analysis can predict nephropathy [6, 25, 26]. Despite these developments, the majority of proteomic studies have relied on two-dimensional (2D) differential in-gel electrophoresis for protein separation. Resulting samples, particularly in CKD subjects, contain large amounts of highly abundant plasma proteins due to nonspecific leakage through the glomerular filtration barrier. Targeted analyses of low-abundance proteins will likely lead to more disease-specific and clinically relevant protein biomarkers.

We therefore focused our attention on the urinary N-linked glycoproteome. Glycoproteins are an important protein subfraction accounting for up to 50% of the human proteome at any given time [27]. Due to their critical role in cell-cell interactions and signaling cascades, glycoproteins are promising markers for identifying kidney disease activity and progression. In this study we present an initial examination of the urinary N-linked glycoproteome in CKD subjects compared to healthy control subjects. We successfully isolated N-linked glycoproteins from twelve urine samples utilizing a hydrazide capture technique. 122 unique glycosylated proteins were detected amongst the twelve subjects (Table 3). This number is similar to other recent glycoproteome analyses. Ahn et al. recently reported isolating 164–174 unique proteins from human diabetic plasma using a multi-lectin column enrichment technique [16]. Yang et al. isolated 265 urinary glycoproteins from bladder cancer subjects and healthy controls also utilizing a multi-lectin column for enrichment, but larger sample sizes were used than in our current study [15]. These

Table 5: Component location and function associated with significantly up- or downregulated proteins identified in CKD patients.

Location					Function				
Gene ontology term	Cluster frequency	*P*-value	FDR	Proteins annotated to the GO Term	Gene ontology term	Cluster frequency	*P* value	FDR	Proteins annotated to the GO term
Extracellular region	18 of 23 proteins, 78.3%	7.36*E*−15	0.00%	EGF, IGHG1, LRG, AMBP, AZGP1, KNG1, SERPING1, TF, LGALS3BP, FETUA, CP, HP, SERPINC1, IGHG2, ORM1, APOD, PTGDS, UMOD	Binding	20 of 23 proteins, 87.0%	0.00051	0.89%	EGF, IGHG1, AMBP, CUBN, AZGP1, KNG1, SERPING1, TF, LGALS3BP, GAA, FETUA, CP, SERPINC1, HP, IGHG2, GLB1, ORM1, APOD, PTGDS, UMOD
Extracellular space	11 of 23 proteins, 47.8%	5.47*E*−11	0.00%	EGF, LGALS3BP, CP, FETUA, SERPINC1, APOD, ORM1, PTGDS, UMOD, KNG1, SERPING1	Protein binding	15 of 23 proteins, 65.2%	0.00066	0.80%	TF, EGF, IGHG1, LGALS3BP, CP, FETUA, AMBP, HP, SERPINC1, APOD, ORM1, GLB1, CUBN, KNG1, SERPING1
Extracellular region part	11 of 23 proteins, 47.8%	1.13*E*−09	0.00%	EGF, LGALS3BP, CP, FETUA, SERPINC1, APOD, ORM1, PTGDS, UMOD, KNG1, SERPING1	Enzyme regulator activity	6 of 23 proteins, 26.1%	0.00024	0.00%	EGF, SERPINC1, FETUA, KNG1, AMBP, SERPING1
Cytoplasmic vesicle	6 of 23 proteins, 26.1%	6.81*E*−05	0.20%	TF, EGF, CUBN, UMOD, KNG1, SERPING1	Endopeptidase inhibitor activity	5 of 23 proteins, 21.7%	3.78*E*−07	0.00%	SERPINC1, FETUA, KNG1, AMBP, SERPING1
Vesicle	6 of 23 proteins, 26.1%	8.57*E*−05	0.18%	TF, EGF, CUBN, UMOD, KNG1, SERPING1	Endopeptidase regulator activity	5 of 23 proteins, 21.7%	4.29*E*−07	0.00%	SERPINC1, FETUA, KNG1, AMBP, SERPING1
Cell fraction	5 of 23 proteins, 21.7%	0.00405	2.29%	EGF, IGHG2, IGHG1, CUBN, AMBP	Peptidase inhibitor activity	5 of 23 proteins, 21.7%	4.84*E*−07	0.00%	SERPINC1, FETUA, KNG1, AMBP, SERPING1
Cytoplasmic membrane-bounded vesicle	5 of 23 proteins, 21.7%	0.00055	0.15%	TF, EGF, CUBN, KNG1, SERPING1	Peptidase regulator activity	5 of 23 proteins, 21.7%	1.10*E*−06	0.00%	SERPINC1, FETUA, KNG1, AMBP, SERPING1
Membrane-bounded vesicle	5 of 23 proteins, 21.7%	0.00062	0.29%	TF, EGF, CUBN, KNG1, SERPING1	Enzyme inhibitor activity	5 of 23 proteins, 21.7%	8.50*E*−06	0.00%	SERPINC1, FETUA, KNG1, AMBP, SERPING1

Table 5: Continued.

Location					Function				
Gene ontology term	Cluster frequency	*P*-value	FDR	Proteins annotated to the GO Term	Gene ontology term	Cluster frequency	*P* value	FDR	Proteins annotated to the GO term
Cytoplasmic vesicle part	4 of 23 proteins, 17.4%	0.00024	0.17%	EGF, CUBN, KNG1, SERPING1	Transporter activity	5 of 23 proteins, 21.7%	0.00527	2.88%	CUBN, APOD, AZGP1, PTGDS, AMBP
Stored secretory granule	4 of 23 proteins, 17.4%	5.33*E*−05	0.22%	TF, EGF, KNG1, SERPING1	Carbohydrate binding	3 of 23 proteins, 13.0%	0.00656	2.50%	SERPINC1, GAA, KNG1
Membrane fraction	4 of 23 proteins, 17.4%	0.00796	2.69%	IGHG2, IGHG1, CUBN, AMBP	Lipid binding	3 of 23 proteins, 13.0%	0.0092	2.82%	APOD, AZGP1, PTGDS
Insoluble fraction	4 of 23 proteins, 17.4%	0.00897	2.81%	IGHG2, IGHG1, CUBN, AMBP	Serine-type endopeptidase inhibitor activity	3 of 23 proteins, 13.0%	9.99*E*−05	0.00%	SERPINC1, AMBP, SERPING1, E9KL23
Apical plasma membrane	3 of 23 proteins, 13.0%	0.00069	0.27%	TF, CUBN, UMOD	Hemoglobin binding	2 of 23 proteins, 8.7%	1.36*E*−05	0.00%	HP, CUBN
Perinuclear region of cytoplasm	3 of 23 proteins, 13.0%	0.00565	2.58%	TF, GLB1, PTGDS	Fatty acid binding	2 of 23 proteins, 8.7%	0.00076	0.73%	AZGP1, PTGDS
Apical part of cell	3 of 23 proteins, 13.0%	0.00146	0.94%	TF, CUBN, UMOD	Cysteine-type endopeptidase inhibitor activity	2 of 23 proteins, 8.7%	0.00084	0.83%	FETUA, KNG1
Lytic vacuole	3 of 23 proteins, 13.0%	0.0026	1.89%	CUBN, GLB1, GAA	Monocarboxylic acid binding	2 of 23 proteins, 8.7%	0.00161	0.92%	AZGP1, PTGDS
Lysosome	3 of 23 proteins, 13.0%	0.0026	1.79%	CUBN, GLB1, GAA	Hydrolase activity, hydrolyzing O-glycosyl compounds	2 of 23 proteins, 8.7%	0.00354	2.29%	GLB1, GAA
Platelet alpha granule lumen	3 of 23 proteins, 13.0%	2.55*E*−05	0.00%	EGF, KNG1, SERPING1	Protein kinase regulator activity	2 of 23 proteins, 8.7%	0.00418	2.13%	EGF, FETUA

Table 5: Continued.

Location					Function				
Gene ontology term	Cluster frequency	*P*-value	FDR	Proteins annotated to the GO Term	Gene ontology term	Cluster frequency	*P* value	FDR	Proteins annotated to the GO term
Secretory granule lumen	3 of 23 proteins, 13.0%	1.37*E*−05	0.00%	EGF, KNG1, SERPING1	Kinase regulator activity	2 of 23 proteins, 8.7%	0.00541	2.71%	EGF, FETUA
Cytoplasmic membrane-bound vesicle lumen	3 of 23 proteins, 13.0%	1.55*E*−05	0.00%	EGF, KNG1, SERPING1	Hydrolase activity, acting on glycosyl bonds	2 of 23 proteins, 8.7%	0.0056	2.78%	GLB1, GAA
Vesicle lumen	3 of 23 proteins, 13.0%	1.74*E*−05	0.00%	EGF, KNG1, SERPING1	Tetrapyrrole binding	2 of 23 proteins, 8.7%	0.00841	2.67%	CUBN, AMBP
Platelet alpha granule	3 of 23 proteins, 13.0%	2.55*E*−05	0.00%	EGF, KNG1, SERPING1	Heparin binding	2 of 23 proteins, 8.7%	0.00608	2.63%	SERPINC1, KNG1
Vacuole	3 of 23 proteins, 13.0%	0.00423	2.27%	CUBN, GLB1, GAA					
Extrinsic to membrane	2 of 23 proteins, 8.7%	0.00532	2.61%	CUBN, UMOD					
Lysosomal membrane	2 of 23 proteins, 8.7%	0.00699	2.64%	CUBN, GAA					
Endocytic vesicle	2 of 23 proteins, 8.7%	0.00332	2.10%	TF, CUBN					
Coated pit	2 of 23 proteins, 8.7%	0.00113	0.88%	TF, CUBN					
Cell projection membrane	2 of 23 proteins, 8.7%	0.00899	2.71%	CUBN, UMOD					

results support a successful hydrazide based technique for glycoprotein isolation in human urine. Further studies are required to identify optimal extraction strategies.

We detected 8 glycoproteins unique to CKD subjects and 35 unique to healthy controls (Table 2). Additionally, of the 122 total proteins identified, 23 glycoproteins were differentially expressed in CKD subjects versus healthy controls. 18 were upregulated in CKD while 5 were downregulated (Table 4). Many of the differentially expressed proteins have been previously linked to kidney disease supporting a potential role as a CKD biomarker. Two of the most significantly upregulated proteins in our CKD samples were AZGP1 and LRG, both of which are established inflammatory mediators. Alteration of AZGP1 and LRG expression is predictive of acute kidney injury in postsurgical patients [28]. AZGP1 has also been shown to be increased in diabetes and diabetic nephropathy [13, 29]. PTGDS, a known extracellular transporter for lipophilic molecules, is formed *de novo* in renal tubules [30]. PTGDS is upregulated in early diabetes [31] and is a marker of hypertension and latent renal injury [32]. SERPING1, an extracellular matrix regulator, is increased in acute renal allograft rejection perhaps suggesting an important role for collagen remodeling [33]. KNG1, a bradykinin precursor, has also been shown to be upregulated in acute renal allograft rejection [34], and gene variation induces altered aldosterone sensitivity in hypertensive subjects [35]. Interestingly, LUM, a proteoglycan, is a protein unique to CKD but without statistically significant upregulation. Altered regulation of LUM has been linked with abnormal collagen fibril morphology as a mediator of fibrotic disease in diabetic nephropathy [36, 37]. CUBN, an apical protein in proximal tubule cells, was unique and downregulated in CKD. Recent investigation supports a role of CUBN in albumin reabsorption with genetic variance at this locus predicting microalbuminuria [38]. The decreased urinary CUBN excretion found in our CKD population may represent a dysfunctional variant or potentially a causative factor responsible for increasing proteinuria.

We used annotations by the GO Consortium and GO Tools to connect the complex array of proteins identified in CKD subjects to biological processes, protein function, and cellular location. Many of the multiprotein pathways differentially expressed in CKD are involved in coagulation, inflammation, and acute-phase response (Table 5, Figure 3). Twenty proteins were linked to protein-protein interactions and binding. Remarkably, there were altered levels of proteins that were involved in acute-phase response and immune/stress response proteins (18 out of 23), implicating a possible mechanistic role for these pathways in CKD. Our detection of the several extracellular proteins and matrix remodeling proteases likely reflects matrix remodeling that occurs in CKD. These findings are consistent with previous literature, as CKD is known to have increased propensity for atherosclerosis, endothelial dysfunction, increased basal inflammation, and altered stress response [39, 40].

In this study, we have established normalization techniques which will be essential to future urine glycoproteome analyses. To account for variations in the glycoprotein extraction efficiency of individual samples, yeast invertase (yeast glycoprotein with several glycopeptides) was added to each sample prior to extraction. In this way, glycopeptides derived from invertase serve as an internal marker for the extraction efficiency in each sample. Our samples were also normalized for urine creatinine content. This is of particular importance as marked intersubject variability can exist in creatinine excretion in random urine specimens consistent with different concentrations due to hydration status. Indeed, such normalization would be essential to extrapolate net excretion rates of a given protein in 24 hours and is commonly employed in clinical practice to quantify albumin excretion rates [23].

In summary, we have utilized a hydrazide-based approach to enrich the urinary glycoproteome with subsequent identification of the urinary glycoproteins in a human CKD population for the first time. Our results indicate that urine carries a distinct population of glycoproteins that function in proteinase inhibition, protein binding, and the acute-phase/immune-stress response in subjects with CKD. It will be of interest to study a larger number of subjects to determine whether urinary levels of these proteins might be useful indicators of CKD and to investigate the proposal that these proteins could be markers of disease progression.

Abbreviations

CKD:	Chronic kidney disease
ESRD:	End-stage renal disease
FDR:	False discovery rate
GO:	Gene ontology
LC-ESI-MS/MS:	Liquid chromatography-electrospray ionization tandem MS analysis
MS:	Mass spectrometry.

Acknowledgments

This work is supported in part by Grants from the National Institutes of Health (HL094230, DK079912, HL102334, GM49500, CA108597, DK082841, and K12HD001438), the Doris Duke Foundation Clinical Scientist Development Award (S. Pennathur), the American Diabetes Association (1-08-CR-48; R. Pop-Busui), the George O'Brien Kidney Center (DK089503), and the Molecular Phenotyping Core of the Michigan Nutrition and Obesity Research Center (DK089503). A. Vivekanandan-Giri and J. L. Slocum contributed equally to this work.

References

[1] J. Coresh, E. Selvin, L. A. Stevens et al., "Prevalence of chronic kidney disease in the United States," *Journal of the American Medical Association*, vol. 298, no. 17, pp. 2038–2047, 2007.

[2] United States Renal Data System, "Atlas of chonic kidney disease and end-stage renal disease in the United States," Annual Data Report USRDS 2010, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Md, USA, 2010.

[3] P. Ruggenenti, A. Perna, L. Mosconi, R. Pisoni, and G. Remuzzi, "Urinary protein excretion rate is the best independent predictor of ESRF in non-diabetic proteinuric chronic

nephropathies," *Kidney International*, vol. 53, no. 5, pp. 1209–1216, 1998.

[4] D. Fliser, J. Novak, V. Thongboonkerd et al., "Advances in urinary proteome analysis and biomarker discovery," *Journal of the American Society of Nephrology*, vol. 18, no. 4, pp. 1057–1071, 2007.

[5] J. Klein, P. Kavvadas, N. Prakoura et al., "Renal fibrosis: insight from proteomics in animal models and human disease," *Proteomics*, vol. 11, no. 4, pp. 805–815, 2011.

[6] K. Rossing, H. Mischak, M. Dakna et al., "Urinary proteomics in diabetes and CKD," *Journal of the American Society of Nephrology*, vol. 19, no. 7, pp. 1283–1290, 2008.

[7] M. L. Merchant, B. A. Perkins, G. M. Boratyn et al., "Urinary peptidome may predict renal function decline in type 1 diabetes and microalbuminuria," *Journal of the American Society of Nephrology*, vol. 20, no. 9, pp. 2065–2074, 2009.

[8] H. A. Shui, T. H. Huang, S. M. Ka, P. H. Chen, Y. F. Lin, and A. Chen, "Urinary proteome and potential biomarkers associated with serial pathogenesis steps of focal segmental glomerulosclerosis," *Nephrology Dialysis Transplantation*, vol. 23, no. 1, pp. 176–185, 2008.

[9] M. Haubitz, S. Wittke, E. M. Weissinger et al., "Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy," *Kidney International*, vol. 67, no. 6, pp. 2313–2320, 2005.

[10] E. Banon-Maneus, F. Diekmann, M. Carrascal et al., "Two-dimensional difference gel electrophoresis urinary proteomic profile in the search of nonimmune chronic allograft dysfunction biomarkers," *Transplantation*, vol. 89, no. 5, pp. 548–558, 2010.

[11] R. P. Woroniecki, T. N. Orlova, N. Mendelev et al., "Urinary proteome of steroid-sensitive and steroid-resistant idiopathic nephrotic syndrome of childhood," *American Journal of Nephrology*, vol. 26, no. 3, pp. 258–267, 2006.

[12] X. Zhang, M. Jin, H. Wu et al., "Biomarkers of lupus nephritis determined by serial urine proteomics," *Kidney International*, vol. 74, no. 6, pp. 799–807, 2008.

[13] S. Jain, A. Rajput, Y. Kumar, N. Uppuluri, A. S. Arvind, and U. Tatu, "Proteomic analysis of urinary protein markers for accurate prediction of diabetic kidney disorder," *Journal of Association of Physicians of India*, vol. 53, pp. 513–520, 2005.

[14] J. B. Lowe, "Glycosylation, immunity, and autoimmunity," *Cell*, vol. 104, no. 6, pp. 809–812, 2001.

[15] N. Yang, S. Feng, K. Shedden et al., "Urinary glycoprotein biomarker discovery for bladder cancer detection using LC/MS-MS and label-free quantification," *Clinical Cancer Research*, vol. 17, no. 10, pp. 3349–3359, 2011.

[16] J. M. Ahn, B. G. Kim, M. H. Yu, I. K. Lee, and J. Y. Cho, "Identification of diabetic nephropathy-selective proteins in human plasma by multi-lectin affinity chromatography and LC-MS/MS," *Proteomics—Clinical Applications*, vol. 4, no. 6-7, pp. 644–653, 2010.

[17] H. Zhang, J. Saha, J. Byun et al., "Rosiglitazone reduces renal and plasma markers of oxidative injury and reverses urinary metabolite abnormalities in the amelioration of diabetic nephropathy," *American Journal of Physiology—Renal Physiology*, vol. 295, no. 4, pp. F1071–F1081, 2008.

[18] H. Zhang, X. J. Li, D. B. Martin, and R. Aebersold, "Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry," *Nature Biotechnology*, vol. 21, no. 6, pp. 660–666, 2003.

[19] T. Vaisar, S. Pennathur, P. S. Green et al., "Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL," *Journal of Clinical Investigation*, vol. 117, no. 3, pp. 746–756, 2007.

[20] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology. The gene ontology consortium," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[21] E. I. Boyle, S. Weng, J. Gollub et al., "GO::termfinder—open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes," *Bioinformatics*, vol. 20, no. 18, pp. 3710–3715, 2004.

[22] R. C. Miller, E. Brindle, D. J. Holman et al., "Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations," *Clinical Chemistry*, vol. 50, no. 5, pp. 924–932, 2004.

[23] G. Eknoyan, T. Hostetter, G. L. Bakris et al., "Proteinuria and other markers of chronic kidney disease: a position statement of the National Kidney Foundation (NKF) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)," *American Journal of Kidney Diseases*, vol. 42, no. 4, pp. 617–622, 2003.

[24] M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, and T. Ideker, "Cytoscape 2.8: new features for data integration and network visualization," *Bioinformatics*, vol. 27, no. 3, pp. 431–432, 2011.

[25] K. Sharma, S. Lee, S. Han et al., "Two-dimensional fluorescence difference gel electrophoresis analysis of the urine proteome in human diabetic nephropathy," *Proteomics*, vol. 5, no. 10, pp. 2648–2655, 2005.

[26] P. V. Rao, X. Lu, M. Standley et al., "Proteomic identification of urinary biomarkers of diabetic nephropathy," *Diabetes Care*, vol. 30, no. 3, pp. 629–637, 2007.

[27] D. Vanderschaeghe, N. Festjens, J. Delanghe, and N. Callewaert, "Glycome profiling using modern glycomics technology: technical aspects and applications," *Biological Chemistry*, vol. 391, no. 2-3, pp. 149–161, 2010.

[28] F. Aregger, C. Pilop, D. E. Uehlinger et al., "Urinary proteomics before and after extracorporeal circulation in patients with and without acute kidney injury," *Journal of Thoracic and Cardiovascular Surgery*, vol. 139, no. 3, pp. 692–700, 2010.

[29] S. Riaz, S. S. Alam, S. K. Srai, V. Skinner, A. Riaz, and M. W. Akhtar, "Proteomic identification of human urinary biomarkers in diabetes mellitus type 2," *Diabetes Technology and Therapeutics*, vol. 12, no. 12, pp. 979–988, 2010.

[30] N. Nagata, K. Fujimori, I. Okazaki et al., "De novo synthesis, uptake and proteolytic processing of lipocalin-type prostaglandin D synthase, β-trace, in the kidneys," *FEBS Journal*, vol. 276, no. 23, pp. 7146–7158, 2009.

[31] H. Oda, Y. Shiina, K. Seiki, N. Sato, N. Eguchi, and Y. Urade, "Development and evaluation of a practical ELISA for human urinary lipocalin-type prostaglandin D synthase," *Clinical Chemistry*, vol. 48, no. 9, pp. 1445–1453, 2002.

[32] N. Hirawa, Y. Uehara, M. Yamakado et al., "Lipocalin-type prostaglandin D synthase in essential hypertension," *Hypertension*, vol. 39, no. 2, part 2, pp. 449–454, 2002.

[33] X. B. Ling, T. K. Sigdel, K. Lau et al., "Integrative urinary peptidomics in renal transplantation identifies biomarkers for acute rejection," *Journal of the American Society of Nephrology*, vol. 21, no. 4, pp. 646–653, 2010.

[34] G. V. Cohen Freue, M. Sasaki, A. Meredith et al., "Proteomic signatures in plasma during early acute renal allograft rejection," *Molecular and Cellular Proteomics*, vol. 9, no. 9, pp. 1954–1967, 2010.

[35] M. Barbalic, G. L. Schwartz, A. B. Chapman, S. T. Turner, and E. Boerwinkle, "Kininogen gene (KNG) variation has a

consistent effect on aldosterone response to antihypertensive drug therapy: the GERA study," *Physiological Genomics*, vol. 39, no. 1, pp. 56–60, 2009.

[36] S. Chakravarti, T. Magnuson, J. H. Lass, K. J. Jepsen, C. LaMantia, and H. Carroll, "Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican," *Journal of Cell Biology*, vol. 141, no. 5, pp. 1277–1286, 1998.

[37] L. Schaefer, I. Raslik, H. J. Grone et al., "Small proteoglycans in human diabetic nephropathy: discrepancy between glomerular expression and protein accumulation of decorin, biglycan, lumican, and fibromodulin," *FASEB journal*, vol. 15, no. 3, pp. 559–561, 2001.

[38] C. A. Böger, M. H. Chen, A. Tin et al., "CUBN is a gene locus for albuminuria," *Journal of the American Society of Nephrology*, vol. 22, no. 3, pp. 555–570, 2011.

[39] A. S. Go, G. M. Chertow, D. Fan, C. E. McCulloch, and C. Y. Hsu, "Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization," *The New England Journal of Medicine*, vol. 351, no. 13, pp. 1296–1305, 2004.

[40] A. Recio-Mayoral, D. Banerjee, C. Streather, and J. C. Kaski, "Endothelial dysfunction, inflammation and atherosclerosis in chronic kidney disease—a cross-sectional study of predialysis, dialysis and kidney-transplantation patients," *Atherosclerosis*, vol. 216, no. 2, pp. 446–451, 2011.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 502845, 8 pages
doi:10.1155/2011/502845

Research Article

Adsorption of Urinary Proteins on the Conventionally Used Urine Collection Tubes: Possible Effects on Urinary Proteome Analysis and Prevention of the Adsorption by Polymer Coating

Iwao Kiyokawa,[1,2] Kazuyuki Sogawa,[1] Keiko Ise,[3] Fumie Iida,[1] Mamoru Satoh,[1] Toshihide Miura,[1,2] Ryo Kojima,[1,2] Katsuhiro Katayama,[1,2] and Fumio Nomura[1,3,4]

[1] *Clinical Proteomics Research Center, Chiba University Hospital, Chiba, Japan*
[2] *R&D Department, Nittobo Medical Co., Ltd., Fukushima, Koriyama City 963-8061, Japan*
[3] *Department of Clinical Laboratory, Chiba University Hospital, Chiba, Japan*
[4] *Department of Molecular Diagnosis (F8), Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan*

Correspondence should be addressed to Fumio Nomura, fnomura@faculty.chiba-u.jp

Received 1 June 2011; Accepted 5 July 2011

Academic Editor: David E. Misek

Copyright © 2011 Iwao Kiyokawa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One possible factor determining recovery of trace amount of protein biomarker candidates during proteome analyses could be adsorption on urine tubes. This issue, however, has not been well addressed so far. Recently, a new technical device of surface coating by poly(2-methacryloyloxyethyl phosphorylcholine (MPC)-*co*-n-butyl methacrylate (BMA)) (poly(MPC-*co*-BMA)) has been developed mainly to prevent the adsorption of plasma proteins. We assessed whether conventionally used urine tubes adsorb trace amount of urinary proteins and, if any, whether the surface coating by poly(MPC-*co*-BMA) can minimize the adsorption. Proteinuric urine samples were kept in poly(MPC-*co*-BMA)-coated and noncoated urine tubes for 15 min and possibly adsorbed proteins and/or peptides onto urine tubes were analyzed by SDS-PAGE, 2-DE, and the MALDI-TOF MS. It was found that a number of proteins and/or peptides adsorb on the conventionally used urine tubes and that surface coating by poly(MPC-*co*-BMA) can minimize the adsorption without any significant effects on routine urinalysis test results. Although it remains to be clarified to what extent the protein adsorption can modify the results of urinary proteome analyses, one has to consider this possible adsorption of urinary proteins when searching for trace amounts of protein biomarkers in urine.

1. Introduction

Urine has now become one of the most attractive biological fluids in clinical proteomics [1, 2]. A number of urinary proteomic studies have been conducted and have revealed urinary biomarker candidates for renal systemic diseases and malignancies of urinary tract [3–5]. Proteomic analysis of urines can be applied to biomarker search in nonrenal diseases as well [6–8].

Although urinary proteome analyses have been conducted by various gel-based and gel-free techniques [9], comprehensive urinary proteome analysis is not an easy task because the urine has very diluted protein concentration with high levels of salts. Sample preparation, processing, and storage for urinary proteomics have been reviewed [10–13]. More recently, an optimized quantitative proteomic strategy for urine biomarker discovery was described [14]. In any event, maximal protein recovery from urine is essential for detecting trace quantities of proteins present in urine for potential biomarker discovery. For this purpose, protein loss during sample preparation should be avoided. One possible factor responsible for loss of trace amounts of urinary proteins could be adsorption to sample tubes, but this issue has not been well addressed so far to our knowledge.

Many attempts have been made to prevent the adsorption of plasma proteins and to improve blood compatibility by surface modification [15, 16]. Ishihara and coworkers reported on rapid development of hydrophilicity and protein adsorption resistance by polymer surfaces bearing poly(2-methacryloyloxyethyl phosphorylcholine (MPC)-*co*-n-butyl methacrylate) (poly(MPC-*co*-BMA)) [17, 18].

We took advantage of this coating method in the present study and assessed whether conventionally used tubes adsorb trace amount of urinary proteins and, if any, whether the surface coating by poly(MPC-*co*-BMA) can minimize the adsorption.

2. Materials and Methods

2.1. Urine Collection Tubes and Coating Method. Poly(MPC-*co*-BMA) was obtained from AI BIO-CHIPS CO., LTD (Tokyo, Japan). A total of 6 different types of conventional urine collection tubes were used in this study. Tubes made from polystyrene (PS) (Cat# 10200), polypropylene(PP) Cat# 72200, polyethylene terephthalate(PET) (Cat# 23540), and styrene-butadiene copolymers/methyl methacrylate-styrene (SBC/MS) (Cat# 17300) were purchased from TOYO KAGAKU KIZAI Co., LTD., Japan. Tubes made from acrylonitrile-styrene (AS) copolymers (Cat# 479511373) were from Nittobo Medical Co., LTD., Japan and those made from styrene-butadiene copolymers (SBC) (Cat# 3324A000A-10) were from ASIAKIZAI Co., LTD., Japan. The conventional tubes made by AS were coated by poly(MPC-*co*-BMA) as described by Futamura et al. [18].

2.2. Samples. Urine samples obtained from outpatients in Chiba University Hospital were used. An aliquot of the samples was taken for routine urinalysis, and the rest of the samples were centrifuged (700 ×g, 5 min at room temperature), and the supernatant was subjected to assess protein adsorption on test tubes as described below. All these procedures were carried out within 2 hours after collection of the samples.

2.3. Urine Sample Preparation for Electrophoresis (SDS-PAGE and 2-DE) and MALDI-TOF MS. One mL of two different levels of pooled proteinuric urines (equivalent to 15 mg/dL and 50 mg/dL, resp.) obtained from 10 patients with renal disease were put into poly(MPC-*co*-BMA)-coated and noncoated urine correction tubes and were kept at room temperature for 15 min. After aspiration of the urines, the tubes were washed with 200 μL of PBS three times. After the third wash and PBS being aspirated, 100 μL of PAGE sample buffer (electrophoresis) (50 mM Tris-HCl, pH 6.8 containing 50 mM dithiothreitol, 0.5% SDS, and 10% glycerol) or 1% TCA aqueous solution (MALDI-TOF MS analysis) was added and the tubes were vortexed for 30 sec to dissolve possibly adsorbed proteins.

2.4. Gel-Based Analysis. The solution which contained proteins possibly adsorbed on the urine tubes was then analyzed using SDS-PAGE (Perfect NT Gel W, 10–20% acrylamide, 20 wells; DRC Co., Ltd., Tokyo, Japan) according to the manufacturer's protocol. The gel was stained with CBB (PhastGel Blue R; GE Healthcare, Little Chalfont, UK). The proteins separated by SDS-PAGE were identified by in-gel tryptic digestion of the proteins followed by MS. In-gel tryptic digestion was performed as described previously [19]. Molar quantities of recovered peptide fragments were estimated from the staining intensity of the SDS-PAGE bands that were digested in-gel with trypsin. Digested peptides roughly equivalent up to 1 pmol of protein were injected into a trap column: 0.3 × 5 mm L-trap column (Chemicals Evaluation and Research Institute, Saitama, Japan), and an analytical column: 0.1 × 50 mm Monolith column (AMR, Tokyo, Japan), which was attached to a HPLC system (Nanospace SI-2; Shiseido Fine Chemicals, Tokyo, Japan). The flow rate of a mobile phase was 1 μL/min. The solvent composition of the mobile phase was programmed to change in 35 min cycles with varying mixing ratios of solvent A (2% v/v CH_3CN and 0.1% v/v HCOOH) to solvent B (90% v/v CH_3CN and 0.1% v/v HCOOH): 5–50% B 20 min, 50–95% B 1 min, 95% B 3 min, 95–5% B 1 min, 5% B 10 min. Purified peptides were introduced from HPLC to an LTQ-XL (Thermo Scientific, Calif, USA), an ion trap mass spectrometer (ITMS), via an attached Pico Tip (New Objective, Mass, USA). The MS and MS/MS peptide spectra were measured in a data-dependent manner according to the manufacturer's operating specifications. The Mascot search engine (Matrix science, London, UK) was used to identify proteins from the mass and tandem mass spectra of peptides. Peptide mass data were matched by searching the Human International Protein Index database (IPI, July 2008, 72079 entries, European Bioinformatics Institute) using the MASCOT engine. The minimum criterion of the probability-based MASCOT/MOWSE score was set with 5% as the significant threshold level.

For 2-DE analysis, we used the method described by Oh-Ishi et al. [20] and Kawashima et al. [21]. Briefly, one mL aliquots of urine samples kept at room temperature for 15 min in poly(MPC-*co*-BMA)-coated and noncoated urine correction tube were concentrated up to 20-fold by BJP Concentrator (ProChem, MA, USA) to 50 μL and lyophilized. The lyophilizate was resuspended by 200 μL of Immobiline reagent (7 M urea, 2 M thiourea, 4% CHAPS, 2% DTT, 2% Pharmalyte, broad range pH 3–10). Finally, 50 μL of the 5-fold urine sample was applied to the IEF agarose gel. The agarose gel was then transferred to the Perfect NT Gels W (10–20% gradient of polyacrylamide gel; DRC. Co. Ltd, Tokyo, Japan) and the second electrophoresis was performed. Protein spots on 2-DE gels were stained with CBB. The protein spots were detected, quantified, and matched with the 2-DE gel view analysis software, Progenesis SameSpots (Nonlinear Dynamics Ltd., UK). The protein spots were excised from the gel and identified, as we previously described [19].

2.5. MS-Based Analysis. One mL aliquots of urine samples (containing 50 mg/mL protein) were kept at room temperature for 15 min in poly(MPC-*co*-BMA)-coated and noncoated urine collection tube. Proteins possibly adsorbed on the tubes were collected as described above for the gel-based method and were analyzed by the MALDI-TOF MS. To obtain quantitative data of the possibly adsorbed proteins, we used stable isotope-labeled 5.9 kDa fibrinogen alpha C chain fragment (FIC 5.9) as an internal standard as described by Sogawa et al. [22]. We obtained the stable isotope-labeled synthetic FIC 5.9

Table 1: Proteins adsorbed on poly(MPC-*co*-BMA)-uncoated urine tubes.

No.	ID	M.W.[a]	Score	Queries matched	pI[b]
1	Tamm-Horsfall urinary glycoprotein	69,761 Da	250	15	4.96
2	Albumin	69,321 Da	353	32	5.67
3	Semenogelin-1	52,131 Da	114	6	9.26
4	Alpha-1-antichymotrypsin	47,651 Da	281	11	5.32
5	Alpha-1-antitrypsin	46,737 Da	85	3	5.37
6	Apolipoprotein A1	27,891 Da	82	4	5.27
7	IGKV1-5 protein	25,765 Da	313	7	5.74–6.30
8	Prostaglandin-H2 D-isomerase	21,029 Da	98	5	8.37
9	Apolipoprotein C3	10,846 Da	93	3	4.72
10	Protein S100-A8	10,835 Da	86	3	6.51
11	SH3 domain-binding glutamic acid-rich-like protein 3	10,438 Da	80	2	4.82

[a,b] Theoretical Mr and pI, as resulted from Compute pI/Mw tool of Expasy (http://us.expasy.org/tools/pi_tool.html), are also indicated.

from the AnyGen Co., Ltd. (Kwangju, Korea). The amino acid sequence of the peptide was SSSYSKQFTSSTSYNRG DSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPV. (The underlined amino acids were synthesized with ^{13}C, ^{15}N uniformly labeled FMOC amino acids.). In urine analysis, ten microliters of SID (stable isotope-labeled) -FIC 5.9 solution (0.5 pmol/μL SID-FIC 5.9, MB-WCX binding solution) and 5 μL of urine sample were transferred to a 200 μL PCR tube (Thermo Fisher Scientific K.K., Kanagawa, Japan). In analysis of urine samples kept in tube, ten microliters of SID-FIC 5.9 solution (0.025 pmol/μL SID-FIC 5.9, MB-WCX binding solution) and 5 μL of extracted samples (urine samples kept in tube) were transferred to a 200 μL PCR tube. A 10 μL homogenous magnetic particle solution was added, mixed with the other solutions, and allowed to sit for 5 min. The PCR tubes were placed in a magnetic bead separator (MBS; Bruker Daltonics GmbH) for 30 s for magnetic fixation of the MB-WCX particles. The supernatant was aspirated, and the tubes were removed from the MBS device. We added 100 μL of the washing solution and carefully mixed it with the magnetic beads. We then replaced the tube into the MBS device and moved it back and forth between adjacent wells on each side of the device's magnetic bar. After fixation of the magnetic beads for 30 s, the supernatant was aspirated. We repeated this washing procedure three times. After the final wash, we eluted the bound molecules by incubating them for 1 min with 5 μL MB-WCX elution solution and then used the MBS device to collect the eluate. For the final step, we added 5 μL of the MB-WCX stabilization solution to the eluate. We then mixed 1 μL of the eluate with 5 μL of a matrix solution (0.3 g/L a-cyano-4-hydroxycinnamic acid in ethanol:acetone, 2 : 1). We spotted 1 μL of the mixture onto an AnchorChip target plate (Bruker Daltonics GmbH) and allowed it to dry. Protein Calibration standard (Protein Calibration standard 1, Bruker Daltonics GmbH) was dissolved in 125 1 μL. We applied 0.5 1 μL of the solution to target spots in proximity to the urine samples for external calibration.

We placed the AnchorChip target plate into the AutoFlex II TOF/TOF mass spectrometer (Bruker Daltonics GmbH) and into the UltraFlex III TOF/TOF mass spectrometer (Bruker Daltonics GmbH), which is controlled by Flexcontrol software 3.0 (Bruker Daltonics GmbH). The instrument was externally calibrated by standard procedures. The automated acquisition method included in the instrument software generated all acquisitions. The automated acquisition laser power was set between 25% and 35%. Spectra were acquired in a positive linear mode in the mass range of 600 to 10,000 Da.

We used FlexAnalysis software 3.0 to perform baseline correction and smoothing. The concentration of the proteins adsorbed to urine collection tube was estimated from the ratio of the peak intensity of adsorbed proteins to the peak intensity of SID-FIC 5.9. For identification of peptides as we previously described [23], the AnchorChip target plate was also placed in an UltrafleXtreme TOF/TOF mass spectrometer (Bruker Daltonics) and the MALDI-TOF/TOF MS/MS spectrum was recorded in LIFT mode. Five hundred laser shots from a total of 3000 laser shots were summed. The MALDI-TOF/TOF MS/MS spectrum was subjected to a database search using the Mascot (Matrix Science, London, UK) database search engine. The search parameters were as follows: no enzyme specificity, 25 ppm mass tolerance for the parent mass, and 0.2 Da for fragment masses. No fixed or variable modifications were selected. The NCBInr database was used for the search.

2.6. Urinalysis Testing

2.6.1. Quantitative Study. One hundred urine samples requested for urinalysis on routine basis at the Division of Laboratory Medicine and Clinical Genetics, Chiba University Hospital were used. The urine samples were aliquoted (10 mL) to poly(MPC-*co*-BMA)-coated and noncoated collection tubes and were kept at room temperature for 15 min before use. Nine different quantitative urinalysis such as protein, glucose, creatinine, microalbumin, beta-2-microglobulin, amylase, *N*-acetyl-β-D-glucosaminidase, urea nitrogen, uric acid, and six kinds of electrolytes were conducted using BioMajesty JCA-BM6010 (JEOL Ltd., Tokyo, Japan).

Table 2: Proteins which were reduced when kept in poly(MPC-*co*-BMA) noncoated urine tubes.

No.	ID	M.W.[a]	Score	Queries matched	pI[b]
1	Ceruloplasmin	122,128 Da	812	52	5.41
2	Lysosomal alpha-glucosidase	105,271 Da	152	8	5.41
3	Alpha-N-acetylglucosaminidase	82,115 Da	163	6	6.21
4	Serotransferrin	77,000 Da	1500	95	6.70
5	Alpha-1-antitrypsin	46,707 Da	253	12	5.37
6	Cell adhesion molecule 4	42,759 Da	134	6	5.59
7	Prostate-specific antigen	28,723 Da	80	2	7.26
8	IGK protein	26,218 Da	1416	4	5.74–6.30
9	IGL protein	24,777 Da	430	31	5.74–6.30
10	Alpha-1-acid glycoprotein 1	23,497 Da	239	7	5.00
11	Prostaglandin 2D synthase	22,932 Da	596	31	7.66
12	Prostaglandin-H2 D-isomerase	21,029 Da	861	42	7.66
13	Transthyretin	15,877 Da	856	37	5.35
14	Rheumatoid factor D5 light chain	12,758 Da	273	6	5.74–6.30
15	Rheumatoid factor D6 light chain	12,520 Da	273	6	5.74–6.30

[a,b] Theoretical Mr and pI, as resulted from Compute pI/Mw tool of Expasy (http://us.expasy.org/tools/pi_tool.html), are also indicated.

2.6.2. Dipstick Urinalysis. The Uriflet S-9UB (Arkray Inc., Tokyo, Japan) and AUTION MAX AX-4030 (Arkray Inc., Tokyo, Japan) analyzers were used. Ten different parameters are assessed: specific gravity (SG, measured via a built-in refractometer), erythrocytes, leukocytes, nitrite, pH, protein, glucose, ketones, bilirubin, and urobilinogen.

2.6.3. Urinary Sediments. The AUTION IQ IQ-5210 analyzer (Arkray Inc., Tokyo, Japan) was used to determine urinary sediments. This equipment includes digital imaging and Auto-Particle Recognition (APR) (Chatsworth, CA, USA) software to classify urine particles and quantitatively report results. In this study, 4 categories red blood cells (RBC), white blood cells (WBC), squamous epithelial cells (SEC), and casts were classified by the APR software.

2.7. Statistical Analysis. The numerical data are presented as the mean ± standard deviation (SD). We evaluated the statistical significance using IBM SPSS Statistics 18 software (SPSS Inc., IL, USA). A $P < 0.05$ was considered significant using the Mann-Whitney U test.

3. Results

3.1. Detection of Urinary Proteins Adsorbed on Urine Tubes by SDS-PAGE. Proteins adsorbed on the poly(MPC-*co*-BMA)-coated and noncoated tubes were analyzed by SDS-PAGE. As shown in Figure 1(a), a few distinct protein bands (60 kDa, 66 kDa and 100 kDa) were noted in samples obtained from noncoated AS tubes. No clear bands were visible in samples obtained from poly(MPC-*co*-BMA)-coated AS tubes under these experimental conditions. LC-MS analysis of trypsin digests of these bands identified 11 proteins as listed in Table 1. Protein adsorption on the tubes was observed in 6 different types of conventionally used urine collection tubes as shown in Figure 1(b).

3.2. 2-DE Analysis of Urine Samples Kept in Poly(MPC-co-BMA)-Coated and Noncoated Tubes. Urine specimens kept in poly(MPC-*co*-BMA)-coated and noncoated tubes were subjected to the agarose 2-DE as described in the Methods section. The representative patterns were presented in Figure 1(c). Nine protein spots the intensities of which were significantly greater ($P < 0.05$) in samples kept at poly(MPC-*co*-BMA)-coated tubes compared with those kept at noncoated tubes were selected based on the results obtained in seven different experiments.

These differences were most likely as the results of more protein adsorption on noncoated tubes. These spots were excised and subjected to in-gel trypsin digestion followed by LC-MS. A total of 15 proteins were identified as listed in Table 2.

3.3. MALDI-TOF MS Analysis of Urinary Proteins Adsorbed on Poly(MPC-co-BMA)-Coated and Noncoated Tubes. Proteins and/or peptides adsorbed on the conventional urine tubes were also detectable by the MALDI-TOF MS. Figure 2(a) shows a representative spectrum of the adsorbed proteins and peptides. The intensities of the two peaks (2556 m/z and 2654 m/z) were notably greater in samples obtained from poly(MPC-*co*-BMA) noncoated tubes. Similar results were obtained in 7 different experiments; the expression levels of the two peaks expressed as the ratio to the internal standard, SID-FIC 5.9, were significantly greater ($P < 0.001$) in poly(MPC-*co*-BMA) noncoated tubes than in coated tubes. Using MALDI-TOF/TOF MS/MS technology, we successfully identified the two peaks (2556 m/z and 2654 m/z) as internal sequences of the fibrinogen alpha C chain fragment. The peptide sequences of the 2556 m/z and 2654 m/z were DEAGSEADHEGTHSTKRGHAKSRP and DEAGSEADHEGTHSTKRGHAKSRPV, respectively. The mean value of the ratio of m/z 2654 to SID-FIC5.9 was 5.70 (Figure 2(a) Right panel) in poly(MPC-*co*-BMA) noncoated

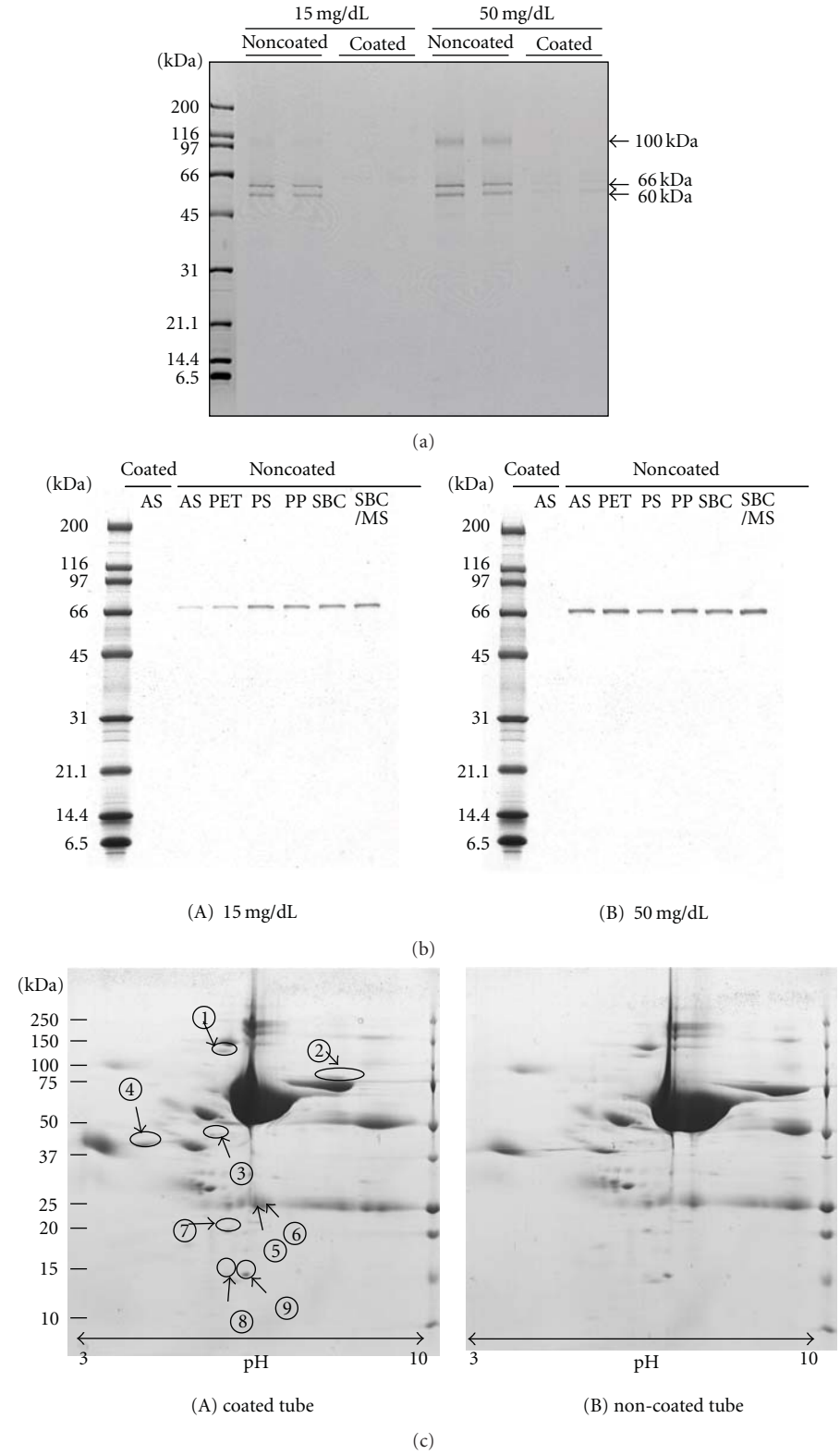


Figure 1: (a) SDS-PAGE of adsorbed proteins to poly(MPC-*co*-BMA)-coated and noncoated urine tubes in two different grades of proteinuric samples. A few distinct protein bands are noted in samples obtained from noncoated tubes. Similar results were obtained in 9 different experiments. (b) SDS-PAGE of adsorbed proteins to 6 different types of poly(MPC-*co*-BMA) noncoated conventionally used urine collection tubes. AS: poly(acrylonitrile-styrene), PET: polyethylene terephthalate, PS: polystyrene, PP: polypropylene, SBC: styrene-butadiene, and SBC/MS: styrene-butadiene copolymers/methyl methacrylate-styrene. (c) 2-DE of urinary proteins obtained from samples kept at poly(MPC-*co*-BMA)-coated and noncoated urine tubes. Nine protein spots the intensities of which were significantly greater ($P < 0.05$) in samples kept at poly(MPC-*co*-BMA)-coated tubes compared with those kept at noncoated tubes were selected based on the results obtained in seven different experiments. The 2-DE gels are shown for pH 3–10.

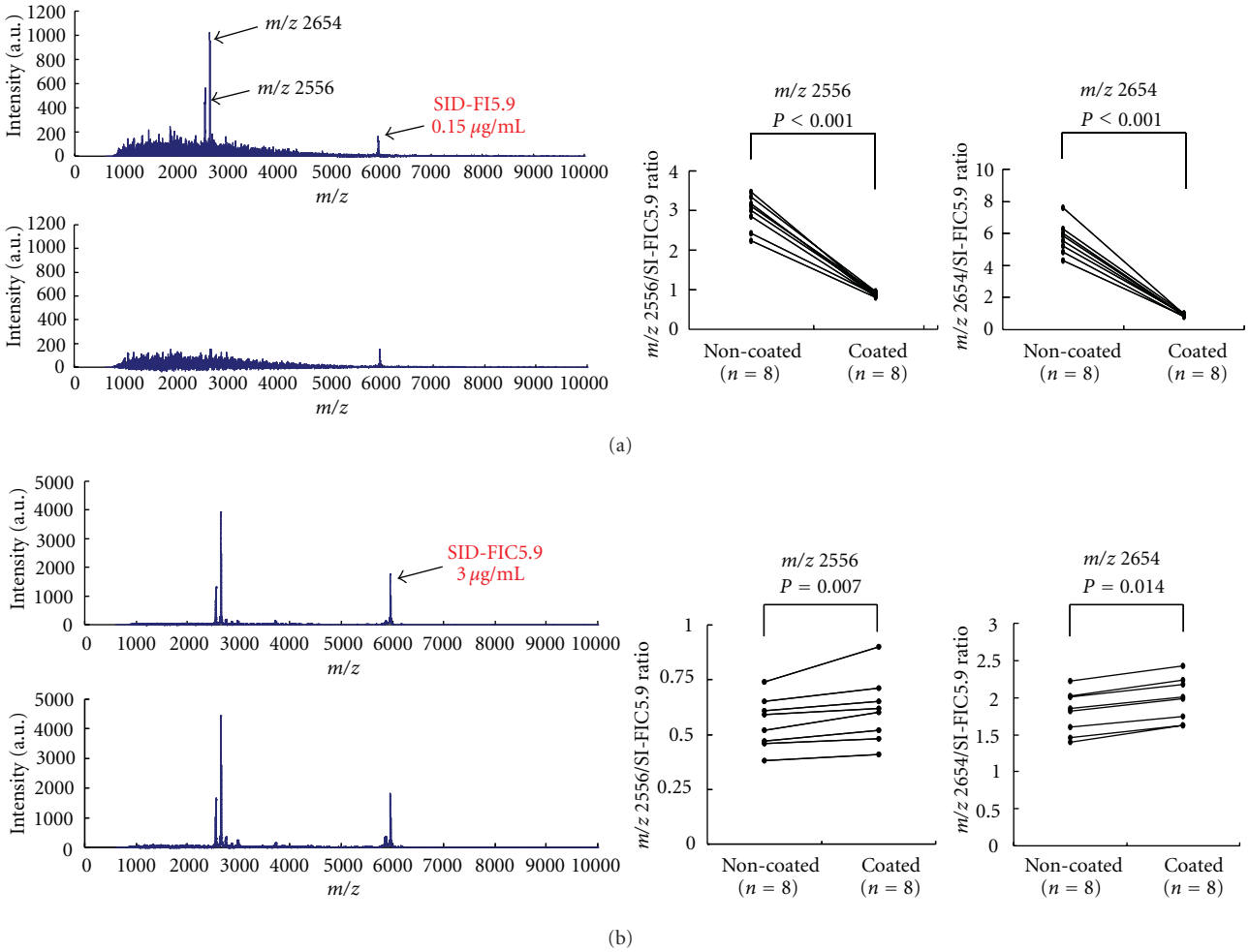


Figure 2: Representative spectra of the MALDI-TOF MS of urinary proteins and/or peptides adsorbed on urinary tubes with and without poly(MPC-*co*-BMA) coating (a) and those of urine samples kept in the coated and noncoated tubes (b). (a) Left: it was notable that 2556 m/z and 2645 m/z peaks were observed in proteinuric samples only when obtained from the poly(MPC-*co*-BMA) non-coated tubes (the upper panel). Similar results were obtained in 7 different experiments. Right: the two peaks detectable in noncoated tubes were significantly attenuated in poly(MPC-*co*-BMA)-coated tubes ($P < 0.001$). (b) When urinary samples were kept in the coated and uncoated urine tubes, relative peak intensities of the 2556 m/z and 2645 m/z peaks were attenuated in poly(MPC-*co*-BMA)-coated tubes (the upper panel), which was confirmed by the quantitative study using the internal standard (2556 m/z; $P < 0.007$, 2653 m/z; $P < 0.014$).

urine collection tube. It is 0.85 μg/mL when converting it into the protein concentration.

3.4. MALDI-TOF MS Analysis of Urine Samples Kept in Poly(MPC-co-BMA)-Coated and Noncoated Tubes. Figure 2(b) shows representative view of the spectrum of urine samples kept in urine tubes. The relative intensities of the two peaks (2556 m/z and 2654 m/z) were greater in the samples kept in poly(MPC-*co*-BMA)-coated tubes than those kept in noncoated tubes. Similar results were obtained in 7 other experiments; the differences quantified using the SID-FIC 5.9 were statistically significant ($P < 0.007$ for 2556 m/z and $P < 0.014$ for 2653 m/z).

3.5. Routine Urinalysis. The quantitative values of urinalysis parameters in samples kept in poly(MPC-*co*-BMA)-coated and noncoated collection tubes are presented in Table 3. They were all comparable between the two groups in linear regression equation, slope of linearity, correlation coefficients ranged from 0.997 to 1.014. The results of dipstick urinalysis and urinary sediment determinations were also comparable between the two groups.

4. Discussion

The issue of preanalytical factors affecting sample integrity is often overlooked and yet is critically important. Although preanalytical factors for serum or plasma proteome analysis have been extensively studied, the impact of adsorption of proteins and peptides on urine tubes on biomarker discovery using urinary proteomics is not well investigated.

The results of this study indicate that conventionally used urine collection tubes adsorb proteins and/or peptides and

Table 3: Quantitative values for urinalysis parameters in samples kept in poly(MPC-*co*-BMA)-coated and noncoated urine collection tubes.

Parameters	Noncoated tube	Coated tube
pH	6.17 ± 0.58	6.17 ± 0.58
Protein (mg/dL)	57.8 ± 136.5	58.5 ± 138.2
Glucose (mg/dL)	182.4 ± 556.3	184.2 ± 561.0
Creatinine (mg/mL)	1.2 ± 0.7	1.2 ± 0.7
Microalbumin (mg/L)	212.0 ± 323.4	212.4 ± 324.2
Beta-2-microglobulin (μg/L)	1771.3 ± 3269.7	1772.4 ± 3275.0
Amylase (IU/L)	298.2 ± 245.2	300.3 ± 246.5
N-acetyl-β-D-glucosaminidase (U/L)	17.4 ± 19.2	17.4 ± 19.2
Urea nitrogen (mg/dL)	620.2 ± 315.9	623.3 ± 317.8
Uric acid (mg/dL)	49.8 ± 27.5	50.0 ± 27.7
Calcium (mg/dL)	10.7 ± 9.7	10.8 ± 9.7
Magnesium (mg/dL)	6.8 ± 3.8	6.8 ± 3.8
Sodium (mEq/L)	106.9 ± 48.9	107.1 ± 49.1
Potassium (mEq/L)	45.4 ± 26.3	45.5 ± 26.4
Chlorine (mEq/L)	169.5 ± 85.4	169.7 ± 85.2
Inorganic phosphorus (mg/dL)	53.1 ± 30.3	53.4 ± 30.6

that the surface coating of the tubes by poly(MPC-*co*-BMA) can minimize the adsorption without any significant effects on routine chemical determinations.

In this study, urine samples were kept in poly(MPC-*co*-BMA)-coated and noncoated tubes for 15 min. This is because it is generally known that proteins adsorb onto a surface within a few minutes when the material contacts body fluids such as blood, plasma, and tears [24–26].

Protein adsorption is one of the most important phenomena in determination of the biocompatibility of materials [16, 18]. Several methods have been proposed to reduce protein adsorption on medical devices.

Polymers composed of MPC and hydrophobic alkylmethacrylate units have been extensively used in many medical devices as coating materials to improve the blood compatibility of these devices [15–18]. However, this coating requires a long wetting pretreatment time to achieve equilibrium hydration by the reorientation of the phosphorylcholine groups [16, 27]. In this study, urinary proteins were found to be adsorbed on poly(MPC-*co*-BMA) noncoated urine collection tubes made from six different types of materials. Recently, Futamura et al. [18] succeeded in rapid development of hydrophilicity and protein adsorption resistance poly(ethylene terephthalate) (PET) surfaces bearing poly(MPC-*co*-2-vinylnaphthalene(vN)) (PMvN). It should be considered, however, that coating effects on the plastic tubes appear to be dependent on the initial properties of the plastic tubes. We took advantage of this coating method and showed that protein adsorption can be reduced in urine samples as well.

Most of the proteins listed in Table 1 are representative protein in urine and have theoretical isoelectric points between 4.7 and 7.0, suggesting that proteins with isoelectric point of this range are likely to be adsorbed on the conventional urine tubes employed in this study. Since the pH of the urine samples kept in poly(MPC-*co*-BMA)-coated and noncoated tube was similar, it is unlikely that the differences obtained in this study were due to pH difference. It has been reported that the factor responsible for protein adsorption to the plastic tube might depend on the relation of sample pH and protein pI [28].

Three proteins (alpha-1-antitrypsin, IGKV1-5 protein, prostaglandin-H2 D-isomerase) were detected in common for two different comparisons.

Ceruloplasmin, one of the proteins listed in Table 1, is a biomarker of uranium nephrotoxicity [29].

The use of the coated tubes did not have any impact on the urine analysis of routine parameters. Since there were no significant differences in the quantitative data of abundant urinary proteins including albumin and beta-2-microglobulin, the effects of adsorption on abundant proteins may be minimal. But, in searching for urinary protein biomarkers with low concentration, possible adsorption on conventional urine tubes should be considered. Since the material used in the conventional and noncoated tubes employed in the present study is widely used around the world, possible adsorption of trance amount of proteins to urine collection tubes should be considered in proteome analyses of urine samples.

5. Summary

Urine is one of the attractive biofluids in clinical proteomics. In chasing very low abundance urinary proteins and peptides, however, loss of biomarker candidates by adsorption on urine tubes should be considered. In this study, we found that protein adsorption on the conventionally used urine collection tubes is not negligible, and that the adsorption can be reduced by using a tube coated by hydrophilic polymers without any effects on routine urinalysis.

I believe that these findings should be shared by those who are interested in urinary proteomic study.

Abbreviations

MPC: 2-methacryloyloxyethyl phosphorylcholine,
poly(MPC-*co*-BMA): poly(MPC-*co*-n-butyl methacrylate(BMA)).

Acknowledgment

Iwao Kiyokawa and Kazuyuki Sogawa equally contributed to this work.

References

[1] A. R. Vaezzadeh, H. Steen, M. R. Freeman, and R. S. Lee, "Proteomics and opportunities for clinical translation in urological disease," *Journal of Urology*, vol. 182, no. 3, pp. 835–843, 2009.

[2] S. Decramer, A. G. de Peredo, B. Breuil et al., "Urine in clinical proteomics," *Molecular and Cellular Proteomics*, vol. 7, no. 10, pp. 1850–1862, 2008.

[3] S. Schaub, D. Rush, J. Wilkins et al., "Proteomic-based detection of urine proteins associated with acute renal allograft rejection," *Journal of the American Society of Nephrology*, vol. 15, no. 1, pp. 219–227, 2004.

[4] A. Vlahou, P. F. Schellhammer, S. Mendrinos et al., "Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine," *American Journal of Pathology*, vol. 158, no. 4, pp. 1491–1502, 2001.

[5] K. Rossing, H. Mischak, M. Dakna et al., "Urinary proteomics in diabetes and CKD," *Journal of the American Society of Nephrology*, vol. 19, no. 7, pp. 1283–1290, 2008.

[6] D. Theodorescu, D. Fliser, S. Wittke et al., "Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine," *Electrophoresis*, vol. 26, no. 14, pp. 2797–2808, 2005.

[7] C. Z. Von Muhlen, E. Schiffer, P. Zuerblg et al., "Evaluation of urine proteome pattern analysis for its potential to reflect coronary artery atherosclerosis in symptomatic patients," *Journal of Proteome Research*, vol. 8, no. 1, pp. 335–345, 2009.

[8] B. Ye, S. Skates, S. C. Mok et al., "Proteomic-based discovery and characterization of glycosylated eosinophil-derived neurotoxin and COOH-terminal osteopontin fragments for ovarian cancer in urine," *Clinical Cancer Research*, vol. 12, no. 2, pp. 432–441, 2006.

[9] T. Pisitkun, R. Johnstone, and M. A. Knepper, "Discovery of urinary biomarkers," *Molecular and Cellular Proteomics*, vol. 5, no. 10, pp. 1760–1771, 2006.

[10] V. Thongboonkerd, "Urinary proteomics: towards biomarker discovery, diagnostics and prognostics," *Molecular BioSystems*, vol. 4, no. 8, pp. 810–815, 2008.

[11] V. Thongboonkerd, S. Chutipongtanate, and R. Kanlaya, "Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: quantity, quality, and variability," *Journal of Proteome Research*, vol. 5, no. 1, pp. 183–191, 2006.

[12] V. Thongboonkerd, "Practical points in urinary proteomics," *Journal of Proteome Research*, vol. 6, no. 10, pp. 3881–3890, 2007.

[13] S. Schaub, J. Wilkins, T. Weiler, K. Sangster, D. Rush, and P. Nickerson, "Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry," *Kidney International*, vol. 65, no. 1, pp. 323–332, 2004.

[14] M. Afkarian, M. Bhasin, S. T. Dillon et al., "Optimizing a proteomics platform for urine biomarker discovery," *Molecular and Cellular Proteomics*, vol. 9, no. 10, pp. 2195–2204, 2010.

[15] K. Ishihara, R. Aragaki, T. Ueda, A. Watenabe, and N. Nakabayashi, "Reduced thrombogenicity of polymers having phospholipid polar groups," *Journal of Biomedical Materials Research*, vol. 24, no. 8, pp. 1069–1077, 1990.

[16] K. Ishihara, H. Nomura, T. Mihara, K. Kurita, Y. Iwasaki, and N. Nakabayashi, "Why do phospholipid polymers reduce protein adsorption?" *Journal of Biomedical Materials Research*, vol. 39, no. 2, pp. 323–330, 1998.

[17] R. R. Palmer, A. L. Lewis, L. C. Kirkwood et al., "Biological evaluation and drug delivery application of cationically modified phospholipid polymers," *Biomaterials*, vol. 25, no. 19, pp. 4785–4796, 2004.

[18] K. Futamura, R. Matsuno, T. Konno, M. Takai, and K. Ishihara, "Rapid development of hydrophilicity and protein adsorption resistance by polymer surfaces bearing phosphorylcholine and naphthalene groups," *Langmuir*, vol. 24, no. 18, pp. 10340–10344, 2008.

[19] M. Satoh, E. Haruta-Satoh, A. Omori et al., "Effect of thyroxine on abnormal pancreatic proteomes of the hypothyroid rdw rat," *Proteomics*, vol. 5, no. 4, pp. 1113–1124, 2005.

[20] M. Oh-Ishi, M. Satoh, and T. Maeda, "Preparative two-dimensional gel electrophoresis with agarose gels in the first dimension for high molecular mass proteins," *Electrophoresis*, vol. 21, no. 9, pp. 1653–1669, 2000.

[21] Y. Kawashima, T. Fukuno, M. Satoh et al., "A simple and highly reproducible method for discovering potential disease markers in low abundance serum proteins," *Journal of Electrophoresis*, vol. 53, pp. 13–18, 2009.

[22] K. Sogawa, Y. Kodera, K. Noda et al., "The measurement of a fibrinogen α C-chain 5.9 kDa fragment (FIC5.9) using MALDI-TOF MS and a stabel isotope-labeled peptide standard dilution," *Clinica Chimica Acta*, vol. 412, pp. 1094–1099, 2011.

[23] H. Umemura, A. Togawa, K. Sogawa et al., "Identification of a high molecular weight kininogen fragment as a marker for early gastric cancer by serum proteome analysis," *Journal of Gastroenterology*, vol. 46, pp. 577–585, 2011.

[24] I. Aoike, "Clinical significance of protein adsorbable membranes–long-term clinical effects and analysis using a proteomic technique," *Nephrology, Dialysis, Transplantation*, vol. 22, pp. v13–19, 2007.

[25] P. Kingshott, H. A. W. S. John, R. C. Chatelier, and H. J. Griesser, "Matrix-assisted laser desorption ionization mass spectrometry detection of proteins adsorbed in vivo onto contact lenses," *Journal of Biomedical Materials Research*, vol. 49, no. 1, pp. 36–42, 2000.

[26] L. Santos, D. Rodrigues, M. Lira et al., "The influence of surface treatment on hydrophobicity, protein adsorption and microbial colonisation of silicone hydrogel contact lenses," *Contact Lens and Anterior Eye*, vol. 30, no. 3, pp. 183–188, 2007.

[27] K. Nishizawa, T. Konno, M. Takai, and K. Ishihara, "Bioconjugated phospholipid polymer biointerface for enzyme-linked immunosorbent assay," *Biomacromolecules*, vol. 9, no. 1, pp. 403–407, 2008.

[28] A. Gotoh, K. Uchida, Y. Hamano, S. Mashiba, I. Kawabata, and Y. Itoh, "Evaluation of adsorption of urine cystatin C to the polymer materials on the microplate by an antigen capture enzyme-linked immunosorbent assay," *Clinica Chimica Acta*, vol. 397, no. 1-2, pp. 13–17, 2008.

[29] V. Malard, J. C. Gaillard, F. Berenguer, N. Sage, and E. Quemeneur, "Urine proteomic profiling of uranium nephrotoxicity," *Biochimica et Biophysica Acta*, vol. 1794, pp. 882–891, 2009.

Hindawi Publishing Corporation
International Journal of Proteomics
Volume 2011, Article ID 739615, 11 pages
doi:10.1155/2011/739615

Research Article

Combined Use of a Solid-Phase Hexapeptide Ligand Library with Liquid Chromatography and Two-Dimensional Difference Gel Electrophoresis for Intact Plasma Proteomics

Tatsuo Hagiwara,[1,2] Yumi Saito,[1] Yukiko Nakamura,[1,3] Takeshi Tomonaga,[3] Yasufumi Murakami,[2] and Tadashi Kondo[1]

[1] *Division of Pharmacoproteomics, National Cancer Center Research Institute, Chuo-ku, Tokyo 104-0045, Japan*
[2] *Laboratory of Genome Biology, Department of Biological Science and Technology, Tokyo University of Science, Tokyo 278-8510, Japan*
[3] *Laboratory of Proteome Research, National Institute of Biomedical Innovation, Osaka 567-0085, Japan*

Correspondence should be addressed to Tadashi Kondo, takondo@ncc.go.jp

Received 2 May 2011; Accepted 9 June 2011

Academic Editor: David E. Misek

Copyright © 2011 Tatsuo Hagiwara et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The intact plasma proteome is of great interest in biomarker studies because intact proteins reflect posttranslational protein processing such as phosphorylation that may correspond to disease status. We examined the utility of a solid-phase hexapeptide ligand library in combination with conventional plasma proteomics modalities for comprehensive profiling of intact plasma proteins. Plasma proteins were sequentially fractionated using depletion columns for albumin and immunoglobulin, and separated using an anion-exchange column. Proteins in each fraction were treated with a solid-phase hexapeptide ligand library and compared to those without treatment. Two-dimensional difference gel electrophoresis demonstrated an increased number of protein spots in the treated samples. Mass spectrometric studies of these protein spots with unique intensity in the treated samples resulted in the identification of high- and medium-abundance proteins. Our results demonstrated the possible utility of a solid-phase hexapeptide ligand library to reveal greater number of intact plasma proteins. The characteristics of proteins with unique affinity to the library remain to be clarified by more extensive mass spectrometric protein identification, and optimized protocols should be established for large-scale plasma biomarker studies.

1. Introduction

The plasma proteome has been extensively investigated with the aim of biomarker development [1, 2]. Plasma is the most accessible clinical material, and plasma biomarkers for early diagnosis and monitoring the response to therapy and disease recurrence would be beneficial for patients with cancer. Because proteins released by tumors, particularly early-stage tumors, are expected to exist in very low concentrations and plasma contains various proteins with considerable heterogeneity between and within patients, the identification of novel plasma biomarkers represents a substantial challenge.

Global expression studies on intact plasma proteins are of special interest in biomarker studies as the intact proteins reflect the functional features of protein structure. Those include posttranslational processing such as phosphorylation and glycosylation. Peptide subsets from complex digests have been analyzed for plasma proteomics, resulting in the identification of low-abundance proteins such as tissue leakage proteins [3] and biomarker candidates [4]. However, analysis of peptide digests may not be sensitive to posttranslational protein processing, and may therefore not reveal many relevant protein isoforms associated with disease status. To date, much effort has been devoted to detect trace intact proteins in complex plasma samples.

The utility of a combinatorial hexapeptide ligand library immobilized on a solid-phase matrix has been reported, introduced to intact plasma proteomics [5–9], and

commercialized as ProteoMiner (Bio-Rad Laboratories, Hercules, CA, USA). ProteoMiner contains millions of randomly synthesized hexapeptide ligands that are equally represented with a selected number of targets. When a complex plasma protein extract is exposed, the hexapeptide ligands for high-abundance proteins are saturated, but the majority remains unbound. In contrast, the proteins which do not saturate the corresponding hexapeptide ligands and usually not observed by the conventional methods will appear in the proteome data. The approach of using a combinatorial hexapeptide ligand library is different from that of using depletion and separation; thus, it reveals a novel aspect of the plasma proteome. A recent report demonstrated that prefractionation using a hexapeptide ligand library for shotgun mass spectrometry studies identified plasma proteins not recorded in the Human Plasma Proteome Project [10]. The combined use of a hexapeptide ligand library with depletion and separation methods has also been a challenge in deeper plasma proteomics [11], and the resulting protein contents are examined by gel electrophoresis and mass spectrometry [12, 13]. ProteoMiner has been used for disease biomarker studies in lung cancer [14] and liver cancer [15]. Considering that it will potentially visualize the unique plasma proteome aspects, the application and optimization of a solid-phase hexapeptide ligand library for disease biomarker studies should be further investigated.

In this study, we examined the utility of a solid-phase hexapeptide ligand library in combination with a depletion column, an anion-exchange column, and 2D-DIGE that allows an instant visual comparison of the protein patterns. Protein spots exhibiting prominent differences between samples treated with and without the library were subjected to mass spectrometry. Our study clearly demonstrated that the combined use of the ProteoMiner and the other proteomics modalities can visualize unique plasma proteome.

2. Materials and Methods

2.1. Sample Preparation. Frozen human plasma was purchased from Cosmobio KOJ (Tokyo, Japan). After the plasma was placed on ice, 40 mL plasma was centrifuged and 30 mL supernatant was recovered for the following experiments.

2.2. Albumin Depletion. Albumin and other proteins were separated using a HiTrap Blue HP column (5 mL resin, GE, Uppsala, Sweden) with the AKTA Explorer system (GE) at a flow rate of 1.0 mL/min. The separation was initiated by washing the column with rinse buffer (50 mM KH_2PO_4/Na_2HPO_4, pH 7.0) for 5 min. Plasma (30 mL) was diluted with 60 mL 50 mM KH_2PO_4/Na_2HPO_4 (pH 7.0), and 9 mL of the diluted plasma was injected. The column was then washed with binding buffer (50 mM KH_2PO_4/Na_2HPO_4, pH 7.0) for 35 min, and the flow-through fraction was collected. Bound proteins were eluted from the column with elution buffer (50 mM $KHPO_4$, 1.5 M KCl, pH 7.0) for 45 min, and the bound fraction was collected. The column was neutralized with rinse buffer for 20 min. This process was repeated 10 times for a total of 90 mL of diluted plasma.

One-third of the flow-through and bound fractions, approximately 150 mL of each, was concentrated to 1.2 mL using a VIVA Spin 20 column (10 K MWCO, 20 mL capacity, Sartorius, Gotteingen, Germany). Then, 1.0 mL and 0.20 mL of the concentrated samples were subjected to treatment with the solid-phase hexapeptide ligand library and 2D-DIGE, respectively. Two-thirds of the flow-through fraction, approximately 300 mL, was subjected to an immunodepletion column.

2.3. Immunoglobulin Depletion. Immunoglobulin was depleted using the HiTrap Protein G HP column (1 mL resin, GE) with the AKTA Explorer system (GE) at a flow rate of 1.0 mL/min. The depletion was initiated by washing the column with rinse buffer (50 mM KH_2PO_4/Na_2HPO_4, pH 7.0) for 4 min. After 15 mL of the flow-through fraction from the HiTrap Blue HP column was injected, the column was washed with binding buffer (50 mM KH_2PO_4/Na_2HPO_4, pH 7.0) for 5 min, and the flow-through fraction was collected. Bound proteins were eluted from the column with elution buffer (0.1 M glycine-HCl, pH 2.2) for 8 min and collected as the bound fraction. The collected bound fraction was immediately neutralized with neutralizing buffer (1.0 M Tris-HCl, pH 9.0). The column was equilibrated with rinse buffer for 5 min for reuse. This process was repeated 20 times for a total of two-thirds of the flow-through fraction from the HiTrap Blue HP column (approximately 300 mL).

Half of the flow-through and bound fractions (approximately 200 mL and 80 mL, resp.) were concentrated to 1.2 mL and 0.25 mL, respectively, using VIVA Spin 20 columns (Sartorius). Then, 1.0 mL of the concentrated flow-through fraction and 0.20 mL of the concentrated bound fraction were subjected to treatment with the solid-phase ligand library, and the remaining samples were subjected to 2D-DIGE. Another half of the flow-through fraction (approximately 200 mL) was concentrated to 2.0 mL using the VIVA Spin 20 column (Sartorius). After diluting with 38 mL of 25 mM Tris-HCl (pH 9.0), the sample was subjected to separation on an anion-exchange column.

2.4. Anion Exchange. The flow-through fraction from the HiTrap Protein G HP column was separated using the Resource Q column (1 mL resin, 6.4 mm id $\times$ 30 mm, GE) with the AKTA Explorer system (GE) at a flow rate of 3.0 mL/min. The separation was initiated by washing the column with rinse buffer (25 mM Tris-HCl, pH 9.0) for 4 min, and 5 mL of the flow-through fraction from the HiTrap Protein G HP column was injected. The separations were performed using a stepwise NaCl gradient as follows: 0, 100, 150, 200, 250, and 1000 mM for 5 min each. All samples contained 25 mM Tris-HCl, pH 9.0. The column was washed with rinse buffer (25 mM Tris-HCl, pH 9.0) for 5 min. This process was repeated 8 times for a total of 40 mL of the diluted flow-through fraction from the HiTrap Protein G HP column.

The collected samples were concentrated to 0.25 mL, and the buffer was exchanged gradually with 25 mM Tris-HCl (pH 9.0) using the VIVA Spin 20 column (Sartorius). Then,

0.2 mL and 0.05 mL were subjected to treatment with the solid-phase ligand library and 2D-DIGE, respectively.

2.5. Treatment with the Solid-Phase Ligand Library. A solid-phase combinatorial library of hexapeptides was purchased from Bio-Rad Laboratories (ProteoMiner kit). Unprocessed plasma (1 mL) and the flow-through fractions from the HiTrap Blue HP and HiTrap Protein G HP columns were treated using the ProteoMiner large-capacity kit, and 0.2 mL of the bound fraction from the HiTrap Protein G HP column, and all fractions from the Resource Q column were treated using the ProteoMiner small-capacity kit. After 2 h of incubation at room temperature, the unbound fraction was washed out by centrifugation. After rinsing, the bound sample was eluted with an elution reagent containing 8 M urea, 2% CHAPS, and 5% acetic acid, according to the manufacturer's instructions.

2.6. Measurement of Protein Concentration. Protein concentration was measured using a protein assay kit (Bio-Rad), according to the manufacturer's instructions (Table 1).

2.7. SDS-PAGE. Protein samples (1 μg) were examined by electrophoresis using 18-well precast 12.5% polyacrylamide gel plates (e-PAGEL, ATTO, Tokyo, Japan). Electrophoresis was performed at a constant current of 40 mA for 80 min and using the page Run AE6531 system [16]. Silver staining was performed using the Silver Stain KANTO III kit (Kanto Chemical, Tokyo, Japan), according to the manufacturer's instructions.

2.8. 2D-DIGE. 2D-DIGE was performed as described previously [17]. Briefly, protein samples (20 μg) were labeled with the Cy3 or Cy5 fluorescent dye (CyDye DIGE Fluor saturation dye, GE), and differentially labeled protein samples were mixed. After dividing into 3, the labeled protein samples were separated by 2D-PAGE. The first-dimension separation was performed using a 24 cm length immobiline gel (IPG, pI 4–7, GE) and Multiphor II (GE) whereas the second-dimension separation was performed using gradient gels prepared in house and EttanDalttwelve (GE). The gels were scanned using a laser scanner (Typhoon Trio, GE) at an appropriate wavelength for Cy3 or Cy5. The Cy3 and Cy5 intensities were compared in the same gel using the Progenesis SameSpots software (version 4.0; Nonlinear Dynamics, Newcastle, UK). ProteoMiner-treated and untreated samples were labeled with Cy3 and Cy5, respectively, or with Cy5 and Cy3, respectively. Six gels were run for each sample. The average value of the intensity ratio was calculated among the triplicate gels for all protein spots and then averaged between the 2 samples for further study. Spot intensity data were exported from the Progenesis SameSpots software as Excel files amenable to numerical data analysis.

2.9. Mass Spectrometric Protein Identification. Proteins were extracted from the protein spots by in-gel digestion, as reported previously [17]. Briefly, protein samples (100 μg) were labeled with Cy3 and separated by 2D-PAGE. The protein spots were then recovered from the gel pieces using an automated spot recovery machine. The recovered protein spots were extensively washed with a solution containing acetonitrile and ammonium bicarbonate minimum and treated with trypsin (Promega, Madison, WI, USA) at 37°C overnight. The tryptic digests were recovered from the gel pieces, concentrated by vacuum, and resolubilized with 0.1% trifluoroacetic acid. The final tryptic digests were subjected to mass spectrometry, which was performed using the LXQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA, USA). The Mascot software (version 2.3.0; Matrix Science, London, UK) was used to search for the mass of the peptide ion peaks against the SWISS-PROT database (Homo sapiens, 471472 sequences in Sprot_57.5 fasta file). The search parameters were as follows: trypsin digestion allowing up to 3 missed tryptic cleavages, fixed modifications of

Table 1: List of the identified proteins and their reported concentration.

Protein name	Normal concentration μg/mL
Adiponectin	2–17
Albumin	35000–52000
Alpha-1-antitrypsin	900–2000
Alpha-1B-glycoprotein	150–300
Alpha-2-macroglobulin	1300–3000
Apolipoprotein A-I	1000–2000
Apolipoprotein A-II	190–300
Apolipoprotein A-IV	110–220
Apolipoprotein D	60–90
Apolipoprotein E	30–60
Carboxypeptidase N	30
Ceruloplasmin	190–370
Clusterin	250–420
Coagulation factor X	10
Complement C3	900–1800
Complement C4-A	25–90
Fibrinogen beta chain	520–1420
Fibrinogen gamma chain	490–1340
Ficolin-2	1–12
Ficolin-3	3–54
Haptoglobin	200–2000
Haptoglobin-related protein	32–41
Inter-alpha-trypsin inhibitor (heavy chain H3)	100–200
Paraoxonase/arylesterase 1	58–61
Prothrombin	100
Serotransferrin	2000–3600
Transthyretin	200–400
Vitronectin	240–530
Zinc-alpha-2-glycoprotein	60–80

The table with the references for the protein concentration is shown in Supplementary Table 7 in Supplemrntary material available online at doi: 10.1155/2011/39615.

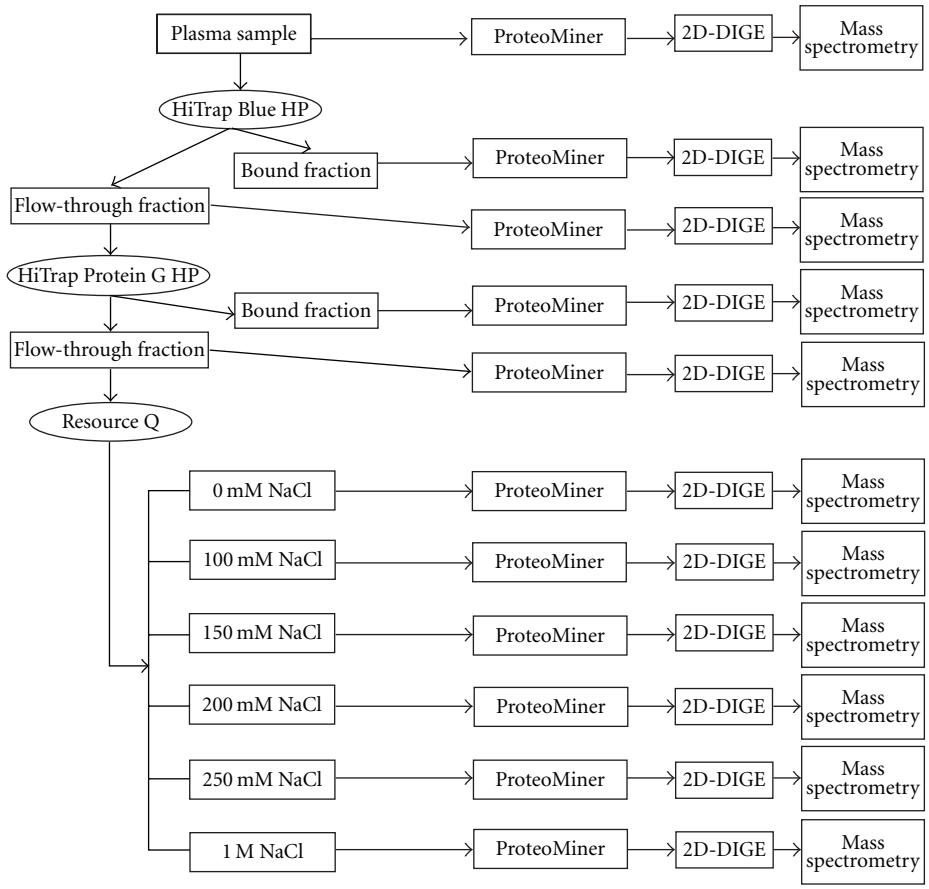


Figure 1: Overview of protein fractionation by sequential use of 3 different columns to separate plasma proteins. All fractions were subjected to ProteoMiner and 2D-DIGE.

carbamidomethyl, variable modifications of oxidation, 1^+, 2^+, and 3^+ peptide charge, peptide mass tolerance of 2.0 Da, and use of MS/MS tolerance of 1.0 Da for all tryptic-mass searches.

3. Results and Discussion

We previously reported the utility of combining multidimensional chromatography and 2D-DIGE for intact plasma proteomics. Extensive fractionation by the different separation modes increased the number of protein spots on 2D-DIGE and allowed a quantitative comparison between the plasma samples from healthy donors and those from patients with lung adenocarcinoma [18] and pancreatic cancer [19]. However, mass spectrometric protein identification revealed that protein spots with a significant difference between the sample groups corresponded to high- and medium-abundance proteins such as acute-phase proteins, but no known plasma tumor markers were detected. Thus, we concluded that further investigations are needed to reveal low-abundance proteins for biomarker studies. In this study, we examined whether a novel technology, a solid-phase hexapeptide ligand library could improve the linkage of multidimensional chromatography and 2D-DIGE.

3.1. Overall View of Protein Fractionation: Comparison and Detection. The overall view of sequential protein separation is shown in Figure 1. A sample equivalent to 10 mL plasma was separated using 3 different columns and then treated with the solid-phase hexapeptide ligand library ProteoMiner. The ProteoMiner-treated and untreated samples were compared using 2D-DIGE by labeling them with different fluorescent dyes and separating the labeled proteins on an identical gel. The protein spots with significantly different intensities between the ProteoMiner-treated and untreated samples were subjected to mass spectrometry to identify the proteins.

The number of observable low-abundance proteins was affected by the initial amount of plasma sample and the sensitivity of the final quantification method. We used a relatively large volume of plasma sample (10 mL) as the initial material. The immunodepletion columns allow the use of only a small volume of plasma sample for separation. Furthermore, a significantly larger number of plasma samples should be examined to obtain conclusive results for biomarker development. A larger volume of samples can be manipulated by repeatedly using the same immunodepletion column. Although it is quite feasible, special attention may be required to maintain reproducibility during a long period of use. In this study, we used Blue Sepharose and Protein

G-Sepharose columns in a sequential manner to deplete albumin and subsequently immunoglobulin and to minimize repeated use of the same column. Although these columns may have less sensitivity than an immunodepletion column and deplete nontargeted proteins that may bind to albumin and immunoglobulin, a larger volume of plasma sample can be treated in individual procedures. A previous study indicated that Cibacron Blue beads remove a major portion of the albumin but with concomitant loss of potentially important peptides and proteins [20]. Thus, we examined both the column-bound and flow-through fractions (Figure 1). Although the specificity of Cibacron Blue beads was not validated in this study, as the purpose of Cibacron Blue was to reduce the complexity of plasma sample, it should not be problem.

To avoid possible redundant proteins in the neighboring fractions as much as possible when utilizing the anion-exchange column, we used stepwise elution and fractionation; once all proteins were eluted, the next elution buffer was applied to the column (Figure 1). Considering the complexity of the samples and resolution of an anion-exchange column, extensive fractionation with a gradient buffer system may result in redundant contents among the fractions. We employed 6 stepwise fractionations by monitoring the fraction contents using SDS-PAGE (data not shown).

3.2. High Reproducibility of Protein Fractionation by Chromatography. The ultraviolet detection (280 nm) trace for each run demonstrated consistent separation of albumin and immunoglobulin from the depletion and anion-exchange columns. This high reproducibility may suggest the possible utilities of this approach for biomarker studies (Supplementary Figure 1). High quantitative and qualitative reproducibility of the solid-phase hexapeptide ligand library ProteoMiner has been confirmed in previous reports [21, 22].

3.3. Demonstration of the Effects of Fractionation and Dynamic Range Reduction. We examined the effects of sequential plasma protein fractionation using 3 columns and the reduction of dynamic range by ProteoMiner (Figure 2). The contents of the fractionated samples were apparently different from each other. Notably, the protein sample bound to the Blue Sepharose and Protein G-Sepharose columns included many proteins that should be different from the targeted proteins, according to their molecular weights. Treatment of the fractionated samples with ProteoMiner enhanced the proteins that were not observed, except for those bound to the Protein G-Sepharose column. Treatment of the bound fraction from the Protein G-Sepharose column with ProteoMiner did not result in a greater number of observable proteins. This may have been due to the low complexity and narrow dynamic range of proteins in the bound fraction from the Protein G-Sepharose column. There was one order of magnitude in concentration difference for the observed protein bands in the ProteoMiner-treated sample. These observations may reflect that the affinity of proteins for the peptide may not be equal and even the number of peptides bound on the beads is equal, the amount of proteins bound to the ProteoMiner may be different

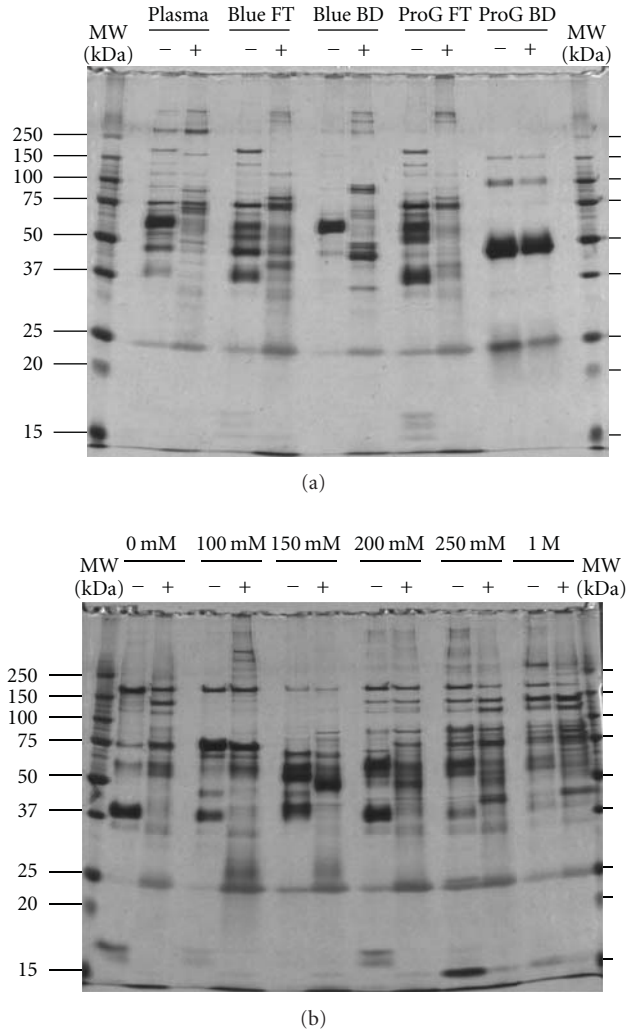


Figure 2: Overview of the protein contents fractionated by liquid chromatography and ProteoMiner. The fractionated protein samples were loaded onto SDS-PAGE, and the protein contents were visualized by silver staining.

depending on their affinity. This fraction contained similar amounts of only 4 major proteins, as revealed by SDS-PAGE (Figure 2), and they may have been absorbed to ProteoMiner in proportion to their original amount.

The concentration and amount of protein samples before and after ProteoMiner treatment are summarized in Supplementary Table 1. The recovery rate from ProteoMiner was between 0.54 and 6.33%, suggesting that a unique population of protein species selectively bound to ProteoMiner. This assumption was supported by the SDS-PAGE data, except the bound fraction from the Protein G-Sepharose column which included only 4 major proteins that were bound to ProteoMiner (Figure 2).

3.4. Higher Separation of Fractionated Protein Samples and an Evaluation of the Effects of ProteoMiner Treatment. Although SDS-PAGE separated individual proteins with higher

resolution than chromatography in this study, using it for a quantitative comparison in a biomarker study may be troublesome because many protein bands obviously overlapped (Figure 2). Thus, we subjected the fractionated samples to 2D-DIGE in order to separate the proteins with higher resolution. Bandow compared ProteoMiner-treated and untreated plasma samples using conventional 2D-PAGE and demonstrated substantial differences between unprocessed and immunodepleted plasma samples [11]. In 2D-DIGE, 2 protein samples were labeled with different fluorescent dyes, mixed, and separated by 2D-PAGE. Because the 2 samples were separated on an identical 2D-PAGE, gel-to-gel variation was compensated. In addition, the wide dynamic range of the fluorescent dyes enabled a quantitative comparison. 2D-DIGE has been applied to compare the performance of ProteoMiner with an immunodepletion column [12]. We further extended the evaluation of the utility of ProteoMiner by loading a high amount of protein and examining the proteins separated by an anion-exchange column.

The fluorescent 2D-PAGE images of the ProteoMiner-treated and untreated samples were overlaid with different colors, so that the unique protein contents were visualized (Figures 3 and 4). The results of experiments in which the fluorescent dyes were swapped are shown in Supplementary Figure 2. Consistent with the SDS-PAGE results (Figure 2), Figure 3 demonstrates that the approach involving depletion of high-abundance proteins and multidimensional separation was an effective prefractionation method to increase the number of protein spots, and the use of ProteoMiner treatment also contributed to reveal more plasma proteins. Because these fractionation methods are based on different binding properties of proteins, their combined use revealed additional plasma proteins.

The number of observed protein spots on 2D gel electrophoresis is summarized in Supplementary Table 2. Overall, the total number of protein spots increased by treating the samples with ProteoMiner, except for the bound fraction from the Protein G-Sepharose column. This observation suggests that ProteoMiner may be a useful tool to observe a greater number of protein spots in prefractionated samples.

We compared the protein spots of the samples with and without ProteoMiner treatment (Supplementary Table 3). Depending on the criteria, different numbers of protein spots showed significantly different intensities. Although the total number of protein spots increased by treating the samples with ProteoMiner (Supplementary Table 2), many protein spots revealed decreased intensity with treatment, suggesting the selective enrichment by ProteoMiner.

3.5. Mass Spectrometric Identification of Proteins with Different Affinities to ProteoMiner. To reveal the characteristics of proteins with a particularly high or low affinity to ProteoMiner, among the protein spots with greater than 5-fold differences (Supplementary Table 3), we selected those with the top 10% different intensities between the ProteoMiner-treated and untreated samples in each fraction and subjected them to mass spectrometric identification. A total of 200 protein spots were subjected to mass spectrometry, and a positive identification was obtained for 128 (Supplementary Table 4). A list of the identified proteins is provided in Supplementary Table 5, and data supporting protein identification are shown in Supplementary Table 6. These 128 protein spots corresponded to 29 unique proteins. Because the fold difference of the protein spots in the bound fraction from the Protein G-Sepharose column was less than 4, we did not examine them. Of the original plasma samples, vitronectin and albumin were most affected by ProteoMiner treatment and disappeared after depletion and fractionation using the anion-exchange column. The other proteins were identified as enriched (or nonenriched) by ProteoMiner treatment. Proteins bound to ProteoMiner have been reported in previous studies in which the proteins were globally identified by mass spectrometry. Dwivedi et al. demonstrated that albumin, alpha 1-antitrypsin, alpha 2-macroglobulin, apolipoprotein A-I, apolipoprotein A-II, haptoglobin-related protein, and serotransferrin have high affinity to ProteoMiner [21]. In addition, Beseme et al. identified apolipoprotein A-IV, apolipoprotein D, apolipoprotein E, ceruloplasmin, complement C3, fibrinogen beta, fibrinogen gamma, ficolin-2, ficolin-3, paroxonase I, prothrombin, transthyretin, and vitronectin [23]. The protein concentrations identified in this study are summarized in Supplementary Table 6. According to the literatures, the identified proteins were classified as high- and medium-abundance proteins. Adiponectin and the carboxypeptidase N catalytic chain are not reported in previous studies, in which ProteoMiner-treated samples were examined by 2D-PAGE and mass spectrometry.

Adiponectin is an adipocytokine [24–27] and plays a protective role against obesity-related disorders such as metabolic syndrome [28], type 2 diabetes [29], and cardiovascular disease [30]. Low levels of plasma adiponectin are associated with obesity [31] and many types of malignancies such as liver cancer [32], breast cancer [33], pancreatic cancer [34], and endometrial cancer [35]. An epidemiological study suggested that adiponectin is involved in early colorectal carcinogenesis [36], and that a low circulating adiponectin level is correlated with a poor prognosis in patients with colorectal cancer [37]. The molecular backgrounds of these observations may be attributable to the antiproliferative effects of adiponectin on cancer cells [38].

Carboxypeptidase N (CPN), which is also known as kininase I, arginine carboxypeptidase, and anaphylatoxin inactivator, is a zinc finger metalloprotease. It cleaves basic lysine and arginine residues from the carboxy terminal of proteins [39]. CPN is produced in the liver and secreted into the plasma. It modulates the activity of cytokines such as stromal cell-derived factor-1 alpha [40]. The association of CPN1 with malignancy and other diseases has not been reported, and the clinical utility of CPN1 has not been suggested.

The working hypothesis of this study was that the combined use of different separation methods, including a solid-phase hexapeptide ligand library, would increase the number of observable proteins, and finally visualize the proteome that may not be observed otherwise. By loading a high amount of protein and using extensive prefractionation techniques prior to using ProteoMiner, trace proteins became visible in SDS-PAGE, and the number of protein spots on

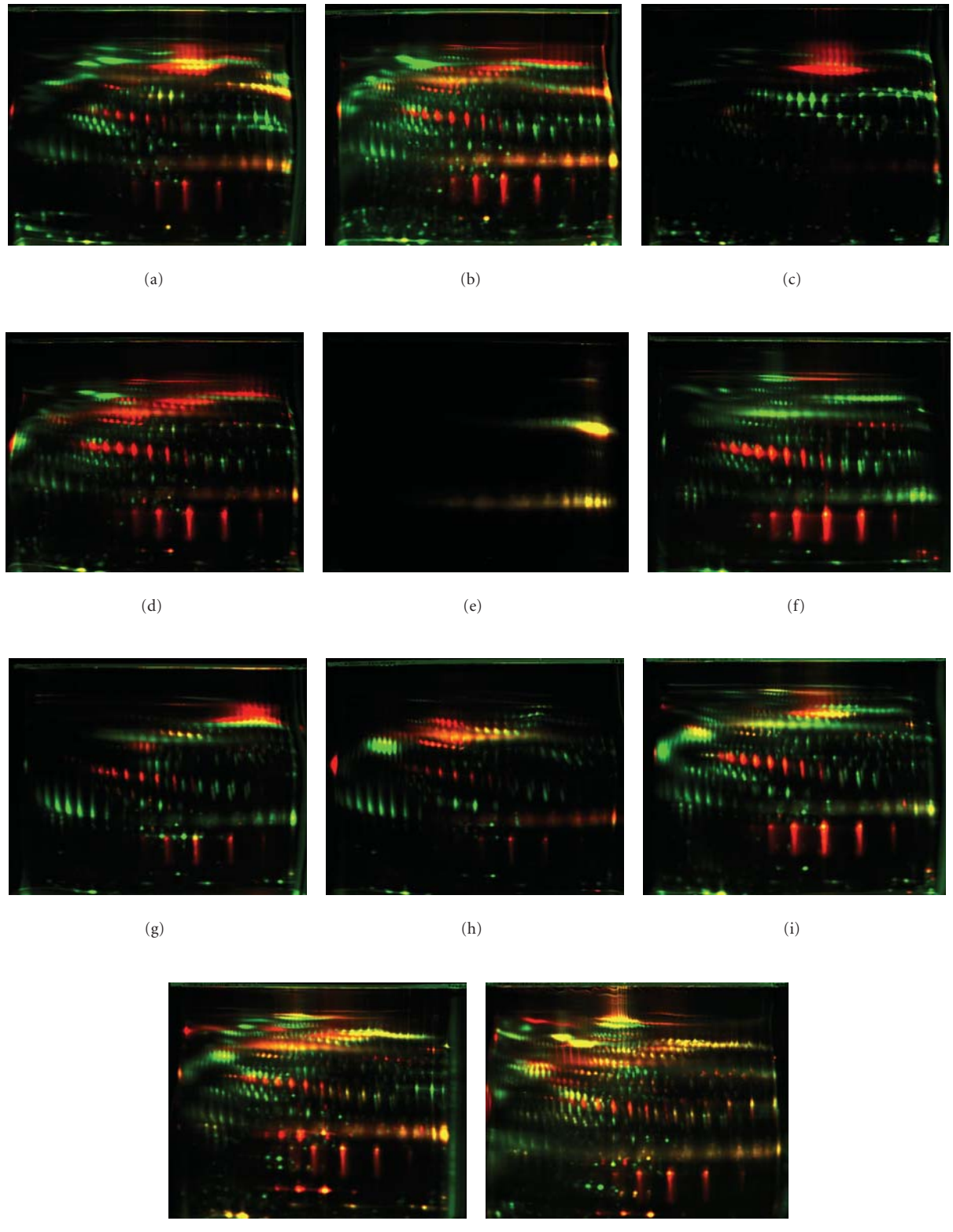


Figure 3: Effects of ProteoMiner treatment were examined by 2D-DIGE. The ProteoMiner-treated and untreated samples were labeled with Cy3 and Cy5, respectively, mixed, and separated by 2D gel electrophoresis. Note that a significant number of protein spots showed different intensities between the 2 samples. The dye-swapped images are shown in Supplementary Figure 2. (a) Original plasma; (b) flow-through fraction of HiTrap Blue HP column; (c) binding fraction of HiTrap Blue HP column. (d) Flow-through fraction of HiTrap Protein G HP column. (e) Binding fraction of HiTrap Protein G HP column; 0 mM fraction. (f) 100 mM fraction. (g) 150 mM. (h) 200 mM. (i) 250 mM. (j) 1 M fraction. (k) Resource Q column.

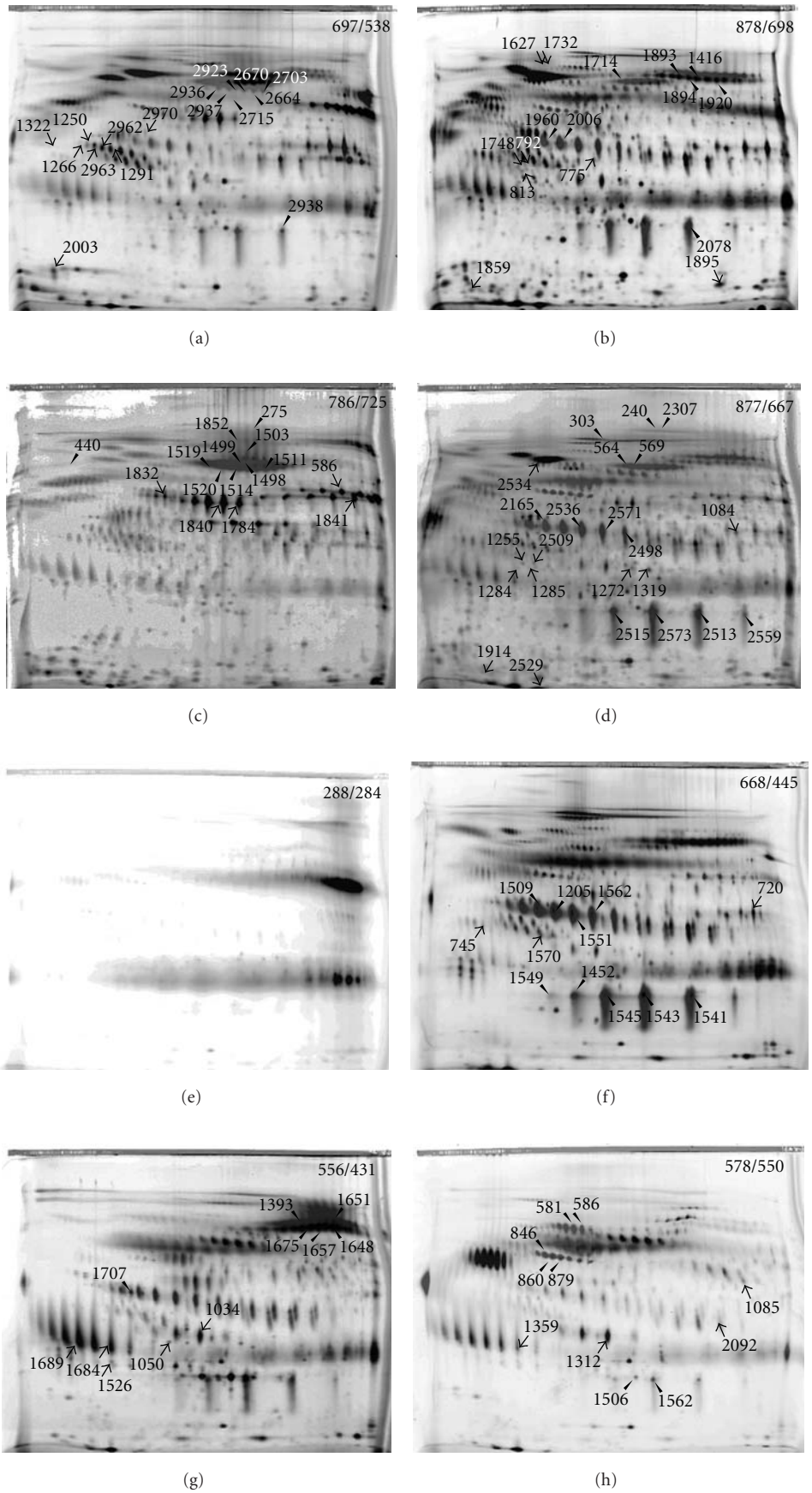


Figure 4: Continued.

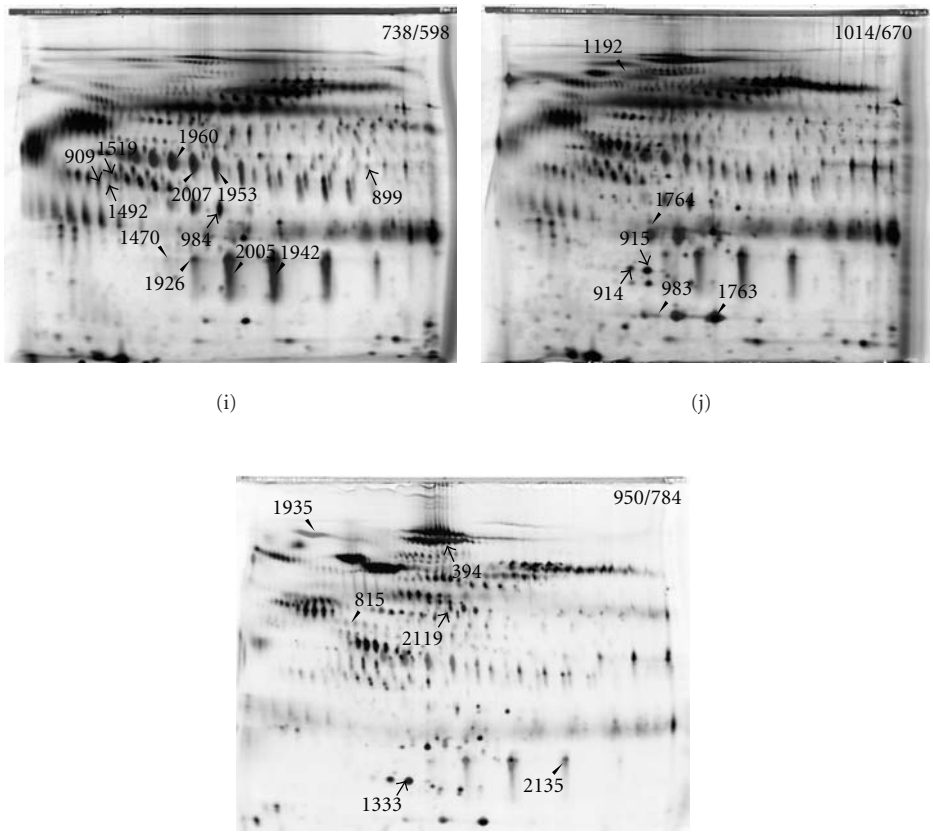


Figure 4: Localization of protein spots showing different intensities between the ProteoMinor-treated and untreated samples. Panels (a–k) correspond to those in Figure 3. The protein spot numbers corresponds to those in Supplementary Tables 5 and 6. −/−: number of protein spots without ProteoMiner treatment/those with ProteoMiner treatment.

2D-DIGE increased significantly. This approach may pave a way to a novel strategy for intact plasma proteomics. In contrast, the present results of mass spectrometric protein identification did not support the use of a solid-phase hexapeptide ligand library to enrich low-abundance proteins. It may be because our present approach had 3 limitations. First, mass spectrometric identification was performed for proteins with a greater prominent difference between the samples with or without ProteoMiner treatment, and only 128 of 200 proteins were successfully identified (Supplementary Table 4), probably because of the low protein amount. Proteins with a smaller difference or amount may include trace proteins. Although we optimized the protocols for mass spectrometric protein identification because the sensitivity of the fluorescent dye in the 2D-DIGE was very high, not all protein spots on 2D-DIGE could be identified by mass spectrometry. To evaluate enriched proteins, the complementary use of an LC-MS/MS shotgun approach may be worth considering. Second, proteins from ProteoMiner were recovered by a single-step elution with 8 M urea, 2% CHAPS, and 5% acetic acid, according to the manufacturer's instructions (Bio-Rad). However, because proteins may have interacted with hexapeptide ligand libraries in all possible modes, the absolute elution process may require sequential steps or more stringent buffer conditions such as boiling 10% SDS with 3% DTE [41]. Furthermore, various binding conditions may also be worth considering to capture whole binding proteins [9]. Third, considering the practical use of trace proteins in a biomarker study, we used as much sample as possible for identifying them and examined 10 mL plasma samples as an initial source. However, a larger volume of plasma sample, such as 100 mL, might be needed to collect rare proteins. In practice, such a high volume of plasma is rarely obtained for many cases in biomarker studies, and we may need to optimize the protocols for use of 10 mL plasma. For instance, we identify the biomarker candidates using 100 mL plasma, and using specific antibody against the identified candidate, we will be able to screen a relatively large number of samples with 10 mL volume or less.

The combined use of the ProteoMiner and the proteomic modalities in this study may enable the quantitative comparison for biomarker studies. We demonstrated that the liquid chromatography was quantitatively reproducible (Supplementary Figure 1), and the quantitative reproducibility of the ProteoMiner and 2D-DIGE was previously reported [10, 17]. We may further need to examine how the combined

use of such reproducible methods generate the results in a reproducible way, considering the degree of differences that we expect between the samples to be compared.

4. Conclusions

The use of ProteoMiner in combination with conventional proteomic modalities such as depletion and anion-exchange columns significantly enhanced trace proteins on SDS-PAGE and increased the number of protein spots on 2D-DIGE, suggesting that the use of a solid-phase hexapeptide ligand library has great potential for intact plasma proteomics. Mass spectrometric protein identification revealed that high- and middle-abundance proteins were enriched by ProteoMiner, and the characteristics of proteins with unique affinity to a solid-phase hexapeptide ligand library remain to be clarified by more extensive mass spectrometric protein identification. Although use of ProteoMiner for biomarker studies is quite feasible and attractive, more extensive characterization of binding proteins and optimized protocols are required for large-scale biomarker studies.

Acknowledgments

This paper was supported by the Ministry of Health, Labor, and Welfare and by the Program for Promotion of Fundamental Studies in Health Sciences of the Organization for Pharmaceutical Safety and Research of Japan, T. Kondo was a principal funding recipient. The authors report no conflicts of interests.

References

[1] S. M. Hanash, S. J. Pitteri, and V. M. Faca, "Mining the plasma proteome for cancer biomarkers," *Nature*, vol. 452, no. 7187, pp. 571–579, 2008.

[2] Z. Zhang and D. W. Chan, "The road from discovery to clinical diagnostics: lessons learned from the first FDA-cleared in vitro diagnostic multivariate index assay of proteomic biomarkers," *Cancer Epidemiology Biomarkers and Prevention*, vol. 19, no. 12, pp. 2995–2999, 2010.

[3] H. Zhang, A. Y. Liu, P. Loriaux et al., "Mass spectrometric detection of tissue proteins in plasma," *Molecular and Cellular Proteomics*, vol. 6, no. 1, pp. 64–71, 2007.

[4] Q. Zhang, V. Faca, and S. Hanash, "Mining the plasma proteome for disease applications across seven logs of protein abundance," *Journal of Proteome Research*, vol. 10, no. 1, pp. 46–50, 2011.

[5] P. G. Righetti, E. Boschetti, L. Lomas, and A. Citterio, "Protein equalizer technology : the quest for a "democratic proteome"," *Proteomics*, vol. 6, no. 14, pp. 3980–3992, 2006.

[6] V. Thulasiraman, S. Lin, L. Gheorghiu et al., "Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands," *Electrophoresis*, vol. 26, no. 18, pp. 3561–3571, 2005.

[7] P. G. Righetti, A. Castagna, P. Antonioli, and E. Boschetti, "Prefractionation techniques in proteome analysis: the mining tools of the third millennium," *Electrophoresis*, vol. 26, no. 2, pp. 297–319, 2005.

[8] P. G. Righetti, A. Castagna, F. Antonucci et al., "Proteome analysis in the clinical chemistry laboratory: myth or reality?" *Clinica Chimica Acta*, vol. 357, no. 2, pp. 123–139, 2005.

[9] P. G. Righetti, E. Boschetti, A. Zanella, E. Fasoli, and A. Citterio, "Plucking, pillaging and plundering proteomes with combinatorial peptide ligand libraries," *Journal of Chromatography A*, vol. 1217, no. 6, pp. 893–900, 2010.

[10] L. Sennels, M. Salek, L. Lomas, E. Boschetti, P. G. Righetti, and J. Rappsilber, "Proteomic analysis of human blood serum using peptide library beads," *Journal of Proteome Research*, vol. 6, no. 10, pp. 4055–4062, 2007.

[11] J. E. Bandow, "Comparison of protein enrichment strategies for proteome analysis of plasma," *Proteomics*, vol. 10, no. 7, pp. 1416–1425, 2010.

[12] C. Sihlbom, I. Kanmert, H. Von Bahr, and P. Davidsson, "Evaluation of the combination of bead technology with SELDI-TOF-MS and 2-D DIGE for detection of plasma proteins," *Journal of Proteome Research*, vol. 7, no. 9, pp. 4191–4198, 2008.

[13] E. Ernoult, A. Bourreau, E. Gamelin, and C. Guette, "A proteomic approach for plasma biomarker discovery with iTRAQ labelling and OFFGEL fractionation," *Journal of Biomedicine & Biotechnology*, vol. 2010, Article ID 927917, 8 pages, 2010.

[14] J. S. K. Au, W. C. S. Cho, T. T. Yip et al., "Deep proteome profiling of sera from never-smoked lung cancer patients," *Biomedicine and Pharmacotherapy*, vol. 61, no. 9, pp. 570–577, 2007.

[15] C. Marrocco, S. Rinalducci, A. Mohamadkhani, G. M. D'Amici, and L. Zolla, "Plasma gelsolin protein: a candidate biomarker for hepatitis B-associated liver cirrhosis identified by proteomic approach," *Blood Transfusion*, vol. 8, supplement 3, pp. s105–s112, 2010.

[16] M. Gatto, M. C. Bragazzi, R. Semeraro et al., "Cholangiocarcinoma: update and future perspectives," *Digestive and Liver Disease*, vol. 42, no. 4, pp. 253–260, 2010.

[17] T. Kondo and S. Hirohashi, "Application of highly sensitive fluorescent dyes (CyDye DIGE Fluor saturation dyes) to laser microdissection and two-dimensional difference gel electrophoresis (2D-DIGE) for cancer proteomics," *Nature Protocols*, vol. 1, no. 6, pp. 2940–2956, 2007.

[18] T. Okano, T. Kondo, T. Kakisaka et al., "Plasma proteomics of lung cancer by a linkage of multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis," *Proteomics*, vol. 6, no. 13, pp. 3938–3948, 2006.

[19] T. Kakisaka, T. Kondo, T. Okano et al., "Plasma proteomics of pancreatic cancer patients by multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis (2D-DIGE): Up-regulation of leucine-rich alpha-2-glycoprotein in pancreatic cancer," *Journal of Chromatography B*, vol. 852, no. 1-2, pp. 257–267, 2007.

[20] N. Zolotarjova, J. Martosella, G. Nicol, J. Bailey, B. E. Boyes, and W. C. Barrett, "Differences among techniques for high-abundant protein depletion," *Proteomics*, vol. 5, no. 13, pp. 3304–3313, 2005.

[21] R. C. Dwivedi, O. V. Krokhin, J. P. Cortens, and J. A. Wilkins, "Assessment of the reproducibility of random hexapeptide peptide library-based protein normalization," *Journal of Proteome Research*, vol. 9, no. 2, pp. 1144–1149, 2010.

[22] E. Mouton-Barbosa, F. Roux-Dalvai, D. Bouyssié et al., "In-depth exploration of cerebrospinal fluid by combining peptide ligand library treatment and label-free protein quantification," *Molecular and Cellular Proteomics*, vol. 9, no. 5, pp. 1006–1021, 2010.

[23] O. Beseme, M. Fertin, H. Drobecq, P. Amouyel, and F. Pinet, "Combinatorial peptide ligand library plasma treatment: advantages for accessing low-abundance proteins," *Electrophoresis*, vol. 31, no. 16, pp. 2697–2704, 2010.

[24] P. E. Scherer, S. Williams, M. Fogliano, G. Baldini, and H. F. Lodish, "A novel serum protein similar to C1q, produced exclusively in adipocytes," *Journal of Biological Chemistry*, vol. 270, no. 45, pp. 26746–26749, 1995.

[25] E. Hu, P. Liang, and B. M. Spiegelman, "AdipoQ is a novel adipose-specific gene dysregulated in obesity," *Journal of Biological Chemistry*, vol. 271, no. 18, pp. 10697–10703, 1996.

[26] K. Maeda, K. Okubo, I. Shimomura, T. Funahashi, Y. Matsuzawa, and K. Matsubara, "cDNA cloning and expression of a novel adipose specific collagen-like factor, apM1 (adipose most abundant gene transcript 1)," *Biochemical and Biophysical Research Communications*, vol. 221, no. 2, pp. 286–289, 1996.

[27] Y. Nakano, T. Tobe, N. H. Choi-Miura, T. Mazda, and M. Tomita, "Isolation and characterization of GBP28, a novel gelatin-binding protein purified from human plasma," *Journal of Biochemistry*, vol. 120, no. 4, pp. 803–812, 1996.

[28] Y. Matsuzawa, T. Funahashi, S. Kihara, and I. Shimomura, "Adiponectin and metabolic syndrome," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 24, no. 1, pp. 29–33, 2004.

[29] J. Spranger, A. Kroke, M. Möhlig et al., "Adiponectin and protection against type 2 diabetes mellitus," *The Lancet*, vol. 361, no. 9353, pp. 226–228, 2003.

[30] D. M. Maahs, L. G. Ogden, G. L. Kinney et al., "Low plasma adiponectin levels predict progression of coronary artery calcification," *Circulation*, vol. 111, no. 6, pp. 747–753, 2005.

[31] E. E. Calle, C. Rodriguez, K. Walker-Thurmond, and M. J. Thun, "Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. Adults," *New England Journal of Medicine*, vol. 348, no. 17, pp. 1625–1638, 2003.

[32] T. Arano, H. Nakagawa, R. Tateishi et al., "Serum level of adiponectin and the risk of liver cancer development in chronic hepatitis C patients," *International Journal of Cancer*, vol. 129, no. 9, pp. 2226–2235, 2011.

[33] C. Duggan, M. L. Irwin, L. Xiao et al., "Associations of insulin resistance and adiponectin with mortality in women with breast cancer," *Journal of Clinical Oncology*, vol. 29, no. 1, pp. 32–39, 2011.

[34] T. Krechler, M. Zeman, M. Vecka et al., "Leptin and adiponectin in pancreatic cancer: connection with diabetes mellitus," *Neoplasma*, vol. 58, pp. 58–64, 2011.

[35] P. T. Soliman, X. Cui, Q. Zhang, S. E. Hankinson, and K. H. Lu, "Circulating adiponectin levels and risk of endometrial cancer: the prospective nurses' health study," *American Journal of Obstetrics and Gynecology*, vol. 204, no. 2, pp. 167.e1–167.e5, 2011.

[36] T. Yamaji, M. Iwasaki, S. Sasazuki, and S. Tsugane, "Interaction between adiponectin and leptin influences the risk of colorectal adenoma," *Cancer Research*, vol. 70, no. 13, pp. 5430–5437, 2010.

[37] P. Ferroni, R. Palmirotta, A. Spila et al., "Prognostic significance of adiponectin levels in non-metastatic colorectal cancer," *Anticancer Research*, vol. 27, no. 1 B, pp. 483–489, 2007.

[38] G. Li, L. Cong, J. Gasser, J. Zhao, K. Chen, and F. Li, "Mechanisms underlying the anti-proliferative actions of adiponectin in human breast cancer cells, MCF7-dependency on the cAMP/protein Kinase-A pathway," *Nutrition and Cancer*, vol. 63, no. 1, pp. 80–88, 2011.

[39] R. A. Skidgel, G. B. McGwire, and X. Y. Lix, "Membrane anchoring and release of carboxypeptidase M: implications for extracellular hydrolysis of peptide hormones," *Immunopharmacology*, vol. 32, no. 1-3, pp. 48–52, 1996.

[40] D. A. Davis, K. E. Singer, M. De La Luz Sierra et al., "Identification of carboxypeptidase N as an enzyme responsible for C-terminal cleavage of stromal cell-derived factor-1α in the circulation," *Blood*, vol. 105, no. 12, pp. 4561–4568, 2005.

[41] G. Candiano, V. Dimuccio, M. Bruschi et al., "Combinatorial peptide ligand libraries for urine proteome analysis: investigation of different elution systems," *Electrophoresis*, vol. 30, no. 14, pp. 2405–2411, 2009.