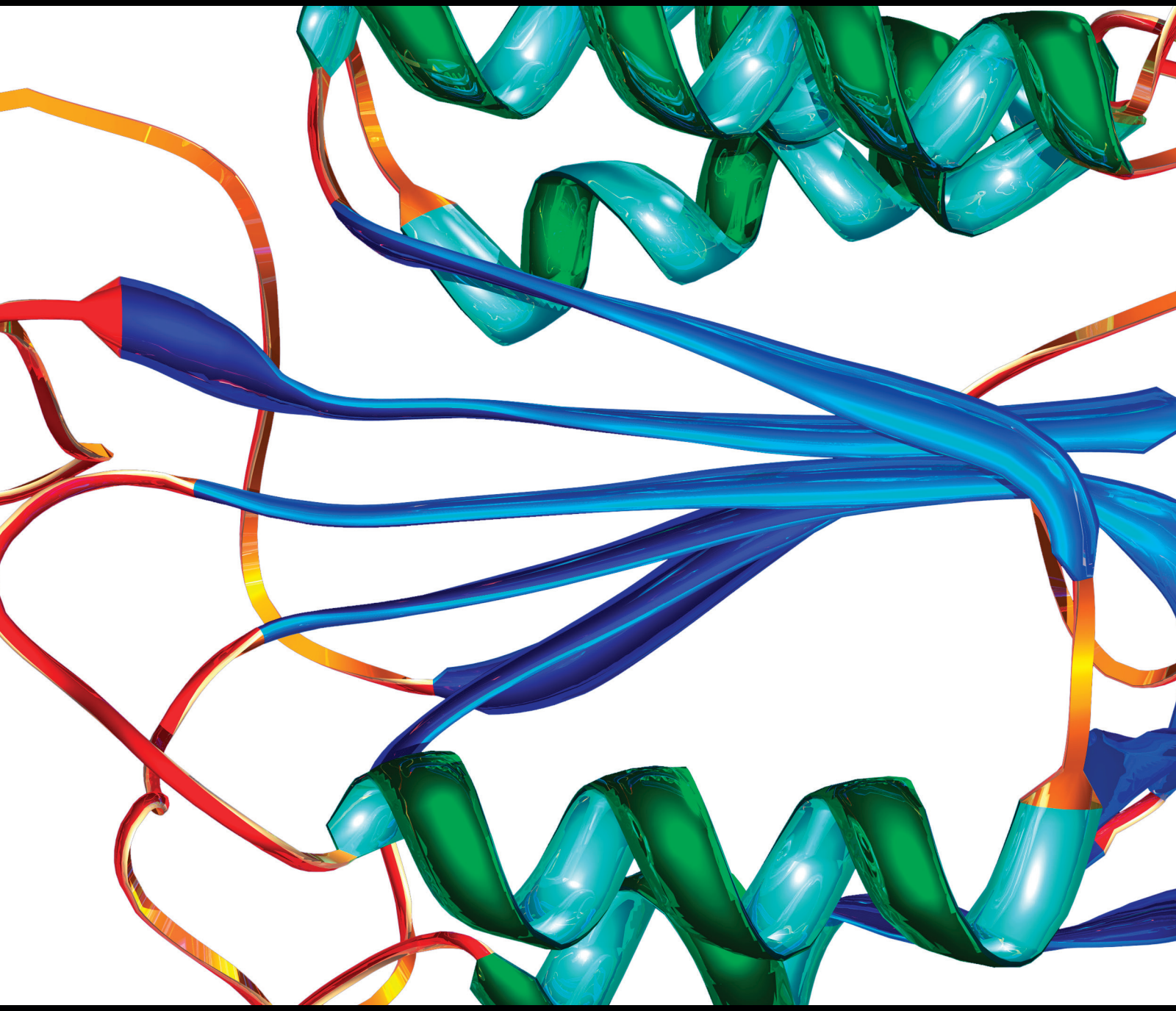
Disease Markers

High-Throughput Screening for Biomarker Discovery

Guest Editors: Tavan Janvilisri, Haruo Suzuki, Joy Scaria, Jenn-Wei Chen, and Varodom Charoensawan



High-Throughput Screening for Biomarker Discovery

Disease Markers

High-Throughput Screening for Biomarker Discovery

Guest Editors: Tavan Janvilisri, Haruo Suzuki, Joy Scaria, Jenn-Wei Chen, and Varodom Charoensawan

Editorial Board

Contents

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 108064, 2 pages
http://dx.doi.org/10.1155/2015/108064

Editorial

High-Throughput Screening for Biomarker Discovery

Tavan Janvilisri,[1] Haruo Suzuki,[2] Joy Scaria,[3] Jenn-Wei Chen,[4] and Varodom Charoensawan[1]

[1]*Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand*
[2]*Graduate School of Science and Engineering, Yamaguchi University, Yamaguchi 753-8511, Japan*
[3]*Department of Veterinary & Biomedical Sciences, College of Agriculture & Biological Sciences, South Dakota State University, Brookings, SD 57006, USA*
[4]*Center of Infectious Disease and Signaling Research, National Cheng Kung University, Tainan 701, Taiwan*

Correspondence should be addressed to Tavan Janvilisri; tavan.jan@mahidol.ac.th

Received 17 March 2015; Accepted 17 March 2015

Biomarkers are indicators of disease occurrence and progression. They can be used to assess normal biological and pathogenic processes or pharmacologic responses to a therapeutic intervention and, in some cases, may serve as potential drug and prognostic targets. In the postgenomic era, the omics techniques have been applied and served as high-throughput screening for identification of potential biomarkers for pathogenesis of several diseases including cancers and metabolic diseases, as well as infectious diseases. These biomarkers would lead to the improved understanding of the mechanism of pathogenesis and might be clinically useful as the molecular targets for better diagnosis, prognosis, and treatment.

In this special issue, there are six articles: two reviews and four original research articles. The first review article focuses on the application of proteomics in the discovery of cancer biomarkers, while the others focus specifically on biomarkers in nasopharyngeal carcinoma derived through omics approaches. Two research articles focus on protein biomarkers in cholangiocarcinoma using proteomics, one of which attempted to search for serum diagnostic markers and the other investigated the biomarkers for metastasis. Finally, the other two original papers address the relationship between gene polymorphic and miRNA biomarkers and diseases. Thus, the articles in this special issue represent a broad spectrum of experimental approaches and areas of investigation and demonstrate a wide array of molecular biomarker research.

In "Novel Serum Biomarkers to Differentiate Cholangiocarcinoma from Benign Biliary Tract Diseases Using a Proteomic Approach," T. Janvilisri et al. identified proteins in the serum that can potentially discriminate the patients with cholangiocarcinoma from individuals with benign biliary tract diseases through proteomic approach using highly stringent analysis with cross-validation. They identified potential serum molecular markers that are worth further validation and could be useful for distinguishing between these two diseases with similar appearance, leading to better therapeutic measures.

In "Proteomics in Cancer Biomarkers Discovery: Challenges and Applications," R. M. Sallam provided us with a perspective review that summarizes the potential use of proteomics approach to identify molecular targets in cancer research with examples of three of the most studied cancers including lung, breast, and ovarian cancers. This review describes the appreciation of proteomics in cancer research, which makes us understand tumor biology better, facilitates the development of biomarkers, and, most importantly, makes us move towards bedside applications in cancer management.

In "Analysis of Serum MicroRNAs as Potential Biomarker in Coronary Bifurcation Lesion," Y. Liu et al. investigated circulating miRNAs for coronary bifurcation lesion in order to identify potential biomarkers. They performed miRNA profiling using microarray to screen the serum miRNAs profiles of patients with coronary bifurcation lesion and coronary nonbifurcation lesions. The miRNAs identified in this study may play a part in pathogenesis of coronary bifurcation lesion and could serve as novel biomarkers for the diagnosis and prognosis of this disease in the future.

In "Comparative Proteomic Analysis of Human Cholangiocarcinoma Cell Lines: S100A2 as a Potential Candidate Protein Inducer of Invasion," K. Wasuworawong et al. attempted to identify metastatic protein markers in cholangiocarcinoma through protein expression profiling of highly invasive cell line, KKU-M213, and lowly invasive one, KKU-100. They proposed that S100A2 could potentially be a key protein involved in the progression of cholangiocarcinoma and may pose as a biomarker as well as a novel therapeutic target for this type of cancer.

In the review "Omics-Based Identification of Biomarkers for Nasopharyngeal Carcinoma," T. Janvilisri provided an overview of the discovery of molecular biomarkers for nasopharyngeal carcinoma through the emerging omics technologies including genomics, miRNA-omics, transcriptomics, proteomics, and metabolomics. A large number of potential biomarkers for this disease related to various pathophysiological states have been discussed. This review article, thus, could be a reference point for downstream research in the field of nasopharyngeal carcinoma biomarkers.

In "Polymorphisms in *C-Reactive Protein* and *Glypican-5* Are Associated with Lung Cancer Risk and *Gartrokine-1* Influences Cisplatin-Based Chemotherapy Response in a Chinese Han Population," S. Zhang et al. investigated polymorphic variations in seven genes including *CRP*, *GPC5*, *ACTA2*, *AGPHD1*, *SEC14L5*, *RBMS3,* and *GKN1*, which have previously been associated with lung cancer. They found the relationship between *CRP* and *GPC5* variants and risks for lung cancer. Variation in *GKN1* has also been shown to correlate to chemotherapy response in the Chinese population.

At present, the biomarker data are blooming through the omics research. It will be a challenge to merge a vast amount of data and acquire biological meanings to resolve a daunting conundrum of different diseases. Integrative approaches together with further validations will be necessary. The editorial team hopes that this special issue will be useful for investigators in the field.

Acknowledgments

Finally, we would like to thank the authors for their contributions in this special issue and all the reviewers for critical review of the manuscripts.

Tavan Janvilisri
Haruo Suzuki
Joy Scaria
Jenn-Wei Chen
Varodom Charoensawan

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 105358, 11 pages
http://dx.doi.org/10.1155/2015/105358

Research Article

Novel Serum Biomarkers to Differentiate Cholangiocarcinoma from Benign Biliary Tract Diseases Using a Proteomic Approach

Tavan Janvilisri,[1] Kawin Leelawat,[2] Sittiruk Roytrakul,[3] Atchara Paemanee,[3] and Rutaiwan Tohtong[1]

[1]*Department of Biochemistry, Faculty of Science, Mahidol University, 272 Rama VI Road, Phayathai, Rajdhevi, Bangkok 10400, Thailand*
[2]*Department of Surgery, Rajavithi Hospital, Bangkok 10400, Thailand*
[3]*Proteomics Research Laboratory, Genome Institute, National Center for Genetic Engineering and Biotechnology, NSTDA, Pathumthani 12120, Thailand*

Correspondence should be addressed to Rutaiwan Tohtong; rutaiwan.toh@mahidol.ac.th

Received 20 August 2014; Accepted 13 October 2014

Academic Editor: Joy Scaria

Background and Aim. Cholangiocarcinoma (CCA) is the most frequent biliary malignancy, which poses high mortality rate due to lack of early detection. Hence, most CCA cases are present at the advanced to late stages with local or distant metastasis at the time of diagnosis. Currently available tumor markers including CA19-9 and CEA are inefficient and of limited usage due to low sensitivity and specificity. Here, we attempt to identify serum tumor markers for CCA that can effectively distinguish CCA from benign biliary tract diseases (BBTDs). *Methods*. Serum samples from 19 CCA patients and 17 BBTDs were separated by SDS-PAGE followed with LC-MS/MS and were subjected to statistical analysis and cross-validation to identify proteins whose abundance was significantly elevated or suppressed in CCA samples compared to BBTDs. *Results*. In addition to identifying several proteins previously known to be differentially expressed in CCA and BBTDs, we also discovered a number of molecules that were previously not associated with CCA. These included FAM19A5, MAGED4B, KIAA0321, RBAK, and UPF3B. *Conclusions*. Novel serum biomarkers to distinguish CCA from BBTDs were identified using a proteomic approach. Further validation of these proteins has the potential to provide a biomarker for differentiating CCA from BBTDs.

1. Introduction

Cholangiocarcinoma (CCA) is one of the highly aggressive malignant tumors that arise from the cholangiocytes lining biliary trees [1]. The incidence and mortality of the disease continue to increase worldwide, and the highest incidence has been observed in the Southeast Asia, especially in Thailand [2, 3]. The prognosis of this malignancy is poor due to its silent clinical characteristics, difficulties in early diagnosis, and limited therapeutic measures. At present, radiotherapy and chemotherapy do not significantly improve the survival rate, while the resection of detected tumors at the early stage offers the best curative treatment [4]. Clinical presentations of most CCA patients include biliary tract obstruction; however, many cases of benign biliary tract diseases (BBTDs) are also presented with similar clinical symptoms [5]. Differences in the treatment and prognosis between CCA and BBTDs urge us a need to identify accurate tumor biomarkers that can differentially diagnose the CCA from BBTDs. As CCA typically grows along the bile duct without protruding outward as a forming mass, therefore current imaging techniques including ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI) are not efficient to reveal this lesion [6]. Laboratory assessments for CCA are often not sensitive, nor specific enough. Distinguishing between benign and malignant causes of biliary tract obstruction based on biopsies is rather difficult and usually inadequate to provide an accurate measure. Currently, determination of the serum marker carbohydrate antigen 19-9 (CA19-9) concentration is routinely applied in most laboratories for

CCA detection. However, a wide range of sensitivity (50–90%) and specificity (54–98%) of this biomarker for CAA has been reported [7–9], and the elevated serum CA19-9 has also been observed in patients with BBTDs [10, 11]; therefore, the use of CA19-9 for differentiating CCA and BBTDs is not reliable. Other serum markers including carcinoembryogenic antigen (CEA) and cancer antigen 125 (CA125) have also been used for detecting CCA, but these markers are not satisfactory for CCA detection due to low specificity and sensitivity for screening [12–14]. Hence, identification of new tumor markers in the serum would be beneficial in the clinical management of this disease.

In recent years, quantitative proteomics has gained considerable attention and investment in order to identify diagnostic biomarkers for several diseases, including a variety of cancers [15]. In the present study, the proteome of serum samples from CCA patients were quantitatively compared with that of patients with BBTDs, who have shared many molecular and imaging features with CCA. A large-scale quantitative global protein profiling of serum coupled with bioinformatic analyses would identify a proteomic signature for effectively differentiating CCA from BBTDs. Patterns of differentially serum protein expression between CCA and BBTD patients were exploited for development of diagnostic or prognostic tool for this type of cancer.

2. Methods

2.1. Serum Samples. Serum samples were collected from obstructive jaundice patients who underwent endoscopic retrograde cholangiography (ERCP) or biliary tract surgery at Rajavithi Hospital. The use of human materials was approved by the research ethics committee of Rajavithi Hospital. Seventeen patients with BBTDs and 19 CCA patients were enrolled in this study. The diagnosis of CCA was carried out using one of the following criteria: (i) tissue biopsy; (ii) cytology plus radiological (CT scan or MRI) and clinical observation to identify tumor progression at a follow-up of at least two months. Serum samples from these patients were separated by centrifugation and stored at −80°C within 1 h. The biochemical determinations of serum markers, including CEA and CA19-9, were performed using routine automated methods in the Pathological Laboratory at Rajavithi Hospital.

2.2. Sample Preparation, Electrophoresis, and Trypsin Digestion. Samples were treated with protease inhibitor cocktail and protein extraction from serum was carried out in lysis buffer containing 8 M urea and 10 mM dithiothreitol (DTT). Protein concentration was determined using Bradford protein assay with bovine serum albumin as a standard. Fifty micrograms of total serum proteins were resolved on 12.5% SDS-PAGE. The gel was then fixed for 30 min in a fixing solution containing 50% methanol, 12% acetic acid, and 0.05% formaldehyde, washed twice for 20 min in 35% ethanol, and then sensitized in 0.02% (w/v) sodium thiosulfate for 2 min with mild agitation. After washing twice for 5 min each with deionized water, the gel was then stained with 0.2% (w/v) silver nitrate for 20 min and washed twice prior to the detection in a developing solution (6% (w/v) sodium carbonate, 0.02% (w/v) sodium thiosulfate and 0.05% formalin). The staining was stopped by incubation in 1.5% Na_2 EDTA solution for 20 min. Finally, the stained gel was washed three times for 5 min each with deionized water. The gel was scanned using a GS-710 scanner (Bio-Rad, Benicia, CA) before being stored in 0.1% acetic acid until in-gel tryptic digestion.

The gel lanes were divided into 5 fractions according to the standard protein markers and then subdivided into 15 ranges. Each gel range was chopped into pieces (1 mm^3/piece), which were dehydrated in 100% acetonitrile (ACN) for 5 min with agitation and dried at room temperature for 15 min. Subsequently, the cysteine residues were blocked with 10 mM DTT in 10 mM NH_4HCO_3 for 1 h at room temperature and alkylated with 100 mM iodoacetamide in 10 mM NH_4HCO_3 for 1 h at room temperature in the dark. The gel pieces were dehydrated twice in 100% ACN for 5 min and then were incubated with 0.20 μg trypsin in 50% ACN/10 mM NH_4HCO_3 for 20 min. Purified peptide fractions were dried and reconstituted in 2% ACN and 0.1% formic acid for subsequent LC-MS/MS.

2.3. Liquid Chromatography-Tandem Mass Spectrometry (LC/MS-MS). The LC-MS/MS analysis was carried out using a Waters nanoACQUITY ultra performance liquid chromatography coupled with a SYNAPT HDMS mass spectrometer. A 5-μL aliquot of peptide fractions was injected using a built-in nanoACQUITY auto sampler onto a Symmetry C18 trapping column (200 μm × 180 mm, 5 μm particle size; Waters) at 10 μL/min flow rate for on-line desalting and then separated on a C-18 RP nano-BEH column (75 μm id × 200 mm, 1.7 μm particle size, Waters) and eluted in a 30 min gradient of 2% to 40% ACN in 0.1% formic acid (FA) at 350 nL/min, followed by a 10-min ramping to 80% ACN-0.1% FA and a 5-min holding at 80% ACN-0.1% FA. The column was reequilibrated with 2% ACN-0.1% FA for 20 min prior to the next run. The MS nanoion source contained a 10-μm analyte emitter (New Objective, Woburn, MA) and an additional 20-μm reference sprayer through which a solution of 200 fmol/μL Glu Fibrinopeptide B (Glufib) in 25% ACN-0.1% FA was constantly infused at 200 nL/min for external lock mass correction at 30 s intervals. For all measurements, the MS instrument was operated in V mode (at 10,000 resolution) with positive nanoES ion mode. The instrument was tuned and calibrated by infusion of 200-fmol/μL Glufib and set up for a spray voltage at 2.7 kV and sample cone voltage at 45 eV. The spectral acquisition time was 0.6 sec. In MS expression mode, low energy of trap was set at a constant collision energy of 6 V. In elevated energy of MS expression mode, the collision energy of trap was ramped from 15 to 40 V during each 0.6-s data collection cycle with one complete cycle of low and elevated energy. In transfer collision energy control, 4 V and 7 V were set for low and high energy, respectively. The quadrupole mass analyzer was adjusted such that ions from m/z 200 to 1990 were efficiently transmitted.

Table 1: Clinical characteristics of patients with benign biliary tract diseases (BBTDs) and cholangiocarcinoma (CCA) in this study.

Characteristics	BBTDs	CCA	*P* values
Number of patients	17	19	—
Sex			
(Male : female)	10 : 7	10 : 9	0.085
Age (years)			
Mean ± S.D.	52.6 ± 13	60.9 ± 13	0.285
CEA (U/mL)			
Median	8.70	9.87	0.711
(Min–max)	(0.62–118.30)	(1.47–410.40)	
CA19-9 (ng/mL)			
Median	48.03	4355.50	0.015
(Min–max)	(0.60–10000)	(0.60–10000)	

2.4. Data Processing, Protein Identification, and Data Analysis. Continuum LC-MS data were processed using ProteinLynx Global Server version 2.4 (Waters) for ion detection, clustering, and mass correction. Protein identification was performed with the embedded ion accounting algorithm against NCBI human protein database with the minimum cutoffs of two peptides/proteins. The relative quantitation ratios were $\log_2$-transformed, processed with median normalization for each sample and rank normalization across the data set. The data were subjected to a 6-fold cross-validation. A differentially expressed (DE) protein was defined as having a *P* value of <0.01, based on *t*-distribution with Welch approximation, in all data sets in the fold validation. The visualization and statistical analyses were performed using the MultiExperiment Viewer (MeV) in the TM4 suite software [16]. Other information including protein categorization and biological function was analyzed according to protein analysis through evolutionary relationships (Panther) protein classification [17]. Known and predicted functional interaction networks of identified proteins were derived from the STRING database version 9.1 [18].

2.5. Statistical Analysis. Comparisons between the quantitative variables were performed using either the Mann-Whitney *U* or Student's *t*-test, where appropriate. Qualitative variables were reported as counts, and comparisons between independent groups were performed using Pearson Chi-squared tests. *P* values of less than 0.05 were considered significant.

3. Results

3.1. Patient Characteristics. A total of 36 subjects were included in this serum proteome study, of which 17 were diagnosed as having BBTDs and 19 were diagnosed as having CCA. The BBTD cases included intrahepatic duct stones, common bile duct stones, and benign bile duct strictures. The CCA cases included perihilar cholangiocarcinoma, intrahepatic cholangiocarcinoma, and middle and distal common bile duct cancer. The clinical characteristics of the patients in this study are shown in Table 1. No statistically significant differences were found among the data of the BBTD patients and those with CCA regarding gender, age, and CEA. Although the level of CA19-9 in the serum of patients with CCA was significantly higher when compared to the control patients, the range of detection in both groups was exactly the same (0.60–10000).

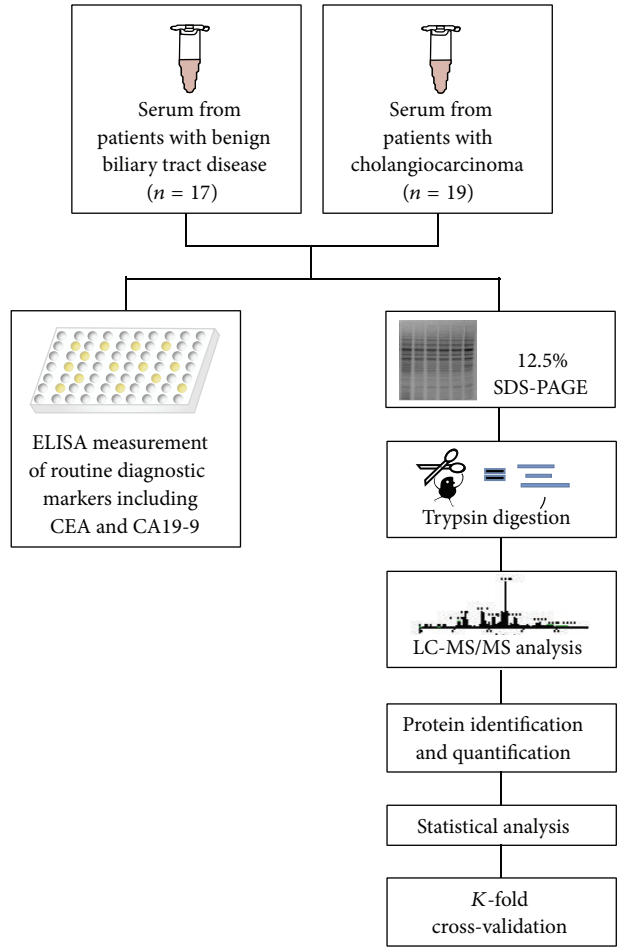


Figure 1: Schematic diagram of the experimental workflow. Serum samples were collected from 17 BBTD patients and 19 CCA patients, which were then subjected to routine ELISA for CEA and CA19-9. Purified proteins from these samples were then separated by SDS-PAGE. After migration, entire lanes were divided into 5 sections, which were excised into slices and treated with in-gel digestion. The resulting tryptic peptides were subjected to reverse-phase LC-MS/MS, from which the mass spectrometric results were then analyzed for protein identification and quantification. The relative quantitation ratios were subjected to statistical analyses and 6-fold cross-validation to retrieve the DE proteins between BBTDs and CCA.

3.2. Serum Proteome Profiling. An overview of the experimental strategy conducted in this study is shown in Figure 1. The proteome of serum samples from CCA patients was compared with the serum proteome of the BBTD controls in order to identify the proteins in serum, in particular those that are secreted or leaked from tissues including potential differential protein biomarkers from tumor cells. A total of

Table 2: A list of differentially expressed serum proteins between CCA and BBTDs. The protein expression measurements were averaged and represented as $\log_2$-transformed intensity values with standard deviation. The *P* values are also indicated.

Protein names	ID details	GI accession	CCA mean ± SD	BBTD mean ± SD	*P* values
CCA > BBTD					
ABHD11	Alpha/beta hydrolase domain-containing protein 11	23200008	1.20 ± 0.4	0.82 ± 0.3	0.002
—	Antioxidized LDL immunoglobulin light chain variable region	62868476	0.88 ± 0.3	0.55 ± 0.3	0.004
—	Chain L, crystal structure of the Fab fragment of nimotuzumab. An antiepidermal growth factor receptor antibody	255311843	1.93 ± 0.2	1.72 ± 0.3	0.010
—	Complement factor H	758073	−0.86 ± 0.3	−1.51 ± 0.5	0.001
COG7	Conserved oligomeric Golgi complex subunit 7	23957690	−0.50 ± 0.4	−1.23 ± 0.5	<0.001
DHDDS	Dehydrodolichyl diphosphate synthase	13177736	0.14 ± 0.6	−0.61 ± 0.5	0.005
MLH1	DNA mismatch repair protein	4557757	−0.49 ± 0.5	−1.00 ± 0.5	0.004
EIF3J	Eukaryotic translation initiation factor 3 subunit J	83281438	2.40 ± 0.4	1.95 ± 0.4	0.002
FAM19A5	FAM19A5 protein	71052198	0.98 ± 0.3	−0.69 ± 0.7	<0.001
HBZ	Hemoglobin subunite zeta	4885397	−0.23 ± 0.3	−0.78 ± 0.4	0.001
V4-34	IgG	2632200	0.54 ± 0.4	0.20 ± 0.2	0.001
—	Immunoglobulin heavy chain variable region	37694587	1.46 ± 0.4	0.93 ± 0.5	0.003
IGK	Immunoglobulin kappa light chain VLJ region	21669309	1.98 ± 0.3	1.68 ± 0.1	<0.001
IL16	Interleukin 16	119619506	0.15 ± 0.4	−0.36 ± 0.6	0.008
KIAA0321	KIAA0321 protein	2224583	−0.25 ± 0.8	−1.70 ± 0.9	0.002
KIAA0612	KIAA0612 protein	34327964	−1.17 ± 0.4	−1.76 ± 0.5	0.006
KIAA0896	KIAA0896 protein	71891755	0.18 ± 0.6	−0.50 ± 0.6	0.003
MAGED4B	Melanoma-associated antigen D4	29337296	0.33 ± 0.3	−1.17 ± 0.9	0.002
NXF3	Nuclear RNA export factor 3	11545757	0.59 ± 0.4	−0.29 ± 0.5	<0.001
PAXBP1	PAX3- and PAX7-binding protein 1	22035565	0.64 ± 0.1	0.39 ± 0.2	0.001
LOC390791	Peptidyl-prolyl cis-trans isomerase A-like	310113085	−0.50 ± 0.3	−1.24 ± 0.4	0.002
PLEKHO2	Pleckstrin homology domain-containing family O member 2	33457316	−0.29 ± 0.4	−0.91 ± 0.4	0.001
PLEKHM2	PLEKHM2 protein	26251859	−0.46 ± 0.4	−1.06 ± 0.5	0.002
RBAK	RB-associated KRAB zinc finger protein	13430850	0.17 ± 0.5	−0.96 ± 0.7	<0.001
PTPRG	Receptor tyrosine phosphatase gamma	1263069	0.90 ± 0.3	0.54 ± 0.2	<0.001
RPS10	Ribosomal protein S10	3088338	0.99 ± 0.2	0.67 ± 0.1	<0.001
NOB1	RNA-binding protein NOB1	7661532	−0.81 ± 0.4	−1.26 ± 0.4	0.002
VAT1	Synaptic vesicle membrane protein VAT-1	18379349	0.57 ± 0.4	0.15 ± 0.3	0.003
TRAC-1	T3 receptor-associating cofactor-1	1911770	0.12 ± 0.3	−0.28 ± 0.2	<0.001
—	Unnamed protein product	10433849	1.67 ± 0.5	0.77 ± 0.3	<0.001
—	Unnamed protein product	21752201	0.27 ± 0.2	−0.07 ± 0.3	0.001
CCA < BBTD					
GCAT	2-Amino-3-ketobutyrate coenzyme A ligase	7657118	−1.62 ± 0.5	−1.17 ± 0.3	0.002
ALB	Albumin	119626083	1.05 ± 0.3	1.45 ± 0.3	<0.001
SERPINA1	Alpha-1-antitrypsin	1703025	0.84 ± 0.4	1.24 ± 0.3	<0.001
A2M	Alpha-2-macroglobulin	177872	−0.04 ± 0.4	0.52 ± 0.3	<0.001
AGT	Angiotensinogen	4261988	0.09 ± 0.6	0.79 ± 0.2	<0.001
apo AII	Apolipoprotein	671882	0.04 ± 0.3	0.31 ± 0.2	0.003
APOB	Apolipoprotein B-100	105990532	−1.67 ± 0.4	−0.86 ± 0.5	<0.001
ARID5B	AT-rich interactive domain-containing protein 5B	74136549	0.08 ± 0.5	0.71 ± 0.7	0.006
BAZ2	BWSCR2 associated zinc-finger protein BAZ2	6002480	−0.50 ± 0.9	0.23 ± 0.3	0.005
C1orf87	C1orf87 protein	27503780	−0.91 ± 0.4	−0.49 ± 0.3	0.001

Table 2: Continued.

Protein names	ID details	GI accession	CCA mean ± SD	BBTD mean ± SD	*P* values
PDEA	cGMP phosphodiesterase	2366987	0.11 ± 0.4	0.73 ± 0.3	<0.001
	Chain D, The Nucleosome Containing A Testis-Specific Histone Variant	296863399	−2.56 ± 0.4	−2.04 ± 0.5	0.002
C4A	Complement C4-A	476007827	−0.81 ± 0.7	−0.02 ± 0.3	<0.001
DCAF15	DDB1- and CUL4-associated factor 15	78486540	0.87 ± 0.4	1.31 ± 0.2	<0.001
FN1	Fibronectin 1	53791223	−0.98 ± 0.5	−0.29 ± 0.4	<0.001
FLJ00044	FLJ00044 protein	10440418	−1.02 ± 0.4	−0.59 ± 0.2	0.004
FLJ16008	FLJ16008 protein, isoform CRA_b	119615716	1.79 ± 0.1	1.94 ± 0.2	0.006
GNG5	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5	4885287	0.94 ± 0.3	1.24 ± 0.1	<0.001
hCG_1817987	hCG1817987	119612015	−0.09 ± 0.3	0.36 ± 0.3	<0.001
hCG_1981701	hCG1981701	119572460	−1.14 ± 0.5	−0.52 ± 0.5	0.002
hCG_2008076	hCG2008076	119592316	0.05 ± 0.4	0.42 ± 0.2	0.012
hCG_2008267	hCG2008267	119592800	−0.07 ± 0.4	0.35 ± 0.2	0.001
hCG_201157	hCG201157	119576573	0.64 ± 0.3	0.89 ± 0.3	0.009
hCG_2020343	hCG2020343	119629275	−0.95 ± 0.4	−0.27 ± 0.4	<0.001
—	Hypothetical protein	12224988	−1.46 ± 0.5	−0.90 ± 0.4	<0.001
FLJ22688	Hypothetical protein FLJ22688, isoform CRA_b	119572924	−0.93 ± 0.6	−0.17 ± 0.4	<0.001
LOC286076	Hypothetical protein LOC286076	119602615	1.19 ± 0.4	1.52 ± 0.2	0.003
IgA1	Ig Aalpha1 Bur	223099	1.84 ± 0.6	2.28 ± 0.3	0.006
	Immunoglobulin heavy chain variable region	37694587	1.48 ± 0.3	1.70 ± 0.1	0.007
ITIH1	Interalpha (globulin) inhibitor H1	825681	−0.74 ± 0.4	0.08 ± 0.6	0.001
ITIH2	Interalpha (globulin) inhibitor H2	119606784	−1.44 ± 0.5	−0.54 ± 0.4	<0.001
KRT1	Keratin 1	11935049	−1.12 ± 0.3	−0.56 ± 0.4	<0.001
KRT10	Keratin-10	307086	−2.40 ± 0.5	−1.50 ± 0.6	<0.001
KIAA0366	KIAA0366 protein	2224673	−1.45 ± 0.4	−0.61 ± 0.6	<0.001
KIAA0920	KIAA0920 protein	40788986	1.05 ± 0.3	1.37 ± 0.2	<0.001
KIAA1234	KIAA1234 protein	6330736	−0.70 ± 0.4	−0.38 ± 0.3	0.007
KIAA1529	KIAA1529 protein	7959325	1.33 ± 0.3	2.38 ± 0.4	<0.001
MAGEB2	Melanoma-associated antigen B2	222418639	−0.48 ± 0.3	−0.07 ± 0.4	0.001
MTDH	Metadherin	119612168	−0.45 ± 0.3	−0.09 ± 0.3	<0.001
MUC16	Mucin-16	74716283	−0.49 ± 0.5	0.41 ± 0.4	<0.001
MYOT	Myotilin	5803106	−0.41 ± 0.5	0.30 ± 0.3	<0.001
NPTX1	Neuronal pentraxin 1	1438954	1.51 ± 0.3	1.74 ± 0.2	0.010
PLG	Plasminogen	38051823	1.67 ± 0.3	2.17 ± 0.4	<0.001
GALNT2	Polypeptide N-acetylgalactosaminyltransferase 2	4758412	−0.99 ± 0.3	−0.50 ± 0.3	<0.001
FAM83E	Protein FAM83E	153251792	0.70 ± 0.5	1.33 ± 0.3	0.006
LOC100131107	Putative UPF0607 protein ENSP00000383783	239741331	1.03 ± 0.7	2.07 ± 0.7	0.002
RAB-R	RAB-R protein	4102709	−0.83 ± 0.4	−0.05 ± 0.3	0.002
UPF3B	Regulator of nonsense transcripts 3B	18375528	−2.17 ± 0.5	−0.99 ± 0.6	<0.001
RIMBP3	RIMBP3 protein	71052030	0.07 ± 0.4	0.69 ± 0.4	<0.001
	Suppressor of cytokine signaling 3	54695958	0.85 ± 0.4	1.23 ± 0.2	0.004
	Testis specific kinase-1	21886788	−1.65 ± 0.3	−1.07 ± 0.7	0.005
TTC34	Tetratricopeptide repeat protein 34	239741018	−0.54 ± 0.5	0.34 ± 0.3	<0.001
TPO	Thyroid peroxidase	4680721	0.69 ± 0.4	1.15 ± 0.2	<0.001
—	Ubiquitously transcribed tetratricopeptide repeat protein Y-linked transcript	148733192	−1.43 ± 0.4	−0.88 ± 0.5	0.001

Table 2: Continued.

Protein names	ID details	GI accession	CCA mean ± SD	BBTD mean ± SD	*P* values
—	Unnamed protein product	34531956	0.84 ± 0.3	1.17 ± 0.2	<0.001
—	Unnamed protein product	10435479	0.76 ± 0.4	1.15 ± 0.3	0.003
—	Unnamed protein product	194384842	1.24 ± 0.3	1.67 ± 0.2	<0.001
—	Unnamed protein product	22760231	1.14 ± 0.5	1.77 ± 0.3	<0.001
—	Unnamed protein product	194381130	−2.08 ± 0.4	−1.34 ± 0.5	<0.001
DBP	Vitamin D-binding protein	455970	0.88 ± 0.5	1.30 ± 0.3	0.007
ZNF410	Zinc finger protein 410	119601547	−0.27 ± 0.4	0.46 ± 0.3	<0.001
ZnF_RBZ	ZIS1	4191327	0.03 ± 0.4	0.55 ± 0.4	0.001

951 proteins were identified in all samples. Among these, the ones with altered expression levels in the serum of CCA patients compared to those of BBTD patients were identified. To reduce the effect of biological and experimental variations and the possibility of false-positive protein identification, 6-fold cross-validations were performed. In each fold, BBTD and CCA samples were randomly split into a training set (30 cases with 13–15 BBTD and 15–17 CCA) and an independent validation set (6 cases with 2–4 BBTD and 2–4 CCA). Only proteins identified and quantifiable in all folds in cross-validation were further analyzed, allowing for stringent and sensitive protein identification and quantification of differential proteins.

3.3. Identification of Differentially Expressed Proteins between CCA and BBTDs. Applying a *P* value cutoff of <0.01 yielded a total of 94 candidate proteins, with 32 of them up and 62 down in observed abundance for the serum samples from CCA patients comparing to the BBTD controls (Table 2 and Figure 2(a)). We also tested the discriminatory power of these differentially expressed proteins using unsupervised hierarchical clustering. As shown in Figure 2(c), the spectral counts for these proteins resulted in near complete separation of the CCA cases from the BBTD control cases with only two exceptions where BBTD cases were clustered with the CCA samples. However, the PCA scores plot based on the normalized data of serum samples showed a clear separation between the CCA patients and BBTD controls (Figure 2(b)).

The Panther classification system was used to identify the functional attributes of the 94 potential CCA-selective proteins. The analysis of the abundance of each functional category revealed substantial differences in CCA serum proteome compared to the BBTD serum proteome. The number of each functional class of differentially expressed proteins is schematically depicted in Figure 3. The analysis revealed significant enrichment of proteins related to a number of various biological functions such as cell adhesion molecules, cytoskeletal proteins, defense/immunity proteins, enzymes and the modulators, extracellular matrix proteins, membrane traffic proteins, nucleic acid-binding proteins, receptors, signaling molecules, structural proteins, transcription factors, transfer/carrier proteins, and transporters. To gain an overview of the biological interaction among the identified proteins, we also constructed the protein-protein functional networks using String database (Figure 4). The protein network analysis provides us a clearer view of a complex framework of proteins that might result in the differences in CCA and BBTDs.

To determine the distinguishing performance of the top five differentially expressed proteins in terms of fold-change, the comparison of the averaged $\log_2$ folds of family with sequence similarity 19 (chemokine (C-C motif)-like), member A5 (FAM19A5) protein, KIAA0321 protein, melanoma-associated antigen D4 (MAGED4B), RB-associated KRAB zinc finger protein (RBAK), and regulator of nonsense transcripts 3B (UPF3B), between CCA and BBTD cases from all cross-validation cohorts was shown in Figure 5. However, due to the limited resources and the lack of availability of an independent validation set, the diagnostic relevance of such molecules for CCA requires further investigation.

4. Discussion

CCA is the second most prevalent primary hepatobiliary malignancy and represents about 3% of all gastrointestinal cancers [1]. It is associated with inflammatory conditions in the biliary system, and patients with risk factors such as primary sclerosing cholangitis and liver fluke infestations have a higher risk for CCA development [1–3]. The generally late clinical presentation of CCA results in a high mortality. At present, the most commonly studied and routinely used serum biomarkers for detecting CCA include CEA and CA19-9 [6]. However, they are nonspecific to CCA and can be elevated in the setting of other gastrointestinal malignancies or other benign conditions, such as cholangitis, cirrhosis, and hepatolithiasis [7–14]. Based on the results in this study, both CEA and CA19-9 could not also distinguish the patients with CCA and BBTDs in our sample cohort as both appeared to be nonspecific for either case. Hence, there is an urgent need for new diagnostic targets. In this study, we evaluated the differential proteome in the serum between the BBTD controls and CCA patients and identified potential biomarker panels to aid in the diagnosis of these common liver diseases.

Total proteins were retrieved from the whole serum without the depletion of high abundant proteins due to the fact that additional steps may not help enrich the level of low abundant proteins and may reduce reproducibility from one sample to the others [19]. Among the identified

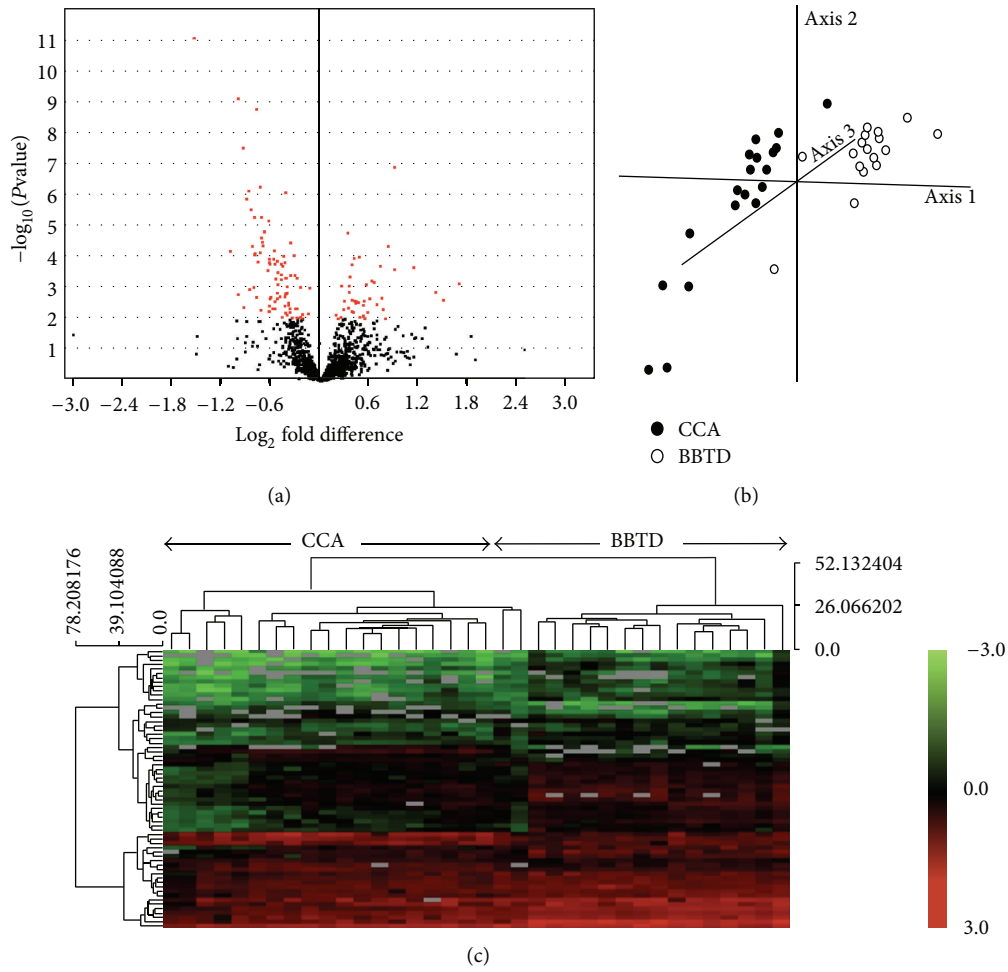


Figure 2: An overview of the DE proteins between BBTDs and CCA. (a) Volcano plots on $\log_2$ fold change and probability values between BBTDs and CCA cases. Red dots correspond to the identified DE proteins that were cut off at $P < 0.01$. (b) Principle component analysis for DE proteins. Three-dimensional scatter plot represents two clusters of BBTDs and CCA cases based on the DE proteins. Each dot represents a patient with either BBTDs or CCA, as indicated. (c) A hierarchical clustering analysis was carried out on the basis of the expression pattern. The DE proteins were linked together according to their expression (dendrogram on left). BBTD and CCA patients were also clustered (dendrogram on top). The protein-expression intensities were standardized between −3.0 (green) and 3.0 (red).

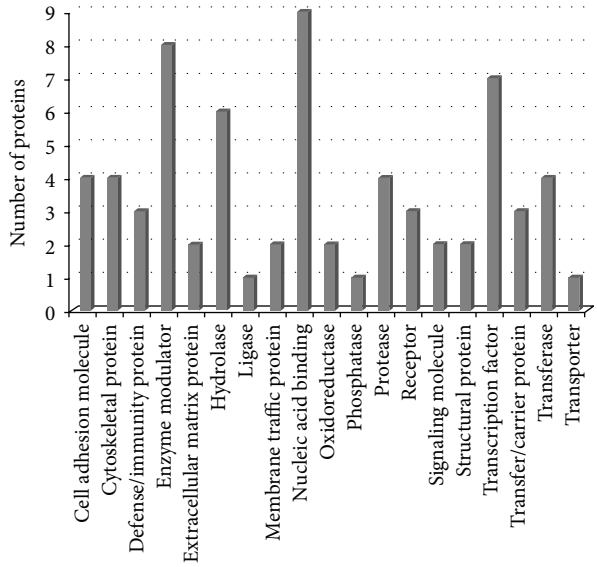


Figure 3: Distribution of DE proteins between BBTDs and CCA according to PANTHER protein classes. The bar chart shows the number of DE proteins in each functional class.

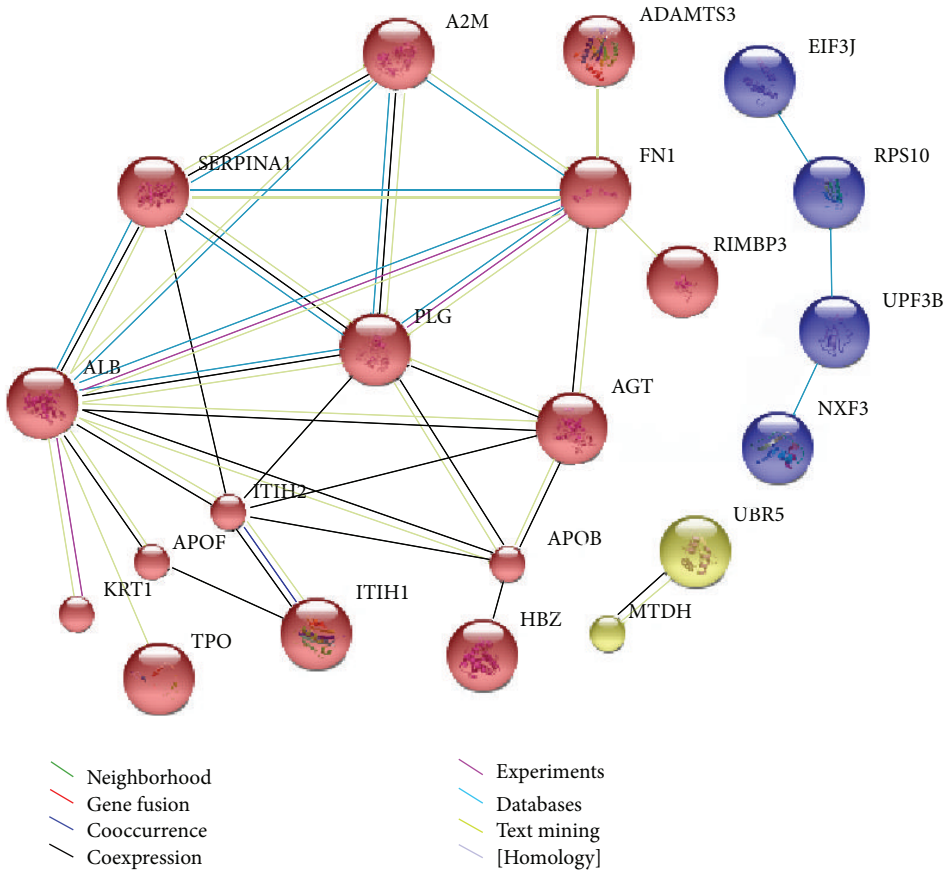


Figure 4: Visualization of protein interaction networks of the DE proteins. Inputting all DE proteins into STRING yielded a network visualizing linkages. The network nodes are proteins, whereas the edges represent the functional associations. Different line colors of edges represent different types of evidence for the association, as indicated in the figure.

proteins, we found that a number of them had previously been described in the context of CCA, confirming the validity of our quantitative proteomic approach. These included overexpression of MAGED4 [20] and DNA mismatch repair protein (MLH1) [21, 22], downregulation of albumin (ALB) [20], apolipoprotein B (APOB) [20], apolipoprotein A-II (APOA2) [20], and interalpha (globulin) inhibitor H1 (ITIH1) [20, 23]. Expression of serum alpha 1-macroglobulin (A2M) was found to be significantly higher in BBTD compared to CCA patients. Consistently, it has also been reported that the serum A2M increased in patients with liver malignancies including CCA but markedly elevated in hepatic cirrhosis [24]. Fibronectin 1 (FN1) in serum of CCA patients seemed to be lower than that of BBTD patients. Biliary FN1 has been reported as a differential biomarker of benign and malignant diseases [25]. Similarly, serum plasminogen (PLG) of CCA cases was significantly lower than that of BBTD controls. PLG in malignant livers including CCA has been demonstrated to be lower than that of the cirrhosis patients [26]. Other serum proteins were also found differentially expressed between CCA and BBTD including angiotensinogen (AGT), ADAM metallopeptidase with thrombospondin type 1 motif 3 (ADAMTS3), hemoglobin, zeta (HBZ), keratin-1 (KRT1), keratin-10 (KRT10), and serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), and member 1 (SERPINA1). However, the validation of these identified proteins is needed in order to determine if they can be clinically useful as differential biomarkers for CCA and BBTD.

The top five proteins which exhibited the maximal fold change between CCA and BBTD consisted of FAM19A5, MAGED4B, KIAA0321, RBAK, and UPF3B. FAM19A5 belongs to the TAFA family of small secreted proteins, which are brain-specific and distantly related to MIP-1 alpha, a member of the CC-chemokine family [27]. This family of proteins has been postulated to function as brain-specific chemokines or neurokines that act as regulators of immune and nervous cells, although the association of this protein and CCA pathogenesis has yet to be evaluated. For MAGED4B, its overexpression has been linked to malignant tumors and poor patient outcome in many types of cancer including breast [28], oral squamous cell carcinoma [29], and hepatocellular carcinoma [30]. However, there are no data available on the expression and the diagnostic or prognostic relevance of MAGED4B in CCA and BBTDs. KIAA0321 is a zinc finger FYVE domain-containing protein, which mediates binding of these proteins to membrane lipids and may be involved in the abscission step of cytokinesis. However, the relevance of

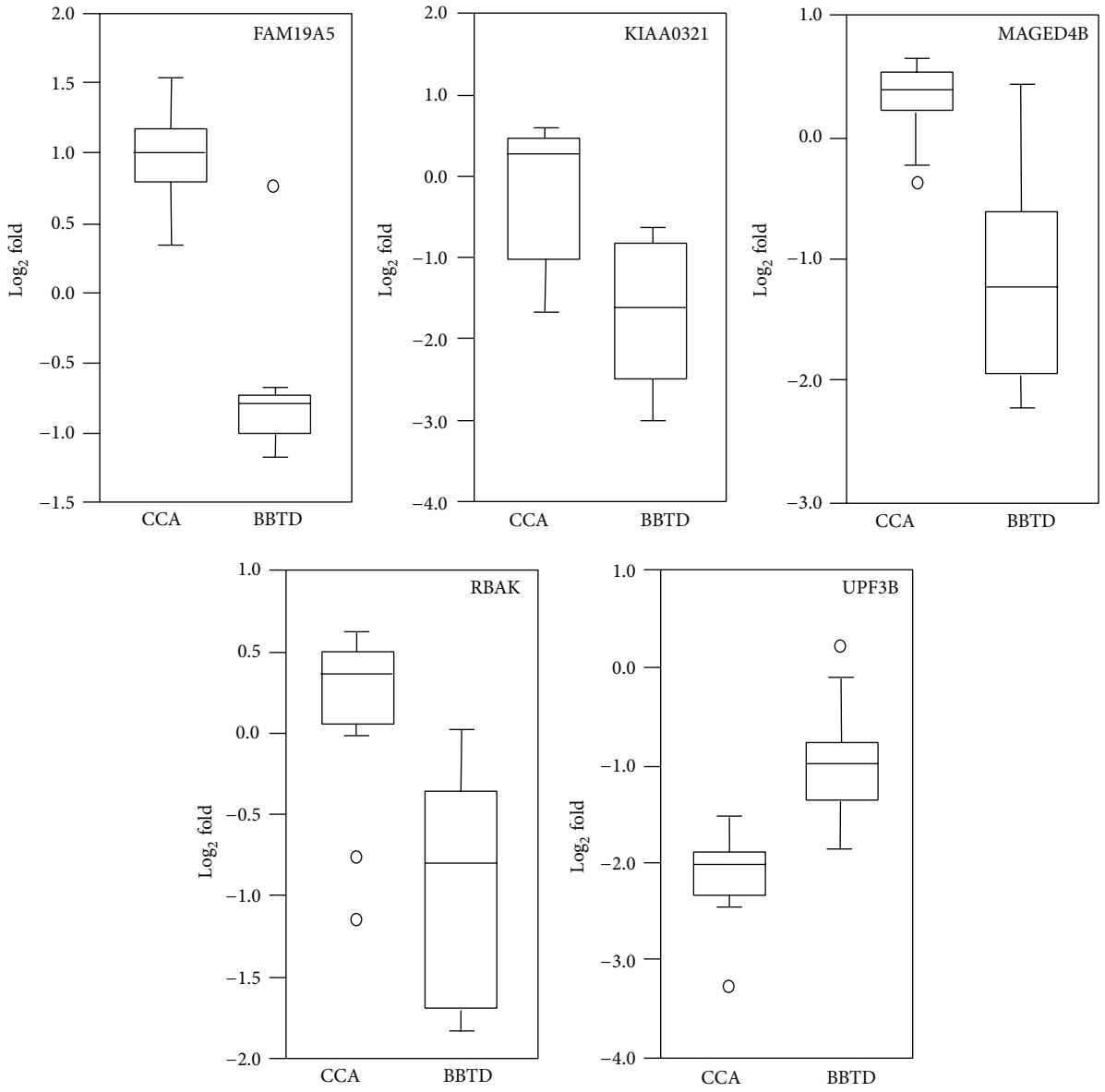


Figure 5: Comparison of the top five differentially expressed proteins between BBTDs and CCA. Normalized $\log_2$-transformed data were used to create box plots, in which the horizontal lines of each box correspond to the first, second, and third quartiles (25%, 50%, and 75%, resp.) and the whiskers correspond to the maximum and minimum values.

this protein and cancer development is yet to be elucidated [31]. RBAK is a member of a known family of transcriptional repressors that contain zinc fingers of the Kruppel type, which interacts with the tumor suppressor retinoblastoma 1. It has been shown that RBAK is expressed ectopically in human fibroblast cells [32]. Since fibroblasts in the stroma of desmoplastic cancers provide optimal microenvironment for CCA progression and they usually become susceptible for apoptosis [33], it would therefore be possible that overexpression of serum RBAK in CCA patients may be from apoptogenic cancer-associated fibroblasts. UPF3B has been reported to be overexpressed in the patients with alcoholic hepatitis [34], but there is currently no link on UPF3B and cancer yet.

In conclusion we identified proteins in the serum that can potentially discriminate patients with CCA from BBTD individuals through proteomic approach using highly stringent analysis with cross-validation. These proteins will be clinically useful to prevent misdiagnosis between CCA and BBTD as they have similar clinical symptoms. Further independent validation of these biomarkers is certainly required using greater numbers of samples from patients with CCA and a wider range of BBTD conditions to test its robustness and obtain the ones with the greatest diagnostic power for differentiating patients with CCA from BBTD controls.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the Faculty of Science, Mahidol University, with funding from a research grant for midcareer university faculty from Thailand Research Fund and Office of the Higher Education Commission (RMU5080066) to RT.

References

[1] B. Blechacz and G. J. Gores, "Cholangiocarcinoma: advances in pathogenesis, diagnosis, and treatment," *Hepatology*, vol. 48, no. 1, pp. 308–321, 2008.

[2] Y. Shaib and H. B. El-Serag, "The epidemiology of cholangiocarcinoma," *Seminars in Liver Disease*, vol. 24, no. 2, pp. 115–125, 2004.

[3] B. Sripa and C. Pairojkul, "Cholangiocarcinoma: lessons from Thailand," *Current Opinion in Gastroenterology*, vol. 24, no. 3, pp. 349–356, 2008.

[4] T. Nakagohri, T. Kinoshita, M. Konishi, S. Takahashi, and N. Gotohda, "Surgical outcome and prognostic factors in intrahepatic cholangiocarcinoma," *World Journal of Surgery*, vol. 32, no. 12, pp. 2675–2680, 2008.

[5] F.-T. Deng, Y.-X. Li, L. Ye, L. Tong, X.-P. Yang, and X.-Q. Chai, "Hilar inflammatory pseudotumor mimicking hilar cholangiocarcinoma," *Hepatobiliary & Pancreatic Diseases International*, vol. 9, no. 2, pp. 219–221, 2010.

[6] B. E. Van Beers, "Diagnosis of cholangiocarcinoma," *HPB*, vol. 10, no. 2, pp. 87–93, 2008.

[7] A. H. Patel, D. M. Harnois, G. G. Klee, N. F. Larusso, and G. J. Gores, "The utility of CA 19-9 in the diagnoses of cholangiocarcinoma in patients without primary sclerosing cholangitis," *The American Journal of Gastroenterology*, vol. 95, no. 1, pp. 204–207, 2000.

[8] X.-L. Qin, Z.-R. Wang, J.-S. Shi, M. Lu, L. Wang, and Q.-R. He, "Utility of serum CA19-9 in diagnosis of cholangiocarcinoma: in comparison with CEA," *World Journal of Gastroenterology*, vol. 10, no. 3, pp. 427–432, 2004.

[9] C. Levy, J. Lymp, P. Angulo, G. J. Gores, N. Larusso, and K. D. Lindor, "The value of serum CA 19-9 in predicting cholangiocarcinomas in patients with primary sclerosing cholangitis," *Digestive Diseases and Sciences*, vol. 50, no. 9, pp. 1734–1740, 2005.

[10] S. L. Ong, A. Sachdeva, G. Garcea et al., "Elevation of carbohydrate antigen 19.9 in benign hepatobiliary conditions and its correlation with serum bilirubin concentration," *Digestive Diseases and Sciences*, vol. 53, no. 12, pp. 3213–3217, 2008.

[11] A. Principe, M. Del Gaudio, G. L. Grazi, U. Paolucci, and A. Cavallari, "Mirizzi syndrome with cholecysto-choledocal fistula with a high CA19-9 level mimicking biliary malignancies: a case report," *Hepato-Gastroenterology*, vol. 50, no. 53, pp. 1259–1262, 2003.

[12] Y. Nakanuma and M. Sasaki, "Expression of blood group-related antigens in the intrahepatic biliary tree and hepatocytes in normal livers and various hepatobiliary diseases," *Hepatology*, vol. 10, no. 2, pp. 174–178, 1989.

[13] A. Nakeeb, P. A. Lipsett, K. D. Lillemoe et al., "Biliary carcinoembryonic antigen levels are a marker for cholangiocarcinoma," *The American Journal of Surgery*, vol. 171, no. 1, pp. 147–152, 1996.

[14] C.-Y. Chen, S.-C. Shiesh, H.-C. Tsao, and X.-Z. Lin, "The assessment of biliary CA 125, CA 19-9 and CEA in diagnosing cholangiocarcinoma—the influence of sampling time and hepatolithiasis," *Hepato-Gastroenterology*, vol. 49, no. 45, pp. 616–620, 2002.

[15] Y. Zhao, W.-N. P. Lee, and G. G. Xiao, "Quantitative proteomics and biomarker discovery in human cancer," *Expert Review of Proteomics*, vol. 6, no. 2, pp. 115–118, 2009.

[16] A. I. Saeed, N. K. Bhagabati, J. C. Braisted et al., "TM4 microarray software suite," *Methods in Enzymology*, vol. 411, pp. 134–193, 2006.

[17] H. Mi, A. Muruganujan, J. T. Casagrande, and P. D. Thomas, "Large-scale gene function analysis with the PANTHER classification system," *Nature Protocols*, vol. 8, no. 8, pp. 1551–1566, 2013.

[18] A. Franceschini, D. Szklarczyk, S. Frankild et al., "STRING v9.1: protein-protein interaction networks, with increased coverage and integration," *Nucleic Acids Research*, vol. 41, pp. D808–D815, 2013.

[19] J. Bruix and J. M. Llovet, "Prognostic prediction and treatment strategy in hepatocellular carcinoma," *Hepatology*, vol. 35, no. 3, pp. 519–524, 2002.

[20] M.-H. Chen, K.-J. Lin, W.-L. R. Yang et al., "Gene expression-based chemical genomics identifies heat-shock protein 90 inhibitors as potential therapeutic drugs in cholangiocarcinoma," *Cancer*, vol. 119, no. 2, pp. 293–303, 2013.

[21] T. Limpaiboon, P. Khaenam, P. Chinnasri et al., "Promoter hypermethylation is a major event of *hMLH1* gene inactivation in liver fluke related cholangiocarcinoma," *Cancer Letters*, vol. 217, no. 2, pp. 213–219, 2005.

[22] U. Liengswangwong, A. Karalak, Y. Morishita et al., "Immunohistochemical expression of mismatch repair genes: a screening tool for predicting mutator phenotype in liver fluke infection-associated intrahepatic cholangiocarcinoma," *World Journal of Gastroenterology*, vol. 12, no. 23, pp. 3740–3745, 2006.

[23] I. Subrungruang, C. Thawornkuno, C.-P. Porntip, C. Pairojkul, S. Wongkham, and S. Petmitr, "Gene expression profiling of intrahepatic cholangiocarcinoma," *Asian Pacific Journal of Cancer Prevention*, vol. 14, no. 1, pp. 557–563, 2013.

[24] S. Changbumrung, P. Migasena, V. Supawan, P. Juttijudata, and T. Buavatana, "Serum protease inhibitors in opisthorchiasis, hepatoma, cholangiocarcinoma, and other liver diseases," *The Southeast Asian Journal of Tropical Medicine and Public Health*, vol. 19, no. 2, pp. 299–305, 1988.

[25] C.-Y. Chen, X.-Z. Lin, H.-C. Tsao, and S.-C. Shiesh, "The value of biliary fibronectin for diagnosis of cholangiocarcinoma," *Hepato-Gastroenterology*, vol. 50, no. 52, pp. 924–927, 2003.

[26] H. Alkim, S. Ayaz, N. Sasmaz, P. Oguz, and B. Sahin, "Hemostatic abnormalities in cirrhosis and tumor-related portal vein thrombosis," *Clinical and Applied Thrombosis/Hemostasis*, vol. 18, no. 4, pp. 409–415, 2012.

[27] Y. T. Tang, P. Emtage, W. D. Funk et al., "TAFA: a novel secreted family with conserved cysteine residues and restricted expression in the brain," *Genomics*, vol. 83, no. 4, pp. 727–734, 2004.

[28] S. Germano, S. Kennedy, S. Rani et al., "MAGE-D4B is a novel marker of poor prognosis and potential therapeutic target involved in breast cancer tumorigenesis," *International Journal of Cancer*, vol. 130, no. 9, pp. 1991–2002, 2012.

[29] C. E. Chong, K. P. Lim, C. P. Gan et al., "Over-expression of MAGED4B increases cell migration and growth in oral squamous cell carcinoma and is associated with poor disease outcome," *Cancer Letters*, vol. 321, no. 1, pp. 18–26, 2012.

[30] H. Takami, M. Kanda, H. Oya et al., "Evaluation of MAGE-D4 expression in hepatocellular carcinoma in Japanese patients," *Journal of Surgical Oncology*, vol. 108, no. 8, pp. 557–562, 2013.

[31] T. G. Kutateladze, "Phosphatidylinositol 3-phosphate recognition and membrane docking by the FYVE domain," *Biochimica et Biophysica Acta—Molecular and Cell Biology of Lipids*, vol. 1761, no. 8, pp. 868–877, 2006.

[32] S. X. Skapek, D. Jansen, T.-F. Wei et al., "Cloning and characterization of a novel Kruppel-associated box family transcriptional repressor that interacts with the retinoblastoma gene product, RB," *The Journal of Biological Chemistry*, vol. 275, no. 10, pp. 7212–7223, 2000.

[33] J. C. Mertens, C. D. Fingas, J. D. Christensen et al., "Therapeutic effects of deleting cancer-associated fibroblasts in cholangiocarcinoma," *Cancer Research*, vol. 73, no. 2, pp. 897–907, 2013.

[34] S. Affò, M. Dominguez, J. J. Lozano et al., "Transcriptome analysis identifies TNF superfamily receptors as potential therapeutic targets in alcoholic hepatitis," *Gut*, vol. 62, no. 3, pp. 452–460, 2013.

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 762128, 10 pages
http://dx.doi.org/10.1155/2015/762128

Review Article

Omics-Based Identification of Biomarkers for Nasopharyngeal Carcinoma

Tavan Janvilisri

Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

Correspondence should be addressed to Tavan Janvilisri; tavan.jan@mahidol.ac.th

Received 7 November 2014; Accepted 10 March 2015

Academic Editor: Lance A. Liotta

Nasopharyngeal carcinoma (NPC) is a head and neck cancer that is highly found in distinct geographic areas, such as Southeast Asia. The management of NPC remains burdensome as the prognosis is poor due to the late presentation of the disease and the complex nature of NPC pathogenesis. Therefore, it is necessary to find effective molecular markers for early detection and therapeutic measure of NPC. In this paper, the discovery of molecular biomarker for NPC through the emerging omics technologies including genomics, miRNA-omics, transcriptomics, proteomics, and metabolomics will be extensively reviewed. These markers have been shown to play roles in various cellular pathways in NPC progression. The knowledge on their function will help us understand in more detail the complexity in tumor biology, leading to the better strategies for early detection, outcome prediction, detection of disease recurrence, and therapeutic approach.

1. Introduction

Nasopharyngeal carcinoma (NPC) is a cancer of the head and neck region that arises from the squamous epithelial cells covering the surface of nasopharynx, the uppermost part of the pharynx extending from the base of the skull to the upper surface of the soft plate [1]. The incidence of NPC varies greatly on the basis of ethnic and geographical backgrounds. While NPC is a rare malignancy in most parts of the world, it is one of the most common cancers in the East and Southeast Asia including China, Hong Kong, Taiwan, Singapore, Malaysia, and Thailand [2]. The annual incidence of NPC in the United States is about 5 per 100,000. The annual incidence of the NPC in the southern part of China including Taiwan is more than 10 per 100,000 and is up to 30 per 100,000 in Hong Kong. The annual incidence of the NPC in Southeast Asia such as Malaysia and Thailand is ~20 per 100,000 and ~7 per 100,000, respectively [2]. The etiological factors for NPC include the Epstein-Barr virus (EBV) infection, ethnics, genetic susceptibility, environmental factors, and consumption of food with volatile nitrosamines [3, 4].

NPC can be diagnosed and staged by a biopsy of the tissue mass, together with positron emission tomography (PET) and computed tomography (CT). However, most of NPC patients tend to present at a more advanced stage of the disease because the primary anatomical site of tumor growth is located in the silent painless area. Moreover, NPC in advanced stages exhibits higher metastatic potential than other head and neck squamous cell carcinomas [5]. On the basis of local anatomic constraints of NPC and its tendency to present with cervical lymph node metastasis, surgery has no role for definitive therapy. At present, radiotherapy represents the standard treatment for NPC. The disease tends to be more sensitive to radiation than other cancers, but the success depends mostly on the tumors stages, which tend to be in the advanced stages at the point of diagnosis. The 5-year survival rate of stages I and II NPC ranges from 72 to 90%. However, the 5-year survival rates of stages III and IV NPC are ~55% and 30%, respectively, due to a relatively high incidence of locoregional recurrence or metastasis [6]. In case of advanced tumors, both regional-control and distant metastatic tumors, the patients are usually treated with systemic therapy. Concurrent chemotherapy is generally accepted to have a role in management of locally advanced disease. The combination chemotherapy has been used with concurrent cisplatin and radiation followed by adjuvant cisplatin and 5-fluorouracil [7, 8].

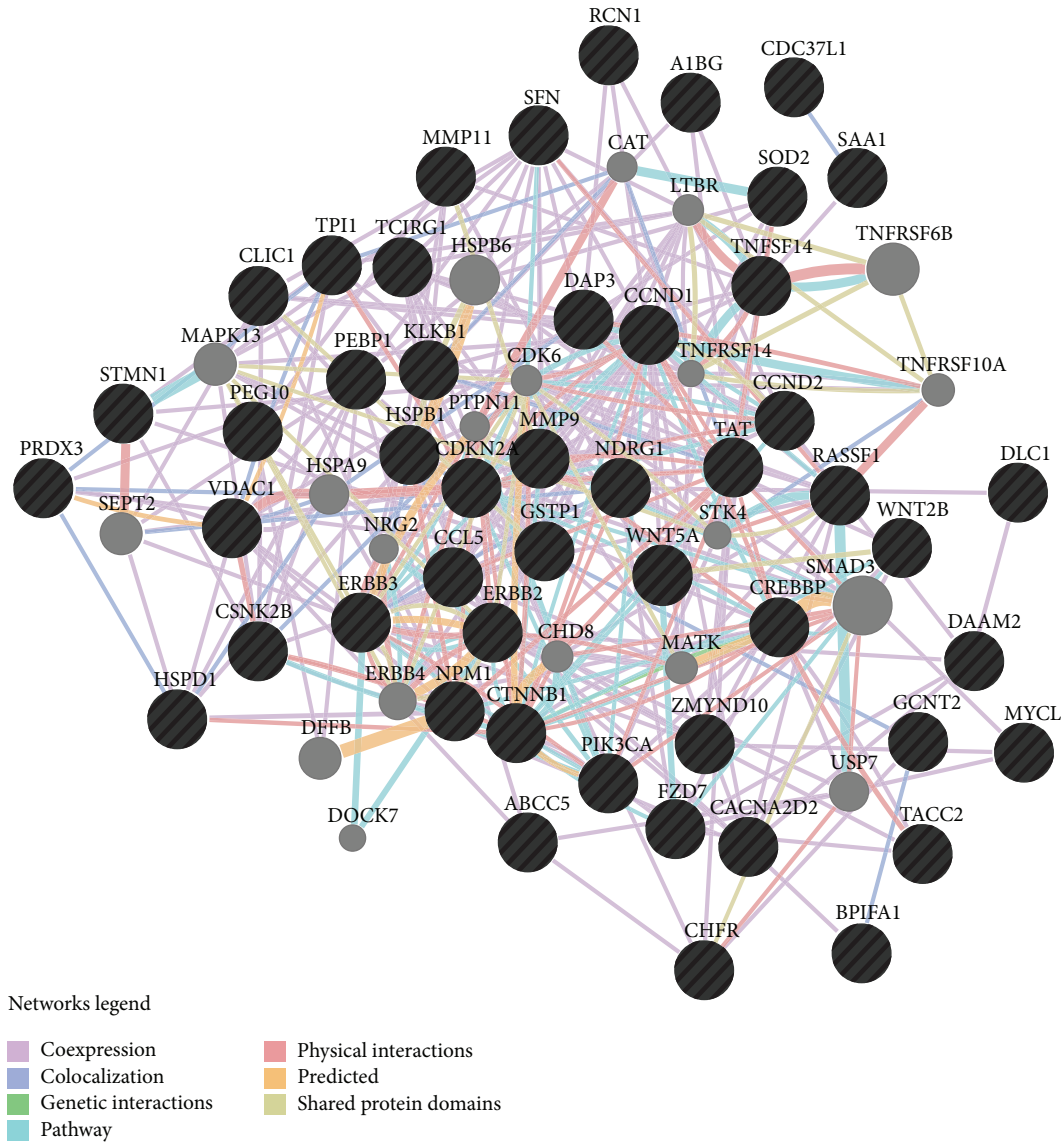


Figure 1: A network map of molecular biomarkers for nasopharyngeal carcinoma identified through omics technologies. Detailed information is described in the text.

NPC patients mostly appear in advanced stages of the disease and have a poor prognosis because of late presentation of lesions, limited knowledge of molecular pathogenesis, lack of reliable and robust biomarkers for early detection, and poor response to available therapies [9]. One of the reasons for the lack of effective molecular markers is that NPC is a highly complex multifactorial disease caused by an interaction of host genetics with the macro- and microenvironment that is influenced by EBV chronic infection and other environmental factors, in a multistep process of tumorigenesis [10]. In-depth understanding of the molecular alterations in and across the cellular pathways involved in NPC carcinogenesis can certainly facilitate the integration of diagnosis, anticancer drug discovery, and therapy for NPC. In the postgenomics era, an exponential growth of our knowledge on the disease etiology, carcinogenesis, and progression has been gained through an adoption of high-throughput technologies including genomics, transcriptomics, proteomics, metabolomics, and bioinformatics together with integration and application of systems biology. An increasing mass of these omics data has leads us to identify potential molecular targets for diagnosis, prognosis, and therapeutic treatment. The scope of this review is to shed light on the current findings of NPC biomarker discovery through the omics approaches. An overview of NPC biomarkers identified through omics approaches described herein is illustrated in Figure 1 using GeneMania (http://www.genemania.org/).

2. Genomics

Biomarkers at the genomic level can be retrieved by comparative genomic hybridization (CGH), exome sequencing, and

whole genome sequencing. These biomarkers can identify genomic alterations including single-nucleotide polymorphisms (SNPs), copy number variations (CNVs), and other structural variations in the genome and may have functional significance in the pathophysiology of a defined phenotype. As genomic instability including amplification of oncogenes and/or deletion of tumor suppressor genes, together with dysfunction of the gene by point mutations, can be an early event marker in carcinogenesis of NPC, and there are several CGH studies to analyze the gain and the loss of genetic materials in the genome. Chen and colleagues [11] performed CGH on a total of 51 NPC cases including 25 primary and 26 recurrent tumors. They reported the chromosomal hotspots for copy number gains including the chromosome arms 12p, 1q, 11q, 12q, and 17q and losses including 3p, 9p, 11q, 13q, and 14q. They also showed that there was no additional chromosomal alteration in the recurrent tumors compared to the primary cancers. A few other studies based on CGH approach have also been reported [12–15]. The patterns of genomic imbalances in NPC from these CGH data appeared to be largely consistent with those identified in banding analysis and loss of heterozygosity studies. However, the discrepancy between these studies exists, which may be due to the different cohorts of samples with different clinicopathological backgrounds, reflecting variations in distinct types of carcinogens to the oncogenic process. Furthermore, the existing CGH data of 103 NPC cases were integrated and input into evolutionary tree models, which revealed the chromosomal loss of 3p and the gain of chromosome 12 as an important hallmark for an early event for NPC carcinogenesis [16].

In the past decade, microarray technology has served as an essential tool for examination of genetic profiles of biological samples and enables us to analyze more than ten thousand genes at a time, which can reveal genetic abnormalities in cancers at a genome-wide level. The principle of microarray is based on the complementary hybridization between nucleotide chains such as DNA-to-RNA strands and DNA-to-DNA strands [17]. A microarray is basically a microscopic slide with up to hundreds of thousands of DNA fragments, which are dotted on its surface with ~50–150 μm diameter. The fragments are robotically printed or synthesized *in situ*. Each DNA fragment has a corresponding complementary DNA that binds to it. The genomic DNA can be isolated from tumor and/or normal samples, which can then be labeled with fluorophores such as cyanine-3 (Cy3; green) and cyanine-5 (Cy5; red) prior to hybridization. These labeled DNAs are added to the slide and thousands of hybridization reactions occur between input DNA samples and DNA probes on the microarray slides. After microarray laser scanning, the fluorescence values at each spot reveal the relative levels of copy number of the corresponding region [18]. Array-based CGH has therefore been used extensively to detect and quantify genomic aberrations in NPC and map onto chromosomal positions to identify relevant oncogenes or tumor suppressor genes. Hui and coworkers [19] utilized array-CGH to simultaneously investigate amplification of 58 oncogenes throughout the genome of 15 NPC samples including five cell lines, two xenografts, and eight primary tumours. The frequency of oncogenes including *MYCL1*, *TERC*, *ESR*, and *PIK3CA* were found to be amplified in the NPC samples. Other array-CGH experiments on different NPC samples have also been reported and revealed similar geographic variations in the frequencies of chromosome aberrations [15, 20, 21].

Recently, a genome-wide analysis of chromosome copy number was performed in the C666-1 cell line and from 15 NPC biopsies using high-density microarrays [22]. The data are in broad agreement with the data from conventional CGH, in which the copy loss at 3p, 9p, and 11q was observed. It has been revealed that several tumor suppressor genes such as *CDKN2A*, *ZMYND10*, *RASSF1*, *NDRG1*, *TACC2*, and *CACNA2D2* are significantly enriched within genomic regions that are frequently deleted; however, no significant correlation is established between the presence of potential tumor promoting genes and the genomic regions exhibiting gain of copy number [22]. As aberrant DNA methylation has been recognized to be associated with the transcriptional inactivation of genes related to cancer development, the application of microarrays has also been extended to study genome-wide DNA methylation patterns in NPC. Zhang and collaborators [23] investigated the methylation alterations in the genome of taxol-resistant NPC cell lines. The differential methylation profiles between the taxol-sensitive and taxol-resistant cells have been demonstrated, where the global hypermethylation was found in the latter case. The hypermethylated genes, namely, *DLC1*, *PEG10*, and the hypomethylated genes, namely, *ABCC5*, *CHFR*, *ERBB2*, and *GSTP1*, were identified and confirmed as downregulated and upregulated, respectively, in the resistant cells. Yang and colleagues [24] also applied the microchip containing ~27 k CpG loci covering more than 14,000 genes at single-nucleotide resolution to evaluate the effect of trichostatin A, one of the most potent HDAC inhibitors, on genome-wide DNA methylation pattern of a NPC cell line CNE2. Their data showed that the DNA methylation in trichostatin A-treated cells appeared to be higher in total compared to the controls. The hypermethylation of genes, namely *DAP3*, *HSPB1,* and *CLDN,* was identified in the treated group and the results were validated through quantitative reverse transcription polymerase chain reaction to confirm them as downregulated genes upon the treatment.

The discovery of variations in the DNA sequence of tumor cells associated with clinical significance has been hurled ahead by next-generation sequencing technologies. A combination of whole-exome and targeted deep sequencing, as well as SNP array analysis has been applied in order to characterize the mutational landscape of 128 NPC cases [25]. The results revealed multiple recurrent copy number variations with the most frequent deletion region covering the gene *CDKN2A* on 9p21. The loss of this chromosomal region has also been identified in the conventional CGH [11–15], providing further support for this NPC hotspot. Differential copy numbers in the genes, namely, *CCND1*, *AKT2*, *MYC*, and *TP53,* have also been observed. Interrogating pathway analyses also highlighted the dysregulation of cellular pathways involving in chromatin modification and ERBB-PI3K signaling pathway. Furthermore, the data

indicated that the alterations in ERBB-PI3K pathway were linked to the more advanced stages and the survival of NPC patients with ERBB-PI3K mutations was shorter than the patients without such mutations [25]. Our recent work also supported this notion, where we demonstrated the abnormal expression of ERBB proteins and showed that the expression of ERBB3 was associated with patient survival and could serve as a novel and valuable predictor for prognostic evaluation of patients with NPC [26].

3. Transcriptomics

Expression biomarkers are traditionally derived through the measurement of a single gene or a cluster of biochemical and histopathological molecules in a given pathway. Transcriptomics or gene expression profiling offers evaluation of the levels of gene expression of all transcripts in a given sample at the same time. The conceptual idea of transcriptomics is that the genes involved in a particular pathophysiology often function in a concerted fashion and therefore the genes with similar expression patterns may be functionally associated and/or under similar molecular regulation [27]. Initially, suppression subtractive hybridization has been applied on a cohort of libraries of PCR-amplified cDNA fragments that differ between control (normal) and experimental (cancer) transcriptome [28]. Zhang and coworkers constructed the human embryo nasopharynx cDNA library in order to isolate and screen tissue-specific genes of human nasopharynx and new tumor suppressor genes of NPC [29, 30]. Microarray-assisted analysis of subtracted cDNA libraries constructed by suppression subtractive hybridization has been performed to search for differentially expressed genes and screen candidate molecular markers in NPC [31]. The differential transcriptomes of 9 NPC cases, 3 NPC cell lines, and 10 chronic inflammation of nasopharyngeal mucosa tissue samples and the result validation using real-time quantitative reverse transcription polymerase chain reaction and *in situ* hybridization techniques revealed that the palate, lung, and nasal epithelium carcinoma (*PLUNC*) and *Homo sapiens* cell division cycle 37 Homo sapien cell division cycle 37 homolog (*Saccharomyces cerevisiae*)-like 1 (*CDC37L1*) might serve as the potential molecular biomarkers for NPC [31].

High-throughput technologies based on the well-established DNA microarray represent the most cost-effective and convenient means to assess the gene expression profiles. However, a number of biological replicates or samples of the same condition as well as additional validation through qRT-PCR are necessary to eventually identify the biomarkers for class prediction on an independent validation set as the true changes in gene expression are often underestimated [32]. As the chronic EBV infection poses as one of the causative risk factors for NPC, the application of the microarray platform to distinguish transcriptome of the EBV− and EBV+ NPC cells has enabled us to gain more understanding of EBV-specific signals for NPC tumorigenesis [33]. A set of EBV-regulated genes has been identified, involved in cellular processes such as cell proliferation, cell cycle control, and cell mobility [33]. Because NPC tissue is heterogeneous comprising cancer cells, infiltrating inflammatory cells, and nonneoplastic nasopharyngeal epithelium and stroma, tissue microdissection of NPC and normal epithelial nasopharynx has been applied to select specific types of cells on the slides prior to being subjected to the gene expression profiling analysis. Collectively, these data point to the differential NPC genes involving in the cell cycle, apoptosis, tumor suppressors, cell adhesion, and motility [34–37]. The Wnt pathways, such as wingless-type MMTV integration site family, member 5A (WNT5A), FZD7, casein kinase IIβ (CSNK2B), β-catenin (CTNNB1), CREB-binding protein (CREBBP), and dishevelled-associated activator of morphogenesis 2 (DAAM2), transforming growth factor β (TGFβ), and mitogen-activated protein kinase signaling pathway, have been found to be induced in NPC [35, 38]. Furthermore, among these genes, cyclin D1 has been shown to be the prognostic biomarker for NPC patients [37].

Microarray technology has also been used to explore the biological functions of novel genes in NPC at different metastatic features, clinical stages, and aggressive states. According to the metastatic states, the comparison of global gene expression patterns in NPC cells lines 5–8F (high tumorigenic and metastatic) and 6–10B (low tumorigenic and metastatic) revealed a cohort of genes involving in cell cycle, apoptosis, metastasis, chemokine, and immunomodulation, which potentially mediate their differential metastatic characteristics. Among them, PTHLH has been suggested to regulate the WNT pathway through the *DKK1* gene to affect metastasis and the apoptosis processes of NPC [39]. However, the validation in an independent set of samples is required to confirm this finding. Su et al. [40] identified a number of transcription factors including ATF1 and ATF2 to be associated with clinical stages. The potential downstream molecules for these transcription factors include the epithelial growth factor receptor (EGFR/ERBB1) and matrix metalloproteinase 2 (MMP-2). As the main pathological type of NPC appears to be nonkeratinizing carcinoma, gene expression profile changes have been evaluated among differentiated-type nonkeratinizing NPC cases, which revealed possible molecular subtypes [41]. It has been shown that the expression of cyclin D2 (CCND2) could serve as a molecular marker for the more aggressive tumor subtype and a strong predictor for survival time in this group of NPC patients.

The more recent RNA sequencing (RNASeq) approach utilizes deep-sequencing technologies to identify differential expression of an entire genome at any specific sample in any given time point, albeit rather expensive at present [42]. Szeto and colleagues characterized the transcriptomes of undifferentiated EBV-positive NPC xenograft X666 and its derived cell line C666, well-differentiated NPC cell line HK1, and the immortalized nasopharyngeal epithelial cell line NP460 using Solexa sequencing [43]. A total of 2,812 differentially expressed genes were identified among these samples and together with gene enrichment analysis, the extracellular matrix organization, beta-1 integrin cell surface interactions, and the PI3K/AKT, ERBB, and Wnt pathways were dysregulated in NPC [43]. In agreement with these findings, comparison of the gene expression of tumor cells and normal controls in recent studies also revealed that

the Wnt, PI3K/AKT, and ERBB signalling pathways were dysregulated [22, 25].

A large number of NPC gene expression profiles have emerged in public databases. It is challenging to integrate these data from several datasets to yield maximal information. Researchers have employed meta-analysis of transcriptomic data by integrating them from multiple studies to successfully identify new prognostic and diagnostic markers for cancer and other diseases [44]. It involves a systematic search for proper datasets and data retrieval, filtering, reprocessing, integration, and analysis [45]. However, the common problems in meta-analysis exist and are challenging. Identification of proper studies for meta-analysis is a time-consuming process as experimental information is often stored in a free-text format. The completeness and correctness of information largely depend on the thoroughness of the authors, and this issue constitutes a major challenge for microarray meta-analysis. Recently, there are a few published works on the meta-analyses of nasopharyngeal carcinoma using microarrays. Chen and collaborators combined the bioinformatics with evidence from biological experiments as a new way to gain more insights into the molecular mechanism of EBV-regulated neoplastic transformation [46]. By using a meta-analysis approach, they separated the sample into 2 metasets. The meta-A set was meta-analyzed to identify gene commonly activated or deactivated on EBV infection/reactivation in NPC (EBV reactivation in NPC versus EBV+/EBV—NPC). The meta-B set was meta-analyzed to obtain differentially expressed genes that are common in NPC and primary effusion lymphoma or PEL (EBV+/EBV—NPC versus EBV+/EBV—PEL). The meta-A and meta-B analyses revealed 23 and 45 differentially expressed genes, respectively. Then they integrated meta-A, meta-B, and related transcription factors into an interaction network using acquired information. A network of 23 meta-A genes in EBV-infected cells linked by some related transcription factors, of which the main nodes involve transcription factors JUN, CD9 and HOXA9. The 45 genes of meta-B network are connected by few related transcription factors CDKNIA, NFKBI, and MYC. The genes in meta-A and meta-B sets have been mapped into connected regulatory networks. There are 3 common genes between 2 sets including DEK, ITGA6, and DUSP1 [46]. Moreover, the regulatory network of genes involved in the EBV-dependent NPC reveals that NPC transformation depends timely on the regulation of DEK, CDK inhibitor, p53, RB, and several transcriptional cascades, which are interrelated by E2F, AP-1, NK-κB, and STAT3 among others during latent and lytic cycles [46]. The meta-analysis of EBV-related tumor data may lead to further understanding of the EBV-related neoplastic transformation.

4. MicroRNA Omics

MicroRNAs (miRNAs) are a family of small noncoding nucleotide sequences which are able to complementarily bind to and negatively regulate gene expression at the posttranscriptional level, leading to either mRNA degradation or translational repression [47]. Primary miRNAs are usually transcribed from introns or noncoding regions and are cleaved in the nucleus by Drosha enzyme to yield hairpin precursor miRNAs (pre-miRNAs). Pre-miRNAs are then translocated into the cytoplasm and are subsequently cleaved by RNase III Dicer, giving rise to miRNA. These miRNA fragments execute their regulatory role as element of the RNA-induced silencing complex (RISC) [48, 49]. Research on miRNA as cancer biomarkers has gain considerable attention as miRNAs have been shown to play a role in fundamental cellular processes including cell proliferation and cell death and negatively control the expression of several cancer promoting proteins. In contrast to other types of molecular markers, miRNAs are relatively stable in the body and tissues, rendering them better candidates for cancer biomarkers [50].

A recent study investigated the miRNA expression profiles of two poorly differentiated NPC cell lines, CNE-2 and 6–10B, and their radioresistant sublines using next-generation deep sequencing [51]. Together with qRT-PCR validation, 3 downregulated miRNAs including miR-324-3p, miR-93-3p, and miR-4501, 3 upregulated miRNAs including miR-371a-5p, miR-34c-5p, and miR-1323, and 2 novel miRNAs have been identified to play a role in NPC radioresistance. One of the downstream targets for miR-324-3p is WNT2B, which has been reported to participate in the mediated NPC radioresistance [52]. However, identification of other downstream targets of these miRNAs needs further investigations. Furthermore, miRNA expression profiles in 312 paraffin-embedded specimens of NPC and 18 specimens of noncancer nasopharyngitis have been assessed. A total of 41 miRNA were differentially expressed between NPC and noncancer counterparts. The authors proposed a signature of five miRNAs with a prognostic value in addition to the TNM staging system [53]. In another study, differentially expressed plasma miRNAs in NPC patients including miR-483-5p, miR-103, and miR-29a were identified by next-generation sequencing as potential prognostic markers for NPC [54].

5. Proteomics

Proteomics approaches have been applied to discover cancer biomarkers. In the early days, the gel-based assay, in which two-dimensional gel electrophoresis (2DE) is coupled with mass spectrometry (MS), is utilized to screen the proteins with differential abundance between samples of different conditions of interest [55]. However, the disadvantages of this method include the time and labor inefficiency as well as low recovery rate of proteins. The gel-independent assay, in which liquid chromatography (LC) is used to separate peptides/proteins instead of 2DE and is combined with MS for protein identification, has later gained popularity as it offers superior protein identification and quantitation [56]. In the cell culture model, several proteomics-based molecular markers have been identified in various experimental settings. Jiang and coworkers [57, 58] reported the differential proteomes of a poorly differentiated squamous NPC cell line, CNE2, upon the treatment with 12-O-tetradecanoylphorbol 13-acetate (TPA), a known potent carcinogen for NPC. The

results revealed upregulation of triosephosphate isomerase (TPI1) and 14-3-3 protein sigma (SFN) as well as downregulation of reticulocalbin 1 precursor (RCN1), nucleophosmin (NPM1), mitochondrial matrix protein p1 precursor (HSPD1), and stathmin (STMN1) in CNE2 cells following TPA treatment. Another study analyzed the proteomic profiles of an EBV-associated NPC cell line, C666-1, and a normal NP cell line, NP69, which showed that annexin II and beta-2-tubulin were suppressed in NPC cells [59]. Validation with immunocytochemistry also revealed that the downregulation of annexin II was positively correlated with lymph node metastasis, pointing to its potential application as a prognostic factor for NPC [59]. The proteins linked to the radioresistant trait of the NPC cells have been identified through proteomics in two independent studies using a radioresistant subclone cell line (CNE2-IR) derived from NPC cell line CNE2 [60, 61]. Feng et al. found the reduced expression of 14-3-3σ and the increased expression of Maspin, GRP78, and Mn-SOD in CNE2-IR cells compared to the control CNE2. The results were confirmed by Western blot and immunohistochemistry, suggesting that these proteins could serve as predicting biomarkers for patient response to radiotherapy and their dysregulation might be involved in the radioresistance of NPC [60]. On the other hand, Li et al. identified 16 differentially expressed proteins including upregulation of Nm23 H1 and downregulation of annexin A3 in the radioresistant NPC cells [61]. The different observations may arise from the fact that these two studies may have two different radioresistant sublines of CNE2 cells. Another study using the highly differentiated CNE1 cells and its radioresistant CNE1-IR subline demonstrated that the elevated level of heat shock protein 27 (HSP27) might play a role in radioresistance [62]. Moreover, differential proteomics of the CNE-2 and its highly metastatic subclone, S-18, and the knockdown experiment also suggested that HSP27 plays an important role in cancer metastasis and the corresponding downstream molecules could be NF-κB, MMP9, and MMP11 [63]. Therefore, HSP27 could serve as prognostic and therapeutic target.

Comparative proteomics has been performed to identify differential expression proteins between the EBV− and EBV+ NPC cells [64]. Upon the EBV infection, a total of 12 proteins were identified as being significantly upregulated and associated with (i) signal transduction including voltage-dependent anion-selective channel protein 1 (VDAC1), S100-A2, hsc-70 interacting protein (Hip-70), ubiquitin, TPT1-like protein, and 4F2 cell surface antigen; (ii) cytoskeleton formation including keratin-75, tubulin beta-8 chain B, and dynein light chain 1; (iii) metabolic pathways including l-lactate dehydrogenase B chain (LDH-B) and triosephosphate isomerase (TIM); and (iv) DNA bindings including high mobility group protein B1 (HMG-1) [64]. These proteins provide a hint on the EBV-related mechanisms of NPC carcinogenesis and pose as potential biomarkers for the interaction of NPC-EBV. As cancer cells usually secrete biomolecules to enhance their proliferation, reduce apoptosis, and invade immune system [65], a few studies on differential secretomes for NPC have attempted to identify the secreted proteins that might be useful as cancer biomarkers and therapeutic targets. The secreted proteomes of two NPC cell lines including NPC-TW02 and NPC-TW04 cell lines were analyzed and a total of 23 proteins retrieved in both cell lines. Validation with Western blotting and immunohistochemistry confirmed their results, which indicated that fibronectin, Mac-2 BP, and PAI-1 might be potential molecular markers for NPC diagnosis [66]. Other secretome studies [67–70] also identified a cohort of proteins that might be useful as NPC biomarkers including chloride intracellular channel 1 (CLIC1) and C-C motif chemokine 5 (CCL5).

Several studies combined the laser capture microdissection of NPC tissues and proteomic analysis to identify protein markers for NPC. RKIP, a member of the phosphatidylethanolamine-binding protein family, has been identified to be a NPC metastasis suppressor and its suppression has been associated with the aggressiveness through the activation of MAPK pathway [71, 72]. The expression of stathmin, 14-3-3, and annexin I in NPC tissues has been shown to be correlated with differentiation and/or metastatic potential of the NPC cells; thus the dysregulation of these proteins might play a role in NPC development [73]. Among all identified proteins, cathepsin D [74, 75], cytokeratin 18 [76], L-plastin, S100A9 [77], a stroma-associated protein periostin [78], galectin-1 [79], keratin-8, SFN, and stathmin-1 [75] have been suggested to be biomarkers for NPC differentiation, progression, and prognosis.

As human blood holds a large reservoir of proteins and provides a less invasive mean of analytes for diagnosis, differential serum proteomics have been performed to identify even slight changes of certain proteins, which could potentially be biomarkers for NPC. Serum amyloid A protein (SAA) has been identified to be useful for invigilating the recurrent NPC cases [80]. Elevated levels of blood coagulation-related proteins including plasma kallikrein (KLKB1) and thrombin-antithrombin III complex (TAT) have been observed in NPC and could provide a diagnostic value for NPC cases [81]. A glycoprotein component of fibrinogen FGA in the serum has also been associated with NPC [82]. Apart from the individual protein markers, the MS signatures of the serum proteome in normal controls and NPC patients at different stages [83–85] and NPC with different levels of radiosensitivity [86] have been shown to be distinct.

It has been shown that the ERBB signaling pathway is dysregulated in NPC [25, 26]. This pathway is known to be tightly regulated by phosphorylation and dephosphorylation. Ruan and collaborators [87] have attempted to identify the downstream proteins, which are affected by stimulation of epithelial growth factor (EGF), by evaluating the phosphoproteome of CNE2 cells. A total of 33 proteins were identified in CNE2 upon the treatment with EGF. Among the identified proteins, glutathione S-transferase P1 has been validated using Western immunoblotting and knockdown experiments and has been linked to drug resistant trait in NPC cells [87]. Mitochondrial proteomes of the NPC cell lines 5–8F and 6–10B have been compared in order to find a clue for molecular mechanism of NPC metastasis and biomarkers related to metastasis [88]. A total of 16 mitochondrial proteins including PRDX3 and SOD2 were identified and serve as

potential biomarkers for NPC. As these proteins are involved in the cellular response to reactive oxygen species, their abnormal function would play a role in oxidative stress, which could in turn mediate NPC metastasis [88].

6. Metabolomics

Metabolomics is considered to be a relatively new field of omics that simultaneously monitors many hundreds and thousands of small molecule metabolites from biofluids and tissue samples [89]. In any given conditions, a concerted function of metabolic processes occurs within a cell, which is readily changing in different physiological conditions. Hence, metabolomics represents a biochemical footprint of a physiological state of a cell. Metabolic profiles can be measured using nuclear magnetic resonance (NMR) spectroscopy and MS-based assays coupled with gas chromatography (GC-MS) or liquid chromatography (LC-MS) [90]. Differential metabolomes between case and control samples will lead to a cohort of molecules that has potential for early diagnosis, therapy, and understanding of the pathogenesis of many diseases. The metabolomics for NPC is still in its infancy. Recently, the metabolites of sera samples from 40 normal controls and 39 NPC patients were analyzed to find novel metabolic biomarkers [91]. Three novel candidate biomarkers including glucose, glutamate, and pyroglutamate were identified with the high specificity, suggesting that glycolysis and glutamate metabolism are involved in NPC carcinogenesis. Further validation of these molecules is warranted with larger cohorts of patients to prove their usefulness in terms of diagnosis. Yi et al. [92] performed a GC-MS-based metabolic profiling of 402 serum samples from NPC patients and normal controls. Metabolites including glucose, linoleic acid, stearic acid, arachidonic acid, proline, b-hydroxybutyrate, and glycerol 1-hexadecanoate were shown to have high distinguishing power of NPC from the healthy controls. Moreover, the metabolic signatures of the NPC patients who received radiotherapy appeared to resemble those of the normal controls, pointing to the possibility of applying metabolomics in assessing therapeutic effects.

7. Concluding Remarks

The omics technology enables the high-throughput profiling in the levels of genomics, epigenomics, transcriptomics, proteomics, and metabolomics, which lead to the large amount of data and together with bioinformatic tools can retrieve novel biomarkers. Current omics research in NPC has been reviewed, focusing on the biomarker discovery. A large number of potential biomarkers for NPC related to various pathophysiological states have been identified. However, extensive validation of these molecules in a larger cohort and in a multicenter platform is essential to verify their usefulness as biomarkers. In the future, it will be challenging to integrate the vast amount of multiomics data to gain better understanding of molecular basis of this complex malignancy.

Conflict of Interests

The author declares that he has no conflict of interests.

Acknowledgments

This work is supported by the Faculty of Science, Mahidol University, and a research grant for mid-career university faculty from Thailand Research Fund and Office of the Higher Education Commission (RMU5380009).

References

[1] W. I. Wei and J. S. T. Sham, "Nasopharyngeal carcinoma," *The Lancet*, vol. 365, no. 9476, pp. 2041–2054, 2005.

[2] E. T. Chang and H.-O. Adami, "The enigmatic epidemiology of nasopharyngeal carcinoma," *Cancer Epidemiology Biomarkers and Prevention*, vol. 15, no. 10, pp. 1765–1777, 2006.

[3] S. H. Chan, "Aetiology of nasopharyngeal carcinoma," *Annals of the Academy of Medicine Singapore*, vol. 19, no. 2, pp. 201–207, 1990.

[4] Q. Tao and A. T. C. Chan, "Nasopharyngeal carcinoma: molecular pathogenesis and therapeutic developments," *Expert Reviews in Molecular Medicine*, vol. 9, no. 12, pp. 1–24, 2007.

[5] A. S. Chan, K. F. To, K. W. Lo et al., "High frequency of chromosome 3p deletion in histologically normal nasopharyngeal epithelia from Southern Chinese," *Cancer Research*, vol. 60, no. 19, pp. 5365–5370, 2000.

[6] J. T.-C. Chang, J.-Y. Ko, and R.-L. Hong, "Recent advances in the treatment of nasopharyngeal carcinoma," *Journal of the Formosan Medical Association*, vol. 103, no. 7, pp. 496–510, 2004.

[7] A. T. C. Chan, S. F. Leung, R. K. C. Ngan et al., "Overall survival after concurrent cisplatin-radiotherapy compared with radiotherapy alone in locoregionally advanced nasopharyngeal carcinoma," *Journal of the National Cancer Institute*, vol. 97, no. 7, pp. 536–539, 2005.

[8] A. T. C. Chan, P. M. L. Teo, R. K. Ngan et al., "Concurrent chemotherapy-radiotherapy compared with radiotherapy alone in locoregionally advanced nasopharyngeal carcinoma: progression-free survival analysis of a phase III randomized trial," *Journal of Clinical Oncology*, vol. 20, no. 8, pp. 2038–2044, 2002.

[9] A. R. A. Razak, L. L. Siu, F.-F. Liu, E. Ito, B. O'Sullivan, and K. Chan, "Nasopharyngeal carcinoma: the next challenges," *European Journal of Cancer*, vol. 46, no. 11, pp. 1967–1978, 2010.

[10] W. C.-S. Cho, "Nasopharyngeal carcinoma: molecular biomarker discovery and progress," *Molecular Cancer*, vol. 6, article 1, 2007.

[11] Y.-J. Chen, J.-Y. Ko, P.-J. Chen et al., "Chromosomal aberrations in nasopharyngeal carcinoma analyzed by comparative genomic hybridization," *Genes, Chromosomes & Cancer*, vol. 25, no. 2, pp. 169–175, 1999.

[12] Y. Fang, X.-Y. Guan, Y. Guo et al., "Analysis of genetic alterations in primary nasopharyngeal carcinoma by comparative genomic hybridization," *Genes, Chromosomes & Cancer*, vol. 30, no. 3, pp. 254–260, 2001.

[13] G. Chien, P. W. Yuen, D. Kwong, and Y. L. Kwong, "Comparative genomic hybridization analysis of nasopharygeal carcinoma: consistent patterns of genetic aberrations and clinicopathological correlations," *Cancer Genetics and Cytogenetics*, vol. 126, no. 1, pp. 63–67, 2001.

[14] J. Yan, Y. Fang, Q. Liang, Y. Huang, and Y. Zeng, "Novel chromosomal alterations detected in primary nasopharyngeal carcinoma by comparative genomic hybridization," *Chinese Medical Journal*, vol. 114, no. 4, pp. 418–421, 2001.

[15] S. Rodriguez, A. Khabir, C. Keryer et al., "Conventional and array-based comparative genomic hybridization analysis of nasopharyngeal carcinomas from the Mediterranean area," *Cancer Genetics and Cytogenetics*, vol. 157, no. 2, pp. 140–147, 2005.

[16] L. Shih-Hsin Wu, "Construction of evolutionary tree models for nasopharyngeal carcinoma using comparative genomic hybridization data," *Cancer Genetics and Cytogenetics*, vol. 168, no. 2, pp. 105–108, 2006.

[17] B. A. Bejjani, A. P. Theisen, B. C. Ballif, and L. G. Shaffer, "Array-based comparative genomic hybridization in clinical diagnosis," *Expert Review of Molecular Diagnostics*, vol. 5, no. 3, pp. 421–429, 2005.

[18] S. Wang and Q. Cheng, "Microarray analysis in drug discovery and clinical applications," *Methods in Molecular Biology*, vol. 316, pp. 49–65, 2006.

[19] A. B.-Y. Hui, K.-W. Lo, P. M. L. Teo, K.-F. To, and D. P. Huang, "Genome wide detection of oncogene amplifications in nasopharyngeal carcinoma by array based comparative genomic hybridization," *International Journal of Oncology*, vol. 20, no. 3, pp. 467–473, 2002.

[20] X. Guo, W.-O. Lui, C.-N. Qian et al., "Identifying cancer-related genes in nasopharyngeal carcinoma cell lines using DNA and mRNA expression profiling analyses," *International Journal of Oncology*, vol. 21, no. 6, pp. 1197–1204, 2002.

[21] Y. Y. Or, A. B. Y. Hui, K. Y. Tam, D. P. Huang, and K. W. Lo, "Characterization of chromosome 3q and 12q amplicons in nasopharyngeal carcinoma cell lines," *International journal of oncology*, vol. 26, no. 1, pp. 49–56, 2005.

[22] C. Hu, W. Wei, X. Chen et al., "A global view of the oncogenic landscape in nasopharyngeal carcinoma: an integrated analysis at the genetic and expression levels," *PLoS ONE*, vol. 7, no. 7, Article ID e41055, 2012.

[23] X. Zhang, W. Li, H. Li, Y. Ma, G. He, and G. Tan, "Genomic methylation profiling combined with gene expression microarray reveals the aberrant methylation mechanism involved in nasopharyngeal carcinoma taxol resistance," *Anti-Cancer Drugs*, vol. 23, no. 8, pp. 856–864, 2012.

[24] X. L. Yang, C. D. Zhang, H. Y. Wu et al., "Effect of trichostatin A on CNE2 nasopharyngeal carcinoma cells—genome-wide DNA methylation alteration," *Asian Pacific Journal of Cancer Prevention*, vol. 15, no. 11, pp. 4663–4670, 2014.

[25] D. C. Lin, X. Meng, M. Hazawa et al., "The genomic landscape of nasopharyngeal carcinoma," *Nature Genetics*, vol. 46, no. 8, pp. 866–871, 2014.

[26] W. Tulalamba, N. Larbcharoensub, and T. Janvilisri, "ERBB3 as an independent prognostic marker for nasopharyngeal carcinoma," *Journal of Clinical Pathology*, vol. 67, no. 8, pp. 667–672, 2014.

[27] G. Bucca, G. Carruba, A. Saetta, P. Muti, L. Castagnetta, and C. P. Smith, "Gene expression profiling of human cancers," *Annals of the New York Academy of Sciences*, vol. 1028, pp. 28–37, 2004.

[28] O. D. von Stein, "Isolation of differentially expressed genes through subtractive suppression hybridization," *Methods in Molecular Biology*, vol. 175, pp. 263–278, 2001.

[29] B.-C. Zhang, L. Cao, J. Qian et al., "Construction of directional cDNA library from human embryo nasopharyngeal epithelia and screeming a candidate tumor suppressor gene related with nasopharyngeal carcinoma," *Progress in Biochemistry and Biophysics*, vol. 29, no. 2, pp. 302–306, 2002.

[30] B. C. Zhang, X. M. Nie, B. Y. Xiao et al., "Identification of tissue-specific genes in nasopharyngeal epithelial tissue and differentially expressed genes in nasopharyngeal carcinoma by suppression subtractive hybridization and cDNA microarray," *Genes Chromosomes and Cancer*, vol. 38, no. 1, pp. 80–90, 2003.

[31] Y. Zhou, Z. Zeng, W. Zhang et al., "Identification of candidate molecular markers of nasopharyngeal carcinoma by microarray analysis of subtracted cDNA libraries constructed by suppression subtractive hybridization," *European Journal of Cancer Prevention*, vol. 17, no. 6, pp. 561–571, 2008.

[32] M. Schena, R. A. Heller, T. P. Theriault, K. Konrad, E. Lachenmeier, and R. W. Davis, "Microarrays: biotechnology's discovery platform for functional genomics," *Trends in Biotechnology*, vol. 16, no. 7, pp. 301–306, 1998.

[33] Y.-C. G. Lee, Y.-C. Hwang, K.-C. Chen et al., "Effect of Epstein-Barr virus infection on global gene expression in nasopharyngeal carcinoma," *Functional & Integrative Genomics*, vol. 7, no. 1, pp. 79–93, 2007.

[34] V. Sriuranpong, A. Mutirangura, J. W. Gillespie et al., "Global gene expression profile of nasopharyngeal carcinoma by laser capture microdissection and complementary DNA microarrays," *Clinical Cancer Research*, vol. 10, no. 15, pp. 4944–4958, 2004.

[35] Z. Y. Zeng, Y. H. Zhou, W. Xiong et al., "Analysis of gene expression identifies candidate molecular markers in nasopharyngeal carcinoma using microdissection and cDNA microarray," *Journal of Cancer Research and Clinical Oncology*, vol. 133, no. 2, pp. 71–81, 2007.

[36] Z.-Q. Liu, Y.-Q. Tian, Y.-F. Hu, X.-L. Li, F.-R. Ma, and G.-Y. Li, "Alteration of gene expression during nasopharyngeal carcinogenesis revealed by oligonucleotide microarray after microdissection of tumor tissue and normal epithelia from nasopharynx," *Chinese Medical Journal (Engl)*, vol. 122, no. 4, pp. 437–443, 2009.

[37] W. Zhang, Z. Zeng, Y. Zhou et al., "Identification of aberrant cell cycle regulation in Epstein-Barr virus-associated nasopharyngeal carcinoma by cDNA microarray and gene set enrichment analysis," *Acta Biochimica et Biophysica Sinica*, vol. 41, no. 5, Article ID gmp025, pp. 414–428, 2009.

[38] Z. Y. Zeng, Y. H. Zhou, W. L. Zhang et al., "Gene expression profiling of nasopharyngeal carcinoma reveals the abnormally regulated Wnt signaling pathway," *Human Pathology*, vol. 38, no. 1, pp. 120–133, 2007.

[39] J. Li, Y. Fan, J. Chen, K.-T. Yao, and Z.-X. Huang, "Microarray analysis of differentially expressed genes between nasopharyngeal carcinoma cell lines 5–8F and 6–10B," *Cancer Genetics and Cytogenetics*, vol. 196, no. 1, pp. 23.e1–23.e13, 2010.

[40] B. Su, H.-L. Tang, M. Deng et al., "Stage-associated dynamic activity profile of transcription factors in nasopharyngeal carcinoma progression based on protein/DNA array analysis," *OMICS*, vol. 15, no. 1-2, pp. 49–60, 2011.

[41] S. Wang, X. Li, Z. G. Li et al., "Gene expression profile changes and possible molecular subtypes in differentiated-type nonkeratinizing nasopharyngeal carcinoma," *International Journal of Cancer*, vol. 128, no. 4, pp. 753–762, 2011.

[42] S. Marguerat and J. Bähler, "RNA-seq: from technology to biology," *Cellular and Molecular Life Sciences*, vol. 67, no. 4, pp. 569–579, 2010.

[43] C. Y.-Y. Szeto, C. H. Lin, S. C. Choi et al., "Integrated mRNA and microRNA transcriptome sequencing characterizes sequence variants and mRNA-microRNA regulatory network in nasopharyngeal carcinoma model systems," *FEBS Open Bio*, vol. 4, pp. 128–140, 2014.

[44] D. R. Rhodes and A. M. Chinnaiyan, "Integrative analysis of the cancer transcriptome," *Nature Genetics*, vol. 37, no. 6, pp. S31–S37, 2005.

[45] W.-C. Cheng, M.-L. Tsai, C.-W. Chang et al., "Microarray meta-analysis database (M^2DB): a uniformly pre-processed, quality controlled, and manually curated human clinical microarray database," *BMC Bioinformatics*, vol. 11, article 421, 2010.

[46] X. Chen, S. Liang, W. L. Zheng, Z. J. Liao, T. Shang, and W. L. Ma, "Meta-analysis of nasopharyngeal carcinoma microarray data explores mechanism of EBV-regulated neoplastic transformation," *BMC Genomics*, vol. 9, article 322, 2008.

[47] M. T. McManus, "MicroRNAs and cancer," *Seminars in Cancer Biology*, vol. 13, no. 4, pp. 253–258, 2003.

[48] D. P. Bartel, "MicroRNAs: genomics, biogenesis, mechanism, and function," *Cell*, vol. 116, no. 2, pp. 281–297, 2004.

[49] D. P. Bartel, "MicroRNAs: target recognition and regulatory functions," *Cell*, vol. 136, no. 2, pp. 215–233, 2009.

[50] G. A. Calin and C. M. Croce, "MicroRNA signatures in human cancers," *Nature Reviews Cancer*, vol. 6, no. 11, pp. 857–866, 2006.

[51] G. Li, Y. Qiu, Z. Su et al., "Genome-wide analyses of radioresistance-associated miRNA expression profile in nasopharyngeal carcinoma using next generation deep sequencing," *PLoS ONE*, vol. 8, no. 12, Article ID e84486, 2013.

[52] G. Li, Y. Liu, Z. Su et al., "MicroRNA-324-3p regulates nasopharyngeal carcinoma radioresistance by directly targeting WNT2B," *European Journal of Cancer*, vol. 49, no. 11, pp. 2596–2607, 2013.

[53] N. Liu, N.-Y. Chen, R.-X. Cui et al., "Prognostic value of a microRNA signature in nasopharyngeal carcinoma: a microRNA expression analysis," *The Lancet Oncology*, vol. 13, no. 6, pp. 633–641, 2012.

[54] H.-Y. Wang, L.-X. Yan, Q. Shao et al., "Profiling plasma microRNA in nasopharyngeal carcinoma with deep sequencing," *Clinical Chemistry*, vol. 60, no. 5, pp. 773–782, 2014.

[55] Y. Iwadate, "Clinical proteomics in cancer research-promises and limitations of current two-dimensional Gel electrophoresis," *Current Medicinal Chemistry*, vol. 15, no. 23, pp. 2393–2400, 2008.

[56] W. H. McDonald and J. R. Yates III, "Shotgun proteomics and biomarker discovery," *Disease Markers*, vol. 18, no. 2, pp. 99–105, 2002.

[57] P. Jiang, M. Gan, H. Huang, X. Shen, S. Wang, and K. Yao, "Proteome analysis of antiproliferative mechanism of 12-O-tetradecanoylphorbol 13-acetate on cultured nasopharyngeal carcinoma CNE2 cells," *Journal of Proteome Research*, vol. 4, no. 2, pp. 599–605, 2005.

[58] P.-Z. Jiang, M. Gan, H. Huang, X.-M. Shen, S. Wang, and K.-T. Yao, "Proteomics-based identification of proteins with altered expression induced by 12-*O*-tetradecanoylphorbol 13-acetate in nasopharyngeal carcinoma CNE2 cells," *Acta Biochimica et Biophysica Sinica*, vol. 37, no. 2, pp. 97–106, 2005.

[59] C. M. L. Chan, S. C. C. Wong, M. Y. Y. Lam et al., "Proteomic comparison of nasopharyngeal cancer cell lines C666-1 and NP69 identifies down-regulation of annexin II and β2-tubulin for nasopharyngeal carcinoma," *Archives of Pathology and Laboratory Medicine*, vol. 132, no. 4, pp. 675–683, 2008.

[60] X.-P. Feng, H. Yi, M.-Y. Li et al., "Identification of biomarkers for predicting nasopharyngeal carcinoma response to radiotherapy by proteomics," *Cancer Research*, vol. 70, no. 9, pp. 3450–3462, 2010.

[61] L. Li, S. Huang, X. Zhu et al., "Identification of radioresistance-associated proteins in human nasopharyngeal carcinoma cell lines by proteomic analysis," *Cancer Biotherapy & Radiopharmaceuticals*, vol. 28, no. 5, pp. 380–384, 2013.

[62] B. Zhang, J.-Q. Qu, L. Xiao et al., "Identification of heat shock protein 27 as a radioresistance-related protein in nasopharyngeal carcinoma cells," *Journal of Cancer Research and Clinical Oncology*, vol. 138, no. 12, pp. 2117–2125, 2012.

[63] G.-P. Li, H. Wang, Y.-K. Lai et al., "Proteomic profiling between CNE-2 and its strongly metastatic subclone S-18 and functional characterization of HSP27 in metastasis of nasopharyngeal carcinoma," *Proteomics*, vol. 11, no. 14, pp. 2911–2920, 2011.

[64] X. Feng, J. Zhang, W. N. Chen, and C. B. Ching, "Proteome profiling of Epstein-Barr virus infected nasopharyngeal carcinoma cell line: identification of potential biomarkers by comparative iTRAQ-coupled 2D LC/MS-MS analysis," *Journal of Proteomics*, vol. 74, no. 4, pp. 567–576, 2011.

[65] H. Xue, B. Lu, and M. Lai, "The cancer secretome: a reservoir of biomarkers," *Journal of Translational Medicine*, vol. 6, article 52, 2008.

[66] C.-C. Wu, K.-Y. Chien, N.-M. Tsang et al., "Cancer cell-secreted proteomes as a basis for searching potential tumor markers: nasopharyngeal carcinoma as a model," *Proteomics*, vol. 5, no. 12, pp. 3173–3182, 2005.

[67] S. Ge, Y. Mao, Y. Yi, D. Xie, Z. Chen, and Z. Xiao, "Comparative proteomic analysis of secreted proteins from nasopharyngeal carcinoma-associated stromal fibroblasts and normal fibroblasts," *Experimental and Therapeutic Medicine*, vol. 3, no. 5, pp. 857–860, 2012.

[68] S.-J. Lin, K.-P. Chang, C.-W. Hsu et al., "Low-molecular-mass secretome profiling identifies C-C motif chemokine 5 as a potential plasma biomarker and therapeutic target for nasopharyngeal carcinoma," *Journal of Proteomics*, vol. 94, pp. 186–201, 2013.

[69] Y.-H. Chang, C.-C. Wu, K.-P. Chang, J.-S. Yu, Y.-C. Chang, and P.-C. Liao, "Cell secretome analysis using hollow fiber culture system leads to the discovery of CLIC1 protein as a novel plasma marker for nasopharyngeal carcinoma," *Journal of Proteome Research*, vol. 8, no. 12, pp. 5465–5474, 2009.

[70] H.-Y. Wu, Y.-H. Chang, Y.-C. Chang, and P.-C. Liao, "Proteomics analysis of nasopharyngeal carcinoma cell secretome using a hollow fiber culture system and mass spectrometry," *Journal of Proteome Research*, vol. 8, no. 1, pp. 380–389, 2009.

[71] Y. Chen, G.-L. Ouyang, H. Yi et al., "Identification of RKIP as an invasion suppressor protein in nasopharyngeal carcinoma by proteomic analysis," *Journal of Proteome Research*, vol. 7, no. 12, pp. 5254–5262, 2008.

[72] Y. Chen, C.-E. Tang, G.-L. Ouyang et al., "Identification of RKIP as a differentially tyrosine-phosphorylated protein in nasopharyngeal carcinoma and normal nasopharyngeal epithelial tissues by phosphoproteomic approach," *Medical Oncology*, vol. 26, no. 4, pp. 463–470, 2009.

[73] A.-L. Cheng, W.-G. Huang, Z.-C. Chen et al., "Identification of novel nasopharyngeal carcinoma biomarkers by laser capture microdissection and proteomic analysis," *Clinical Cancer Research*, vol. 14, no. 2, pp. 435–445, 2008.

[74] A.-L. Cheng, W.-G. Huang, Z.-C. Chen et al., "Identificating cathepsin d as a biomarker for differentiation and prognosis

of nasopharyngeal carcinoma by laser capture microdissection and proteomic analysis," *Journal of Proteome Research*, vol. 7, no. 6, pp. 2415–2426, 2008.

[75] Z. Xiao, G. Li, Y. Chen et al., "Quantitative proteomic analysis of formalin-fixed and paraffin-embedded nasopharyngeal carcinoma using iTRAQ labeling, two-dimensional liquid chromatography, and tandem mass spectrometry," *Journal of Histochemistry and Cytochemistry*, vol. 58, no. 6, pp. 517–527, 2010.

[76] X.-M. Li, W.-G. Huang, H. Yi, A.-L. Cheng, and Z.-Q. Xiao, "Proteomic analysis to identify cytokeratin 18 as a novel biomarker of nasopharyngeal carcinoma," *Journal of Cancer Research and Clinical Oncology*, vol. 135, no. 12, pp. 1763–1775, 2009.

[77] M.-X. Li, Z.-Q. Xiao, Y.-F. Liu et al., "Quantitative proteomic analysis of differential proteins in the stroma of nasopharyngeal carcinoma and normal nasopharyngeal epithelial tissue," *Journal of Cellular Biochemistry*, vol. 106, no. 4, pp. 570–579, 2009.

[78] M. Li, C. Li, D. Li et al., "Periostin, a stroma-associated protein, correlates with tumor invasiveness and progression in nasopharyngeal carcinoma," *Clinical & Experimental Metastasis*, vol. 29, no. 8, pp. 865–877, 2012.

[79] C.-E. Tang, T. Tan, C. Li et al., "Identification of Galectin-1 as a novel biomarker in nasopharyngeal carcinoma by proteomic analysis," *Oncology Reports*, vol. 24, no. 2, pp. 495–500, 2010.

[80] W. C. S. Cho, T. T. C. Yip, C. Yip et al., "Identification of serum amyloid a protein as a potentially useful biomarker to monitor relapse of nasopharyngeal cancer by serum proteomic profiling," *Clinical Cancer Research*, vol. 10, no. 1, part 1, pp. 43–52, 2004.

[81] P.-H. Peng, C.-C. Wu, S.-C. Liu et al., "Quantitative plasma proteome analysis reveals aberrant level of blood coagulation-related proteins in nasopharyngeal carcinoma," *Journal of Proteomics*, vol. 74, no. 5, pp. 744–757, 2011.

[82] Y.-L. Tao, Y. Li, J. Gao et al., "Identifying FGA peptides as nasopharyngeal carcinoma-associated biomarkers by magnetic beads," *Journal of Cellular Biochemistry*, vol. 113, no. 7, pp. 2268–2278, 2012.

[83] Y.-S. Wei, Y.-H. Zheng, W.-B. Liang et al., "Identification of serum biomarkers for nasopharyngeal carcinoma by proteomic analysis," *Cancer*, vol. 112, no. 3, pp. 544–551, 2008.

[84] Y.-J. Huang, C. Xuan, B.-B. Zhang et al., "SELDI-TOF MS profiling of serum for detection of nasopharyngeal carcinoma," *Journal of Experimental and Clinical Cancer Research*, vol. 28, article 85, 2009.

[85] S.-M. Cao, J.-K. Yu, Q.-Y. Chen et al., "Detection of nasopharyngeal carcinoma using surface-enhanced laser desorption and ionization mass spectrometry profiles of the serum proteome," *Chinese Journal of Cancer*, vol. 29, no. 8, pp. 721–728, 2010.

[86] F. Su, X. Zhu, Z. Liang et al., "Analysis of serum proteome profiles in nasopharyngeal carcinoma with different radiosensitivity," *Clinical and Translational Oncology*, vol. 16, no. 2, pp. 147–152, 2014.

[87] L. Ruan, X.-H. Li, X.-X. Wan et al., "Analysis of EGFR signaling pathway in nasopharyngeal carcinoma cells by quantitative phosphoproteomics," *Proteome Science*, vol. 9, article 35, 2011.

[88] J. Liu, X. Zhan, M. Li et al., "Mitochondrial proteomics of nasopharyngeal carcinoma metastasis," *BMC Medical Genomics*, vol. 5, article 62, 2012.

[89] C. J. Clarke and J. N. Haselden, "Metabolic profiling as a tool for understanding mechanisms of toxicity," *Toxicologic Pathology*, vol. 36, no. 1, pp. 140–147, 2008.

[90] G. A. N. Gowda, S. Zhang, H. Gu, V. Asiago, N. Shanaiah, and D. Raftery, "Metabolomics-based methods for early disease diagnostics," *Expert Review of Molecular Diagnostics*, vol. 8, no. 5, pp. 617–633, 2008.

[91] L. Yi, N. Dong, S. Shi et al., "Metabolomic identification of novel biomarkers of nasopharyngeal carcinoma," *RSC Advances*, vol. 4, no. 103, pp. 59094–59101, 2014.

[92] L. Yi, C. Song, Z. Hu et al., "A metabolic discrimination model for nasopharyngeal carcinoma and its potential role in the therapeutic evaluation of radiotherapy," *Metabolomics*, vol. 10, no. 4, pp. 697–708, 2014.

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 629367, 6 pages
http://dx.doi.org/10.1155/2015/629367

Research Article

Comparative Proteomic Analysis of Human Cholangiocarcinoma Cell Lines: S100A2 as a Potential Candidate Protein Inducer of Invasion

Kasima Wasuworawong,[1] Sittiruk Roytrakul,[2] Atchara Paemanee,[2] Kattaleeya Jindapornprasert,[1] and Waraporn Komyod[1]

[1]*Department of Biochemistry, Faculty of Science, Mahidol University, 272 Rama VI Road, Bangkok 10400, Thailand*
[2]*Proteomics Research Laboratory, Genome Institute, National Science and Technology Development Agency, 111 Thailand Science Park, Phahonyothin Road, Pathum Thani 12120, Thailand*

Correspondence should be addressed to Waraporn Komyod; waraporn.kom@mahidol.ac.th

Received 7 November 2014; Accepted 9 March 2015

Academic Editor: Joy Scaria

Cholangiocarcinoma (CCA) is a bile duct cancer, commonly found in Asia including Thailand and especially in the northeastern region of Thailand. To identify the proteins involved in carcinogenesis and metastasis of CCA, protein expression profiles of high-invasive KKU-M213 and low-invasive KKU-100 cell lines were compared using a comparative GeLC-MS/MS proteomics analysis. A total of 651 differentially expressed proteins were detected of which 27 protein candidates were identified as having functions involved in cell motility. A total of 22 proteins were significantly upregulated in KKU-M213, whereas 5 proteins were downregulated in KKU-M213. S100A2, a calcium-binding protein in S100 protein family, is upregulated in KKU-M213. S100A2 is implicated in metastasis development in several cancers. The protein expression level of S100A2 was verified by Western blot analysis. Intriguingly, high-invasive KKU-M213 cells showed higher expression of S100A2 than KKU-100 cells, consistent with proteomic data, suggesting that S100A2 may be a key protein involved in the progression of CCA. However, the biological function of S100A2 in cholangiocarcinoma remains to be elucidated. S100A2 might be a potential biomarker as well as a novel therapeutic target in CCA metastasis.

1. Introduction

Cholangiocarcinoma (CCA) is a malignant tumor that originates from epithelial cells of the bile duct. CCA is difficult to diagnose and the curative treatment remains challenging. Due to its late clinical manifestation, morbidity and mortality rates of CCA are high and its incidence has been increasing over the past three decades, especially in northeastern Thailand [1]. CCA is often associated with metastasis which is a highly complicated process that involves cell motility, invasion, angiogenesis, intravasation of tumor cells into the blood stream, and finally extravasation and colonization of tumor cells at secondary sites [2]. The migration and invasion properties have been a hallmark of cancer [3] including CCA, in incrimination of disease severity. In particular, metastasis is one of the major hindrances to the treatment of CCA and many cancer types that cause more than 90% of cancer-associated mortality [4]. Moreover, CCA is resistant to radio- and chemotherapy, and surgical resection is the only effective therapy against this type of cancer [5–7]. Hence, understanding the mechanism of invasion and metastasis will be important in identifying key players involved, which may lead to development of effective targeted therapy against this deadly disease.

Here, we compared the protein profiles of two human CCA cell lines with different metastatic abilities, KKU-M213 and KKU-100. KKU-M213, a high-invasive cell line, originated from adenosquamous CCA with well differentiation and KKU-100, a low-invasive cell line, was isolated from adenocarcinoma CCA with poor differentiation [8]. Studying the differential protein patterns of these cell lines allowed us to identify several proteins which might be the key

determinant of the metastatic properties of the CCA cells and might be beneficial as a future drug target.

Proteomics analysis is currently considered to be a powerful tool for global evaluation of protein expression, and proteomics has been widely applied in analysis of diseases, especially in fields of cancer research. In this study, we employed a comparative SDS-PAGE coupled with LC-MS/MS (GeLC-MS/MS) based proteomics approach [9] to compare the protein expression profile of the high-invasive KKU-M213 cell line with low-invasive KKU-100 cell line to better understand the development and metastasis of CCA. MS/MS spectra of obtained proteins were identified based on NCBI human database. This technique can identify potential candidate proteins that might be involved in the different degrees of invasiveness displayed by the two CCA cell lines. Differential expression at transcription and protein expression levels of a candidate protein was further confirmed by quantitative real-time PCR and Western blot analysis.

2. Materials and Methods

2.1. Cell Culture. Human cholangiocarcinoma cell lines, KKU-M213 and KKU-100, were kindly provided by Professor Banchob Sripa (Khon Kaen University, Khon Kaen, Thailand). Cells were cultured in Ham's F-12 nutrient mixture medium (Invitrogen Corp., Auckland, NZ) supplemented with 10% fetal bovine serum (FBS), 100 U/mL penicillin, 100 μg/mL streptomycin sulfate (Invitrogen Corp., Auckland, NZ), and 15 mM HEPES (USB Corp., OH, USA). Cells were incubated at 37°C in a humidified atmosphere with 5% CO_2.

2.2. Invasion Assay. Invasion assay was determined by Matrigel transwell *in vitro* invasion assay as previously described [10] with some modification. In brief, the upper chamber of a transwell unit (6.5 mm diameter polycarbonate membrane with 8 μm pore size) (Corning Incorporated Life Science, Corning, NY) was coated with 30 μg of Matrigel (BD Biosciences, Bedford, MA). Cells (80% confluent) were harvested using 0.25% Trypsin-EDTA (Invitrogen Corp., Auckland, NZ) and resuspended in serum-free media. A 200 μL aliquot of cell suspension (10^5 cells) was added to the upper chamber. The lower chamber was filled with 600 μL of media containing 1% FBS. After 12 hours of incubation at 37°C under CO_2 atmosphere, noninvading cells in the upper chamber were removed and cells that invaded the Matrigel and had attached to the lower surface of the transwell membrane were fixed with 25% methanol for 15 min and stained with 0.5% crystal violet. Invaded cells were counted in 5 random fields under light microscope at 10x magnification. The reported values represent mean ± SE of the results obtained from three independent experiments.

2.3. Preparation of Cell Lysates. Cells were washed twice with cold PBS containing 100 μM Na_3VO_4, trypsinized and collected by centrifugation. Cell pellets were kept in −80°C prior to use. The pellets were lysed in Tris-lysis buffer (20 mM Tris-HCl, pH 7.5; 150 mM NaCl; 10 mM NaF; 1% (v/v) Triton-X) supplemented with 1 mM Na_3VO_4, 1 mM PMSF, and protease inhibitors (Sigma, UK) and chilled on ice for 30 min. Lysates were centrifuged at 13000 rpm for 10 min at 4°C to discard cell debris and protein concentration in the supernatant was determined by Lowry assay [11]. The cell lysates were stored at −20°C.

2.4. One-Dimension Gel Electrophoresis and Tryptic In-Gel Digestion. Cell lysates (30 μg) were mixed in loading buffer (312.5 mM Tris-Cl, pH 6.8, 50% glycerol, 10% SDS, 0.05% bromophenol blue, 12.5% 2-B-mercaptoethanol) and boiled for 5 min before being applied on a 12.5% SDS-polyacrylamide gel (BioRad, Hercules, CA) using Hoefer apparatus. After Coomassie blue staining, gel slices were excised, cut into 1 mm^3 cubes, and subjected to in-gel tryptic digestion. The excised gel slices were reduced with 10 mM DTT/10 mM NH_4HCO_3, alkylated with 100 mM IAA/10 mM NH_4HCO_3, and digested with 1 ng protein per 20 ng sequencing grade trypsin (Promega, Germany) at 37°C overnight.

2.5. Protein Identification Using LC-MS/MS. Tryptic peptides were protonated with 0.1% formic acid before injection into NanoAcquity system (Waters Corp., Milford, MA) equipped with a Symmetry C_{18} 5 μm, 180-μm × 20-mm Trap column and a BEH130 C_{18} 1.7 μm, 100-μm × 100-mm analytical reversed phase column (Waters Corp., Milford, MA). [Glu^1] fibrinopeptide B was used as the reference sprayer of the NanoLockSpray source of the mass spectrometer. Analysis of tryptic peptides was performed using a SYNAPT HDMS mass spectrometer (Waters Corp., Manchester, UK). The time-of-flight analyzer of the mass spectrometer was externally calibrated with [Glu^1] fibrinopeptide B. The quadrupole mass analyzer was adjusted such that ions from m/z 300 to 1,800 were efficiently transmitted. BSA, used for normalization, was performed along with the samples.

MS intensities of individual LC-MS analysis were differentially quantified by using DeCyder MS Differential Analysis Software (GE Healthcare, USA). PepMatch module was used for evaluating the average abundance ratio of each sample peptide, allowing for automated detection of peptides and assignment of charge states. The MS/MS data was searched against the NCBInr database and identified by using Mascot software (Matrix Science, London, UK). Database interrogation was implemented as follows: taxonomy—homo sapiens; database—NCBInr; enzyme—trypsin; fixed modification—carbamidomethyl; variable modification—oxidation of methionine residues; mass values—monoisotopic; peptide mass tolerance—2 Da; peptide charge state—1+, 2+, and 3+. Protein accession numbers were classified according to their biological function by PANTHER Classification system version 8.1 (http://www.pantherdb.org/geneListAnalysis.do).

2.6. Western Blot Analysis. A total of 30 μg of protein lysates was separated by 12% SDS-PAGE and then transferred to PVDF membrane (Pall, Germany) by semidry electroblotting at constant voltage (25 V) for 60 min. The membranes were blocked with 5% BSA in 1x TBS-N for 1 hr and then incubated with anti-S100A2 primary antibody (Abcam, UK) at 4°C overnight. The blots were washed three times for 5 min

with TBS-N buffer and incubated with anti-rabbit HRP-conjugated secondary antibody (Santa Cruz Biotechnology Inc., Santa Cruz, CA) at room temperature for 30 min. Signals were detected using Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, USA).

2.7. Quantitative Real-Time PCR. Total RNA was extracted using illustra RNAspin Mini kit (GE Healthcare, USA) as described by the manufacturer. 1 μg of total RNA was converted to cDNA using ImProm-II Reverse Transcription System kit (Promega, Germany) using random hexamer primers according to the manufacturer's description. The PCR reaction was performed in a final volume of 20 μL containing 100 ng of cDNA, 300 nM of each primer, and 10 μL FastStart Universal SYBR Master (Roche, Germany). Specific primers were as follows: S100A2 forward 5′-CTGGGTCTG-TCTCTGCCACC-3′, S100A2 reverse 5′-GCAGGAGTA-CTTGTGGAAGGTAGTG-3′ and β-actin forward 5′-CTC-TTCCAGCCTTCCTTCCT-3′, β-actin reverse 5′-AGCACT-GTGTTGGCGTACAG-3′. Thermal cycling conditions were as follows: denaturing at 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 60°C for 30 s. Real-time PCR was performed on Mx3000P QPCR System (Agilent Technologies, USA). All PCR amplifications were conducted in triplicate. The $2^{-\Delta\Delta CT}$ method [12] was used to calculate the relative gene expression level.

2.8. Statistical Analysis. Statistical analysis was performed using Student's t-test with $P < 0.05$ considered to be significant.

3. Results and Discussion

Cell invasion using the Boyden-transwell migration assay revealed that the KKU-M213 cells displayed approximately 8-fold higher level of invasiveness than KKU-100 cells (Figure 1). The results were consistent with previous report [13] that KKU-M213 is a high-invasive cell line while KKU-100 is a low-invasive cell line. A comparative SDS-PAGE of protein lysates of both cell types initially indicated differences in intensity of protein bands as shown in Figure 2. Upon in-gel tryptic digestions coupled with LC-MS/MS (GeLC-MS/MS) analysis of the protein lysates of high-invasive KKU-M213 cells and low-invasive KKU-100 cells, six hundred and fifty-one differentially expressed proteins were identified. These proteins were classified into 8 groups of functional proteins according to their biological processes synergized by UniProtKB, using PANTHER classification system. These proteins were categorized as cellular component (32%), transcription and translation process (15%), metabolic process (12%), signal transduction (11%), immune response (10%), cell motility (4%), unknown (13%), and others (3%) (Figure 3). Among these, 27 proteins were identified belonging to the cell motility group. The relative expression of these proteins was determined by reciprocal common fraction between peptide intensities of the two cell lines. With this approach, a total of 22 proteins were found to be significantly

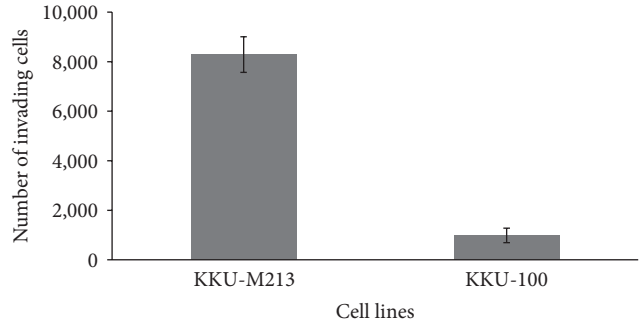


Figure 1: *In vitro* invasion assays of KKU-M213 and KKU-100 cells were conducted in a transwell unit coated with Matrigel. Cells in serum-free medium were plated in the upper chamber of a transwell unit. After 12 hours of incubation, cells invading to the lower compartment of the transwell unit were stained and counted. The numbers of invading cells are presented as mean ± SE of results obtained from three independent experiments, $P < 0.01$, compared with KKU-100 cell line.

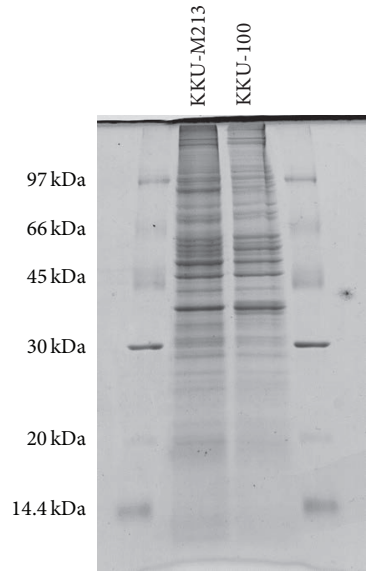


Figure 2: Differential expression of proteins comparing between KKU-M213 and KKU-100 cell lines. 30 μg of protein lysates was separated in 12.5% SDS-PAGE with constant amplitude (20 mA/gel). The gel was stained with Coomassie blue R-250 to visualize protein bands.

upregulated in KKU-M213 (Table 1), whereas 5 proteins were downregulated in KKU-M213 (Table 2).

As shown in Table 1, the expression of S100A2 is notably higher in KKU-M213 than in KKU-100. The S100A2 protein is a calcium-binding protein in S100 protein family and has been implicated in the initiation and progression of human cancers such as epithelial ovarian cancer, pancreatic cancer, and gastric cancer [14, 15] and is associated with cancer metastasis process [15–17]. Further verification of the protein expression level of S100A2 by Western blot analysis using an antibody specific to S100A2 confirmed that the expression of S100A2 in KKU-M213 is obviously higher than KKU-100 cells

Table 1: Overexpression of proteins in KKU-M213 compared to KKU-100.

Protein	ID details	Sequence	Score	Fold[a]
gi|5174661	Protein S100A2	ELPSFVGEK	22.51	1.80
gi|18088719	Tubulin, beta	ISVYYNEATGGK	60.59	1.68
gi|12667788	Myosin-9	IAQLEEQLDNETK	33.10	1.58
gi|4885583	Rho-associated protein kinase 1	SVAMCEMEK	2.96	1.47
gi|66346662	Rho GTPase-activating protein 8 isoform 1	KDGDLTMWPR	20.28	1.42
gi|46249758	Ezrin	IALLEEAR	49.55	1.33
gi|5174735	Tubulin beta-4B chain	INVYYNEATGGK	64.96	1.29
gi|116063573	Filamin-A isoform 1	SPFEVYVDK	36.32	1.27
gi|4502101	Annexin A1	TPAQFDADELR	93.72	1.24
gi|33469929	Pikachurin isoform 1 precursor	QKIVEGMAEGGFTQIK	3.99	1.23
gi|47059046	Protocadherin-23 isoform 1	AVPPRMPAVNLGQVPPK	9.20	1.23
gi|4501891	Alpha-actinin-1 isoform b	AGTQIENIEEDFRDGLK	20.13	1.21
gi|336020355	Mitogen-activated protein kinase 4 isoform 2	LTANETQSASSTLQK	10.91	1.21
gi|50845388	Annexin A2 isoform 1	GVDEVTIVNILTNR	79.32	1.20
gi|28372535	Tctex1 domain-containing protein 3	VQQILTESLK	30.94	1.20
gi|40788018	Rho GTPase-activating protein 11A isoform 2	MSSNTEKK	8.47	1.18
gi|105990514	Filamin-B isoform 2	VLFASQEIPASPFR	40.73	1.15
gi|224451128	Protein eyes shut homolog isoform 1	ISDISFHYEFHLK	13.75	1.15
gi|122937398	Cytoplasmic dynein 2 heavy chain 1 isoform 2	AADLKDLNSR	15.59	1.13
gi|4503355	Dedicator of cytokinesis protein 1	KVTAKIDYGNR	3.12	1.10
gi|7662284	Protein-methionine sulfoxide oxidase MICAL2	AAHLASMFGHGDFPQNK	11.55	1.08
gi|7657532	Protein S100A6	LMEDLDR	36.48	1.08

[a]Statistical significance was determined by Student's t-test ($P < 0.05$).

Table 2: Overexpression of proteins in KKU-100 compared to KKU-M213.

Protein	ID details	Sequence	Score	Fold[a]
gi|4504981	Galectin-1	SFVLNLGK	30.26	2.33
gi|4506091	Mitogen-activated protein kinase 6	RLDHDNIVK	9.03	1.39
gi|62548860	Matrilin-2 isoform a precursor	NFNSAKDMK	12.44	1.21
gi|53828924	Neuropeptides B/W receptor type 1	TYSAAR	9.92	1.22
gi|6005810	Mitogen-activated protein kinase 1 isoform 2	LGGGTYGEVFK	8.48	1.12

[a]Statistical significance was determined by Student's t-test ($P < 0.05$).

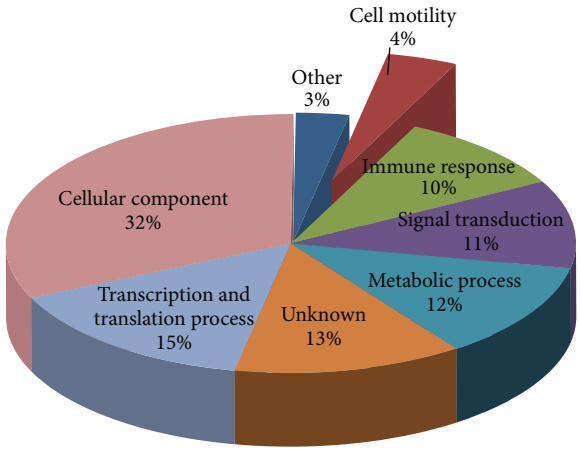


Figure 3: Gene ontology pie-charts showed categorization of 651 identified proteins from MS/MS spectra according to their biological processes using the PANTHER classification system.

(Figure 4(a)), in compliance with the expression profile of proteomic analysis. Moreover, S100A2 had higher expression in KKU-M213 than MMNK-1, a normal cholangiocyte cell line, since S100A2 protein expression was not detected in MMNK-1 cells (data not shown). The expression of S1000A2 at transcription level as determined by quantitative real-time PCR also elucidated about 850-fold higher expression in KKU-M213 than KKU-100 (Figure 4(b)). Altogether the expression levels of S100A2 in correlation with the invasiveness of KKU-M213 cells implied that S100A2 might be a key protein involved in the progression of CCA. S100A2 has been reported to be a potential biomarker for diagnosis and prognosis in many types of cancer [14], with overexpression and downregulation in various types of cancer [15, 18–22]. The role of S100A2 in promoting NSCLC metastasis [18] and migration/invasion in hepatocellular carcinoma [23] has also been documented. The S100A2 expression has been shown to be necessary for TGF-β-mediated migration/invasion [23]

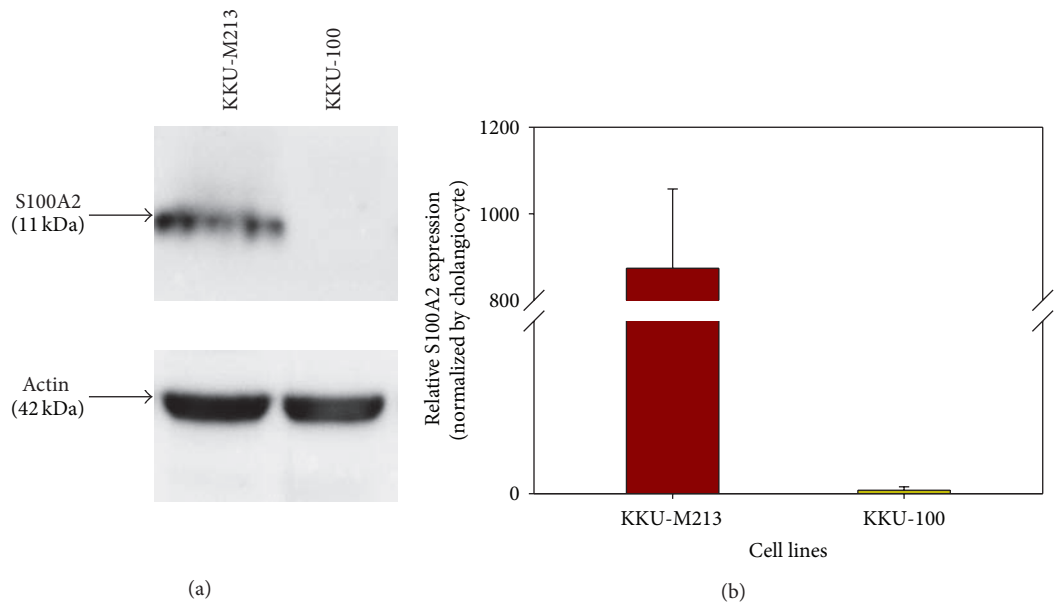


Figure 4: Validation of S100A2 in cholangiocarcinoma KKU-M213 and KKU-100 cell lines. (a) Western blot analysis. (b) Quantitative real-time PCR. Data are presented as mean ± SE of S100A2 mRNA level normalized with β-actin mRNA obtained from three independent experiments, $P = 0.003$, compared with KKU-100 cell line.

and was regulated by TGF-β-induced MEK/ERK signaling. Ectopic expression of S100A2 also elucidated that S100A2 regulated PI3K/Akt signaling, a potent pathway involved in epithelial-mesenchymal transition (EMT) [24]. Furthermore, S100A2 can interact with p53 to modulate its transcriptional activity in the calcium-dependent manner [25] while Smad3 does not need calcium ion to interact with S100A2 [24]. Taken together our findings of the S100A2 differential overexpression in KKU-M213 signified its role in invasive ability in CCA. Recently, immunohistochemistry of resected CCA tissues illustrated that S100A2 expression level was correlated with severity of CCA cancer progression and suggested it as a potential biomarker for the diagnosis of cholangiocarcinoma patients [26]. However, the biological function of S100A2 toward invasiveness and progression of cholangiocarcinoma still needs further investigations. Importantly our results suggested that S100A2 could be a candidate biological marker and novel target for diagnosis of CCA metastasis.

4. Conclusions

Our study aims to compare the protein profiles of the two CCA cell lines, KKU-M213 and KKU-100, with an attempt to identify proteins associated with invasiveness of CCA. SDS-PAGE coupled with LC-MS/MS (GeLC-MS/MS) is a potential initial technique to obtain an entire protein expression profile, followed by further verification steps. With this method, we showed a profile of proteome alterations in the two CCA cells with different invasive ability. We have identified 651 proteins that were found to be differentially expressed between the two cell lines and could be categorized into at least 6 functional groups including cellular component (32%), transcription and translation process (15%), metabolic process (12%), signal transduction (11%), immune response (10%), and cell motility (4%). In cell motility group, S100A2, a calcium-binding protein, which had pronouncedly 1.8-fold higher expression in high-invasive KKU-M213 cells, has been identified. Higher expression of S100A2 was confirmed at both transcription and protein expression levels. Our results suggested that S100A2 could be a significant candidate marker of CCA carcinogenesis and possibly a novel therapeutic target in CCA metastasis. Further investigation of the biological function of S100A2 in CCA could be pursued by overexpressing of S100A2 in S100A2-depleted cell line and by an approach using knock-down protein expression in S100A2-expressing cell line. Finally our observations by proteomic approach provided useful insights for understanding the mechanism involved in CCA carcinogenesis and could have implications in improved CCA diagnosis and prognosis capability.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by Faculty of Science, Mahidol University, the Office of the Higher Education Commission, and Thailand Research Fund (MRG5380073). The authors wish to thank Associate Professor Pornpimol Rongnoparat for critical reading of the paper and for her helpful comments.

References

[1] B. Sripa, S. Kaewkes, P. Sithithaworn et al., "Liver fluke induces cholangiocarcinoma," *PLoS medicine*, vol. 4, no. 7, article e201, 2007.

[2] A. F. Chambers, A. C. Groom, and I. C. MacDonald, "Dissemination and growth of cancer cells in metastatic sites," *Nature Reviews Cancer*, vol. 2, no. 8, pp. 563–572, 2002.

[3] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell*, vol. 144, no. 5, pp. 646–674, 2011.

[4] W. G. Stetler-Stevenson, "Matrix metalloproteinases in angiogenesis: a moving target for therapeutic intervention," *Journal of Clinical Investigation*, vol. 103, no. 9, pp. 1237–1241, 1999.

[5] C. D. Anderson, C. W. Pinson, J. Berlin, and R. S. Chari, "Diagnosis and treatment of cholangiocarcinoma," *Oncologist*, vol. 9, no. 1, pp. 43–57, 2004.

[6] G. J. Gores, "Cholangiocarcinoma: current concepts and insights," *Hepatology*, vol. 37, no. 5, pp. 961–969, 2003.

[7] A. E. Sirica, "Cholangiocarcinoma: molecular targeting strategies for chemoprevention and therapy," *Hepatology*, vol. 41, no. 1, pp. 5–15, 2005.

[8] B. Sripa, S. Leungwattanawanit, T. Nitta et al., "Establishment and characterization of an opisthorchiasis-associated cholangiocarcinoma cell line (KKU-100)," *World Journal of Gastroenterology*, vol. 11, no. 22, pp. 3392–3397, 2005.

[9] S. R. Piersma, M. O. Warmoes, M. de Wit, I. de Reus, J. C. Knol, and C. R. Jiménez, "Whole gel processing procedure for GeLC-MS/MS based proteomics," *Proteome Science*, vol. 11, no. 1, article 17, 2013.

[10] A. Albini, Y. Iwamoto, and H. K. Kleinman, "A rapid in vitro assay for quantitating the invasive potential of tumor cells," *Cancer Research*, vol. 47, no. 12, pp. 3239–3245, 1987.

[11] O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, "Protein measurement with the Folin phenol reagent," *The Journal of Biological Chemistry*, vol. 193, no. 1, pp. 265–275, 1951.

[12] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001.

[13] K. Ratthaphol, K. Imtawil, S. Wongkham, and C. Wongkham, "Development of homemade cholangiocarcinoma DNA array," *Srinagarind Medical Journal*, pp. 203–205, 2007.

[14] I. Salama, P. S. Malone, F. Mihaimeed, and J. L. Jones, "A review of the S100 proteins in cancer," *European Journal of Surgical Oncology*, vol. 34, no. 4, pp. 357–364, 2008.

[15] S. Wolf, C. Haase-Kohn, and J. Pietzsch, "S100A2 in cancerogenesis: a friend or a foe?" *Amino Acids*, vol. 41, no. 4, pp. 849–861, 2011.

[16] S. K. Mishra, H. R. Siddique, and M. Saleem, "S100A4 calcium-binding protein is key player in tumor progression and metastasis: preclinical and clinical evidence," *Cancer and Metastasis Reviews*, vol. 31, no. 1-2, pp. 163–172, 2012.

[17] E. Lukanidin and J. P. Sleeman, "Building the niche: the role of the S100 proteins in metastatic growth," *Seminars in Cancer Biology*, vol. 22, no. 3, pp. 216–225, 2012.

[18] E. Bulk, B. Sargin, U. Krug et al., "S100A2 induces metastasis in non-small cell lung cancer," *Clinical Cancer Research*, vol. 15, no. 1, pp. 22–29, 2009.

[19] R. Yao, A. Lopez-Beltran, G. T. Maclennan, R. Montironi, J. N. Eble, and L. Cheng, "Expression of S100 protein family members in the pathogenesis of bladder tumors," *Anticancer Research*, vol. 27, no. 5, pp. 3051–3058, 2007.

[20] H.-L. Hsieh, B. W. Schäfer, N. Sasaki, and C. W. Heizmann, "Expression analysis of S100 proteins and RAGE in human tumors using tissue microarrays," *Biochemical and Biophysical Research Communications*, vol. 307, no. 2, pp. 375–381, 2003.

[21] C. D. Hough, K. R. Cho, A. B. Zonderman, D. R. Schwartz, and P. J. Morin, "Coordinately up-regulated genes in ovarian cancer," *Cancer Research*, vol. 61, no. 10, pp. 3869–3876, 2001.

[22] D. B. Villaret, T. Wang, D. Dillon et al., "Identification of genes overexpressed in head and neck squamous cell carcinoma using a combination of complementary DNA subtraction and microarray analysis," *Laryngoscope*, vol. 110, no. 3, part 1, pp. 374–381, 2000.

[23] S. Naz, P. Ranganathan, P. Bodapati, A. H. Shastry, L. N. Mishra, and P. Kondaiah, "Regulation of S100A2 expression by TGF-β-induced MEK/ERK signalling and its role in cell migration/invasion," *Biochemical Journal*, vol. 447, no. 1, pp. 81–91, 2012.

[24] S. Naz, M. Bashir, P. Ranganathan, P. Bodapati, V. Santosh, and P. Kondaiah, "Protumorigenic actions of S100A2 involve regulation of PI3/Akt signaling and functional interaction with Smad3," *Carcinogenesis*, vol. 35, no. 1, pp. 14–23, 2014.

[25] A. Muelleri, B. W. Schäfer, S. Ferrari et al., "The Calcium-binding protein S100A2 interacts with p53 and modulates its transcriptional activity," *The Journal of Biological Chemistry*, vol. 280, no. 32, pp. 29186–29193, 2005.

[26] Y. Sato, K. Harada, M. Sasaki, and Y. Nakanuma, "Clinicopathological significance of S100 protein expression in cholangiocarcinoma," *Journal of Gastroenterology and Hepatology*, vol. 28, no. 8, pp. 1422–1429, 2013.

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 351015, 5 pages
http://dx.doi.org/10.1155/2015/351015

Research Article

Analysis of Serum MicroRNAs as Potential Biomarker in Coronary Bifurcation Lesion

Yan Liu, Shaoliang Chen, Junjie Zhang, Shoujie Shan, Liang Chen, Rong Wang, Jing Kan, and Tian Xu

Department of Cardiology, Nanjing First Hospital, Nanjing Medical University, Changle Road No. 68, Nanjing, Jiangsu 210029, China

Correspondence should be addressed to Shaoliang Chen; chenshaliang@gmail.com

Received 31 October 2014; Revised 11 January 2015; Accepted 17 February 2015

Academic Editor: Varodom Charoensawan

Recent evidence suggests that cell-derived circulating miRNAs may serve as the biomarkers of cardiovascular diseases. However, no study has investigated the potential of circulating miRNAs as biomarker for coronary bifurcation lesion. In this study, we aimed to characterize the miRNA profiles that could distinguish coronary bifurcation lesion and identify potential miRNAs as biomarkers of coronary bifurcation lesion. We employed miRNA microarray to screen serum miRNAs profiles of patients with coronary bifurcation lesion and coronary nonbifurcation lesions. We identified 197 miRNAs differentially expressed, including 150 miRNAs upregulated and 47 miRNAs downregulated. We chose 3 miRNAs with significant differences for further testing in 200 patients. RT-PCR analysis of serum samples confirmed that miR30d was upregulated and miR1246 was downregulated in the serum of coronary bifurcation lesion patients compared with nonbifurcation lesion patients. Our findings suggest that these miRNAs may have a role in the pathogenesis of coronary bifurcation lesion and may represent novel biomarkers for the diagnosis and prognosis of coronary bifurcation lesion.

1. Introduction

Coronary artery disease (CAD) is one of the leading causes of death worldwide. The pathogenesis of CAD remains incompletely understood and both genetic and environmental factors are involved in this complex disease [1, 2]. Coronary bifurcation lesion is the most challenging lesion in percutaneous coronary interventional medicine because the rate of restenosis and major adverse cardiac event is significantly higher than nonbifurcation lesion [3]. MicroRNAs (miRNAs) are key regulators of gene expression that have been widely associated with a variety of diseases [4, 5]. Recent evidence suggests that cell-derived circulating miRNAs may serve as the biomarkers of cardiovascular diseases including CAD [6]. However, no study has investigated the potential of circulating miRNAs as biomarker for coronary bifurcation lesion.

In this study, we aimed to characterize the miRNA profiles that could distinguish coronary bifurcation lesion and identify potential miRNAs as biomarkers of coronary bifurcation lesion.

2. Methods

2.1. Patients. In the matched case-control study, we recruited 5 patients with coronary bifurcation lesion aged 45–87 years (average 72.4 ± 16.4 years) and 5 control patients with coronary nonbifurcation lesion aged 46–75 years (average 59.6 ± 11.7 years). For large sample validation, we recruited 100 patients with coronary bifurcation lesion aged 45–88 years (average 67.8 ± 10.9 years) and 100 control patients with coronary nonbifurcation lesion aged 40–84 years (average 65.5 ± 10.2 years). Serum samples of all subjects were collected.

Coronary bifurcation lesion was defined as a junction of a main vessel and a side branch (with a minimal diameter of 1.5 mm) [7]. Coronary bifurcation lesion patients and control subjects had no other concomitant diseases, including severe cardiomyopathy or valvular heart disease, lung disease, significant cardiac dysfunction or liver and kidney dysfunction, thrombotic disease or blood disease, and connective tissue disease or malignancy or active infection such as hepatitis and tuberculosis.

2.2. Serum Samples. The blood was collected from each subject into EDTA anticoagulant tube and centrifuged at 3,000 rpm for 10 min, and the supernatant was stored at −80°C.

2.3. RNA Extraction. Total RNA was isolated using TRIzol (Invitrogen) and miRNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. RNA was dissolved in nuclease-free water by passing a few times through a pipette tip. RNA quality and quantity were measured by using nanodrop spectrophotometer (ND-1000, Nanodrop Technologies) and RNA integrity was determined by gel electrophoresis.

2.4. RNA Labeling. RNA was labeled using miRCURY Hy3/Hy5 Power Labeling Kit (Exiqon, Vedbaek, Denmark) according to the manufacturer's guideline. One microgram of each sample was 3′-end-labeled with Hy3TM fluorescent label, using T4 RNA ligase. The mixture was incubated for 30 min at 37°C, and the reaction was terminated by incubation for 5 min at 95°C. Then, 3.0 μL of labeling buffer, 1.5 μL of fluorescent label (Hy3TM), 2.0 μL of DMSO, and 2.0 μL of labeling enzyme were added into the mixture. The labeling mixture was incubated for 1 h at 16°C, and the reaction was terminated by incubation for 15 min at 65°C.

2.5. miRNA Microarray. Hy3TM-labeled samples were hybridized on the miRCURYTM LNA Array (v.18.0) (Exiqon) according to the manufacturer's manual. Following the hybridization, the slides were achieved, washed several times using Wash Buffer Kit (Exiqon), and finally dried by centrifugation for 5 min at 400 rpm. Then, the slides were scanned using the Agilent Microarray Scanner (part number G2505C).

2.6. Quantitative RT-PCR. Quantitative PCR was performed on an ABI 7500 system (Applied Biosystems, Foster City, CA). RNA was reverse transcribed with the TaqMan miRNA Reverse Transcription Kit (ABI) according to the manufacturer's instructions. Subsequently, 2.5 ul of the product was used for detecting miRNA expression by quantitative polymerase chain reaction with TaqMan miRNA Assay Kits (ABI).

2.7. Data Analysis. The intensity of green signal was calculated after background subtraction and four replicated spots of each probe on the same slide have been averaged. We used Median Normalization Method to obtain "Normalized Data," Normalized Data = (Foreground − Background)/median; the median was 50 percent quantile of microRNA intensity which was larger than 30 in all samples after background correction. After normalization, the statistical significance of differentially expressed miRNA was analyzed by t-test. A threshold cut-off value was set to 1 fold change which indicates that the expression of a given miRNA is uniform in both case and control groups. The fold change of greater than 1 indicated upregulated miRNAs, whereas the fold change of less than 1 indicated downregulated miRNAs.

Table 1: The characteristics of the patients for microarray analysis.

	Coronary bifurcation lesion $n = 5$	Coronary nonbifurcation lesion $n = 5$
Age (years)	72.4 ± 16.4	59.6 ± 11.7
Males	4 (80%)	2 (40%)
Hypertension	4 (80%)	4 (80%)
Diabetes	0 (0%)	0 (0%)
Stroke	2 (40%)	2 (40%)
Current smoker	2 (40%)	2 (40%)
Lesion site LAD	4 (80%)	1 (20%)
Lesion site LM	0 (0%)	4 (80%)
Lesion site RCA	1 (20%)	0 (0%)
Lesion calcification	0 (0%)	3 (60%)
Needing rotablation	0 (0%)	1 (20%)
Restenosis	0 (0%)	1 (20%)
Thrombus	1 (20%)	1 (20%)
TIMI grade 0	3 (60%)	2 (40%)
TIMI grade 1	2 (40%)	3 (60%)
Medina classification 111	0 (0%)	4 (80%)

Unsupervised hierarchical clustering and correlation analysis was performed on miRNA data. Data were expressed as mean ± SD and analyzed using SPSS13.0 software. Categorical variables were compared by χ^2 test, and continuous variables were compared by t-test. A value of $P < 0.05$ was considered significant.

3. Results

3.1. The Characteristics of the Patients. To characterize the miRNAs profiles of coronary bifurcation lesion, we recruited 5 patients with coronary bifurcation lesion and 5 patients with coronary nonbifurcation lesion as the control. The characteristics of these patients were listed in Table 1.

3.2. Identification of miRNAs Expression Patterns of Patients with Coronary Bifurcation Lesion. We performed microarray analysis to identify miRNAs expression patterns of patients with coronary bifurcation lesion. We made a heat map to visualize the results of the two-way hierarchical clustering of miRNAs (Figure 1). The color scale shown at the top illustrated the relative expression level of a miRNA: red represented a high relative expression level while green represented a low relative expression level [8].

Furthermore, we made a Volcano plot to illustrate miRNAs differentially expressed between coronary bifurcation lesion and coronary nonbifurcation lesion (Figure 2). We identified 197 miRNAs differentially expressed between coronary bifurcation lesion and coronary nonbifurcation lesion, including 150 miRNAs upregulated and 47 miRNAs downregulated (Supplemental File 1 in Supplementary Material available online at http://dx.doi.org/10.1155/2015/351015).

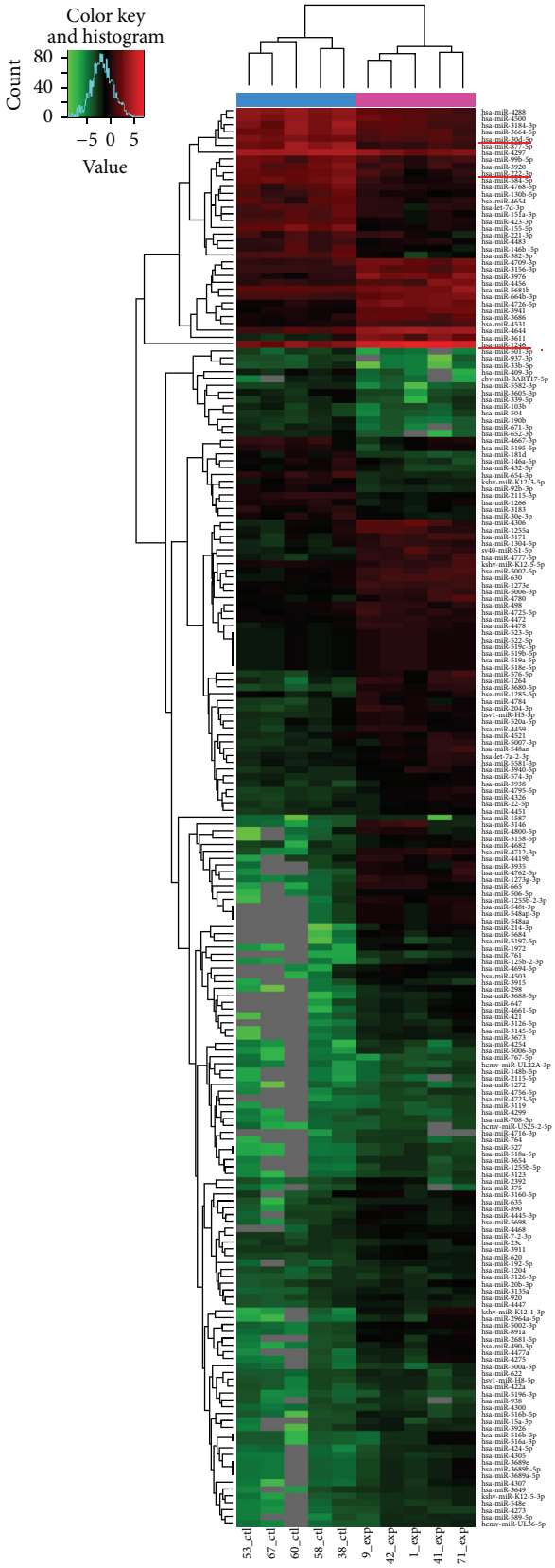


Figure 1: Heat map illustrating the expression patterns of upregulated and downregulated miRNAs in patients with bifurcation lesion. Upregulated miRNAs were indicated by red while downregulated miRNAs were indicated by green. The three candidate miRNA markers miR30d, miR222, and miR1246 were underlined in red.

Table 2: Confirmation of differential miRNA expression in 200 patients.

	Bifurcation lesion ($n = 100$)	Nonbifurcation lesion ($n = 100$)	P
Age (years)	67.8 ± 10.9	65.5 ± 10.2	0.377
Males	52	56	0.55
Smoker	52	50	0.82
Hypertension	61	60	0.76
Diabetes	28	30	0.51
miR30d level	0.0258 ± 0.0566	0.0017 ± 0.0006	0.000
miR222 level	0.1024 ± 0.0616	0.0953 ± 0.0693	0.881
miR1246 level	0.0346 ± 0.0567	0.3004 ± 0.2469	0.000

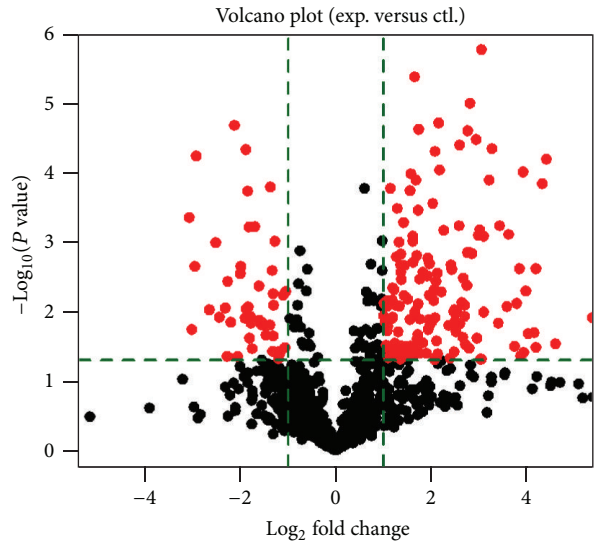


Figure 2: miRNAs differentially expressed in patients with bifurcation lesion and patients with nonbifurcation lesion. The volcano plots illustrated miRNAs differentially expressed: dots in black indicted the miRNAs that did not reach significant changes of expression; dots in red on the left indicated the miRNAs that had significant downregulation of expression; and dots in red on the right indicated the miRNAs that had significant upregulation of expression.

3.3. Confirmation of Different miRNAs in Serum of Patients with Coronary Bifurcation Lesion. To confirm our results, we chose three of the most differentially expressed miRNAs, including 2 upregulated (miR30d and miR222) and one downregulated (miR1246), which showed difference by more than 10 times. We enrolled 100 patients with bifurcation lesion and 100 patients with nonbifurcation lesion. As illustrated in Table 2, there were no significant differences in the two groups, including the age, the gender, the status of smoking, hypertension, and diabetes.

As shown in Figure 3, RT-PCR analysis showed that circulating serum level of miR30d was profoundly elevated in patients with coronary bifurcation lesion compared to nonbifurcation lesion patients ($P < 0.05$). Circulating level of

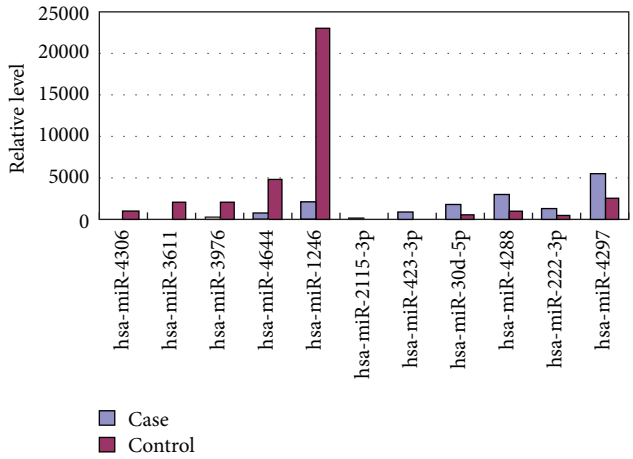


Figure 3: qRT-PCR analysis of several miRNAs differentially expressed in case and control groups. Case: bifurcation lesion patients; control: nonbifurcation lesion patients. Shown were representative data from three independent experiments.

miR222 was modestly but not significantly increased in coronary bifurcation lesion patients compared to nonbifurcation lesion patients ($P = 0.881$). Moreover, serum level of miR1246 was significantly lower in bifurcation lesion patients than in nonbifurcation lesion patients ($P < 0.05$). These data confirmed our results of microarray analysis.

4. Discussion

Coronary artery disease (CAD) is a multifactorial disease that can be influenced by a multitude of environmental and heritable risk factors. CAD has serious impact on human life and health. Coronary atherosclerosis is a main reason that causes CAD. Atherosclerosis is a chronic and progressive pathologic process characterized by the accumulation of lipid and fibrous elements in the large arteries, which causes a number of cardiovascular-related diseases. The development of atherosclerosis involves the following steps: foam cell formation, fatty streak accumulation, migration and proliferation of vascular smooth muscle cells (VSMCs), and fibrous cap formation. Finally, the rupture of the unstable fibrous cap causes thrombosis that leads to unstable coronary syndromes, myocardial infarction, and stroke.

Coronary intervention is an effective means of coronary heart disease treatment. Coronary bifurcation is prone to develop atherosclerotic plaque due to turbulent blood flow and high shear stress. Treatment of coronary bifurcation lesion represents a challenging area in interventional cardiology but recent advances in percutaneous coronary interventions (PCI) have led to the dramatic increase in the number of patients successfully treated percutaneously. Compared with nonbifurcation interventions, bifurcation interventions have a lower rate of procedural success, higher procedural costs, longer hospitalization, and a higher clinical and angiographic restenosis. Introduction of drug-eluting stents (DES) has resulted in a lower event rate and reduction of main vessel (MV) restenosis. However, side branch (SB) ostial residual stenosis and long-term restenosis remain a problem [9].

MicroRNAs (miRNAs) are a class of short, noncoding, single stranded RNA molecules, approximately 22 nucleotides in length. They negatively regulate gene expression either through inhibition of mRNA translation or by promoting mRNA degradation. Emerging evidence suggests that miRNAs are pivotal regulators of various processes including cell proliferation, differentiation, apoptosis, survival, motility, and morphogenesis [10]. Recently, specific miRNA expression profiles have been reported as a prognostic factor or a predictive factor for disease progression. In particular, serum miRNAs may be used as a biomarker in diagnosis [6]. In pathological process, miRNAs are linked to myocardial hypertrophy, myocardial fibrosis, heart failure, and arrhythmias. A series of miRNAs are involved in pathological progression of coronary heart disease and play pivotal roles in the development of the disease [11–13].

Hoekstra et al. investigated the potential of miRNAs as biomarkers for CAD and reported that unstable angina pectoris patients could be discriminated from stable patients based on the relatively high expression levels of miR-134, miR-198, and miR-370 in peripheral blood mononuclear cells [14]. Therefore, we hypothesized that in CAD serum miRNA levels could be changed and those miRNAs can be used as the biomarkers. In this study, we performed microarray analysis to identify 197 miRNAs differentially expressed in the serum of patients with coronary bifurcation lesion, including 150 miRNAs upregulated and 47 miRNAs downregulated. After a rigorous selection, we screened 3 miRNAs differentially expressed, including downregulated miR1246 and upregulated miR30d and miR222. Further large sample validation in 200 patients demonstrated a distinct serum miRNA expression pattern in patients with coronary bifurcation lesion; miR30d was upregulated and miR1246 was downregulated compared with nonbifurcation lesion patients. Consistent with our results, a recent study reported that miR30d was upregulated in diabetic cardiomyopathy [15]. Therefore, miR30d may be a promising marker and therapeutic target for various cardiovascular diseases. miR222 has been shown to play an important role in the regulation of vascular inflammation [16]. In this study, we found that circulating level of miR222 was not significantly increased in coronary bifurcation lesion patients compared to nonbifurcation lesion patients. It will be important to further confirm the function of miR222 in coronary bifurcation lesion. Up to now, miR1246 is mainly reported to be involved in cancer development [17, 18]. The role of miR1246 in cardiovascular diseases needs further studies.

5. Conclusions

This is the first study using microarray method to investigate the association between serum miRNAs and coronary bifurcation lesion. Evaluation of the screened up- and downregulated miRNAs according to their target mRNAs and biological significance will give some clues for their functional role in coronary bifurcation lesion. Our findings suggest that these

miRNAs may have a role in the pathogenesis of coronary bifurcation lesion and may represent novel biomarkers for the diagnosis and prognosis of coronary bifurcation lesion.

Conflict of Interests

The authors declare that they have no conflict of interests.

Authors' Contribution

Yan Liu and Shaoliang Chen designed the study and wrote the paper. Junjie Zhang, Shoujie Shan, Liang Chen, Rong Wang, Jing Kan, and Tian Xu performed the experiments and analyzed the data. All authors read and approved the final paper.

Acknowledgments

This study is supported by Jiangsu Provincial Special Program of Medical Science (BL2013001). The authors thank the patients who participated in the study with great enthusiasm, the clinic doctors and nurses, and Wu Jie (Shanghai Kangcheng Biotechnology companies) and Zhang Lei (Shanghai Invitrogen Biotechnology companies) for technical assistance.

References

[1] J. Lian, J. Guo, Z. Chen et al., "Positive association between GCKR rs780093 polymorphism and coronary heart disease in the aged Han Chinese," *Disease Markers*, vol. 35, no. 6, pp. 863–868, 2013.

[2] J. Wu, R.-X. Yin, Q.-Z. Lin et al., "Two polymorphisms in the Fractalkine receptor *CX3CR1* gene influence the development of atherosclerosis: a meta-analysis," *Disease Markers*, vol. 2014, Article ID 913678, 13 pages, 2014.

[3] C. Frangos, S. Noble, N. Piazza et al., "Impact of bifurcation lesions on angiographic characteristics and procedural success in primary percutaneous coronary intervention for ST-segment elevation myocardial infarction," *Archives of Cardiovascular Diseases*, vol. 104, no. 4, pp. 234–241, 2011.

[4] X. Zhou, P. Yuan, and Y. He, "Role of microRNAs in peripheral artery disease (review)," *Molecular Medicine Reports*, vol. 6, no. 4, pp. 695–700, 2012.

[5] A. S. M. Sayed, K. Xia, T. L. Yang, and J. Peng, "Circulating microRNAs: a potential role in diagnosis and prognosis of acute myocardial infarction," *Disease Markers*, vol. 35, no. 5, pp. 561–566, 2013.

[6] S. de Rosa, A. Curcio, and C. Indolfi, "Emerging role of micrornas in cardiovascular diseases," *Circulation Journal*, vol. 78, no. 3, pp. 567–575, 2014.

[7] G. Sianos, M. A. Morel, A. P. Kappetein et al., "The SYNTAX score: an angiographic tool grading the complexity of coronary artery disease," *EuroIntervention*, vol. 1, pp. 219–227, 2005.

[8] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14863–14868, 1998.

[9] D. Siu, "A new way of targeting to treat coronary artery disease," *Journal of Cardiovascular Medicine*, vol. 11, no. 1, pp. 1–6, 2010.

[10] V. Portnoy, V. Huang, R. F. Place, and L.-C. Li, "Small RNA and transcriptional upregulation," *Wiley Interdisciplinary Reviews: RNA*, vol. 2, no. 5, pp. 748–760, 2011.

[11] S. Gal-Ben-Ari, J. W. Kenney, H. Ounalla-Saad et al., "Consolidation and translation regulation," *Learning and Memory*, vol. 19, no. 9, pp. 410–422, 2012.

[12] J. M. Lorenzen, F. Martino, and T. Thum, "Epigenetic modifications in cardiovascular disease," *Basic Research in Cardiology*, vol. 107, no. 2, article 245, 2012.

[13] S. Fichtlscherer, A. M. Zeiher, and S. Dimmeler, "Circulating microRNAs: biomarkers or mediators of cardiovascular diseases?" *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 31, no. 11, pp. 2383–2390, 2011.

[14] M. Hoekstra, C. A. C. van der Lans, B. Halvorsen et al., "The peripheral blood mononuclear cell microRNA signature of coronary artery disease," *Biochemical and Biophysical Research Communications*, vol. 394, no. 3, pp. 792–797, 2010.

[15] X. Li, N. Du, Q. Zhang et al., "MicroRNA-30d regulates cardiomyocyte pyroptosis by directly targeting foxo3a in diabetic cardiomyopathy," *Cell Death and Disease*, vol. 5, no. 10, Article ID e1479, 2014.

[16] T. Staszel, B. Zapała, A. Polus et al., "Role of microRNAs in endothelial cell pathophysiology," *Polskie Archiwum Medycyny Wewnetrznej*, vol. 121, no. 10, pp. 361–367, 2011.

[17] N. Takeshita, I. Hoshino, M. Mori et al., "Serum microRNA expression profile: miR-1246 as a novel diagnostic and prognostic biomarker for oesophageal squamous cell carcinoma," *British Journal of Cancer*, vol. 108, no. 3, pp. 644–652, 2013.

[18] B. Zhang, J. Chen, Z. Ren et al., "A specific miRNA signature promotes radioresistance of human cervical cancer cells," *Cancer Cell International*, vol. 13, no. 1, article 118, 2013.

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 321370, 12 pages
http://dx.doi.org/10.1155/2015/321370

Review Article

Proteomics in Cancer Biomarkers Discovery: Challenges and Applications

Reem M. Sallam[1,2,3]

[1]*Clinical Chemistry Unit, Department of Pathology, College of Medicine, King Saud University, Riyadh 11461, Saudi Arabia*
[2]*Obesity Research Center, College of Medicine, King Saud University, Riyadh 11461, Saudi Arabia*
[3]*Department of Medical Biochemistry and Molecular Biology, Faculty of Medicine, Ain Shams University, Cairo 11566, Egypt*

Correspondence should be addressed to Reem M. Sallam; sallam@ksu.edu.sa

Received 1 October 2014; Revised 15 January 2015; Accepted 18 February 2015

Academic Editor: Tavan Janvilisri

With the introduction of recent high-throughput technologies to various fields of science and medicine, it is becoming clear that obtaining large amounts of data is no longer a problem in modern research laboratories. However, coherent study designs, optimal conditions for obtaining high-quality data, and compelling interpretation, in accordance with the evidence-based systems biology, are critical factors in ensuring the emergence of good science out of these recent technologies. This review focuses on the proteomics field and its new perspectives on cancer research. Cornerstone publications that have tremendously helped scientists and clinicians to better understand cancer pathogenesis; to discover novel diagnostic and/or prognostic biomarkers; and to suggest novel therapeutic targets will be presented. The author of this review aims at presenting some of the relevant literature data that helped as a step forward in bridging the gap between bench work results and bedside potentials. Undeniably, this review cannot include all the work that is being produced by expert research groups all over the world.

1. Introduction

In the -omics era, the nature of high-throughput technologies, their capabilities, limitations, performance quality, and applicability are among factors determining their significance and influence not only in pure exploratory research, but also in potential clinical use.

Advances to the field of genomics and related computational tools are constantly being produced and applied in cancer-related research [1]. However, other fields are needed to complement the limitations of the genomics approach.

Proteomics-based strategy in studying diseases is considered one of the dynamic and innovative tools that could confirm, complement, or quite often provide more elaborate information beyond that obtained by other high-throughput approaches. While several genes were identified by genomics technologies to be specifically related to cancers [2], the function of such genes and the data interpretation in the context of functional networks require the power of proteomics. Moreover, although studies focusing on detecting the differential expression of mRNA have been extremely informative, they do not necessarily correlate with the functional protein concentrations. Macromolecules, in general, and proteins, in particular, are highly dynamic molecules. Mechanistically, proteins can be subjected to extensive functional regulation by various processes such as proteolytic degradation, post-translational modification, involvement in complex structures, and compartmentalization. Proteomics is concerned with studying the whole protein repertoire of a defined entity, be it a biological fluid, an organelle, a cell, a tissue, an organ, a system, or the whole organism. Therefore, in-depth studying of proteomics profiles of various biospecimens obtained from cancer patients are expected to increase our understanding of tumor pathogenesis, monitoring, and the identification of novel targets for cancer therapy. In addition, an essential goal for applying proteomics to study cancers is to adapt its high-throughput tools for regular use in clinical laboratories for the purpose of diagnostic and prognostic categorization of cancers, as well as in assessing various cancer therapeutic regimens.

Similar to other high-throughput technologies, proteomics has been generating a vast amount of data in the form of lists of hundreds or thousands of proteins that are differentially expressed, whether increase or decrease, as a cause or consequence of ongoing physiological, developmental, or pathological events. Interpretation and analysis of such flood of information depend on building on existing data stored in constantly updated databases. Obviously, researchers have to be extra-cautious in designing their work in the first place, ensuring that good analytical tracks are being undertaken, to avoid snow ball effect and erroneous outcomes [3]. Scientifically sound analysis of the information flow as it represents complex networks and interactions of intra-, inter-, and extra-cellular environments should be the ultimate goal. Unraveling such complexity is the focus of interest for several research groups. For instance, a mass spectroscopy- (MS-) based draft of human proteome has been recently reported, which incorporated huge amount of proteomics data both from public accessed databases as well as from several research groups' work [4].

The complexity of proteomics technologies when applied to cancer research increases even more due to the current concept of cancer heterogeneity. As a matter of fact, cancer heterogeneity and biospecimen variables are considered by some researchers the most crucial and challenging point for all –omics technologies at their application in cancer studies [5].

Moreover, an integrated approach for research performed on cancers and diseases, in general, is recommended when designing studies with the intention of discovering disease biomarkers as argued by George Poste: "...The dismal patchwork of fragmented research on disease-associated biomarkers should be replaced by a coordinated 'big science' approach" [6]. Such study designs have to comply with standardized and validated guidelines.

2. Mechanisms of Proteomic Changes in Cancer

Although exact causes of most cancers are not clearly defined, cancer is thought to result from a combination of genetic and environmental abnormalities. Several genomic defects have been implicated, including mutations, variation in copy number, chromosomal anomalies, and alternative splicing. One potential mechanism for the proteomic variation in cancer is the ubiquitous aneuploidy, which is defined as an imbalanced chromosomal content [7]. Aneuploid cells are thought to be under proteotoxic stress as a result of defective proteostasis; the latter is the state of dynamic equilibrium in which protein synthesis and correct folding are balanced with protein degradation. This state is a manifestation of several machineries that cooperatively ensure proper protein turnover while allowing for the conformational flexibility that is critical for proteins' biological functions. Therefore, defective proteostasis will result in not only proteotoxic stress, but also cellular dysfunction and subsequent pathologies [8]. Recent findings have shed some light into the yet-not-fully-understood mechanisms underlying the association between aneuploidy, proteotoxic stress, and abnormal cellular proliferation and tumorigenesis [7]. However, this association is still a matter of controversy and is lacking straightforward relationship pattern; for instance, an extra chromosome that results in increased gene expression and a theoretical increased protein production is not necessarily translated into an actual elevation of circulating protein levels, since there is high possibility of overwhelming the cellular protein folding apparatus, leading to chronic protein misfolding and subsequent protein degradation. It is proposed that certain proteins, such as various kinases and multimeric protein complexes, have increased requirements for the cellular protein folding apparatus, and hence they are more susceptible to misfolding than others. This and other relevant examples are comprehensively reviewed by Donnelly and Storchová [7]. Emerging evidence linking aneuploidy, defective proteome, and cancer development is of obvious significance as it provides potential for treating aneuploid cancer cells using suitable antineoplastic agents targeting the proteostatic machinery [9]. This will be discussed in more details later.

Another potential mechanism for proteomics changes in cancers is the consequence of defective protein structure and hence function. Mutations in cancer-associated genes can be manifested in defective protein structure. These defects can exert their deleterious impact through changing protein stability and causing the protein to be more susceptible for degradation; changing the protein's functional site residues; or changing the affinity controlling protein-protein interactions [10].

Genomic and proteomic changes in cancer could be further followed up by the recently emerged field of "interactome profiling" focusing on network-centered approach, that is providing an enormous amount of data representing protein interactions and the influence of protein structures. This is reviewed recently by Gulati and coworkers and is beyond the scope of the current review [11].

3. Cancer Biomarkers' Applications: Challenges and Recommended Solutions

3.1. Cancer Heterogeneity. The current concept of cancer heterogeneity and biospecimen variables is considered by some researchers as one of the most crucial and challenging points for proteomics as well as for other -omics technologies, at their application in cancer studies. Recently, intratumoral heterogeneity was examined in invasive breast cancer, comparing biospecimens obtained by intraoperative image-guided, core-needle biopsies to surgical biopsies taken from the center and the periphery of cancer breast. Proteomics techniques undertaken in that study have demonstrated that even though most biomarkers studied did not manifest significant intratumoral heterogeneity, protein and phosphoprotein levels were affected by biospecimen type, as well as by other preanalytic variables, including surgical manipulation and the duration of cold ischemia [5]. A recent approach to circumvent the challenge of tumor heterogeneity and to extract meaningful data from formalin-fixed tissue blocks has

been suggested, combining matrix-assisted laser desorption ionization (MALDI) with imaging (MALDI imaging mass spectroscopy; MALDI-IMS). This approach is unique as it allows proteomics-based studies to provide both patient-specific and cancer-specific information as a means for biomarker discovery and cancer tissue classification. It also provides morphology-based proteomics analysis for cancer tissue [12]. In addition, studies using MALDI-IMS analysis of specific cancer tissues generate peptide reference datasets to facilitate peptide identification in future studies on the same cancer type. However, several technical challenges still exist including low signal to noise ratio and low mass accuracy [13]. In a recent work studying prostate cancer, Shipitsin and coworkers have developed a biopsy simulation procedure by tissue microarrays aiming at exaggerating prostate cancer tissues' variation that is expected in clinical practice. Their approach has provided a useful model for predicting cancer aggressiveness through reliable biomarkers, regardless of sample variation [14].

3.2. Cancer Early Detection. Detecting cancer at an early stage, when there is a better chance for its treatment, is a real challenge to the scientific and medical communities as most clinical blood biomarkers assays do not have the required sensitivity and specificity necessary for that purpose. In an interesting approach focusing on ovarian cancer, Hori and Gambhir have recently developed a mathematical model looking at the estimated time at which ovarian cancer can be detected by measuring the amount of the cancer antigen 125 (CA 125) shed from the tumor during its growth. Surprisingly and despite the reported sensitivity of the CA 125 measuring assays, the authors reported that a tumor could grow unnoticed for more than 10 years and reach a size of more than 2.5 cm before becoming detectable. This mathematical approach might yield similar finding in other tumor types, and the model can be extended to virtually any solid cancer and associated biomarkers, according to the authors' suggestions [15]. Nevertheless, a lot of debate has emerged regarding the applicability of this approach in other types of tumor and the sort of assumptions used in its calculation [16]. This example illustrates a unique approach to test the applicability of circulating biomarkers' assays in early cancer detection, cancer prognosis, and therapeutic response and monitoring. Combining panels of circulating biomarkers, rather than a single molecule, with newly developed or newly updated technologies such as imaging procedures might be more informative in terms of early diagnosis, accurate assessment of the prognosis, and response to therapy in cancer patients [16, 17].

3.3. Protocols for Developing Tumor Biomarkers. More than a decade ago, several research groups have formulated multisteps strategies for developing tumor biomarkers. Hammond and Taube's phased approach involved the following steps/phases: the biomarker discovery, the development of an assay system, the performance of preliminary analysis for the biomarker's clinical potential, the standardization and assessment of the biomarker's measuring assay, and finally the validation of that assay for clinical use [18]. Despite the strict step-wise analytical criteria of this strategy, pre-analytical issues were not sufficiently addressed. About the same period of time, Pepe and colleagues suggested another strategy that focused on the need for accurate definition of the study's aims and its outcomes together with strict criteria for specimen selection, sample size calculation, and experimental methods [19]. Several years later, the same group has suggested a more rigorous study design for the development of tumor biomarkers, emphasizing that the described design would maintain a high research quality and improve the possibility of obtaining a clinically promising biomarker ready for subsequent rigorous scrutiny [20]. Common biases that plagued the process of biomarker discovery research were claimed to be avoided if this design, which was called "nested case-control study design" was strictly followed. The design included prospective collection of specimens before outcome ascertainment from a case-control study cohort that is relevant to the clinical application under study, and blind assessment of the biomarker in specimens obtained from randomly selected case and control subjects. The authors described various aspects of their design in relation to the clinical context, biomarker performance criteria, biomarker test, and study size [20].

3.4. General Guidelines for a Good Study Design for Biomarkers' Discovery. In order to plan a good study design for cancer biomarkers discovery, several aspects have to be meticulously tackled. This was reviewed in more details elsewhere [21, 22] and is summarized in the following section. Firstly, careful planning starting with the formulation of a research question supported by convincing evidence for its importance and relevance to a clinically pressing problem. A rational choice of the most suitable analytical tests to approach this research question is of equal significance. The performance characteristics for such test(s), in terms of specificity, sensitivity, and positive and negative predictive power, should be appropriate for the experimental design and clearly described. In addition, theverification and validation strategies of the method(s) performed and the clear and detailed description of the samples' nature, collection, and storage protocols have to be openly defined. Details of the samples' source, as the subjects' age, gender, disease stage, medications taken, and lifestyle are necessary to be highlighted as well. Furthermore, in cancer tissues' biomarkers-related research, the sampling procedures are of critical importance due to their heterogeneity. Therefore, collecting a representative sample is important in order to obtain reliable data. Likewise, sample size calculation is a crucial component of the study coherence and if carefully conducted will average out sample heterogeneity. Moreover, protocols of executing the experiments should maintain basic and critical points, such as incorporating proper blank(s), positive and negative control samples, and reference compound(s) within *each* run for the analytical procedure. Details of the quality performance of instrumentation and their calibration are equally important for the procedure's validation. Collectively,

every step in the study design and execution has to be clearly described in sufficient details to allow for reproducing the work and/or comparing the data.

The scientific communities have been working diligently to standardize the procedures of proteomics-generated data optimum utilization. Useful data repositories have been constructed such as Panorama (https://panoramaweb.org/) that, together with portals for proteomics assays involved in targeting cancer-related proteins and peptides, will enable researchers interested in specific protein or peptide to obtain the standard operating procedures (SOPs) for those assays, their quality assessment, and validation proofs [22].

4. Proteomics Techniques Used in Cancer Research

Research studying protein alterations in cancer existed for more than 70 years [23]; however it was only in the last 3 decades or so that recent proteomics technologies have been extensively utilized in deciphering protein differential expression in human cancers [24]. Various approaches have been carried out, taking advantage of the recent analytical techniques and advanced bioinformatics. In general, two main proteomics tracks can be undertaken, the "shotgun" or "bottom-up" methods and the targeted proteomics methods. A recent set of "best practices" for MS-based assay development using the concept of "fit-for-purpose" was recently published following a workshop that was held in mid-2013 in the United States of America's National Institutes of Health (NIH) with representatives from different institutes involved in the development and/or utilization of targeted proteomics assays [22]. The following section starts by briefly describing basic techniques such as 2D gel electrophoresis, difference in gel electrophoresis (DIGE), and MS, followed by introducing more recent technologies and combined applications such as protein microarray and combined proteomics and imaging methods.

Polyacrylamide gel electrophoresis (PAGE) and isoelectric focusing (IEF) have been the basis for the 2D PAGE techniques resolving proteins based on their molecular mass and isoelectric points, respectively. This approach has been frequently applied to analyze cancer cells proteins for more than 2 decades [25] and is still in use [26]. Further advancement in this approach has been the result of introducing fluorescent dyes and in-gel comparative proteomic analysis in the technique of 2D-DIGE. This is usually coupled to protein spot analysis by fluorescence gel scanners, spots' picking, and enzyme digestion, followed by identification by one of the MS-based available techniques.

Advancements in MS resulted in optimal performance in the low mass range of proteins. In-depth profiling of plasma and other biofluids proteomes results in identification of proteins that span more than six logs of protein abundance. As such, it has been the method of choice in many cancer applications [24]. To detect low-abundance proteins, an initial samples' preclearing step might be performed to remove the high-abundance proteins, such as albumin and immunoglobulins. However, this carries the risk of depleting the samples from the low-abundance and low-molecular weight proteins that are bound to the circulating carrier proteins. The latter have been demonstrated to act as a reservoir storing diagnostic information within the accumulated bound low-molecular weight potential biomarkers [27]. Incorporation of bead-based immunoassays may also be used to better identify low abundance proteins [17].

MS use in protein analysis has undergone several stages of technical advancement and improved instrumentation efficiency. MALDI-MS [28] and electrospray ionization (ESI) MS, combined with advancements in protein fractionation and separation, as liquid chromatography (LC) and gas chromatography (GC) and labeling techniques, are examples of such technological developments. This has been thoroughly reviewed in several articles [24, 29–31]. More recently, proteomics approach has been extended to involve studying of epigenetic processes in cancer research. The use of MS-based proteomics in studying various aspects of chromatin biology and in evaluating specific histone posttranslational modifications resulted in the discovery of chromatin-associated proteins and multisubunit complexes that can be considered epigenetic biomarkers with future potential in cancer diagnosis and therapy. This has gained a wide attention and was recently reviewed by Bartke et al. [32].

Microarray is considered one of the most exciting developments in high-throughput technologies. Simultaneous measuring of the expression of thousands of genes (gene microarray) or proteins (protein microarray) and detection of genomic or proteomic biomarkers, respectively, that are tightly linked to cancer development and/or progression have revolutionized the cancer research studies. Recently, such technologies have been applied to study a relatively uncommon category of cancer patients who are presenting with metastatic cancer without any obvious anatomically detectable primary tumor, the so called cancer of unknown primary or CUP [33].

In addition, the targeted proteomic approach of selected reaction monitoring (SRM) has been developed and widely applied, for instance, to detect mutant proteins in the colorectal cancer tissue and in the fluid obtained from potential precancerous pancreatic cysts [34]. Other recent approaches have been described in the literature, including a multiplexed microfluidic immunohistochemistry-/immunocytochemistry-based quantitative proteomics profiling of cancer samples [35].

Combining proteomics and imaging-based methods has been recently described. Shipitsin and coworkers were able to identify a panel of 5 protein biomarkers for prostate cancer lethality using an automated, integrated quantitative multiplex immunofluorescence in situ imaging approach [36]. Such combination is thought to produce more clinically representing data in terms of the actual *in vivo* environment where the active proteins exert their functions. This is because such approach was designed to measure the levels and activity status of protein biomarkers in defined intact tissue regions, avoiding the need to lyse the tissues of interest that is commonly performed in the traditional proteomics approaches. Wider range of applications and comparative studies to more established proteomics approaches is still in

progress. In a different context, integrating proteomics and imaging tools to gain more insight into the pathogenesis of cancer progression and penetrability at the molecular level is recently experimented. An article describing such mechanistic-oriented approach was recently published by Oh and colleagues [37]. This group used their advanced integrative tools to study caveolae at the blood-solid tumor interface *in vivo* aiming to reveal molecular portals to infiltrate solid tumors of mammary, prostate, and lung origins. They were able to reveal a transvascular pumping system and define some of its component proteins, as caveolin 1 and annexin A1, that are affecting tumor uptake of various agents. Such approach will probably get a large scale attention as it can be applied to assess the effectiveness of therapeutic agents based on their ability to cross the biological barriers *in vivo* and find their way into the solid tumors.

5. Examples of Proteomics Research Applications in Various Cancer Types

In various types of cancer, the biomarkers discovery is expected to improve one or more of the following critical applications: early diagnosis and prognosis and monitoring of disease progression, its response to therapy, and its recurrence. High-throughput hypothesis-generating methods have revealed hundreds to thousands of cancer associated proteins (CAPs). This implies that hundreds to thousands of potential protein biomarkers have been suggested in the literature and are awaiting proper validation. It is only after validating these molecules that they can be considered for application in diverse clinical setting such as diagnosis, prognosis, disease staging, and patients' categorization. This is a critical aspect in translational cancer research [38, 39]. Classically, hypothesis-testing has been performed using antibody-based methods, such as enzyme-linked immunosorbent assays (ELISA). However the limited availability of validated ELISAs and their high cost and time-consuming nature, together with the technical challenges of assay multiplexing, all have been obstacles hindering the use of these assays in validating the rapidly evolved lists of potential cancer protein biomarkers [40].

Due to the lagging in high-throughput hypothesis-testing methods, these CAPs cannot yet be applied in clinical setting. Therefore, a pressing need for accurate, precise, and sensitive validation assays has been the driving force for an ongoing extensive recent research. One promising track has generated selected reaction monitoring (SRM) assays of targeted proteomics. SRM assays have been recently developed and refined for many human CAPs that are functionally related to cancer driver mutations. They have been used to measure the detectability of target proteins in the circulation or urine and have resulted in reproducible quantification across cohort of cancer patients' samples. Therefore, these assays are thought to represent a valuable resource for accelerating and planning biomarker verification studies [41].

The following section describes some of the proteomics research outcomes in three of the most-studied cancers: lung, breast, and ovarian cancers. A more detailed discussion will be presented for the ovarian cancer aiming at emphasizing few points of critical significance. For instance, various perspectives in approaching the subject of cancer biomarkers, the need to standardize and optimize study design, preanalytical and analytical assays components, and strict validation strategies are among the points to be discussed in more detail for ovarian cancer. A description of how proteomics has helped in clarifying the ovarian cancer markers for carcinogenesis, cancer progression, diagnosis, prognosis, and targets for therapeutic treatment will be presented as well.

5.1. Lung Cancer Biomarkers: Implications from Proteomics. Lung cancer signatures in plasma have been studied both in mouse models and in human with data implying concordance between findings in both species [42]. Circulating levels of EGFR as well as other biomarkers as SFTPB and WFDC2 were significantly different in lung cancer cases relative to control. As with the case of other types of cancer, finding a marker or a panel of markers, which has a screening power, if measured in prediagnostic biological sample is an important goal, as it carries the potential for application in early detection or even screening strategies of lung cancer in addition to the potential use for monitoring subjects diagnosed with the disease [42]. Unfortunately, this goal is not yet achieved. As mentioned in previous sections, recent technologies integrating proteomics and imaging tools are being used, with promising results, to gain better insight for the pathogenesis of lung cancer at the molecular levels. This is expected to improve our understanding of the effectiveness of anticancer therapies in terms of their ability to be successfully delivered in the required dosage to solid tumors [37].

5.2. Breast Cancer Biomarkers: Implications from Proteomics. Proteomics approach in studying breast cancer has also been progressing and yielding promising findings with both diagnostic and therapeutic applications. An example of combined *in vitro* and *in vivo* approaches involving deep analysis of cultured breast cancer cell lines was obtained from tumors of defined breast cancer stages and validated using human breast cancer tissue. This approach has demonstrated that the tumor stage-specific proteomic signatures extracted from the *in vitro* study were validated on tissue microarrays. Transformed cells showed proteomic signatures characterizing the loss of tissue architecture and the cellular metabolic changes [43]. Remarkably, recent work has shown that the plasma proteome in breast cancer also indicates the tumor microenvironment-derived proteins involved in a number of innate physiologic processes such as wound repair, immune response, and tissue remodeling [44].

5.3. Ovarian Cancer Biomarkers: Implications from Proteomics. Recent epidemiological studies have demonstrated that ovarian cancer remains a serious condition that is considered the most lethal gynecological malignancy [45]. Unfortunately, very few cases are diagnosed at clinically early stages, and the vast majority are diagnosed at late stages with the tumor already spread distantly [45]. Moreover, due to the low prevalence of ovarian cancer, no screening

test is available for population screening. In fact, with the condition's low prevalence, a screening test has to be of an extremely high specificity to possess an acceptable positive predictive value [46]. The condition is pathologically not-very-well understood, and, few years ago, the use of molecular profiling has confirmed that ovarian cancer represents a heterogeneous class of diseases that are sharing a common organ [47]. Therefore, there is a pressing need for discovering novel biomarkers to improve the outcomes of such a serious disease.

Ovarian cancer-induced altered biologic processes are expressed as aberrant molecules that belong to various biochemical families, such as DNA, mRNA, proteins (and related subfamilies as glycosylated proteins, peptides, and autoantibodies), and metabolites. The recent technology breakthroughs in genomics and proteomics fields have positively influenced our understanding of the pathophysiology of the disease.

To date, only 2 individual circulating biomarkers, CA 125 and human epididymis protein 4 (HE4), are approved by the American Food and Drug Administration (FDA) for monitoring treatment and detecting recurrence in ovarian cancer patients. In addition, the FDA has recently approved two algorithms to be used clinically as a supplement for decision making for preoperative adnexal mass patients [46].

Several research groups worldwide focus on studying the altered biology in ovarian cancer and to discover promising molecular mediator(s) as biomarkers or as therapeutic targets, using proteomics tools. Such tools target, beside the proteins repertoire, other related biochemical entities, for example, the glycosylated proteins (glycomics), the low molecular weight peptides (peptidomics), the metabolites (metabolomics), and the antitumor antibodies (immunoproteomics). These entities have been reviewed recently by Leung et al. [46]. Mechref and coworkers have described major advances in both preanalytical separation methods and MS that allowed for increasingly comprehensive characterization of glycosylated proteins repertoire (the glyome) and cancer-specific glycoproteins in various types of cancers including ovarian cancer [48]. Although the detection and characterization of aberrantly glycosylated proteins in biospecimens still face technical challenges, recent advances in MALDI-MS and in the preanalytical enrichment methods such as peptide-N-glycosidase digestion and chromatographic separation have enabled glycoproteomics techniques to positively add to the list of cancer-specific glycoproteins [49–51]. Glycosylation as a posttranslational modification is described as heterogeneous, structurally complex, widespread, and cell- and protein-specific process. Therefore, in studying the cancer-specific glycans, researchers are faced with both technical limitations and uncertainty in the biological interpretation. Examples of the technical challenges are the heterogeneity of the glycans resulting in a collection of glycoforms and isomers for each glycoprotein and the limited ability of most current proteomics technologies to precisely differentiate these forms and isomers. Moreover, following the discovery of candidate glycan biomarker(s), there should be reliable quantitative validation assays, with good specificity for the glycan epitope, as well as good sensitivity. Currently, there have been trials to develop such assays, using lectin or antibody capturing technology; however these are not yet sufficiently available for strict validation. To complicate the matters even more, researchers struggle to determine the biological implications for the aberrant glycoproteins' profiles in cancer states. In the context of ovarian cancer, it is not clear whether the components of glycomic profiles published in the literature are unique to this cancer or, alternatively, are a consequence of cancer-related metabolic defect(s). Therefore, more rigorous investigations are clearly needed in this field [52].

The study of the global metabolites population in biospecimens, the metabolome, by MS-based assays has been increasingly utilized in the field of cancer biomarkers discovery. Biological fluids as urine and serum or plasma are the usual specimens used. Urine specimens are sometimes preferred in proteomics and related technologies for the biomarkers discovery over serum or plasma. The reasons for this preference include the relatively low total protein concentration in normal urine and the noninvasive nature for urine sample collection. Urine sample is relatively free from high molecular mass proteins making it less complicated than the serum/plasma samples [53]. Proteomics technologies as ultraperformance LC quadrupole time-of-flight MS (UPLC-Q-TOF MS), hydrophilic interaction chromatography, and reversed-phase LC MS were able to identify several metabolites in the urine of ovarian cancer patients as compared to healthy control subjects. Interestingly, some of these metabolites were discriminatory between early and late clinical stages of those patients [54, 55]. Recently the metabolomics profiles of plasma samples obtained from epithelial ovarian cancer (EOC), benign ovarian tumor (BOT), uterine fibroid, and healthy controls using UPLC were published. Fifty-three metabolites were identified in this work as specific biomarkers for EOC. Again, these metabolites were able to discriminate EOC from BOT and uterine fibroids, as well as early-stage from late-stage EOC. The critical analysis of the aberrant metabolites has identified unique metabolic pathways that were disturbed in cancer cases, namely, those of phospholipids metabolism, tryptophan catabolism, and fatty acid β-oxidation. These findings are expected to increase our understanding of ovarian cancer pathophysiology [56]. Despite the noticeable advances in this approach, a number of confounding variables are still hindering the introduction of metabolomics for full clinical application. Technical limitations include the biases related to preanalytical factors as sample collection and storage conditions. Biological limitations involve the unstable nature of metabolites that may be extensively transformed during transition from the cancer site to the biospecimen collected or even after collection. Moreover, other confounding factors include the subjects' age, smoking habits, sleep patterns, and lifestyle. Hence, standardized and robust protocols are needed to eliminate such biases and to allow for assay's precision [46].

Ascites fluid has been studied as a source for proteomics and metabolomics potential biomarkers in ovarian cancer, with an advantage over plasma or serum due to its close proximity to the site of the tumor. Comparing malignant ascites with cirrhosis ascites' metabolomes has identified 41 metabolites that differed significantly between both pathologies.

Detailed analysis of these metabolites has revealed that most of the cancer-specific metabolites belong to signaling pathways. Similarly, proteomic analysis has identified even more molecules discriminating the ovarian cancer from cirrhosis ascites. Interestingly, spliceosomal proteins and RNAs were found in the ovarian cancer ascites, a finding suggesting that these molecules might play an essential role in intercellular communications between cancer cells [57].

More recently, low-molecular weight proteomics or peptidomics has been used in studying biospecimens as blood, urine, ascites, or even tumor tissue, seeking to identify unique biomarkers for the ovarian cancer [58–60]. This approach, although still in the beginning, is promising, and, if standardized, is suggested to complement the conventional proteomics approach as it reflects the cancer-related protease activity. However, the lack of standardized protocols and of robust quantifying validation assays is hindering its widespread use.

The last decade has also witnessed a novel approach in cancer biomarkers discovery that targets identifying cancer related populations of antitumor antibodies, the so called immunoproteome [61, 62]. However, similar to peptidomics, this approach is still lacking appropriate validation assays before any applications for the identified potential biomarkers can be suggested.

As mentioned above, the proteomic profiling of plasma is quite challenging due to the high dynamic range of protein concentration which makes it hard to identify low-abundance proteins. Some researchers have turned their attention into more proximal biofluids such as the ovarian-tumor-tissue-interstitial fluid as more promising sample sources [63]. However, because a good sample for clinical use should be easily accessible, results of the biomarkers candidates produced by this approach must be assessed and rigorously tested in more clinically relevant body fluids, such as serum, urine, or saliva, before being considered as a tumor specific biomarker [64]. As previously mentioned, screening assays for ovarian cancer among healthy individuals are lacking and are indeed seriously needed due to the aggressive course of the late-diagnosed disease. This is not-yet-feasible despite the ongoing active research. For instance, Moore and coworkers have combined an immunoassay for CA 125 with a proteomic approach: surface enhanced laser desorption ionization time of flight MS (SELDI TOF MS) to assess and quantify a panel of 7 biomarkers (apolipoprotein A1, truncated transthyretin, transferrin, hepcidin, β-2-microglobulin, connective tissue activating protein III, and inter-alpha trypsin inhibitor heavy chain 4), aiming at improving the specificity and sensitivity of detecting EOC in preclinical cases using prediagnostic serum samples [65]. The experimental design was based on previous published work by the same research group demonstrating that using this combination in postdiagnostically collected sera has increased the sensitivity for detecting ovarian cancer beyond CA 125 alone [66]. However, addition of these biomarkers to CA 125 failed to enhance the sensitivity for preclinical diagnosis [65]. The need is still pressing for a biomarker or a panel of biomarkers that could be relied on in screening for ovarian cancer. Strong and evidence-based data can be a real challenge but should not be rushed to produce. More than a decade ago, serum proteomic patterns capable of discriminating normal subjects from ovarian cancer patients were published and got a lot of attention, both in the scientific community and among the decision makers and sponsoring agents [67]. However, more scrutinizing analysis for those findings has demonstrated a defect in the experimental design that prevented their reproducibility and that emphasized on the critical importance of a good design to obtain reproducible data [3].

Essentially, extensive advances in the traditional proteomics and its more recent related technologies are producing vast amounts of data for ovarian cancer. Appropriate standardization and validation of assaying these potential biomarkers, whether individually or in combinations, are critical, prior to introducing them as clinical determinants, for screening, diagnosis, prognosis, and monitoring for the response of treatment or for recurrence.

From the previous section, it is evident that results obtained from proteomics and related technologies contributed positively in the past and are expected to remain capable of doing so in the future, to obtain better understanding for ovarian cancer. The following discussion highlights few examples of various aspects of this contribution.

(1) Ovarian cancer pathogenesis: Proteomics has resulted in better insight into the molecular bases of ovarian cancer pathogenesis. For instance, overexpression of particular signaling pathways' molecules within ovarian cancer cells have been described in the literature as a possible mechanism underlying or associated with the condition. Signaling pathways involved in cancer cell differentiation, survival (proliferation or apoptosis), migration, and metabolism are most commonly affected during the pathogenesis of cancer ovary. Examples of these pathways include the lysophosphatidic acid, the phosphatidylinositol 3-kinase, NF κB, the MAPK, and the vascular endothelial growth factor signaling pathways [68, 69]. These findings provide essential information about potential diagnostic and prognostic markers, as well as therapeutic targets for future pharmacotherapeutic-oriented ovarian cancer research. Furthermore, recent publications demonstrating the results obtained from a large Gynecologic Oncology Group trial are producing promising data. For instance, specific patterns of glycans were found to be discriminatory in distinguishing epithelial ovarian cancer and low malignant potential ovarian tumor cases from normal individuals. The candidate glycan biomarkers demonstrated sensitivity and specificity high enough to suggest further in-depth validation prior to using them as diagnostic markers for early detection of ovarian cancer [51].

(2) Etiologically, ovarian cancers can be sporadic or hereditary. Risk factors that increase women's susceptibility to ovarian cancers include genetic mutations as those reported in BRCA1 and BRCA2 genes and the mutations in the DNA mismatch repair genes characterizing Lynch syndrome [70]. Proteomics techniques

can be performed to detect the profile(s) characterizing these mutations. For instance, a proteomic signature predicting the malignant transformation of conditions with high risk of developing ovarian cancers, such as ovarian endometriosis and pelvic inflammation during ovarian carcinogenesis, is of great significance [71]. Increasing awareness of the hereditary aspect of gynecological tumors such as breast and ovarian cancer has resulted in a remarkable interest in screening populations at high risk for these malignancies. Specialized cancer centers and institutes have been formulating programs aiming at multidisciplinary coordinated approach for evaluating women with high risk of breast and ovarian cancers, organizing appropriate clinical care, updating relevant recommendations and guidelines, providing support to patients, and facilitating enrollment in appropriate research studies and registries [72]. Proteomics analysis has been performed for samples obtained during the surgical procedure of risk-reducing bilateral salpingooophorectomy (RRBSO) that is undertaken for women of high risk category. LC/MS MS and protein network database algorithms were used to evaluate the proteomic profiles characterizing the pathological changes in this group of high risk women. Few years ago, a high-throughput workflow for analyzing the proteomes of pelvic tissues (peritoneal, fallopian tube, and ovarian surface epithelial samples collected at the time of this surgery) has been described. The aim for this approach was to discover novel biomarkers that could have predictive or diagnostic value in the pelvic tissues to identify precancerous and cancerous proteomic changes of high risk deleterious mutations carriers [73].

(3) Ovarian cancer progression: The transition of benign ovarian tissue into its early malignant transformed state is such a critical step that should be extensively studied aiming to obtain a descriptive profile for it, since, as already mentioned, the ovarian cancers have notoriously poor prognosis and a highly aggressive clinical course. Proteomics technologies have been involved in following up ovarian cancer progression by evaluating the protein expression profiles in cancers of different clinical and pathological stages and in normal ovarian epithelium tissues. By performing 2D electrophoresis combined with MALDI-TOF/TOF techniques, Li and coworkers have identified 54 aberrantly expressed proteins in serous ovarian cancers. The expression of one of those proteins, the glia maturation factor beta (GMFB), was further analyzed in large cohort of patients with various stages of ovarian cancers and was found to be significantly increased as compared to normal, benign, or borderline ovarian tissues. The statistically significant positive correlation between the expression of GMFB and the FIGO staging of the tumor, and the association between this protein's expression and a poor disease-free survival and overall survival, together with the multivariate analysis results, all have suggested that this protein is an independent prognostic factor for disease-free survival and overall survival in the studied serous ovarian cancer patients [74]. Other research groups have performed slightly different approaches on various biospecimens. For instance, combining shotgun proteomics and SRM MS, Elschenbroich and colleagues have published the results of in-depth proteomics analysis of ovarian cancer ascites as compared to ascites from benign ovarian tumors. They have designed an analysis pipeline that included discovery-based proteomics, bioinformatics, and targeted proteomics quantification of the detected cancer biomarkers candidates [75]. Combined 2-DE and MS/MS analysis has been used to study the ovarian cancer tissue, interstitial fluid, and peritoneal effusion, as compared to normal tissue and fluid, in specimens obtained surgically [76]. This comparative analysis has revealed differential expression of six proteins that are involved in cell cycle progression and apoptosis, as well as in signal transduction pathways. One of those proteins, the calgranulin, was reported to be significantly overexpressed in all pathological samples and to represent a potential diagnostic and/or prognostic biomarker. Other studies have reported changes in N-linked glycan structures and their expression as diagnostic signature in ovarian cancer patients [50]. A shotgun quantitative proteomic evaluation of benign and malignant epithelial ovarian tumors as compared to normal tissue, using iTRAQ technology with LC-MALDI-TOF/TOF and LC-ESI-QTOF MS/MS was published two years ago. The PI3K/Akt signaling pathway was reported as a significant pathway capable of discriminating the clinicopathologically different tissues studied [77]. More recently, MS analysis of the secretome from *ex vivo* coculturing of ovarian cancer cells and peritoneal cells to detect proteomic markers of their interactions was suggested to reflect the metastasizing nature of ovarian cancers. A protein, Mucin 5AC was suggested as a potential biomarker for the invasiveness of ovarian cancers since its expression was significantly elevated in the ovarian-peritoneal cells coculture as compared to monoculture of each type of cells [78]. Furthermore, overexpression of class III β-tubulin within the ovarian tumor microenvironment was recently demonstrated to have prognostic power predicting poor overall survival in patients treated with neoadjuvant chemotherapy [79].

(4) Targets for therapeutic means: A rare histologic subset of ovarian cancer, clear cell ovarian cancer, is known to have low survival relative to other types of ovarian cancers. Genomics and immunohistochemical studies have demonstrated similar gene and protein expression profiles to clear cell cancers in other organs, specifically the kidney and uterus. Therefore, it might be recommended to consider therapeutic

approach of this serious cancer histotype based on the protein expression profile, rather than on the organ affected [80]. Few years ago, Anglesio and coworkers have demonstrated that women with clear cell ovarian cancer had shown a positive response to Sunitinib, a drug used with relatively successful outcome in patients with renal cancers [81]. Additional new perspectives for novel targets in ovarian cancer therapy are being examined utilizing data obtained from various high-throughput technologies [82].

6. Can Proteomics Research Findings in Cancer Be Translated into Clinically Oriented Research?

As already mentioned, massive applications of recent -omics technologies in cancer research have started since the last century and have been constantly evolving so far. These have been translated into genomics and proteomics cancer signatures. The translation of biomarker discoveries into potential anticancer agents is highly dependent on the quality of data generated, which is influenced by several factors as mentioned above [83]. Wilhelm and colleagues have recently reported an MS-based draft of human proteome. Among their findings of human proteome expression, they confirmed high levels of expression of functional proteins in relation to specific cancer. For instance, the protooncogene EGFR, which was discovered in the eighties of the last century [84], was recently found to be highly expressed in a confined manner to certain cancerous tissue as in breast cancer. Beta-catenin, a member of the Wnt signaling pathway, was also highly expressed in colon cancer cells, where it participated in the development of the malignancy [4]. These findings and others represent a rich source of information and a platform, based on which researchers can design projects aiming at discovering novel anticancer agents. The following section summarizes information from 2 research groups working on example of such agents, the EGFR kinase inhibitors and the heat shock protein 90 (Hsp90) inhibitors.

6.1. EGFR Kinase Inhibitors. Studying the cellular mechanisms of cancer in general and of drug action in particular has been a hot area in proteomic cancer research. This area holds a promising outcome of clinical significance and hence the hope of moving cancer proteomics from bench to bed side [24]. Using cancer cell line panel developed by the National Cancer Institute (NCI) as a model system for different tissue types and genetic diversity of human cancers, and analyzing the massive amount of information obtained by bioinformatics, Moghaddas and coworkers have shown a strong cell line clusters based on tissue type. Hundreds of differentially expressed proteins were demonstrated in this model system, which are potential biomarkers for different tumor properties. Moreover, by integrating their proteomic data to the publicly accessed transcriptomic data for this model system, the authors have shown consistency between mRNA and protein expression. They were also capable of demonstrating that protein expression can be correlated to many FDA-approved anticancer drug response, both drug sensitivity and resistance [85]. Of special importance as anticancer drug targets are various families of cellular protein kinases. Kinases represent important oncogene classes and are key players in intracellular signaling; subsequently their differential expression and/or functional dysregulation can be a cause or consequence of tumorigenesis. Therefore, not surprisingly, kinases are important anticancer therapeutic targets [86, 87]. The EGFR kinase inhibitors erlotinib and lapatinib have been used in cancer therapy. Recently, proteomics approach in cancer cell lines using elastic net analysis has been utilized for the identification of markers for drug sensitivity (positive-effect-size) or resistance (negative-effect-size) [4].

6.2. HSP90 Inhibitors. Hsp90 is a molecular chaperone that is essential for the correct folding, stability, and hence functions of many proteins. As such, it is part of a system that functions in both physiological and pathological states [88]. Cancer cells are considered chaperone addict, since they have special requirement for the protein folding machinery components to deal with the surplus of proteins being synthesized. The significance of targeting Hsp90 in cancer therapy lies in the nature of its clients, since many of them belong to the family of oncogenes, including tyrosine kinases, transcription factors, and cell cycle regulatory proteins. Therefore, inhibiting Hsp90 leads to degradation of such proteins through the proteasome machinery. The use of Hsp90 inhibitors in treating cancer has been promising in certain solid tumors as well as in hematological malignancies. This has been recently reviewed by Garcia-Carbonero and coworkers [89].

7. Conclusion and Perspectives

Proteomics approach in studying many diseases including cancer is producing data that complements those produced by other high-throughput technologies. Such technologies should aim beyond the mere generation of lists of differentially expressed macromolecules and their derivatives, as a cause or consequence of the studied pathology. For instance, careful interpretation of proteomics data has shed some light on the underlying mechanisms leading to cancer formation. Examples discussed in the present review include the association between aneuploidy, proteotoxic stress, abnormal cellular proliferation, and tumorigenesis; the defective proteins' structure and hence function secondary to gene mutations; and the consequent aberrant networks interactions of abnormal protein repertoire in cancer states. Nevertheless, the field is faced with numerous biological and technical challenges as a result of the concepts of cancer heterogeneity, samples variables, and poor study designs. These challenges can be minimized by proper study designs, implementing strict protocols paying attention to every step in the process, establishing robust validation assays, and exploring innovative tools or even combinations of tools. Besides the traditional proteomics techniques that are constantly being advanced, more recent approaches combining proteomics with other technologies such as imaging are unraveling

the complexity of the proteomics changes in cancer and are producing data that are thought to be more representing to the *in vivo* situation and tumor environment. Examples of three of the most studied cancers, lung, breast, and ovarian cancers, have been discussed illustrating various perspectives in approaching the subject of cancer biomarkers, the need to standardize and optimize study design, preanalytical and analytical assays components, and strict validation strategies. Overall, common objectives for proteomics studies in cancer are to better understand tumor biology, to facilitate the development of biomarkers and, most importantly, to move towards bedside applications in cancer management. Refining the huge amount of information obtained from proteomics and related technologies is required to enable transition to clinical validation, which is an ultimate goal for many proteomics-centered studies.

Conflict of Interests

The author declares no conflict of interests.

References

[1] L. Ding, M. C. Wendl, J. F. McMichael, and B. J. Raphael, "Expanding the computational toolbox for mining cancer genomes," *Nature Reviews Genetics*, vol. 15, no. 8, pp. 556–570, 2014.

[2] M. S. Lawrence, P. Stojanov, C. H. Mermel et al., "Discovery and saturation analysis of cancer genes across 21 tumour types," *Nature*, vol. 505, no. 7484, pp. 495–501, 2014.

[3] M. W. Duncan, "Good mass spectrometry and its place in good science," *Journal of Mass Spectrometry*, vol. 47, no. 6, pp. 795–809, 2012.

[4] M. Wilhelm, J. Schlegl, H. Hahne et al., "Mass-spectrometry-based draft of the human proteome," *Nature*, vol. 509, no. 7502, pp. 582–587, 2014.

[5] F. Meric-Bernstam, A. Akcakanat, H. Chen et al., "Influence of biospecimen variables on proteomic biomarkers in breast cancer," *Clinical Cancer Research*, vol. 20, no. 14, pp. 3870–3883, 2014.

[6] G. Poste, "Bring on the biomarkers," *Nature*, vol. 469, no. 7329, pp. 156–157, 2011.

[7] N. Donnelly and Z. Storchová, "Dynamic karyotype, dynamic proteome: buffering the effects of aneuploidy," *Biochimica et Biophysica Acta—Molecular Cell Research*, vol. 1843, no. 2, pp. 473–481, 2014.

[8] W. E. Balch, R. I. Morimoto, A. Dillin, and J. W. Kelly, "Adapting proteostasis for disease intervention," *Science*, vol. 319, no. 5865, pp. 916–919, 2008.

[9] J. Adams, "The proteasome: a suitable antineoplastic target," *Nature Reviews Cancer*, vol. 4, no. 5, pp. 349–360, 2004.

[10] R. Mosca, A. Céol, and P. Aloy, "Interactome3D: adding structural details to protein networks," *Nature Methods*, vol. 10, no. 1, pp. 47–53, 2013.

[11] S. Gulati, T. M. K. Cheng, and P. A. Bates, "Cancer networks and beyond: interpreting mutations using the human interactome and protein structure," *Seminars in Cancer Biology*, vol. 23, no. 4, pp. 219–226, 2013.

[12] O. J. R. Gustafsson, J. S. Eddes, S. Meding, S. R. McColl, M. K. Oehler, and P. Hoffmann, "Matrix-assisted laser desorption/ionization imaging protocol for in situ characterization of tryptic peptide identity and distribution in formalin-fixed tissue," *Rapid Communications in Mass Spectrometry*, vol. 27, no. 6, pp. 655–670, 2013.

[13] S. Meding, K. Martin, O. J. R. Gustafsson et al., "Tryptic peptide reference data sets for MALDI imaging mass spectrometry on formalin-fixed ovarian cancer tissues," *Journal of Proteome Research*, vol. 12, no. 1, pp. 308–315, 2013.

[14] M. Shipitsin, C. Small, S. Choudhury et al., "Identification of proteomic biomarkers predicting prostate cancer aggressiveness and lethality despite biopsy-sampling error," *British Journal of Cancer*, vol. 111, no. 6, pp. 1201–1212, 2014.

[15] S. S. Hori and S. S. Gambhir, "Mathematical model identifies blood biomarker-based early cancer detection strategies and limitations," *Science Translational Medicine*, vol. 3, no. 109, pp. 109–116, 2011.

[16] D. Konforte and E. P. Diamandis, "Is early detection of cancer with circulating biomarkers feasible?" *Clinical Chemistry*, vol. 59, no. 1, pp. 35–37, 2013.

[17] B. M. Nolen and A. E. Lokshin, "Protein biomarkers of ovarian cancer: the forest and the trees," *Future Oncology*, vol. 8, no. 1, pp. 55–71, 2012.

[18] M. E. H. Hammond and S. E. Taube, "Issues and barriers to development of clinically useful tumor markers: a development pathway proposal," *Seminars in Oncology*, vol. 29, no. 3, pp. 213–221, 2002.

[19] 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. 1054–1061, 2001.

[20] M. S. Pepe, Z. Feng, H. Janes, P. M. Bossuyt, and J. D. Potter, "Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design," *Journal of the National Cancer Institute*, vol. 100, no. 20, pp. 1432–1438, 2008.

[21] M. A. Baldwin, "Protein identification by mass spectrometry: issues to be considered," *Molecular and Cellular Proteomics*, vol. 3, no. 1, pp. 1–9, 2004.

[22] S. A. Carr, S. E. Abbatiello, B. L. Ackermann et al., "Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach," *Molecular & Cellular Proteomics*, vol. 13, no. 3, pp. 907–917, 2014.

[23] L. G. Longsworth, T. Shedlovsky, and D. A. Macinnes, "Electrophoretic patterns of normal and pathological human blood serum and plasma," *Journal of Experimental Medicine*, vol. 70, no. 4, pp. 399–413, 1939.

[24] S. Hanash and A. Taguchi, "The grand challenge to decipher the cancer proteome," *Nature Reviews Cancer*, vol. 10, no. 9, pp. 652–660, 2010.

[25] S. M. Hanash, L. J. Baier, L. McCurry, and S. A. Schwartz, "Lineage-related polypeptide markers in acute lymphoblastic leukemia detected by two-dimensional gel electrophoresis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 3, pp. 807–811, 1986.

[26] J. M. A. Moreira, G. Ohlsson, P. Gromov et al., "Bladder cancer-associated protein, a potential prognostic biomarker in human bladder cancer," *Molecular and Cellular Proteomics*, vol. 9, no. 1, pp. 161–177, 2010.

[27] L. A. Liotta, M. Ferrari, and E. Petricoin, "Clinical proteomics: written in blood," *Nature*, vol. 425, no. 6961, p. 905, 2003.

[28] K. Fujii, T. Kondo, M. Yamada, K. Iwatsuki, and S. Hirohashi, "Toward a comprehensive quantitative proteome database: protein expression map of lymphoid neoplasms by 2-D DIGE and MS," *Proteomics*, vol. 6, no. 17, pp. 4856–4876, 2006.

[29] B. F. Cravatt, G. M. Simon, and J. R. Yates III, "The biological impact of mass-spectrometry-based proteomics," *Nature*, vol. 450, no. 7172, pp. 991–1000, 2007.

[30] N. Siuti and N. L. Kelleher, "Decoding protein modifications using top-down mass spectrometry," *Nature Methods*, vol. 4, no. 10, pp. 817–821, 2007.

[31] A. I. Nesvizhskii, O. Vitek, and R. Aebersold, "Analysis and validation of proteomic data generated by tandem mass spectrometry," *Nature Methods*, vol. 4, no. 10, pp. 787–797, 2007.

[32] T. Bartke, J. Borgel, and P. A. DiMaggio, "Proteomics in epigenetics: new perspectives for cancer research," *Briefings in Functional Genomics*, vol. 12, no. 3, pp. 205–218, 2013.

[33] E. Agwa and P. C. Ma, "Overview of various techniques/platforms with critical evaluation of each," *Current Treatment Options in Oncology*, vol. 14, no. 4, pp. 623–633, 2013.

[34] 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.

[35] M. S. Kim, S. Kwon, T. Kim, E. S. Lee, and J.-K. Park, "Quantitative proteomic profiling of breast cancers using a multiplexed microfluidic platform for immunohistochemistry and immunocytochemistry," *Biomaterials*, vol. 32, no. 5, pp. 1396–1403, 2011.

[36] M. Shipitsin, C. Small, E. Giladi et al., "Automated quantitative multiplex immunofluorescence in situ imaging identifies phospho-S6 and phospho-PRAS40 as predictive protein biomarkers for prostate cancer lethality," *Proteome Science*, vol. 12, article 40, 2014.

[37] P. Oh, J. E. Testa, P. Borgstrom, H. Witkiewicz, Y. Li, and J. E. Schnitzer, "*In vivo* proteomic imaging analysis of caveolae reveals pumping system to penetrate solid tumors," *Nature Medicine*, vol. 20, no. 9, pp. 1062–1068, 2014.

[38] J. A. Ludwig and J. N. Weinstein, "Biomarkers in cancer staging, prognosis and treatment selection," *Nature Reviews Cancer*, vol. 5, no. 11, pp. 845–856, 2005.

[39] R. Schiess, B. Wollscheid, and R. Aebersold, "Targeted proteomic strategy for clinical biomarker discovery," *Molecular Oncology*, vol. 3, no. 1, pp. 33–44, 2009.

[40] J. R. Whiteaker, C. Lin, J. Kennedy et al., "A targeted proteomics-based pipeline for verification of biomarkers in plasma," *Nature Biotechnology*, vol. 29, no. 7, pp. 625–634, 2011.

[41] R. Huttenhain, M. Soste, N. Selevsek et al., "Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics," *Science Translational Medicine*, vol. 4, no. 142, p. 142ra94, 2012.

[42] A. Taguchi, K. Politi, S. J. Pitteri et al., "Lung cancer signatures in plasma based on proteome profiling of mouse tumor models," *Cancer Cell*, vol. 20, no. 3, pp. 289–299, 2011.

[43] T. Geiger, S. F. Madden, W. M. Gallagher, J. Cox, and M. Mann, "Proteomic portrait of human breast cancer progression identifies novel prognostic markers," *Cancer Research*, vol. 72, no. 9, pp. 2428–2439, 2012.

[44] S. J. Pitteri, K. S. Kelly-Spratt, K. E. Gurley et al., "Tumor microenvironment-derived proteins dominate the plasma proteome response during breast cancer induction and progression," *Cancer Research*, vol. 71, no. 15, pp. 5090–5100, 2011.

[45] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2013," *CA: Cancer Journal for Clinicians*, vol. 63, no. 1, pp. 11–30, 2013.

[46] F. Leung, E. P. Diamandis, and V. Kulasingam, "Ovarian cancer biomarkers: current state and future implications from high-throughput technologies," *Advances in Clinical Chemistry*, vol. 66, pp. 25–77, 2014.

[47] S. Vaughan, J. I. Coward, R. C. Bast et al., "Rethinking ovarian cancer: recommendations for improving outcomes," *Nature Reviews Cancer*, vol. 11, no. 10, pp. 719–725, 2011.

[48] Y. Mechref, Y. Hu, A. Garcia, and A. Hussein, "Identifying cancer biomarkers by mass spectrometry-based glycomics," *Electrophoresis*, vol. 33, no. 12, pp. 1755–1767, 2012.

[49] B. Adamczyk, T. Tharmalingam, and P. M. Rudd, "Glycans as cancer biomarkers," *Biochimica et Biophysica Acta*, vol. 1820, no. 9, pp. 1347–1353, 2012.

[50] W. R. Alley Jr., J. A. Vasseur, J. A. Goetz et al., "N-linked glycan structures and their expressions change in the blood sera of ovarian cancer patients," *Journal of Proteome Research*, vol. 11, no. 4, pp. 2282–2300, 2012.

[51] K. Kim, L. R. Ruhaak, U. T. Nguyen et al., "Evaluation of glycomic profiling as a diagnostic biomarker for epithelial ovarian cancer," *Cancer Epidemiology Biomarkers and Prevention*, vol. 23, no. 4, pp. 611–621, 2014.

[52] J. N. Arnold, R. Saldova, U. M. Abd Hamid, and P. M. Rudd, "Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation," *Proteomics*, vol. 8, no. 16, pp. 3284–3293, 2008.

[53] A. A. Alfadda, A. A. Turjoman, A. S. Moustafa et al., "A proteomic analysis of excreted and circulating proteins from obese patients following two different weight-loss strategies," *Experimental Biology and Medicine*, vol. 239, no. 5, pp. 568–580, 2014.

[54] T. Zhang, X. Wu, C. Ke et al., "Identification of potential biomarkers for ovarian cancer by urinary metabolomic profiling," *Journal of Proteome Research*, vol. 12, no. 1, pp. 505–512, 2013.

[55] J. Chen, L. Zhou, X. Zhang et al., "Urinary hydrophilic and hydrophobic metabolic profiling based on liquid chromatography-mass spectrometry methods: differential metabolite discovery specific to ovarian cancer," *Electrophoresis*, vol. 33, no. 22, pp. 3361–3369, 2012.

[56] C. Ke, Y. Hou, H. Zhang et al., "Large-scale profiling of metabolic dysregulation in ovarian cancer," *International Journal of Cancer*, vol. 136, no. 3, pp. 516–26, 2015.

[57] V. O. Shender, M. S. Pavlyukov, R. H. Ziganshin et al., "Proteome-metabolome profiling of ovarian cancer ascites reveals novel components involved in intercellular communication," *Molecular & Cellular Proteomics*, vol. 13, no. 12, pp. 3558–3571, 2014.

[58] C. R. Smith, I. Batruch, J. Bauça et al., "Deciphering the peptidome of urine from ovarian cancer patients and healthy controls," *Clinical Proteomics*, vol. 11, article 23, 2014.

[59] Z. Xu, C. Wu, F. Xie et al., "Comprehensive quantitative analysis of Ovarian and breast cancer tumor peptidomes," *Journal of Proteome Research*, vol. 14, no. 1, pp. 422–433, 2015.

[60] A. Bery, F. Leung, C. R. Smith, E. P. Diamandis, and V. Kulasingam, "Deciphering the ovarian cancer ascites fluid peptidome," *Clinical Proteomics*, vol. 11, no. 1, article 13, 2014.

[61] K. S. Anderson, D. W. Cramer, S. Sibani et al., "Autoantibody signature for the serologic detection of ovarian cancer," *Journal of Proteome Research*, vol. 14, no. 1, pp. 578–586, 2015.

[62] Y. S. Cho-Chung, "Autoantibody biomarkers in the detection of cancer," *Biochimica et Biophysica Acta—Molecular Basis of Disease*, vol. 1762, no. 6, pp. 587–591, 2006.

[63] L. Gortzak-Uzan, A. Ignatchenko, A. I. Evangelou et al., "A proteome resource of ovarian cancer ascites: integrated proteomic and bioinformatic analyses to identify putative biomarkers," *Journal of Proteome Research*, vol. 7, no. 1, pp. 339–351, 2008.

[64] E. R. Hoskins, B. L. Hood, M. Sun, T. C. Krivak, R. P. Edwards, and T. P. Conrads, "Proteomic analysis of ovarian cancer proximal fluids: validation of elevated peroxiredoxin 1 in patient peripheral circulation," *PLoS ONE*, vol. 6, no. 9, Article ID e25056, 2011.

[65] L. E. Moore, R. M. Pfeiffer, Z. Zhang, K. H. Lu, E. T. Fung, and R. C. Bast Jr., "Proteomic biomarkers in combination with CA 125 for detection of epithelial ovarian cancer using prediagnostic serum samples from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial," *Cancer*, vol. 118, no. 1, pp. 91–100, 2012.

[66] L. E. Moore, E. T. Fung, M. McGuire et al., "Evaluation of apolipoprotein A1 and posttranslationally modified forms of transthyretin as biomarkers for ovarian cancer detection in an independent study population," *Cancer Epidemiology Biomarkers and Prevention*, vol. 15, no. 9, pp. 1641–1646, 2006.

[67] E. F. Petricoin, A. M. Ardekani, B. A. Hitt et al., "Use of proteomic patterns in serum to identify ovarian cancer," *The Lancet*, vol. 359, no. 9306, pp. 572–277, 2002.

[68] R. Longuespée, C. Boyon, A. Desmons et al., "Ovarian cancer molecular pathology," *Cancer and Metastasis Reviews*, vol. 31, no. 3-4, pp. 713–732, 2012.

[69] A. Toss, E. de Matteis, E. Rossi et al., "Ovarian cancer: can proteomics give new insights for therapy and diagnosis?" *International Journal of Molecular Sciences*, vol. 14, no. 4, pp. 8271–8290, 2013.

[70] T. Walsh, S. Casadei, K. H. Coats et al., "Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer," *The Journal of the American Medical Association*, vol. 295, no. 12, pp. 1379–1388, 2006.

[71] C. Fuseya, A. Horiuchi, A. Hayashi et al., "Involvement of pelvic inflammation-related mismatch repair abnormalities and microsatellite instability in the malignant transformation of ovarian endometriosis," *Human Pathology*, vol. 43, no. 11, pp. 1964–1972, 2012.

[72] N. J. Engel, P. Gordon, D. L. Thull et al., "A multidisciplinary clinic for individualizing management of patients at increased risk for breast and gynecologic cancer," *Familial Cancer*, vol. 11, no. 3, pp. 419–427, 2012.

[73] B. Rungruang, B. L. Hood, M. Sun, E. Hoskins, T. P. Conrads, and K. K. Zorn, "Novel surgical approaches for sampling the ovarian surface epithelium and proximal fluid proteome," *Journal of Proteome Research*, vol. 9, no. 11, pp. 6071–6076, 2010.

[74] Y. L. Li, F. Ye, X. D. Cheng et al., "Identification of glia maturation factor beta as an independent prognostic predictor for serous ovarian cancer," *European Journal of Cancer*, vol. 46, no. 11, pp. 2104–2118, 2010.

[75] S. Elschenbroich, V. Ignatchenko, B. Clarke et al., "In-depth proteomics of ovarian cancer ascites: combining shotgun proteomics and selected reaction monitoring mass spectrometry," *Journal of Proteome Research*, vol. 10, no. 5, pp. 2286–2299, 2011.

[76] L. Cortesi, E. Rossi, L. D. Casa et al., "Protein expression patterns associated with advanced stage ovarian cancer," *Electrophoresis*, vol. 32, no. 15, pp. 1992–2003, 2011.

[77] S. Waldemarson, M. Krogh, A. Alaiya et al., "Protein expression changes in ovarian cancer during the transition from benign to malignant," *Journal of Proteome Research*, vol. 11, no. 5, pp. 2876–2889, 2012.

[78] N. Musrap, G. S. Karagiannis, P. Saraon, I. Batruch, C. Smith, and E. P. Diamandis, "Proteomic analysis of cancer and mesothelial cells reveals an increase in Mucin 5AC during ovarian cancer and peritoneal interaction," *Journal of Proteomics*, vol. 103, pp. 204–215, 2014.

[79] D. M. Roque, N. Buza, M. Glasgow et al., "Class III beta-tubulin overexpression within the tumor microenvironment is a prognostic biomarker for poor overall survival in ovarian cancer patients treated with neoadjuvant carboplatin/paclitaxel," *Clinical and Experimental Metastasis*, vol. 31, no. 1, pp. 101–110, 2014.

[80] K. K. Zorn, T. Bonome, L. Gangi et al., "Gene expression profiles of serous, endometrioid, and clear cell subtypes of ovarian and endometrial cancer," *Clinical Cancer Research*, vol. 11, no. 18, pp. 6422–6430, 2005.

[81] M. S. Anglesio, J. George, H. Kulbe et al., "IL6-STAT3-HIF signaling and therapeutic response to the angiogenesis inhibitor sunitinib in ovarian clear cell cancer," *Clinical Cancer Research*, vol. 17, no. 8, pp. 2538–2548, 2011.

[82] D. S. P. Tan, R. E. Miller, and S. B. Kaye, "New perspectives on molecular targeted therapy in ovarian clear cell carcinoma," *British Journal of Cancer*, vol. 108, no. 8, pp. 1553–1559, 2013.

[83] J. W. Lee, D. Figeys, and J. Vasilescu, "Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers," *Advances in Cancer Research*, vol. 96, pp. 269–298, 2007.

[84] S. Cohen, R. A. Fava, and S. T. Sawyer, "Purification and characterization of epidermal growth factor receptor/protein kinase from normal mouse liver," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 79, no. 20, pp. 6237–6241, 1982.

[85] A. M. Gholami, H. Hahne, Z. Wu et al., "Global proteome analysis of the NCI-60 cell line panel," *Cell Reports*, vol. 4, no. 3, pp. 609–620, 2013.

[86] P. Cohen, "Protein kinases—the major drug targets of the twenty-first century?" *Nature Reviews Drug Discovery*, vol. 1, no. 4, pp. 309–315, 2002.

[87] S. Knapp, P. Arruda, J. Blagg et al., "A public-private partnership to unlock the untargeted kinome," *Nature Chemical Biology*, vol. 9, no. 1, pp. 3–7, 2013.

[88] H. Wiech, J. Buchner, R. Zimmermann, and U. Jakob, "Hsp90 chaperones protein folding in vitro," *Nature*, vol. 358, no. 6382, pp. 169–170, 1992.

[89] R. Garcia-Carbonero, A. Carnero, and L. Paz-Ares, "Inhibition of HSP90 molecular chaperones: moving into the clinic," *The Lancet Oncology*, vol. 14, no. 9, pp. e358–e369, 2013.

Hindawi Publishing Corporation
Disease Markers
Volume 2015, Article ID 824304, 8 pages
http://dx.doi.org/10.1155/2015/824304

Research Article

Polymorphisms in *C-Reactive Protein* and *Glypican-5* Are Associated with Lung Cancer Risk and *Gartrokine-1* Influences Cisplatin-Based Chemotherapy Response in a Chinese Han Population

Shuo Zhang,[1] Asmitananda Thakur,[1] Yiqian Liang,[1] Ting Wang,[1] Lei Gao,[1] Tian Yang,[1] Yang Li,[1] Tingting Geng,[2] Tianbo Jin,[2] Tianjun Chen,[1] Johnson J. Liu,[3] and Mingwei Chen[1]

[1]*Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of School of Medicine of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, China*
[2]*National Engineering Research Center for Miniaturized Detection Systems, School of Life Sciences, Northwest University, Xi'an 710069, China*
[3]*School of Pharmacy, Faculty of Health Science, University of Tasmania, Hobart, TAS 7001, Australia*

Correspondence should be addressed to Mingwei Chen; chenmingwei64@hotmail.com

Received 21 August 2014; Revised 24 December 2014; Accepted 15 January 2015

Academic Editor: Varodom Charoensawan

The role of genetics in progression of cancer is an established fact, and susceptibility risk and difference in outcome to chemotherapy may be caused by the variation in low-penetrance alleles of risk genes. We selected seven genes (*CRP, GPC5, ACTA2, AGPHD1, SEC14L5, RBMS3*, and *GKN1*) that previously reported link to lung cancer (LC) and genotyped single nucleotide polymorphisms (SNPs) of these genes in a case-control study. A protective allele "C" was found in rs2808630 of the C-reactive protein (*CRP*). Model association analysis found genotypes "T/C" and "C/C" in the dominant model and genotype "T/C" in the overdominant model of rs2808630 associated with reduced LC risk. Gender-specific analysis in each model showed that genotypes "T/T" and "C/C" in rs2352028 of the *Glypican 5* (*GPC5*) were associated with increased LC risk in males. Logistic regression analysis showed "C/T" genotype carriers of rs4254535 in the *Gastrokine 1* (*GKN1*) had less likelihood to have chemotherapy response. Our results suggest a potential association between *CRP* and *GPC5* variants with LC risk; variation in *GKN1* is associated with chemotherapy response in the Chinese Han population.

1. Introduction

Lung cancer is the most common malignancy in the world and is reported to have an increasing incidence in developing countries [1, 2]. According to the global cancer statistics, in 2008 approximately 1.6 million people were diagnosed with lung cancer, and there were 1.4 million deaths [3]. Tobacco smoke, environmental pollution, occupational exposures, and preexisting lung disease increase the risk of lung cancer. However, patients have been diagnosed with lung cancer in the absence of these risk factors [4–6]. Genetic susceptibility to lung cancer independent of established risk factors has not yet been clearly defined.

Despite considerable advances in the field of tumor biology, the majority of patients with lung cancer are diagnosed at an already advanced stage and thus surgical resection is not a feasible treatment option. Platinum-based doublet chemotherapy is the current standard of therapy in this situation. However, the response to chemotherapy among lung cancer patients has significant variation. We hypothesize that the susceptibility risk and variation in outcome to chemotherapy may be caused by the variation in low-penetrance alleles.

In this study, we selected single nucleotide polymorphisms (SNPs) from seven different genes *(CRP (C-reactive protein), GPC5 (Glypican 5), ACTA2 (actin, alpha 2, smooth muscle, aorta), AGPHD1 (aminoglycoside phosphotransferase*

domain containing 1), SEC14L5 (SEC14-like 5), RBMS3 (RNA binding motif, single stranded interacting protein 3), and GKN1 (Gastrokine 1)) that have been linked to lung cancer [7–13]. We analyzed each tag single nucleotide polymorphism (tSNP) for lung cancer risk in a case-control study involving Chinese population. Multivariate logistic regression analysis was used to test the association between gene polymorphisms and chemotherapy response.

2. Materials and Methods

2.1. Study Participants. A case-control study involving the Chinese study population of 309 lung cancer patients and 310 controls was conducted at the First Affiliated Hospital of Xi'an Jiaotong University. All included patients had recently diagnosed and histopathologically confirmed primary lung cancer. The control subjects were recruited from the health check-up center of the First Affiliated Hospital of Xi'an Jiaotong University, which they had visited for an annual health examination. Patients were ascertained to be free from any acute or chronic pathology. Their cancer-free status was reconfirmed by testing for plasma levels of carcinoembryonic antigen and alpha-fetoprotein. Blood samples from the patients were collected before initiation of chemotherapy or radiotherapy. Demographic and related clinical data of the study population was collected by a face-to-face questionnaire and medical case record. Patients were categorized as smokers or nonsmokers. The smokers were defined as those who smoked one cigarette/pipe per day for twelve months or longer at any time in their life. All of the participants were genetically unrelated ethnic Han Chinese from Shaanxi Province and provided written informed consent for their participation in the present study. The protocols for this study were conducted according to the Declaration of Helsinki and were approved by the Institutional Review Boards of both the First Affiliated Hospital of Xi'an Jiaotong University and Northwest University.

Five milliliters of whole blood were collected from each subject into tubes containing ethylenediaminetetraacetic acid (EDTA) at the time of initial diagnosis. After centrifugation, the samples were stored at −80°C until further use.

2.2. Evaluation of Cisplatin-Based Chemotherapeutic Response. There are all together 113 lung cancer patients who received cisplatin based first-line chemotherapy and satisfied the following criteria: Eastern Cooperative Oncology Group (ECOG) performance status (PS) ≤ 1, age > 18 years, and adequate bone marrow reserve, as well as satisfactory liver and renal function. These patients were in clinical stage III or IV and had a measurable lesion on computed tomography scan at the beginning of treatment. The patients received chemotherapy every 3 weeks, for a maximum of six cycles or until disease progression or unacceptable toxicity occurred. Response to treatment was determined according to the Response Evaluation Criteria in Solid Tumor Group (RECIST) guidelines after two cycles of chemotherapy and every two cycles thereafter [14]. For data analysis, patients achieving complete response (CR) or partial response (PR) were considered "responders," and patients with stable disease (SD) or progressive disease (PD) were defined as "nonresponders" [15]. Multivariate logistic regression analysis was used to test the association between gene polymorphisms and chemotherapy response.

2.3. tSNP Selection and Genotyping. All seven tSNPs in the selected genes were associated with lung cancer and with minor allele frequencies (MAF) greater than 5% in the HapMap CHB (Chinese Han Beijing) population. DNA was extracted from whole blood by GoldMag-Mini Whole Blood Genomic DNA Purification Kit (GoldMag Co., Ltd., Xi'an City, China). The concentration was measured by NanoDrop 2000 (Thermo Scientific, Waltham, Massachusetts, USA). The design of primers, SNP genotyping, and data processing were performed by Sequenom MassARRAY platform Software (Sequenom Co., Ltd., San Diego, California, USA) [16, 17].

2.4. Statistical Analysis. Statistical analysis was undertaken using statistical software (SPSS 16.0; Chicago, IL) and Microsoft Excel. A two-sided P value < 0.05 was considered the threshold for statistical significance. Hardy-Weinberg equilibrium (HWE) of each tSNP in control group was tested by Fisher's exact test. The differences in allelic frequencies between case and control groups were compared via the Chi-squared test [18]. Associations between genotypes and lung cancer risk were tested in different genetic models (codominant, dominant, recessive, overdominant, and log-additive) by SNPStats website software http://bioinfo.iconcologia.net/snpstats/start.htm [19]. Testing of odds ratios (ORs) with 95% confidence intervals (CIs) was performed by unconditional logistic regression analysis with adjustment for gender and age [20]. Akaike's Information Criterion and Bayesian Information Criterion were applied to estimate the best-fit model for each SNP. Association between genotypes and lung cancer risk was determined by SNPStats for gender-specific populations under each model [19].

3. Results

We recruited 309 patients (74 females and 235 males, mean age at diagnosis 58 years, range 25–85, SD ± 10) and 310 healthy (113 females and 197 males, mean age at diagnosis 50 years, range 29–75, SD ± 8) individuals into our study (Table 1). The genotype profiles of our study patients are shown in Supplementary Table S1 in the Supplementary Material available online at http://dx.doi.org/10.1155/2015/824304. The SNPs and primers used in the multiplexed SNP MassEXTENED assay are presented in Table 2. None of the tSNPs that we evaluated among the control group deviated from HWE (Table 3). We hypothesized that the minor allele of each SNP was a risk factor compared with the wild-type allele.

A significant protective allele "C" was found in rs2808630 of the *CRP* gene based on the crude P value of 0.05 (OR = 0.66; 95% CI, 0.48–0.91; $P = 0.01$) by Chi-square test (Table 3). Various genetic models were applied to calculate genetic risk. Reduced risk for lung cancer was associated with

Table 1: Characteristics of patients and controls.

Characteristics	Lung cancer ($n = 309$)		Control ($n = 310$)	
Age (means ± SD, year)	58.2 ± 10.2		50.3 ± 8.1	
Sex				
Male	235	76.1	197	63.5
Female	74	23.9	113	36.5
Smoking status				
Never	94	30.4	188	60.6
Ever	215	69.6	122	39.4
	No.	%		
Histology				
Adenocarcinoma	110	35.6		
Squamous cell carcinoma	116	37.5		
Small-cell carcinoma	66	21.3		
Large-cell carcinoma	2	0.6		
Unspecified lung cancer	15	5.0		
Stage				
I	67	21.7		
II	52	16.8		
III	69	22.3		
IV	118	38.2		
Data uncertain	3	1.0		

the genotypes "T/C" and "C/C" in rs2808630 (OR = 0.66, 95% CI, 0.44–0.98; $P = 0.036$) in the dominant model and the genotype "T/C" (OR = 0.65, 95% CI, 0.43–0.98; $P = 0.037$) in the overdominant model (Table 4). Each tSNP was analyzed in a gender-specific population under each model. We found that the genotypes "T/T" and "C/C" in rs2352028 of the *GPC5* gene were associated with increased lung cancer risk in males in the overdominant model (Table 5). rs2808630 in *CRP* and rs2352028 in *GPC5* were both associated with lung cancer risk.

"C/T" genotype distribution in the rs4254535 of the *GKN1* gene was significantly higher in nonresponders than in responders (34.62% versus 14.29%, $P = 0.029$) (Table 6). Logistic regression analysis showed that "C/T" genotype carriers had poor response for chemotherapy as compared to "T/T" genotype carriers (OR 3.287, 95% CI, 1.135–9.522; $P = 0.029$) after adjustment for age, gender, smoking status, histology, stage, and chemotherapy regimens.

However, as shown in Tables 3, 4, and 6, the significance levels were attenuated after applying a strict Bonferroni correction, indicating a likely association between positive tSNPs and risk of lung cancer and chemotherapy response.

4. Discussion

In this case-control study, we selected tSNPs with MAF greater than 5% in the HapMap CHB population to ensure that the statistical power was sufficient for data analysis. Our results firstly suggest that polymorphisms in *CRP* and *GPC5* genes have an association with susceptibility risk of lung cancer in the Chinese Han population. The multivariate logistic regression analysis shows that polymorphism in *GKN1* influences chemotherapy response.

The *CRP* gene, located in 1q23.2, encodes CRP protein which has several host defense-related functions, including recognition and elimination of foreign pathogens and damaged host cell. CRP is an acute-phase protein that increases during the host response to tissue injuries, including infection, trauma, surgery, myocardial infarct, and cancer [8, 21]. There are three potential mechanisms linking CRP to cancers. One is that tumor growth promotes tissue inflammation and increases the level of CRP. Another possibility is that cancer cells increase production of inflammatory proteins, which leads to high CRP levels in cancer patients. Besides, CRP may promote tumor growth in chronic inflammation [22]. Elevated CRP levels are associated with poor prognosis of lung, hepatic, renal, colorectal, and ovarian cancers [23–29].

Our study found that rs2808630, an intronic SNP within the *CRP* gene, was significantly linked with lung cancer risk in both allelic and genotypic association analysis of a Chinese population. We also ascertained a significant allele "C" and genotypes "T/C" and "C/C" in rs2808630 in the dominant model and genotype "T/C" in the overdominant model that is protective against lung cancer development. We hypothesize that rs2808630 variant of the *CRP* gene could have decreased the level of CRP or reduced the activity of CRP in the presence of allele "C". A recent study by Xu et al. [30] found that 5 SNPs in the *CRP* gene (including rs2808630) were uncorrelated with lung cancer risk. They recruited 96 lung cancer patients and 124 controls of different races. This disparity in findings could be attributed to the small sample size and racial or regional differences in study populations. To our knowledge, our study is the first genotype/allele-based study that describes the association between SNPs within the *CRP* locus and lung cancer risk in a Chinese population.

The *GPC5* gene is a member of the glypican gene family and has eight exons encoding 572 amino acids in a large genomic region (1.47 Mb) of chromosome 13q31.3. Reduction of GPC5 protein is linked to lung cancer [7]. A previous study involving American population reported an association (OR = 1.46, 95% CI 1.26–1.70, $P = 5.94 \times 10^{-6}$) between the single nucleotide polymorphism rs2352028 and lung cancer risk in never smokers [31] but failed to replicate in Caucasian [32] and Chinese [33] populations, indicating that the sensitivity and specificity of rs2352028 in terms of smoking status may not be similar in between races. Our study observed the variation between gender and found that genotypes "T/T" and "C/C" in rs2352028 of the *GPC5* gene are associated with increased lung cancer risk in males (under the overdominant model, after adjusting for age).

The *GKN1* gene is located in 2p13.3 and has a protective function on gastric antral mucosa by facilitating restoration and proliferation after injury. As it is expressed in normal gastric tissue but absent in gastric cancer tissues, GKN1 protein is treated as a potential biomarker for gastric cancer [34]. It is also found downregulated in placental tissue and cell [35]. Although current research focuses on the potential clinical use of GKN1 in the treatment of tumor, little is

Table 2: PCR primers.

SNP_ID	Forward primer	Reverse primer	UEP_SEQ
rs2808630	ACGTTGGATGGGGATGTAGGTTGAGCTAAT	ACGTTGGATGTAAGGCCAGAGGCTGTCTAC	tTTGCTTGCATCTTACTATAC
rs1926203	ACGTTGGATGAATCCACGTTACCTAAGCCC	ACGTTGGATGGGCTCTGATACCTGATTTGG	cgtgACCTAAGCCCAAATTATTAC
rs2352028	ACGTTGGATGATGACCCTGAACAGTAGTGG	ACGTTGGATGTGAAGGGTTTAACATGAAT	AGGGAAAGTCCATCTTTT
rs8034191	ACGTTGGATGCCACAAGTCCCCTTAGTTAC	ACGTTGGATGAGTGGTTAGAGCCCAATGTG	tTGTCAGGGCCTTTCT
rs9635542	ACGTTGGATGGAAGGTTGGTGGAAATTGCG	ACGTTGGATGTACAAACATGTACCCGGGTC	ttggcAATTGCGTGAGGAAAAG
rs4254535	ACGTTGGATGGAGACTGAAATAGAGTCTGC	ACGTTGGATGGATAGTTAGGACTCAACTGG	agAGAGTCTGCATGAAGGGAC
rs1530057	ACGTTGGATGTTCCATGAACAAAATGGAC	ACGTTGGATGTCAACATTATGGGCCACTCC	ggtggACAAAATGGACATGAACATGCAG

UEP_SEQ: unextended minisequencing primer.

Table 3: Candidate tSNPs.

SNP ID	Gene name	Chromosome position	Position	Allele	Minor allele	MAF (case)	MAF (control)	P value for HWE test	ORs	95% CI	P value from x^2	P value adj.*
rs2808630	*CRP*	1q23.2	159680868	C/T	C	0.135	0.191	0.982	0.66	0.48 0.91	0.010	0.070
rs1926203	*ACTA2*	10q23.31	90727334	G/T	T	0.167	0.150	0.999	1.13	0.83 1.53	0.436	1
rs2352028	*GPC5*	13q31.3	92445229	C/T	T	0.198	0.225	0.319	0.85	0.65 1.12	0.248	1
rs8034191	*AGPHD1*	15q25.1	78806023	C/T	C	0.034	0.032	0.841	1.05	0.56 1.96	0.874	1
rs9635542	*SEC14L5*	16p13.3	5001380	A/G	G	0.463	0.437	0.897	1.11	0.89 1.39	0.360	1
rs4254535	*GKN1*	2p13.3	69198388	C/T	C	0.204	0.217	0.317	0.93	0.70 1.22	0.577	1
rs1530057	*RBMS3*	3p24.1	29575463	A/C	A	0.065	0.078	0.788	0.82	0.53 1.27	0.368	1

* P value was adjusted by Bonferroni correction.

Table 4: Relationship between rs2808630 of *CRP* and lung cancer risk (adjusted by gender and age).

Model	Genotype	Control (N, %)	Case (N, %)	OR (95% CI)	P value	P value adj.*	AIC	BIC
Codominant	T/T	189 (65.4%)	218 (75.4%)	1.00	0.100	0.500	710.2	732.0
	T/C	90 (31.1%)	64 (22.1%)	0.64 (0.43–0.97)				
	C/C	10 (3.5%)	7 (2.4%)	0.80 (0.28–2.30)				
Dominant	T/T	189 (65.4%)	218 (75.4%)	1.00	0.036	0.180	708.3	725.8
	T/C-C/C	100 (34.6%)	71 (24.6%)	0.66 (0.44–0.98)				
Recessive	T/T-T/C	279 (96.5%)	282 (97.6%)	1.00	0.850	1	712.7	730.1
	C/C	10 (3.5%)	7 (2.4%)	0.90 (0.32–2.58)				
Overdominant	T/T-C/C	199 (68.9%)	225 (77.8%)	1.00	0.037	0.185	708.3	725.8
	T/C	90 (31.1%)	64 (22.1%)	0.65 (0.43–0.98)				
Log-additive				0.72 (0.51–1.02)	0.061	0.305	709.2	726.6

AIC: Akaike's information criterion; BIC: Bayesian information criterion.
* P value was adjusted by Bonferroni correction.

Table 5: rs2352028 of *GPC5* and gender cross-classification interaction.

Genotype	Female			Male			P value
	Control	Case	OR (95% CI)	Control	Case	OR (95% CI)	
C/C-T/T	82	47	1.00	110	163	1.80 (1.12–2.88)	0.019
C/T	31	27	1.42 (0.72–2.80)	87	71	0.98 (0.58–1.64)	

(N = 618, adjusted by age) under over-dominant model.

known about its expression and function in other organ systems or the significance of *GKN1* polymorphisms in cancer. Our study firstly reports that polymorphismin *GKN1* is influence cisplatin based chemotherapy response in lung cancer patients. The SNPs from the other four genes *(ACTA2, AGPHD1, SEC14L5,* and *RBMS3)* included in this study did not reach any statistically significant association with lung cancer risks or cisplatin based chemotherapy response in our study population.

There are certain intrinsic limitations in our study and must be noted. The sample size was not as large as some other lung cancer association studies. We performed Bonferroni correction in our statistical analysis and found no statistical significant associations between *CRP* and *GPC5* SNPs and lung cancer risk, neither in *GKN1* polymorphisms nor in response to cisplatin-based chemotherapy, which could be attributed to the relatively small sample size that may not satisfy all the seven independent hypotheses at the same time. Adjustments for multiple tests, like Bonferroni correction, are needed for medical association studies but may create more problems. The main weakness of Bonferroni correction is that the results depend on the number of other tests performed. True important differences may be deemed nonsignificant since the likelihood of type II errors is also increased [36]. Cumulatively, our findings provide evidence that polymorphisms in *C-reactive protein* and *Glypican 5* genes are associated with lung cancer risk, and *GKN1* determines chemotherapy response in Chinese population. We believe our results will encourage further studies to understand the function of these genes.

Abbreviations

tSNP: Tag single nucleotide polymorphism
CRP: *C-reactive protein* gene
GPC5: *Glypican 5* gene
ACTA2: *Actin, alpha 2, smooth muscle, aorta* gene
AGPHD1: *Aminoglycoside phosphotransferase domain containing 1* gene
SEC14L5: *SEC14-like 5* gene

Table 6: Genotype and the allele frequencies of candidate genes in chemotherapy patients.

Genotype/allele	Responder		Nonresponder		OR[a]	95% CI		*P* value[a]	*P* value adj.[*]
	N	%	*N*	%					
rs2808630									
T/T	25	71.43	64	82.05	1.000				
T/C	9	25.71	10	12.82	0.434	0.158	1.194	0.142	0.994
C/C	1	2.86	4	5.13	1.652	0.166	14.670	0.791	1
T	59	84.29	138	88.46					
C	11	15.71	18	13.04	0.700	0.311	1.572	0.386	1
rs1926203									
G/G	23	65.71	55	70.51	1.000				
G/T	12	34.29	17	21.79	0.592	0.245	1.435	0.244	1
T/T	0	0	6	7.70	#	#	#	0.999	1
G	58	82.86	127	81.41					
T	12	17.14	29	18.59	1.104	0.526	2.316	0.989	1
rs2352028									
C/C	22	72.86	57	73.08	1.000				
C/T	13	37.14	21	26.92	0.623	0.267	1.457	0.221	1
T/T	0	0	0	0	—	—	—	—	
C	57	81.43	135	86.54					
T	13	18.57	21	13.46	0.682	0.320	1.455	0.269	1
rs8034191									
T/T	34	97.14	74	94.87	1.000				
T/C	1	2.86	4	5.13	1.838	0.198	17.068	0.477	1
C/C	0	0	0	0	—	—	—	—	
T	69	98.57	152	97.44					
C	1	1.43	4	2.56	1.816	0.199	16.547	0.485	1
rs9635542									
A/A	10	28.57	20	25.64	1.000				
G/A	17	48.57	37	47.44	1.088	0.420	2.819	0.866	1
G/G	8	22.86	21	26.92	1.312	0.431	3.996	0.809	1
A	38	54.29	76	48.72					
G	32	45.71	80	51.28	1.250	0.710	2.200	0.619	1
rs4254535									
T/T	28	80.00	46	58.97	1.000				
C/T	5	14.29	27	34.62	3.287	1.135	9.522	0.029	0.203
C/C	2	5.71	5	6.41	1.522	0.276	8.378	0.901	1
T	61	87.14	119	76.28					
C	9	12.86	37	23.72	2.107	0.955	4.649	0.109	0.763
rs1530057									
C/C	33	94.29	71	91.03	1.000				
C/A	2	5.71	7	95.51	1.627	0.320	8.260	0.659	1
A/A	0	0	0	0	—	—	—	—	
C	68	97.14	149	95.51					
A	2	2.86	7	4.49	1.597	0.323	7.891	0.671	1

P value ≤ 0.05 indicates statistical significance; OR: odds ratio; CI: confidence interval.

[a]Adjusted by age, gender, smoke status, histology, stage, and chemotherapy regimens.

[#]When a factor cell associated with the odds ratio is zero, extremely high odds ratios may occur, and it is the same with extremely low odds ratios. It is because the algorithm estimating the logistic coefficient (and hence also exp., the odds ratio) is unstable, failing to converge while attempting to move iteratively toward positive infinity (or negative infinity).

— Some of the mutated genotypes do not exist in the study subjects, so the relative statistics cannot be calculated.

[*] *P* value was adjusted by Bonferroni correction.

RBMS3: *RNA binding motif, single stranded interacting protein 3* gene
GKN1: *Gastrokine 1* gene
LC: Lung cancer
MAF: Minor allele frequency
HWE: Hardy-Weinberg equilibrium
OR: Odds ratio
CI: Confidence intervals.

Conflict of Interests

The authors have no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Science and Technology Major Project of the Ministry of Science and Technology of China (no. 2012ZX09506001) and Collaboration and Communication Project of International Science and Technology of Shaanxi Province (Grant no. 2014KW24-03).

References

[1] L. Yang, D. M. Parkin, L. D. Li, Y. D. Chen, and F. Bray, "Estimation and projection of the national profile of cancer mortality in China: 1991–2005," *British Journal of Cancer*, vol. 90, no. 11, pp. 2157–2166, 2004.

[2] J. She, P. Yang, Q. Y. Hong, and C. X. Bai, "Lung cancer in China: challenges and interventions," *Chest*, vol. 143, no. 4, pp. 1117–1126, 2013.

[3] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA: Cancer Journal for Clinicians*, vol. 61, no. 2, pp. 69–90, 2011.

[4] P. de Groot and R. F. Munden, "Lung cancer epidemiology, risk factors, and prevention," *Radiologic Clinics of North America*, vol. 50, no. 5, pp. 863–876, 2012.

[5] M. Hills and P. M. Lansdorp, "Short telomeres resulting from heritable mutations in the telomerase reverse transcriptase gene predispose for a variety of malignancies," *Annals of the New York Academy of Sciences*, vol. 1176, pp. 178–190, 2009.

[6] K. C. Seng and C. K. Seng, "The success of the genome-wide association approach: a brief story of a long struggle," *European Journal of Human Genetics*, vol. 16, no. 5, pp. 554–564, 2008.

[7] Y. F. Li and P. Yang, "GPC5 gene and its related pathways in lung cancer," *Journal of Thoracic Oncology*, vol. 6, no. 1, pp. 2–5, 2011.

[8] B. Zhou, J. Liu, Z.-M. Wang, and T. Xi, "C-reactive protein, interleukin 6 and lung cancer risk: a meta-analysis," *PLoS ONE*, vol. 7, no. 8, Article ID e43075, 2012.

[9] T. Rafnar, P. Sulem, S. Besenbacher et al., "Genome-wide significant association between a sequence variant at 15q15.2 and lung cancer risk," *Cancer Research*, vol. 71, no. 4, pp. 1356–1361, 2011.

[10] P. Broderick, Y. F. Wang, J. Vijayakrishnan et al., "Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study," *Cancer Research*, vol. 69, no. 16, pp. 6633–6641, 2009.

[11] M. L. Gu, X. Q. Dong, X. Z. Zhang et al., "Strong association between two polymorphisms on 15q25.1 and lung cancer risk: a meta-analysis," *PLoS ONE*, vol. 7, no. 6, Article ID e37970, 2012.

[12] H. W. Lee, Y. M. Park, S. J. Lee et al., "Alpha-smooth muscle actin (ACTA2) is required for metastatic potential of human lung adenocarcinoma," *Clinical Cancer Research*, vol. 19, no. 21, pp. 5879–5889, 2013.

[13] J. Chen, L. Fu, L.-Y. Zhang, D. L. Kwong, L. Yan, and X.-Y. Guan, "Tumor suppressor genes on frequently deleted chromosome 3p in nasopharyngeal carcinoma," *Chinese Journal of Cancer*, vol. 31, no. 5, pp. 215–222, 2012.

[14] P. Therasse, S. G. Arbuck, E. A. Eisenhauer et al., "New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada," *Journal of the National Cancer Institute*, vol. 92, no. 3, pp. 205–216, 2000.

[15] J. L. Ramirez, R. Rosell, M. Taron et al., "14-3-3σ methylation in pretreatment serum circulating DNA of cisplatin-plus-gemcitabine-treated advanced non-small-cell lung cancer patients predicts survival: the Spanish Lung Cancer Group," *Journal of Clinical Oncology*, vol. 23, no. 36, pp. 9105–9112, 2005.

[16] R. K. Thomas, A. C. Baker, R. M. DeBiasi et al., "High-throughput oncogene mutation profiling in human cancer," *Nature Genetics*, vol. 39, no. 3, pp. 347–351, 2007.

[17] S. Gabriel, L. Ziaugra, and D. Tabbaa, "SNP genotyping using the sequenom MassARRAY iPLEX platform," in *Current Protocals in Human Genetics*, vol. 1, chapter 2, unit 2.12, 2009.

[18] C. Adamec, "Example of the use of the nonparametric test. Test X2 for comparison of 2 independent examples," *Ceskoslovenská Zdravotnictví*, vol. 12, pp. 613–619, 1964.

[19] X. Solé, E. Guinó, J. Valls, R. Iniesta, and V. Moreno, "SNPStats: a web tool for the analysis of association studies," *Bioinformatics*, vol. 22, no. 15, pp. 1928–1929, 2006.

[20] J. M. Bland and D. G. Altman, "Statistics notes. The odds ratio," *The British Medical Journal*, vol. 320, no. 7247, p. 1468, 2000.

[21] M. Bolayirli, H. Turna, T. Orhanoglu, R. Ozaras, M. Ilhan, and M. Özgüroglu, "C-reactive protein as an acute phase protein in cancer patients," *Medical Oncology*, vol. 24, no. 3, pp. 338–344, 2007.

[22] K. Heikkilä, S. Ebrahim, and D. A. Lawlor, "A systematic review of the association between circulating concentrations of C reactive protein and cancer," *Journal of Epidemiology and Community Health*, vol. 61, no. 9, pp. 824–832, 2007.

[23] L. A. Hefler, N. Concin, G. Hofstetter et al., "Serum C-reactive protein as independent prognostic variable in patients with ovarian cancer," *Clinical Cancer Research*, vol. 14, no. 3, pp. 710–714, 2008.

[24] J. E. M. Crozier, R. F. McKee, C. S. McArdle et al., "Preoperative but not postoperative systemic inflammatory response correlates with survival in colorectal cancer," *The British Journal of Surgery*, vol. 94, no. 8, pp. 1028–1032, 2007.

[25] G. W. A. Lamb, D. C. McMillan, S. Ramsey, and M. Aitchison, "The relationship between the preoperative systemic inflammatory response and cancer-specific survival in patients undergoing potentially curative resection for renal clear cell cancer," *The British Journal of Cancer*, vol. 94, no. 6, pp. 781–784, 2006.

[26] P. I. Karakiewicz, G. C. Hutterer, Q.-D. Trinh et al., "C-reactive protein is an informative predictor of renal cell carcinoma-specific mortality—a European study of 313 patients," *Cancer*, vol. 110, no. 6, pp. 1241–1247, 2007.

[27] K. Hashimoto, Y. Ikeda, D. Korenaga et al., "The impact of preoperative serum C-reactive protein on the prognosis of patients with hepatocellular carcinoma," *Cancer*, vol. 103, no. 9, pp. 1856–1864, 2005.

[28] I. Gockel, K. Dirksen, C. M. Messow, and T. Junginger, "Significance of preoperative C-reactive protein as a parameter of the perioperative course and long-term prognosis in squamous cell carcinoma and adenocarcinoma of the oesophagus," *World Journal of Gastroenterology*, vol. 12, no. 23, pp. 3746–3750, 2006.

[29] K. H. Allin and B. G. Nordestgaard, "Elevated C-reactive protein in the diagnosis, prognosis, and cause of cancer," *Critical Reviews in Clinical Laboratory Sciences*, vol. 48, no. 4, pp. 155–170, 2011.

[30] M. Xu, M. L. Zhu, Y. G. Du et al., "Serum C-reactive protein and risk of lung cancer: a case-control study," *Medical Oncology*, vol. 30, no. 1, article 319, 2013.

[31] Y. F. Li, C.-C. Sheu, Y. Q. Ye et al., "Genetic variants and risk of lung cancer in never smokers: a genome-wide association study," *The Lancet Oncology*, vol. 11, no. 4, pp. 321–330, 2010.

[32] M. T. Landi, N. Chatterjee, N. E. Caporaso et al., "GPC5 rs2352028 variant and risk of lung cancer in never smokers," *The Lancet Oncology*, vol. 11, no. 8, pp. 714–716, 2010.

[33] Y. Zheng, M. Kan, L. Yu et al., "GPC5 rs2352028 polymorphism and risk of lung cancer in Han Chinese," *Cancer Investigation*, vol. 30, no. 1, pp. 13–19, 2012.

[34] J. H. Yoon, Y. J. Choi, W. S. Choi et al., "GKN1-miR-185-DNMT1 axis suppresses gastric carcinogenesis through regulation of epigenetic alteration and cell cycle," *Clinical Cancer Research*, vol. 19, no. 17, pp. 4599–4610, 2013.

[35] F. B. Fahlbusch, M. Ruebner, H. Huebner et al., "The tumor suppressor gastrokine-1 is expressed in placenta and contributes to the regulation of trophoblast migration," *Placenta*, vol. 34, no. 11, pp. 1027–1035, 2013.

[36] T. V. Perneger, "What's wrong with Bonferroni adjustments," *The British Medical Journal*, vol. 316, no. 7139, pp. 1236–1238, 1998.