Review:
Proteomics and Biomarkers for Ovarian Cancer Diagnosis

Bei Zhang,1 Zeinab Barekati,1 Corina Kohler,1 Ramin Radpour,1 Reza Asadollahi,1 Wolfgang Holzgreve,2 and Xiao Yan Zhong1
1Department of Biomedicine, Women’s Hospital, University of Basel, Basel, Switzerland
2University Medical Center Freiburg, Freiburg, Germany

Abstract. Ovarian cancer remains a leading cause of death from gynecological malignancy. Early diagnosis is the most important determinant of survival. Current diagnostic tools have had very limited success in early detection. In recent years, the advancing techniques for proteomics have accelerated the discovery of ovarian cancer biomarkers. Numerous proteomics-based molecular biomarkers/panels have been identified and hold great potential for diagnostic applications, but they need further development and validation. This article reviews recently published data on the diagnosis of ovarian cancer with proteomics, including the major proteomics technologies and promising strategies for biomarker discovery and development.

Keywords: proteomics, ovarian cancer, mass spectrometry, microarrays, 2-dimensional electrophoresis

Introduction

Ovarian cancer is the most lethal of all common gynecologic malignancies, with more than 204,000 new cases and 125,000 deaths each year, accounting for 4% of all cancer cases and 4.2% of all cancer deaths in women around the world [1]. Data from the Swiss Association of Cancer Registries (1986-2005) show that ovarian cancer is the seventh most common cancer and the fifth most frequent cause of death from cancer in Swiss women [2,3].

Contributing to the poor prognosis of ovarian cancer is the lack of symptoms in the early stages of the disease [4-8]. Current diagnostic methods, such as CA-125 assay, for the detection of early stage ovarian cancer are inadequate. Only 25% of all ovarian cancer is found at an early stage [4,6,9,10]. Given the prevalence of ovarian cancer, strategies for early detection require high sensitivity for early stage disease (>75%), and must have extremely high specificity (99.6%) to attain a positive predictive value of at least 10% [6,9-11]. There is no single screening test nor any existing screening paradigm that currently has such high specificity [11,12]. Thus, discovery of specific molecular biomarkers/panels is emerging as an important requirement for early detection of ovarian cancer. With the progress in techniques for proteomics, numerous proteomics-based molecular biomarkers/panels have been identified and show promise for ovarian cancer diagnosis, but they need validation. The present review highlights the discovery of biomarkers for early detection of ovarian cancer using proteomics, including the major proteomics technologies, promising strategies for biomarker discovery and development, and proteomics-based biomarkers/panels that have already been described.

Proteomic Technology for Cancer Biomarkers

Proteomics has emerged as a powerful technology to decipher biological processes. It denotes the
large-scale characterization of proteins, including complicated features like isoforms, modifications, interactions, and functional structures [13]. One of the main goals of proteomics is the identification of biomarkers for diseases. As the proteome in any cell represents a subset of all possible gene products, the genome is simpler than the proteome; any protein may exist in multiple forms that vary within a particular cell or in different cells. Modifications may derive from translational, post-translational, regulatory, and degradative processes that affect protein structure, localization, function, and turnover [14]. Thus, protein biomarkers may be more specific in respect to cancer type and status than gene-based biomarkers. For diagnostic applications, protein-based biomarkers in body fluids provide easy-access as well.

With proteomics defined as the study of all proteins in a biological system, it is obvious that much is demanded from the analytical technologies. Fortunately, progress in proteomic technologies such as protein separation, quantification, and identification makes it possible to analyze proteins more intensively and to understand their functions [15,16]. Therefore a systemic overview of expressed proteins may lead to improved cancer diagnosis, treatment, and prognosis.

In cancer biomarker discovery, two-dimensional gel electrophoresis, mass spectrometry (MS), and protein microarrays [17], in conjunction with advanced bioinformatics, have become powerful tools to identify proteins.

**Two-dimensional gel electrophoresis.** Traditionally, proteomic analyses have been performed using two-dimensional gel electrophoresis (2DE), which separates proteins according to two distinct protein characteristics, size and charge [17]. In recent years, the introduction of immobilized pH gradients and advanced bioinformatics have vastly improved the reproducibility and comparability of this technique, although low-throughput remains a serious obstacle to 2DE becoming routine in clinical laboratories [14,17]. However, 2DE is extremely valuable for research.

In ovarian cancer, many studies have shown that 2DE can detect differences between normal and cancer sample proteomes [18-23] (Table 1). In addition, the introduction of fluorescent two-dimensional differential in-gel electrophoresis (2D-DIGE) is helping to solve the major inherent drawbacks of 2DE and to enable accurate protein separation, detection, and quantitation. 2D-DIGE is now becoming an important tool for biomarker discovery of cancers, including ovarian cancer [24].

**Mass spectrometry (MS).** The major technique that fundamentally supports the discovery of cancer biomarkers is MS, which can determine the precise mass and charge of proteins, and thus identify the actual precursor proteins or protein profiles. In general, MS instruments are made up of three primary components: the source, which produces ions for analysis; the mass analyser, which identifies the ions based on their mass-to-charge ratios; and the detector, which quantifies the ions resolved by the analyser [25]. In recent years, MS instruments have been greatly modified and improved; highly sensitive, robust instruments have been developed to analyze biomolecules, proteins and peptides in particular. This technique is now sensitive to the picomole to femtomole range required for the detection of cancer biomarkers like oligonucleotides, small polar molecules, peptides, proteins, and posttranslationally modified proteins, such as glycoproteins and phosphoproteins. However, MS is not helpful for detecting high molecular weight and heavily glycosylated proteins [25].

Among the MS-based proteomics approaches, matrix-assisted laser desorption and ionization time-of-flight (MALDI–TOF) and surface-enhanced laser desorption and ionization time-of-flight (SELDI–TOF) techniques are two of the most frequent methods for new biomarker discovery. MALDI–TOF can identify nano- to pico-molar amounts of protein. The process involves producing a co-precipitate of an ultraviolet light-absorbing matrix and the protein being studied. The coprecipitate is irradiated with a laser, and then the ionized biomolecules are accelerated in an electric field and enter the flight tube. Within the tube, the molecules are separated according to their mass, producing distinct signals [25]. SELDI-TOF involves chromatography, protein chip application, and MS-based detection. It includes the capture of proteins on a resin ‘chip’ that
<table>
<thead>
<tr>
<th>Authors</th>
<th>Methods (and sample)</th>
<th>Identified Biomarker</th>
<th>Regulated in Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petri et al (2009)[44]</td>
<td>SELDI-TOF MS (urine)</td>
<td>fibrinogen alpha fragment</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>collagen alpha 1 (III) fragment</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fibrinogen beta NT fragment</td>
<td>√</td>
</tr>
<tr>
<td>Li et al (2009)[18]</td>
<td>2DE &amp; MALDI-TOF MS (tissue)</td>
<td>prx-II</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>prx-III</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hsb27</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hsb60</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mitochondrial short-chain enoyl-CoA hydratase</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>prohibitin</td>
<td>√</td>
</tr>
<tr>
<td>Helleman et al (2008)[45]</td>
<td>SELDI-TOF MS (serum)</td>
<td>amyloid A1</td>
<td>√</td>
</tr>
<tr>
<td>Scholler et al (2008)[19]</td>
<td>1DE &amp; 2DE; LC-MS/MS (serum)</td>
<td>Catabolic fragments of complement factors</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EMILIN2</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>von Willebrand factor</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PEBP1 (RKIP)</td>
<td>√</td>
</tr>
<tr>
<td>Lemaire et al (2007)[46]</td>
<td>MALDI-TOF MS (tissue)</td>
<td>PA28 (Reg-alpha fragment)</td>
<td>√</td>
</tr>
<tr>
<td>An et al (2006)[20]</td>
<td>2DE &amp; MALDI-TOF MS (tissue)</td>
<td>NM23-H1</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>annexin-1</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>protein phosphatase-1</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ferritin light chain</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>proteasome alpha-6</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAGK (N-acetyl glucosamine kinase)</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>COOH-terminal osteopontin fragments</td>
<td>√</td>
</tr>
<tr>
<td>Kozak et al (2005)[47]</td>
<td>SELDI-TOF MS (serum)</td>
<td>transthyretin</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>beta-hemoglobin</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>apolipoprotein A1</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>transferrin</td>
<td>√</td>
</tr>
<tr>
<td>Ahmed et al (2005)[22]</td>
<td>2DE, MALDI-TOF MS (serum)</td>
<td>haptoglobin</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>transferrin</td>
<td>√</td>
</tr>
<tr>
<td>Moshkovskii et al (2005)[48]</td>
<td>SELDI-TOF MS (plasma)</td>
<td>amyloid A1</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>truncated form of transthyretin</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cleavage fragment of inter-α trypsin inhibitor heavy chain H4</td>
<td>√</td>
</tr>
<tr>
<td>Bergen et al (2003)[50]</td>
<td>nanoLC/ESI-TOF, FT-ICR MS (serum)</td>
<td>fibrinopeptide-A</td>
<td>√</td>
</tr>
<tr>
<td>Ye et al (2003)[51]</td>
<td>SELDI-TOF MS (serum)</td>
<td>haptoglobin-derived alpha subunit</td>
<td>√</td>
</tr>
</tbody>
</table>

**SELDI-TOF MS**: surface-enhanced laser desorption/ionization time-of-flight mass spectrometry; **MALDI-TOF MS**: matrix-assisted laser desorption ionisation time-of-flight mass spectrometry; **1DE**: one dimensional gel electrophoresis; **2DE**: two dimensional gel electrophoresis; **LC-MS/MS**: liquid chromatography-mass spectrometry; **2D-DIGE**: two-dimensional differential in-gel electrophoresis; **nanoLC/ESI-TOF-MS**: nano-liquid chromatography/electrospray ionization time-of-flight mass spectrometry; **FT-ICR MS**: Fourier-transform ion cyclotron resonance mass spectrometry.
fractionates and isolates proteins based on charge, hydrophobicity, and other characteristics. The proteins are then assayed using laser desorption ionization time-of-flight mass spectroscopy [26, 27]. In particular, the capture of proteins of interest on a protein chip array, directly from the original source material and without previous sample preparation, makes this technique promising for clinical applications [14]. This sensitive technique allows the identification of thousands of proteins bound to a single site in minute concentrations.

The two aforementioned MS techniques have the advantage of being adaptable to high-throughput settings and can detect parts of the proteome, such as low molecular weight proteins, that could not be studied easily in the past [28]. They have assumed key roles in most proteomic workflows for ovarian cancer biomarker discovery (Table 1).

Researchers have recently emphasized the importance of reliability and reproducibility of MS methods for protein profiling [29]. The reproducibility of findings in studies of ovarian cancer and normal tissues was discussed, with the conclusion that differences in the proteomic profiles uncovered in an experiment [30] were due to sample processing and not the underlying biology of cancer [31]. Another review summarized studies of reproducibility of MALDI protein profiling and current approaches to improve its analytical performance, including automated sample processing, prefractionation strategies, immuno-capture, prestructured target surfaces, standardized matrix (co)crystallization, improved MALDI-TOF MS instrument components, internal standards, quality-control samples, replicate measurements, and algorithms for normalization and peak detection. The author recommended further evaluation and optimization of MALDI-TOF MS before it is used for routine analysis [32]. These findings attest that quality assessment of advanced MS proteomic techniques is important in the field of proteomics.

**Protein microarrays.** Protein microarrays, similar to gene arrays, have emerged as a promising technique to analyze at a high-throughput level the abundance of proteins and their modifications (such as phosphorylation status). Protein microarrays are now being used to profile the proteome of cell populations using antigen-antibody interactions [28]. Protein microarray formats can be divided into two major classes: (i) forward-phase arrays (FPAs), where antibodies are arrayed and probed with cell lysates, or (ii) reverse-phase arrays (RPAs), where cell lysates are arrayed and probed with antibodies [17,33]. In contrast to FPAs, RPAs do not require labelling of cellular protein lysates, and constitute a sensitive high-throughput platform for marker screening, pathophysiologic studies, and therapeutic monitoring [34]. Furthermore, the RPA has a unique ability to analyze signaling pathways using small numbers of cultured cells or cells isolated by laser capture microdissection (LCM) from human tissues procured during clinical trials [33]. However, the major limitation for RPAs is the specific antibodies that are required. In ovarian cancer, the RPA platform has been used to study disease progression and profile signaling pathways, identify therapeutic targets, and suggest prognostic indicators [35].

**Ovarian Cancer Biomarker Discovery**

**Sample collection.** Biomaterials, mostly from cancer tissues and body fluids, are usually used for protein biomarker discovery in ovarian cancer. In biopsied ovarian tissues, not only cancer cells but also other molecules involved in self-defense mechanisms of the human body can be found, such as immune cells, cytokines, and derivatives from immune or inflammation responses.

Blood is the most common body fluid used in biomarker studies because of its easy access and routine use for blood chemistry measurements in patients. Many biomarkers present in biopsied cancer tissues can also be found in blood. This includes circulating protein fragments generated in the diseased tissue and its microenvironment. Because the ultimate goal of biomarkers is the specific, early, and non-invasive diagnosis and post-therapy monitoring of cancer, blood is an appropriate biological material. Thus, many biomarker discoveries are carried out with blood-based strategies [36,37].
Several biomarkers of interest for ovarian cancer have been identified in urine. The strengths of urine for biomarker studies are that the protein profile may be less complex in urine than in blood, and that proteins in urine may be more stable than those in blood, as well as the convenience that a urine test has when compared to a more invasive blood test. However, care is necessary to ensure that urine collection and storage are uniform in trials of urinary biomarkers, because cleavage of proteins may be generated by specific proteases present in the urine of patients [38].

**Protein preparation and separation.** Applications of MS-based technologies are limited by the high dynamic range of proteins present in the serum [39]. Thus, prior to searching for biomarker proteins, proteomes in the body fluids need to be separated by specific characteristics, such as by glycoproteome or phosphoproteome enrichments, ion charges, hydrophobicity or hydrophilicity by strong cation exchange/strong anion exchange (SCX/SAX), reverse phase or normal phase chromatography, or by molecular weights, etc.

Glycoproteome analysis in the body fluids has great advantages in the area of cancer biomarker discovery. About half of the serum proteins are known to be glycosylated and the glycosylation status of glycoproteins, their degrees and forms, is also altered by disease conditions including cancers. Also, various known biomarkers are glycoproteins, such as the ovarian cancer biomarker, CA-125.

**Essential procedures for ovarian cancer biomarker discovery.** After proper preparation, each protein can be analyzed by similar procedures of identification, verification, and validation. For protein identification, prepared samples are analyzed by MS. Potential biomarkers, elevated or decreased in their expression levels, are confirmed by Western blot, ELISA, or by recently developed multiple reaction monitoring. Validation of potential markers must be carried out initially with sera from patients with early-stage ovarian cancer obtained at the time of conventional diagnosis. Ultimately, however, multiple biomarkers must be tested for their ability to detect disease in asymptomatic women before detection of cancer prompted by symptoms or an abnormal pelvic examination [6]. For biomarker validation, multiplex assays should be developed to permit the simultaneous assay of multiple markers in the small volumes of sera that are likely to be available from retrospective trials. For example, the Luminex LabMap technology, which combines the principle of a sandwich immunoassay with a fluorescent bead-based technology, allowing individual and multiplex analysis of up to 100 different analytes with as little as 50 μl of serum, has been applied to measurement of multiple markers [6]. Additionally, if multiple markers are used to increase sensitivity without sacrificing specificity, a more sophisticated mathematical analysis will be required to analyze the trends of the biomarker levels over time [9].

**Proteomic Analysis in Ovarian Cancer Diagnosis**

Proteomic analyses in ovarian cancer diagnosis have followed two paths [14]: one, called “proteomic pattern diagnostics” or “serum proteomic profiling,” is based on complex mass spectrometric differences between proteomic patterns of samples with and without cancer identified by bioinformatics. The “proteomic pattern diagnostics” approach to discover new biomarkers by MS is based on two premises: (i) the low-molecular-weight serum proteome contains an enormous wealth of biomarker information, which has not yet been explored, and (ii) a pattern of multiple biomarkers may contain a higher level of discriminatory information than a single biomarker alone across large heterogeneous patient populations [40]. Published studies show that proteomic pattern analysis in ovarian cancer has the potential to be a novel, highly sensitive diagnostic tool for detection at an early stage [30,41,42]. However, despite the promising results in terms of specificity and sensitivity in ovarian cancer detection, some criticisms regarding instrument reproducibility, quality control, and standard operating procedures for sample collection, handling, and shipping have been raised [40,43].

An alternative or integrative proteomic approach to ovarian cancer biomarkers is the identification of single, novel biomarkers and the subsequent development of new assays [14].
Workflow of ovarian cancer biomarker discovery using proteomics was briefly described above. In recent years many promising biomarkers discovered by proteomic analysis for ovarian cancer diagnosis have been published (Table I) [18-24,44-51].

Among the various markers identified by proteomic analysis, some, such as the cleavage fragment of inter-alpha-trypsin inhibitor heavy chain H4 [49], are normal serum proteins that have undergone posttranslational modification by proteases and reflect the protease profiles of particular cancers. Other biomarkers, such as transferrin [22], are acute phase proteins and have been associated with systemic inflammation as well as other non-neoplastic conditions. They are not cancer-specific markers and are not derived directly from the ovarian cancers. So for proteomics-based biomarkers, their significance and degree of specificity for ovarian cancer remain to be explored. In fact, to date no single test or modality has met the criterion (positive predictive value of 10%) for early diagnosis of ovarian cancer [6].

Given the complexity and heterogeneity of ovarian cancer, it is unlikely that a single biomarker will detect all subtypes and stages of the disease with high specificity and high sensitivity. Many studies show that combining several biomarkers dramatically improves the sensitivity of CA-125 in ovarian cancer patients [52]. Markers have generally been analyzed only two or three at a time. The increased sensitivity achieved with markers in combination has generally been associated with decrease in specificity [9]. A couple of biomarker panels have been published with adaptable sensitivity and specificity ranges, which may hold great potential for the detection of ovarian cancer [53]. However, most of the impressive sensitivities and specificities for biomarker panels have been based on relatively small numbers of samples (especially a few cases of stage I disease) without independent validation studies.

Before biomarker tests are adopted for routine diagnostic use, more research, such as retrospective and prospective clinical trials, is needed to evaluate the overall clinical utility of the tests. In the future, it will be crucial to develop panels of biomarkers not only for early detection of ovarian cancer but also for guidance of ovarian cancer therapy.

Challenges in Proteomics of Biomarkers

Since cancer biomarker discovery has started extensively with the progress of proteomic technology, many protein molecules or proteomic patterns have been identified as potential ovarian cancer biomarkers. However, critical assessment of the results has shown significant shortcomings and uncertainties in regard to the reproducibility of the findings, identity of the proteins behind the pattern peaks, and validation of the results. Interlaboratory SELDI experiments performed recently alleviated some of the reproducibility concerns [54]. In fact, validation of the newly discovered biomarkers remains the most challenging aspect of clinical proteomics.

Critical issues that need to be addressed for the validation studies include the specificity and reproducibility of the biomarker, the experimental design, and appropriate controls, such as specimen collection, handling, study design, and data analysis. Furthermore, it has yet to be established whether the patterns identified reflect cancer-specific phenomena, or epiphenomena related to general inflammatory responses or metabolic disturbances. Nor is it certain that the results from different laboratories are comparable; the effects of sample handling on the obtained patterns should be investigated [29]. Another challenge is that validation of MS data has been performed with the use of antibodies in a Western blot or ELISA platform, but these methods may require a large quantity of the specific antigen being identified. The lack of antibodies specific to small peptide fragments that do not cross-react with the larger “parental” peptide is a great problem [55].

Conclusion and Perspectives

During the last decade, with the development of high-throughput technologies in proteomics, several proteomics-based ovarian cancer biomarkers have shown promise in a variety of studies and have provided new insights into ovarian cancer diagnosis, but few have turned out to be useful in the clinic. Whether serum proteomic profiling or the measurement of peptide markers/panels will yield the most accurate approaches to ovarian cancer
References


