

RESEARCH ARTICLE

Open Access



MALDI-TOF-MS analysis in discovery and identification of serum proteomic patterns of ovarian cancer

Agata Swiatly^{1†}, Agnieszka Horal^{2†}, Joanna Hajduk¹, Jan Matysiak¹, Ewa Nowak-Markwitz² and Zenon J. Kokot^{1*}

Abstract

Background: Due to high mortality and lack of efficient screening, new tools for ovarian cancer (OC) diagnosis are urgently needed. To broaden the knowledge on the pathological processes that occur during ovarian cancer tumorigenesis, protein-peptide profiling was proposed.

Methods: Serum proteomic patterns in samples from OC patients were obtained using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF). Eighty nine serum samples (44 ovarian cancer and 45 healthy controls) were pretreated using solid-phase extraction method. Next, a classification model with the most discriminative factors was identified using chemometric algorithms. Finally, the results were verified by external validation on an independent test set of samples.

Results: Main outcome of this study was an identification of potential OC biomarkers by applying liquid chromatography coupled with tandem mass spectrometry. Application of this novel strategy enabled the identification of four potential OC serum biomarkers (complement C3, kininogen-1, inter-alpha-trypsin inhibitor heavy chain H4, and transthyretin). The role of these proteins was discussed in relation to OC pathomechanism.

Conclusions: The study results may contribute to the development of clinically useful multi-component diagnostic tools in OC. In addition, identifying a novel panel of discriminative proteins could provide a new insight into complex signaling and functional networks associated with this multifactorial disease.

Keywords: Epithelial ovarian cancer, Ovarian cancer, Biomarkers, MALDI-TOF, Protein-peptide profiling

Background

Ovarian cancer (OC) is one of the leading causes of death among all gynecological malignancies [1]. As there are no early specific symptoms, OC is diagnosed in advanced clinical stages in more than 70% cases when, despite appropriate treatment, 5-year survival rate drops to 30% [2]. Early diagnosis improves treatment outcomes and also dramatically reduces mortality rate [3]. However, adequate diagnostic methods are lacking and therefore novel technologies that would allow early detection of OC are urgently needed.

Serum measurement of cancer antigen 125 (CA125) and transvaginal ultrasound examination have become the most widely used methods in OC diagnosis [4]. Nonetheless, they are characterized by low specificity, especially in early stage cancer and in women before menopause [5]. Extensive efforts to identify other OC biomarkers led to the discovery of human epididymis protein 4 (HE4). Usefulness of HE4 in diagnosis of OC has been widely explored [6–8]. As single cancer biomarkers were insufficient to detect a tumor in its early stages, many studies focused on the development of multi-marker serum panels [3, 9]. Food and Drug Administration (FDA) cleared for use two multiple biomarker tests: Risk of Ovarian Malignancy Algorithm (ROMA) and OVA1 - a multivariate index assay (MIA). ROMA combines serum CA125 and HE4 levels with menopausal status. This predictive probability algorithm

* Correspondence: zkokot@ump.edu.pl

[†]Equal contributors

¹Department of Inorganic and Analytical Chemistry, Poznan University of Medical Sciences, ul. Grunwaldzka 6, 60-780 Poznań, Poland
Full list of author information is available at the end of the article



allows for classifying patients into high and low risk OC groups [10]. The OVA1 test is a proprietary algorithm that combines serum concentrations of five markers (CA125, apolipoprotein A-1, β 2-microglobulin, transthyretin and transferrin) and calculates a malignancy risk index score [9].

Despite the use of multi-marker diagnostic strategies, early detection of OC remains far from satisfactory. Thus, new strategies based on novel methodology such as proteomic research have been employed in OC research [11]. In recent years, untargeted proteomics, such as protein-peptide profiling, has emerged as an interesting tool for clinical diagnostics [12–14]. Identification of distinctive pattern of protein expression is a promising strategy for understanding molecular alterations during pathological processes [15]. Subsequently, the obtained information could be useful in detection of specific biomarkers and could increase the efficacy of early diagnosis [16]. One of the most frequently used tools in proteomic research (besides ESI - electrospray ionization) is matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) [17]. MALDI-TOF instruments have been reported sensitive and robust for clinical trials [18]. However, in the studies based on mass spectrometry analyses of complex biological samples like blood, serum or plasma, application of enrichment strategies seems to be necessary for generating good quality mass spectra [19]. Highly abundant proteins, as well as the presence of lipids and salts, mask other low abundant compounds, including cancer-related biomarkers [20]. Therefore, many different strategies have been proposed to pretreat plasma or serum samples. Currently, MALDI-TOF MS combined with ZipTip micropipette tips based on solid phase extraction proved successful. Moreover, several studies explored robustness and reliability of this methodology in protein-peptide profiling [20, 21].

The aim of this study was to characterize MALDI-TOF-MS-based serum proteomic patterns of OC and to identify differences in those patterns between OC samples and healthy control group. As far as we know, the combination of solid phase extraction pretreatment with MALDI-TOF-MS in OC research was presented for the first time. The MS data obtained were further processed and analyzed with advanced chemometric tools. A classification model containing the most discriminative peaks was calculated based on the obtained spectra and verified using an independent test set. Potential OC serum biomarkers were identified using nano-liquid chromatography (nano-LC) coupled with MALDI-TOF-MS/MS, since they might provide a new insight into the multifactorial processes that occur during OC tumorigenesis. To the best of our knowledge this is the first study in

which novel OC protein patterns have been both discovered and identified based on MALDI-TOF MS techniques.

Methods

Characteristics of the study groups

Blood samples were collected from 89 patients operated in Gynecologic Oncology Department of Poznan University of Medical Sciences, Poland, on the day before surgery, between August 2014 and December 2015. Blood samples were incubated for 30 min at room temperature for clotting and centrifuged for 15 min at 4000 rpm. The resulting sera were isolated and stored at -80°C until analysis. All serum samples were handled using the same laboratory equipment and stored in the same type of plastic vials and boxes. Based on histopathological result the patients were divided into two groups: OC (including borderline ovarian tumors) ($N = 44$) and no pathology of the ovaries - further referred to as “control group” ($N = 45$). The control group consisted of patients operated (hysterectomy with bilateral salpingo-oophorectomy) due to reasons other than ovarian tumors and in which the final histopathological examination confirmed no existing ovarian pathology. All participants were after overnight fasting. The patients were selected according to the following exclusion criteria: other than epithelial OC, other cancers currently or in anamnesis, chronic metabolic diseases (diabetes, dyslipidemia), previous or current cancer treatment (radiotherapy, chemotherapy, hormonal therapy), relevant concomitant medication (anti-diabetic agents, statins, hormonal replacement therapy, oral contraception). Additionally two markers, CA124 and HE4, were measured in the OC group with an electrochemiluminescence immunoassay (Roche Diagnostics, Indianapolis, IN, USA). Detailed characterization of the studied groups, including demographic and clinical profiles, is presented in Table 1 and Additional file 1 Table S1. The project was approved by the Bioethics Committee of Poznan University of Medical Sciences, Poland (Decision No. 165/16).

Serum samples pretreatment

Each sample was diluted in 0.1% trifluoroacetic acid (TFA) in water (1:5). In order to desalt and concentrate the samples, solid phase extraction method based on ZipTip C18 pipette tips was used according to the manufacturer's protocol (Millipore, Bedford, MA, USA). The tips were conditioned with acetonitrile (ACN) and 0.1% TFA. The prepared samples were loaded onto the tips and the peptides were bound. After washing with 0.1% TFA, sample fractions were eluted using 50% ACN solution in 0.1% TFA.

MALDI-TOF-MS protein and peptide profiling

Each eluent sample was mixed with matrix solution of α -cyano-4-hydroxycinnamic acid (0.3 g/L HCCA in a

Table 1 Study group characteristics

Patient group	Number of samples	Median age (min-max)	Median BMI (min-max)	% of postmenopausal	Average concentration of CA125 (U/mL)	Average concentration of HE4 (pmol/L)
OC training set	33	57 (36–72)	26.81 (17.29–38.37)	23 (70%)	2381.29	1025.10
- Type I OC	10					
* borderline	*5					
- Type II	23					
OC test set	11	65 (32–78)	24.75 (22.27–31.62)	9 (82%)	2177.67	1261.00
- Type I OC	3					
* borderline	*1					
- Type II	8					
Control training set	33	58 (19–73)	26.06 (21.15–40.06)	22 (67%)	not determined	not determined
Control test set	12	55 (31–63)	27.56 (22.43–35.70)	7 (58%)	not determined	not determined

solution containing 2:1 ethanol:acetone, *v/v*) at the ratio of 1:10. One microliter of the sample/matrix solution was spotted onto the MALDI target (AnchorChip 800 μm , Bruker Daltonics, Bremen, Germany) and left to crystallize at room temperature. The samples from both study groups were analyzed in a random order and the disease status of the women was blinded to minimize variability and systematic errors. UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) was used to perform MS analyses in the linear positive mode. Positively charged ions were detected in the *m/z* range of 1000–10,000 Da and 2000 shots were accumulated per one spectrum. The MS spectra were externally calibrated with the mixture of Peptide Calibration Standard and Protein Calibration Standard I at the ratio of 1:5. The average mass deviation was less than 100 ppm. The matrix suppression mass cut off was *m/z* 700 Da. The following ion source parameters were used: ion source 1, 25.09 kV; ion source 2, 23.80 kV. Other settings for MALDI-TOF MS analysis were as follows: pulsed ion extraction, 260 ns and lens, 6.40 kV. FlexControl 3.4 software (Bruker Daltonics, Bremen, Germany) was applied for the acquisition and processing of the spectra. Each sample was analyzed in three repetitions. Inter-day and intra-day reproducibility of the applied procedure was evaluated in our previous study [22].

nanoLC-MALDI-TOF-TOF MS/MS identification of discriminative peaks

The sample was prepared with ZipTip technique. The obtained eluent was further subjected to nano-LC separation using: nanoflow HPLC set (EASY-nano LC II, Bruker Daltonics, Germany) and fraction collector (ProteinExpert II, Bruker Daltonics, Germany). The nano-LC system consisted of a trap column, NS-MP-10 BioSphere C18, (20 mm \times 100 μm I.D., particle size 5 μm , pore size 120 \AA) (NanoSeparations, Nieuwkoop, the Netherlands) and Thermo Scientific Acclaim PepMap 100 column C18 (150 mm \times 75 μm

I.D., particle size 3 μm , pore size 100 \AA) (Thermo Scientific: Sunnyvale, CA, USA). Linear gradient was 2%–50% of ACN during 96 min. Two mobile phases were used: mobile phase A (0.05% TFA in water) and mobile phase B (0.05% TFA 90% ACN). The volume of injected sample eluent was 4 μL . The separation was performed with a flow rate 300 nL/min. A total of 384 fractions, 80 nL each, were obtained. Each eluent was automatically mixed with 420 nL of matrix solution that was prepared by mixing 36 μL of HCCA saturated solution of 0.1% TFA and ACN (90:10 *v/v*), 784 μL ACN and 0.1% TFA (95:5 *v/v*), 8 μL of 10% TFA and 8 μL of 100 mM ammonium phosphate monobasic and spotted onto the MALDI target (AnchorChip 800 μm) using a fraction collector. The system was controlled by HyStar 3.2 software (Bruker Daltonics, Germany). MALDI-TOF/TOF mass spectrometer (UltrafleXtreme, Bruker Daltonics, Germany) operated in a reflector mode was used in further analysis of the sample. The MS spectra were externally calibrated using Peptide Calibration Standard mixture (Bruker Daltonics, Germany). A list of precursor peaks was obtained using WARP-LC software (Bruker Daltonics, Germany). The chosen discriminative *m/z* were analyzed with MS/MS mode for protein identification. The parameters for MS and MS/MS mode were described in our previous study [22]. FlexControl 3.4 software (Bruker Daltonics, Germany) was applied for the acquisition of spectra. Processing and evaluation of the data was achieved using FlexAnalysis 3.4 (Bruker Daltonics, Germany). BioTools 3.2 (Bruker Daltonics, Germany) was used to perform protein database searches. Proteins were identified using the SwissProt database and Mascot 2.4.1 search engine (Matrix Science, London, UK) with taxonomical restriction to "*Homo sapiens*". The following general protein search parameters were used: precursor-ion mass tolerance ± 50 ppm; fragment-ion mass tolerance ± 0.7 Da; no enzyme; monoisotopic mass; peptide charge +1.

Data analysis

Data analysis of each spectrum was performed with ClinProTools version 3.0 software (Bruker Daltonics, Germany). In order to let the software group all analyzed sample replicates into one biological replicate, spectra grouping function was applied. This option provided improved measurement quality. Before any analysis or spectra processing, the multiple measurements were averaged. Further steps were processed upon one averaged spectrum per sample. Comparison of the obtained data was achieved through a standard workflow. Each spectrum was first normalized to the total ion current (TIC) and recalibrated with the prominent common m/z values. "Top hat" baseline subtraction with the minimum baseline width set to 10% was used to remove broad structures. Spectra were also smoothed and processed in the mass range of 1000–10,000 Da. The signal-to-noise ratio was greater than or equal to 5. Peak picking and average peak calculation procedures were used. A total average spectrum was calculated from the preprocessed spectra. Averaging of the spectra allowed us to improve the signal to noise during peak picking procedure. Due to average peak list calculation, small peaks that might be missed on a single spectrum, were included in the overall profile. All reproducible peaks were detected according to this procedure.

Comparisons between patients with OC and healthy individuals were evaluated with Wilcoxon test. Statistical significance was attained when p -value was ≤ 0.02 . All p -

values were internally corrected with the Benjamini-Hochberg algorithm. Evaluation of the discrimination ability of each peak was achieved by calculating receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) (Fig. 1). Chemometric algorithms: supervised neural network (SNN), genetic algorithm (GA), and quick classifier (QC) were used for model analysis and selection of peptide/protein peak clusters. Each model indicated a combination of the differentiating peaks. The studied groups were randomly subdivided into a training set (containing 33 ovarian cancer patients and 33 healthy controls) and a test set (containing 11 ovarian cancer patients and 12 healthy controls). The use of these two sets allowed for testing robustness of the obtained models. For the training set two parameters, 20% leave one out cross validation and recognition capability, were calculated. For the model with the best performance of these two indicators, an external validation using the test set was calculated. The values of sensitivity and specificity were used to define discriminative ability of the model. Peaks that indicated the best discrimination between the studied groups were further identified as fragments of defined proteins.

Results

Eighty nine serum samples derived from ovarian cancer patients ($n = 44$) and healthy individuals ($n = 45$) were pretreated with ZipTips and analyzed in triplicate by MALDI-TOF MS. The reproducibility and reliability of

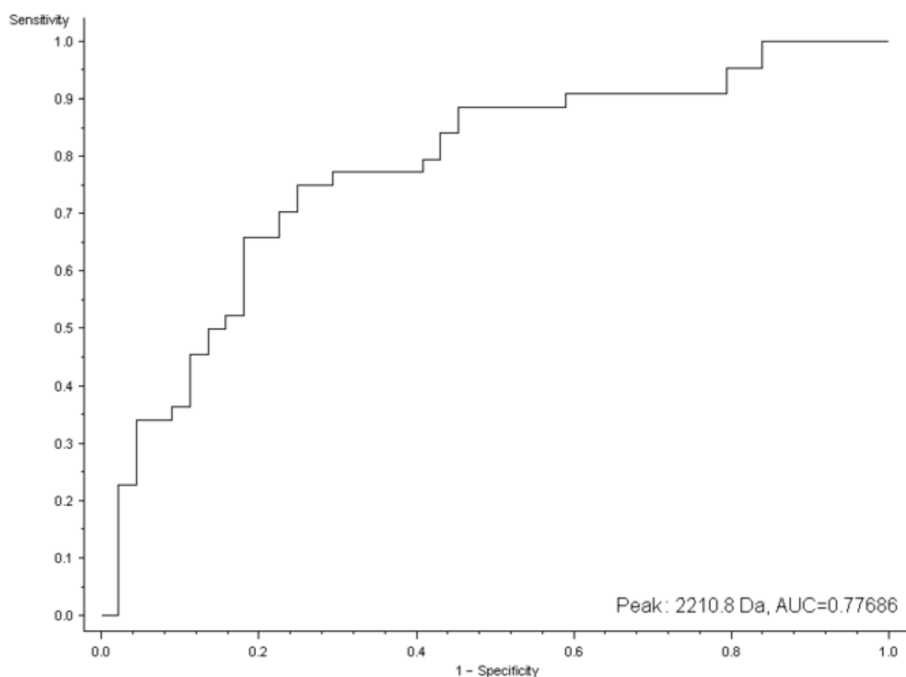


Fig. 1 Receiver operating characteristic (ROC) curve representing sensitivity and specificity of m/z peak 2210.8 Da. Area under the ROC curve (AUC) is 0.78

the used methodology were evaluated and described in our previous report by calculating inter-day and intra-day variability [22]. The average coefficient of variation (CV) for inter-day study was 20.0% and for intra-day study it was 6.9%. The combination of ZipTips technology and MALDI-TOF MS analysis allowed us to generate a total of 170 spectral components (m/z unique peaks) from the serum samples. Univariate statistical analysis based on Wilcoxon test identified 98 peaks as significantly different between the studied groups. Moreover, discriminatory power of the obtained peaks was further analyzed by calculating the ROC curve, which represents a graphical relation between sensitivity and specificity (Fig. 1). Based on univariate tests, discriminative ability of the detected peaks was examined (Additional file 2, Table S2).

Panels of multiple disease markers manifest more powerful discriminative abilities than a single uncorrelated marker. Therefore, three mathematical algorithms (SNN, GA and QC) were used in order to generate prediction models based on the training set with randomly selected samples (cancer patients $n = 33$ and healthy controls $n = 33$). Combinations of peaks used by these algorithms are shown in the Table 2. Six peaks (m/z) are present in more than one model. However, only the peak of 2082.75 Da occurs in all three discriminatory panels. Two parameters (recognition capability and cross validation) were calculated for all used discriminative models (Table 3). Cross validation of the established models reached 63.64% (SNN), 54.55% (GA) and 68.18% (QC), while recognition capability rates were 80.30% (SNN), 93.94% (GA) and 72.72% (QC). External validation was proceeded using independent data set (cancer patients $n = 11$ and healthy controls $n = 12$). The highest values of sensitivity (71.00%) and specificity (68.60%) were associated with SNN (Table 3). This model was composed of 25 different peaks. According to the univariate tests (Wilcoxon test and ROC curve) 10 of them revealed statistically significant variation between studied groups with p -values <0.02 and AUC in the range 0.67–0.78.

In order to identify the peaks that, according to statistical analyses, had the highest diagnostic efficacy (linear positive mode m/z 1505.24; 1945.38; 2023.17; 2082.73; 2116.08; 2210.80; 3158.75; 6560.82; 7567.69 and 7830.60 Da) (Table 4), serum samples were pretreated with ZipTips and examined using tandem mass spectrometry nano-LC-MALDI-TOF/TOF-MS/MS. The spectra were analyzed in the mass range of 700–3500 Da in the reflector mode, which requires using sufficient resolution. It enables proper baseline separation of the analyzed peaks and highly accurate determination of their mass [15]. Therefore, discriminative peaks with mass of m/z: 6560.82; 7567.69 and 7830.60 were not detected. The MS/MS analysis of precursor ions m/z

Table 2 Combinations of peaks (m/z) used in calculated algorithms (SNN, QC and GA)

SNN (Da)	QC (Da)	GA (Da)
1364.60	1466.76	1207.36
1419.84	1945.40	1229.32
1435.84	2082.75	1419.84
1505.28	2116.16	1509.45
1509.45	2604.30	1520.10
1538.10	3158.79	1639.82
1741.40	3507.55	1888.54
1897.69	4075.32	2082.75
1945.40	4112.79	4249.16
2023.30	4151.55	5753.62
2082.75	4209.96	
2116.16	4231.58	
2210.84	4249.16	
2453.10	4268.99	
2770.01	4282.73	
2863.35	4644.22	
3158.79	4663.57	
3192.56	4680.02	
3263.59	4712.28	
3284.04	4755.77	
3302.35	5065.16	
6631.04	6376.98	
7692.26	6395.79	
7767.39	6585.81	
8602.82		

1504.8231 and 2021.1246 resulted in identification of Complement C3 protein (CO3_HUMAN) based on the peptide sequence G.SPMYSIITPNILR.L and R.SSKITH RIHWESASLLR.S, respectively, with significant hit in the Mascot search. Another discriminative peak, precursor ion m/z 1943.9257, was identified according to the MS/MS fragmentation as sequence H.NLGHGHKHER DQGHGHQ.R with high score in the Mascot database to Kininogen-1 protein (KNG1_HUMAN). Identification of both precursors m/z 2083.0695 and 3156.5613

Table 3 Results of recognition capability, cross validation, sensitivity and specificity for discriminative models (SNN, QC and GA)

	SNN	QC	GA
Recognition capability (%)	80.30	72.73	93.94
Cross validation (%)	63.64	68.18	54.55
Sensitivity (%)	71.00	77.40	87.10
Specificity (%)	68.60	51.40	48.60

Table 4 The most discriminative peaks (m/z signals) according to Wilcoxon test (*p*-values), ROC curve (AUC) and mathematical model (SNN) with their identification

Mass (m/z)	<i>p</i> -value	AUC	Peptide sequence	Identification	Ref.
1505.24	0.00937	0.676	G.SPMYSIITPNILRL	CO3_HUMAN	[22, 24, 34, 35]
1945.38	0.00290	0.725	H.NLGHGHKHERDQGHGHQR	KNG1_HUMAN	[36–40]
2023.17	0.01280	0.667	R.SSKITHRIHWESASLLR.S	CO3_HUMAN	[22, 24, 34, 35]
2082.73	0.00056	0.767	P.GVLSSRQLGLPGPPDVPDHAA.Y	ITIH4_HUMAN	[37, 41–45]
2116.08	0.00171	0.736	Hypothetical TSSTSYNRGDSTFESKSY	Hypothetical FIBA_HUMAN	-
2210.80	0.00056	0.777	G.ISPFHEHAEVVFTANDSGPR.R	TTHY_HUMAN	[9, 46–48]
3158.75	0.00171	0.738	R.NVHSGSTFFKYLLQGAKIPKPEASFSPR.R	ITIH4_HUMAN	[37, 41–45]
6560.82	0.00689	0.680	-	-	-
7567.69	0.01180	0.671	-	-	-
7830.60	0.00451	0.700	-	-	-

allowed for obtaining the sequences: P.GVLSSRQLGLPGPPDVPDHAA.Y and R.NVHSGSTFFKYLLQGAKIPKPEASFSPR.R, respectively, with significant hit in the Mascot database to Inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4_HUMAN). The fragmentation of signal m/z 2210.0565 allowed us to identify the following peptide sequence: G.ISPFHEHAEVVFTANDSGPR.R. It gave a significant score in the Mascot search to trans-thyretin protein (TTHY_HUMAN). Despite our efforts to extend the time of nano-LC gradient to 150 min, the obtained MS/MS spectrum of signal m/z 2016.8865 was contaminated with other neighborhood fragments or contained some unknown modification. For these reasons, its identification failed. For the hypothetical peptide sequence: TSSTSYNRGDSTFESKSY of the m/z 2016.8865, no results in the Mascot search were obtained. However, homology to fibrinogen alpha chain isoform alpha-E preproprotein (FIBA_HUMAN) was shown according to the MS-BLAST database. Therefore, the identification of this peak would require further analysis.

Discussion

OC is the most deadly gynecological cancer [23]. Due to scarceness of symptoms and lack of effective screening tests, this disease remains undetected until advanced stages. Therefore, in an attempt to discover high sensitivity biomarkers, a rapid development of novel approaches is observed. In the literature, there are several studies focusing on plasma and serum proteomic patterns of ovarian cancer obtained by MALDI-TOF MS. In order to detect low abundance proteins and peptides, biomarker enrichment kits [24], immunodepletion [25] and magnetic beads [26] have been applied in OC studies. Due to significant impact of sample pretreatment on the MS spectra, the discriminative peaks proposed as candidates for OC biomarkers depend on efficiency of

the enrichment strategy. Thus, in the present study a solid phase extraction technology - micropipette tips ZipTips - was proposed as a depletion method with the aim of low molecular peptide/protein characterization of the serum protein-peptide profiles of the OC. The applied methodology constitutes an objective tool for identification of the OC indicators, which may contribute to understanding the pathological processes and may facilitate the development of both novel diagnostic tools and molecular targeted therapies.

The serum proteomic patterns of OC were obtained by MALDI-TOF MS. All spectra were analyzed using univariate tests including ROC curve and Wilcoxon test. However, according to the literature typing only single disease biomarker is not sufficient and multi-component combinations may lead to the design of new diagnostic tools characterized by significant sensitivity and specificity [27, 28]. Therefore, in this study discrimination models were calculated using three different mathematical algorithms. Differences in their combinations of components are caused by various calculation mechanisms [29]. SNN allows for an identification of the most characteristic spectra for all the studied groups. They are called prototypes and reflect prototypical samples of each class [22, 30]. GA is based on natural evolution, which enables selection of the most important variables. A cost function leads to significant class selection [29]. QC, a univariate sorting algorithm, calculates average peak areas for each class and stores them together with other data like *p*-values at defined peak positions. Prediction models are created based on weighted average derived from all peaks [30].

For the calculated models two parameters (cross validation and recognition capability) were determined. Discriminatory models tend to achieve better results on data on the basis of which they were originally constructed than on data derived from a

new set of samples [31]. Thus, cross validation might be insufficient to assess the power of the obtained multi-component panel. For that reason, external validation seems an important step in defining accuracy of a created model [32]. Nevertheless, some clinical studies focus only on the internal validation and leave behind the need of the external validation [31]. In this study, independent test sets were used to evaluate robustness of the prediction models. Due to difficulties in selecting a model with the best diagnostic efficacy based on recognition capability and cross validation, external validation was performed for all three classification algorithms (Table 3). The best differentiating capabilities and satisfactory values of sensitivity (71.00%) and specificity (68.60%) were associated with SNN.

Protein-peptide profiling studies based on MALDI-TOF MS analyses are often limited to a list of the most discriminative *m/z* peaks as potential disease indicators [15, 25, 26]. Nevertheless, the identification step is crucial for understanding the pathological processes that occur during cancer development and it should not be omitted. However, protein identification using MALDI-TOF without a digestion step might be challenging. Thus, the subject literature contains reports on combining MALDI-TOF profiling with other mass spectrometry platforms [24, 33]. This study proposes a novel approach that enables protein-peptide profiling as well as identification of clusters of ions with diagnostic capability using tandem mass spectrometry nano-LC-MALDI-TOF/TOF-MS/MS.

Application of developed strategy allowed for the identification of four potential OC serum biomarkers (complement C3, kininogen-1, inter-alpha-trypsin inhibitor heavy chain H4, and transthyretin). Complement C3 plays a key role in both immunological and inflammatory processes. Recent findings suggest that it may promote tumor growth, angiogenesis, cellular proliferation and regeneration [34]. Thus, a new concept of cancer treatment based on blocking the complement system was proposed [35]. Moreover, a number of studies reported that patients with cancer (including OC) produce altered levels of complement C3 as compared with healthy subjects [22, 24]. It might be caused by an inflammatory response to the tumor development. However, this marker should be further validated with the use of controls from inflammatory conditions.

Another identified protein – kininogen-1 takes part in blood coagulation and in the kinin-kallikrein system. It shows antiangiogenic properties and it also inhibits proliferation of endothelial cells. The role of this protein in cancer development might be associated with survival of the cancer cells [36, 37]. Changes in the levels of

kininogen-1 were observed in urine samples of OC patients [37]. Its expression was altered in the serum and plasma in patients with proliferative vitreoretinopathy [38] and colorectal cancer [36]. Moreover, other diseases like interstitial cystitis [39] or IgA nephropathy [40] are also related to non-standard urine concentrations of kininogen-1.

Protein also identified as potential OC marker is ITIH4, which belongs to the inter-alpha-trypsin inhibitor (ITI) family and it is an acute-phase reactant [41]. There are a few reports that proposed this protein as an OC marker [37, 42]. Changes in the levels of ITIH4-derived peptides were also observed in urine of early prostate cancer patients [43] and in the serum of breast cancer patients [44] and gastric adenocarcinoma patients [41]. A correlation was also suggested between different fragmentation of ITIH4 and disease conditions [45].

The last protein proposed as discriminatory marker of OC is transthyretin. Transthyretin plays an essential role in a transport of thyroxine and tri-iodothyronine. It also takes part in the transfer of retinol. Differences in the expression of transthyretin during OC development were already reported [46, 47]. Moreover, changes in the cellular retinol binding protein levels in OC patients were observed [48]. What is worth emphasizing, transthyretin is one of the five biomarkers the concentrations of which are measured in OVA1 multivariate index assay [9].

The applied methodology allowed us to identify four different serum proteins (Table 4) with an essential role in OC development, according to the literature. Complement C3, inter-alpha-trypsin inhibitor heavy chain H4 and transthyretin were also identified using a combination of carrier protein-bound affinity enrichment strategy with MALDI-TOF MS to characterize OC samples, which is in agreement with our findings [24]. Unfortunately, other OC studies based on MALDI-TOF MS profiling lack the identification step [25, 26]. It is unknown whether other depletion methods are capable of detecting relevant proteins.

Naturally, MALDI-TOF MS is a very sensitive technique of qualitative analysis [49] that should be complemented with a quantitative approach to confirm the results of peptide-protein profiling. Thus, the clinical utility of complement C3, kininogen-1, inter-alpha-trypsin inhibitor heavy chain H4 and transthyretin should be examined by quantitative analysis in a larger set of samples. Moreover, further studies are planned to identify the remaining discriminative peaks since they might extend our knowledge on the pathological processes that occur during OC.

Conclusions

To conclude, proteomic profiling of serum samples based on the solid phase extraction enrichment technology coupled with MALDI-TOF MS demonstrated differences

in the serum protein expression in patients with OC compared with the healthy control group. The SNN classification algorithm yielded a discriminative model characterized by significant sensitivity (71%) and specificity (68.6%) in the external validation. The novel approach, which enabled protein-peptide profiling as well as identification of four potential OC biomarkers (complement C3, kininogen-1, inter-alpha-trypsin inhibitor heavy chain H4 and transthyretin) using MALDI-TOF MS, may contribute to the creation of new effective multi-component diagnostic tools. Additionally, a panel of discriminative proteins could provide an explanation of complex signaling and functional networks associated with this multifactorial disease.

Additional files

Additional file 1: Table S1. Study group characterization according to histopathological type and FIGO stage at diagnosis. (DOCX 12 kb)

Additional file 2: Table S2. Masses (m/z) and intensities of the peaks with the highest values of the univariate statistical tests: Wilcoxon test and the ROC curve (p -value <0.05; AUC > 0.7). (DOCX 14 kb)

Abbreviations

ACN: acetonitrile; AUC: area under the ROC curve; CA125: cancer antigen 125; CO3_HUMAN: Complement C3 protein; CV: coefficient of variation; FDA: Food and Drug Administration; FIBA_HUMAN: fibrinogen alpha chain isoform alpha-E preproprotein; GA: genetic algorithm; HCCA: α -cyano-4-hydroxycinnamic acid; HE4: human epididymis protein 4; ITH4_HUMAN: Inter-alpha-trypsin inhibitor heavy chain H4; KNG1_HUMAN: Kininogen-1 protein; MALDI-TOF-MS: Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry; MIA: multivariate index assay; nano-LC: nano-liquid chromatography; OC: Ovarian carcinoma; QC: quick classifier; ROC: receiver operating characteristic; ROMA: Risk of Ovarian Malignancy Algorithm; SNN: supervised neural network; TFA: trifluoroacetic acid; TIC: total ion current; TTHY_HUMAN: transthyretin protein

Acknowledgements

Not applicable.

Funding

The project was supported by the Polish National Science Centre (2014/15/B/NZ7/00964). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article (and its additional files). More datasets of the current study are available from the corresponding author on reasonable request.

Authors' contributions

A.S. and A.H. contributed equally to this work as first authors. A.S., A.H., E.N.M. and Z.J.K. designed the research. A.S., J.H. and J.M. performed the experiments. A.H. and E.N.M. contributed important samples. A.S., A.H., J.H. collected the data. A.S., A.H., J.H., J.M. analyzed the data. A.S., A.H., J.H., E.N.M. and Z.J.K. wrote the manuscript. All authors read and approved the submitted manuscript.

Ethics approval and consent to participate

The study was approved by the Bioethics Committee of Poznan University of Medical Sciences, Poland (Decision No. 165/16). All participants of the study

signed informed consent to publish their (anonymized) data for scientific purposes.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Inorganic and Analytical Chemistry, Poznan University of Medical Sciences, ul. Grunwaldzka 6, 60-780 Poznań, Poland. ²Gynecologic Oncology Department, Poznan University of Medical Sciences, ul. Polna 33, 60-535 Poznań, Poland.

Received: 26 October 2016 Accepted: 30 June 2017

Published online: 06 July 2017

References

- Das PM, Bast RC. Early detection of ovarian cancer. *Biomark Med*. 2008;2:291–303.
- Havrilesky LJ, Sanders GD, Kulasingam S, Myers ER. Reducing ovarian cancer mortality through screening: is it possible, and can we afford it? *Gynecol Oncol*. 2008;111:179–87.
- Romagnolo C, Leon AE, Fabricio ASC, Taborelli M, Polesel J, Del Pup L, et al. HE4, CA125 and risk of ovarian malignancy algorithm (ROMA) as diagnostic tools for ovarian cancer in patients with a pelvic mass: an Italian multicenter study. *Gynecol Oncol*. 2016;141:303–11.
- Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *Lancet*. 2014;384:1376–88.
- Terlikowska KM, Dobrzycka B, Witkowska AM, Mackowiak-Matejczyk B, Sledziwski TK, Kinalski M, et al. Preoperative HE4, CA125 and ROMA in the differential diagnosis of benign and malignant adnexal masses. *J Ovarian Res*. 2016;9:43.
- Karlan BY, Thorpe J, Watabayashi K, Drescher CW, Palomares M, Daly MB, et al. Use of CA125 and HE4 serum markers to predict ovarian cancer in elevated-risk women. *Cancer Epidemiol Prev Biomarkers*. 2014;23
- Chang X, Ye X, Dong L, Cheng H, Cheng Y, Zhu L, et al. Human epididymis protein 4 (HE4) as a serum tumor biomarker in patients with ovarian carcinoma. *Int J Gynecol Cancer*. 2011;21:852–8.
- Moore RG, Miller MC, Eklund EE, Lu KH, Bast RC, Lambert-Messerlian G. Serum levels of the ovarian cancer biomarker HE4 are decreased in pregnancy and increase with age. *Am J Obstet Gynecol*. 2012;206:349–7.
- Eskander RN, Carpenter BA, Wu HG, Wolf JK. The clinical utility of an elevated-risk multivariate index assay score in ovarian cancer patients. *Curr Med Res Opin*. 2016;32:1161–5.
- Moore RG, Miller MC, Disilvestro P, Landrum LM, Gajewski W, Ball JJ, et al. Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. *Obstet Gynecol*. 2011;118:280–8.
- Elzek MA, Rodland KD. Proteomics of ovarian cancer: functional insights and clinical applications. *Cancer Metastasis Rev*. 2015;34:83–96.
- Widlak P, Pietrowska M, Polanska J, Marczyk M, Ros-Mazurczyk M, Dziadziuszko R, et al. Serum mass profile signature as a biomarker of early lung cancer. *Lung Cancer*. 2016;99:46–52.
- Opstal-van Winden AWJ, Krop EJM, Käredal MH, Gast M-CW, Lindh CH, Jeppsson MC, et al. Searching for early breast cancer biomarkers by serum protein profiling of pre-diagnostic serum; a nested case-control study. *BMC Cancer*. 2011;11:381.
- Yang J, Yang J, Gao Y, Zhao L, Liu L, Qin Y, et al. Identification of potential serum proteomic biomarkers for clear cell renal cell carcinoma. *PLoS One*. 2014;9:e111364.
- Hajduk J, Matysiak J, Kokot ZJ. Challenges in biomarker discovery with MALDI-TOF MS. *Clin Chim Acta*. 2016;458:84–98.
- Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*. 2002;359:572–7.

17. Sandanayake NS, Camuzeaux S, Sinclair J, Blyuss O, Andreola F, Chapman MH, et al. Identification of potential serum peptide biomarkers of biliary tract cancer using MALDI MS profiling. *BMC Clin Pathol*. 2014;14:7.
18. Callesen AK, Madsen JS, Vach W, Kruse TA, Mogensen O, Jensen ON. Serum protein profiling by solid phase extraction and mass spectrometry: A future diagnostics tool. *Proteomics*. 2009;14:28–41.
19. Bellei E, Bergamini S, Monari E, Fantoni LI, Cuoghi A, Ozben T, et al. High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids*. 2011;40:145–56.
20. Hajduk J, Matysiak J, Kokot P, Nowicki P, Dereziński P, Kokot ZJ. The application of fuzzy statistics and linear discriminant analysis as criteria for optimizing the preparation of plasma for matrix-assisted laser desorption/ionization mass spectrometry peptide profiling. *Clin Chim Acta*. 2015;448:174–81.
21. Tiss A, Smith C, Camuzeaux S, Kabir M, Gayther S, Menon U, et al. Serum peptide profiling using MALDI mass spectrometry: avoiding the pitfalls of coated magnetic beads using well-established ZipTip technology. *Proteomics*. 2007;7(Suppl 1):77–89.
22. Klupczynska A, Swiatly A, Hajduk J, Matysiak J, Dyszkiewicz W, Pawlak K, et al. Identification of serum peptidome signatures of non-small cell lung cancer. *Int J Mol Sci*. 2016;17.
23. Marcus CS, Maxwell GL, Darcy KM, Hamilton CA, Mcguire WP. Current approaches and challenges in managing and monitoring treatment response in ovarian cancer. *J Cancer*. 2014;5:25–30.
24. Lopez MF, Mikulskis A, Kuzdzal S, Golenko E, Petricoin EF, Liotta LA, et al. A novel, high-throughput workflow for discovery and identification of serum carrier protein-bound peptide biomarker candidates in ovarian cancer samples. *Clin Chem*. 2007;53:1067–74.
25. Periyasamy A, Gopisetty G, Veluswami S, Joyimallaya Subramaniam M, Thangarajan R. Identification of candidate biomarker mass (m/z) ranges in serous ovarian adenocarcinoma using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry profiling. *Biomarkers*. 2015;20:292–8.
26. Wu S, Xu K, Chen G, Zhang J, Liu Z, Xie X. Identification of serum biomarkers for ovarian cancer using MALDI-TOF-MS combined with magnetic beads. *Int J Clin Oncol*. 2012;17:89–95.
27. Yin X, Subramanian S, Hwang S-J, O'Donnell CJ, Fox CS, Courchesne P, et al. Protein biomarkers of new-onset cardiovascular disease. *Arterioscler Thromb Vasc Biol*. 2014;34(4):939–45.
28. Horala A, Swiatly A, Matysiak J, Banach P, Nowak-Markwitz E, Kokot Z. Diagnostic value of serum angiogenesis markers in ovarian cancer using multiplex immunoassay. *Int J Mol Sci*. 2017;18:123.
29. Basics on data preparation model generation and spectra classification. In: *ClinProTools 3.0 Software for biomarker detection and evaluation User Manual*. Bruker Daltonics; Bremen, Germany: 2011.; p. 51–96.
30. Zhang Y, Chen X, Hu Y, Du S, Shen L, He Y, et al. Preliminary characterizations of a serum biomarker for sarcoidosis by comparative proteomic approach with tandem-mass spectrometry in ethnic Han Chinese patients. *Respir. Res. BioMed Central*; 2013;14:18.
31. Bleeker S, Moll H, Steyerberg E, Donders AR, Derksen-Lubsen G, Grobbee D, et al. External validation is necessary in prediction research: A clinical example. *J Clin Epidemiol*. 2003;56:826–32.
32. Ioannidis JPA. A roadmap for successful applications of clinical PROTEOMICS. *PROTEOMICS - Clin. Appl*. 2011;5:241–7.
33. Chinello C, Cazzaniga M, De Sio G, Smith AJ, Gianazza E, Grasso A, et al. Urinary Signatures of Renal Cell Carcinoma Investigated by Peptidomic Approaches. Eberini I, editor. *PLoS One*. 2014;9:e106684.
34. Rutkowski MJ, Sughrue ME, Kane AJ, Mills SA, Parsa AT, Balkwill F, et al. Cancer and the complement cascade. *Mol Cancer Res*. 2010;8:1453–65.
35. Pio R, Ajona D, Lambris JD. Complement inhibition in cancer therapy. *Semin Immunol*. 2013;25:54–64.
36. Wang J, Wang X, Lin S, Chen C, Wang C, Ma Q, et al. Identification of Kininogen-1 as a serum biomarker for the early detection of advanced colorectal adenoma and colorectal cancer. *PLoS One*. 2013;8:e70519.
37. Abdullah-Soheimi SS, Lim B-K, Hashim OH, Shuib AS. Patients with ovarian carcinoma excrete different altered levels of urine CD59, kininogen-1 and fragments of inter-alpha-trypsin inhibitor heavy chain H4 and albumin. *Proteome Sci*. 2010;8:58.
38. Yu J, Peng R, Chen H, Cui C, Ba J, Wang F. Kininogen 1 and insulin-like growth factor binding protein 6: candidate serum biomarkers of proliferative vitreoretinopathy. *Clin Exp Optom*. 2014;97:72–9.
39. Canter MP, Graham CA, Heit MH, Blackwell LS, Wilkey DW, Klein JB, et al. Proteomic techniques identify urine proteins that differentiate patients with interstitial cystitis from asymptomatic control subjects. *Am. J. Obstet. Gynecol*. 2008;198:553.e1–6.
40. Rocchetti MT, Centra M, Papale M, Bortone G, Palermo C, Centonze D, et al. Urine protein profile of IgA nephropathy patients may predict the response to ACE-inhibitor therapy. *Proteomics*. 2008;8:206–16.
41. Subbannaya Y, Mir SA, Renuse S, Manda SS, Pinto SM, Puttamalles VN, et al. Identification of differentially expressed serum proteins in gastric adenocarcinoma. *J Proteome*. 2015;127:80–8.
42. Mohamed E, Abdul-Rahman PS, Doustjalali SR, Chen Y, Lim B, Omar SZ, et al. Lectin-based electrophoretic analysis of the expression of the 35 kDa inter-alpha-trypsin inhibitor heavy chain H4 fragment in sera of patients with five different malignancies. *Electrophoresis*. 2008;29:2645–50.
43. Jayapalan JJ, Ng KL, Shuib AS, Razack AHA, Hashim OH. Urine of patients with early prostate cancer contains lower levels of light chain fragments of inter-alpha-trypsin inhibitor and saposin B but increased expression of an inter-alpha-trypsin inhibitor heavy chain 4 fragment. *Electrophoresis*. 2013;34:1663–9.
44. van den Broek I, Sparidans RW, van Winden AWJ, Gast M-CW, van Dulken EJ, Schellens JHM, et al. The absolute quantification of eight inter-alpha-trypsin inhibitor heavy chain 4 (ITI4)-derived peptides in serum from breast cancer patients. *Proteomics Clin Appl*. 2010;4:931–9.
45. Song J, Patel M, Rosenzweig CN, Chan-Li Y, Sokoll LJ, Fung ET, et al. Quantification of fragments of human serum inter-alpha-trypsin inhibitor heavy chain 4 by a surface-enhanced laser desorption/ionization-based immunoassay. *Clin Chem*. 2006;52:1045–53.
46. Moore LE, Fung ET, McGuire M, Rabkin CC, Molinaro A, Wang Z, et al. Evaluation of apolipoprotein A1 and Posttranslationally modified forms of transthyretin as biomarkers for ovarian cancer detection in an independent study population. *Cancer Epidemiol Biomark Prev*. 2006;15:1641–6.
47. Husseinzadeh N. Status of tumor markers in epithelial ovarian cancer has there been any progress? A review *Gynecol Oncol*. 2011;120:152–7.
48. Doldo E, Costanza G, Agostinelli S, Tarquini C, Ferlosio A, Arcuri G, et al. Vitamin a, cancer treatment and prevention: the new role of cellular retinol binding proteins. *Biomed Res Int*. 2015;2015:624627.
49. Liu C, Shen J, Pan C, Yang L, Mou S, Wang H, et al. MALDI-TOF MS combined with magnetic beads for detecting serum protein biomarkers and establishment of boosting decision tree model for diagnosis of hepatocellular carcinoma. *Am J Clin Pathol*. 2010;134:235–41.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

