Published in final edited form as:

Proteomics. 2011 December ; 11(24): 4677–4687. doi:10.1002/pmic.201000811.

Identification of glycoproteins associated with different histological subtypes of ovarian tumors using quantitative glycoproteomics

Yuan Tian¹, Zhihao Yao², Richard Roden¹, and Hui Zhang¹

¹Department of Pathology, Johns Hopkins University, Baltimore, MD, 21231

²Merck Serono, Beijing, China 10022

Abstract

Ovarian cancer is the most lethal gynecologic malignancy in adult women. The origin of epithelial ovarian tumors is both morphologically and biologically heterogeneous, and different subtypes of ovarian tumors have different clinical outcomes. In spite of the heterogeneous nature of ovarian carcinoma, the current biomarkers and treatments for this disease are not subtype-specific. To discover the molecular basis of the ovarian tumor subtypes, we analyzed extracellular glycoproteins of seven common subtypes and normal ovary tissues using quantitative glycoproteomic analysis. Glycoproteins for different ovarian tumor subtypes were identified by liquid chromatography-tandem mass spectrometry and quantitated by spectral counting and then verified by iTRAQ labeling and Western blotting. Glycoproteins uniquely expressed in different subtypes of ovarian tumors or commonly expressed in most subtypes were identified. Using Western blots, we verified that mesothelin was overexpressed in serous carcinoma and transitional-cell carcinoma, CEA5 and CEA6 were overexpressed only in mucinous carcinoma, while versican and periostin were overexpressed in most subtypes of ovarian tumors. This study represents the first proteomic characterization of different ovarian tumor subtypes. The identified glycoproteins for histological subtypes of ovarian tumors will facilitate the understanding of the molecular basis, diagnosis of ovarian tumor subtypes, and predictions for treatment responses to therapeutic agents.

Keywords

ovarian tumor; histological subtypes; quantitative glycoproteomics; mass spectrometry; Western blot

Introduction

Ovarian cancer is the most lethal gynecologic malignancy in adult women and represents 30% of cancers of the female genital tract [1]. Chemotherapy is the common treatment for ovarian cancer for the past several decades, but the overall survival rate of women with this disease has not much improved due to several factors: 1) There are no simple preventive measures to significantly reduce the risk for developing ovarian cancer. 2) There is no reliable screening test for the early detection of ovarian cancer; thus, approximately two-thirds of women with epithelial ovarian tumors already have advanced disease at diagnosis.

Requests for reprints: Hui Zhang, Department of Pathology, Johns Hopkins University, 1550 Orleans Street, CRBII, Room 3M-03, Baltimore, MD 21231. Phone: (410)502-8149; hzhang32@jhmi.edu.

The authors have declared no conflict of interest.

3) The prognosis for women with advanced ovarian cancer is very poor, with a 5-year overall survival rate of 30% [2, 3].

Epithelial ovarian tumors are morphologically and biologically heterogeneous: They can be histologically subclassified into subtypes including high grade and low grade serous, mucinous, high grade and low grade endometrioid, clear-cell, transitional-cell, squamous-cell, mixed, and undifferentiated subtypes [4]. Serous carcinomas represent the majority of ovarian tumors (~53%) and have the lowest five-year survival rate (20–35%) [5–7]. Mucinous and endometrioid ovarian tumors each represent ~10% of epithelial ovarian cancers and have five-year survival rates of 40–60% while clear-cell tumors represent ~5% of epithelial ovarian tumors and have a five-year survival rate of 35–50% [7–9].

It is increasingly recognized that different histological subtypes of ovarian tumors have different responses to treatments, and that low-grade serous, mucinous, and clear-cell carcinomas are intrinsically resistant to standard chemotherapeutic agents [10–13]. Tumor biology appears to vary among ovarian tumor subtypes [14–17]. The molecular study supports the notion that different subtypes likely represent distinct diseases [16]. Therefore, the identification of specific markers for different histology subtypes of ovarian cancer is necessary for the diagnosis of ovarian tumors. A multiple-marker panel consisting of proteins from different subtypes can also be developed to detect most subtypes and potentially improve the specificity and sensitivity for ovarian cancer diagnosis as well as predict the outcomes for different treatments.

Extracellular proteins, including cell surface proteins, transmembrane proteins, and secreted proteins, are mostly glycosylated and account for about one-third of total human proteins. Therefore, glycoproteomic analysis will target the extracellular proteins. Since they are located outside of the cell, these proteins are easily accessible by therapeutic reagents and molecular imaging probes. The cancer-associated extracellular proteins, likely secreted by cancer cells or shed from the cell surface, enter the bloodstream, representing a rich source of potential disease markers for blood test. In the present study, extracellular glycoproteins from seven major subtypes of ovarian tumors were analyzed using quantitative glycoproteomic technology, and the extracellular protein profiles were compared to those of normal ovary tissues to identify glycoproteins that are differentially expressed in different ovarian tumor subtypes. The identified ovarian tumor-specific proteins were further verified by Western blot.

Materials and methods

Materials

Hydrazide resin and sodium periodate were from Bio-Rad (Hercules, CA); sequencing-grade trypsin was from Promega (Madison, WI); PNGase F was from New England Biolabs (Ipswich, MA); C18 columns were from Waters (Milford, MA); the mouse anti-CEA5/CEA6 antibody (CEACAM 1,5,6,8) was from ABR Affinity BioReagents (Golden, CO); the mouse anti-mesothelin antibody [K1], rabbit anti-versican antibody, rabbit anti-periostin antibody, and mouse anti-LGALS3BP (galectin-3-binding protein) antibody were from Abcam (Cambridge, UK); and the BCA assay kit, HRP-labeled secondary antibodies, and Novex ECL Chemiluminescent Substrate Reagent Kit were from Pierce (Rockford, IL). All other chemicals were from Sigma-Aldrich (St. Louis, MO).

Tissue samples and clinical information were obtained with informed consent. This study was performed with the prior approval of the Johns Hopkins Medicine Institutional Review Board. Fresh ovarian tumors and adjacent normal ovary tissues were obtained by surgery at the Johns Hopkins Hospital. These tissues were stored at -80°C until used.

Glycopeptide isolation

Frozen tissue (~100 mg) was sliced into 1–3-mm-thick sections, vortexed in 100 μ L of 5-mM phosphate buffer for 2–3 min, and homogenized by sonication in an ice-water bath for 5 min. The resulting tissue homogenate was incubated in 100 μ L of trifluoroethanol (TFE) for 2 h at 60°C, followed by sonication for 2 min to denature the proteins. Protein disulfide bonds in the tissues were reduced by a 30-min incubation in 5-mM tributylphosphine (TBP) at 60°C and then alkylated by a 30-min incubation in 10-mM iodoacetamide at room temperature (RT) in the dark. Finally, the proteins were diluted 5-fold with 50-mM NH_4HCO_3 (pH 7.8) and digested with trypsin (enzyme-to-protein ratio of 1:50 w/w) overnight at 37°C with gentle shaking.

As a control experiment, a small portion of proteins (5 μ g) before and after tryptic digestion was resolved on a protein gel and silver-stained to determine the completeness of trypsin digestion. The complete digestion of the proteins was indicated by the disappearance of high-molecular-weight protein bands and the appearance of low-molecular-weight (<10 kDa) peptide bands. After tryptic digestion, the samples were centrifuged at 13,000 rpm for 5 min to remove any particulate matter, and the peptide concentrations were measured by BCA assay. *N*-linked glycopeptides were isolated from 2 mg of peptides by solid-phase extraction of glycopeptides (SPEG) [18, 19]. Briefly, the peptides were oxidized by 10 mM sodium periodate in 5 % ACN in 0.1 % TFA with 1-hour incubation at RT. The oxidized samples were applied to C18 columns which were preconditioned with 800 μ L of 80 % ACN in 0.1 % TFA, then 800 μ L 0.1 % TFA, and the peptides were eluted with 800 μ L of 80 % ACN in 0.1 % TFA. The 50 μ L of (50 % slurry) hydrazide resins were prewashed with 1 mL deionized water then added to the peptide mixture. The hydrazide resins were incubated with the oxidized samples overnight for coupling reaction at RT with gentle shaking. The resins were then washed three times with 800 μ L of 1.5 M NaCl followed by three times with 800 μ L of water to remove the non-glycosylated peptides. 2 μ L of PNGase F (500,000 units/mL, New England Biolabs) was added and incubated at 37 °C overnight with shaking. The supernatant was cleaned and concentrated using C18 columns as described above and resuspended in 40 μ L of 0.4% acetic acid prior to iTRAQ labeling and/or LC-MS/MS analysis.

iTRAQ labeling of peptides

Formerly *N*-linked glycopeptides (10 μ L/sample) were labeled with iTRAQ 8plex (ABI) according to the manufacturer's instructions: The peptides were dried and resuspended in 20 μ L of dissolution buffer provided in the iTRAQ kit. Each iTRAQ 8-plex reagent was dissolved in 60 μ L of isopropanol, vortexed for 1 min, and then added to the corresponding sample as follows: reagents 113, 114, 115, 116, 117, 118, 119, and 121 were added to glycopeptides of normal ovary, clear-cell carcinoma, high-grade endometrioid carcinoma, high-grade serous carcinoma, low-grade endometrioid carcinoma, low-grade serous carcinoma, mucinous carcinoma, and transitional carcinoma, respectively. Each mixture was incubated at RT for 2 h. Labeled peptides from different samples were mixed and then purified using a strong cation exchange column.

LC-MS/MS analysis

Formerly *N*-linked glycopeptides (~1.5 μ g) were analyzed using a linear ion trap mass spectrometer (LTQ, Thermo Fisher, Waltham, MA) after separation with a 15 cm $\times$ 75 μ m C18 column (5 μ m particles with 100 angstrom pore size). A nanoAquity UPLC at 300 nL/min with a 100-minute linear acetonitrile gradient (from 5–32% B over 100 min; A= 0.1% formic acid in water, B=0.1% formic acid in acetonitrile) was used. Top 8 data dependent MS/MS spectra with exclusion for 20 seconds and a repeat count of 2 were set. The exclusion window was –1 Da to +1.5 Da, and the isolation width for precursors was set to 2

Da. CID in the ion trap was used with a collision energy setting of 35%. The voltage was set at 2.0kV. Each sample was analyzed three times to identify and quantify the formerly *N*-linked glycopeptides using spectral counting (see data analysis below).

iTRAQ-labeled former glycopeptides (~3 µg) were analyzed by LC-MS/MS using an LTQ-Orbitrap velos (Thermo Fisher, Waltham, MA) coupled with the same C18 column described above. A nano Aquity UPLC at 300 nL/min with a 90-minute linear acetonitrile gradient (from 5–32% B over 90 min; A= 0.1% formic acid in water, B=0.1% formic acid in acetonitrile) was used. Top 10 data dependent MS/MS spectra with exclusion for 20 seconds were set. The samples were run with HCD fragmentation at normalized collision energy of 45 and an isolation width of 1.2 Da. Monoisotopic Precursor Selection (MIPS) was enabled and the dynamic exclusion was set to 30 sec with a repeat count of 1 and +/- 10ppm mass window. Source voltage was 2.0 kV. A lock mass of the polysiloxane peak at 371.10123 was used to correct the mass in MS and MS/MS. Target values were 1e6 ions at a resolution setting of 30,000 in MS and 1e5 ions at a resolution setting of 7,500 in MS2.

Peptide and protein identifications

Raw MS data from LTQ were converted into mzXML files by a MassWolf file converter [20]. MS/MS spectra were searched with SEQUEST [21] against a human International Protein Index (IPI, version 2.28) database containing 40,110 entries. For this database search, the peptide mass tolerance was set at 3.0 Da, MS/MS tolerance was ~0.5 Da, and flexible parameters were set as follows: cysteine modification (add cysteine with 57 Da), methionine oxidation (add Met with 16 Da), and a (PNGase F-catalyzed) conversion of Asn to Asp (add Asn with 1Da). One missed tryptic end and a maximum of 2 missed cleavage sites were permitted. The assigned peptides were evaluated by Peptide Prophet, and only peptides with a minimum probability score of 0.8 (with error rate less than 0.027) were reported in this study [22, 23].

MS/MS analysis of iTRAQ labeled peptides using LTQ-Orbitrap was searched with MASCOT (version 2.2.0) using Proteome Discoverer (version 1.0) (Thermo Fisher) against human subdatabase of NCBI Reference Sequence (RefSeq) (version 40, released at April 16, 2010) containing 29,704 sequences. Integration window tolerance was set at 20 ppm for peak integration. Peptide cut off score was set at 10. For this database search, the precursor mass tolerance and fragment mass tolerance was set at 15ppm and 0.05 Da respectively, and other database-searching parameters were set as flexible modifications as follows: oxidized methionines (add Met with 15.99 Da), a (PNGase F-catalyzed) conversion of Asn to Asp (add Asn with 0.984 Da), and cysteine modification (add Cys with 57.02 Da). The False Discovery Rate was set at 0.01 so that low-probability protein identifications could be filtered out.

Subcellular location of identified proteins

Signal peptides were predicted using SignalP 2.0 [24]. Transmembrane (TM) regions were predicted using TMHMM (version 2.0) [25]. The TMHMM program predicts protein topology and the number of TM helices. Information from SignalP and TMHMM were combined to separate proteins into the following categories: i) *cell surface* - proteins that contain predicted non-cleavable signal peptides but no predicted transmembrane segments; ii) *secreted* - proteins that contain predicted cleavable signal peptides but no predicted transmembrane segments; iii) *transmembrane* - proteins that contain predicted transmembrane segments and extracellular loops and intracellular loops; and iv) *intracellular* - proteins that contain neither predicted signal peptides nor predicted transmembrane regions. All protein sequences were taken from IPI human protein database (version 2.28).

Protein quantitation

Spectral counting, a label-free quantitation method by calculating the number of MS/MS spectra [26], was used to analyze the LC-MS/MS data obtained from LTQ to determine the protein relative abundance in different subtypes of tumors and normal ovarian tissues as described previously [27]. The statistical analysis of spectral counting data was performed using our previously described procedure [27]. In addition, quantitation of iTRAQ labeled peptides from the same set of ovarian tumors and normal tissues were achieved by LTQ-Orbitrap and Proteome Discoverer software (version 1.0) from Thermo Fisher.

Western blot analysis

Proteins (20 µg) were resolved by SDS-PAGE and transferred electrophoretically onto a nitrocellulose membrane. The membrane was blocked with 5% non-fat milk/0.1% TBS-Tween 20 at RT for 2 h and then probed with primary antibody (mouse anti-CEA5/CEA6 monoclonal antibody at 1:1000; mouse anti-mesothelin monoclonal antibody at 1:1000, rabbit anti-versican polyclonal antibody at 1:1000, rabbit anti-periostin polyclonal antibody at 1:1000, and mouse anti-LGALS3BP (galectin-3-binding protein) monoclonal antibody at 1:500) at 4°C overnight, followed by three washes with 0.1% TBS-Tween 20. HRP-conjugated secondary antibody was added at 1:2000 and incubated at RT for 1 h, followed by three washes with 0.1% TBS-Tween 20. The signal was visualized using SuperSignal Substrate (Pierce). The densitometry analysis of the western blot bands was performed and normalized to beta-actin.

Results

Quantitative analysis of glycoproteins from different subtypes of ovarian tumors

To identify extracellular proteins commonly or uniquely expressed in different ovarian tumor subtypes, we performed quantitative glycoproteomic analysis (Figure 1) based on the fact that extracellular proteins are mostly glycosylated. SPEG [18, 19] was used to isolate formerly *N*-linked glycopeptides from 3 cases of normal ovary tissues and 3 cases from each of the seven major subtypes of human ovarian tumors, including high-grade serous, low-grade serous, mucinous, high-grade endometrioid, low-grade endometrioid, clear-cell, and transitional-cell carcinomas. Glycopeptides were isolated from three cases of each subtype and combined prior to being analyzed by mass spectrometry and each sample was analyzed in three technical replicates. The detailed information of specimens is listed in Supplementary Table 1. The relative abundance of glycoproteins was normalized by β -actin to identify the protein changes associated with different subtypes of ovarian tumors.

With a minimum probability of 0.8 (less than 0.027 of error rate), 959 out of 1037 peptide identifications contained consensus *N*-linked glycosylation motif (NXS/T, where X is any amino acid except P, ~92.48%), which resulted in 368 unique *N*-linked glycosites (containing consensus *N*-linked glycosylation NXS/T motif), representing 286 unique glycoproteins (Supplementary Tables 2). We used the number of spectra assigned to each glycosite of glycoprotein to determine the relative abundance of glycosites from different subtypes of ovarian tumors and normal tissues [26, 27]. The results represent the glycoprotein changes.

From the quantitative proteomic data, we are interested in identifying protein changes commonly altered in most ovarian tumor subtypes comparing to normal ovary tissues as well as those uniquely altered in specific subtypes of ovarian tumor. Proteins commonly altered in most ovarian tumor subtypes but not in normal tissues, might be useful in detecting ovarian tumors with improved sensitivity and specificity. Proteins uniquely altered in specific subtypes of ovarian tumor might be useful for diagnosis of ovarian tumor

subtypes. Table 1 shows the ratio of spectral counts of ovarian cancer subtypes to normal tissue. The ratio of the proteins which were identified only in ovarian cancer but not in normal tissues was arbitrary assigned to 100. Eleven proteins including versican, periostin, desmoglein2, 150 kDa oxygen-regulated protein, and tetraspanin 1, were identified showing increased expression in most ovarian tumor subtypes but not in normal tissues (Table 1 and Supplementary Table 2). Additional 13 proteins were identified as uniquely overexpressed in specific ovarian tumor subtypes; among these were carcinoembryonic antigen-related cell adhesion molecule 5 (CEA 5) and CEA 6 for mucinous carcinoma, mesothelin for high-grade serous, low-grade serous, and transitional carcinomas, and integrin alpha-M for high-grade endometrioid and high-grade serous carcinomas (Table 1 and Supplementary Table 2, Figure 2a, 2b, 2c and 2d).

Orthogonal quantitative proteomic approach can be used as a high-content method to verify the changes identified with spectral counting method as well as to identify and quantify additional protein changes [27]. The captured glycopeptides from tissues of seven different ovarian tumor subtypes and normal ovary were labeled with 8-plex iTRAQ reagents and quantitatively analyzed by LTQ-Orbitrap. The ratio of the protein level in each subtype to the protein level in normal ovary was calculated, and proteins with at least a 2-fold change were considered to have altered expression. Several proteins determined to be over-expressed in most ovarian tumor subtypes using spectral counting, such as desmoglein2 and versican, yielded consistent results using iTRAQ. Additional protein changes in most ovarian tumor subtypes or specific subtypes were also identified using iTRAQ and LC-Orbitrap (Table 2 and Supplementary Table 3, Figure 2e, 2f, 2g, 2h, 2i).

Validation by Western blot

To verify the expression of candidate proteins in ovarian tumor subtypes, Western blots were performed and relative protein abundance was determined using normalized densitometry data of Western blotting (Figure 3). Proteins for which antibodies were commercially available were used to probe the proteins pooled from three individuals of each subtype of ovarian tumors and normal tissues. The results for mesothelin, CEA5, and CEA6 support the proteomics data: Mesothelin was detected at higher levels in high-grade and low-grade serous carcinomas than other ovarian tumor subtypes and normal ovary tissue, while CEA5 and CEA6 were elevated in ovarian mucinous carcinoma (Figure 3a). In addition, to perform biological replicates, proteins from mucinous carcinoma tissues of four individual patients were probed with CEA5 antibody, and CEA5 was found to be upregulated in all four cases of ovarian mucinous carcinoma compared to the three patient-matched normal ovary tissues (Figure 3b).

To verify the proteins commonly expressed in most ovarian tumor subtypes, tissues from two individual ovarian cancer cases from each subtype, which were from different individuals than the three cases per subtype used for the discovery study using mass spectrometry, were probed with antibodies against versican, periostin, and galectin-3-binding protein. Enhanced expression of these three proteins was observed in ovarian tumors compared to normal tissues; however, the expression level of a particular protein differed in individuals within each ovarian tumor subtype. For example, periostin was elevated in most ovarian tumor subtypes compared to normal tissues but to different degrees (Figure 3c). It may be feasible to combine several potential markers, such as periostin and versican, to distinguish all the ovarian tumors from normal ovary.

Discussion

Epithelial ovarian tumors are both morphologically and biologically heterogeneous: they can be classified into several different subtypes, which have different clinical outcomes and may

require different treatments. In this study, we identified extracellular proteins either specifically expressed in a specific subtype or commonly expressed among the different subtypes. The former could be used as candidate biomarkers to discriminate between ovarian tumor subtypes, while the latter could be used in combination with current ovarian tumor biomarkers to increase the sensitivity and specificity for ovarian tumor diagnosis. We also identified 13 proteins with decreased expression or absent in ovarian cancer subtypes compared to normal ovary tissue (Supplementary Table 2).

Our search for candidate ovarian tumor biomarkers involved quantitatively analyzing *N*-linked glycoproteins using the SPEG method, which greatly enriched extracellular glycoproteins, and mass spectrometry, which led us to identify several candidate extracellular glycoproteins for ovarian tumor subtypes or ovarian tumor in general. The spectral counting for quantitation was performed on three repeated analyses of each sample. The candidate proteins were further verified by orthogonal quantitative proteomic approach using iTRAQ labeling and Western blots. We used two quantitative methods for this study to increase the confidence of the protein changes [27]. Even though the scale was not exactly the same, a trend of spectral counting was consistent with iTRAQ quantitation, which was similar to the conclusion drawn by Neilson in his review paper [28]. To our knowledge, this is the first *N*-linked glycoproteomic study to characterize proteins in different ovarian tumor subtypes.

Mesothelin was one of the altered proteins identified by both spectral counting and iTRAQ quantitation methods in this study, although there was a slight difference between the two quantification results: Mesothelin was overexpressed in high-grade serous, low-grade serous, and transitional-cell carcinomas by spectral counting, while it was overexpressed in these three subtypes plus, to a lesser extent, clear-cell and mucinous carcinomas, by the iTRAQ quantitation method (Tables 1 and 2, Supplementary Tables 2 and 3, Figure 3a). The detection of increased mesothelin glycopeptide in clear-cell and mucinous carcinomas by iTRAQ quantitation using LTQ-Orbitrap but not spectral counting using LTQ may be due to the different sensitivity of the two instruments in detection and identification of mesothelin glycopeptide in complex peptide mixtures from different cancer subtypes.

Mesothelin is a 40-kDa glycosylphosphatidylinositol (GPI)-linked glycoprotein that has been reported to be overexpressed in ovarian cancer [29–31] as well as other malignancies, including pancreatic cancer [32, 33], gastrointestinal stromal tumors [34], mesothelioma [35], biliary carcinomas [36], endometrial adenocarcinomas, and lung and stomach/esophagus carcinoma [30]. Approximately 70% of ovarian cancers [29, 37] express increased levels of mesothelin, and 55% of serous ovarian carcinomas present mesothelin immunoreactivity [38]. The fact that mesothelin was identified in this study as being overexpressed in ovarian tumors indicates that our discovery strategy using glycopeptides isolation and quantitative proteomic analysis is capable of identifying the known biomarkers for ovarian tumors and suitable for discovery of new proteins associated with different subtypes of ovarian tumors.

CA125 has been utilized as a tumor marker in monitoring the response of patients to therapy; however, it is not an optimal marker for screening ovarian tumor, as it is not expressed in some histological ovarian tumor subtypes, such as mucinous carcinoma [12, 36]. In our study, CEA5 and CEA6 were elevated only in mucinous carcinoma (Table 1, Supplementary Table 2, and Figures 3a and 3b). Further studies will be performed to evaluate the utility of the proteins for the diagnosis and treatment of mucinous carcinoma. CEA5, and CEA6, play roles in cell adhesion, invasion, and metastasis [39, 40], all of which are inhibited *in vitro* by an anti-CEA6 antibody [41]. The CEA5 gene, which is also known as CD66e, codes for the CEA protein and was originally described in 1965 as an antigen

expressed by gastrointestinal carcinomas [42]. Increased CEA levels in plasma and tissue correlate with reduced survival in patients with gastrointestinal carcinoma [43]. CEA6 is expressed on epithelia and granulocytes from various organs and by many human cancers, including many breast cancer, colon cancer, pancreatic cancer, and non-small-cell lung cancer cell lines [44–46]. In addition, Blumenthal *et al* demonstrated that CEA6 was elevated by almost 3-fold in mucinous ovarian adenocarcinomas versus serous ovarian adenocarcinomas, as determined using a tissue microarray [47].

Although further studies are needed for large-scale validation of the proteins we identified, this study provides the first discovery of candidate proteins for ovarian tumor subtypes. Our findings will help researchers and physicians understand the mechanisms of ovarian tumorigenesis and predict responses to targeted therapeutic agents. Furthermore, given the heterogeneous nature of other human tumors, our study may support the need for similar molecular characterizations of tumor subtypes of other organs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by HERA Foundation Ovarian Cancer Outside-the-box (OSB1) Seed Grant and with federal funds from National Institutes of Health, by grants U01CA152813 and RO1 CA122581. We thank Xiaer Sun from Johns Hopkins University for technique assistant.

Abbreviations

TFE	Trifluoroethanol
TBP	Tributylphosphine
LTD	Linear ion-trap mass spectrometer
SPEG	Solid-phase extraction of glycopeptides

References

1. Christie M, Oehler MK. Molecular pathology of epithelial ovarian cancer. *J Br Menopause Soc.* 2006; 12:57–63. [PubMed: 16776856]
2. Einhorn N, Trope C, Ridderheim M, Boman K, et al. A systematic overview of radiation therapy effects in ovarian cancer. *Acta Oncol.* 2003; 42:562–566. [PubMed: 14596514]
3. Zhu Y, Wu R, Sangha N, Yoo C, et al. Classifications of ovarian cancer tissues by proteomic patterns. *Proteomics.* 2006; 6:5846–5856. [PubMed: 17068758]
4. Jaffe, NS. Pathology and Genetics: Tumours of Haematopoietic and Lymphoid Tissues. Lyon: International Agency for Research in Cancer; 2003.
5. Seidman, J.; Russell, P.; Kurman, RJ. Surface epithelial tumors of the ovary. In: Kurman, RJ., editor. *Blaustein's Pathology of the Female Genital Tract.* 5th edn. Springer-Verlag; New York: 2002.
6. du Bois A, Luck HJ, Meier W, Adams HP, et al. A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. *J Natl Cancer Inst.* 2003; 95:1320–1329. [PubMed: 12953086]
7. Biade S, Marinucci M, Schick J, Roberts D, et al. Gene expression profiling of human ovarian tumours. *Br J Cancer.* 2006; 95:1092–1100. [PubMed: 16969345]
8. Seidman JD, Horkayne-Szakaly I, Haiba M, Boice CR, et al. The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *Int J Gynecol Pathol.* 2004; 23:41–44. [PubMed: 14668549]

9. McGuire V, Jessor CA, Whittemore AS. Survival among U.S. women with invasive epithelial ovarian cancer. *Gynecol Oncol.* 2002; 84:399–403. [PubMed: 11855877]
10. Schmandt RE, Bennett M, Clifford S, Thornton A, et al. The BRK tyrosine kinase is expressed in high-grade serous carcinoma of the ovary. *Cancer Biol Ther.* 2006; 5:1136–1141. [PubMed: 16855388]
11. Sugiyama T, Kamura T, Kigawa J, Terakawa N, et al. Clinical characteristics of clear cell carcinoma of the ovary: a distinct histologic type with poor prognosis and resistance to platinum-based chemotherapy. *Cancer.* 2000; 88:2584–2589. [PubMed: 10861437]
12. Itamochi H, Kigawa J, Sultana H, Iba T, et al. Sensitivity to anticancer agents and resistance mechanisms in clear cell carcinoma of the ovary. *Jpn J Cancer Res.* 2002; 93:723–728. [PubMed: 12079522]
13. Goff BA, Sainz de la Cuesta R, Muntz HG, Fleischhacker D, et al. Clear cell carcinoma of the ovary: a distinct histologic type with poor prognosis and resistance to platinum-based chemotherapy in stage III disease. *Gynecol Oncol.* 1996; 60:412–417. [PubMed: 8774649]
14. Singer G, Kurman RJ, Chang HW, Cho SK, Shih Ie M. Diverse tumorigenic pathways in ovarian serous carcinoma. *Am J Pathol.* 2002; 160:1223–1228. [PubMed: 11943707]
15. Singer G, Stohr R, Cope L, Dehari R, et al. Patterns of p53 mutations separate ovarian serous borderline tumors and low- and high-grade carcinomas and provide support for a new model of ovarian carcinogenesis: a mutational analysis with immunohistochemical correlation. *Am J Surg Pathol.* 2005; 29:218–224. [PubMed: 15644779]
16. Aunoble B, Sanches R, Didier E, Bignon YJ. Major oncogenes and tumor suppressor genes involved in epithelial ovarian cancer (review). *Int J Oncol.* 2000; 16:567–576. [PubMed: 10675491]
17. Feeley KM, Wells M. Precursor lesions of ovarian epithelial malignancy. *Histopathology.* 2001; 38:87–95. [PubMed: 11207821]
18. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol.* 2003; 21:660–666. [PubMed: 12754519]
19. Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H. Solid-phase extraction of N-linked glycopeptides. *Nat Protoc.* 2007; 2:334–339. [PubMed: 17406594]
20. Pedrioli PG, Eng JK, Hubley R, Vogelzang M, et al. A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol.* 2004; 22:1459–1466. [PubMed: 15529173]
21. Eng J, McCormack AL, Yates JR 3rd. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 1994; 5:976–989.
22. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem.* 2002; 74:5383–5392. [PubMed: 12403597]
23. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem.* 2003; 75:4646–4658. [PubMed: 14632076]
24. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst.* 1997; 8:581–599. [PubMed: 10065837]
25. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001; 305:567–580. [PubMed: 11152613]
26. Liu H, Sadygov RG, Yates JR 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem.* 2004; 76:4193–4201. [PubMed: 15253663]
27. Tian Y, Tan A, Sun X, Olson MT, Xie Z, Jinawath N, Chan DW, Shih Ie M, Zhang Z, Zhang H. Quantitative Proteomic Analysis of Ovarian Cancer Cells Identified Mitochondrial Proteins Associated with Paclitaxel Resistance Proteomics. *Proteomics Clin Appl.* 2009; 3:1288–1295. [PubMed: 21113235]

28. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, et al. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics*. 11:535–553. [PubMed: 21243637]
29. Ordenez NG. Value of mesothelin immunostaining in the diagnosis of mesothelioma. *Mod Pathol*. 2003; 16:192–197. [PubMed: 12640097]
30. Frierson HF Jr, Moskaluk CA, Powell SM, Zhang H, et al. Large-scale molecular and tissue microarray analysis of mesothelin expression in common human carcinomas. *Hum Pathol*. 2003; 34:605–609. [PubMed: 12827615]
31. Schaner ME, Ross DT, Ciaravino G, Sorlie T, et al. Gene expression patterns in ovarian carcinomas. *Mol Biol Cell*. 2003; 14:4376–4386. [PubMed: 12960427]
32. Sato N, Fukushima N, Maitra A, Iacobuzio-Donahue CA, et al. Gene expression profiling identifies genes associated with invasive intraductal papillary mucinous neoplasms of the pancreas. *Am J Pathol*. 2004; 164:903–914. [PubMed: 14982844]
33. Argani P, Iacobuzio-Donahue C, Ryu B, Rosty C, et al. Mesothelin is overexpressed in the vast majority of ductal adenocarcinomas of the pancreas: identification of a new pancreatic cancer marker by serial analysis of gene expression (SAGE). *Clin Cancer Res*. 2001; 7:3862–3868. [PubMed: 11751476]
34. Antonescu CR, Viale A, Sarraf L, Tschernyavsky SJ, et al. Gene expression in gastrointestinal stromal tumors is distinguished by KIT genotype and anatomic site. *Clin Cancer Res*. 2004; 10:3282–3290. [PubMed: 15161681]
35. Robinson BW, Creaney J, Lake R, Nowak A, et al. Mesothelin-family proteins and diagnosis of mesothelioma. *Lancet*. 2003; 362:1612–1616. [PubMed: 14630441]
36. Swierczynski SL, Maitra A, Abraham SC, Iacobuzio-Donahue CA, et al. Analysis of novel tumor markers in pancreatic and biliary carcinomas using tissue microarrays. *Hum Pathol*. 2004; 35:357–366. [PubMed: 15017593]
37. Ordenez NG. Application of mesothelin immunostaining in tumor diagnosis. *Am J Surg Pathol*. 2003; 27:1418–1428. [PubMed: 14576474]
38. Yen MJ, Hsu CY, Mao TL, Wu TC, et al. Diffuse mesothelin expression correlates with prolonged patient survival in ovarian serous carcinoma. *Clin Cancer Res*. 2006; 12:827–831. [PubMed: 16467095]
39. Yoshioka T, Masuko T, Kotanagi H, Aizawa O, et al. Homotypic adhesion through carcinoembryonic antigen plays a role in hepatic metastasis development. *Jpn J Cancer Res*. 1998; 89:177–185. [PubMed: 9548445]
40. Thomas P, Gangopadhyay A, Steele G Jr, Andrews C, et al. The effect of transfection of the CEA gene on the metastatic behavior of the human colorectal cancer cell line MIP-101. *Cancer Lett*. 1995; 92:59–66. [PubMed: 7757961]
41. Blumenthal RD, Hansen HJ, Goldenberg DM. Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen). *Cancer Res*. 2005; 65:8809–8817. [PubMed: 16204051]
42. Gold P, Freedman SO. Demonstration of Tumor-Specific Antigens in Human Colonic Carcinomata by Immunological Tolerance and Absorption Techniques. *J Exp Med*. 1965; 121:439–462. [PubMed: 14270243]
43. Kim DY, Kim HR, Shim JH, Park CS, et al. Significance of serum and tissue carcinoembryonic antigen for the prognosis of gastric carcinoma patients. *J Surg Oncol*. 2000; 74:185–192. [PubMed: 10951413]
44. Scholzel S, Zimmermann W, Schwarzkopf G, Grunert F, et al. Carcinoembryonic antigen family members CEACAM6 and CEACAM7 are differentially expressed in normal tissues and oppositely deregulated in hyperplastic colorectal polyps and early adenomas. *Am J Pathol*. 2000; 156:595–605. [PubMed: 10666389]
45. Hinoda Y, Saito T, Takahashi H, Itoh F, et al. Induction of nonspecific cross-reacting antigen mRNA by interferon-gamma and anti-fibronectin receptor antibody in colon cancer cells. *J Gastroenterol*. 1997; 32:200–205. [PubMed: 9085168]
46. Kuroki M, Matsushita H, Matsumoto H, Hirose Y, et al. Nonspecific cross-reacting antigen-50/90 (NCA-50/90) as a new tumor marker. *Anticancer Res*. 1999; 19:5599–5606. [PubMed: 10697625]

47. Blumenthal RD, Leon E, Hansen HJ, Goldenberg DM. Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. *BMC Cancer*. 2007; 7:2. [PubMed: 17201906]

Figure 1. Flowchart of the quantitative analysis of ovarian tumor-specific and ovarian tumor subtype-specific glycoproteins

2a

2b

2c

2d

2e

2f

2g

2h

Figure 2. The MS/MS spectra of the identified glycopeptides from the glycoproteins associated with ovarian cancer

2a–d were identified by spectral counting, 2e–i were identified by iTRAQ labeling. Lower case c represents carbamidomethyl modification, lower case m represents the oxidation modification, lower case f, e, a, and k represent the iTRAQ labeled N-termini and Lys, lower case n in the nXT/S motif represents the formerly glycosylated Asp and deaminated after SPEG isolation.

2a. CEA5, peptide 1 (AYVcGIQNSVSAⁿR);

2b. CEA5, peptide 2 (ITPNNnGTyAcFVSNLATGR);

2c. CEA6, peptide 1 (LQLSNGnmTLTLLSVK);

2d. CEA6, peptide 2 (NDAGSYEcEIQNPASAnR);

2e. versican, peptide (fEnQTGFPPPSR);

2f. periostin, peptide (eVnDTLLVNELk);

2g. mesothelin, peptide (aLSQQnVSmDLATFmk);

2h. Galectin-3-binding protein, peptide 1 (aLGFEnATQALGR);

2i. Galectin-3-binding protein, peptide 2 (aAIPsALDTnSSk);

3a

	Nr/Nr	1/Nr	2/Nr	3/Nr	4/Nr	5/Nr	6/Nr	7/Nr
CEA5	1.00	0.64	3.76	0.48	3.61	0.81	22.70	0.73
CEA6	1.00	0.92	2.65	1.47	2.12	1.14	9.59	1.66
Mesothelin	1.00	1.86	1.68	2.18	1.72	4.15	2.05	1.84

3b

3c

Figure 3. Verification of glycoprotein expression in ovarian tumor subtypes using Western blot analyses

3a, Western blot analyses of CEA5, CEA6, and mesothelin with the pooled tumor and normal tissues used in discovery study. Nr, normal ovary tissue; 1, clear-cell carcinoma; 2, high-grade endometrioid carcinoma; 3, high-grade serous carcinoma; 4, low-grade endometrioid carcinoma; 5, low-grade serous carcinoma; 6, mucinous carcinoma; 7, transitional-cell carcinoma. 3b, Western blot analysis of CEA5 with individual samples. N1-3, three cases of normal ovary tissues; M1-4, four cases of ovarian mucinous carcinoma. 3c, Western blot analyses of versican, periostin, and galectin-3-binding protein (gal3BP) with additional cases of tumor and normal tissues independent of discovery study. The cases

used were different in blot 1 versus blot 2. Nr, normal ovary tissue; 1, clear-cell carcinoma; 2, high-grade endometrioid carcinoma; 3, high-grade serous carcinoma; 4, low-grade endometrioid carcinoma; 5, low-grade serous carcinoma; 6, mucinous carcinoma; 7, transitional-cell carcinoma.

Extracellular glycoproteins altered in most ovarian tumor subtypes or in particular subtypes, as determined by spectral counting

Table 1

IP1	Protein Name	Number of unique peptides	Clear cell *	H-Endo *	L-Endo *	H-Serous *	L-Serous *	Muc-Inous *	Transitional-cell *	Nor-Mal *
expressed in most subtypes										
IP100028931	Desmoglein 2	1	8.00	30.00	15.00	17.00	42.00	21.00	26.00	1.00
IP100000877	150 kDa oxygen-regulated protein	3	1.60	5.40	2.60	2.40	2.60	1.40	3.00	1.00
IP100218585	perlestin	1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.00
IP100021230	CD44 antigen	1	3.67	4.17	2.83	1.50	1.17	1.83	3.83	1.00
IP100215631	Versican	1	3.08	2.43	2.50	3.00	1.08	1.18	2.83	1.00
IP100103356	Integrin-like protein	2	2.13	3.75	2.63	2.75	1.88	1.00	2.75	1.00
IP100031131	Adipocyte plasma membrane-associated protein	2	8.00	12.00	17.00	9.00	7.00	7.00	6.00	1.00
IP100032292	Metalloproteinase inhibitor 1	1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	0.00
IP100030936	Tetraspanin 1	1	14.00	7.00	12.00	8.00	6.00	57.00	8.00	1.00
IP100011229	Cathepsin D	1	100.00	100.00	100.00	0.00	100.00	100.00	100.00	0.00
IP100385428	Biliary glycoprotein	1	100.00	100.00	100.00	100.00	0.00	100.00	100.00	0.00
subtype specific										
IP100002406	Lutheran blood group glycoprotein	1	0.00	0.00	0.00	3.30	2.00	0.00	2.00	1.00
IP100003648	Poliovirus receptor related protein 1(Nectin 1)	1	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00
IP100012165	Mucin 5B	3	0.00	100.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100015872	Transmembrane 4 superfamily, member 3	1	1.00	8.30	1.30	2.30	0.00	19.00	4.30	1.00
IP100022255	BA209J19.1.1	2	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100025110	Mesothelin	2	0.00	0.00	0.00	100.00	100.00	0.00	100.00	0.00
IP100027201	Mucin 2	1	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100027486	CEA5	2	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100396094	CEA6	2	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100029153	Angiotensin converting enzyme-like protein	2	0.00	100.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100217987	Integrin alpha-M	3	0.00	100.00	100.00	100.00	100.00	0.00	0.00	0.00
IP100257928	hypothetical protein MGC44287	1	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
IP100220216	Tenascin	1	0.00	100.00	0.00	0.00	0.00	0.00	100.00	0.00

* Ratio of spectral counts of ovarian cancer subtypes to normal tissue. The ratio of the proteins which were identified only in ovarian cancer but not in normal tissues was arbitrary assigned to 100.

Table 2

Extracellular glycoproteins altered in ovarian tumors, as determined by iTRAQ quantitation

Accessions	Protein name	Number of unique peptides	Clear cell	H-endo	L-endo	H-serous	L-serous	Mucinous	Transitional-cell	Normal
Expressed in most subtypes										
gi4557485	Cenuloplasmin	1	41375.82	0.00	100.99	95.33	0.00	43319.69	177.06	1.00
gi5031863	Galectin-3-binding protein (Gal3BP)	2	2.18	2.69	2.75	1.87	3.09	5.71	4.21	1.00
gi31377806	Polymeric immunoglobulin receptor	2	60.55	164.52	0.00	38.56	11.09	216.97	1.60	1.00
gi189163485	Cathepsin A isoform b	1	22.05	10.34	0.00	11.14	17.86	36.59	31.57	1.00
gi7656967	Cadherin EGF LAG seven-pass G-type receptor 1	1	5.16	1.43	0.33	2.34	4.01	1.41	3.16	1.00
gi153946395	Tenascin	1	2.21	4.27	1.91	3.25	4.43	0.78	16.17	1.00
gi187607300	Versican	1	5.78	1.17	0.33	3.08	5.93	1.68	2.98	1.00
gi189217428	Laminin subunit alpha-3	1	2.05	2.08	0.36	11.57	5.26	1.64	2.35	1.00
gi189011550	Acid ceramidase	1	1.69	2.72	0.00	0.73	5.11	0.84	2.19	1.00
gi4557759	Myeloperoxidase	1	4.05	5.44	1.67	2.77	2.20	2.86	2.90	1.00
gi48255943	CD44 antigen	1	2.61	1.77	0.14	1.15	2.18	1.01	2.51	1.00
gi22202619	Cathepsin L1	1	2.03	1.05	0.15	0.85	2.11	0.64	2.23	1.00
gi119393891	Acid alpha-glucosidase preproprotein	1	3.22	3.04	1.35	1.42	3.61	1.42	3.11	1.00
gi53988378	Mesothelin	1	14.15	0.00	0.00	34.42	138.25	6.83	63.63	1.00
gi224586817	Golgi apparatus protein 1	1	2.08	2.38	0.52	1.30	2.69	1.48	2.78	1.00
gi148664211	Cell adhesion molecule 1	1	1.53	4.34	1.70	1.84	2.85	0.45	2.49	1.00
gi4504957	Lysosomal-associated membrane protein 2	1	1.85	2.09	0.46	1.57	2.44	1.47	3.39	1.00
gi57242798	Acid sphingomyelinase-like phosphodiesterase 3b	1	1.73	0.85	0.00	1.51	1.52	2.62	2.83	1.00
gi5453832	Hypoxia up-regulated 1	1	2.23	3.71	0.86	2.41	3.01	1.67	5.65	1.00
gi89191865	Integrin, beta 2	1	4.42	1.27	0.65	1.47	2.22	1.35	3.63	1.00
gi116534898	Desmoglein 2	2	1.38	1.95	0.05	1.98	4.68	2.29	4.50	1.00
Subtype specific										
gi50659080	Serpin peptidase inhibitor, clade A, member 3	1	2.44	0.72	0.40	0.98	1.39	1.26	0.83	1.00
gi167614504	Laminin, beta 1	2	1.42	0.85	0.00	1.32	0.83	0.58	2.12	1.00
gi223468595	Integrin alpha-V	1	2.56	0.68	0.00	1.32	2.00	0.40	1.55	1.00

Accessions	Protein name	Number of unique peptides	Clear cell	H-endo	L-endo	H-serous	L-serous	Mucinous	Transitional-cell	Normal
gi56711308	Protein GPR107	1	2.18	1.10	0.00	1.01	1.31	1.14	2.59	1.00
gi68161541	CEA1	1	1.49	2.06	0.26	1.30	0.91	1.76	1.84	1.00
gi42740907	Clusterin	1	0.69	1.05	0.22	13.02	0.73	0.63	1.07	1.00
gi16933553	Anthrax toxin receptor 1	1	1.06	1.25	0.14	1.16	2.24	1.00	1.72	1.00
gi225543438	Complement component 2	1	1.01	0.92	0.51	1.96	3.33	0.47	1.44	1.00
gi41350214	Asporin	1	1.25	0.31	0.02	0.10	3.04	0.16	0.44	1.00
gi209863034	Periostin	1	1.89	1.44	0.00	0.40	2.50	0.00	2.27	1.00
gi110611231	Chloride channel accessory 1	1	1.20	0.05	0.00	0.11	0.00	109.07	0.13	1.00
gi4759238	Tetraspanin-8	1	0.95	0.91	0.17	0.61	0.63	4.97	1.13	1.00
gi109633039	Protein tyrosine phosphatase, receptor type, F	1	1.74	1.51	0.22	2.00	1.29	1.04	6.04	1.00
gi157419122	Laminin, alpha 4	1	1.69	0.86	0.61	1.13	1.26	1.17	2.14	1.00
gi4507677	Endoplasmin	1	1.35	1.29	0.28	1.68	1.68	0.73	6.31	1.00