

An Integrated Clinical-Genomics Approach Identifies a Candidate Multi-Analyte Blood Test for Serous Ovarian Carcinoma

Ivo Meinhold-Heerlein,¹ Dirk Bauerschlag,¹ Yingyao Zhou,² Lisa M. Sapinoso,² Keith Ching,² Henry Frierson, Jr.,³ Karen Bräutigam,¹ Jalid Sehouli,⁴ Elmar Stickeler,⁵ Dominique Könsgen,⁴ Felix Hilpert,¹ Constantin S. von Kaisenberg,¹ Jacobus Pfisterer,¹ Thomas Bauknecht,⁶ Walter Jonat,¹ Norbert Arnold,¹ and Garret M. Hampton²

Abstract **Purpose:** Cancer of the ovary confers the worst prognosis among women with gynecologic malignancies, underscoring the need to develop new biomarkers for detection of early disease, particularly those that can be readily monitored in the blood. **Experimental Design:** We developed an algorithm to identify secreted proteins encoded among ~22,500 genes on commercial oligonucleotide arrays and applied it to gene expression profiles of 67 stage I to IV serous papillary carcinomas and 9 crudely enriched normal ovarian tissues, to identify putative diagnostic markers. ELISAs were used to validate increased levels of secreted proteins in patient sera encoded by genes with differentially high expression. **Results:** We identified 275 genes predicted to encode secreted proteins with increased/decreased expression in ovarian cancers (<0.5- or >2-fold, $P < 0.001$). The serum levels of four of these proteins (matrix metalloproteinase-7, osteopontin, secretory leukoprotease inhibitor, and kallikrein 10) were significantly elevated in a series of 67 independent patients with serous ovarian carcinomas compared with 67 healthy controls ($P < 0.001$, Wilcoxon rank sum test). Optimized support vector machine classifiers with as few as two of these markers (osteopontin or kallikrein 10/matrix metalloproteinase-7) in combination with CA-125 yielded sensitivity and specificity values ranging from 96% to 98.7% and 99.7% to 100%, respectively, with the ability to discern early-stage disease from normal, healthy controls. **Conclusions:** Our data suggest that this assay combination warrants further investigation as a multi-analyte diagnostic test for serous ovarian adenocarcinoma.

Cancer of the ovary accounts for the highest tumor-related mortality among women diagnosed with gynecologic malignancy. Of the 22,200 new cases of ovarian cancer estimated by the American Cancer Society in 2005, 16,100 are predicted to

die of disease.⁷ With a relatively uniform incidence of ovarian cancer over the past 10 years,⁷ these data suggest that ~75% of patients will succumb to disease progression. The high ratio of death to incidence for patients with ovarian carcinoma is largely due to late-stage diagnosis, at a time when the disease has typically spread into the peritoneum beyond the pelvic region. The 5-year survival rate for patients with International Federation of Gynecology and Obstetrics stage I disease, which is confined to the ovary, is >90% (1, 2). However, survival over the same period of time for patients diagnosed with International Federation of Gynecology and Obstetrics stage III or IV disease is <20%. Late-stage diagnosis of epithelial ovarian cancer can be attributed to the fact that the disease is relatively "asymptomatic" in its early stages, and that the symptoms of late-stage disease, such as abdominal discomfort, weight loss, diarrhea or constipation, vaginal bleeding, and shortness of breath, are nonspecific complaints. The relatively nonspecific nature of these symptoms underscores the need for disease diagnostics (1, 2).

The discovery of biomarkers for the detection of ovarian cancer, as in other cancers, has traditionally been driven by the serendipitous discovery of elevated tumor-associated antigen/

Authors' Affiliations: ¹Department of Gynecology and Obstetrics, University Hospital Schleswig-Holstein, Kiel, Germany; ²Genomics Institute of the Novartis Research Foundation, San Diego, California; ³Department of Pathology, University of Virginia Health Sciences System, Charlottesville, Virginia; ⁴Department of Gynecology and Obstetrics, University Hospital Charité, Berlin, Germany; ⁵Department of Gynecology and Obstetrics, University Hospital, Freiburg, Germany; and ⁶Lilly Deutschland, Bad Homburg, Germany. Received 3/21/06; revised 11/3/06; accepted 11/8/06.

Grant support: Deutsche Krebshilfe/German Cancer Aid grant 70-3099-Me1 and D/98/02291 (I. Meinhold-Heerlein), Lilly GmbH Deutschland, and Novartis Research Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

I. Meinhold-Heerlein, D. Bauerschlag, and Y. Zhou contributed equally to this study. Current address for G.M. Hampton: Celgene Corp., San Diego, CA.

Requests for reprints: Ivo Meinhold-Heerlein, Department of Gynecology and Obstetrics, University of Schleswig-Holstein, Kiel, Germany. E-mail: imeinhold@email.uni-kiel.de.

©2007 American Association for Cancer Research.
doi:10.1158/1078-0432.CCR-06-0691

⁷ <http://www.cancer.org>.

protein levels in patients with cancer compared with healthy controls. Antibodies recognizing several tumor-associated antigens, notably OC-125, which binds the CA-125/MUC-16 protein (3, 4), have been successfully used to aid in the diagnosis of ovarian cancer, in combination with other noninvasive techniques, such as transabdominal ultrasonography or transvaginal sonography (reviewed in ref. 5). As cancer of the ovary occurs in ~1 of 2,500 postmenopausal women in the United States, screening methodologies must attain at least 75% sensitivity and >99.6% specificity to yield a recommended positive predictive value (PPV) of 10% (where 1 of 10 surgical interventions would lead to the diagnosis of one ovarian cancer). Although trials of transvaginal sonography screening are at the margin of acceptable positive predictive value (9.9%), the use of transabdominal ultrasonography on patients with serum concentrations of CA-125 >30 units/mL, increases PPV to 21%. The diagnostic usefulness of CA-125, and perhaps of other serum markers, may be significantly enhanced by serial measurements taken over time. For example, a longitudinal analysis of serum from postmenopausal women in Stockholm using linear regression increased the PPV to 16%, substantially greater than the PPV of a single CA-125 assay (reviewed in ref. 5). However, these assays, or combinations of assays, are still insufficient for broad clinical acceptance and population-based screening.

The advent of large-scale gene expression analysis of tumors has generated a substantial number of potential new markers, several of which have been validated in small test populations. These include prostasin (6), osteopontin (OPN; refs. 7, 8), and the whey acidic protein HE4 (9). A recent review tabulated 29 different serum markers, including many of these newly described candidates (5). Although no single assay seems to provide the necessary PPV, many of these markers are reported to increase the specificity and sensitivity of CA-125 (5). Moreover, the conjoint analysis of multiple protein assays, such as the combination of leptin, prolactin, OPN, and the insulin-like growth factor-2, have shown encouraging results in training/test studies, yielding a test specificity and sensitivity of 95% (10). Thus, the discovery of additional biomarkers and methodologies for combined analyses hold promise for the diagnosis of early-stage disease.

Here, we developed an algorithm to identify secreted proteins encoded by genes represented on the ~22,000-member Affymetrix HU133a array, reasoning that genes with elevated expression in cancers may reflect increased levels of the encoded proteins in patient sera. To test this hypothesis, the algorithm was applied to the expression profiles of 67 serous ovarian adenocarcinomas from 64 patients, alongside a series of nine crudely enriched normal ovarian samples. Among 275 differentially expressed genes encoding putative secreted proteins, we show that a combination of ELISA assays for four proteins [matrix metalloproteinase-7 (MMP-7), OPN, kallikrein 10 (KLK10), secretory leukoprotease inhibitor (SLPI)], in combination with CA-125, leads to very high positive and negative predictive values, with correspondingly high sensitivity and specificity.

Materials and Methods

Tissue samples. For oligonucleotide microarray analysis, fresh-frozen samples of 67 tumors were obtained from 64 patients undergoing surgery for ovarian carcinoma at the University Hospital

of Freiburg, Germany ($n = 31$); the Tumor Bank Ovarian Cancer of the University Hospital Charité, Berlin, Germany ($n = 12$); and the University Hospital Schleswig-Holstein, Kiel ($n = 24$). For several patients at the University of Freiburg, we obtained multiple samples, typically from one or more ovaries and the peritoneal cavity. Institutional review board approvals were granted before accession of the tissues. The histologic characteristics of each sample were assessed on formalin-fixed and paraffin-embedded H&E sections, and approximately similar areas of the frozen tumor samples were used for RNA and DNA extraction. Most specimens contained 80% to 90% tumor cells, whereas no specimen contained <60% tumor cells.

RNA preparation and GeneChip hybridization. Frozen tissues were homogenized with a rotary homogenizer (Omni International, Marietta, GA) in RNeasy lysis buffer (Qiagen, Valencia, CA). Total RNA was prepared from tissues and cells using the RNeasy Mini kit (Qiagen). Hybridization on oligonucleotide microarrays (U133a GeneChip, Affymetrix, Inc., Santa Clara, CA) was done as described (11, 12). GeneChip hybridization data were processed and scaled as described (11, 12).

Data processing and prefiltering. Microarray data was processed using MAS5.0 (Affymetrix) and scaled to a nominal value of 200 (~3-5 copies of transcript per cell; refs. 13, 14) to facilitate inter-chip comparison. The primary data have been deposited and are accessible on the Gene Express Omnibus.⁸

Delineation of genes encoding putative secreted proteins. Several databases were used to identify genes encoding proteins with characteristics of secretion. First, we mapped U133a probe sets to REFSEQ protein identifiers via the NetAffix web site.⁹ Next, we mapped the Swissprot database to REFSEQ using Blastp, thus resulting in a common map for Swissprot annotations and U133a probe set identifiers. The subset of mapped proteins was queried for the "subcellular location," "secreted," and "extracellular" annotations. We then downloaded the Genome Ontology database (MySQL release and Perl library modules)¹⁰ and mapped them to the REFSEQ database. This subset of proteins were queried for the following level 3 and 4 terms: "extracellular space," "extracellular," "plasma protein," "cell-cell signaling," "extracellular matrix," "hormone activity," "extracellular matrix structural constituent," "cytokine activity," and "receptor binding." Proteins with the following Genome Ontology annotations were discarded from the list: "membrane fraction," "nucleus," "cytosol," "cytoplasm," "integral to membrane," "integral to plasma membrane," and "transmembrane receptor activity." Lastly, we queried REFSEQ proteins for the presence of signal peptides using signalP.¹¹ The criteria set was that at least one of the metrics for a signal peptide sequence was called by the signalnn method and called secreted by the signalhmm method. The combination of these methods led to a subset of 1,441 genes with evidence for encoding secreted proteins by one or more methods.

Differential gene expression. Differentially expressed genes were defined and ranked using a previously described method (11), which optimizes the selection of genes with a combination of high fold-change and significant P .

Statistical analysis. Differences in serum protein levels between groups of patients was calculated using the Wilcoxon rank sum test implemented in R (version 2.1.0).¹²

Split point, logistic regression, and support vector machine analyses. The split-point scoring procedure was described previously by Mor et al. (10). The method constructs a simple vote classifier, where each marker contributes a vote by comparing the observation value to a split-point value. The optimal split-point value for each individual marker is determined by maximizing the accuracy of its corresponding binary

⁸ <http://www.ncbi.nlm.nih.gov/geo/>.

⁹ <http://www.affymetrix.com/analysis/index.affx>.

¹⁰ <http://www.godatabase.org/dev/database/>.

¹¹ <http://www.cbs.dtu.dk/services/SignalP/>.

¹² <http://www.r-project.org>.

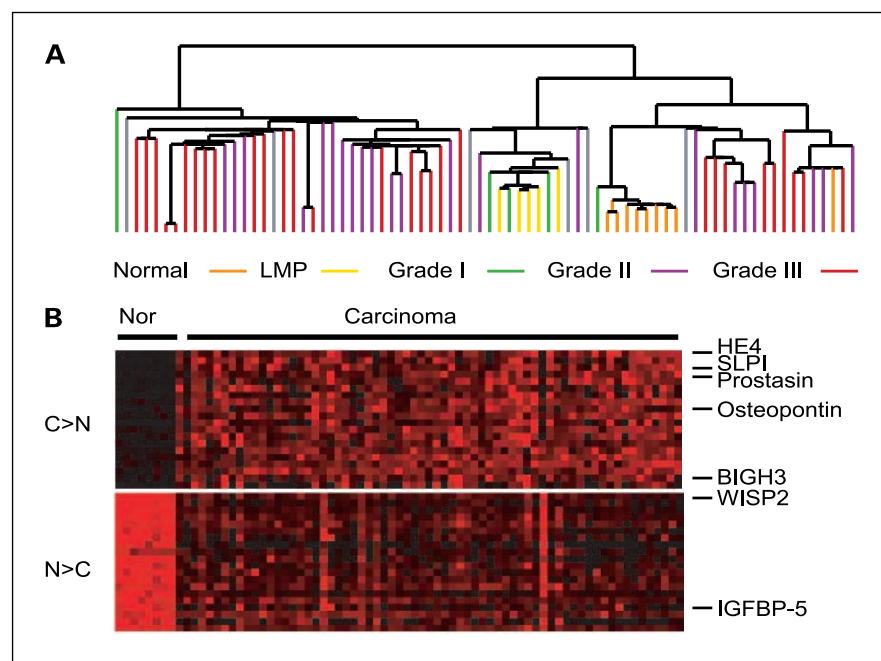


Fig. 1. Clustering of serous ovarian carcinomas and identification of secreted proteins. *A*, the 67 serous ovarian carcinomas and 8 normal tissues were subjected to hierarchical clustering as described (27). Note the separation of normal and grade 1 tissues. *B*, a Treeview (27) visualization of the relative expression of genes predicted to encode secreted proteins. Gene expression was normalized to a mean of 0 and a variance of 1. C > N, cancer expression > normal expression; N > C, normal expression > cancer expression.

classifier (essentially a single-node decision tree). Then, the optimal cutoff value for the vote classifier is determined based on cross-validation runs. Notice, in this method, that votes from each marker are counted equally.

Logistic regression analysis linearly combines measurements from each marker with individual assigned weights and then compares the resultant value to a split point. The implementation of logistic regression described by Smyth was done using MATLAB (Smyth GK, Statbox Toolbox for MATLAB).¹³

Support vector machine (SVM) analyses in this study were carried out using the LIBSVM implementation.¹⁴ Based on the authors' recommendation, we chose radial basis kernel, with both constrain variable *C* and kernel variable γ optimized via an exhaustive grid searching protocol. Whenever class sizes in a training set were unequal, we used a weighting argument “-w,” to rebalance the data set. To apply the LIBSVM method to three classes, we used the one-against-one approach described by Hsu et al. (15).

ELISA. The sera of tumor patients were taken before surgery in the Department of Gynecology of the University Hospital Schleswig-Holstein/Campus Kiel and in the Department of Gynecology of the University Hospital Charité Berlin. Negative control samples came from age-matched healthy women taken in the Department of Gynecology of the University Hospital of Freiburg. Blood was spun at $1,500 \times g$ for 5 min (room temperature), and the serum supernatant was carefully retained and directly frozen at -20°C . For quantification of the serum protein levels of OPN, MMP-7, SLPI, and KLK10, ELISAs were done using blood sera of 67 patients with stage I to IV serous ovarian carcinoma, 67 healthy women, and a small series of 15 serum samples from women with cystadenofibromas of the ovary. Measurements of OPN, MMP-7, and SLPI protein levels were carried out with the appropriate Quantikine Immunoassay kits according to the manufacturer's recommendations (R&D Systems; Wiesbaden-Nordenstadt, Germany). KLK10 protein levels were determined with the IBEX hk 10 mikrotiter ELISA (IBEX, Quebec, Canada). Before analyses, serum samples were diluted 1:10 for OPN, 1:4 for MMP-7, 1:20 for SLPI, and 1:1 for KLK10.

Serum CA-125 serum levels were determined automatically using the AxSYM CA-125 test on the Abbott AxSYM analyzer (Abbott Diagnostics,

Abbott Park, IL). A value of >35 units/mL was considered elevated. All immunoassay kits were used according to the manufacturer's protocols, and each serum sample was analyzed in duplicate.

Results

RNA from a series of 67 serous ovarian carcinomas and 9 crudely enriched normal tissue samples from the surface of the ovary were hybridized to oligonucleotide arrays containing probe-sets for $\sim 22,500$ human transcripts. The quality of the data set was assessed in two ways. First, agglomerative clustering of the most variably expressed genes was used to evaluate the extent to which normal and malignant tissues could be discerned (Fig. 1A). With the exception of one normal sample (N35S) and a single grade I tumor, all of the normal tissues could be clearly delineated from malignant tissues. Among the serous carcinomas, a major distinction was based on tumor grade and malignant potential, an observation we recently reported (16). Second, we used a previously described algorithm (*t* test/FOLD metric), which optimizes selection of genes with the highest fold change and most uniform differential expression, to identify genes with increased or decreased expression in carcinomas relative to normal tissue (11). The genes identified included the whey acidic protein *HE4* and the cell surface proteins *GA733-1*, *CD24*, *CD9*, and *mesothelin*, among others, which we and others have previously reported to be differentially expressed in ovarian carcinomas (reviewed in ref. 17). In aggregate, these results show the quality of the data in terms of previously reported observations and independent validation.

To identify putative diagnostic biomarkers within the data set that could be measured by protein assays in patient serum, we developed an algorithm based on previous studies (18) to identify genes encoding proteins with evidence of secretion. This algorithm combines keywords from the Genome Ontology consortium that imply secretion (e.g., “secreted,” “extracellular,” and “cytokine”), as well as sequence characteristics of

¹³ <http://www.statsci.org/matlab/contents.html>.

¹⁴ <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

Table 1. Genes encoding putative secreted proteins with increased expression in serous ovarian carcinomas

RefSeg	Symbol	Gene name
NM_006103	<i>WFDC2</i>	WAP four-disulfide core domain 2
NM_000096	<i>CP</i>	Ceruloplasmin
NM_003064	<i>SLPI</i>	Antileukoproteinase
NM_002773	<i>PRSS8</i>	Prostasin
NM_002087	<i>GRN</i>	Granulin
NM_001562	<i>IL18</i>	Interleukin-18
NM_001275	<i>CHI3L1</i>	Cartilage glycoprotein-39
NM_006010	<i>ARMET</i>	Arginine-rich, mutated in early-stage tumors
NM_000582	<i>SPP1</i>	Osteopontin
NM_003710	<i>SPINT1</i>	Serine protease inhibitor, Kunitz type 1
NM_001275	<i>CHI3L1</i>	Cartilage glycoprotein-39
NM_021102	<i>SPINT2</i>	Serine protease inhibitor, Kunitz type, 2
NM_002658	<i>PLAU</i>	Plasminogen activator, urokinase
NM_005576	<i>LOXL1</i>	Lysyl oxidase-like 1
NM_004636	<i>SEMA3B</i>	Semaphorin 3B
NM_002405	<i>MFNG</i>	Manic fringe homologue (<i>Drosophila</i>)
NM_015926	<i>ZSIG11</i>	Putative secreted protein ZSIG11
NM_014265	<i>ADAM28</i>	A disintegrin and metalloproteinase domain 28
NM_000358	<i>TGFBI</i>	Transforming growth factor, β -induced, 68 kDa
NM_001854	<i>COL11A1</i>	Collagen, type XI, α 1

secreted proteins, such as the presence of a signal peptide and the lack of transmembrane domains. Of the 22,383 probe sets on the array, we identified 1,441 with evidence of encoding secreted proteins. Of these, 275 were differentially expressed in serous ovarian carcinomas (see Supplementary Data). The genes were ranked by the *t* test/FOLD metric, and the top 20 up-regulated and down-regulated genes are listed in Tables 1 and 2 and depicted in Fig. 1B. Genes were also ranked simply based on fold change (Table 3), as many of these display high expression in a significant but lower fraction of cancers compared with those in Tables 1 and 2. We reasoned that measurement of several of these encoded proteins in combination may also yield a useful diagnostic panel.

We chose to focus on the measurement of four proteins encoded by genes listed in Tables 1 and 3 [*SLPI* (also known as antileukoprotease, *ALP*), *OPN*, *MMP-7*, and *KLK10*] for the following reasons: (a) each of the genes encoding these proteins exhibited highly restricted expression among diverse human tissues (data not shown¹⁵), suggesting that secretion from tissues other than cancers may be low, thus minimizing background protein levels; (b) cancer-associated expression was also evident in multiple cell lines of ovarian and other anatomic origins (data not shown⁹), suggesting continued selection for their expression and thus implying potential biological roles for these proteins in disease etiology; and (c) commercial ELISA assays were available to measure each

protein in patient plasma, thus allowing rapid measurement of these analytes without assay development and, if relevant, enabling rapid clinical translation.

The concentration of each of these proteins, alongside CA-125, was initially measured in a series of 67 serum samples from patients with stage I to IV serous ovarian carcinoma ($n = 7, 2, 49,$ and 9 with stage I, II, III, and IV disease, respectively) and compared with levels in a series of 67 serum samples from healthy controls. The levels of each protein in healthy and cancer patients were significantly different in each case (Fig. 2; Table 4). We then assessed the diagnostic performance of each individual marker using three classification methods: logistic regression, split-point analysis, and SVM, each computed with 1,000 runs of 10-fold cross-validation (Table 5). For single marker classifiers, all three methods yielded similar results. Because the logistic regression method is optimized against the log-odds of whether the dependent variable occurs or not (i.e., normal versus cancer in this case), and because the split-point method (see Materials and Methods) is directly optimized against the accuracy function, the latter tended to give the best single marker performance. However, when we combined all of the markers into a single classifier, the split-point method yielded the lowest accuracy (92.8%), which was lower than that obtained with OPN alone. The most likely explanation for this is that the method counts votes from each marker equally, despite the fact that the diagnostic power of each marker varies widely. In contrast, the logistic regression method assigns appropriate linear weights to individual markers and thus achieved a higher accuracy

Table 2. Genes encoding putative secreted proteins with decreased expression in serous ovarian carcinomas

RefSeg	Symbol	Gene name
NM_001393	<i>ECM2</i>	Extracellular matrix protein 2
NM_003881	<i>WISP2</i>	WNT1 inducible signaling pathway protein 2
NM_002937	<i>RNASE4</i>	Ribonuclease, RNase A family, 4
NM_025208	<i>SCDGF-B</i>	Spinal cord-derived growth factor-B
NM_000186	<i>HF1</i>	H factor 1 (complement)
NM_002937	<i>RNASE4</i>	Ribonuclease, RNase A family, 4
Hs.180324	<i>IGFBP5</i>	Insulin-like growth factor binding protein 5
NM_014057	<i>OGN</i>	Osteoglycin
NM_019105	<i>TNXB</i>	Tenascin XB
Hs.323511	<i>none</i>	Homo sapiens cDNA: FLJ23176 fis, clone LNG10452
NM_001920	<i>DCN</i>	Decorin
NM_006216	<i>SERPINE2</i>	Nexin, plasminogen activator inhibitor type 1
Hs.180324	<i>IGFBP5</i>	Insulin-like growth factor binding protein 5
NM_005410	<i>SEPP1</i>	Selenoprotein P, plasma, 1
NM_003243	<i>TGFBR3</i>	Transforming growth factor, β receptor III
NM_001847	<i>COL4A6</i>	Collagen, type IV, α 6
NM_000599	<i>IGFBP5</i>	Insulin-like growth factor binding protein 5
NM_005228	<i>EGFR</i>	Epidermal growth factor receptor
NM_000587	<i>C7</i>	Complement component 7
NM_016134	<i>PGCP</i>	Plasma glutamate carboxypeptidase

¹⁵ <http://symatlas.gnf.org>.

Table 3. Genes encoding putative secreted proteins with increased expression in serous ovarian carcinomas based on fold change

RefSeg	Symbol	Gene name
NM_001275 Hs.380020	<i>CHI3L1</i> None	Cartilage glycoprotein-39 DKFZp666A038 (from clone DKFZp666A038)
NM_001275 NM_006551	<i>CHI3L1</i> <i>SCGB1D2</i>	Cartilage glycoprotein-39 Secretoglobin, family 1D, member 2
NM_001854 NM_002421 NM_006681 NM_000096 NM_014265	<i>COL11A1</i> <i>MMP-1</i> <i>NMU</i> <i>CP</i> <i>ADAM28</i>	Collagen, type XI, α 1 Matrix metalloproteinase 1 Neuromedin U Ceruloplasmin A disintegrin and metalloproteinase domain 28
NM_004385 NM_006103 NM_002773 NM_000612 NM_020995 NM_001562 NM_003064 NM_000295 NM_000141	<i>CSPG2</i> <i>WFDC2</i> <i>PRSS8</i> <i>IGF2</i> <i>HPR</i> <i>IL18</i> <i>SLPI</i> <i>SERPINA1</i> <i>FGFR2</i>	Versican HE4 Prostasin Insulin-like growth factor 2 Haptoglobin-related protein Interleukin-18 Antileukoproteinase α -1 antiproteinase Fibroblast growth factor receptor 2
NM_002343 NM_004038 NM_002546 NM_005764 NM_000582 NM_002087 NM_006475 NM_002160 NM_004385 NM_005143 NM_000612 NM_001048 NM_000331 NM_002985 NM_005576 NM_003376	<i>LTF</i> <i>AMY1A</i> <i>TNFRSF11B</i> <i>MAP17</i> <i>SPP1</i> <i>GRN</i> <i>OSF-2</i> <i>TNC</i> <i>CSPG2</i> <i>HP</i> <i>IGF2</i> <i>SST</i> <i>SAA1</i> <i>CCL5</i> <i>LOXL1</i> <i>VEGF</i>	Lactotransferrin Amylase, α 1A; salivary Osteoprotegerin Membrane-associated protein 17 Osteopontin Granulin Osteoblast specific factor 2 Tenascin C Versican Haptoglobin Insulin-like growth factor 2 Somatostatin Serum amyloid A1 Chemokine (C-C motif) ligand 5 Lysyl oxidase-like 1 Vascular endothelial growth factor
NM_002776 NM_001845 NM_006010	<i>KLK10</i> <i>COL4A1</i> <i>ARMET</i>	Kallikrein 10 Collagen, type IV, α 1 Arginine-rich, mutated in early-stage tumors
NM_003710	<i>SPINT1</i>	Serine protease inhibitor, Kunitz type 1
NM_000358	<i>TGFBI</i>	Transforming growth factor, β -induced, 68 kDa
NM_001854 NM_002423 NM_001275 NM_006853 NM_002658	<i>COL11A1</i> <i>MMP-7</i> <i>CHGA</i> <i>KLK11</i> <i>PLAU</i>	Collagen, type XI, α 1 Matrix metalloproteinase 7 Chromogranin A Kallikrein 11 Plasminogen activator, urokinase
NM_032237 NM_021102	<i>FLJ23356</i> <i>SPINT2</i>	Hypothetical protein FLJ23356 Serine protease inhibitor, Kunitz type, 2
NM_004636 NM_000096 NM_006419	<i>SEMA3B</i> <i>CP</i> <i>CXCL13</i>	Semaphorin 3B Ceruloplasmin Chemokine (C-X-C motif) ligand 13
NM_004385	<i>CSPG2</i>	Versican

(98.0%). Finally, the radial kernel-based SVM classifier is a nonlinear method, which typically works better for multiple markers. For this reason, we chose to use SVM as the method of choice for further analyses. Results from logistic regression

analysis and split-point analysis are presented as supplementary materials (Supplementary Tables S1 and S2).

To quantitatively assess the diagnostic usefulness of multiple markers, we used SVM analysis on all 31 possible combinations of the five markers. Figure 3 shows the 10-fold cross-validation accuracies for the set of SVM results that can be obtained by starting from any individual marker and sequentially recruiting additional markers to achieve the next best-performing SVM. For all of the markers except SLPI, this heuristic feature selection routine quickly converged to the maximum performance of 99.3% with only three markers (i.e., OPN, CA-125, and MMP 7/KLK-10). With SLPI included, the SVM accuracy dropped to 98.5% and leveled off after three markers. Notably, OPN increased the CA-125-based SVM accuracy from $87.8 \pm 3.3\%$ to a near optimal accuracy of $97.6 \pm 1.5\%$, with a P of 1×10^{-68} based on the t test. Correspondingly, the PPV was also increased significantly from $94.3 \pm 2.3\%$ to $99.3 \pm 0.9\%$ ($P = 1 \times 10^{-48}$).

Because the SVM method is trained against optimal accuracy rather than PPV, specificity values tend not to be optimized in these classifiers. To check if the combination of multiple markers could potentially lead to higher-specificity SVM classifiers (>99.6%), we increased the cost of misclassifying controls as cancer samples (i.e., false positives) to be 250 times greater using the weighting option in LIBSVM (see Materials and Methods). We observed that with three markers, an SVM classifier based on OPN, CA-125, and MMP-7 reached a sensitivity of 98.7% and specificity of 99.7%. An SVM classifier based on OPN, CA-125, and KLK10 reached a sensitivity of 96.0% and specificity of 99.8%. With all five markers, a sensitivity of 95.7% and a specificity of 100% could be obtained. Although more samples would be required to minimize the statistical uncertainties in these estimations, the current results show that with multiple markers, it is possible to develop a clinically viable diagnosis tool with further improvements.

Receiver operator characteristic (ROC) analysis was carried out to compare the performance of classifiers over a range of internal settings (Fig. 4). When a single marker was used in the classification model, both split-point and logistic regression methods make use of a single internal threshold value. We therefore obtained the ROC curve by taking every observation value in the sample set as a possible internal threshold. Thus, both methods shared the same set of ROC curves. ROC curves were only generated for logistic regression analysis when multiple markers were involved to preserve the clarity of the curves illustrated in Fig. 4. Sensitivity-specificity pairs were collected within each cross-validation run by varying the internal threshold. Results from all of the 1,000 runs were then averaged to yield the ROC curve. It is clear from the ROC curves that for a wide range of classifier configurations, performance of individual markers decreases in the order of OPN, CA-125, MMP 7, KLK10, and SLPI, using either split-point or logistic regression. Multiple markers perform significantly better than single markers, with little difference between four and five marker combinations. This observation is also consistent with the SVM results in Table 5.

The collection of patient serum used here included nine from patients with early-stage disease (seven stage I and two stage II samples, respectively). We did SVM analysis to ask whether (a) early-stage serum levels could be distinguished from the levels in healthy controls; (b) early-stage serum protein levels could

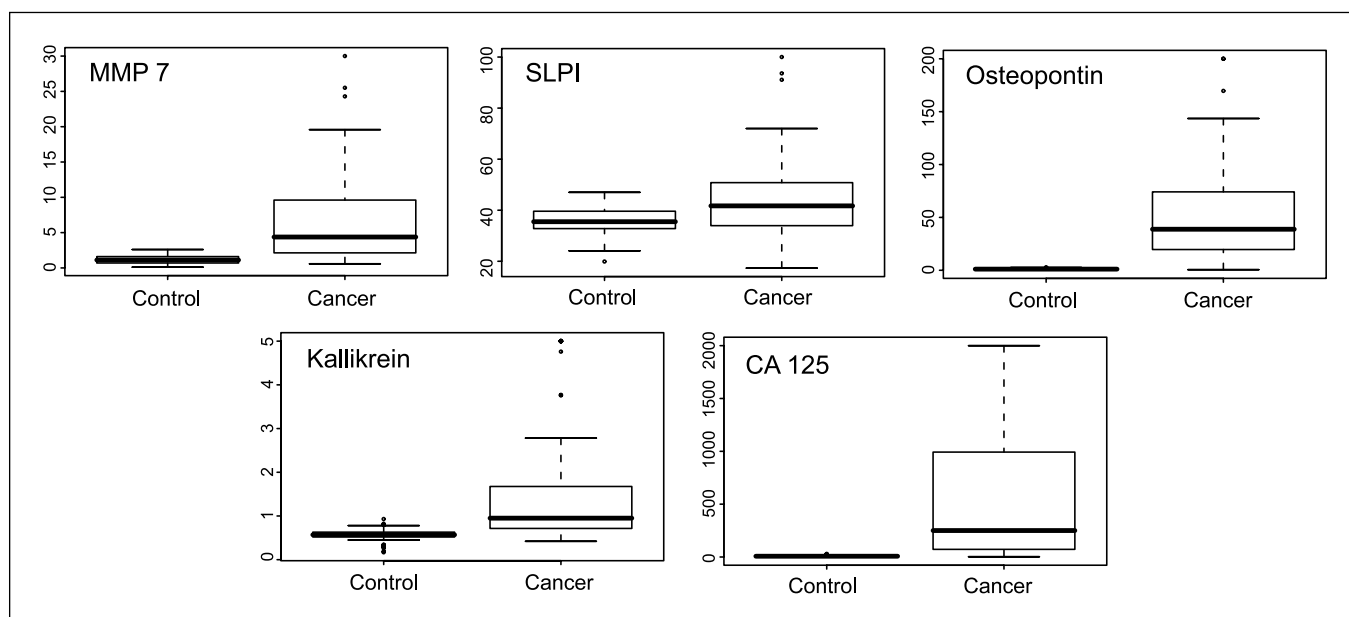


Fig. 2. Distribution of serum protein levels in healthy and malignant cases. Distributions are represented by box plots (thick black line, median; 95% confidence interval shown). Protein concentrations are expressed as ng/mL.

be distinguished from benign, and (c) early-stage and late-stage protein serum levels could be distinguished from one another? SVM correctly classified all of the early-stage cancers from normal samples but was less accurate in the early-stage benign distinction, correctly classifying 61% and 80% of early-stage and benign cases (3-fold cross-validation), respectively. SVM could not distinguish any of the early-stage serum protein levels from those of late-stage disease.

To further assess the potential of these marker combinations for the diagnosis of ovarian carcinomas, we measured the concentrations of each protein in a small series of serum from patients diagnosed with benign cystadenofibromas ($n = 15$) and evaluated whether these could be distinguished from protein levels in sera from normal and malignant patients (Fig. 5). Although three of five of the biomarker serum levels were significantly different between normal and benign samples (Table 3; Fig. 4), only OPN and CA-125 had any apparent power to differentiate serum from patients with cystadenofibromas from those with carcinomas (Table 3; Fig. 4). We used a three-class SVM based on a "one-against-one" approach, where binary SVM classifiers are separately built and require two-third votes for class assignment. This resulted in non-weighted and weighted accuracies of 90.6% and 75.5% (5-fold cross-validation), respectively. However, examination of the cases contributing to the overall accuracy revealed that cancer and normal samples account for the majority of correct

votes. In contrast, a significant number of benign cases were misclassified (data not shown).

Discussion

Ovarian carcinomas are typically diagnosed at International Federation of Gynecology and Obstetrics stage III/IV, a point at which tumors are localized beyond the pelvic region into the peritoneal surface and retroperitoneal lymph nodes. The lack of cancer-specific symptoms is a major reason for delayed diagnosis and, coupled with the poor outcome associated with late-stage disease, has fuelled significant efforts to identify diagnostic biomarkers for early disease detection. These efforts have centered on the discovery and validation of increased gene or protein expression in malignant tumors relative to normal tissues as well as proteomic profiling of low molecular weight fragments, typically by surface-enhanced laser desorption/ionization-time of flight. The former approach has yielded a large number of candidate proteins (reviewed in ref. 5), with widely variable diagnostic efficacy.

Here, we undertook a combined clinical-genomics approach toward the identification of new diagnostics, focusing on differentially expressed genes encoding putative secreted proteins. Validation of our approach comes from the identification of several proteins that exhibit increased levels in ovarian

Table 4. Wilcoxon sum rank values for the differences in serum protein levels in normal, benign, and cancer patients

	MMP-7	SLPI	OPN	KLK10	CA-125
Normal vs cancer	<2.2e-16	1.64e-04	<2.2e-16	8.48e-13	<2.2e-16
Normal vs benign	4.35e-06	0.07293	1.85e-05	3.15e-07	0.000191
Benign vs cancer	0.1587	0.378	0.02452	0.4534	7.54e-05

Table 5. Accuracies of three classification methods using all five markers individually or combined

	Accuracy (10-fold cross-validation), %		
	Logistic regression	Split point	SVM
OPN	94.0	94.8	94.0
CA-125	89.9	91.3	91.0
MMP-7	85.6	86.6	85.8
KLK10	79.9	81.7	80.6
SLPI	65.8	71.0	69.3
All five markers	98.0	92.8	99.3

cancer sera [e.g., HE4 (9), prostaticin (6), OPN (7, 8), several members of the kallikrein protein family (19), and insulin-like growth factor-2 (10)]. In this study, we focused on four genes encoding secreted proteins (MMP-7, SLPI, OPN, and KLK10), with the aim of assessing their potential usefulness as a multi-analyte diagnostic test.

Analysis of the serum data with several classification methods pointed to SVM as the most accurate for combinations of markers. SVM was sufficient to achieve >99% diagnostic accuracy with as few as three markers (CA-125, OPN, and MMP-7/KLK10) and increased the diagnostic accuracy of CA-125 (91%) by 7.5% with the addition of OPN (98.5%). Further modifications, using a penalty for false-positive classifications, yielded sensitivity and specificity values of 95.7% and 100%, respectively, with all five markers combined.

Although the numbers of patients with early-stage disease used in this preliminary study were low (seven stage I and two stage II), it is nonetheless noteworthy that our classifiers could readily distinguish these from control serum protein levels. However, the distinction between these early-stage cases and those with benign cystadenofibromas (*n* = 15) was not particularly good. However, detection of these lesions by the diagnostic combination described here may be advantageous because patients with this condition will normally undergo surgery for differential histologic diagnosis or to remove bulky benign disease.

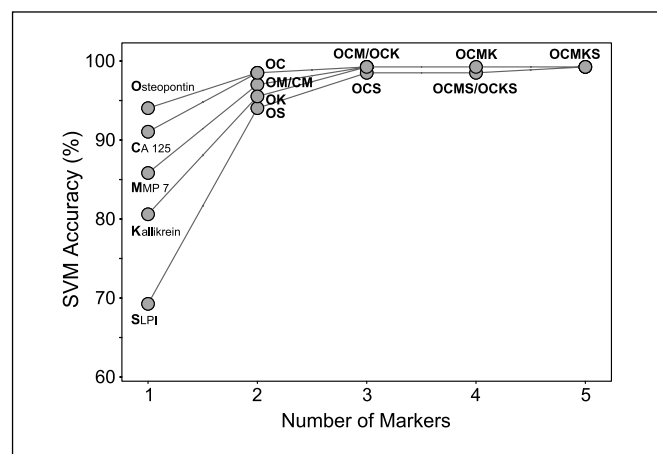


Fig. 3. SVM accuracy as a function of marker combinations. Ten-fold cross-validation accuracies for the set of SVM results that can be obtained by starting from any individual marker and sequentially recruiting additional markers to achieve the next best-performing SVM. O, OPN; K, KLK10; C, CA-125; M, MMP-7; S, SLPI.

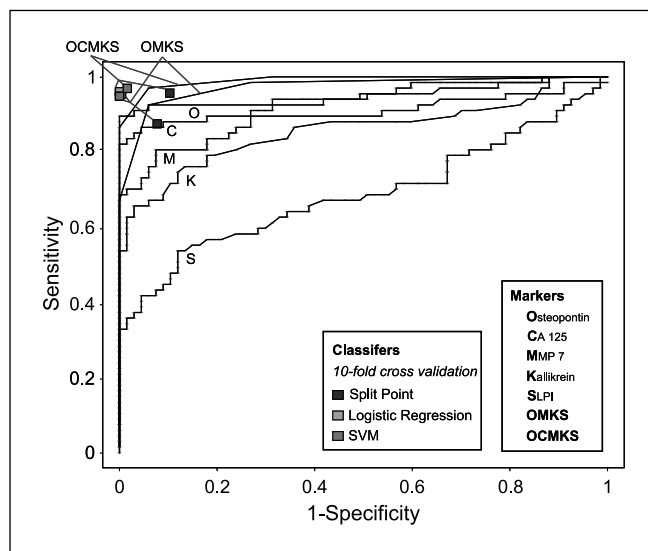


Fig. 4. ROC analysis of putative diagnostic markers.

Several studies have previously shown potential usefulness for OPN as diagnostic marker in ovarian cancer. In their original report, Mok et al. showed a specificity of 80.4% and early-stage (stage I/II) and late-stage (stage III/IV) sensitivities of 80.4% and 85.7%, respectively, noting significantly higher levels of serum OPN in ovarian carcinomas compared with other gynecologic malignancies and benign pelvic disease (7). Subsequently, Schorge et al. showed that OPN, while diagnostically inferior to CA-125 in predicting response to therapy, increased earlier in 90% of patients developing recurrent disease (20). In a study of 296 ovarian cancers, of which 65 were found to have weak or absent CA-125 immunoreactivity, OPN was immunoreactive in 60 of 65 cases, second to human KLK6 and KLK10 (98% and 100% positive, respectively; ref. 21). Thus, OPN seems to provide additional diagnostic information to CA-125. In the current study, OPN showed the highest sensitivity and specificity with 10-fold cross-validation (88% and 100%, respectively) and was notably different in benign and malignant cases, along with CA-125.

We found that SLPI done poorly in terms of differentiating serum protein levels in normal, benign, and malignant cases. This contrasts with a recent report in which the levels of protein in malignancy (mean = 63 and 67 ng/mL for stage I/II and III/IV disease, respectively) were almost twice those in benign (39 ng/mL) and normal patients (33 ng/mL; ref. 22). In the current study, the equivalent numbers were 45, 39, and 36 ng/mL.

KLK10 has been proposed as a potential marker for ovarian cancer, although the diagnostic sensitivity at 90% specificity is 54% (23). However, the combination of KLK10 with CA-125 increases the specificity by 21% at the same sensitivity. The same report also showed that KLK10 was elevated in 35% of CA-125-negative cases at 90% specificity. KLK10 was independently reported as elevated in 100% of 65 CA-125-negative cases (21); however, the latter measured KLK10 by immunohistochemistry, whereas the former was measured by ELISA at a defined specificity. Nonetheless, like OPN, these data point toward a complementary role for KLK10 in ovarian cancer diagnosis.

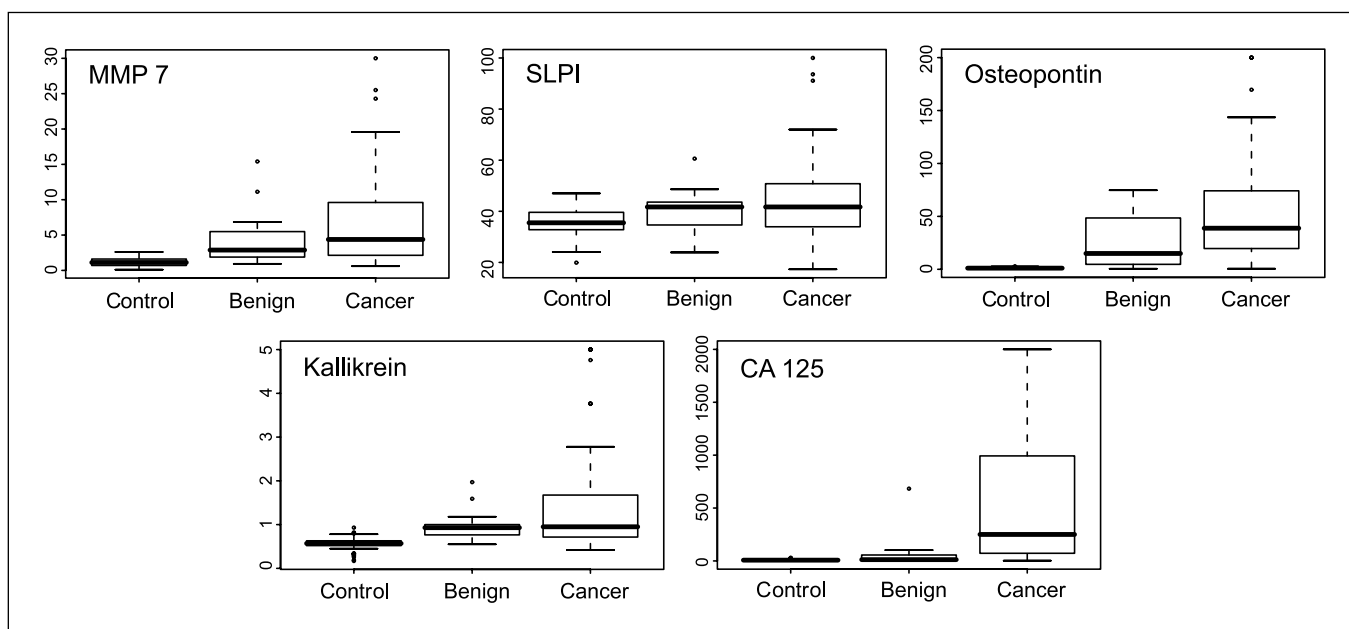


Fig. 5. Distribution of serum protein levels in healthy, benign, and malignant cases. Distributions are represented by box plots (thick black line, median; 95% confidence interval shown). Protein concentrations are expressed as ng/mL.

To our knowledge, MMP-7 has not been reported as elevated in the serum of ovarian cancer patients, although overexpression in ovarian carcinoma is well documented (24), and it has been shown to mechanistically influence ovarian cancer cell invasion (25). Our data suggest that MMP-7 is a useful adjunct to CA-125, with independent specificity and sensitivity of 79% and 82%, respectively (10-fold cross-validation). The combination of MMP-7, OPN, and CA-125 was shown to be sufficient for a decision tree model, leading to 96.3% diagnostic accuracy in the current study (10-fold cross-validation).

The algorithm used to identify secreted proteins yielded a significant number of candidate proteins for which assays are not yet available but may nonetheless prove diagnostically useful. In combination with other markers that have been recently described (e.g., HE4, insulin-like growth factor-2, prolactin, prostaticin, and leptin; see ref. 5 for a recent review),

a concerted effort to develop new assays, most likely based upon solid surface supports methodologies (e.g., protein or antibody arrays; ref. 26), is warranted. Large retrospective studies should be capable of identifying the subset of markers with the highest diagnostic accuracy and discrimination of benign and early-stage disease (stage I/II), which can then be validated in equally large multicenter prospective studies. With improving technologies, a multi-analyte test that incorporates 10 to 15 biomarkers should become financially viable and would have a profound effect on the current ovarian carcinoma mortality rates in the United States and elsewhere.

Acknowledgments

We thank Rosemarie Grams, Frank Rösel, Sigrid Hamann, and Ursula Petzold for excellent technical assistance.

References

- Memarzadeh S, Berek JS. Advances in the management of epithelial ovarian cancer. *J Reprod Med* 2001; 46:621–9; discussion 9–30.
- Holschneider CH, Berek JS. Ovarian cancer: epidemiology, biology, and prognostic factors. *Semin Surg Oncol* 2000;19:3–10.
- Verheijen RH, von Mensdorff-Pouilly S, van Kamp GJ, Kenemans P. CA 125: fundamental and clinical aspects. *Semin Cancer Biol* 1999;9:117–24.
- Bast RC, Jr., Xu FJ, Yu YH, Barnhill S, Zhang Z, Mills GB. CA 125: the past and the future. *Int J Biol Markers* 1998;13:179–87.
- Bast RC, Jr. Status of tumor markers in ovarian cancer screening. *J Clin Oncol* 2003;21:200–5.
- Mok SC, Chao J, Skates S, et al. Prostaticin, a potential serum marker for ovarian cancer: identification through microarray technology. *J Natl Cancer Inst* 2001;93:1458–64.
- Kim JH, Skates SJ, Uede T, et al. Osteopontin as a potential diagnostic biomarker for ovarian cancer. *JAMA* 2002;287:1671–9.
- Brakora KA, Lee H, Yusuf R, et al. Utility of osteopontin as a biomarker in recurrent epithelial ovarian cancer. *Gynecol Oncol* 2004;93:361–5.
- Hellstrom I, Raycraft J, Hayden-Ledbetter M, et al. The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res* 2003;63:3695–700.
- Mor G, Visintin I, Lai Y, et al. Serum protein markers for early detection of ovarian cancer. *Proc Natl Acad Sci U S A* 2005;102:7677–82.
- Welsh JB, Zarrinkar PP, Sapinoso LM, et al. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc Natl Acad Sci U S A* 2001;98:1176–81.
- Welsh JB, Sapinoso LM, Su AI, et al. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res* 2001;61:5974–8.
- Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotechnol* 1997;15:1359–67.
- Lockhart DJ, Dong H, Byrne MC, et al. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 1996;14:1675–80.
- Hsu C-W, Lin C-J. A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks* 2002;13:415–25.

16. Meinhold-Heerlein I, Bauerschlag D, Hilpert F, et al. Molecular and prognostic distinction between serous ovarian carcinomas of varying grade and malignant potential. *Oncogene* 2005;24:1053–65.
17. Le Page C, Provencher D, Maugard CM, Ouellet V, Mes-Masson AM. Signature of a silent killer: expression profiling in epithelial ovarian cancer. *Expert Rev Mol Diagn* 2004;4:157–67.
18. Welsh JB, Sapinoso LM, Kern SG, et al. Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc Natl Acad Sci U S A* 2003;100:3410–5.
19. Diamandis EP, Yousef GM. Human tissue kallikreins: a family of new cancer biomarkers. *Clin Chem* 2002;48:1198–205.
20. Schorge JO, Drake RD, Lee H, et al. Osteopontin as an adjunct to CA125 in detecting recurrent ovarian cancer. *Clin Cancer Res* 2004;10:3474–8.
21. Rosen DG, Wang L, Atkinson JN, et al. Potential markers that complement expression of CA125 in epithelial ovarian cancer. *Gynecol Oncol* 2005;99:267–77.
22. Tsukishiro S, Suzumori N, Nishikawa H, Arakawa A, Suzumori K. Use of serum secretory leukocyte protease inhibitor levels in patients to improve specificity of ovarian cancer diagnosis. *Gynecol Oncol* 2005;96:516–9.
23. Luo LY, Katsaros D, Scorilas A, et al. The serum concentration of human kallikrein 10 represents a novel biomarker for ovarian cancer diagnosis and prognosis. *Cancer Res* 2003;63:807–11.
24. Tanimoto H, Underwood LJ, Shigemasa K, et al. The matrix metalloprotease pump-1 (MMP-7, Matrilysin): a candidate marker/target for ovarian cancer detection and treatment. *Tumour Biol* 1999;20:88–98.
25. Wang FQ, So J, Reierstad S, Fishman DA. Matrilysin (MMP-7) promotes invasion of ovarian cancer cells by activation of progelatinase. *Int J Cancer* 2005;114:19–31.
26. Bogen SA, Sompuram SR. Recent trends and advances in immunodiagnostics of solid tumors. *Bio-Drugs* 2004;18:387–98.
27. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95:14863–8.