research articles Journal of proteome research

A Proteome Resource of Ovarian Cancer Ascites: Integrated Proteomic and Bioinformatic Analyses To Identify Putative Biomarkers

Limor Gortzak-Uzan,[†,▽,¶] Alex Ignatchenko,[†,¶] Andreas I. Evangelou,[†,¶] Mahima Agochiya,[‡] Kevin A. Brown,[‡] Peter St.Onge,[§] Inga Kireeva,[†] Gerold Schmitt-Ulms,[△] Theodore J. Brown,[⊥,#] Joan Murphy,[⊥,▽] Barry Rosen,[⊥,▽] Patricia Shaw,[△] Igor Jurisica,*[,‡,○,◆] and Thomas Kislinger*[,†,○]

Ontario Cancer Institute, Division of Cancer Genomics and Proteomics, Ontario Cancer Institute, Division of Signaling Biology, Department of Cell and Systems Biology, University of Toronto, Department of Laboratory Medicine and Pathology, University of Toronto, Department of Obstetrics and Gynecology, University of Toronto, Samuel Lunenfeld Research Institute, Division of Gynecological Oncology, Princess Margaret Hospital, Department of Medical Biophysics, University of Toronto, and Department of Computer Science, University of Toronto, Toronto, Canada

Received May 28, 2007

Epithelial ovarian cancer is the most lethal gynecological malignancy, and disease-specific biomarkers are urgently needed to improve diagnosis, prognosis, and to predict and monitor treatment efficiency. We present an in-depth proteomic analysis of selected biochemical fractions of human ovarian cancer ascites, resulting in the stringent and confident identification of over 2500 proteins. Rigorous filter schemes were applied to objectively minimize the number of false-positive identifications, and we only report proteins with substantial peptide evidence. Integrated computational analysis of the ascites proteome combined with several recently published proteomic data sets of human plasma, urine, 59 ovarian cancer related microarray data sets, and protein–protein interactions from the Interologous Interaction Database I^{2}D (http://ophid.utoronto.ca/i2d) resulted in a short-list of 80 putative biomarkers. The presented proteomics analysis provides a significant resource for ovarian cancer research, and a framework for biomarker discovery.

Keywords: Ascites • Proteomics • Bioinformatics • Integrative Computational Biology • Microarray • Body fluids • Biomarker • Ovarian cancer

Introduction

Epithelial ovarian cancer accounts for over 15 000 deaths per year in the United States of America[1] and over 125 000 worldwide,[2] making it the most lethal gynecological malignancy. Abdominal distention is a common symptom that presents as a result of the accumulation of ascites fluid in the peritoneal cavity. The mechanisms by which ascites fluid is formed are complex and include lymphatic obstruction, activation of mesothelial cells by the malignant metastatic process, and increased vascular permeability derived by the secretion of vascular endothelial growth factors and other cytokines.[3–7]

Ascites fluid represents the local microenvironment that is secreted by ovarian tumors and contains various cell types, including malignant cell that shed from the tumor, and soluble growth factors.[8–10] There is evidence that malignant ascites can stimulate the growth and invasive behavior of ovarian cancer cells both *in vivo* and *in vitro*;[11] thus, understanding the exact protein content of human ascites may provide critical information regarding ovarian tumor growth and progression.

Proteomic analysis of ascites can serve as a valuable tool in the ongoing search for biomarkers of ovarian carcinoma. The accumulation of ascites fluid in the peritoneal cavity of patients usually occurs as a manifestation of advanced disease, and therefore, identification of disease-related protein expression signatures may not improve the detection of early stage disease. However, comparison to other existing data sets, such as urine and plasma proteome profiles,[12,13] and publicly available ovarian cancer microarray data sets will enable integrated bioinformatics data mining to increase the specificity of proteomics analysis, and to identify putative ovarian cancer

* To whom correspondence should be addressed. Thomas Kislinger, Ph.D., tel, 416-581-7627; fax, 416-581-7629; e-mail, thomas.kislinger@utoronto.ca. Igor Jurisica, Ph.D., tel, 416-581-7437; fax, 416-581-7629; e-mail, juris@ai.utoronto.ca.
† Ontario Cancer Institute, Division of Cancer Genomics and Proteomics.
▽ Division of Gynecological Oncology, Princess Margaret Hospital.
¶ Authors contributed equally to this work.
‡ Ontario Cancer Institute, Division of Signaling Biology.
# Samuel Lunenfeld Research Institute.
§ Department of Cell and Systems Biology, University of Toronto.
△ Department of Laboratory Medicine and Pathology, University of Toronto.
⊥ Department of Obstetrics and Gynecology, University of Toronto.
○ Department of Medical Biophysics, University of Toronto.
◆ Department of Computer Science, University of Toronto.

10.1021/pr0703223 CCC: $40.75

Published on Web 12/13/2007

Table 1. Patient Information of Analyzed Ovarian Cancer Ascites

patient no.	age at diagnosis	histology	grade	stage
1	61	serous	3	IIIc
2	57	serous	n/a	IIIc
3	69	serous	n/a	IV
4	59	serous	3	IIIc

biomarkers with potential use in the prediction of outcome and treatment response, disease surveillance, and potentially early detection.

Although many attempts have been made to provide molecular insight into the content of ascites, no in-depth analysis has been reported to date. Xu et al. characterized an ovarian cancer activating factor in ascites from ovarian cancer patients.[14] The authors used mass spectrometry (MS) to identify an activating factor, purified from ascites of diseased subjects. The results demonstrated that a major component of ovarian cancer activating factor was palmitoyl lysophosphatidic acid. In another study, Gericke et al.[15] used MALDI-TOF-MS to characterize differences of transthyretin between ascitic fluid and plasma of women affected with ovarian cancer.

In this study, we present an in-depth, high quality proteome resource of human ascites (fluid and cellular fractions) from four ovarian cancer patients. All ascites samples were analyzed by a gynecologic pathologist, confirming the diagnosis of high grade serous carcinoma. None of the patients had previously been treated with chemotherapy (Table 1). Extensive MS-based proteome profiling by multidimensional protein identification technology (MudPIT)[16–19] and gel-enhanced LC−MS[12] resulted in the confident identification of over 2500 proteins.

Rigorous analysis of the MS-data was applied on the peptide and protein level to reduce the false discovery rate and to minimize protein inference by only accepting proteins with substantial peptide evidence. Data mining of this rich resource of ascites proteins revealed a large number of proteins involved in essential biological processes including immune function, metabolism, cytoskeletal dynamics, and cell motility.[20] Interestingly, over 600 proteins detected in this study have recently been reported in large-scale proteomics data sets of human plasma and urine.[12,13] To further minimize these *secreted* proteins to a panel of candidate ovarian cancer markers, we mapped these data against 59 available ovarian cancer microarrays and to protein–protein interaction data from the Interologous Interaction Database - I^2D (http://ophid.utoronto.ca/i2d).[21,22] This integrated bioinformatics data mining strategy reduced the large number of detected proteins to a more manageable number of putative biomarkers for future investigation.

Materials and Methods

Materials. Ultrapure grade urea, ammonium acetate, calcium chloride, TRIS, HEPES, and Triton-X-100 were from BioShop Canada, Inc. (Burlington, ON, Canada). Ultrapure grade iodoacetamide, DTT, and formic acid were obtained from Sigma. HPLC grade solvents (methanol, acetonitrile, and water) were obtained from Fisher Scientific. Recombinant, proteomics grade trypsin was from Roche Diagnostics (Montreal, QC, Canada). OMIX solid phase extraction pipet tips were from Varian (Mississauga, ON, Canada).

Fractionation of Human Ovarian Cancer Ascites. Ascites samples were obtained from the University Health Network Tissue Bank according to a protocol approved by the Institutional Research Ethics Board. Ascites were centrifuged at 2000 rpm for 10 min at 4 °C to separate the fluid (**Fraction S1**) from cellular components (**S2/S3**). The cells were incubated in mild hypotonic lysis buffer (10 mM HEPES, pH 7.4), to enrich for soluble proteins, for 30 min on ice. The suspension was briefly sonicated, and debris was sedimented by centrifugation at 14 000 rpm for 30 min at 4 °C. The supernatant was labeled **Fraction S2**. The cellular debris pellet was resuspended and further incubated for 30 min at 4 °C in hypotonic lysis buffer containing 1.5% Triton-X-100. Incubation with nonionic detergents will increase the extraction of membrane proteins. Following incubation, the suspension was centrifuged at 14 000 rpm for 30 min at 4 °C. The supernatant was labeled **Fraction S3**.

In-Solution Protein Digestion. An aliquot of 150 μg of total protein (determined by Bradford assay) from each fraction was precipitated overnight with 5 vol of ice-cold acetone followed by centrifugation at 14 000 rpm for 15 min. The protein pellet was solubilized in 8 M urea, 2 mM DTT, 50 mM Tris-HCl, pH 8.5, at 37 °C for 1 h, followed by carboxyamidomethylation with 10 mM iodoacetamide for 1 h at 37 °C in the dark. Samples were diluted with 50 mM ammonium bicarbonate, pH 8.5, to ~1.5 M urea. Calcium chloride was added to a final concentration of 1 mM and the protein mixture digested with a 1:30 molar ratio of recombinant, proteomics grade trypsin at 37 °C overnight. The resulting peptide mixtures were solid phase-extracted with Varian OMIX cartridges (Mississauga, ON, Canada) according to the manufacturer's instructions and stored at −80 °C until further use.

Multidimensional Protein Identification Technology (MudPIT). A fully automated 9-cycle, 16-h MudPIT procedure was set up essentially as described.[16–19] A quaternary HPLC-pump was interfaced with a linear ion-trap mass spectrometer (LTQ Thermo Fisher Scientific, San Jose, CA) equipped with a nanoelectrospray source (Proxeon Biosystems, Odense, Denmark). A 100 μm inner diameter fused silica capillary (Innova-Quartz, Phoenix, AZ) was pulled to a fine tip using a P-2000 laser puller (Sutter Instruments, Novato, CA) and packed with ~7 cm of Jupiter 4 μm Proteo 90 Å C_{12} reverse phase resin (Phenomenex, Torrance, CA), followed by ~5 cm of Luna 5 μm SCX 100 Å strong cation exchange resin (Phenomenex, Torrance, CA). Samples were loaded manually on separate columns using an in-house pressure vessel. As peptides eluted from the microcapillary columns, they were sprayed directly into the MS. A distal 2.3 kV spray voltage was applied to the microsplitter tee (Proxeon Biosystems). The MS operated in a cycle of one full-scan mass spectrum (400–1400 m/z), followed by 6 data-dependent MS/MS spectra at 35% normalized collision energy, which was continuously repeated throughout the entire MudPIT separation. The MS functions and the HPLC solvent gradients were controlled by the Xcalibur data system (Thermo Fisher Scientific, San Jose, CA).

Gel-Enhanced LC−MS Analysis. For gel separations, approximately 50–100 μg of total protein of each ascites fraction was loaded onto 10–14.5% gradient polyacrylamide precast Tris-HCl gels (Bio-Rad), followed by Coomassie staining (Bio-Safe Coomassie stain, Bio-Rad). Each gel lane was cut into 11 slices, which were subject to in-gel proteolysis, essentially as described[23] (Supplementary Figure S1 in Supporting Information). Briefly, destained gel blocks were washed with 100 mM ammonium bicarbonate and shrunk with acetonitrile. Reduction and alkylation was performed with 10 mM DTT and 55

mM iodoacetamide, respectively. Gel slices were washed twice with ammonium bicarbonate/acetonitrile and dried down in a speed-vac centrifuge. Each gel block was rehydrated in 12.5 ng/μL trypsin in 50 mM ammonium bicarbonate at 4 °C for 15 min. The solution was removed and replaced by 50 mM ammonium bicarbonate without trypsin and incubated overnight at 37 °C. The next day, the supernatant was removed and the gel slices were extracted twice with 5% formic acid/acetonitrile (50:50). The three supernatants were combined, speed-vac-concentrated to low volumes, diluted with 5% acetonitrile and 0.1% formic acid, and stored at −80 °C until analyzed by LC−MS. For gel-enhanced LC−MS analysis, a 100 μm inner diameter fused silica capillary was pulled to a fine tip using a P-2000 laser puller and packed with ∼7 cm of Jupiter 4 μm Proteo 90 Å C_{12} reverse phase resin. An aliquot of the in-gel-digested peptide mixture was pressure-loaded on a microcapillary column and analyzed by a 1-h LC−MS gradient. The MS scan functions were set exactly as described above.

Western Blotting. The soluble ascites fraction S1 was used for Western blotting. Total protein was quantified by Bio-Rad Protein Assay (Bio-Rad Laboratories, Mississauga, ON). Total protein (3–50 μg) was separated by 7.5% or 16% SDS-PAGE and transferred onto nitrocellulose membranes (Bio-Rad Laboratories). Blots were probed for cofilin-1 (Abcam, Cambridge, MA), profilin-1 (Abcam), GAPDH (Abcam), and IQGAP-1 (Santa Cruz Biotechnology, Santa Cruz, CA). Immunoreactive proteins were detected with secondary HRP-coupled antibodies (Sigma) and visualized with ECL Western Blotting Detection Reagents (Amersham).

Protein Identification, Validation, and Grouping. Raw files were converted to *m*/*z* XML using ReAdW and searched by X!Tandem[24] against a human IPI (International Protein Index; http://www.ebi.ac.uk/IPI) protein sequence database (v3.20).[25] To estimate and minimize our false-positive rate, the protein sequence database also contained every IPI protein sequence in its reversed amino acid orientation (*target-decoy strategy*) as recently described.[16,18,26,27] Search parameters were the following: Parent ion Δ-mass of 4 Da, fragment mass error of 0.4 Da, and complete carbaminomethyl modification of cystein by IAA. Only peptides passing a default log-expectation value of −1 were further evaluated (see below).

A rigorous peptide quality control strategy was applied to effectively minimize false-positive identifications. Briefly, a Perl-based tool was written to calculate, on the peptide level, a user-defined, false-positive rate (Perl version 5). Identified peptides were binned into three charge states (+1, +2, +3), and individual X!Tandem expectation values were calculated for each charge state to minimize the number of peptides mapping to decoy sequences to a user-defined percentage. For the presented study, we set the value of total reverse spectra to total forward spectra to 0.2%, resulting in a low number of decoy sequences in the final output (<0.5%; 13 reverse proteins in 2737 forward proteins). The calculated expectation values were −3.22 for +1-ions, −2.17 for +2-ions, and −2.89 for +3-ions.[28,29] Only proteins identified with ≥2 unique, fully tryptic peptides with ≥7 amino acids matching these criteria were accepted to generate the final list of identified proteins.

To minimize protein inference, we developed a database grouping scheme and only report proteins with substantial peptide information. Briefly, the algorithm comprises several grouping steps performed on the database level and applies a heuristic approach by favoring proteins within a parsimonious cluster, identified with the largest number of peptides. Similar strategies for the inference from peptide to protein level have been applied by the HUPO Plasma Proteome Project[30] and more recently in a large-scale investigation of human adipocytes.[31] This strategy includes several consecutive steps and only considers confidently identified peptides that passed our above filter criteria:

(1) **Linking peptides to proteins:** Peptide identifications obtained from X!Tandem searches (passing the above criteria) including all matching IPI protein accessions were parsed into our MySQL database.

(2) **Connecting groups:** These initial assignments of peptides to proteins were arranged by combining peptide groups that identify the same proteins.

(3) **Arranging groups:** These second level groupings were then sorted in order to reflect number of peptides matching to a given protein. This step takes into account proteins with sequence homologies as manifested by shared peptides (protein inference problem).

(4) **Minimal proteins list:** In the final grouping step, proteins are arranged into "*levels*" favoring proteins identified with the largest number of peptides, which were placed onto the highest level. Importantly, only proteins on this top level are reported in our final "*minimal protein inference list*". Although, if highly homologous proteins were identified by this process, having an identical number of diagnostic peptides, an arbitrary choice of the protein having the lowest IPI accession number was reported in our final output (e.g., Actin, cytoplasmic 1, and Actin cytoplasmic 2). Proteins on a lower level within one group were only reported if they were identified by at least one additional *distinct* peptide.

Comparison to the HUPO Plasma Proteome and Urine Proteome Data Sets. Comparison of our ascites proteome resource to the HUPO plasma proteome data (3020 proteins)[13] and a recently published urine proteome data set (1543 proteins)[12] was performed using the ProteinCenter bioinformatics software (Proxeon Biosystems, Odense, Denmark). To link these data sets, we loaded the IPI accession keys for each of the three projects into ProteinCenter. Individual data sets were then linked via accession keys in the ProteinCenter database, which contains over 30 million accessions.

Comparison to Ovarian Cancer Microarray Data Sets. Identified candidate proteins (Figure 4B) were queried against our in-house database for differentially expressed genes assembled from several ovarian cancer microarray studies (Supplementary Table S6 in Supporting Information). Protein/gene pairs found to be significantly changed in at least one of the comparisons were selected.

Protein–Protein Interaction Analysis. To further focus the selection of putative markers, we mapped the set of 18 and 10 potential biomarkers (see Results and Discussion) into I^2D protein–protein interaction databases.[22] This enabled us to (1) identify subnetworks that are significantly deregulated in multiple microarray ovarian cancer studies, and (2) determine additional putative markers that may have been removed initially due to stringent peptide-to-protein filtering steps, or were absent in plasma and urine from healthy subjects. To further increase the confidence in these putative biomarkers, we assessed whether the enrichment can be achieved by chance, as described next.

Random Network Comparisons. To determine if the experimentally derived networks could be obtained by random chance, the following permutation method was employed.

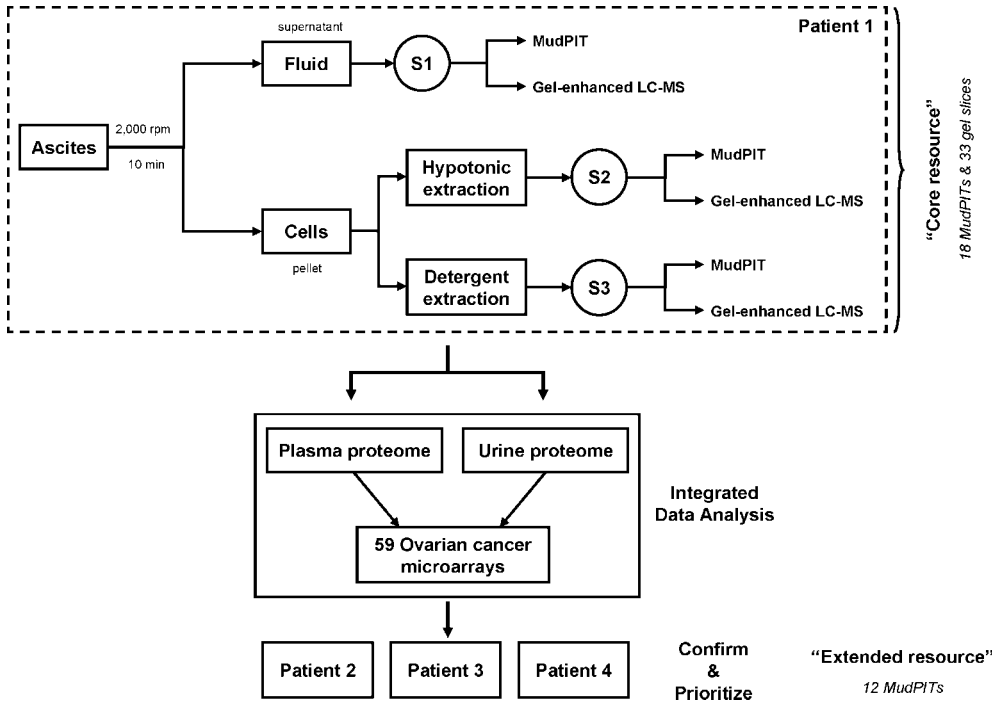

Figure 1. Analysis scheme of human ovarian cancer ascites. Approximately 1 mL of ovarian cancer ascites was centrifuged to separate the soluble (**S1**) from the cellular fraction. The cellular fraction was consecutively extracted with a hypotonic (**S2**) and detergent (**S3**) extraction buffer. Each fraction was extensively analyzed by MudPIT and gel-enhanced LC−MS/MS (for patient 1; *core resource*). Integrated bioinformatics analysis was used to map the ascites proteome data to human plasma[13] and urine,[12] as well as 59 available ovarian cancer microarrays. A limited analysis of the cellular fractions of three additional ovarian cancer ascites samples was used to confirm and prioritize putative biomarkers.

Eighteen (or 10) proteins were chosen at random from a pool of known secreted proteins, or the set of all known human proteins in UniProt (build 8.2). The random proteins (seeds) were then used to search the I^2D database[22] in two passes. On the first pass, all proteins known to interact with the seeds were selected and added to the seed list. On the second pass, protein interactions from the I^2D database among any of the expanded set of seed proteins were selected and added to generate the random network. Local and global network statistics were then calculated. The process was repeated for 1000 iterations (for the 18 seeds) or 10 000 iterations (for the 10 seeds) to generate the random distribution of network properties. The Student's *t* test was then used to compare the properties of the experimentally determined networks against the distributions of random networks.

Results and Discussion

A Systems Biology Analysis Strategy To Identify Putative Biomarkers. In this study, we applied in-depth shotgun proteomics to generate a high confidence protein resource of ovarian cancer ascites. Although MudPIT is an excellent technology for the identification of large numbers of proteins, it lacks the high-throughput capabilities of microarray or SELDI-TOF-MS approaches, required to analyze large numbers of patients for useful clinical insights. Current microarray platforms are highly automated and enable analysis of many samples, albeit lacking the ability to directly identify proteins, the molecular species most likely to be useful as a disease biomarker. Integrative systems biology enables us to merge the two platforms, and diminish negative aspects, while taking advantage of the positive features of each platform. We have combined extensive MudPIT-based profiling with publicly available proteome data sets, predicted and known protein–protein interactions in I^2D, and a large number of ovarian cancer microarray data sets. This strategy enabled us to (1) confidently identify a large number of proteins in a selected number of ovarian cancer ascites, (2) identify proteins previously described as secreted proteins, as these are most likely the best biomarker candidates, and (3) link these putative candidates to changes in the expression profile of several ovarian cancer microarrays (disease context). This integrated approach led to the identification of 80 putative biomarkers. To fully evaluate the potential of these candidate biomarkers, future investigations in large numbers of carefully controlled plasma samples are required. These studies require high-throughput, sensitive, and quantitative methods such as ELISA (Enzyme-Linked ImmunoSorbent Assay) or protein arrays for targets with validated antibodies, or highly specialized MS-based methods such as *Multiple Reaction Monitoring* (MRM).

Biochemical Fractionation of Human Ovarian Cancer Ascites. As ascites are produced in large volumes, and contain both tumor cells and soluble growth factors, we sought to perform an in-depth proteomics analysis of human ovarian cancer ascites to guide the discovery of candidate biomarkers for ovarian cancer. We obtained ascites from four patients with *serous* ovarian carcinoma who were not previously treated by chemotherapy (Table 1). To minimize the protein complexity, we fractionated ascites into cellular and soluble protein fractions, and independently analyzed them by extensive MudPIT-based expression proteomics and gel-enhanced LC−MS (Figure 1). The cellular fraction within ascites typically includes a heterogeneous mixture of many cell types including cancer cells, mesothelial cells, inflammatory cells, and blood cells.[32] For our core database sample, we selected ascites from a

Table 2. Proteins Detected in Ovarian Cancer Ascites

	core resource	extended resource			
	patient 1	patient 2	patient 3	patient 4	total
Unique peptides	18559	4569	5938	6137	22972
Unique proteins	2278	710	976	1057	2737
Reverse proteins	9	0	2	2	13

patient diagnosed with high grade serous carcinoma (patient 1) and that was found to contain largely carcinoma cells as determined from cytological analysis of a cytospin preparation (Supplementary Figure S2 is Supporting Information). For the purpose of this study, we made no attempts to further enrich for cancer cells and purposely decided to analyze the cellular fraction of the ascites as a whole, since factors secreted by other cells might also serve as useful biomarkers of the disease. To compensate for the high sample complexity and large dynamic range of protein abundance, we performed repeat MudPIT-based analyses, as recently suggested.[16,33–35] The combination of a simple biochemical fractionation protocol, with extensive MudPIT-based and gel-enhanced LC−MS, enabled us to confidently detect over 2500 proteins (Table 2), and we present for the first time a comprehensive human ovarian cancer ascites resource.

Protein Identification. In the course of this study, a total of 30 MudPITs and 33 gel-enhanced LC−MS/MS runs were recorded, resulting in the accumulation of over 3 million mass spectra. An essential process of proteomics is to rigorously filter the recorded data to minimize the number of false-positive protein identifications, while aiming at low false-negative rates (see Materials and Methods for details).

For this project, we used the open-source search algorithm X!Tandem,[24] as it provides a cost efficient, fast, and accurate alternative to the commercially available algorithms SEQUEST[36] and Mascot.[37] X!Tandem searches were performed as described in the Materials and Methods. Using the "target/decoy" strategy,[16,27] we filtered the generated data to minimize false-positive identifications (<0.5%) (Figure 2 and Table 2).

In total, 18 559 (patient 1 only; 22 972 for all four patients) unique peptides were identified, mapping to a total of 5629 unique IPI accessions (≥2 peptides per protein). To minimize protein inference, we developed a peptide-centric database scheme (see Materials and Methods) and grouped IPI accessions to a minimal set of reported proteins, similar to a recently reported strategy.[31] After applying the grouping scheme, the 5629 unique IPI accessions were collapsed to a reported *"minimal proteins"* list of 2299 unique IPI IDs. After removing known contaminants (human keratins and trypsin), a total of 2272 proteins are currently present in our ascites resource [Note: The 2272 proteins are derived form the very detailed analysis of patient 1, referred to as '*core ascites resource*'. The analyses of selected cellular fractions of three more patients (2, 3, and 4) increased the total number of identified proteins to 2737, referred to as '*extended ascites resource*']. This data was used to confirm and prioritize some of our initial candidates. The entire set of 2737 proteins is available in Supplementary Table S1 (Supporting Information). We have also deposited the RAW-data for this project on the Tranche (www.proteomecommons.org) Web server.

Comparison of the biochemical fractions revealed that the majority of the detected proteins were present in the mixed cellular fractions (**S2** and **S3**) and only 229 proteins were confidently detected in the soluble **S1** fraction (Figure 2C). This is not surprising as this fraction is very similar in protein composition and concentration to human plasma, containing very high concentrations of serum albumin and immunoglobulines, complicating the detection of low-abundance proteins.[38] Furthermore, MudPIT detected a larger number of proteins compared to gel-enhanced LC−MS. The reason for this increased number of protein identifications is most likely due to the larger number of MudPIT experiments (30 in total; 270 individual files) as compared to 33 gel-enhanced LC−MS experiments (Figure 2D).

Functional Annotation and Comparison to Human Body Fluid Data Sets. To obtain a systematic and functional overview of proteins present in human ovarian cancer ascites, we used an FDA Gene Ontology For Functional Analysis tool (GOFFA).[20] GOFFA is an intuitive tool that calculates statistically enriched GO-terms and allows the user to interactively visualize high-throughput data in the context of biological functions.

In Figure 3, proteins confidently detected in selected ascites fractions were analyzed, and significantly enriched (P-value ≤ 0.01; E-value ≥ 2) GO-terms in the category '*biological process*' are presented. A large number of detected proteins could be mapped to at least one GO-term (67–88%), and a significant number of enriched GO-terms were detected in each ascites fractions, covering a range of biologically relevant processes. As to be expected, fraction **S1** was enriched in Gene Ontology terms related to '*blood*, *plasma* and *immune*' related functions (69%). Several other GO categories, *metabolism* and *death*, were also significantly enriched in the soluble ascites fraction. Interestingly, the fraction of mixed cell populations (**S2** and **S3**) was heavily enriched in '*metabolism*' related GO-terms; this finding is in agreement with other reports that found malignant cells to be enriched with this biological process.[39–41] Other enriched Gene Ontology functions were *translation*, *splicing*, *transport*, and *cytoskeleton/mobility*. Supplementary Table S2 (Supporting Information) contains the most significantly enriched Gene Ontology terms in the category '*cellular component*' for each for the three analyzed fractions.

Bioinformatics Data Mining To Identify Putative Biomarkers. Clinical biomarkers are needed to improve diagnosis, prognosis, and to monitor the effects of therapeutic treatment.[42,43] However, molecular heterogeneity of the tumors, noncancer diseases that reduce biomarker specificity, and low sensitivity of the cancer biomarker due to low concentration in early disease are the three main pitfalls for the discovery of effective biomarkers for ovarian cancer. The most commonly used marker for ovarian cancer is CA-125,[44] which is widely used for the surveillance of ovarian cancer patients whose diagnosis was previously made by histology. A downside of CA-125-based diagnosis is its low sensitivity and specificity, especially for the detection of early disease.[45]

A large number of molecules have been evaluated as possible biomarkers for ovarian cancer. These include vascular endothelial growth factor (VEGF),[46,47] matrix metalloproteinases (MMP's),[48,49] and members of the kallikrein family of serine proteases.[50–52] Because of concerns about low sensitivity and specificity, the search for new and improved biomarkers for ovarian cancer continues.

To identify putative biomarker candidates, we used all proteins identified in ovarian cancer ascites of patient 1 (***core resource***), and mapped them to several previously reported large-scale data sets. These included the HUPO plasma proteome,[13] a recently published high quality urine proteome data set[12] (Figure 4), and 59 available ovarian cancer microarrays

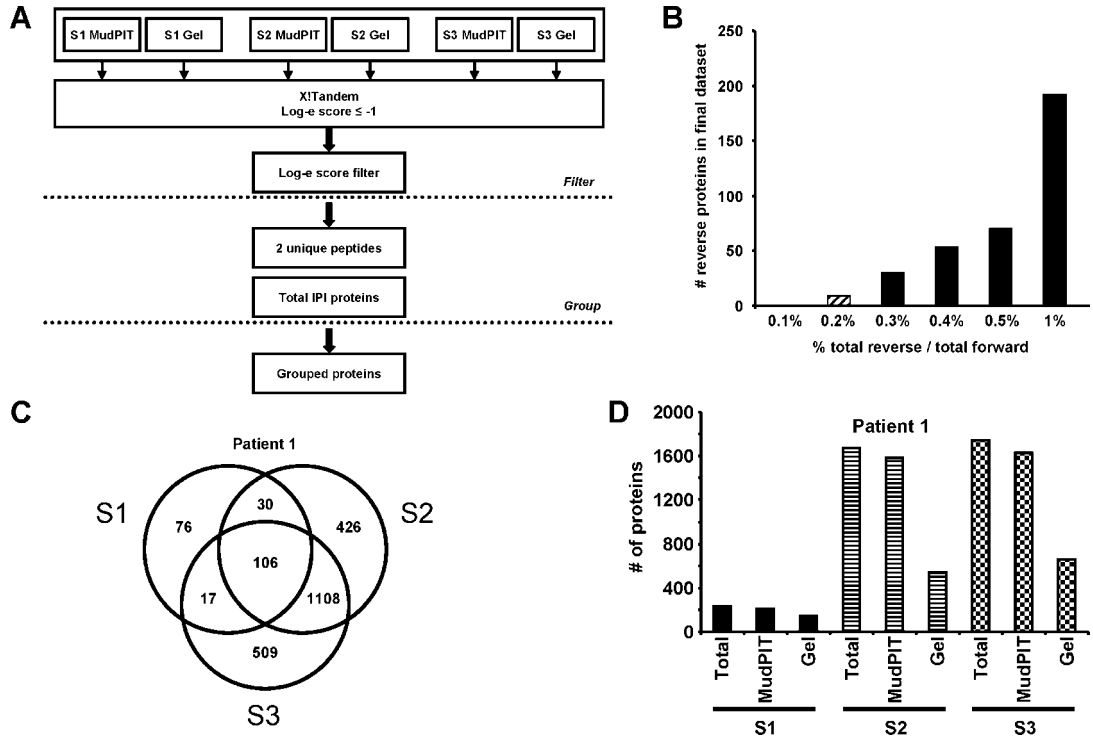

Figure 2. Proteomics identification strategy and number of detected proteins. (A) MS data obtained from individual ascites fractions was searched by X!Tandem and peptides passing a default log-expectation value of ≤−1 were parsed into an in-house MySQL database. A developed filter algorithm was used to calculate specific log-expectation values for individual charge states (+1, +2, +3) of assigned peptide–ions, to minimize the number of reverse peptide sequences to a user defined percentage. Only proteins identified with at least two peptide sequences were further considered for identification. A developed grouping algorithm was used to minimize the protein interference problem and only report proteins with informative peptide evidence. (B) Different user-defined filter criteria are presented. The number of reverse proteins in the final data set is blotted against the percent of total reverse/forward spectra present after applying the filter algorithm, a cutoff value of 0.2% was chosen for this project. (C) Venn diagram of proteins confidently identified in each of the three biochemical fractions. (D) A comparison of proteins detected by MudPIT or gel-enhanced LC−MS/MS.

(for detailed information see Supplementary Tables S3−S6 in Supporting Information). We then applied three different mining strategies to generate our list of putative biomarkers (Figure 4B):

1. Consider only proteins found in human ovarian cancer ascites that were also found in plasma and/or urine, and significantly changed on at least one ovarian cancer microarray → *secreted proteins*.
2. Select proteins found in human ovarian cancer ascites, significantly changed on at least one ovarian cancer microarray, but not detected in healthy body fluids → *putative disease related proteins*.
3. Prioritize protein interacting partners of group (1) according to the I^2D database [21,22] that were also significantly changed on at least one ovarian cancer microarray and found in our ascites proteome. The main reason to include this group is not to miss potential biomarkers that may be false negatives from a stringent proteomic analysis, but are related to ovarian cancer and interact with detected proteins.

To verify and confirm the reproducible detection of these initial candidate proteins, we analyzed three more ovarian cancer ascites (although less extensively; "***extended resource***") and only accepted proteins confidently detected by mass spectrometry in all four ascites samples. This reduced our panel of putative ovarian cancer biomarkers to a nonredundant list of 80 proteins (Supplementary Tables S3−S5 is Supporting Information). Clearly, this strategy is biased against lower abundance proteins, but the goal of this study was to identify the most robustly detected proteins. The entire data set presented in this manuscript is available to the scientific community, and the RAW-MS data was deposited to the Tranche server to enable further analysis.

In the current study, we present for the first time a high quality proteome resource of ovarian cancer ascites. Using extensive MudPIT-based and gel-enhanced LC−MS/MS proteomics, we were able to identify over 2500 proteins in the soluble and mixed cellular fractions of four confirmed cases of ovarian cancer. Please note that this relatively small number of analyzed patients is mandated by the enormous efforts involved in shotgun proteomics analyses. The data presented here contain a large number of confidently identified proteins for future investigations in larger patient cohorts by high-throughput technologies. Applying an integrated informatics platform, we were able to minimize the large number of identified proteins to a small panel of proteins reproducibly detectable in ovarian cancer ascites.

1. Secreted Proteins in Human Plasma/Urine Also Detected in Ascites. A total of 18 proteins were detected in all four ascites and have previously been reported as secreted proteins in large-scale analyses of human plasma[13] and urine[12] (Figure 5). Systematic analysis allowed their further grouping into functional categories such as *cell proliferation, differentiation and apoptosis, signaling to the cytoskeleton, cell adhesion and motility, transport, metabolic processes*, and *proteolysis*. Comparison to 59 available ovarian cancer microarrays further prioritized these proteins as

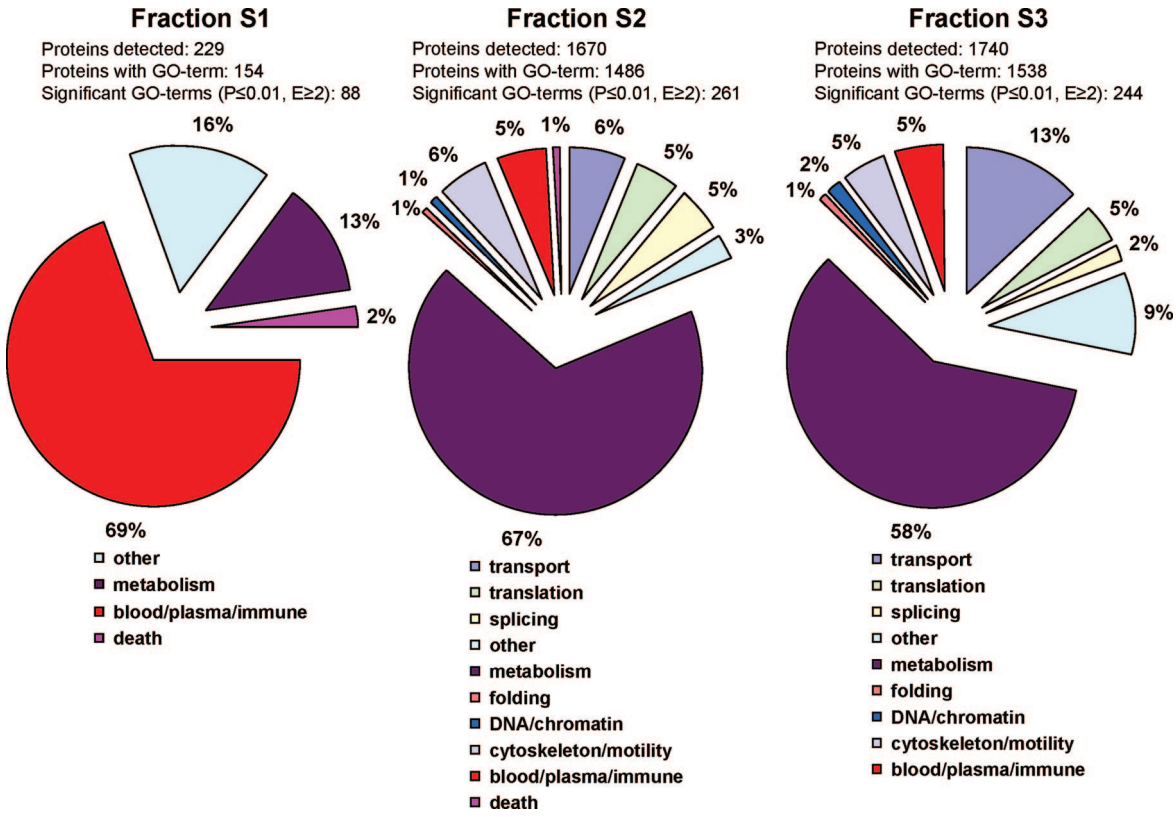

Figure 3. Significantly enriched Gene Ontology terms (*biological process*). Distribution of significantly enriched GO-terms in individual ascites fractions. Numbers were calculated by the GOFFA tool, and only significant categories are reported (*P*-value ≤ 0.01, *E*-value ≥ 2).

putative biomarkers. Although microarray expression and protein abundance do not always correlate, we have recently shown good correlation between mRNA and protein expression in several large-scale profiling projects.[16,53] By adding microarray expression profiles, we are able to add an additional stringency filter, although at a cost of bias against lower abundance proteins.

1.1. Proteins Associated with Cell Proliferation. S100A11 (Calgizzarin), a calcium-binding protein, regulates cell growth by inhibiting DNA synthesis.[54] In epithelial ovarian carcinoma cells, nuclear expression of S100 protein is associated with aggressive behavior of ovarian tumors.[55] We robustly detected S100A11 in all four ovarian cancer ascites, and its expression was up-regulated in ovarian cancer, as compared to normal ovary by microarray analyses (Figure 5 and Supplementary Table S6 in Supporting Information).

1.2. Proteins Involved in Cell Differentiation and Apoptosis. Several candidates are associated with cellular differentiation and apoptosis. These include, glutathione-*S*-transferase P (GSTP1), Cofilin-1 (CFL1), apolipoprotein-E (APOE), Lysozyme C (LYZ), and mitochondrial 10 kDa heat shock protein (HSPE1).

Increased expression of GSTP1 in human cancers confers some protection of malignant cells from antineoplastic drug treatment by preventing apoptosis and poor prognosis.[56,57] APOE is involved in lipid homeostasis, and was found up-regulated in ovarian cancer samples through SAGE analysis.[58] Chen et al. showed that APOE was absent from borderline ovarian tumors but was detected in later stage ovarian cancers.[59] Its inhibition in an ovarian cancer cell line (OVCAR-3) led to cell cycle arrest and apoptosis. CFL1 is a cytoskeletal protein related to cell migration, aggregation, and differentiation. Regulation of cofilin affects morphological changes by remodeling the Actin cytoskeleton associated with differentiation in cancer cells,[60] and has been shown to modulate tumor cell invasion.[61]

The robust detection of these proteins in the present study, its previous detection by mass spectrometry in urine/plasma, and significant changes for each candidate protein on several ovarian cancer microarrays advocates further investigation of these proteins as potential biomarkers for ovarian cancer (Figure 5 and Supplementary Table S6 in Supporting Information).

1.3. Proteins Involved in Signaling to the Cytoskeleton. Several candidate proteins including CFL1, Profilin-1 (PFN1), Rho GDP-dissociation inhibitor 2 (ARHGDIB), Ras GTPase-activating-like protein 1 (IQGAP1), Galectin-3-binding protein (LGALS3BP), S100A11, APOE, and 10 kDa Heat shock protein (HSPE1) are involved in intracellular signaling to the cytoskeleton.

IQGAP1, a scaffold protein, is involved in cytoskeletal rearrangement through interaction with various proteins including Actin, calmodulin, CD44, E-cadherin, the tumor suppressor adenomatuous polyposis coli (APC), CDC42, and Rac1.[62,63] Overexpression of IQGAP1 can significantly increase the migratory and invasive potential of cancer cells.[62] Additionally, IQGAP1 is expressed in endothelial cells and can bind VEGFR2, promoting endothelial cell migration and proliferation, following vascular injury,[64] and contribute to angiogenesis.[65] The connection to invasion, migration, and angiogenesis, as well as its robust detection in ovarian cancer ascites, human plasma,[13] and urine,[12] make it an interesting candidate biomarker for future evaluation. To the best of our knowledge, IQGAP1 has not been proposed as a biomarker candidate for ovarian cancer (Figure 5 and Supplementary Table S6 in Supporting Information).

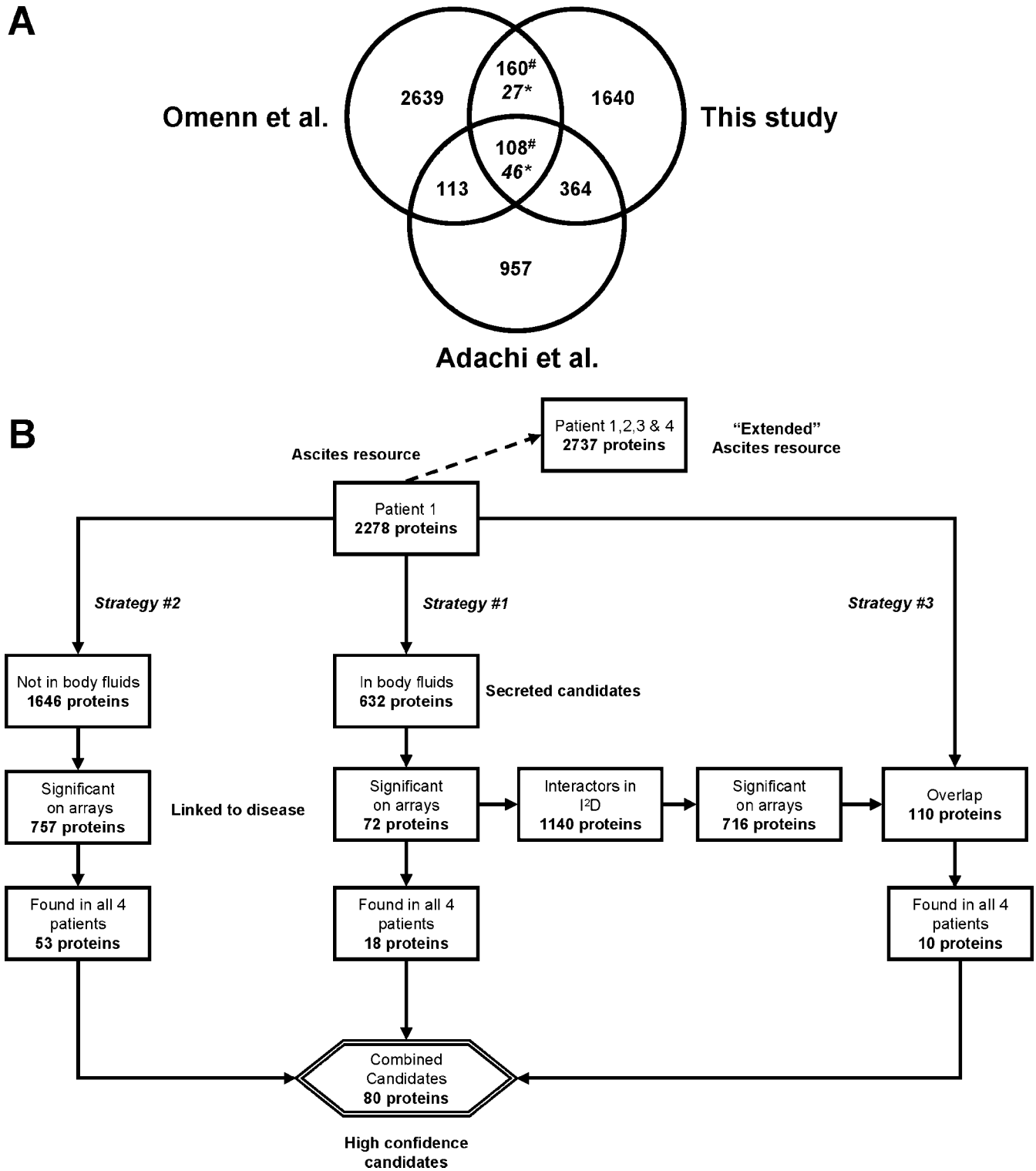

Figure 4. Integrated bioinformatics strategy to identify putative biomarker candidates. (A) Venn diagram of proteins detected in ascites, the HUPO plasma proteome[13] and a high quality human urine data set.[12] #, Overlap to the initial HUPO plasma proteome.[13] *, Overlap to the more rigorous plasma proteome data set of States et al.[85] (B) Proteins from the ascites resource (patient 1) were first mapped to available urine and plasma proteomes. This strategy resulted in 632 proteins found in previously published, body fluid data sets (secreted proteins) and 1640 proteins not detectable in these healthy plasma and urine resources. Next, both of these lists were mapped against 59 available ovarian cancer microarrays, and only proteins showing significant changes on at least one microarray were further evaluated. As a third arm of this integrated data analysis strategy, we used the I^2D protein interaction database (version 1.5) and identified known or predicted protein interaction partners for the 72 secreted proteins also significantly changed by microarray. Finally, we analyzed 3 more ovarian cancer ascites, and only considered proteins passing the above criteria, also found by mass spectrometry in all four ovarian cancer patients. This resulted in a panel of 80 putative ovarian cancer biomarkers for future investigation.

1.4. Proteins Associated with Cell Adhesion and Motility. Oncogenic signaling can stimulate cytoskeletal changes that in turn alter substrate adhesions, essential for cell spreading and motility. Several proteins identified in the present study

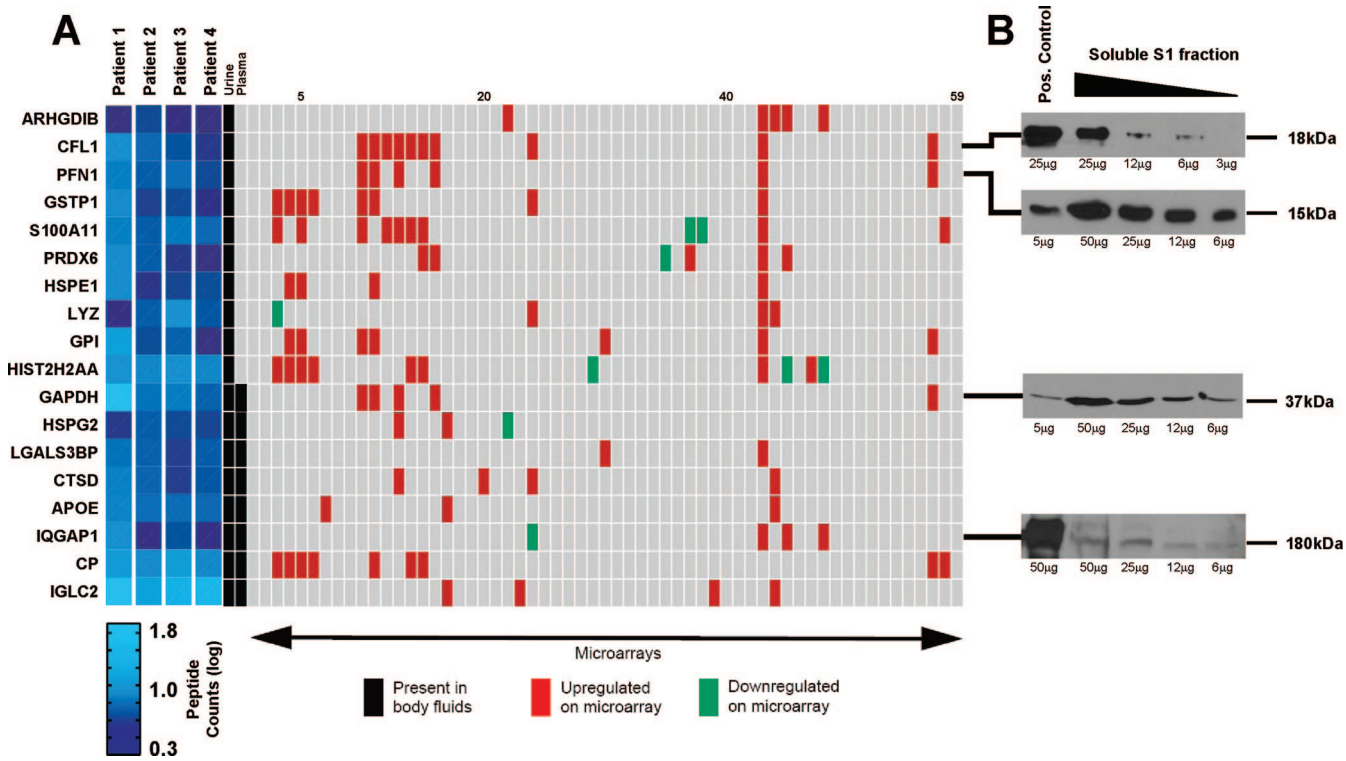

Figure 5. Visualization of 18 putative biomarkers. (A) Proteins detected in all four ovarian cancer ascites and previously reported in human body fluid proteomes (secreted proteins). Mapping to 59 ovarian cancer microarray data sets revealed significant changes on several arrays for all of these candidates. (B) Representative Western blots for candidate markers in ascites fraction S1. The ovarian cancer cell line ES-2 was used as a positive control. Different concentrations of ascites fluid (3–50 μg) were loaded on each gel.

are associated with cellular adhesion and motility. These include ARHGDIB, LGALS3BP, Basement membrane-specific heparan sulfate proteoglycan (HSPG2), and PFN1. Profilins are ubiquitous, 12–15 kDa proteins that regulate Actin polymerization by binding to and sequestering the Actin monomer.[66] Overexpression of PFN1 increases endothelial cell adhesion to fibronectin[67] and reduces the migration of invasive breast cancer cells.[68] Although decreased expression of PFN1 has been reported in invasive and migratory breast cancer cells, in our study, comparison to available microarray data revealed up-regulation of the PFN1 transcript in several grades and stages of ovarian cancer (Figure 5 and Supplementary Table S6 in Supporting Information).

1.5. Proteins Involved in Transport, Metabolic Processes, and Proteolysis. Several identified proteins are associated with transport, metabolic, and proteolytic processes. These include Ceruloplasmin (CP), Glucose-6-Phosphate Isomerase (GPI), Peroxiredoxin-6 (PRX6), and Cathepsin D (CTSD). CP is a plasma glycoprotein that transports copper throughout the body.[69] Interestingly, microarray analysis has linked this gene to tumor invasion and metastasis.[70] High serum levels of CP have been reported in many cancer patients and increase with tumor mass.[71,72] Its transcript was up-regulated in ovarian cancer compared to normal ovary, as well as in malignant versus benign cancer (Figure 5 and Supplementary Table S6 in Supporting Information) making it an interesting candidate for future investigation. Serum levels of GPI are known to be elevated in ovarian cancer patients.[73] It has been linked to tumor invasion and metastasis.[74,75] The robust detection of GPI in this study, its presence in human urine, and up-regulation in ovarian cancer suggest further investigation. The proteinase CTSD has been implicated in tumor invasion, metastasis,[76] tumor cell proliferation, apoptosis, and angiogenesis. It is abundantly detected in invasive ovarian cancer tumors and is associated with poor clinical outcome.[77] In the present study, Cathepsin-D was detected in ovarian cancer ascites and was previously detected in urine and plasma (Figure 5 and Supplementary Table S6 in Supporting Information).

We used Western blotting to detect the presence of four candidates in ascites fraction S1 (soluble fraction). Representative Western blots using commercially available antibodies against CFL1, PFN1, GAPDH, and IQGAP-1 are shown in Figure 5B. Future work will include the validation of these markers in collected and strictly controlled plasma samples, to fully evaluate their potential use as ovarian cancer biomarkers.

2. Proteins Identified in Ovarian Cancer Ascites but Not in Other Body Fluids. These included *Proliferation-associated protein 2G4* (PA2G4), a transcriptional corepressor of the androgen receptor involved in cell growth regulation and transcriptional inhibition of some E2F1-regulated promoters; *Proteasome activator complex subunit 2* (PSME2), implicated in immunoproteasome assembly required for efficient antigen processing; *150 kDa oxygen regulated protein* (HYOU1), which can trigger cytoprotective cellular mechanisms in a hypoxic environment; *Lamin B1* (LMNB1), a structural component of the nuclear lamina thought to play an important role in nuclear architecture, DNA replication, and gene expression (some evidence exists placing this protein as a potential marker for drug resistance and transformation to malignancy);[78] and *Myeloperoxidase* (MPO), which has microbicidal activity against a wide rage of organisms. Importantly, all these proteins were detected in all four ascites samples, showed significant changes on at least one ovarian cancer microarray, or were known or predicted protein interaction partners of secreted ascites

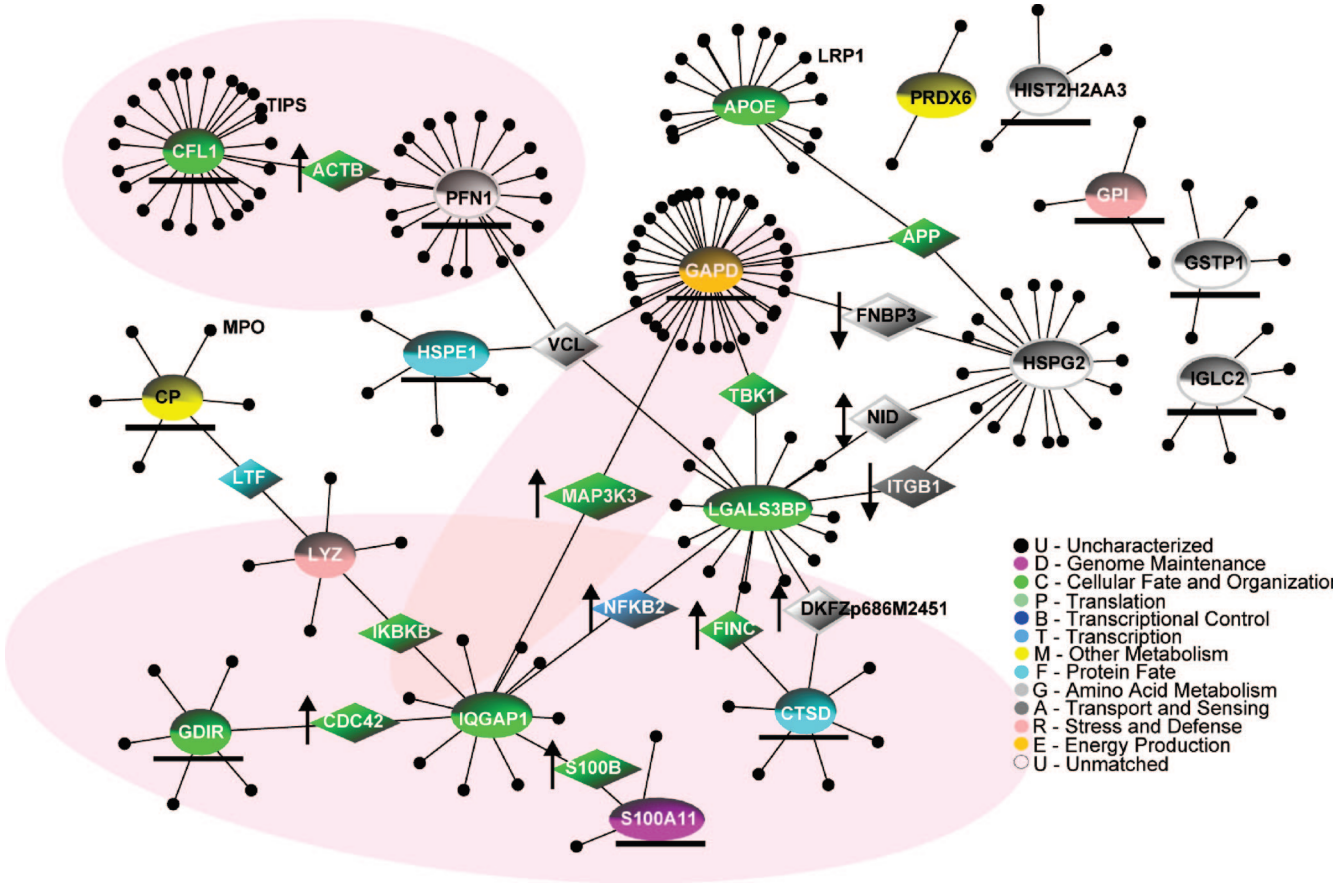

Figure 6. Protein–protein interaction network for 18 candidate biomarkers. The network was generated by querying the I^{2}D database for the 18 proteins, which resulted in a network with 201 proteins and 201 interactions. Importantly, while all 18 proteins have been significant on multiple microarray studies (ovals), we highlighted those that are the most frequent (underline). We have also highlighted subnetworks that are significantly up-regulated (shaded areas). Proteins in the network highlighted as diamonds are important interconnecting proteins, and we also show their up- or down-regulation on multiple microarray studies by arrows. The color represents Gene Ontology function.

proteins (Supplementary Tables S4 and S5 in Supporting Information).

3. Protein–Protein Interaction Networks. The protein–protein interaction network for the 18 candidate biomarkers (Figure 6) was generated by querying the I^{2}D database (version 1.5)[22] (called seeds and highlighted as ovals), which resulted in a sparse network with 201 proteins and 201 interactions. Importantly, while all 18 proteins are significant on multiple microarray studies, we highlighted those that are the most frequent (underlined). We have also highlighted with shaded areas, subnetworks that are significantly up-regulated. Interestingly, this includes both proteins identified in human body fluids and those that were not, but were significantly deregulated on multiple microarrays. There is a strong enrichment of finding the 18 putative markers significantly deregulated on multiple microarrays compared to random chance, where less than half of these seed proteins are significantly deregulated (using a random secreted protein network, $p = 2.6383 \times 10^{-7}$, or using a random protein network $p = 2.1179 \times 10^{-9}$).

While the proteins are not highly interconnected and on average do not have a high degree, they form a large main connected subgraph. To determine if this can happen by chance, we compared the resulting network to random networks generated from 18 random secreted proteins and 18 random proteins. Interestingly, both comparisons show significantly larger size of the largest connected subgraph for our network ($p = 0.000115438$ for random secreted proteins; $p = 4.0451 \times 10^{-5}$ for the random proteins).

Proteins in the network highlighted as diamonds are proteins interconnecting seed proteins, and we show their up- or down-regulation on ovarian cancer microarray studies by arrows. Interestingly, these proteins are significantly enriched for being deregulated on multiple microarray studies; no interconnecting proteins are significantly deregulated in average in random networks (random secreted protein network $p = 4.73864 \times 10^{-11}$; random protein network $p = 2.63083 \times 10^{-54}$). The network for the 10 proteins (Supplementary Figure S3 in Supporting Information) shows similar significant results (data not shown).

Both 18- and 10-seed networks have lower average degree and are not highly interconnected compared to random networks. This could be caused by two phenomena. First, it is possible that these putative biomarkers represent proteins that are less studied in average (although there are exceptions such as CFL1 and PFN1) compared to any randomly selected proteins from the current interaction network. Second, it may be possible that secreted biomarkers form a network of different topology, a network that is more planar and less interconnected, compared to signaling networks. Most importantly, neither of the two types of random networks have many seed proteins significantly deregulated on multiple arrays, and in average none of the interconnected proteins have such

support. This makes our 18- and 10-protein networks much more relevant to ovarian cancer biology.

Conclusion

Although our proteomic investigation of ovarian cancer ascites identified multiple potential biomarkers of ovarian cancer, we are aware that not every known ascites protein was detected in this study. These include growth factors such as EGF, TGF-β, and VEGF. There are many possible reasons these secreted growth factors might have been missed in this study, including low abundance and overall complexity of the sample. More directed technologies, such as MRM, should be applied for the detection of such low-abundance, secreted proteins. Furthermore, we only considered the most robustly detected proteins (e.g., detected in all four patients and supported by microarray analysis) as putative markers for future evaluation, which clearly favors higher abundance proteins. Our goal for this study was the identification of the most robust candidates. The current ovarian cancer biomarker of choice, mucin 16 (CA-125), was detected only in the first patient by gel-enhanced LC−MS. Metalloproteinase-9, a protein previously proposed as a biomarker for ovarian cancer,[79] was detected in three of the four patients, rejecting it from our high-stringency panel of putative biomarkers. Nevertheless, the entire panel of detected proteins (Supplementary Table S1 in Supporting Information) is available to the public for further investigation. We have also deposited the RAW-data to the Tranche (www.proteomecommons.org) server for further data mining by others.

Proteomics of ascites fluid could serve as a valuable tool to study the mechanisms of acquired chemoresistance. Whereas most women respond to the initial platinum-based chemotherapy, resistance to the treatment develops in all cases of recurrent disease. This issue was previously studied by microarray gene chip analyses.[80–82] A limited number of proteomic investigations have addressed the issue of chemoresistance, mainly by comparing chemo-sensitive to chemo-resistant cell lines, and several proteins with differential expression patterns have been identified in these studies.[83,84] Although our study did not aim to discover proteins involved in acquired chemoresistance, we believe that ascites could be a valuable source to address this topic by proteomics.

In summary, we present for the first time a high-quality, in-depth proteomics analysis of ovarian cancer ascites. Integrated bioinformatics analyses of this data with previously published body fluid proteomes, as well as 59 available ovarian cancer microarray data sets, allowed prioritized data analysis, with the goal to identify putative biomarkers for future investigation. We present a panel of 80 proteins, robustly detected in all four cases of ovarian cancer ascites, including several novel markers for future investigation.

Acknowledgment. T.K. and I.J. thank the Canadian Research Chair Program for infrastructure support. This work was supported by a start-up grant from the Ontario Cancer Institute and the Toronto Fashion Show and grant support form the CR Younger Foundation to T.K. and by grant support from the Department of Defense (DOD #W81XWH-05-1-0104), Genome Canada through the Ontario Genomics Institute, and IBM Canada to I.J.

Supporting Information Available: Figures of SDS-PAGE used for gel-enhanced LC−MS, cytological analysis of a cytospin preparation, network for 10 proteins detected in ascites, but not in other body fluids; tables listing proteins detected in human ovarian cancer ascites, the most significantly enriched Gene Onology terms, proteins detected and the number of identified peptides in four ascites patients and previously published urine and plasma datasets, proteins detected and their unique peptides in four ascites samples but not in healthy urine or plasma, proteins detected in four ascites samples, which are known or predicted interactors of the candidates from Supplementary Table S4, and description and number code for the microarray data set used in this study. Supporting Information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Jemal, A.; Siegel, R.; Ward, E.; Murray, T.; Xu, J.; Smigal, C.; Thun, M. J. Cancer statistics, 2006. *CA Cancer J. Clin.* **2006**, *56* (2), 106–130.

(2) Sankaranarayanan, R.; Ferlay, J. Worldwide burden of gynaecological cancer: the size of the problem. *Best Pract. Res. Clin. Obstet. Gynaecol.* **2006**, *20* (2), 207–225.

(3) Abramov, Y.; Anteby, S. O.; Fasouliotis, S. J.; Barak, V. The role of inflammatory cytokines in Meigs' syndrome. *Obstet. Gynecol.* **2002**, *99* (5 Pt 2), 917–919.

(4) Davidson, B.; Reich, R.; Kopolovic, J.; Berner, A.; Nesland, J. M.; Kristensen, G. B.; Trope, C. G.; Bryne, M.; Risberg, B.; van de Putte, G.; Goldberg, I. Interleukin-8 and vascular endothelial growth factor mRNA and protein levels are down-regulated in ovarian carcinoma cells in serous effusions. *Clin. Exp. Metastasis* **2002**, *19* (2), 135–144.

(5) Fang, X.; Yu, S.; Bast, R. C.; Liu, S.; Xu, H. J.; Hu, S. X.; LaPushin, R.; Claret, F. X.; Aggarwal, B. B.; Lu, Y.; Mills, G. B. Mechanisms for lysophosphatidic acid-induced cytokine production in ovarian cancer cells. *J. Biol. Chem.* **2004**, *279* (10), 9653–9661.

(6) Feldman, G. B.; Knapp, R. C.; Order, S. E.; Hellman, S. The role of lymphatic obstruction in the formation of ascites in a murine ovarian carcinoma. *Cancer Res.* **1972**, *32* (8), 1663–1666.

(7) Zebrowski, B. K.; Yano, S.; Liu, W.; Shaheen, R. M.; Hicklin, D. J.; Putnam, J. B., Jr.; Ellis, L. M. Vascular endothelial growth factor levels and induction of permeability in malignant pleural effusions. *Clin. Cancer Res.* **1999**, *5* (11), 3364–8336.

(8) Berchuck, A.; Carney, M. Human ovarian cancer of the surface epithelium. *Biochem. Pharmacol.* **1997**, *54* (5), 541–544.

(9) Bookman, M. A. Biological therapy of ovarian cancer: current directions. *Semin. Oncol.* **1998**, *25* (3), 381–396.

(10) Westermann, A. M.; Beijnen, J. H.; Moolenaar, W. H.; Rodenhuis, S. Growth factors in human ovarian cancer. *Cancer Treat. Rev.* **1997**, *23* (2), 113–131.

(11) Mills, G. B.; May, C.; McGill, M.; Roifman, C. M.; Mellors, A. A putative new growth factor in ascitic fluid from ovarian cancer patients: identification, characterization, and mechanism of action. *Cancer Res.* **1988**, *48* (5), 1066–1071.

(12) Adachi, J.; Kumar, C.; Zhang, Y.; Olsen, J. V.; Mann, M. The human urinary proteome contains more than 1500 proteins including a large proportion of membranes proteins. *GenomeBiology* **2006**, *7* (9), R80.

(13) Omenn, G. S.; States, D. J.; Adamski, M.; Blackwell, T. W.; Menon, R.; Hermjakob, H.; Apweiler, R.; Haab, B. B.; Simpson, R. J.; Eddes, J. S.; Kapp, E. A.; Moritz, R. L.; Chan, D. W.; Rai, A. J.; Admon, A.; Aebersold, R.; Eng, J.; Hancock, W. S.; Hefta, S. A.; Meyer, H.; Paik, Y. K.; Yoo, J. S.; Ping, P.; Pounds, J.; Adkins, J.; Qian, X.; Wang, R.; Wasinger, V.; Wu, C. Y.; Zhao, X.; Zeng, R.; Archakov, A.; Tsugita, A.; Beer, I.; Pandey, A.; Pisano, M.; Andrews, P.; Tammen, H.; Speicher, D. W.; Hanash, S. M. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* **2005**, *5* (13), 3226–3245.

(14) Xu, Y.; Shen, Z.; Wiper, D. W.; Wu, M.; Morton, R. E.; Elson, P.; Kennedy, A. W.; Belinson, J.; Markman, M.; Casey, G. Lysophosphatidic acid as a potential biomarker for ovarian and other gynecologic cancers. *JAMA, J. Am. Med. Assoc.* **1998**, *280* (8), 719–723.

(15) Gericke, B.; Raila, J.; Sehouli, J.; Haebel, S.; Konsgen, D.; Mustea, A.; Schweigert, F. J. Microheterogeneity of transthyretin in serum and ascitic fluid of ovarian cancer patients. *BMC Cancer* **2005**, *5*, 133.

(16) Kislinger, T.; Cox, B.; Kannan, A.; Chung, C.; Hu, P.; Ignatchenko, A.; Scott, M. S.; Gramolini, A. O.; Morris, Q.; Hallett, M. T.; Rossant, J.; Hughes, T. R.; Frey, B.; Emili, A. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* **2006**, *125* (1), 173–186.
(17) Kislinger, T.; Gramolini, A. O.; Pan, Y.; Rahman, K.; MacLennan, D. H.; Emili, A. Proteome dynamics during C2C12 myoblast differentiation. *Mol. Cell. Proteomics* **2005**, *4* (7), 887–901.
(18) Kislinger, T.; Rahman, K.; Radulovic, D.; Cox, B.; Rossant, J.; Emili, A. PRISM, a generic large scale proteomic investigation strategy for mammals. *Mol. Cell. Proteomics* **2003**, *2* (2), 96–106.
(19) Washburn, M. P.; Wolters, D.; Yates, J. R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242–247.
(20) Sun, H.; Fang, H.; Chen, T.; Perkins, R.; Tong, W. GOFFA: Gene Ontology For Functional Analysis—A FDA gene ontology tool for analysis of genomic and proteomic data. *BMC Bioinf.* **2006**, *7 Suppl 2*, S23.
(21) Brown, K. R.; Jurisica, I. Online predicted human interaction database. *Bioinformatics* **2005**, *21* (9), 2076–2082.
(22) Brown, K. R.; Jurisica, I. Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome-Biology* **2007**, *8* (5), R95.
(23) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **1996**, *68* (5), 850–858.
(24) Craig, R.; Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004**, *20* (9), 1466–1467.
(25) Kersey, P. J.; Duarte, J.; Williams, A.; Karavidopoulou, Y.; Birney, E.; Apweiler, R. The International Protein Index: an integrated database for proteomics experiments. *Proteomics* **2004**, *4* (7), 1985–1988.
(26) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2003**, *2* (1), 43–50.
(27) Elias, J. E.; Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **2007**, *4* (3), 207–214.
(28) Kapp, E. A.; Schutz, F.; Connolly, L. M.; Chakel, J. A.; Meza, J. E.; Miller, C. A.; Fenyo, D.; Eng, J. K.; Adkins, J. N.; Omenn, G. S.; Simpson, R. J. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. *Proteomics* **2005**, *5* (13), 3475–3490.
(29) Tabb, D. L.; Fernando, C. G.; Chambers, M. C. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J. Proteome Res.* **2007**, *6* (2), 654–661.
(30) Adamski, M.; Blackwell, T.; Menon, R.; Martens, L.; Hermjakob, H.; Taylor, C.; Omenn, G. S.; States, D. J. Data management and preliminary data analysis in the pilot phase of the HUPO Plasma Proteome Project. *Proteomics* **2005**, *5* (13), 3246–3261.
(31) Adachi, J.; Kumar, C.; Zhang, Y.; Mann, M. In-depth analysis of the adipocyte proteome by mass spectrometry and bioinformatics. *Mol. Cell. Proteomics* **2007**, *6*, 1257–1273.
(32) Barker, S. D.; Casado, E.; Gomez-Navarro, J.; Xiang, J.; Arafat, W.; Mahasreshti, P.; Pustilnik, T. B.; Hemminki, A.; Siegal, G. P.; Alvarez, R. D.; Curiel, D. T. An immunomagnetic-based method for the purification of ovarian cancer cells from patient-derived ascites. *Gynecol. Oncol.* **2001**, *82* (1), 57–63.
(33) Kislinger, T.; Gramolini, A. O.; MacLennan, D. H.; Emili, A. Multidimensional protein identification technology (MudPIT): technical overview of a profiling method optimized for the comprehensive proteomic investigation of normal and diseased heart tissue. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (8), 1207–1220.
(34) Koziol, J. A.; Feng, A. C.; Schnitzer, J. E. Application of capture-recapture models to estimation of protein count in MudPIT experiments. *Anal. Chem.* **2006**, *78* (9), 3203–3207.
(35) Liu, H.; Sadygov, R. G.; Yates, J. R., III. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193–4201.
(36) Eng, J. K.; McCormack, A. L.; Yates, J. R. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *11*, 976–989.
(37) Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551–3567.
(38) Anderson, N. L.; Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **2002**, *1* (11), 845–867.
(39) Gamberoni, G.; Storari, S.; Volinia, S. Finding biological process modifications in cancer tissues by mining gene expression correlations. *BMC Bioinf.* **2006**, *7*, 6.
(40) Patil, M. A.; Chua, M. S.; Pan, K. H.; Lin, R.; Lih, C. J.; Cheung, S. T.; Ho, C.; Li, R.; Fan, S. T.; Cohen, S. N.; Chen, X.; So, S. An integrated data analysis approach to characterize genes highly expressed in hepatocellular carcinoma. *Oncogene* **2005**, *24* (23), 3737–3747.
(41) Chen, E. I.; Hewel, J.; Krueger, J. S.; Tiraby, C.; Weber, M. R.; Kralli, A.; Becker, K.; Yates, J. R., III; Felding-Habermann, B. Adaptation of energy metabolism in breast cancer brain metastases. *Cancer Res.* **2007**, *67* (4), 1472–1486.
(42) Engwegen, J. Y.; Gast, M. C.; Schellens, J. H.; Beijnen, J. H. Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. *Trends Pharmacol. Sci.* **2006**, *27* (5), 251–259.
(43) Gogoi, R.; Srinivasan, S.; Fishman, D. A. Progress in biomarker discovery for diagnostic testing in epithelial ovarian cancer. *Expert Rev. Mol. Diagn.* **2006**, *6* (4), 627–637.
(44) Bast, R. C., Jr.; Klug, T. L.; St John, E.; Jenison, E.; Niloff, J. M.; Lazarus, H.; Berkowitz, R. S.; Leavitt, T.; Griffiths, C. T.; Parker, L.; Zurawski, V. R., Jr.; Knapp, R. C. A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N. Engl. J. Med.* **1983**, *309* (15), 883–887.
(45) Markman, M. Limitations to the use of the CA-125 antigen level in ovarian cancer. *Curr. Oncol. Rep.* **2003**, *5* (4), 263–264.
(46) Fujimoto, J.; Sakaguchi, H.; Hirose, R.; Ichigo, S.; Tamaya, T. Biologic implications of the expression of vascular endothelial growth factor subtypes in ovarian carcinoma. *Cancer* **1998**, *83* (12), 2528–2533.
(47) Yamamoto, S.; Konishi, I.; Mandai, M.; Kuroda, H.; Komatsu, T.; Nanbu, K.; Sakahara, H.; Mori, T. Expression of vascular endothelial growth factor (VEGF) in epithelial ovarian neoplasms: correlation with clinicopathology and patient survival, and analysis of serum VEGF levels. *Br. J. Cancer* **1997**, *76* (9), 1221–1227.
(48) Fishman, D. A.; Bafetti, L. M.; Banionis, S.; Kearns, A. S.; Chilukuri, K.; Stack, M. S. Production of extracellular matrix-degrading proteinases by primary cultures of human epithelial ovarian carcinoma cells. *Cancer* **1997**, *80* (8), 1457–1463.
(49) Moser, T. L.; Young, T. N.; Rodriguez, G. C.; Pizzo, S. V.; Bast, R. C., Jr.; Stack, M. S. Secretion of extracellular matrix-degrading proteinases is increased in epithelial ovarian carcinoma. *Int. J. Cancer* **1994**, *56* (4), 552–559.
(50) Diamandis, E. P.; Yousef, G. M.; Soosaipillai, A. R.; Bunting, P. Human kallikrein 6 (zyme/protease M/neurosin): a new serum biomarker of ovarian carcinoma. *Clin. Biochem.* **2000**, *33* (7), 579–583.
(51) Dong, Y.; Kaushal, A.; Bui, L.; Chu, S.; Fuller, P. J.; Nicklin, J.; Samaratunga, H.; Clements, J. A. Human kallikrein 4 (KLK4) is highly expressed in serous ovarian carcinomas. *Clin. Cancer Res.* **2001**, *7* (8), 2363–2371.
(52) Luo, L. Y.; Bunting, P.; Scorilas, A.; Diamandis, E. P. Human kallikrein 10: a novel tumor marker for ovarian carcinoma. *Clin. Chim. Acta* **2001**, *306* (1–2), 111–118.
(53) Cox, B.; Kislinger, T.; Wigle, D. A.; Kannan, A.; Brown, K.; Okubo, T.; Hogan, B.; Jurisica, I.; Frey, B.; Rossant, J.; Emili, A. Integrated proteomic and transcriptomic profiling of mouse lung development and Nmyc target genes. *Mol. Syst. Biol.* **2007**, *3*, 109.
(54) Makino, E.; Sakaguchi, M.; Iwatsuki, K.; Huh, N. H. Introduction of an N-terminal peptide of S100C/A11 into human cells induces apoptotic cell death. *J. Mol. Med.* **2004**, *82* (9), 612–620.
(55) Kikuchi, N.; Horiuchi, A.; Osada, R.; Imai, T.; Wang, C.; Chen, X.; Konishi, I. Nuclear expression of S100A4 is associated with aggressive behavior of epithelial ovarian carcinoma: an important autocrine/paracrine factor in tumor progression. *Cancer Sci.* **2006**, *97* (10), 1061–1069.
(56) Trachte, A. L.; Suthers, S. E.; Lerner, M. R.; Hanas, J. S.; Jupe, E. R.; Sienko, A. E.; Adesina, A. M.; Lightfoot, S. A.; Brackett, D. J.; Postier,R. G. Increased expression of alpha-1-antitrypsin, glutathione S-transferase pi and vascular endothelial growth factor in human pancreatic adenocarcinoma. *Am. J. Surg.* **2002**, *184* (6), 642−647, discussion 647–648.
(57) Boss, E. A.; Peters, W. H.; Roelofs, H. M.; Boonstra, H.; Steegers, E. A.; Massuger, L. F. Glutathione S-transferases P1–1 and A1–1 in ovarian cyst fluids. *Eur. J. Gynaecol. Oncol.* **2001**, *22* (6), 427–432.
(58) Hough, C. D.; Sherman-Baust, C. A.; Pizer, E. S.; Montz, F. J.; Im, D. D.; Rosenshein, N. B.; Cho, K. R.; Riggins, G. J.; Morin, P. J.

Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. *Cancer Res.* **2000**, *60* (22), 6281–6287.
(59) Chen, Y. C.; Pohl, G.; Wang, T. L.; Morin, P. J.; Risberg, B.; Kristensen, G. B.; Yu, A.; Davidson, B.; Shih Ie, M. Apolipoprotein E is required for cell proliferation and survival in ovarian cancer. *Cancer Res.* **2005**, *65* (1), 331–337.
(60) Tavazoie, S. F.; Alvarez, V. A.; Ridenour, D. A.; Kwiatkowski, D. J.; Sabatini, B. L. Regulation of neuronal morphology and function by the tumor suppressors Tsc1 and Tsc2. *Nat. Neurosci.* **2005**, *8* (12), 1727–1734.
(61) Wang, W.; Mouneimne, G.; Sidani, M.; Wyckoff, J.; Chen, X.; Makris, A.; Goswami, S.; Bresnick, A. R.; Condeelis, J. S. The activity status of cofilin is directly related to invasion, intravasation, and metastasis of mammary tumors. *J. Cell Biol.* **2006**, *173* (3), 395–404.
(62) Mataraza, J. M.; Briggs, M. W.; Li, Z.; Entwistle, A.; Ridley, A. J.; Sacks, D. B. IQGAP1 promotes cell motility and invasion. *J. Biol. Chem.* **2003**, *278* (42), 41237–41245.
(63) Meng, X.; Krokhin, O.; Cheng, K.; Ens, W.; Wilkins, J. A. Characterization of IQGAP1-containing complexes in NK-like cells: evidence for Rac 2 and RACK1 association during homotypic adhesion. *J. Proteome Res.* **2007**, *6* (2), 744–750.
(64) Yamaoka-Tojo, M.; Ushio-Fukai, M.; Hilenski, L.; Dikalov, S. I.; Chen, Y. E.; Tojo, T.; Fukai, T.; Fujimoto, M.; Patrushev, N. A.; Wang, N.; Kontos, C. D.; Bloom, G. S.; Alexander, R. W. IQGAP1, a novel vascular endothelial growth factor receptor binding protein, is involved in reactive oxygen species--dependent endothelial migration and proliferation. *Circ. Res.* **2004**, *95* (3), 276–283.
(65) Yamaoka-Tojo, M.; Tojo, T.; Kim, H. W.; Hilenski, L.; Patrushev, N. A.; Zhang, L.; Fukai, T.; Ushio-Fukai, M. IQGAP1 mediates VE-cadherin-based cell-cell contacts and VEGF signaling at adherence junctions linked to angiogenesis. *Arterioscler., Thromb., Vasc. Biol.* **2006**, *26* (9), 1991–1997.
(66) Babcock, G.; Rubenstein, P. A. Control of profilin and Actin expression in muscle and nonmuscle cells. *Cell Motil. Cytoskeleton* **1993**, *24* (3), 179–188.
(67) Moldovan, N. I.; Milliken, E. E.; Irani, K.; Chen, J.; Sohn, R. H.; Finkel, T.; Goldschmidt-Clermont, P. J. Regulation of endothelial cell adhesion by profilin. *Curr. Biol.* **1997**, *7* (1), 24–30.
(68) Roy, P.; Jacobson, K. Overexpression of profilin reduces the migration of invasive breast cancer cells. *Cell Motil. Cytoskeleton* **2004**, *57* (2), 84–95.
(69) Luza, S. C.; Speisky, H. C. Liver copper storage and transport during development: implications for cytotoxicity. *Am. J. Clin. Nutr.* **1996**, *63* (5), 812S–820S.
(70) Kluger, H. M.; Kluger, Y.; Gilmore-Hebert, M.; DiVito, K.; Chang, J. T.; Rodov, S.; Mironenko, O.; Kacinski, B. M.; Perkins, A. S.; Sapi, E. cDNA microarray analysis of invasive and tumorigenic phenotypes in a breast cancer model. *Lab. Invest.* **2004**, *84* (3), 320–331.
(71) Sergeev, A. I.; Naidich, V. I.; Bunto, T. V. [Study of the kinetics of ascitic tumor development by proton magnetic relaxation]. *Dokl. Akad. Nauk SSSR* **1978**, *238* (4), 974–976.
(72) Vachtenheim, J.; Pavel, S.; Duchon, J. Dopa oxidase activity and ceruloplasmin in the sera of hamsters with melanoma. *Neoplasma* **1981**, *28* (1), 59–65.
(73) Yeshowardhana; Singh, V. S. Significance of serum phosphohexose isomerase, hexokinase and aldolase in carcinoma ovary. *Indian J. Physiol. Pharmacol.* **1985**, *29* (1), 51–54.
(74) Torimura, T.; Ueno, T.; Kin, M.; Harada, R.; Nakamura, T.; Kawaguchi, T.; Harada, M.; Kumashiro, R.; Watanabe, H.; Avraham, R.; Sata, M. Autocrine motility factor enhances hepatoma cell invasion across the basement membrane through activation of beta1 integrins. *Hepatology* **2001**, *34* (1), 62–71.
(75) Haga, A. [Possibility that AMF will serve as a target molecule for the diagnosis and treatment of a metastatic neoplasm]. *Yakugaku Zasshi* **2005**, *125* (2), 169–175.
(76) Leto, G.; Gebbia, N.; Rausa, L.; Tumminello, F. M. Cathepsin D in the malignant progression of neoplastic diseases (review). *Anticancer Res.* **1992**, *12* (1), 235–240.
(77) Losch, A.; Schindl, M.; Kohlberger, P.; Lahodny, J.; Breitenecker, G.; Horvat, R.; Birner, P. Cathepsin D in ovarian cancer: prognostic value and correlation with p53 expression and microvessel density. *Gynecol. Oncol.* **2004**, *92* (2), 545–552.
(78) de Angelis, P. M.; Fjell, B.; Kravik, K. L.; Haug, T.; Tunheim, S. H.; Reichelt, W.; Beigi, M.; Clausen, O. P.; Galteland, E.; Stokke, T. Molecular characterizations of derivatives of HCT116 colorectal cancer cells that are resistant to the chemotherapeutic agent 5-fluorouracil. *Int. J. Oncol.* **2004**, *24* (5), 1279–1288.
(79) Ozalp, S.; Tanir, H. M.; Yalcin, O. T.; Kabukcuoglu, S.; Oner, U.; Uray, M. Prognostic value of matrix metalloproteinase-9 (gelatinase-B) expression in epithelial ovarian tumors. *Eur. J. Gynaecol. Oncol.* **2003**, *24* (5), 417–420.
(80) Bachvarov, D.; L'Esperance, S.; Popa, I.; Bachvarova, M.; Plante, M.; Tetu, B. Gene expression patterns of chemoresistant and chemosensitive serous epithelial ovarian tumors with possible predictive value in response to initial chemotherapy. *Int. J. Oncol.* **2006**, *29* (4), 919–933.
(81) Jazaeri, A. A.; Awtrey, C. S.; Chandramouli, G. V.; Chuang, Y. E.; Khan, J.; Sotiriou, C.; Aprelikova, O.; Yee, C. J.; Zorn, K. K.; Birrer, M. J.; Barrett, J. C.; Boyd, J. Gene expression profiles associated with response to chemotherapy in epithelial ovarian cancers. *Clin. Cancer Res.* **2005**, *11* (17), 6300–6310.
(82) Spentzos, D.; Levine, D. A.; Kolia, S.; Otu, H.; Boyd, J.; Libermann, T. A.; Cannistra, S. A. Unique gene expression profile based on pathologic response in epithelial ovarian cancer. *J. Clin. Oncol.* **2005**, *23* (31), 7911–7918.
(83) Le Moguen, K.; Lincet, H.; Deslandes, E.; Hubert-Roux, M.; Lange, C.; Poulain, L.; Gauduchon, P.; Baudin, B. Comparative proteomic analysis of cisplatin sensitive IGROV1 ovarian carcinoma cell line and its resistant counterpart IGROV1-R10. *Proteomics* **2006**, *6* (19), 5183–5192.
(84) Stewart, J. J.; White, J. T.; Yan, X.; Collins, S.; Drescher, C. W.; Urban, N. D.; Hood, L.; Lin, B. Proteins associated with Cisplatin resistance in ovarian cancer cells identified by quantitative proteomic technology and integrated with mRNA expression levels. *Mol. Cell. Proteomics* **2006**, *5* (3), 433–443.
(85) States, D. J.; Omenn, G. S.; Blackwell, T. W.; Fermin, D.; Eng, J.; Speicher, D. W.; Hanash, S. M. Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat. Biotechnol.* **2006**, *24* (3), 333–338.

PR0703223