Microvesicle Proteomic Profiling of Uterine Liquid Biopsy for Ovarian Cancer Early Detection

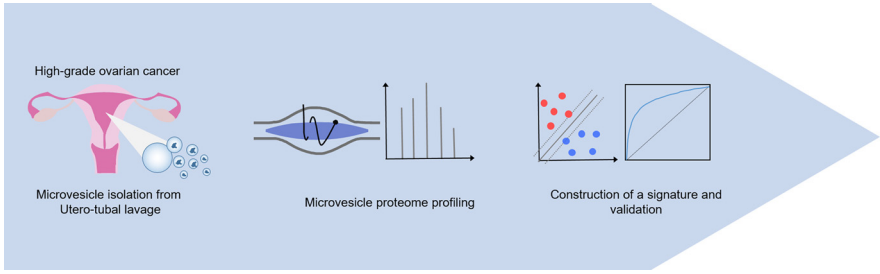
Authors

Georgina D. Barnabas, Keren Bahar-Shany, Stav Sapoznik, Limor Helpman, Yfat Kadan, Mario Beiner, Omer Weitzner, Nissim Arbib, Jacob Korach, Tamar Perri, Guy Katz, Anna Blecher, Benny Brandt, Eitan Friedman, David Stockheim, Ariella Jakobson-Setton, Ram Eitan, Shunit Armon, Hadar Brand, Oranit Zadok, Sarit Aviel-Ronen, Michal Harel, Tamar Geiger, and Keren Levanon

Correspondence

Keren.Levanon@sheba.health.gov.il; geiger@tauex.tau.ac.il

Graphical Abstract



In Brief

High-grade ovarian cancer accounts for higher mortality rates because of ineffective biomarkers for early diagnosis. Deep proteome profiling of the microvesicles from a total of 187 liquid biopsies of Utero-tubal Lavage, combined with support vector machine algorithms, extracted a 9-protein classifier with high accuracy. The signature predicted all the early stage lesions, and outperformed the known markers CA125 and HE4 with 70% sensitivity and 76.2% specificity. Our study reveals UtL-microvesicle proteomics as the potential biomarker source for early diagnosis of HGOC.

Highlights

- Microvesicle proteomics of 187 utero-tubal lavage samples for early diagnosis of HGOC.
- Machine learning-based classification of a 9-protein signature with high predictive power.
- Signature has 70% sensitivity and 76.2% specificity, predicting stage I lesions.

Barnabas et al., 2019, Molecular & Cellular Proteomics *18*, 865–875

https://doi.org/10.1074/mcp.RA119.001362

Microvesicle Proteomic Profiling of Uterine Liquid Biopsy for Ovarian Cancer Early Detection*[S]

Georgina D. Barnabas‡‖‖‖, Keren Bahar-Shany§‖‖‖, Stav Sapoznik§, Limor Helpman¶‖‡‡‡‡, Yfat Kadan¶, Mario Beiner¶, Omer Weitzner**, Nissim Arbib**, Jacob Korach‖‡‡, Tamar Perri‖‡‡, Guy Katz‡‡, Anna Blecher‡‡, Benny Brandt‡‡, Eitan Friedman‖§§, David Stockheim‖¶¶, Ariella Jakobson-Setton‖‖‖, Ram Eitan‖‖‖, Shunit Armon‡‡‡, Hadar Brand§‖, Oranit Zadok§§§, Sarit Aviel-Ronen§§§¶¶¶, Michal Harel‡, Tamar Geiger‡‖‖‖¶¶¶¶, and Keren Levanon§‖¶¶¶‖‖‖§§§§

High-grade ovarian cancer (HGOC) is the leading cause of mortality from gynecological malignancies, because of diagnosis at a metastatic stage. Current screening options fail to improve mortality because of the absence of early-stage-specific biomarkers. We postulated that a liquid biopsy, such as utero-tubal lavage (UtL), may identify localized lesions better than systemic approaches of serum/plasma analysis. Further, while mutation-based assays are challenged by the rarity of tumor DNA within nonmutated DNA, analyzing the proteomic profile, is expected to enable earlier detection, as it reveals perturbations in both the tumor as well as in its microenvironment. To attain deep proteomic coverage and overcome the high dynamic range of this body fluid, we applied our method for microvesicle proteomics to the UtL samples. Liquid biopsies from HGOC patients (n = 49) and controls (n = 127) were divided into a discovery and validation sets. Data-dependent analysis of the samples on the Q-Exactive mass spectrometer provided depth of 8578 UtL proteins in total, and on average ~3000 proteins per sample. We used support vector machine algorithms for sample classification, and crossed three feature-selection algorithms, to construct and validate a 9-protein classifier with 70% sensitivity and 76.2% specificity. The signature correctly identified all Stage I lesions. These results demonstrate the potential power of microvesicle-based proteomic biomarkers for early cancer diagnosis. ***Molecular & Cellular Proteomics 18: 865–875, 2019. DOI: 10.1074/mcp.RA119.001362.***

Overall survival of patients with high-grade ovarian cancer (HGOC)[1] correlates with disease stage at diagnosis: whereas patients with stage I disease have >90% 5-year overall survival, rates for stage IV disease are extremely low. Regrettably, ~75% of HGOC cases are diagnosed at late-stage regardless of adherence to testing recommendations (1). Early-detection of HGOC among high-risk population, such as germline *BRCA1/2* mutation carriers, is of exceptional importance. These women are currently counseled to undergo prophylactic removal of the ovaries and fallopian tubes (risk reducing bilateral salpingo-oophorectomy, RRBSO) at age ~40, therefore there is an urgent unmet need for a personalized risk-assignment modality to guide RRBSO timing and alleviate unnecessary morbidity of early menopause (2, 3). This grim reality stems primarily from the lack of effective screening methods and of early-stage biomarkers. A multitude of biomarkers have been proposed and tested over the years, however even the most established markers, namely serum CA125 and HE4, have not proven to be effective in improving survival (4–8). Several recent large-scale screening trials based on blood CA125-based monitoring, with or without transvaginal ultrasound, showed insignificant stage shift among high-risk population and low specificity and sensitivity (9–11). Recently, blood-miRNA signatures have been proposed as highly sensitive and specific biomarkers, though the technicalities and their utility for early detection are not yet established (12–14).

From the ‡Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel; §Sheba Cancer Research Center, Chaim Sheba Medical Center, Ramat Gan, Israel; ¶Division of Gynecologic Oncology, Meir Medical Center, Kfar Saba, Israel; ‖Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel; **Department of Obstetrics and Gynecology, Meir Medical Center, Kfar Saba, Israel; ‡‡Department of Gynecologic Oncology, Chaim Sheba Medical Center, Ramat Gan, Israel; §§The Susanne-Levy Gertner Oncogenetics Unit, Chaim Sheba Medical Center, Ramat Gan, Israel; ¶¶Department of Gynecology, Chaim Sheba Medical Center, Ramat Gan, Israel; ‖‖‖Department of Gynecologic Oncology, Rabin Medical Center, Petah Tikva, Israel; ‡‡‡Department of Obstetrics & Gynecology, Shaare Zedek Medical Center, Jerusalem, Israel; §§§Department of Pathology, Chaim Sheba Medical Center, Ramat Gan, Israel; ¶¶¶The Talpiot Medical Leadership Program, Chaim Sheba Medical Center, Ramat Gan, Israel

Received January 29, 2019
Published, MCP Papers in Press, February 13, 2019, DOI 10.1074/mcp.RA119.001362

TABLE I
Patient characteristics for UtL samples included in the proteomic analysis

Clinical Characteristics	Discovery set			Validation set		
	No.	(%)	Age (ave.)	No.	(%)	Age (ave.)
Entire cohort	**24**		**57.4**	**152**		**53**
Patient cohort:	**12**	**100**	**60.6**	**37**	**100**	**62.3**
Type of surgery:						
Primary debulking	12	100	60.6	15	40.5	59
Interval debulking	0	0	NA	22	59.5	64.5
Stage:						
Early stage (STIC-I-II)	3	25	57	1	2.7	48
Late stage (III-IV)	9	75	61.8	36	97.3	62.3
BRCA status:						
Germline mutation	0	0	NA	10	24.3	54.1
No mutation	6	50	58.5	12	35.1	63.3
Unknown	6	50	62.7	15	40.5	66.9
Control cohort:	**12**	**100**	**54.2**	**115**	**100**	**50.1**
Indication for surgery:						
Benign ovarian mass	6	50	46	28	23.9	54.8
Endometrial polyp	3	25	61.7	9	7.7	61.7
Menometrorrhagia	0	0	NA	12	10.3	49.8
Uterine prolapse	1	8	74	14	12	62.4
Leiomyomatous uterus	0	0	NA	10	8.5	45.2
Risk reducing BSO	0	0	NA	20	17.1	46.8
Gestational residua	0	0	NA	10	8.5	30.8
Normal Endometrium	2	6.8	58	6	5.1	50.8
Other	0	0	NA	8	6.8	36.9

Blood-based testing for most biomarkers has limited efficacy because of their association with tumor burden, which results in late diagnosis at the metastatic stage. In contrast, intraluminal body fluids are expected to contain the putative biomarkers at an earlier disease stage. High grade serous papillary carcinoma, the most common histological subtype of HGOC, arises from precursor lesions that develop in the epithelium of the fallopian tube fimbriae (FTE, the distal end of the fallopian tube, adjacent to the ovaries) (15–17). Therefore, sampling the cells of the fimbriae or their secreted biological products (via liquid aspirated from the gynecological tract) may reveal markers of the initial lesions. Several gynecologic liquid biopsy methods were recently described, primarily for examining circulating mutant p53 DNA (18–21). All of these showed very low sensitivity (33–60%). Despite technical and conceptual limitations, proteomics may be superior to genomic assays for the specific context of detection of very small tumors, because it is able to capture the expression perturbation of both tumor cells and their complex microenvironment.

The challenge of proteomics-based biomarker discovery lies in the high dynamic range of most body fluids. High levels of extracellular proteins, primarily plasma proteins, mask the proteins secreted from tumor cells, and therefore hamper biomarker identification. To overcome this challenge, as we have previously published, we performed deep proteomic analysis of plasma microvesicles, which allowed us to reach thousands of protein identification in single runs (22). Microvesicles (100 nm-1 μm) form by outward budding of the plasma membrane and are released into body fluids from all cell types (23, 24). Thus, microvesicles can serve as a reservoir of diagnostic biomarkers, which sets the foundations for the development of an assay that could be used as a screening or monitoring tool (25). Because these are largely devoid of highly abundant plasma proteins, their analysis overcomes masking of the potential protein biomarkers. In the current study, we adapted the plasma microparticle analysis to uterotubal lavage (UtL) samples. We combine the virtues of state-of-the-art MS-based proteomics with minimally invasive sampling method, and extract a proteomic signature, as a first step toward early HGOC diagnostics.

EXPERIMENTAL PROCEDURES

Cohort Design and Assembly—All samples were collected in accordance with approvals of the institutional ethics review boards at Chaim Sheba Medical Center, Rabin Medical Center and Meir Medical Center, Israel (ClinicalTrials.gov identifier: NCT03150121). Informed consent was obtained from each participant. Recruited patients underwent gynecological surgical procedures under general anesthesia, including hysteroscopy, hysterectomy and/or RRBSO. Eligible indications included HGOC (primary or interval debulking), suspicious ovarian mass, risk reduction, or various other benign gynecological disorders (Table I and supplemental Table S1). Patients with endometrial and cervical carcinoma were excluded, as well as patients with non-HG serous ovarian tumors. Sample processing was performed in six technical batches; each included both patients and controls, and all three centers. MS analysis was then performed in a blinded manner for each batch. Altogether, we collected 187 UtL samples, and eleven of those were excluded because of missing information. Of the remaining 176 samples, 24 were selected to the discovery set (12 HGOC patients and 12 representative controls),

[1] The abbreviations used are: HGOC, high-grade ovarian cancer; RRBSO, risk-reducing bilateral salpingo-oophorectomy; UtL, uterotubal lavage; RFE, recursive feature elimination.

whereas subsequent samples were regarded as a validation set (n = 152) and analyzed independently in a blinded manner. Additional information can be found in Fig. 1*A*, Table I, and supplemental Table S1.

UtL Collection—UtL samples were collected before surgery, after induction of anesthesia, by surgeons in the participating centers. An intrauterine insemination catheter (Insemi^{TM}-Cath, Cook Inc. Bloomington, IN) or rigid pipelle uterine sampler (Endosampler, MedGyn, Addison, IL) was inserted into the endometrial space through the cervical canal. Ten ml of saline were flushed into the uterine cavity and fallopian tubes and immediately retrieved (at an average volume of 4.6 ml per patient).

Isolation of UtL Microvesicles—Microvesicle isolation was performed as previously described (22). Briefly, the UtL samples were immediately centrifuged at 480 × *g* for 15 min to eliminate cells, and supernatants were stored at −80 °C. Subsequently, processing and analysis was performed in several batches as follows: 1 ml UtL samples were centrifuged at 1000 × *g* for 20 min to remove cell debris, followed by microvesicle precipitation by centrifugation at 20,000 × *g* for 60 min at 4 °C. Pellets were then washed with 1 ml ice-cold PBS and centrifuged again at 20,000 × *g* for 60 min at 4 °C.

In-solution Digestion and LC-MS/MS Analysis—Microvesicle pellets were solubilized in 8 M urea in 100 mM Tris-HCl (pH 8.5), reduced with 1 mM dithiothreitol (DTT) at RT for 30 min and alkylated with 5 mM iodoacetamide (IAA) for 30min in the dark. The lysates were diluted 4-fold with 50 mM ammonium bicarbonate, followed by overnight digestion with Trypsin/Lys-C mix (MS grade Promega, Madison, WI; 1:100 enzyme to protein ratio) and sequencing grade modified trypsin (Promega, 1:50 enzyme to protein ratio). Resulting peptides were acidified with trifluoroacetic acid (TFA), purified on C_{18} StageTips (3 M Empore™, St. Paul, MN) and vacuum dried (26). The dried peptides were resuspended in 2% acetonitrile/0.1% TFA before the LC-MS/MS analysis.

Peptides were analyzed by liquid-chromatography using the EASY-nLC1000 HPLC (Thermo Fisher Scientific) coupled to the Q-Exactive (QE) Plus or Q-Exactive HF mass spectrometers (Thermo Fisher Scientific, Bremen, Germany). Peptides were separated on 75 μm i.d. × 50 cm long EASY-spray PepMap columns (Thermo Fisher Scientific) packed with 2 μm, C_{18} material with 100Å pore size. The peptides were loaded with Buffer A (0.1% formic acid) and eluted with a gradient of 7–28% Buffer B (80% acetonitrile/0.1% formic acid), at a flow rate of 300 nl/min, over a gradient of 210 min. MS acquisition was performed in a data-dependent manner, positive-ion mode with selection of the top 10 peptides from each MS spectrum for fragmentation and MS/MS analysis. Full MS spectra were acquired at a resolution of 70,000 (QE-Plus) or 60,000 (QE-HF), *m/z* range of 300–1800 Th, with AGC target of 3E+06 ions and maximal injection time of 20 ms (QE-Plus) or 100 ms (QE-HF). Peptides were isolated for fragmentation with an isolation window of 1.6 *m/z*. Higher-energy collisional dissociation (HCD) fragmentation was performed with normalized collisional energy of (NCE) 25 (QE-Plus) or 27 (QE-HF). MS/MS spectra were acquired at a resolution of 17,500 (QE-Plus) or 30,000 (QE-HF), with AGC target of 1E+05 (QE-Plus) or 5E+04 (QE-HF) and maximal injection time of 60 ms (QE-Plus) or 50ms (QE-HF). Dynamic exclusion was set to 30 s. Raw files of all the samples are available via ProteomeXchange with identifier PXD009655. Spectra for single-peptide based protein identification are added in the supplementary material.

Experimental Design and Statistical Rationale for Proteomic analysis—Raw MS files were analyzed in the MaxQuant software (version 1.5.2.18) and the Andromeda search engine (27, 28). Separate analyses were performed for the discovery cohort (n = 24) and the validation cohort (n = 152) using the same parameters. MS/MS spectra were searched against the Uniprot database (version Apr2014 with 92,179 entries), a decoy, reverse database of the same size, and a list of common contaminants (245 entries). The peptide search included carbamidomethyl-cysteine as a fixed modification, and N-terminal acetylation and methionine oxidation as variable modifications. MaxQuant search parameters for the initial mass recalibration of the precursors were 20 ppm, and in the main search, the mass tolerance for precursor and fragment ions was 4.5 and 20 ppm, respectively. Trypsin was the specified protease and the maximal number of missed cleavages allowed was two. The minimal peptide length was set to seven amino acids and a minimum of one razor peptide per protein. The search results were filtered with false discovery rate of 0.01 for peptide-spectrum matches and 0.01 for protein identifications. The label-free quantification algorithm (LFQ) in MaxQuant was used for relative quantification, and the "match between runs" feature was enabled. All proteins that could not be discriminated based on the identified peptides were merged into a single protein group.

All statistical analyses were performed with the Perseus software (1.5.1.16) (29). Initially, we filtered out proteins identified in the decoy database, proteins identified only based on the variable modification site, potential contaminants and immunoglobulins. The "protein groups" and the "peptide" output tables from the separate MaxQuant analyses for the discovery and validation cohort are available as supplemental Tables S2–S5. Bioinformatic analysis of the discovery cohort was performed on the log2-LFQ- intensities. Data were filtered to include proteins with valid values in at least 75% of the samples. Missing values were then imputed by replacing them with random, low intensity values that form a normal distribution with a width of 30%, and downshift of 1.8 standard deviations of the general data distribution. The imputed LFQ intensities of the discovery cohort are provided in supplemental Table S6. Machine learning was performed on the imputed LFQ intensities of the discovery cohort. Support vector machines (SVM) algorithm using linear kernel function was employed to extract a predictive signature that can discriminate between the control and ovarian cancer patients. We combined three feature selection algorithms: recursive feature elimination (RFE)-SVM, SVM and ANOVA (30). In each of these processes, cross validation was performed on the discovery set with 250 iterations of random sampling of 85% of the samples as test and 15% as validation. The optimal number of overlapping features of these three analytic methods was calculated to provide highest predictive accuracy with the lowest possible error rates in the discovery set. Filtration and imputation of the validation set was performed in the same manner as for the discovery set. The performance of the extracted classifier was then blindly examined on the (log2) LFQ intensities of the 9-signature proteins in the validation cohort (supplemental Table S7). ROC curve and the AUC calculations were performed in MATLAB. Hierarchical clustering was performed on the z-scored log2 intensities using Euclidean distances between averages. The FDR of the signature proteins was estimated by permuting the sample labels of the discovery cohort 100 times, followed by the same SVM classification and feature selection procedures. FDR calculation was based on the number of times the protein reached the top 15 ranks out of the 100 permutations.

RNA Extraction and RT-PCR—Fresh-frozen HGOC tumors and fresh grossly benign FT fimbriae were obtained from the Chaim Sheba Institutional Tumor Bank. H&E staining was performed to ensure >80% tumor cellularity. The fimbriae were processed as previously described (31, 32). Total RNA was extracted from primary fresh frozen HGOC tumors and dissociated normal FTE cells using QIAzol reagent (Qiagen, Valencia, CA) followed by RNeasy clean-up kit (Qiagen) according to manufacturer's protocol. Gene expression was assessed using FastStart Universal SYBR Green Master (ROX) (Sigma-Aldrich, St. Louis, MO). Primers for the signature-genes are listed in supplemental Table S8 (Sigma-Aldrich).

Immunohistochemistry—Archival tissues were retrieved from the Department of Pathology at the Chaim Sheba Medical Center with the appropriate ethical review board approvals. We constructed TMAs of 30 representative cases (in duplicates) of morphologically benign fimbriae of patients that were removed and submitted to pathological evaluation as part of surgery for the following diagnoses: (1) normal FT adjacent to HGOC (median age = 60, range: 40–74), (2) tubal ectopic pregnancy (EP, median age = 33, range: 20–45), (3) leiomyomatous uterus (LM, benign condition of the uterus which does not involve the fallopian tube epithelium, median age = 52, range: 38–67), and (4) Risk reducing bilateral salpingo-oophorectomy (RRBSO, prophylactic removal of the ovaries and fallopian tubes because of germ-line *BRCA* mutation, median age = 43, range: 35–66). TMA of 46 HGOC tumors (median age = 62, range 30–88) was also constructed. All slides were simultaneously stained and scored for staining intensity and distribution, on a scale of 0–3 (0 - no staining or faint staining in <10% of cells, 1 - faint staining in >10% of cells, 2 - moderate staining of >10% of cells, and 3 - strong staining of >10% of cells).

Primary antibodies used: (1) anti-SERPINB5 (HPA020136, 1:200, positive control: keratinocytes) and (2) anti-S100A14 (HPA027613, 1:1000, positive control: keratinocytes) (Sigma-Aldrich).

Statistical Analysis—Statistical significance ($p < 0.05$) was assessed by Student *t* test for RT-PCR data or by Fisher exact test for IHC intensity scores. Binomial model analysis was used to evaluate the correlation of actual diagnosis, age and menopausal status confounders with the prediction. Pearson correlation test was used to examine the correlation of individual protein expression with age, and Chi square test was used to test the correlation of individual protein expression with menopause.

RESULTS

Our approach to identify early-stage biomarkers with high sensitivity and specificity combines liquid biopsies from the lumen of the gynecologic tract, with deep microvesicles proteomics of the samples. To profile the proteome of the complex utero-tubal body fluid and extract diagnostic biomarkers, we adapted the previously described method for microparticle proteomics, which overcomes masking by highly abundant proteins, and followed with high-resolution MS analysis (22). Briefly, we isolated microvesicles from 1 ml of UtL liquid biopsy samples by high-speed centrifugation, and followed by urea-based denaturation and in-solution digestion (Fig. 1*B*). Peptides were analyzed on the Q-Exactive Plus or Q-Exactive HF MS, and proteins were quantified using the label-free algorithm in MaxQuant. Patient cohorts included samples from HGOC patients and controls (with nonmalignant gynecological conditions) from three medical centers.

Initial data analysis included all samples, in a combined MaxQuant analysis, to evaluate the data quality, and examine whether there are any technical artifacts associated with the sample origin and batch. Combined analysis identified a total of 8760 proteins. Among them, we found known lineage markers of FTE/HGOC, such as MUC16 (CA125), WFDC2 (HE4), and OVGP1 (MUC9), as well as very low-abundance proteins, including cytokines and growth factors, such as IGF1, CXCL12, IL18, and HGF (Fig. 1*C*). The dynamic range of relative abundance of the microvesicle proteome spanned eight orders of magnitude. In agreement with our previous results (31), the amounts of lineage markers CA125 (MUC16), HE4 (WFDC2), and OVGP1 (MUC9) did not discriminate between HGOC patients and control samples (Fig. 1*D*). Moreover, the concentration of CA125 in unfractionated UtL samples was measured with a commercial assay (Access Immunoassay OV Monitor, Beckman Coulter), and demonstrated no significant difference between patients and controls (data not shown). Next, we examined potential "batch effect" or differences in composition of samples (surrogate for UtL sampling technique variations and analysis batches). Principal component analysis (PCA) showed no clear separation between the groups of samples, implying low technical variation between the batches and between the three medical centers (supplemental Fig. S1*A* and S1*B*). Additionally, correlation analysis between samples showed an average correlation of 0.67 within each center and correlation of 0.66 between centers. Reassuringly, we found higher correlations between controls from different centers, than between patients and controls from the same center (supplemental Fig. S1*C*). We therefore concluded that the batch effects and inter-institutional differences are negligible and did not require any correction. We then investigated whether we can identify significantly different proteins between patients and controls.

Diagnostic UtL-Based Proteomic Classifier—Aiming to identify diagnostic markers, we divided the data into a discovery set and a validation set. To eliminate any dependence between the discovery and the validation cohorts, we analyzed each of these sets separately in MaxQuant. A discovery cohort, including a total of 24 patients and controls, was used to construct a protein classifier for HGOC diagnosis. It was designed to include patients from all three medical centers, exclude any previously treated patients and *BRCA* carriers, and have equal numbers of cases and controls. MaxQuant analysis of these samples identified a total of 5565 UtL microvesicle proteins, and an average number of ~2500 proteins per sample (range: 1100–3600; supplemental Fig. S2*A*–S2*B*, supplemental Table S2). To obtain a signature of minimal number of proteins with highest accuracy and robustness, we tested three feature selection algorithms: Support vector machine (SVM), recursive feature elimination (RFE)-SVM and ANOVA. The entire analytical workflow was embedded in a cross validation procedure to reduce over-fitting. To minimize the dependence on the feature selection algorithm, we tested the performance of several sets of top-ranked overlapping signatures, ranging in size from 6 to 19 features (Fig. 2*A* and 2*B*). Optimal sensitivity, specificity, and area under the curve (AUC) of Receiver Operating Characteristic (ROC) curve of sensitivity *versus* 1-specificity were obtained with a 9-protein classifier, 6 of which were higher in the HGOC patients, and 3 that were higher in controls (Fig. 2*C*, supplemental Fig. S2*C*, and Table II). *t* test showed that five of the signature proteins (S100A2, S100A14, SERPINB5, IVL, and CLCA4) were also statistically significant (FDR 0.05; s0 = 0.5) between the control and patient samples in the discovery cohort (Fig. 2*D*). This

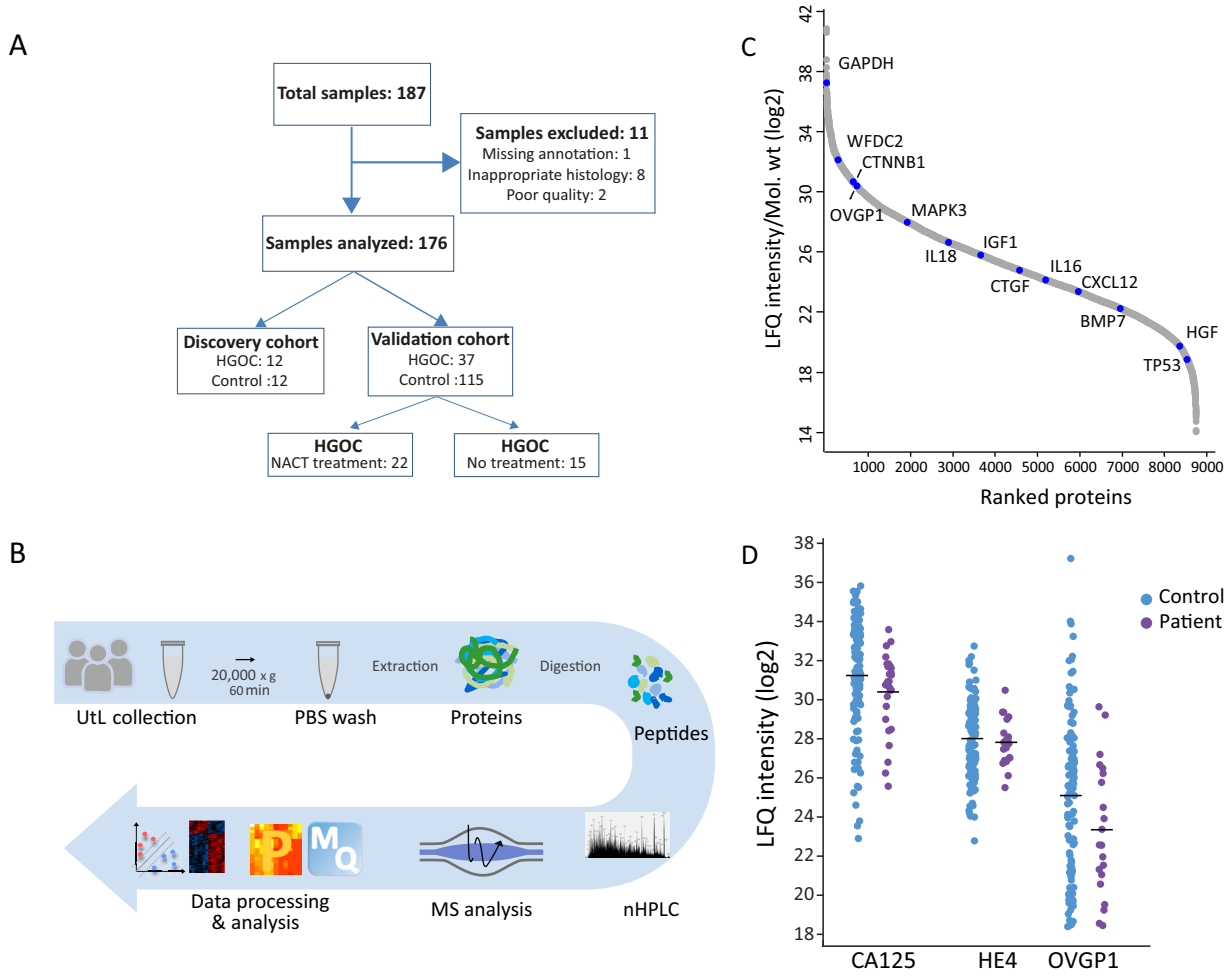


FIG. 1. **UtL microvesicle proteomics.** *A*, Schematic representation of the cohort design. *B*, Workflow from UtL fluid collection, through microvesicle isolation, peptide purification to MS analysis. *C*, Dynamic range of proteins in UtL microvesicles ranging from high abundant known ovarian markers to low abundant cytokines and growth factors. *D*, LFQ intensities of known markers CA125 (MUC16), HE4 (WFDC2) and OVGP1 (MUC9) in log2 normalized intensities.

signature demonstrated 83% sensitivity (95% confidence interval: 51.6–98%), at a specificity of 100% (95% confidence interval: 73.5–100%), and an AUC of 0.99 in the discovery set (Fig. 2*E*). The permutation-based FDR for the signature proteins ranged between 0–0.11 (Supplemental Table S9). The coefficients of variation of the nine signature proteins were below 25% (supplemental Fig. S2*D*). Importantly, this signature correctly predicted all three stage IA HGOC cases included in the discovery set. Intensities of seven of the nine proteins discriminated them from control samples better than they discriminated advanced stage HGOC samples from controls (supplemental Fig. S3*A*), suggesting the potential strength of this signature in identification of early-stage lesions. Last, to control for any influence of confounding factors, we included the age, batch and medical center in the prediction process. Repeating the same machine learning procedures identified age as a predictor of disease, but the other potential confounding factors were not included in the top ranks. Because the HGOC group was on average, significantly older than the control group (61.8 versus 50.5 years old) this result is expected. Moreover, adding these potential confounding variables to the signature resulted in reduced performance of the classifier, thus showing that these do not contribute the signature performance (supplemental Fig. S3*B*).

We next sought to validate the performance of the proteomic signature on the independent validation cohort of patient/control UtL samples (n = 152, Table I). This sample set was analyzed independently in MaxQuant, which led to identification of a total of 8544 proteins, and an average of 3200 per sample. Application of the 9-protein classifier to the validation cohort predicted 84 of the controls and 25 patients correctly, providing 74% sensitivity at 66% specificity (Sensitivity- 95% confidence interval: 48.6–80.3%; Specificity- 95% confidence interval: 64.6–81.5%). ROC curve for the validation cohort showed an AUC of 0.71 (Fig. 3*A*). Of note, one case of an incidental occult precursor lesion of serous tubal intraepithelial carcinoma (STIC) was incorrectly desig-

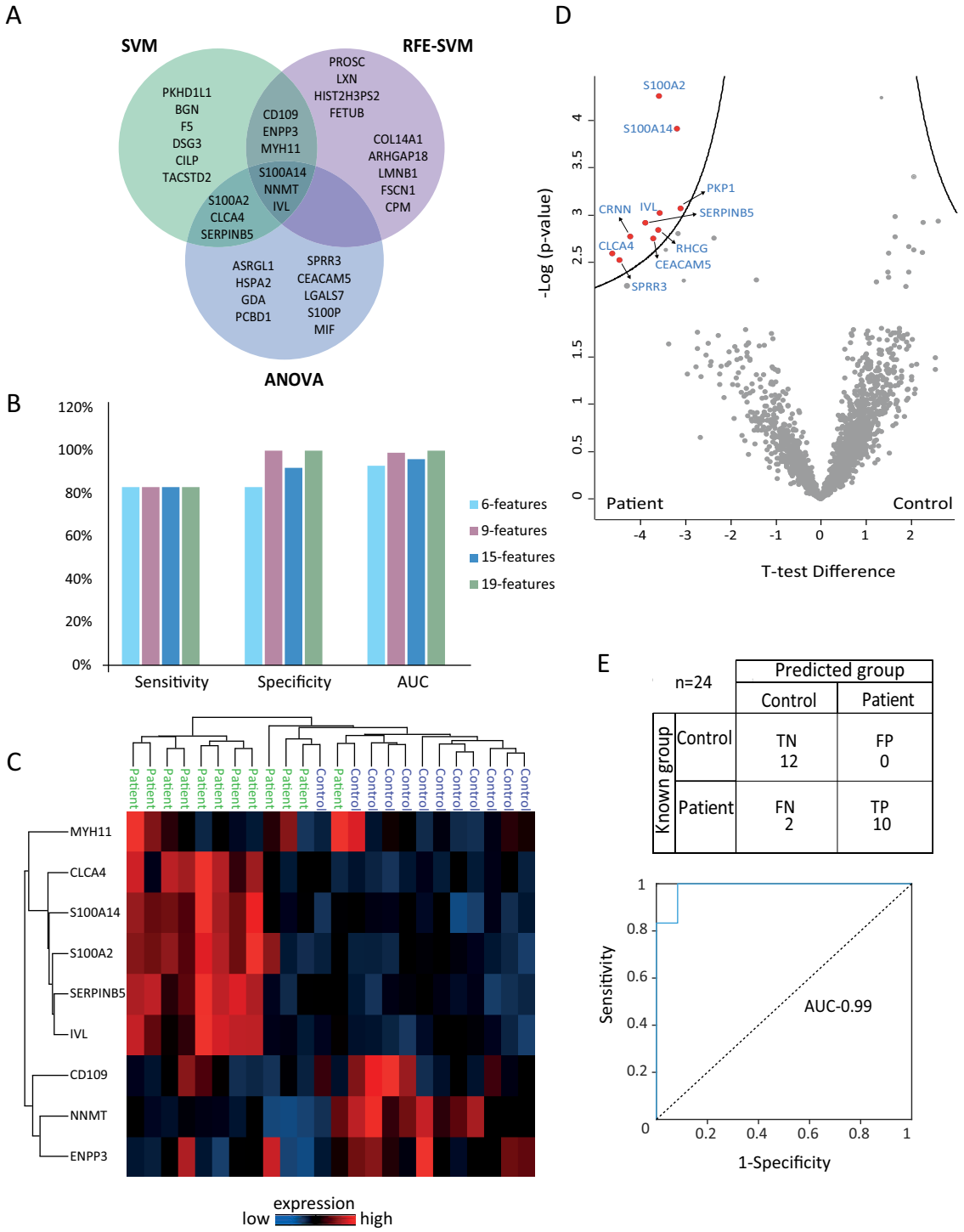


FIG. 2. **Development of a proteomic classifier for diagnosis of HGOC.** *A*, Venn diagram showing the selected nine overlapping features in the top 15 ranks from three different feature selection methods. *B*, Comparison of sensitivity, specificity and AUC for the top ranked 6, 9, 15, and 19 overlapping features from different feature ranking methods. *C*, Heatmap shows the expression of nine signature-proteins across the discovery set of UtL samples. *D*, Volcano plot shows the differentially expressed proteins between the control and patient samples with FDR 0.05 and s0 = 0.5. Statistically significant proteins are highlighted in red. *E*, Confusion matrix and ROC curve show the performance of the 9-protein classifier.

nated as "healthy" by the 9-protein classifier. However, because of the small number of early stage patients, the performance in these cases requires further investigation. Our validation set included 22 UtL samples from HGOC patients who received neo-adjuvant chemotherapy (NACT). PCA shows that the NACT treated samples were highly like the

TABLE II
The overlapping features which compose the 9-protein classifier

#	Gene names	Protein names	UNIPROT ID	SVM rank	RFE-SVM rank	ANOVA rank
1	MYH11	Myosin-11	P35749	1	10	478
2	CLCA4	Calcium-activated chloride channel regulator 4	Q14CN2	3	19	10
3	S100A14	Protein S100-A14	Q9HCY8	14	14	3
4	S100A2	Protein S100-A2	P29034	3	17	2
5	SERPINB5	Serpin B5	P36952	11	21	4
6	IVL	Involucrin	P07476	2	15	5
7	CD109	CD109 antigen	Q6YHK3	8	8	168
8	NNMT	Nicotinamide N-methyltransferase	P40261	9	2	8
9	ENPP3	Ectonucleotide pyrophosphatase/phosphodiesterase family member 3	O14638	6	13	150

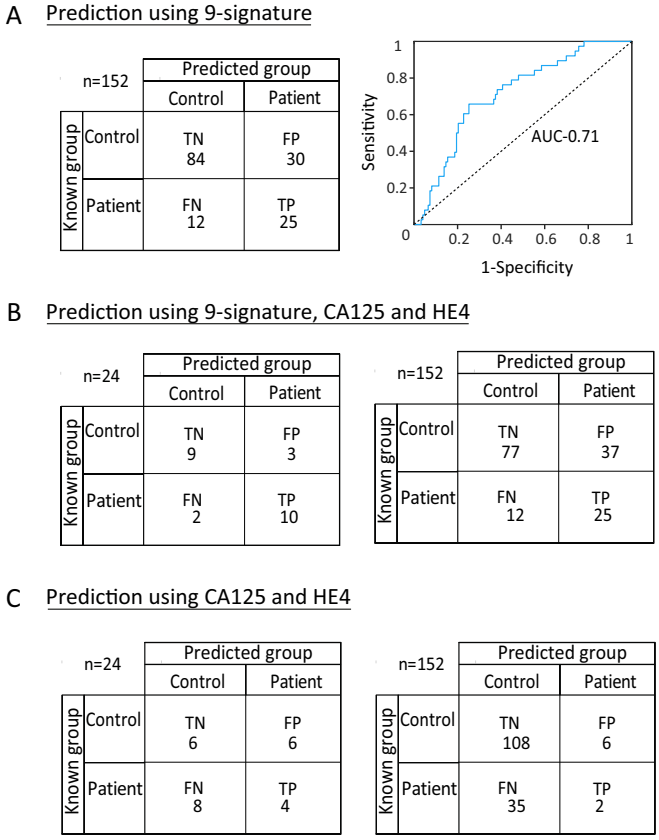


FIG. 3. **Performance of the proteomic signature on a validation set.** *A*, Confusion matrix and ROC curve of the independent validation cohort, with AUC = 0.71. *B*, Confusion matrix shows the performance of the 9-classifier along with the known markers (CA125 & HE4) in the discovery and validation cohorts. *C*, Performance of only the known markers in the discovery and validation cohorts.

samples obtained from HGOC patients during primary debulking (supplemental Fig. S4). Looking at the integration of the two cohorts, the classifier offered 70% sensitivity and 76.2% specificity for diagnosis of HGOC.

Given the long-standing clinical use of CA125 and HE4 as diagnostic markers, we examined whether their combination with our signature has any predictive advantage. The performance of a combination of the 9-classifier and the best-validated protein biomarkers (CA125 and HE4) was calculated in the discovery and validation cohorts. Both sensitivity and specificity were reduced compared with the 9-protein signature alone: sensitivity of 83% and 68% and specificity of 75% and 67.5% in the discovery and validation cohort, respectively (Fig. 3*B*). The performance of CA125+HE4 alone was even worse, with sensitivity of 33% and 5.2% and specificity of 50% and 94.7% in the discovery and validation cohorts, respectively (Fig. 3*C*).

Because the patient age correlated with the prediction, and most HGOC patients were post-menopausal, we tested whether age and menopausal status affect the signature protein expression. Because hormonal status information was not available for all patients, we divided the cohort into age $< = 50$ (pre-menopausal) *versus* age > 50 (post-menopausal). Binomial model multivariate analysis demonstrated no correlation of the signature with age (p value = 0.414). *p value* for regression correlation of 1.45 with menopausal status was 0.01, because diagnosis of HGOC directly and strongly correlates with menopause (p value = 2.5E-06). Reassuringly, the actual diagnosis strongly correlated with the signature prediction (p value = 3.9E-07). Moreover, the LFQ intensities of the individual signature proteins did not significantly correlate with menopausal status. Only one protein, Ectonucleotide Pyrophosphatase/Phosphodiesterase 3 (ENPP3), inversely correlated with age (p value = 0.00078; supplemental Table S9).

Real-time PCR Validation of Differential Expression of Signature Proteins—The UtL liquid biopsy samples proteins that are not necessarily exclusively expressed by the cancer cells, but can also capture stromal response to tumor development, or can result from an increase in specific tissue mass. Some known tumor markers (*e.g.* CA125) directly reflect an increase in mass of a specific tissue type, and are not uniquely expressed by malignant cells, nor do they possess cancer-promoting biological functions. Such markers are expected to detect tumors at an advanced stage, and may not be appropriate for early cancer diagnosis, whereas cancer-specific expression may increase the sensitivity of signature biomarker and increase the chances of diagnosing the disease at an early stage. We therefore examined the expression pat-

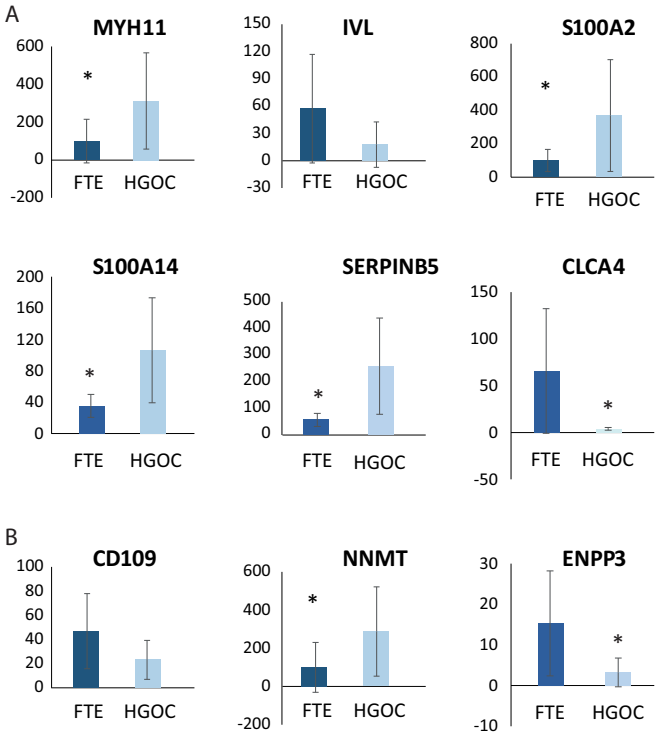


FIG. 4. **RT-PCR of the signature genes in normal FTE and HGOC.** mRNA levels that correspond with the 6 signature proteins that are higher in UtL of HGOC patients (*A*) and the 3 signature proteins that are lower in UtL of HGOC patients (*B*). RNA was extracted from fresh independent normal FTE ($n = 10$) and unmatched HGOC ($n = 10$) specimens, were measured using RT-PCR. Statistical significance of DE marked by * for p value < 0.05.

terns of the signature proteins at the RNA level, by RT-PCR, and the protein level, by IHC. We measured the mRNA expression of all signature genes in HGOC tumors *versus* normal FTE on an independent set of unmatched samples: fresh-frozen advanced HGOC tumors ($n = 10$) and unmatched benign FTE cells harvested from normal fimbriae ($n = 10$). Our results indicate statistically significant transcriptomic differential expression (DE) in accordance with the proteomic analysis of five of the nine genes (Fig. 4). The fact that not all transcripts are DE suggests that some proteins remain relatively consistent through malignant transformation and may also stem from the profound differences in the type of biological materials examined (extracellular microvesicle proteins *versus* cellular mRNA), and the methodologies used (MS *versus* RT-PCR).

Immunohistochemistry (IHC) Validation of Tumor Expression of Signature Proteins—MS and RT-PCR methods lack spatial resolution, thus precluding disclosure of the specific cell-type that expresses each of the classifier's proteins. To explore the localization of selected signature proteins in HGOC tumors and normal FTE, and confirm the DE by tumor cells, we performed IHC for SERPINB5 and S100A14, two selected proteins that were over-represented in UtL samples of HGOC patients. IHC was performed on a tissue microarray (TMA) of HGOC tumors *versus* four control-TMAs representing grossly normal FT fimbriae removed from women with: HGOC, tubal ectopic pregnancy (EP), leiomyomatous uterus (LM, benign condition of the uterus not affecting the fallopian tube), or *BRCA*-mutation carriers undergoing RRBSO.

SERPINB5 is an epithelial-cell-specific member of the SERPIN family that lacks serine protease inhibition activity. Not much is known about its cellular functions in cancer, yet it has been implicated as cancer susceptibility gene and a prognostic factor in several cancer types (33). It has been also attributed a role as an exosomal protein (34). In accordance with the proteomic analysis, IHC exhibits weak cytoplasmic staining in less than 50% of normal FTE specimens (intensity 0–1), and a stronger expression in a subset of HGOC tumors (p value = 1.65E-09; Fig. 5*A*, supplemental Fig. S5).

S100A14 is a member of the S100 family lacking calcium-binding function, known to be involved in the regulation of TP53 protein expression and of cellular motility (35). In FTE, it localized exclusively to the cytoplasm of ciliated cells, with very low staining in secretory cells (intensity 0–1) (Fig. 5*B*, supplemental Fig. S6). In agreement with the proteomic analysis, its expression was significantly higher in HGOC tumor cells compared with the presumed cell-of-origin - secretory FTE (p value = 2.04E-06; Fig. 5*B*) (15).

We further obtained IHC evidence from the Human Protein Atlas database (www.proteinatlas.org (36)) for the expression of three additional proteins. According to publicly available histology images in the database, CLCA4, S100A2 and MYH11 had stronger cytoplasmic staining in HGOC tumor cells than in normal FTE. Overall, the IHC results confirm the DE of the five signature proteins in HGOC tumors compared with normal FTE and localize their expression specifically to tumor cells.

DISCUSSION

In this work, we present the discovery of potential early diagnostic markers for HGOC, using microvesicle proteomics of UtL liquid biopsies. Isolation of microvesicles enabled overcoming the large dynamic range of this body fluid, and untargeted identification of thousands of proteins per sample in single LC-MS runs. As opposed to our original methodological study of plasma microparticle proteomics (22), in the current work we used LFQ rather than SILAC, because we did not find a suitable SILAC standard for this body fluid. The MaxQuant LFQ algorithm enabled general normalization that overcame all potential batch effects; however, even in the separate analysis of the discovery and validation sets, predictive ability of the signature was still high. We believe that this work is the first step toward translation of the signature proteins into a clinical test and envision that such a test will use simpler MS-based targeted assays and shorter analytical times. MS-based clinical tests are expected to increase the accuracy and multiplexing capabilities compared with more commonly used antibody-based tests (*e.g.* ELISA). Further,

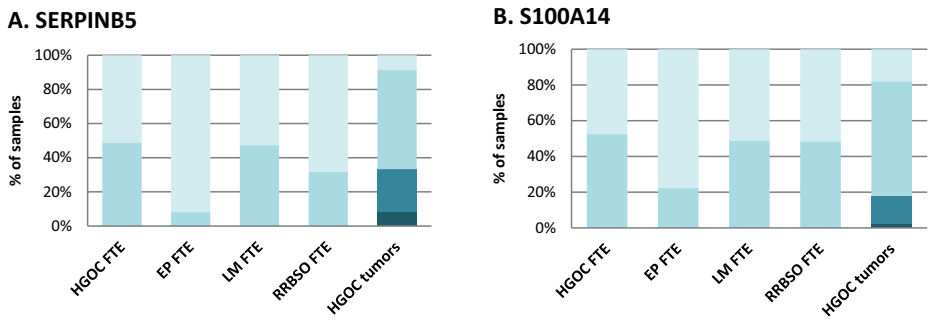


FIG. 5. **Intensity of IHC staining of SERPINB5 (*A*) and S100A14 (*B*) in HGOC tumors and normal FTE.** TMAs of HGOC tumors ($n = 45$), and of cores of normal fimbriae from salpingectomy specimens because of the following indications: HGOC, tubal ectopic pregnancy (EP), leiomyomatous uterus (LM), RRBSO ($n = 60$ each), were immunostained for two signature proteins: SERPINB5 (*A*) and S100A14 (*B*), scored on a 0–3 intensity scale and analyzed. Score scale was as follows: 0 - no staining or faint staining in <10% of cells, 1 - faint staining in >10% of cells, 2 - moderate staining of >10% of cells, and 3 - strong staining of >10% of cells.

these will reduce the cost and assay development times because of the high specificity of the MS results (37). Recent attempts to advance the applicability of the MS-based assays have already simplified the sample preparation and MS analyses, increased the throughput and implemented targeted MS methodologies, combined with absolute quantification (38–41). Future combination of our study with such technologies can potentially lead to implementation of the signature proteins to routine clinical labs.

Early diagnosis of HGOC is of highest importance to women with genetic predisposition, because they are currently counseled to undergo RRBSO around the age of 40, despite the incomplete penetrance and the highly variable age of presentation. This practice gains legitimacy from the exceedingly narrow window-of-opportunity for early-stage diagnosis and the unbearably high mortality rates, thus necessitating extremely cautious management. Recently, evolutionary mutation analyses revealed that the time gap between development of STIC and clinical appearance invasive HGOC is longer than 6 years (17), thus implying that early detection may, after all, be possible once new methodologies become available. UtL liquid biopsy, as opposed to blood, may potentially disclose localized HGOC lesions, which are curable.

Our 9-protein classifier has 70% sensitivity and 76% specificity which outperforms previous results of genomic biomarkers based on gynecological liquid biopsy (18–20). Unlike mutation analysis in UtL samples which looks at a negligible percent of cancer cells, proteomics reflects the complexity of a cancer-associated program that, theoretically, captures expression changes in multiple cell types within the tumor microenvironment, thus can potentially provide a wider array of early-detection biomarkers. Further improvement of the proteomic signature and its predictive power requires analysis of more early-stage HGOC UtL samples or STICs, however, these samples are inherently rare. Coupled with the intrauterine liquid biopsy method, this assay holds promise for clinically significant early detection of HGOC.

The UtL sampling technique that we propose hereby is a simplified version of the originally reported method (18), making it suitable for routine testing of healthy young women at high risk for HGOC, including women who have not undergone vaginal delivery. Fundamental parameters for clinical feasibility, such as patient-reported outcomes, physicians-reported workload and compliance of the target population to undergo routine UtL sampling need to be investigated. Semi-annual monitoring with clinical proteomic assays may be implemented as a measure of reassurance for high-risk populations willing to delay RRBSO until after menopause, and thus become practice changing.

To consolidate the specificity of the signature proteins to HGOC tissues, we examined their expression in independent tissue specimens, comparing FTE and HGOC, using complementary techniques: RT-PCR and IHC. We obtained confirmatory IHC results for five proteins and supportive RT-PCR results for five of the nine genes tested, highlighting the aberrant expression of these proteins in HGOC tissues. These results reinforce the potential of the proteomic signature as a diagnostic test. Of note, discordance between the proteomic predictions and transcriptomic validation results may arise from the differences between mRNA and protein expression patterns, and between the extracellular vesicles and intracellular levels of expression. Alternatively, it is possible that the expression of several proteins is not altered when FTE evolves into HGOC, because they are not directly involved in the cancerous process, but may still be useful biomarkers, like CA125, for aberrant expansion of the cell lineage within HGOC lesions.

Ultimately, the genomic and the proteomic approaches, as well as other possible methodologies of liquid biopsy analysis, may be integrated to yield a multi-modality classifier with an adequate sensitivity and specificity to guarantee early detection of HGOC in both average- and high-risk populations, and potentially enable personalized risk stratification and delay of RRBSO in predisposed women without increasing HGOC incidence.

Acknowledgments—We thank the staff of the Institutional Tumor Bank at Sheba Medical Center for supplying tissue samples and the

staff of the Institute of Pathology for technical assistance. We acknowledge Dr. Michal Barak for assistance with statistical analysis. We thank Dana Silverbush and Jan Daniel Rudolph for statistical advice.

DATA AVAILABILITY

Raw mass spectrometry data of all the samples are available via ProteomeXchange with identifier PXD009655. Spectra for single-peptide based protein identification are available in the supplementary material.

* This research was funded by the Israel Cancer Research Fund (ICRF) - Len & Susan Mark Initiative for Ovarian and Uterine/MMMT Cancers Grant and project grant; Israel Cancer Association Research Grants; Israel Science Foundation (ISF) personal grant 1104/17; Israeli Ministry of Finance Nofar program and I-CORE Centers of Excellence in Gene Regulation in Complex Human Disease, Grant No. 41/11.

S This article contains supplemental material. We declare that they have no conflict of interest.

§§§§ To whom correspondence may be addressed. E-mail: Keren.Levanon@sheba.health.gov.il.

¶¶¶¶ To whom correspondence may be addressed. E-mail: geiger@tauex.tau.ac.il.

‡‡‡‡ Current address: Gynecologic Oncology Division, McMaster University, Juravinski Cancer Center Hamilton, ON, Canada.

|||||| These authors had equal contribution.

Author contributions: G.D.B., K.B.-S., S.S., H.B., O.Z., S.A.-R., and M.H. performed research; G.D.B., K.B.-S., T.G., and K.L. analyzed data; T.G. and K.L. designed research; T.G. and K.L. wrote the paper; L.H., Y.K., M.B., O.W., N.A., J.K., T.P., G.K., A.B., B.B., E.F., D.S., A.J.-S., R.E., and S.A. patient recruitment and sample acquisition.

REFERENCES

1. Vaughan, S., Coward, J. I., Bast, R. C., Berchuck, A., Berek, J. S., Brenton, J. D., Coukos, G., Crum, C. C., Drapkin, R., Etemadmoghadam, D., Friedlander, M., Gabra, H., Kaye, S. B., Lord, C. J., Lengyel, E., Levine, D. A., McNeish, I. A., Menon, U., Mills, G. B., Nephew, K. P., Oza, A. M., Sood, A. K., Stronach, E. A., Walczak, H., Bowtell, D. D., and Balkwill, F. R. (2011) Rethinking ovarian cancer: recommendations for improving outcomes. *Nat. Rev. Cancer* **11,** 719–725
2. Harmsen, M. G., Arts-de Jong, M., Hoogerbrugge, N., Maas, A. H. E. M., Prins, J. B., Bulten, J., Teerenstra, S., Adang, E. M. M., Piek, J. M. J., van Doorn, H. C., van Beurden, M., Mourits, M. J. E., Zweemer, R. P., Gaarenstroom, K. N., Slangen, B. F. M., Vos, M. C., van Lonkhuijzen, L. R. C. W., Massuger, L. F. A. G., Hermens, R. P. M. G., and de Hullu, J. A. (2015) Early salpingectomy (TUbectomy) with delayed oophorectomy to improve quality of life as alternative for risk-reducing salpingo-oophorectomy in BRCA1/2 mutation carriers (TUBA study): a prospective non-randomised multicentre study. *BMC Cancer* **15,** 593–601
3. Arts-de Jong, M., Harmsen, M. G., Hoogerbrugge, N., Massuger, L. F., Hermens, R. P., and de Hullu, J. A. (2015) Risk-reducing salpingectomy with delayed oophorectomy in BRCA1/2 mutation carriers: Patients' and professionals' perspectives. *Gynecol. Oncol.* **136,** 305–310
4. Lu, K. H., Skates, S., Hernandez, M. A., Bedi, D., Bevers, T., Leeds, L., Moore, R., Granai, C., Harris, S., Newland, W., Adeyinka, O., Geffen, J., Deavers, M. T., Sun, C. C., Horick, N., Fritsche, H., and Bast, R. C. (2013) A 2-stage ovarian cancer screening strategy using the Risk of Ovarian Cancer Algorithm (ROCA) identifies early-stage incident cancers and demonstrates high positive predictive value. *Cancer* **119,** 3454–3461
5. Moore, R. G., McMeekin, D. S., Brown, A. K., DiSilvestro, P., Miller, M. C., Allard, W. J., Gajewski, W., Kurman, R., Bast, R. C., Jr, and Skates, S. J. (2009) A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* **112,** 40–46
6. Sölétormos, G., Duffy, M. J., Othman Abu Hassan, S., Verheijen, R. H. M., Tholander, B., Bast, R. C., Gaarenstroom, K. N., Sturgeon, C. M., Bonfrer, J. M., Petersen, P. H., Troonen, H., CarloTorre, G., Kanty Kulpa, J., Tuxen, M. K., and Molina, R. (2015) Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines From the European Group on Tumor Markers. *Int. J. Gynecol. Cancer* **26,** 43–51
7. Karlan, B. Y., Thorpe, J., Watabayashi, K., Drescher, C. W., Palomares, M., Daly, M. B., Paley, P., Hillard, P., Andersen, M. R., Anderson, G., Drapkin, R., and Urban, N. (2014) Use of CA125 and HE4 serum markers to predict ovarian cancer in elevated-risk women. *Cancer Epidemiol. Biomarkers Prev.* **23,** 1383–1393
8. Song, E., Gao, Y., Wu, C., Shi, T., Nie, S., Fillmore, T. L., Schepmoes, A. A., Gritsenko, M. A., Qian, W.-J., Smith, R. D., Rodland, K. D., and Liu, T. (2017) Targeted proteomic assays for quantitation of proteins identified by proteogenomic analysis of ovarian cancer. *Sci. Data* **4,** 170091
9. Jacobs, I. J., Menon, U., Ryan, A., Gentry-Maharaj, A., Burnell, M., Kalsi, J. K., Amso, N. N., Apostolidou, S., Benjamin, E., Cruickshank, D., Crump, D. N., Davies, S. K., Dawnay, A., Dobbs, S., Fletcher, G., Ford, J., Godfrey, K., Gunu, R., Habib, M., Hallett, R., Herod, J., Jenkins, H., Karpinskyj, C., Leeson, S., Lewis, S. J., Liston, W. R., Lopes, A., Mould, T., Murdoch, J., Oram, D., Rabideau, D. J., Reynolds, K., Scott, I., Seif, M. W., Sharma, A., Singh, N., Taylor, J., Warburton, F., Widschwendter, M., Williamson, K., Woolas, R., Fallowfield, L., McGuire, A. J., Campbell, S., Parmar, M., and Skates, S. J. (2016) Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* **387,** 945–956
10. Buys, S. S., Partridge, E., Black, A., Johnson, C. C., Lamerato, L., Isaacs, C., Reding, D. J., Greenlee, R. T., Yokochi, L. A., Kessel, B., Crawford, E. D., Church, T. R., Andriole, G. L., Weissfeld, J. L., Fouad, M. N., Chia, D., O'Brien, B., Ragard, L. R., Clapp, J. D., Rathmell, J. M., Riley, T. L., Hartge, P., Pinsky, P. F., Zhu, C. S., Izmirlian, G., Kramer, B. S., Miller, A. B., Xu, J. L., Prorok, P. C., Gohagan, J. K., and Berg, C. D. (2011) Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA* **305,** 2295–2303
11. Rosenthal, A. N., Fraser, L. S. M., Philpott, S., Manchanda, R., Burnell, M., Badman, P., Hadwin, R., Rizzuto, I., Benjamin, E., Singh, N., Evans, D. G., Eccles, D. M., Ryan, A., Liston, R., Dawnay, A., Ford, J., Gunu, R., Mackay, J., Skates, S. J., Menon, U., Jacobs, I. J., and United Kingdom Familial Ovarian Cancer Screening Study collaborators. (2017) Evidence of stage shift in women diagnosed with ovarian cancer during Phase II of the United Kingdom Familial Ovarian Cancer Screening Study. *J. Clin. Oncol.* **35,** 1411–1420
12. Nakamura, K., Sawada, K., Yoshimura, A., Kinose, Y., Nakatsuka, E., and Kimura, T. (2016) Clinical relevance of circulating cell-free microRNAs in ovarian cancer. *Mol. Cancer* **15,** 48
13. Cohen, J. D., Li, L., Wang, Y., Thoburn, C., Afsari, B., Danilova, L., Douville, C., Javed, A. A., Wong, F., Mattox, A., Hruban, R. H., Wolfgang, C. L., Goggins, M. G., Dal Molin, M., Wang, T.-L., Roden, R., Klein, A. P., Ptak, J., Dobbyn, L., Schaefer, J., Silliman, N., Popoli, M., Vogelstein, J. T., Browne, J. D., Schoen, R. E., Brand, R. E., Tie, J., Gibbs, P., Wong, H.-L., Mansfield, A. S., Jen, J., Hanash, S. M., Falconi, M., Allen, P. J., Zhou, S., Bettegowda, C., Diaz, L. A., Tomasetti, C., Kinzler, K. W., Vogelstein, B., Lennon, A. M., and Papadopoulos, N. (2018) Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* **359,** 926–930
14. Elias, K. M., Fendler, W., Stawiski, K., Fiascone, S. J., Vitonis, A. F., Berkowitz, R. S., Frendl, G., Konstantinopoulos, P., Crum, C. P., Kedzierska, M., Cramer, D. W., and Chowdhury, D. (2017) Diagnostic potential for a serum miRNA neural network for detection of ovarian cancer. *Elife* **6,** e28932
15. Perets, R., and Drapkin, R. (2016) It's totally tubular. Riding the new wave of ovarian cancer research. *Cancer Res.* **76,** 10–17
16. Levanon, K., Crum, C., and Drapkin, R. (2008) New insights into the pathogenesis of serous ovarian cancer and its clinical impact. *J. Clin. Oncol.* **26,** 5284–5293
17. Labidi-Galy, S. I., Papp, E., Hallberg, D., Niknafs, N., Adleff, V., Noe, M., Bhattacharya, R., Novak, M., Jones, S., Phallen, J., Hruban, C. A., Hirsch, M. S., Lin, D. I., Schwartz, L., Maire, C. L., Tille, J.-C., Bowden, M., Ayhan, A., Wood, L. D., Scharpf, R. B., Kurman, R., Wang, T.-L., Shih, I.-M., Karchin, R., Drapkin, R., and Velculescu, V. E. (2017) High grade

serous ovarian carcinomas originate in the fallopian tube. *Nat. Commun.* **8,** 1093

18. Maritschnegg, E., Wang, Y., Pecha, N., Horvat, R., Van Nieuwenhuysen, E., Vergote, I., Heitz, F., Sehouli, J., Kinde, I., Diaz, L. A., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., Speiser, P., and Zeillinger, R. (2015) Lavage of the uterine cavity for molecular detection of Müllerian duct carcinomas: A proof-of-concept study. *J. Clin. Oncol.* **33,** 4293–4300
19. Erickson, B. K., Kinde, I., Dobbin, Z. C., Wang, Y., Martin, J. Y., Alvarez, R. D., Conner, M. G., Huh, W. K., Roden, R. B. S., Kinzler, K. W., Papadopoulos, N., Vogelstein, B., Diaz, L. A., and Landen, C. N. (2014) Detection of somatic TP53 mutations in tampons of patients with high-grade serous ovarian cancer. *Obstet. Gynecol.* **124,** 881–885
20. Kinde, I., Bettegowda, C., Wang, Y., Wu, J., Agrawal, N., Shih, I.-M., Kurman, R., Dao, F., Levine, D. A., Giuntoli, R., Roden, R., Eshleman, J. R., Carvalho, J. P., Marie, S. K. N., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., and Diaz, L. A. (2013) Evaluation of DNA from the papanicolaou test to detect ovarian and endometrial cancers. *Sci. Transl. Med.* **5,** 167ra4–167ra4
21. Wang, Y., Li, L., Douville, C., Cohen, J. D., Yen, T.-T., Kinde, I., Sundfelt, K., Kjaer, S. K., Hruban, R. H., Shih, I.-M., Wang, T.-L., Kurman, R. J., Springer, S., Ptak, J., Popoli, M., Schaefer, J., Silliman, N., Dobbyn, L., Tanner, E. J., Angarita, A., Lycke, M., Jochumsen, K., Afsari, B., Danilova, L., Levine, D. A., Jardon, K., Zeng, X., Arseneau, J., Fu, L., Diaz, L. A., Karchin, R., Tomasetti, C., Kinzler, K. W., Vogelstein, B., Fader, A. N., Gilbert, L., and Papadopoulos, N. (2018) Evaluation of liquid from the Papanicolaou test and other liquid biopsies for the detection of endometrial and ovarian cancers. *Sci. Transl. Med.* **10,** eaap8793
22. Harel, M., Oren-Giladi, P., Kaidar-Person, O., Shaked, Y., and Geiger, T. (2015) Proteomics of microparticles with SILAC quantification (PROMIS-Quan): a novel proteomic method for plasma biomarker quantification. *Mol. Cell. Proteomics* **14,** 1127–1136
23. Becker, A., Thakur, B. K., Weiss, J. M., Kim, H. S., Peinado, H., and Lyden, D. (2016) Extracellular vesicles in cancer: cell-to-cell mediators of metastasis. *Cancer Cell* **30,** 836–848
24. Maas, S. L. N., Breakefield, X. O., and Weaver, A. M. (2017) Extracellular vesicles: unique intercellular delivery vehicles. *Trends Cell Biol.* **27,** 172–188
25. Borrebaeck, C. A. K. (2017) Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nat. Rev. Cancer* **17,** 199–204
26. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop And Go Extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670
27. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–1372
28. Cox, J., Neuhauser, N., Michalski, A., Scheltema Ra Olsen, J. V., and Mann, M. (2011) Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10,** 1794–1805
29. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13,** 731–740
30. Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002) Gene selection for cancer classification using support vector machines. *Mach. Learn.* **46,** 389–422
31. Levanon, K., Ng, V., Piao, H. Y., Zhang, Y., Chang, M. C., Roh, M. H., Kindelberger, D. W., Hirsch, M. S., Crum, C. P., Marto, J. A., and Drapkin, R. (2009) Primary ex vivo cultures of human fallopian tube epithelium as a model for serous ovarian carcinogenesis. *Oncogene* **29,** 1103–1113
32. Bahar-Shany, K., Brand, H., Sapoznik, S., Jacob-Hirsch, J., Yung, Y., Korach, J., Perri, T., Cohen, Y., Hourvitz, A., and Levanon, K. (2014) Exposure of fallopian tube epithelium to follicular fluid mimics carcinogenic changes in precursor lesions of serous papillary carcinoma. *Gynecol. Oncol.* **132,** 322–327
33. Bernardo, M. M., Dzinic, S. H., Matta, M. J., Dean, I., Saker, L., and Sheng, S. (2017) The opportunity of precision medicine for breast cancer with context-sensitive tumor suppressor Maspin. *J. Cell. Biochem.* **118,** 1639–1647
34. Dean, I., Dzinic, S. H., Bernardo, M. M., Zou, Y., Kimler, V., Li, X., Kaplun, A., Granneman, J., Mao, G., and Sheng, S. (2017) The secretion and biological function of tumor suppressor maspin as an exosome cargo protein. *Oncotarget* **8,** 8043–8056
35. Chen, H., Yuan, Y., Zhang, C., Luo, A., Ding, F., Ma, J., Yang, S., Tian, Y., Tong, T., Zhan, Q., and Liu, Z. (2012) Involvement of S100A14 protein in cell invasion by affecting expression and function of matrix metalloproteinase (MMP)-2 via p53-dependent transcriptional regulation. *J. Biol. Chem.* **287,** 17109–17119
36. Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Ponten, F. (2015) Tissue-based map of the human proteome. *Science* **347,** 1260419–1260419
37. Mardamshina, M., and Geiger, T. (2017) Next-generation proteomics and its application to clinical breast cancer research. *Am. J. Pathol.* **187,** 2175–2184
38. Kuhn, E., Whiteaker, J. R., Mani, D. R., Jackson, A. M., Zhao, L., Pope, M. E., Smith, D., Rivera, K. D., Anderson, N. L., Skates, S. J., Pearson, T. W., Paulovich, A. G., and Carr, S. A. (2012) Interlaboratory evaluation of automated, multiplexed peptide immunoaffinity enrichment coupled to multiple reaction monitoring mass spectrometry for quantifying proteins in plasma. *Mol. Cell. Proteomics* **11,** M111.013854
39. Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D., and Mann, M. (2016) Plasma proteome profiling to assess human health and disease. *Cell Syst.* **2,** 185–195
40. Bennike, T. B., Bellin, M. D., Xuan, Y., Stensballe, A., Møller, F. T., Beilman, G. J., Levy, O., Cruz-Monserrate, Z., Andersen, V., Steen, J., Conwell, D. L., and Steen, H. (2018) A cost-effective high-throughput plasma and serum proteomics workflow enables mapping of the molecular impact of total pancreatectomy with islet autotransplantation. *J. Proteome Res.* **17,** 1983–1992
41. Fu, Q., Kowalski, M. P., Mastali, M., Parker, S. J., Sobhani, K., van den Broek, I., Hunter, C. L., and Van Eyk, J. E. (2018) Highly reproducible automated proteomics sample preparation workflow for quantitative mass spectrometry. *J. Proteome Res.* **17,** 420–428