Proteomics-Derived Biomarker Panel Improves Diagnostic Precision to Classify Endometrioid and High-grade Serous Ovarian Carcinoma

Dylan Z. Dieters-Castator1,2, Peter F. Rambau3, Linda E. Kelemen4, Gabrielle M. Siegers2, Gilles A. Lajoie5, Lynne-Marie Postovit1,2,6, and Martin Köbel7

Abstract

Purpose: Ovarian carcinomas are a group of distinct diseases classified by histotypes. As histotype-specific treatment evolves, accurate classification will become critical for optimal precision medicine approaches.

Experimental Design: To uncover differences between the two most common histotypes, high-grade serous (HGSC) and endometrioid carcinoma, we performed label-free quantitative proteomics on freshly frozen tumor tissues (HGSC, n = 10; endometrioid carcinoma, n = 10). Eight candidate protein biomarkers specific to endometrioid carcinoma were validated by IHC using tissue microarrays representing 361 cases of either endometrioid carcinoma or HGSC.

Results: More than 500 proteins were differentially expressed (P < 0.05) between endometrioid carcinoma and HGSC tumor proteomes. A ranked set of 106 proteins was expressed (>0.05) between endometrioid carcinoma and HGSC tumor proteomes. A ranked set of 106 proteins was sufficient to correctly discriminate 90% of samples. IHC validated KIAA1324 as the most discriminatory novel biomarker for endometrioid carcinoma. An 8-marker panel was found to exhibit superior performance for discriminating endometrioid carcinoma from HGSC compared with the current standard of WT1 plus TP53 alone, improving the classification rate for HGSC from 90.7% to 99.2%. Endometrioid carcinoma-specific diagnostic markers such as PLCB1, KIAA1324, and SCGB2A1 were also significantly associated with favorable prognosis within endometrioid carcinoma suggesting biological heterogeneity within this histotype. Pathway analysis of proteomic data revealed differences between endometrioid carcinoma and HGSC pertaining to estrogen and interferon signalling.

Conclusions: In summary, these findings support the use of multi-marker panels for the differential diagnosis of difficult cases resembling endometrioid carcinoma and HGSC.

Introduction

Ovarian cancer affects 1.27% of females in the United States and remains a difficult disease to treat with an overall 5-year survival of only 46% (1). Epithelial ovarian cancer (ovarian carcinoma) is a biologically heterogeneous group of diseases classified into five major histotypes, which are in descending order of frequency: high-grade serous (HGSC), endometrioid, clear cell (CCC), low-grade serous, and mucinous carcinoma (2). These histotypes differ with respect to their underlying molecular alterations and outcomes. Patients with endometrioid carcinoma have the best outcome with a 5-year survival of more than 80%, whereas patients with HGSC have the lowest with 40% (3). Histotypes also differ regarding response to therapy; thus, targeted therapies for managing patients stratified by histotype have begun to emerge (4). The ability to distinguish HGSC from endometrioid carcinoma has significant ramifications for choosing the appropriate predictive test and genetic counseling. Clinically approved PARP inhibitors are a promising treatment option for BRCA1/2 defective or chemosensitive HGSC (5), and the antiangiogenic therapy, bevacizumab (Avastin), slightly increases median progression-free survival in HGSC of the mesenchymal or proliferative molecular subtypes (6). In contrast, hormonal therapy is a treatment option for low-stage endometrioid carcinoma and immunotherapy is emerging as an option for metastatic endometrioid carcinoma that are associated with defects in DNA mismatch repair genes (7–10). HGSC should be tested for BRCA1/2 germline mutations and endometrioid carcinoma for mismatch repair deficiency and underlying Lynch syndrome.

Current histotype diagnosis consists of a combination of morphologic assessment by conventional light microscope with the integration of IHC biomarkers, and this can achieve a precision of >90% (11, 12). The current IHC marker standard is the combination of the presence of WT1 and mutant-type TP53, which is highly specific but not entirely sensitive for HGSC (11, 13).
Translational Relevance

Precision medicine approaches in ovarian cancer require the accurate diagnosis of histotypes. In particular, there is a critical need to distinguish endometrioid from high-grade serous (HGSC) carcinomas, which differ greatly in outcomes, genetic predisposition, and optimal treatment approaches. This approach, involving the IHC detection of WT1 and TP53, performs well in discriminating endometrioid carcinoma from HGSC; however, up to 10% of tumors are still misclassified. Herein we used proteomics to discover biomarkers that may distinguish endometrioid carcinoma from HGSC, and then used IHC to validate these using a contemporarily classified cohort of endometrioid carcinoma and HGSC samples. This pipeline revealed eight biomarkers that can be used to increase the accuracy of current diagnostic methods.

Therefore, in a proportion of ovarian carcinomas cases, discrimination between HGSC and endometrioid carcinoma (mostly grade 2 and 3) remains challenging based on similar morphology and overlapping WT1 and TP53 expression (10, 14–16). Currently, the next best markers to distinguish endometrioid carcinoma from HSGC are progesterone receptor (PGR), which suffers from limited sensitivity (81%) and low specificity (56%), and catenin beta-1 (CTNNB1), which has only 50% sensitivity despite >95% specificity (11, 15). Thus, there is a need for better biomarkers that positively define endometrioid carcinoma.

Mass spectrometry (MS)-based proteomics is a powerful and unbiased technique for characterizing complex biological systems (17). For instance, a large scale study of 169 HGSC tumors by the Clinical Proteomics Tumour Analysis Consortium correlated The Cancer Genome Atlas (TCGA) data with protein expression to identify novel signatures and pathways associated with survival (18). In another study, profiling ovarian (cancer) and fallopian tube cell line proteomes revealed three distinct groups (epithelial, clear cell, and mesenchymal) and a 67-protein signature, which segregated patients with HGSC (TCGA) into good and poor overall survival (19). In addition, multiplexed quantitative proteomic analysis of formalin-fixed, paraffin-embedded HGSC, CCC, and endometrioid carcinoma tumors identified cystathionine γ-lyase as a novel marker for CCC (20).

In an attempt to elucidate histotype-specific markers of endometrioid carcinoma, we undertook a MS-based proteomics approach using fresh frozen tumour samples from patients with HGSC and endometrioid carcinoma. Selected candidates were validated by IHC on a cohort of 361 (HGSC and endometrioid carcinoma) ovarian tumors to assess their clinical potential for discriminating these histotypes.

Materials and Methods

Patient samples

All experiments were conducted in accordance with the principles of the Declaration of Helsinki. Fresh frozen tumor tissue was obtained from the Alberta Cancer Research Biorepository following approval by the Human Research Ethics Board of Alberta (HREBA CC-16-555 and CC-16-0913). Ten recent chemo-naïve HGSC and 10 endometrioid carcinoma were selected on the basis of the diagnosis assessed by routine pathology. These cases showed prototypical features by pathology, as well as the expected clinical features (Supplementary Table S1).

Protein extraction from fresh frozen tumors

To prepare samples for LC-MS/MS, tumor cores were partially thawed on ice and a section corresponding to approximately 100 mg was removed with a razor blade. Tumor sections were immediately wrapped in tinfoil and submerged in liquid nitrogen for approximately 10 minutes. Cryopreserved tumor pieces wrapped in tinfoil were pulverized into a fine powder by hitting with a mallet approximately 3–5 times. One ml of lysis buffer containing 8 mol/L urea, 50 mmol/L ammonium bicarbonate (ABC), 10 mmol/L dithiothreitol (DTT), and 2% sodium dodecyl sulfate was added directly to the dissociated tumor tissue, which segregated patients with HGSC (TCGA) into good and poor survival (Supplementary Table S1)

Protein digestion was performed similarly to Duan and colleagues (22). Briefly, precipitated protein pellets were reconstituted in 100 μL of 50 mmol/L ABC (pH 8) and sonicated (~3 x 0.5 seconds pulses) to break up the pellet. LysC (Wako Chemicals) and Mass Spec Grade Trypsin/LysC Mix (Promega) were added to protein samples at 1:100 and 1:50 ratios, respectively. Protein digestion was carried out at 37°C on a ThermoMixer C (Eppendorf) at 300 rpm overnight (~18 hours). The next day an additional volume of trypsin/LysC mix (1:100) ratio was added to each sample and mixed at 1,400 rpm. After 4 hours, digests were acidic to pH 3–4 with 10 μL of 10% formic acid and centrifuged at 14,000 x g to pellet insoluble material prior to LC-MS/MS.

Chloroform/methanol protein precipitation

A 100 μg aliquot of tumor lystate was reduced in 10 mmol/L DTT for 30 minutes and alkylated in 100 mmol/L iodoacetamide for 30 minutes at room temperature in the dark. Proteins were precipitated in chloroform/methanol in 1.5 mL microfuge tubes according to Wessel and Flugge (21). Briefly, samples in lysis buffer were topped up to 150 μL with 50 mmol/L ABC then mixed with ice-cold methanol (600 μL) followed by addition of ice-cold chloroform (150 μL) and vortexed thoroughly. An additional volume (450 μL) of 4°C distilled water was added followed by vortexing and centrifugation at 14,000 x g for 5 minutes. The top aqueous/methanol phase was carefully removed to avoid disturbing the precipitated protein interphase. A second 450 μL volume of cold methanol was added to each sample followed by vigorous vortexing and centrifugation at 14,000 x g for 5 minutes. The remaining chloroform/methanol supernatant was discarded and the precipitated protein pellet was left to air dry in a fume hood.

On-pellet in-solution digestion

On-pellet in-solution protein digestion was performed similarly to Duan and colleagues (22). Briefly, precipitated protein pellets were reconstituted in 100 μL of 50 mmol/L ABC (pH 8) and sonicated (~3 x 0.5 seconds pulses) to break up the pellet. LysC (Wako Chemicals) and Mass Spec Grade Trypsin/LysC Mix (Promega) were added to protein samples at 1:100 and 1:50 ratios, respectively. Protein digestion was carried out at 37°C on a ThermoMixer C (Eppendorf) at 300 rpm overnight (~18 hours). The next day an additional volume of trypsin/LysC mix (1:100) ratio was added to each sample and mixed at 1,400 rpm. After 4 hours, digests were acidic to pH 3–4 with 10 μL of 10% formic acid and centrifuged at 14,000 x g to pellet insoluble material prior to LC-MS/MS.

LC-MS/MS

Digests were initially analyzed using an M-Class NanoAQUITY UPLC System (Waters) connected to an Orbitrap Elite Mass
Spectrometer (Thermo Fisher Scientific) and then, following IHC validation, subsequently analyzed on a Q Exactive Plus (QE+) Mass Spectrometer (Thermo Fisher Scientific). Buffer A consisted of water/0.1% formic acid and buffer B consisted of acetonitrile/0.1% formic acid. Peptides (~1 μg measured using a Pierce bichinchoninic acid assay) were initially loaded onto an ACQUITY UPLC M-class symmetry C18 trap column (5 μm, 180 μm × 20 mm) and trapped for 6 minutes at a flow rate of 5 μL/min at 99% A/1% B. Peptides were separated on an ACQUITY UPLC M-class peptide BEH C18 column (130 Å, 1.7 μm, 75 μm × 250 mm) operating at a flow rate of 300 nL/min at 35°C using a nonlinear gradient consisting of 1%–7% B over 1 minute, 7%–23% B over 179 minutes (134 for QE+) and 23%–35% B over 60 minutes (45 for QE+) before increasing to 95% B and washing. MS acquisition settings are provided in Supplementary Table S2. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012998 (23, 24).

Data analysis
MS files were searched in MaxQuant (1.5.8.3) using the Human UniProt database (reviewed only; updated May 2017 with 42,183 entries; refs. 25, 26). Missed cleavages were set to three and cysteine carbamidomethylation was set as a fixed modification. Oxidation (M), N-terminal acetylation (protein), and deamidation (NQ) were set as variable amino acid modifications (maximum number of modifications per peptide = 5) and all other settings were left as default. Protein and peptide FDR were set to 0.01 (1%) and the decoy database was set to revert. The “match between runs” feature was utilized to maximize proteome coverage and label-free quantification (LFQ; ref. 27). Datasets were loaded into Perseus (version 1.5.3.5 and 1.6.2.1) and proteins “only identified by site” or “matched to reverse” (decoy) database hits were removed (28). Proteins detected (razor + unique peptides) and quantified (log2-transformed LFQ intensities) in each tumor sample are provided in Supplementary Tables S3 and S4. Protein identifications with quantitative values in ≥2 samples in a least one group (HGSC or endometrioid carcinoma) were retained for downstream analysis unless specified elsewhere. Missing values were imputed using a width of 0.3 and down shift of 1.8. Gene ontology (GO) cellular component analysis and pathway annotation were performed using PantherDB (release 13.1) and Metascape (version 3.0), respectively (29, 30). Gene set enrichment analysis (GSEA) were performed in GSEA v3.0 (Broad Institute, Cambridge, MA) with all settings left default (except minimum gene set size of 10) and the Molecular Signature Database (MSigDB) v6.2 collections: canonical pathways, hallmark, oncogenic signatures, and GO biological processes (31). Support vector machine analysis and Pearson correlation heatmaps were produced using the Bioconductor packages getNetClassifier and complexHeatmap, respectively (32, 33). The Estimation of STromal and Immune cells in MALIGNant Tumours using Expression data (ESTIMATE) score in R was used as a surrogate to assess epithelial (tumor) cell purity (34). Gene lists belonging to each of the four HGSC molecular subtypes and z-scored protein expression data were used to cluster HGSC samples in this study and the CPTAC (19, 35).

IHC
The validation samples were from patients diagnosed with HGSC and endometrioid carcinoma between 1978 and 2010 identified from the Alberta Cancer Registry in Canada. Details of histology review and tissue microarray construction, as well as subsequent biomarker-based histotyping, were described previously (11, 36), resulting in 172 HGSC and 189 endometrioid carcinoma (HREBA CC-16-161; Supplementary Table S1). IHC staining was performed on 4-μm sections from tissue microarrays representing each case by up to three 0.6-mm cores using a DAKO Omnis platform within the Anatomical Pathology Research Laboratory at Calgary Laboratory Services in the Department of Pathology and Laboratory Medicine at the University of Calgary (Calgary, Alberta, Canada). Detailed assay and antibody information, as well as scoring cutoffs are provided in Supplementary Table S5.

Statistical analysis
Two-sided, two sample Student t tests (P < 0.05) were used to identify differentially expressed proteins. ROC-AUCs were calculated from proteomics data in GraphPad Prism (Version 6.01). Fisher exact test was used to determine statistically significant differences (P < 0.05) between the proportion of endometrioid carcinoma and HGSC tumors staining positive and negative for each marker. Nominal logistic regression and Kaplan–Meier survival analysis were performed in JMP v13 (SAS).

Results
Global proteomic analysis of HGSC and endometrioid carcinoma
To identify proteins elevated in or exclusive to endometrioid carcinoma versus HGSC, we performed label-free quantitative proteomics on unfracated digests from freshly frozen tumors (n = 10 endometrioid carcinoma and n = 10 HGSC samples; Supplementary Fig. S1A). Patient characteristics for each subtype are listed in Supplementary Table S1. On average, approximately 4,700 (unique) proteins (3,143–5,101 proteins; Supplementary Fig. S1B) were identified in each sample (~98% contained LFQ intensities) and 6,360 proteins were detected in total (Supplementary Fig. S1C). Most proteins detected were of cytoplasmic origin (cell part) or belonging to macromolecular protein complexes and organelles according to PantherDB (Supplementary Fig. S1D; ref. 29). Within the Perseus computational platform, approximately 8% of the quantified proteins within each sample were similarly annotated with the cellular component (GO slim) term “extracellular matrix” or “extracellular region” and these proteins accounted for a mean 37% of the total MS signal quantified (Supplementary Fig. S2A and S2B; ref. 28). Interestingly, the number of quantified proteins per sample were negatively correlated (r = −0.965; Supplementary Fig. S2C) with the cumulative MS signal (abundance) of these extracellular proteins. Analysis using the ESTIMATE tool revealed minor variations in stromal and immune scores, and epithelial (tumor) purity (0.67–0.87; Supplementary Fig. S2D–S2F; ref. 34).

We next compared HGSC and endometrioid carcinoma samples using unsupervised hierarchical clustering and principal component analysis (PCA) on a filtered list of approximately 5,648 proteins (minimum ≥3 LFQ intensities in a least one group). Although 98% of these proteins were detected in at least one sample within each histotype (Fig. 1A) and tumor proteomes correlated regardless of histotype (Supplementary Fig. S3), most endometrioid carcinoma and HGSC samples (except HGSC-1 and 9) clustered independently based on PCA
Furthermore, 537 proteins were found to be differentially expressed (two-sample t test; \( P < 0.05 \); Fig. 1C; Supplementary Table S6). Among the differentially expressed proteins, we identified known highly expressed markers in endometrioid carcinoma, for example PGR (4.20 log2 fold change), MMP7 (1.45 log2 fold change), and CTNNB1 (0.86 log2 fold change) and highly expressed in HGSC, for example TP53 (−2.06 log2 fold change), MSLN (−2.67 log2 fold change), IGF2 (−1.62 log2 fold change), and CDKN2A (−2.57 log2 fold change), which supported the validity of our approach (2, 37).

Selecting endometrioid carcinoma enriched proteins for IHC validation

To filter candidates of endometrioid carcinoma for validation by IHC, we first tabulated proteins with peptide evidence in \( \geq 80\% \) of tumor samples from one and \( \leq 20\% \) of samples from the other histotype (Supplementary Table S7). Although this strategy identified 15 and seven proteins that were largely specific to endometrioid carcinoma and HGSC tumors, respectively, it may not capture proteins with high differential expression. Therefore, we analyzed our proteomics data using the R package geNetClassifier to rank proteins with the greatest classification power in an unbiased fashion (32). In total, 106 proteins exceeded the posterior probability cutoff (\( \geq 0.95 \)) and were used in training the support vector machine (Supplementary Table S8). The lowest error rate achieved by geNetClassifier was 0.1 (10\%) and corresponded to 69 proteins. Interestingly, the top two ranked proteins, MUC5B (mucin-5B) and PIGR (polymeric immunoglobulin receptor), were not identified by our initial filtering strategy whereas PGR was identified by both approaches. We subsequently performed unsupervised hierarchical clustering utilizing Pearson correlation coefficients calculated from LFQ intensities restricted to the top 106 proteins ranked by geNetClassifier. This clustering approach segregated HGSC and endometrioid carcinoma samples relatively well but revealed a third central cluster comprised of three endometrioid carcinoma and two HGSC samples (Fig. 2A). In agreement with our earlier findings, some tumor proteomes, HGSC-1 and 9 in particular, more closely resembled the opposing histotype suggesting close biological relation.

Validating diagnostic markers for endometrioid carcinoma

On the basis of availability of specific antibodies for IHC validation, we selected six endometrioid carcinoma candidates with large log2 fold changes (MUC5B and PIGR) and/or high specificity (PLCB1, PAM, KIAA1324, and SCGB2A1). ROC-AUCs for the selected candidates for endometrioid carcinoma ranged (Fig. 1B).
from 0.82 to 0.99 when calculated using LFQ intensities (Fig. 2B). Validation by IHC was performed on 361 independent tumor sections (189 endometrioid carcinoma and 172 HGSC) scored using a three-tier system (Fig. 3A; Supplementary Figs. S4 and S5). For comparison, we also performed IHC on standard diagnostic histotype-specific markers of endometrioid carcinoma (PGR and CTNNB1) and HGSC (WT1 and TP53). For all endometrioid carcinoma candidates, we confirmed that expression was significantly greater in endometrioid carcinoma compared with HGSC (Supplementary Fig. S5). Of note, KIAA1324, an estrogen-induced gene (alternative name EIG121; ref. 38), was the only marker that exhibited performance similar to that of PGR with a sensitivity of 88.5% and specificity of approximately 53% (Fig. 3B). Although PLCB1 and PAM yielded the highest ROC-AUC values in our proteomics dataset (0.99 and 0.97, respectively), these proteins were among the least sensitive markers by IHC (62.94 and 56.90%, respectively; Figs. 2B and 3B). SCGB2A1 (mammaglobin-B) was the least specific marker (32.59%), which was unexpected given that its expression, like KIAA1324, was relatively exclusive to endometrioid carcinoma versus HGSC samples in our proteomics analyses (Fig. 3B; Supplementary Table S7). WT1 and TP53, as anticipated, performed exceptionally well with sensitivities of 97.6% and 92.2%, and specificities of 91.4 and 88.7% for HGSC, respectively (Fig. 3B).

Given that individual markers did not outperform known standards such as PGR, we evaluated multi-marker combinations using nominal logistic regression. Our six endometrioid carcinoma candidates (PLCB1, PAM, KIAA1324, SCGB2A1, MUC5B, and PIGR) yielded a ROC-AUC of 0.90016 albeit 32 of 166 (19.3%) were selected as candidate (PLCB1, PAM, KIAA1324, SCGB2A1, MUC5B, and PIGR) to achieve a ROC-AUC of 0.99751 (Model 5; Fig. 4E). The final 8-marker Model 5 could not classify four cases [3/170 (1.8%) endometrioid carcinoma and 1/130 (0.8%) HGSC)] and thus represents an improvement when compared with WT1 and TP53 alone (Fig. 4F).

Next, we evaluated prognostic associations between markers for endometrioid carcinoma and patient survival. Interestingly, PLCB1 [HR = 0.38; 95% confidence interval (CI), 0.18–0.78], KIAA1324 (HR, 0.23; 95% CI, 0.11–0.51), and SCGB2A1 (HR, 0.35; 95% CI, 0.16–0.75) were associated with favorable outcome in endometrioid carcinoma with log-rank P<0.0001, 0.0003, and 0.0120, respectively (Fig. 5).

Deep proteome coverage obtained through complementary MS approaches reveals differences in endometrioid carcinoma and HGSC biology

To explore biological pathways in detail that differ between endometrioid carcinoma and HGSC, we acquired a second proteomics dataset using a more sensitive Q Exactive Plus mass spectrometer and the same unfractionated samples. For reference, using the same filtering criteria, approximately 26% and approximately 32% more proteins were identified (8,041 vs. 6,360) and retained (7,452 vs. 5,648) for quantitative comparisons (endometrioid carcinoma vs. HGSC) with this instrument, respectively (Supplementary Fig. S8A and S8B; Supplementary Table S9). PCA appeared unaffected (near identical components 1 and 2) by the additional depth and, importantly, all candidates we validated, except for PLCB1, were significantly elevated in endometrioid carcinoma, thus indicating minimal technical variability between each instrument (Supplementary Fig. S8C and S8D). In addition, our discovery set of HGSC samples appeared to exhibit protein expression patterns resembling each of the four previously reported molecular subtypes (35), and this observation was further corroborated by clustering our HGSC proteomes with the 169 HGSC samples from the CPTAC (Supplementary Fig. S8; ref. 19).

The differentially expressed proteins and complete proteomes (endometrioid carcinoma and HGSC) from our second proteomics dataset were analyzed using Metascape (www.metascape.org) and GSEA to identify overrepresented and enriched pathways/processes, respectively (30, 31). In endometrioid carcinoma, terms associated with trafficking and localization, estrogen...
signalling, and fatty acid metabolism appeared consistently over-represented or enriched (Fig. 6A; Supplementary Tables S10 and S11). In agreement with the IHC, CTNNB1 activation (BCAT.100_UP.V1_UP) was also significant in endometrioid carcinoma (normalized enriched score = 1.64; FDR q-value = 0.24). Alternatively, immune/antiviral responses, mitotic and cell morphologic processes, and androgen receptor signalling (PID_AR_PATHWAY) were predominately matched to HGSC (Fig. 6B; Supplementary Tables S12 and S13). Following closer inspection, key differences between endometrioid carcinoma and HGSC could be illustrated by heatmap using normalized (z-scored) expression values of core proteins enriched in estrogen response late (Hallmark) and antiviral mechanism by IFN-stimulated genes (Reactome) gene sets (Fig. 6D). Indeed, known estrogen-induced genes were highly expressed in endometrioid carcinoma, whereas the E1 (UBA7), E2 (UBE2L6), and
E3 (HERC5) ubiquitin ligases that mediate ISGylation of many targets (MX1, MX2, IFT1, UBE2E1, STAT1, DDX58, EIF2AK2, and UBE2L6) via conjugation of the ubiquitin-like modifier IFN-stimulated gene 15 (ISG15) were central to the IFN response in HGSC (39). The Ubl carboxyl-terminal hydrolase 18 (USP18) that removes ISGylation was also enriched in HGSC.

**Discussion**

Through robust label-free quantitative proteomics and IHC analyses, we identified and validated an eight-marker panel, which could improve the classification rate for HGSC from 90.7% (standard model consisting of WT1 plus TP53; ref. 11) to 99.2%. This improvement greatly decreases the risk of undertreatment of patients with aggressive HGSC increasing eligibility for targeted therapies (PARP inhibitors) and triaging to the appropriate hereditary screening programs (BRCA1/2 hereditary breast/ovarian cancer as opposed to Lynch syndrome screening).

In the era of precision medicine, wherein accurate diagnosis is of utmost importance, our discovery is of high clinical utility. Using our approach, we were able to identify a number of novel (KIAA1324, SCGB2A1, MUC5B, PLCB1, PAM, and PIGR) and existing (PR, CTNNB1, WT1, and TP53) histotype-specific biomarkers for the differential diagnosis of endometrioid carcinoma and HGSC. Although novel candidates effectively discriminated endometrioid carcinoma and HGSC proteomes, KIAA1324 was the only individual biomarker that achieved good performance by IHC. The performance of individual markers in IHC was generally worse than that predicted by the proteomics data, which was mostly due to a lack of specificity for endometrioid carcinoma using the most widely clinically accepted cutoff for ancillary diagnostic IHC markers (absence vs. presence of expression). SCGB2A1 (mammaglobin-B), for example, is primarily expressed in the breast, uterus, and salivary gland (40), and greater mRNA expression levels have been reported in endometrioid compared with serous ovarian cancer samples (41, 42). We also found SCGB2A1 to be one of the top differentially expressed proteins.
in our study, but IHC demonstrated only 34% specificity for endometrioid carcinoma. This suggests that unlike IHC, low abundance proteins such as transcription factors remain difficult to detect using unfractionated (single-shot) proteomic analysis. Unfortunately, due to lack of quality antibodies, we could not validate additional targets with high specificity for endometrioid carcinoma (e.g., PPM1H); hence the targeted development of quality antibodies may be warranted. Alternatively, our results indicate a multi-marker model for endometrioid carcinoma classification is a promising strategy given the lack of validating of a single defining endometrioid carcinoma marker. This notion is also supported by recent work by Martinez-Garcia and colleagues who employed targeted proteomics on uterine aspirates from patients with and without endometrial cancer, a disease sharing similarities with endometrioid carcinoma (43–45). Of note, SCGB2A1, CTNNB1, and PIGR were significantly elevated in patients with endometrial cancer compared benign conditions, and among others were the best single performing markers for discriminating endometrioid and serous histotypes in the endometrium as well (45, 46). As proteomics technologies and protocols become more robust, easy to use and economical, for example, in the area of plasma biomarkers, there is potential for direct integration into current clinical work flows (47). However, current IHC-based classification is the most efficient first-line approach, which is superior compared with targeted sequencing approaches. Although histotype-specific molecular differences between HGSC and endometrioid carcinoma exist, even their combined sensitivities are much less than that of IHC markers that is, many cases are not informative. For example, BRCA1/2 mutations are only present in about 25% of HGSC, and ARID1A mutations in about 35% of endometrioid carcinoma.

Beyond our biomarker discovery and validation work, pathway analysis revealed interesting biological differences between endometrioid carcinoma and HGSC, for example, with estrogen signalling associated with the former and interferon signalling with the latter. This finding was also corroborated by a previous study by Hughes and colleagues (20). The importance of hormone receptor signalling for endometrioid carcinoma is well established and ESR1 and PGR are validated as prognostic markers for endometrioid carcinoma (7). KIAA1324, an estrogen-induced gene, was highly coexpressed with PGR and was associated with favorable prognosis in endometrioid carcinoma. Therefore, KIAA1324 could be used as a complementary diagnostic, prognostic, and predictive marker for an active hormone axis in endometrioid carcinoma by identifying patients that may benefit from hormonal therapy. Interestingly, two other diagnostic endometrioid carcinoma markers (PLCB1 and SCGB2A1) also showed favorable prognostic associations in endometrioid carcinoma suggesting that these markers segregate a biologically more aggressive subset within endometrioid carcinoma that does not express typical endometrioid carcinoma markers. Indeed, SCGB2A1 expression is associated with favorable prognosis in endometrial carcinomas and longer recurrence-free survival in ovarian carcinomas (41, 48). Lack of these markers might warrant more aggressive therapy, similar to that used for HGSC.

The enrichment of key regulators and targets of IFN responses essential to innate immunity and antiviral defense in HGSC proteomes is interesting given that IFN signalling is commonly associated with antitumoral effects, including cell-cycle arrest, apoptosis, and immune responses (49). Among these proteins was the ubiquitin-like modifier ISG15, which was previously detected in 86% of HGSC and associated with increased overall survival (50). STAT1, which mediates type I (IFNα/β) and II (IFNg) IFN responses, was also elevated in HGSC, and previously associated with increased progression-free survival, response to chemotherapy and CD8+ T-cell infiltration in HGSC (51, 52).

Paradoxically but interestingly, recent evidence implicates IFN/JAK/STAT signaling in resistance to immune checkpoint
Figure 6.
Pathway analysis reveals characteristics of endometrioid carcinoma and HGSC. Proteins significantly elevated in endometrioid carcinoma (A) or HGSC (B) samples were analyzed in Metascape, and the top 20 significant overrepresented pathways are shown. C, Heatmap comparing similarities and differences between highly significant process and pathways. D, Heatmap showing normalized (z-scored) expression of core proteins matching to estrogen response late or antiviral mechanism by IFN-stimulated genes.
blockade, which might explain why immune checkpoint blockade is ineffective in HGSC (53).

There are several potential limitations of our proteomic analysis that have implications for candidate selection. First, the sample size of the discovery analysis was relatively small. However, our goal was to identify highly discriminatory diagnostic biomarkers that work in most samples all the time, therefore, more cases in the discovery cohort would not have yielded a different result. We observed two HGSC samples with proteomes overlapping with endometrioid carcinoma, indicating a relatively close biological relation between these tumor types. Hence, finding a single discriminatory marker is unlikely. However, we believe that the potentially confounding effect of those two HGSC samples is small because we focused our discovery on defining endometrioid carcinoma biomarkers. Our focus on endometrioid carcinoma markers also diminishes the role of intertumoral heterogeneity in HGSC, namely molecular subtypes. While we showed reasonable representation of molecular subtypes in the discovery cohort, the size of validation almost certainly confirms a representation of all HGSC molecular subtypes. Of note, we selected for tumor intrinsic biomarkers (i.e., proteins expressed in epithelial tumor cells), whereas molecular subtype is mainly a reflection of gene expression differences within the microenvironment. Additional validation using even larger sample sizes will be important to confirm our proteomics findings. For example, our marker panel can be applied to the large OTTA consortium, with known molecular subtypes, to refine histotype-specific outcome associations (37). Third, some putative diagnostic proteins may not exhibit high differential expression but may differ predominately in cellular compartmentalization, mutation status, and posttranslational modifications (e.g., CTNNB1 and TP53; ref. 13). Yet we identified CTNNB1 and TP53 as differentially expressed proteins, which supports the validity of our approach. Fourth, the preponderance of stromal, immune, and other cell types in the microenvironment may mask differences in low abundance proteins between endometrioid carcinoma and HGSC. Although our approach yielded the desired outcome to improve upon classification of HGSC and endometrioid carcinoma, we would make the following recommendation for future proteomic studies: To increase the sensitivity for identification of tumor intrinsic protein markers, enrichment techniques such as laser capture microdissection, single-cell approaches, posttranslational modification analysis, and/or subcellular fractionation may reveal additional candidates.

In conclusion, we have discovered and validated a biomarker panel that can improve the ability to discriminate between endometrioid carcinoma and HGSC. Application of this protocol could substantially improve the classification of endometrioid carcinoma and HGSC and could direct more precise histotype-specific treatment options.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: L.-M. Postovit, M. Köbel
Development of methodology: D.Z. Dieters-Castator, G.A. Lajoie, L.-M. Postovit, M. Köbel
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.Z. Dieters-Castator, P.F. Rambau, L.E. Kelemen, G.A. Lajoie, L.-M. Postovit, M. Köbel
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D.Z. Dieters-Castator, P.F. Rambau, L.E. Kelemen, G.A. Lajoie, L.-M. Postovit, M. Köbel
Writing, review, and/or revision of the manuscript: D.Z. Dieters-Castator, P.F. Rambau, L.E. Kelemen, G.M. Siegers, G.A. Lajoie, L.-M. Postovit, M. Köbel
Study supervision: L.-M. Postovit

Acknowledgments
We thank Paula Pittock for technical support and Dr. Miljan Kuljanin for feedback on the article. This work was supported by the Savin-Bladwin Chair in Ovarian Cancer Research and the Dr. Anthony Noujaim Oncology Chair awarded to L.-M. Postovit by the Women and Children Health Research Institute and the Alberta Cancer Foundation, respectively. We thank Shuhong Liu, Young Ou, and Deon Richards for immunohistochemical stains, and Thomas Keyton, BFA, digital imaging specialist for Alberta Public Lab, for scanning the tissue microarrays. The study was supported by Alberta Public Lab internal research support RS17-608 and an NSERC Operating Grant awarded to G. A. Lajoie.

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

Received November 23, 2018; revised February 18, 2019; accepted April 8, 2019; published first April 12, 2019.

References
Proteomics-Derived Biomarker Panel Improves Diagnostic Precision to Classify Endometrioid and High-grade Serous Ovarian Carcinoma
