Proteomics of endometrial cancer diagnosis, treatment and prognosis

Parul Mittal¹, Manuela Klingler-Hoffmann¹, Georgia Arentz¹, Chao Zhang¹, Gurjeet Kaur³, Martin K. Oehler², Peter Hoffmann¹*

1. Adelaide Proteomics Centre, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, South Australia, 5005
2. Department of Gynaecological Oncology, Royal Adelaide Hospital, North Terrace, Adelaide, South Australia, 5000
3. Institute for Research in Molecular Medicine, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia

* To whom all correspondence and requests for reprints should be addressed:
Prof. Peter Hoffmann,
Adelaide Proteomics Centre, University of Adelaide,
Gate 8 Victoria Drive, Adelaide, South Australia, 5005
Phone: +61 (08) 8313 5507; Fax: +61 (08) 0 8313 4362;
Email: peter.hoffmann@adelaide.edu.au

Keywords: Biomarker, Endometrial cancer, Proteomics,

Abbreviations: EC (Endometrial cancer), DDA (Data Dependent Acquisition), DIA (Data Independent Acquisition), FFPE (Formalin Fixed Paraffin Embedded), MALDI-MSI (Matrix Assisted Laser Desorption/Ionisation-Imaging)
Abstract

This review discusses the current status of proteomics technology in endometrial cancer diagnosis, treatment, and prognosis. The first part of this review focuses on recently identified biomarkers for endometrial cancer, their importance in clinical use as well as the proteomic methods used in their discovery. The second part highlights some of the emerging mass spectrometry based proteomic technologies that promise to contribute to a better understanding of endometrial cancer by comparing the abundance of hundreds or thousands of proteins simultaneously.
Introduction

Endometrial cancer (EC) is the most common malignant tumour of the female reproductive tract.

According to the American Cancer Society, an estimated 52,630 new EC were diagnosed and 8,590 patients died from the disease in 2014 in the USA. From 2006 to 2010, incidence rates of EC increased by 1.5% per year among women younger than 50 years and by 2.6% per year among women 50 years and older (1). Based on the classification system introduced by Bokhman in 1983, EC is divided into two histological subtypes: endometroid adenocarcinomas (Type I) and non-endometroid (Type II) carcinomas (2). Type I adenocarcinomas account for 90% of all EC. These cancers are usually low grade, diagnosed at an early stage and have a good prognosis. They are associated with oestrogen excess, obesity and atypical hyperplasia. In contrast, type II non-endometroid tumours are high grade tumours with more aggressive biological behaviour compared to type I disease. Clinically, type II cancers are often diagnosed at advanced stage when prognosis is poor. Type II tumours include serous, and clear cell carcinomas as well as carcinosarcomas.

Proteomics, more precisely the comparative quantitation of protein subsets, holds great promise in improving outcome of patients with EC as it provides unique tools for discovery of new biomarkers and therapeutic targets. To date, biomarkers have proven their tremendous clinical value in early diagnosis (3, 4), categorizing different subtypes of malignancies (5, 6) and in monitoring patient’s response to therapy (7, 8). Regardless of the numerous proteomic studies that have contributed to the standardization of experimental protocols for digestion (9), separation (10), enrichment (11), identification (12) and quantification (13) of less abundant proteins by highly efficient mass spectrometric techniques, proteomic research is still restricted by both technology and bioinformatics tools. Many protein and peptide peaks have been reported to bear significant diagnostic (14), prognostic (15) or predictive value (16) for EC; however, the candidate biomarkers have not yet been validated for use in clinical patient care (17). Some authors have speculated that this might be due to studies using a single proteomics approach, which is not sufficient to gain an in
depth understanding of the protein function and does not eliminate false-negative and false-positive results (18). Nevertheless, implementing ‘omics’ integration approaches and the use of the rapidly emerging mass spectrometry based proteomic technologies will aid in the venture to elucidate protein markers and their function, ultimately providing more reliable, sensitive, and specific biomarkers for EC.

In the first part of this review article we will summarize the various mass spectrometry based proteomic approaches that have been used in EC studies. In the second part, we will introduce emerging proteomic technologies which may not only be used as discovery tools but also allow the implementation of mass spectrometry in the diagnosis and prognosis of EC.

1. Clinical considerations of endometrial cancer

The clinical management of EC remains a challenge with patients presenting with a full spectrum of disease ranging from those with excellent prognosis and high curability to aggressive disease with poor outcome (19). EC is staged according to the International Federation of Gynaecology and Obstetrics (FIGO) system (Table 1) (20). The classification of EC into low and high risk disease is dependent on a number of parameters, only two of which are assessable pre-operatively after a biopsy has been obtained, namely the histological type and grade. However, Jacques et al. have shown a large percentage of cancer will be classified as a higher grade tumour after definitive surgery (21). Therefore, the exact type, grade and stage of EC can only be determined by surgery and subsequent histopathological assessment.

1.1. Diagnosis

EC is frequently diagnosed at an early stage, as it regularly presents with symptoms such as post-menopausal bleeding which usually develops early in the disease process. Nevertheless, the discovery of serum biomarkers for early detection of EC has become a high priority (22). A number of
serum biomarkers has been identified so far, with the most commonly used serum biomarker in gynaecological oncology being Carbohydrate Antigen 125 (CA 125). In EC, an elevated serum CA 125 level have been detected in 11-43% of the cases and is shown to be correlated with advanced stage and with the presence of extraterine disease (23-25). However, due to lack of specificity and sensitivity, CA 125 has limited significance in the diagnosis of EC. Human Epididymis protein 4 (HE4) has recently emerged as a promising biomarker for EC (26). Brennan et al. highlighted the utility of serum HE4 using ELISA for pre-operative risk stratification to identify high-risk patients within low-grade endometrioid EC patients who might benefit from lymphadenectomy (27). For stage I EC, HE4 showed a 17% improvement in sensitivity when compared to CA 125 (28).

Hareyama et al. investigated the immunohistochemical CA72-4 expression in EC. The authors reported an elevated level of serum CA72-4 in 22-32% of the cases, which is associated with depth of myometrial invasion, adnexal metastasis, lymphovascular space invasion, and pelvic and para-aortic lymph node metastasis (29). Konno et al. demonstrated the level of serum soluble Fas (sFas) is significantly higher in EC patients with advanced cancer when compared to localized cancer (p < 0.0001) (30). A study by Sawada et al. showed the raised serum level of immunosuppressive acidic protein (IAP) in 55-76% of EC cases and the level increases with the stage of the disease (31). Hakala et al. reported the serum level of macrophage colony-stimulating factor 1 (mcsf 1), correlates significantly with tumour grade and poor prognosis in 25-73% of EC cases (32) (33). A recent study by Kang-Wai Mu et al. reported differential levels of zinc alpha-2 glycoprotein, alpha 1-acid glycoprotein, and CD59 in the urine of EC patients when compared to urine from healthy controls using Two Dimensional Gel Electrophoresis (2-DIGE) and O-Glycan binging lectin (34).

1.2. Treatment

EC is usually treated by surgery consisting of a hysterectomy, bilateral salpingo-oophorectomy and pelvic lymph node dissection and has a very good prognosis when it is confined to the uterus (28). In contrast, EC with metastasis to the lymph nodes has a high mortality rate. Lymph node metastasis is
therefore a crucial factor in the prognosis and choice of treatment of EC patients. Clinicians presently face the challenge that conventional surgical-pathological variables (e.g. tumour size, depth of invasion and grade of disease) and radiological imaging are unreliable in determining if a gynaecological cancer has spread. Consequently, although only 5% of patients suffer from metastasis, the majority undergo radical treatment including removal of the lymph nodes. Lymph node dissection, however, is associated with significant complications such as lower extremity lymphoedema (35). Predictive tissue markers for lymph node metastasis are therefore warranted to determine the optimal treatment strategy and to avoid morbidity in gynaecological cancer.

1.3. Prognosis

The prognosis of EC depends on various factors including the histological subtype, grade and stage of the disease (which is determined by depths of tumour invasion, lymph node metastasis and spread of the disease to other organs) (36). The depth of myometrial invasion and histological grade strongly correlate with the prevalence of lymph node metastasis and patient survival (37).

Various protein biomarkers have been described in EC that could be of benefit for the prediction of disease outcome. According to Cocco et al. serum amyloid A (SAA) may represent a novel biomarker for EC to monitor disease recurrence and response to therapy (38). Lo SS et al. shown the increased levels of CA125, CA15.3 and CA19.9 were significantly associated with poor prognostic clinical parameters (39). Lambropoulou et al. evaluated the prognostic significance of survivin, c-erbB2, and COX-2 levels in EC and stated that in a 10-year follow-up, patients with tumors expressing more of these three antigens had significantly lower survival rate that those with smaller expression score (40). Yilmaz et al. have also shown higher nuclear expression of survivin in type I when compared to type II EC (p=0.040); but no difference for cytoplasmic survivin and matrix metalloproteinase-2 expressions between type I and type II EC (41). Zeimet et al., identified L1 cell adhesion molecule (L1CAM) as the best variable for predicting recurrence (sensitivity = 0.74; specificity = 0.91) and death (sensitivity = 0.77; specificity = 0.89) (42). Although, the above mentioned biomarkers are able
to identify high risk patients with low grade EC, for clinical application additional biomarkers are
required. New prognostic EC biomarkers will help distinguish patients who are at a low-risk of
developing metastasis as compared to those who are at a high risk. Ideally, novel biomarkers will
also be able to distinguish patients who have already developed metastasis. This would
consequently allow patients to selectively undergo radical surgery while avoiding those who will not
benefit from it and subsequently decreasing the risk of post-surgical morbidity.

2. Molecular Genetics of endometrial cancer

Aside from the morphologic and clinical features, EC can be further distinguished into type I and type
II EC on the basis of genetic alterations (43). Type I and type II ECs are associated with mutations of
independent gene sets (44). Type I endometrial adenocarcinomas are characterized by mutations in
the PTEN, K-ras and β catenin, as well as DNA mismatch repair genes (35) while type II non-
endometrioid EC frequently shows aneuploidy and p53 mutations (35).

PTEN has been reported to be the most frequent genetic alteration of type I EC. PTEN, a tumour
suppressor gene has been altered in up to 83% of endometrial adenocarcinoma and 55% in pre-
cancerous lesions (45). Microsatellite instability (MSI) has been demonstrated in 20% of type I EC
(46). According to Bilbao et al. both PTEN mutations and MSI represent early events in endometrial
carcinogenesis (47). Other genetic alterations that occur in type I EC include mutations in K-ras (48)
and β catenin (49). PTEN, MSIS and K-Ras mutations often coexist with each other, whereas
mutations in β-catenin are usually observed alone (50).

The most common genetic alteration in type II EC is in p53, the tumour suppressor gene with an
occurrence of 93% mutations in p53 in type II EC (48, 51). Other frequent genetic alteration in type II
ECs are inactivation of the tumour suppressor gene p16 and over-expression of the oncogene, HER-
2/neu (52).
3. Metastasis and protein biomarker

Metastatic cancer cells proliferate, lose contact with neighbouring cells, migrate through interstitial matrix, invade blood and lymph vessels and grow out again into lymph nodes or distant organs (53). Most cancer cells fail to undergo metastasis due to deficiency in one of the required steps like invasion, detachment, and survival (54). Discovery of EC biomarkers would significantly aid gynaecological oncologists who currently face the challenge that radiological imaging and conventional surgical-pathological variables such as tumour size, depth of invasion and grade of disease are unreliable in determining if an EC has metastasized.

The working model of metastasis implies that primary tumour cells acquire genetic alterations over time, which enables these cells to metastasize and form new solid tumours at distant sites (55). It has been proposed that the gene expression program of metastasis is actually present in the bulk of primary tumour cells (56). Ramaswamy et al. shown some primary tumours are preconfigured to metastasize, and their susceptibility is detectable at the time of diagnosis. They defined a 17-gene signature pattern associated with metastasis, in which eight genes are up-regulated, while nine genes are down-regulated. They also concluded that metastasis is not dependent on only a single gene but on the complete 17 gene sequence.

Yi et al., shown the expression of cyclooxygenase (COX-2) plays an important role in metastasis of EC (57). They identified the COX-2 inhibitor NS-398 inhibits proliferation, viability and invasion of the EC cell line RL95–2 (57). Maxwell et al. confirmed annexin A2 (ANXA2) and peroxiredoxin (PRDX1) were both being overexpressed in stage 1 EC when compared to normal endometrium (58).

4. Mass Spectrometry based proteomic approaches in endometrial cancer

The use of proteomic technologies are now enabling the identification and relative quantification of multiple proteins simultaneously from a single experiment to identify disease related and specific biomarkers. Proteomics in general deals with the large-scale determination of gene and cellular
function directly at the protein level (59). The proteome has also been defined as the protein complement expressed by a genome (60) but in reality, due to post-translational modifications and alternative splicing, it is estimated that about 22,000 protein coding genes (61) code for more than 500,000 proteins (62) in the human proteome.

In serum it is estimated that cancer derived proteins are 10 million times less abundant than common high abundant proteins secreted by normal cells. Therefore, it is challenging to correctly identify and quantify tumor derived proteins from the whole serum proteome (63). Alternatively, it is obvious that cancer derived proteins will be present in higher concentration in native tissue, organs and their proximal fluids from where the tumor has originated as compared to the distant sites in which the tumor derived proteins may be secreted or leaked. Therefore, targeting those regional sites will dramatically increase the possibility of isolating and identifying tumor specific biomarkers (64).

Proteomic analysis principally relies on mass spectrometry (MS) for protein identification. The uses of MS based techniques have made it possible to build comprehensive profiles of near complete proteomes, comparing the expressions of individual proteins that may serve as biomarkers. Generally, MS based proteomic techniques can be classified as ‘gel based’ and ‘gel free’ (Figure 1). In gel based approaches such as One Dimensional Polyacrylamide Gel Electrophoresis (1D-PAGE), Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) and Two-Dimensional Differential Gel Electrophoresis (2D-DIGE), proteins are separated on polyacrylamide gels via electrophoresis in one or two dimensions, creating unique proteomic patterns based on the mass and/or charge/pH of the proteins.

Gel free MS approaches can be either label based or label free. Label based proteomic techniques are common for mass spectrometry approaches where proteins are tagged either with isotopes such as isotope coded affinity tags (ICAT), isobaric tags for relative and absolute quantification (iTRAQ) or labelled chemically such as stable isotope labelling by amino acids in cell culture (SiLAC). Labelling
facilitates the quantification of proteins during MS analysis and is especially useful when acquiring
data in the standard form of data dependent acquisition (DDA), as labelling allows for the mixing and
analysis of multiple samples at one. This alleviates the stochastic nature of DDA which can otherwise
lead to sampling biases and poor reproducibility. Label free techniques in the form of data
independent acquisition (DIA) help resolve the issue of sampling biases and eliminates any problems
encountered due to artifacts from labelling which can interfere with the identity of the protein,
leading to false conclusions (65, 66). DIA is addressed in more depth later in this article.

MS based proteomics techniques have developed substantially over the past several years.

Previously, protein analysis was reliant on 1D and 2D gel electrophoresis followed by sequencing;
finally resulting in the identification of tens to hundreds of proteins. This limitation has been
overcome with novel mass spectrometric approaches, which can identify many more proteins at
once. Identification and quantification of proteins/peptides by MS can be done either by ‘bottom-
up’ or ‘top-down’ approaches. The term ‘bottom-up’ refers to the reconstruction of the protein
sequence after proteolytic digestion, and thereby, identification and quantification of the peptide
fragments using mass spectrometry and appropriate databases (67). The second approach ‘top-
down’, refers to the identification of the whole proteins directly without enzymatic digestion (68).

To date, the most widely used proteomic technologies applied to the identification of the EC
biomarkers includes liquid chromatography coupled online/offline with MS (LC-MS). This approach is
generally used with ‘bottom-up’ methods. LC-MS has been used to identify and quantify differential
change in protein abundance in healthy versus disease state. Identification of proteins by LC-MS/MS
can be broken down into a few essential components. First, complex protein samples are digested
with an enzyme (protease) into peptides, which are separated by a high resolution chromatographic
technique coupled directly to a MS. Following chromatographic separation samples are sprayed into
the MS whilst being ionised, a process called Electrospray ionisation (ESI). Here the intact peptide
mass is measured, followed by collision with an inert gas to induce fragmentation, upon which the
mass of the fragment ions are measured. The acquired data is then processed to gain peptide sequences and protein identifications, and examined using mathematical algorithms, functions to detect and identify differences within the sample and/or between the samples. In traditional label free approaches, peptide separation is compared across chromatographic profiles, and quantitation is determined either by spectral counting or by peak intensity during MS (69).

A recent study by Alconada et al. has identified and validated annexin (ANXA2) as a reliable biomarker of recurrent disease in EC by 2D DIGE followed by traditional LC-MS/MS. They demonstrated in vitro and with an in vivo mouse model that the increased expression of ANXA2 is associated with an improved ability of the cells to metastasise (70). Monge et al. analysed the role of the transcription factor ERM/ETV5 in myometrial invasion by 2D-DIGE to evaluate the differential expression of proteins in EC cell lines overexpressing ERM/ETV5 (71). The authors have demonstrated that ERM/ETV5 acts by involving matrix metalloproteinase-2 to provide the migratory and invasive capabilities associated with the switch to myometrial infiltration in Hec-1A EC cell line (71). They further characterized a role for ETV5 in a modulated response to oxidative stress associated with the promotion of invasion in EC (71). Ihata et al. analysed the amino acid concentrations in plasma samples by LC-MS and reported that the amino acid profile index (API) is a potential new modality that could eventually play a significant role in the preoperative evaluation of EC and its screening (72).

The matrix assisted laser desorption/ionisation (MALDI) technique has also been widely used for the discovery of gynaecological cancer biomarkers but in a more limited context for EC. Qiu et al. have reported that MALDI-Time of flight (TOF) MS is a high-throughput, sensitive, highly predictive and rapid method for the early detection, diagnosis and prognosis of EC and will be widely used in future clinical screening work (73). In this technique, molecules including proteins, peptides, lipid and metabolites, are first co-crystallized with large excess of suitable matrix (usually a weak organic acid) and then spotted on to a MALDI target plate. MALDI TOF MS is then performed, a process in
which the laser transfers the energy to the aromatic matrix molecules and the matrix transfer a
proton to the analyte, resulting in soft energy transfer without fragmentation. According to the mass
to charge (m/z) ratios, the ionised analytes are then separated in the time-of-flight mass analyser
and the detector transforms the incoming ions into an electric current which is proportional to their
abundance (74). In the case of protein analysis, the generated protein/peptide mass list can then be
compared to protein databases that allow for the matching of peptide masses and identification of
proteins. Casado-Vela et al. used an integrated approach combining three different but
complementary strategies (in-solution digest followed by reverse phase HPLC-MS/MS, protein
separation by SDS-PAGE followed by HPLC-MS/MS analysis, 2D-PAGE followed by MALDI-
TOF/TOF analysis) to describe comprehensive proteomic analysis of EC fluid aspirate which led to the
successful identification of 803 different proteins in the International Protein Index (IPI) human
database (v3.48) (75). Therefore, EC biomarkers can be used for screening, estimating risk of disease,
distinguishing benign from malignant, determining prognosis, and monitoring status of the disease,
either to detect recurrence or response to therapy. Importantly, some biomarkers are only used in a
specific setting, whereas others can serve more than one purpose (76).

5. Emerging techniques for the proteomic analysis of endometrial cancer

Traditional proteomic approaches as described above, continue to be the most widely applied
technologies for protein identification; nonetheless, a major drawback of these proteomics
techniques is that not every peptide from a complex biological sample can be detected. As a result,
the proteomes will be under sampled. New techniques are emerging that offer innovative solutions
to the analysis of low abundant proteins from complex biological samples (77). Those proteomics
technologies have the potential for direct clinical applications including the identification of novel
biomarkers for the monitoring of therapy response, and disease progression, delivering on the great
promise of personalized cancer medicine (78).

5.1. Data Independent Acquisition (DIA)
Protein identification by LC-MS is a straightforward technique, as the detection of only two unique peptides from a protein is considered sufficient. However, relative quantification by LC-MS is a more complex process and the failure to identify or detect a peptide does not necessarily confirm absence, as the peptides may simply be below the threshold of detection (79). Traditional LC-MS or shotgun approaches are based on Data dependent acquisition (DDA), where peptides are selected for isolation and fragmentation based on their relative abundance within the sample. This makes the technique stochastic and can result in under sampling of the injected analyte (80). Abundance measurements can be inaccurate as the quantification of a peptide is highly influenced by co-eluting peptides and their signal to noise ratios, meaning peptides of a higher abundance are quantified more accurately and more often than peptides of low abundance. We do however; know that many cancer-relevant proteins including signalling and regulatory proteins are typically expressed in low concentrations. As a result, the DDA approach tends to miss out on acquiring the most-valuable information (81). Data independent acquisition (DIA) aims to solve the problem of DDA by measuring all precursor ions within a defined mass range of ~25 Da, fragmenting and measuring all of the ions within that mass range (82). Typically a 25 Da mass window is stepped across the entire mass range covered in traditional LC-MS/MS (~150-2200 Da) in cycle times compatible with the coupled chromatographic separation (82). Vowinckel et al. shown the number of precisely quantifiable peptides can be increased up to 5 fold when using DIA compared to DDA, with a coefficient of variation (CV) less than 5% (83).

5.2. MALDI Imaging (MALDI-MSI)

Another innovative technology that can be used for both biomarker discovery and validation is MALDI imaging (MALDI-MSI), which facilitates the direct analysis of protein distribution and abundance in thin tissue sections (84). This technique has the potential to detect and characterize tumour margins and their environment in a spatial context (56). MALDI-MSI is a valuable method for the identification of biomarkers and can complement histology, immunohistochemistry (85) and
molecular pathology in various fields of histopathological diagnostics, especially with regard to identification and grading of tumors (86). Whereas, MALDI-MSI is optimally suited to detect a large number of molecular features in a given tissue sample, the tissue microarray (TMA) method enables expansion of this analysis to hundreds of tissue samples in a single experiment (87, 88). In MALDI-MSI, proteins are directly ionised from the surface of the tissue sample and a characteristic histological pattern is derived for hundreds or thousands of individual peptides simultaneously, which may be further investigated and identified in the same tissue section by in-situ MS/MS (89).

During MS/MS, a desired peptide is isolated based on its $m/z$ ratio and fragmented (i.e. breaking peptide bonds within the peptides), producing a series of fragment ions that are detected as a MS/MS spectrum. The fragmentation pattern is compared to the theoretical fragmentation pattern for every peptide in the proteome to find the closest match. In this way the sequence of the peptide ion is inferred from its fragmentation pattern (90). Realistically in the case of MALDI-MSI, this form of tandem mass spectrometry is limited to highly abundant ionisable peptides whereas the identifications of lower abundance peptides can be made by matching back to fragmentation spectra acquired using LC-MS/MS.

MALDI-MSI has been successfully used to classify HER2 receptor expression in breast cancer (91). Meding et al. have shown the power of MALDI-MSI in classification of metastasis for different tumour entities (92). In addition, their results indicated that even closely related entities such as primary tumour of the colon and its liver metastasis could be classified efficiently (92). Quaas et al. detected several molecular signals associated with phenotypic features of oesophageal cancer cells, highlighting the potential of MALDI-MSI to identify new molecular markers with relevance to oesophageal cancer (87). Another strength of MALDI-MSI is that it can be used in the discovery phase for the identification of tumor derived proteins and then the verification of the presence of those particular proteins can be carried out in body fluid (blood, urine or saliva), aiming to develop non-invasive diagnostic tests. Addona et al. developed a pipeline integrating proteomic technologies from the discovery to the verification stages of plasma biomarker identification and applied it to
identify early biomarkers of cardiac injury (93). Thus, combining the histopathology reports with MALDI-MSI data will help surgeons to make right medical decisions.

In the similar approach, we performed MALDI imaging on EC formalin fixed paraffin embedded (FFPE) tissue samples collected from the Royal Adelaide Hospital (RAH), South Australia. Ethics approval was granted by RAH Human ethics committee.

In this study 6µm FFPE tissue sections were sliced and mounted onto indium tin oxide (ITO) coated conductive glass slides (Bruker Daltonics, Bremen, Germany) and samples were prepared as described by Gustafsson et al. (12). Briefly upon Citric acid antigen retrieval (CAAR), in-situ tryptic digestion was performed using an Image prep station (Bruker Daltonics, Bremen, Germany) at 37°C for 2 hours. α-Cyano-4-hydroxycinnamic acid (CHCA) matrix solution was prepared at 7mg/ml in 50% ACN/ 0.2% TFA and was overlayed onto the tissue section using an Image Prep station. Following matrix deposition, analyses was performed on an UltrafleXtreme MALDI-TOF/TOF (Bruker Daltonics, Bremen, Germany) in positive reflectron mode with FlexControl V3.0.1 and FlexImaging V4.0.1 (Bruker Daltonics) software package. The data was acquired at 100µm lateral resolution with a laser frequency of 2000 Hz over m/z 800–4500. Following MS acquisition, an overlapped MALDI spectral profile of tumour and healthy regions was generated by FlexAnalysis software (V4.0.1, Bruker Daltonics, Bremen, Germany) and as expected, a number of m/z values have been identified which are differentially expressed between healthy and tumour regions (Figure 2).

In order to determine data dependent visualization of tissue morphological regions, the raw data was loaded into SCiLS lab software (V2014b, Bruker Daltonics, Bremen, Germany). Data was processed by baseline removal and total ion count (TIC) normalization. Healthy and tumour regions were then grouped based on annotations done by a pathologist and discriminatory m/z values were calculated by ROC (Receiver operating characteristic) curve. ROC curve compares sensitivity verses specificity across a range of m/z values by plotting a curve called as Area under the ROC curve (AUC). The AUC assumes values between 0 and 1 and expresses the discrimination power of the m/z signal.
A perfect discrimination would yield an AUC value equal to 1 (abundant in group 1) or 0 (abundant in group 2) (94). The AUC closer to 0.5 indicates that there is no discrimination between two sample groups. An example of a discriminatory \( m/z \) value (\( m/z \) 1111.553Da ± 0.250Da) is shown in Figure 3. Here the spatial intensity of \( m/z \) 1111.553Da ± 0.250Da is clearly higher in the healthy tissue regions as compared to the tumour tissue region, shown also in the ROC curve (Figure 4). The results from this study will provide crucial new tools to assist in the diagnosis and prognosis of EC, with the ultimate aim to prevent overtreatment of patients whose primary tumours do not have metastatic potential.

5.3. Targeted Approach

Sensitive detection of low abundant proteins in complex samples has been typically achieved by enzyme-linked immunosorbent assay (ELISA) which require antibodies specific to the protein(s) of interest. However, this poses a problem as development of antibodies is associated with high cost, long development lead times, and high failure rates (95). To overcome this problem, for the past few decades intensive research has been carried out aiming to develop more targeted strategies that have been designed specifically to analyse preselected peptides/proteins of interest within complex samples.

Unlike untargeted approach, which aims to detect all proteins within a complex sample in an unbiased manner, Multiple reaction monitoring (MRM) is a targeted approach, in which the protein(s) of interest are preselected for quantification and analysed using highly sensitive triple quadrupole MS (96). MRM-MS provides high sensitivity and accuracy needed in the discovery phase and offers high reproducibility necessary for clinical validation (97). Due to the high specificity, sensitivity and throughput of MRM, it can be used for the validation of a single protein of interest or a subset of proteins (98). To ensure good quantification by MRM, three of the most intense ion transitions (precursor ion fragmentation to product ions) per protein are monitored. DeSouza et al. shown the MRM of iTRAQ labelled peptides enables absolute quantification of pyruvate kinase in
Elizabeth et al. pioneered a new technique that couples the ability to perform MRM on a triple quadrupole mass spectrometer with a MALDI source (100). The authors have shown that this approach has the feasibility for the precise and accurate quantitation of tissue protein concentrations over 2 orders of magnitude, while maintaining the spatial localization information for the proteins (101). The greatest advantage of the MALDI MRM-based imaging technique is improved sensitivity and selectivity of the analysis, enabling measurement of the accurate protein concentration with the addition of stable isotope labelled internal standards. This technique combines the spatial information gained by traditional MALDI MSI with the accurate quantification achievable by MRM and may develop to be an optimal strategy for the analysis of abundant peptide biomarkers.

6. Conclusion

Above mentioned studies have discovered a number of candidate EC biomarkers, for example CA-125, CA 19-9, CA 15-3, HE4 for the diagnosis and L1CAM, COX-2, Survivin, c-erb B2 for the prognosis of EC. Unfortunately due to their lack in sensitivity and specificity none of those protein biomarkers are currently used in clinical practice. Recently, the major focus of biomarker research has shifted from single biomarker discovery to the multiparametric analysis of proteins, as the proteomic pattern analysis can obtain exceptional results. As part of this review we have introduced two emerging techniques with exceptional sensitivity and dynamic range. These techniques are refining our capability to identify and quantify relative changes in protein expression and will ultimately result in the discovery of new EC biomarkers panels.

Moreover, the integration of the data obtained from various ‘omics’ approaches such as genomics, transcriptomics and proteomics will help to reduce the false positives and false negatives obtained from single ‘omic’ approaches and provide us with a number of potential protein targets. Taken together, this is an exciting time for EC biomarker research and we are hopeful that novel EC biomarker panels will perform well and will be used in the clinic in the future.
7. Conflict of interest statement

The authors have declared no conflict of interest.
8. References


Pertea M, Salzberg SL. Between a chicken and a grape: estimating the number of human genes. Genome biology. 2010;11(5):206. Epub 2010/05/06.


Ong S-E. Whole proteomes as internal standards in quantitative proteomics. Genome Medicine. 2010;2(7):49-.


Christoph W. Sensen BH. Advanced Imaging in Biology and Medicine.

**Figure 1:** Proteomic approaches used in the EC biomarker discovery: Gel based and Gel free. In Gel based proteomic approaches, proteins are separated on the basis of their charge/pH or mass by isoelectric focussing and electrophoresis. Gel free techniques include MS based technologies.

**Figure 2:** In this analysis *in-situ* tryptic digestion was performed on 6µm thick FFPE tissue. The matrix used was α-cyano-4-hydroxycinnamic acid and data was collected in positive reflectron mode using an UltrafleXtreme MALDI-TOF/TOF (Bruker Daltonics, Bremen, Germany). An overlapped MALDI spectral profiles of tumour and healthy regions was generated by FlexAnalysis software (V4.0.1, Bruker Daltonics, Bremen, Germany). (a) The overlapped spectral profile show a multitude of differentially expressed $m/z$ species between healthy and tumor regions simultaneously that allows a straightforward correlation of the expression pattern within the tissue morphology (b) Comparative representative spectra of $m/z$ 1111.553Da ± 0.250Da between healthy and tumor regions”

**Figure 3:** MALDI-MSI of a EC FFPE sample. The expression pattern of $m/z$ 1111.553Da ± 0.250Da is visualized between healthy and tumour regions (B), when comparison is made with corresponding haematoxylin and eosin (H & E) stained annotated image (A). Visualization of difference in the intensities was performed in SCiLS lab software (V2014b, Bruker Daltonics, Bremen, Germany) with edge preserving image denoising and automatic hotspot removal applied. This $m/z$ species is seen to be specifically downregulated in tumour region as comapre to healthy region.

**Figure 4:** The ROC curve of $m/z$ 1111.553Da ± 0.250Da (AUC = 0.985) for healthy versus tumour region
Table 1: Classification of carcinoma of the endometrium, International Federation of Gynaecology and Obstetrics (FIGO 2010), Table modified from Creasman et al. (102)

<table>
<thead>
<tr>
<th>FIGO stage</th>
<th>Tumour localisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>Tumour confined to the corpus uteri</td>
</tr>
<tr>
<td>IA</td>
<td>No or less than half myometrial invasion</td>
</tr>
<tr>
<td>IB</td>
<td>Invasion equal or more than half of the myometrium</td>
</tr>
<tr>
<td>Stage II</td>
<td>Tumour invades cervical stroma, but does not extend beyond the uterus</td>
</tr>
<tr>
<td>Stage III</td>
<td>Local and/or regional spread of the tumour</td>
</tr>
<tr>
<td>IIIA</td>
<td>Tumour invades the serosa of the corpus uteri and/or adnexa</td>
</tr>
<tr>
<td>IIIB</td>
<td>Vaginal and/or parametrial involvement</td>
</tr>
<tr>
<td>IIIC</td>
<td>Metastasis to pelvic and/or para-aortic lymph nodes</td>
</tr>
<tr>
<td>IIIC1</td>
<td>Positive pelvic nodes</td>
</tr>
<tr>
<td>IIIC2</td>
<td>Positive para-aortic lymph nodes with or without positive pelvic lymph nodes</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Tumour invades bladder and/or bowel mucosa, and/or distant metastasis</td>
</tr>
<tr>
<td>IVA</td>
<td>Tumour invades bladder and/or bowel mucosa</td>
</tr>
<tr>
<td>IVB</td>
<td>Distant metastasis, including intra-abdominal metastasis and/or inguinal lymph nodes</td>
</tr>
</tbody>
</table>
Proteomics

Gel Based proteomics
- 1D-PAGE
- 2D-PAGE
- 2D-DIGE (For quantification)
  - Image Analysis
  - Differential expressed protein quantification
    - MS, MS/MS
  - Protein identification

Gel Free proteomics
- Multidimensional HPLC
- Tissue section
- Tissue array
  - LC-MS/MS
  - MALDI-TOF
  - Labelled LC-MS (MS/MS)
    - ICAT, iTRAQ, SiLAC
  - Differential expressed peptide quantification
  - Label-free LC-MS (MS/MS)

Differential expressed protein quantification
Differential expressed peptide quantification

Figure 1
Figure 2b
AUC = 0.985, 1111.553 Da ± 0.250 Da

True positive rate (sensitivity)
False positive rate (1-specificity)

Figure 4