Identifying biomarkers for non-invasive diagnosis of endometriosis

Stella Irungu
Department of Women's Cancer
Institute for Women's Health
University College London

PhD in Biomarker Discovery and Proteomics

This thesis is submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy from University College London

I, Stella Irungu confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated.

Abstract

Endometriosis is a gynaecological disorder occurring when endometrial cells are shed through the fallopian tubes and implant on surfaces in the abdomen and pelvis. There they form lesions that respond to hormones of the cycle and stimulate inflammation. Women with endometriosis experience painful debilitating periods, pain on intercourse and defecation, and may have difficulties conceiving. It is a common disorder, affecting 5-10% of women of reproductive age. Diagnosis of endometriosis is difficult and is often delayed by 5-11 years. Symptoms do not correlate with disease severity and imaging techniques are only sensitive for diagnosing ovarian endometriomas. Definitive diagnosis is surgical, requiring laparoscopy under general anaesthetic, exposing patients to potentially serious complications. With these facts in mind, the aim of this project was to identify biomarkers for the non-invasive diagnosis of endometriosis. This was achieved by defining the protein expression profiles of tissue samples collected from women diagnosed with endometriosis and from control patients who underwent surgery for investigation of chronic pelvic pain or who underwent prophylactic surgery because of familial cancer history. Discovery work involved the use of complementary, quantitative proteomic profiling by 2D difference gel electrophoresis and multiplex mass tagging linked to liquid chromatography-based separation and tandem mass spectrometry. Selected candidate biomarkers (LUM, CPM, TNC, TPM2 and PAEP) were verified using ELISA in serum samples collected from the same women. Biomarkers reported in the literature were also tested. Diagnostic performance of each marker was established. The best single marker in discriminating endometriosis and controls was CA125 (AUC=0.724, *P*=0.002). Multi-marker models were also constructed and the best model in discriminating between endometriosis and healthy controls by cross-validation was CA125, ICAM (AUC=0.744). CA125, ICAM, FST model (AUC=0.75) gave the performance in discriminating between endometriosis and both controls by cross-validation.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Dr John Timms for his continuous patience, guidance and support throughout my PhD research and writing of this thesis. I would also like to thank my secondary supervisor Mr Ertan Saridogan and Dr Dimitri Mavrelos for sharing their knowledge and expertise in endometriosis research.

I am immensely grateful to Dr Jenny Worthington for her valuable practical assistance in mass spectrometry and for her patience and encouragement for the most part of my research. To all the staff of the UCLH Reproductive Medicine Unit, thank you for the assistance during patient recruitment and sample collection. To my colleagues at the proteomics lab (Joy, Harry, Oleg, Richard and Tendayi), thank you for your friendship and support.

I would also like to express my special gratitude to Prof Sam Kariuki (KEMRI-CMR) who encouraged me to pursue a PhD. My sincere gratitude also goes to the Commonwealth Scholarship Commission, whose funding enabled me to undertake my research.

Last but not least, I would like to thank my parents Mwangi Irungu and Jane Irungu for their unconditional support and for being a constant source of encouragement and inspiration throughout my many years of studying. To my brother Dr Eric Irungu, thank you for the unwavering encouragement and support especially the late nights spent on skype editing parts of this thesis. To my sister Nana Irungu, thank you for being a constant source of love and laughter. Finally, I would like to express my sincere gratitude to my friends Zilper, Christine, Maria, Rehema, Caroline, Cynthia, Justin, Job and Brian for the emotional support as well as the fun and happy memories. A special thank you goes to Lizah, Debbie and Anne. Words alone are not enough to express how grateful I am for the kindness and support you have shown me.

TABLE OF CONTENTS

List of tables

List of figures

Abbreviations

ACN- Acetonitrile

AmBiC- Ammonium bicarbonate

AFS- American fertility society

CID- Collision-induced dissociation

CHAPS - 3-[(3-cholamidopropyl) dimethylamminio]-1-propanesulfonate

DTT- Dithiothreitol

ECM- Extracellular matrix

FA- Formic acid

GWAS- Genome wide association studies

HCD- Higher energy collisional dissociation

IAM- Iodoacetamide

LC-MS/MS- Liquid chromatography tandem mass spectrometry

TEAB -Triethylammonium bicarbonate

TCEP -Tris (2-carboxyethyl) phosphine

2D-DIGE- Two-dimension difference in-gel electrophoresis

TMT- Tandem Mass Tags

TFA- Trifluoroacetic acid

SAX- Strong anion exchange

SDS-PAGE- Sodium dodecyl polyacrylamide gel electrophoresis

ES- Eutopic endometrium secretory phase

CS- Control secretory phase

PS- Pain secretory phase

EcS- Ectopic endometrium secretory phase

CP- Control proliferative phase

EP- Eutopic endometrial tissue proliferative phase

CHAPTER ONE: INTRODUCTION

1.1 Endometriosis- A history of the disease

The origin of the disease we call endometriosis today is controversial and has been a subject of debate over the years. The history is also linked with the early history of adenomyosis which up until the mid-1920s was considered the same disease. The evidence of who discovered or identified the disease for the first time has been conflicting with issues arising from origin and early clinical descriptions of lesions and misinterpretation of histological information. The path that led to the identification of the disease is a complex one but it is important to recognize some of the researchers who contributed immensely to the creation of the overall picture of what constitutes the essential features of endometriosis.

1.1.1 Vincent Knapp

Vincent Knapp referred to endometriosis as 'the forgotten disease' and he believed that endometriosis existed 300 years ago (Knapp, 1999). In his publication; How old is endometriosis; Knapp attempts to document the existence of endometriosis as early as the 17th century. He examined various historical manuscripts, published theses and dissertations and several contemporary works describing endometriosis that were housed in the National Library of Medicine in Bethesda, Maryland. One document which he believed contained the first description of endometriosis was published in 1690 and eleven others were published between 1739 and 1797 (Ronald E Batt 2011). In his publication, (Knapp, 1999) Knapp emphasized the 1690 manuscript presented at the University of Jena entitled *Disputatio Inauguralis Medica de Ulceribus Uteri* written by the German physician Daniel Schroen (Schroen 1690), whom he believes was the first person to give a detailed description of the disease.

The manuscript describes what Schroen believed to be endometriosis as sores distributed throughout the peritoneum, bladder, intestines, broad ligaments and outside of the uterus and cervix. He also described these sores as inflammations and

reported that they had a tendancy to form adhesions which linked visceral areas together. He further goes on to report that these sores would constantly expand in size, are vasculated, pus filled, tumour like and susceptible to haemorrhage. He also categorically states that the disorder was very common among sexually maturing females. Knapp believed that this description may have been that of non-ovarian endometriosis. He also reports on the other eleven early manuscripts from around Europe where the investigators reported similar findings to those of Schroen. However, others have refuted (Brosens and Steeno, 2000, Benagiano and Brosens, 2011) that what Schroen described to be indicative of endometriosis or adenomyosis. They state that what was described were actually symptoms and lesions linked to syphilis, abortions, abdominal surgeries and uterine manipulations. They also argue that without microscopic/histological evidence, there was no way of establishing the presence of endometriosis i.e. presence of uterine tissue outside the uterus. The authors however based their views on just a few of the papers that were studied by Knapp and unfortunately Vincent Knapp died shortly after his 1999 publication, therefore questions regarding his findings remain unanswered.

1.1.2 Carl von Rokitansky

Carl von Rokitansky is believed by some, to have given the first pathologic description of endometriosis and adenomyosis (Hudelist et al., 2009). In his publication of 1860 (Rokitansky 1860), three forms of endometriosis are described: i) *sarcoma adenoids uterinum* which invades the muscular wall and *cystosarcoma adenoids uterinum* associated with myometrial hypertrophy, ii) *cystosarcoma adenoids uterinum polyposum* which invades the endometrial cavity forming a polyp and iii) *ovarian cystosarcoma* which invades the ovaries. Some authors (Benagiano and Brosens, 1991) however, have strongly disputed these descriptions as being those of endometriosis and are of the opinion that what Rokitansky described were lesions of a more malignant nature than benign.

1.1.3 Thomas Cullen

Thomas Cullen, a surgeon, is said to have described the morphological and clinical nature of endometriosis and adenomyosis for the first time. He referred to these conditions as uterine adenomyoma (Linda C Giudice, 2012). From his observation of 90 uteri with adenomyomas, Cullen gave a description on the presentation and location of these lesions. He described them as ectopic endometrial-like tissue found in the ovaries, myometrium, recto-vaginal septum, uterine ligaments, bowel and umbilicus.

1.1.3 John Sampson

John A Sampson was a gynaecologist who conducted numerous studies on peritoneal endometriosis and ovarian endometriomas. He performed surgery on women at the time of menstruation. During these surgeries, he found peritoneal lesions that were bleeding in a similar manner to that of the normal endometrium (Clement, 2001). From this observation, he concluded that the tissue within the peritoneum was of endometrial origin. In 1927 (Sampson, 1927b) , he suggested that the presence of endometrial cells outside of the uterus was due to backflow of menstrual fluid through the fallopian tubes and dissemination of this within the peritoneum. He was therefore the first one to provide a theory to explain the pathogenesis of endometriosis, commonly known as Sampson's theory of retrograde menstruation. The word 'endometriosis' was also created by Sampson in 1928 after he observed growing bits of uterine mucosa on the fallopian tubes of women who had undergone sterilization. He believed that the tissue was transplanted surgically, and concluded that if such tissue could grow in that manner, then it was possible to observe similar behaviour in other areas of the body under different circumstances other than after surgery, hence the term endometriosis.

1.2 Disease characterisation

Endometriosis is recognised as a commonly occurring disease. It is defined as the presence of endometrial glandular epithelial cells and stromal cells outside the uterus commonly within the peritoneal cavity. Common locations therefore include the ovaries, fallopian tubes, uterine ligaments, pouch of Douglas, recto-vaginal septum, bowels etc. These endometrial cells invade and proliferate forming endometrial implants which behave in a similar manner to the normal (eutopic) endometrium i.e. they proliferate, bleed and they remain hormonally responsive. These implants also have their own blood supply and are highly invasive (Meehan et al., 2010). The clinical symptoms of endometriosis range from severe menstrual and non-menstruual dysmenorrhoea (painful menstruation), dyspareunia (painful intercourse), dysuria (painful urination), dyschezia (painful bowel movement) to severe chronic pain and finally infertility. The introduction of laparoscopy in the sixties enabled clear description of the disease together with its presenting lesions e.g. red and black lesions, adhesions, ovarian endometriotic cysts etc. It has been reported that the disease exists as three different clinical entities namely peritoneal endometriosis, ovarian endometriosis and deep infiltrative endometriosis (Nisolle and Donnez, 1997).

1.2.1 Peritoneal endometriosis

Peritoneal endometriosis can be explained by the transplantation theory which assumes that menstrual cells are regurgitated via the fallopian tubes into the peritoneum. These cells may also spread via lymphatic or haematogenous routes. The viable cells will then implant onto the peritoneum, have their own blood supply and continue to grow. These implants go through several stages and exhibit different characteristics at each stage.

The early stage of peritoneal endometriosis is characterised by the presence/appearance of red lesions which are consistently located on the peritoneal surface (Donnez et al., 1996a). They penetrate the extracellular matrix and have been

reported as the most active stage of endometriosis (Spuijbroek et al., 1992). A similarity exists between these red lesions and the eutopic endometrium (Nisolle and Donnez, 1997, Donnez et al., 1996c). This is based on the assumption that red lesions are recent cells which have implanted on to the peritoneum as a result of retrograde menstruation. Angiogenesis is also evident on the stroma of recently implanted cells which is characterised by a large vascular network on the peritoneal surface. Gland proliferation of red lesions has been evaluated previously by immunostaining and has shown similarity with eutopic endometrium

Two laparoscopic studies provided detailed reports of these early lesions. One group reported that red lesions appear and disappear like mushrooms giving peritoneal endometriosis a changing appearance (Wiegerinck et al., 1993). Petichial and bleb like implants have been reported as the most common type of red lesion observed in adolescent girls (Hoshiai et al., 1993), while the early serous lesions are said to disappear by the age of 26. These changes in the red lesions suggest that the implantation of viable endometrial cells in endometriosis may not be an on-going process during the reproductive life, but is restricted to the early years after the onset of menarche which is a period characterised by anovulation and low progesterone in peritoneal fluid. This environment favours the implantation of viable endometrial cells that enter the peritoneal cavity (Brosens, 1994).

Red haemorrhagic lesions continue to grow over time, eventually evoking an inflammatory response. A scar is formed which encloses the endometrial implant and this type now becomes a black lesion due to the presence of intraluminal debris (Nisolle and Donnez, 1997, Brosens, 1994). Scar formation causes de-vascularisaton of the endometric foci. White plaques of old collagen eventually become remnants of the ectopic tissue (Linda C Giudice, 2012). This stage represents the latent stages of peritoneal endometriosis. Progression or regression of endometriosis is therefore not reflected by the increase or decrease of these peritoneal implants. This variability is as a result of cycle differentiation, cellular activity and shedding of these early peritoneal implants that emerge on the surface and disappear either by resorption or

heal by fibrosis and scar formation. Despite this, the most important question still remains why endometriosis progresses in some women and not in others.

1.2.2 Ovarian endometriosis

The genesis of ovarian endometriosis remains controversial to date. In Sampson's first paper of 1921 he describes intraoperative and pathologic findings in 23 cases in which he described the endometrial nature of the ovarian cysts he observed (Clement, 2001). He reported these cysts as perforating cysts which upon haemorrhage, the cysts would spill their contents into the peritoneal cavity. As a result of the spillage, which he described as resembling chocolate syrup, there is an irritation of the peritoneal cavity together with formation of adhesions. Sampson also postulated that the endometrial tissue lining the ovarian cysts developed as a result of epithelialisation of a pre-existent cyst by ovarian surface epithelium or by Mullerian remnants of the endometrial type located on the surface of the ovary.

Several authors have disputed Sampson's claim that adhesions develop as a result of spillage of the contents of perforating ovarian cysts (Brosens et al., 1994, Hoshiai et al., 1993, Kennedy et al., 1992). Brosens et al supported an earlier report (Hughesdon, 1957) suggesting that adhesions are not the consequence, but the cause of endometriomas. Through studying ovarian endometriomas *in situ*, Hughesdon concluded that most endometriomas (90%) are formed as a result of invagination of the ovarian cortex after accumulation of menstrual debris from bleeding endometrial implants located on the ovarian surface and adherent to the peritoneum. Brosens et al also observed endometriomas *in situ* by using a double optic endoscopy. They also concluded that the most active implants were found at the site of inversion. Effects of chocolate cysts spillage has not been tested in experimentally. However other workers have attempted to test this effect in the peritoneal fluid of mice (Kennedy et al., 1992). No adhesions were reported as a result of the chocolate cyst fluid being injected into the peritoneum, but the authors state that their study is not conclusive evidence that this phenomenon does not occur.

Coelomic metaplasia of invaginated epithelial inclusions is also believed to contribute to the development of ovarian endometriosis. This assumption is based on the hypothesis that metaplasia of the coelomic epithelium invaginates into the ovarian cortex (Nisolle and Donnez, 1997). The expression of cytokeratin has been shown to indicate a close relationship with mesodermal Mullerian origin and the absence of vimentin expression in epithelial cells can reveal either an absence of glandular function or characteristic of metaplastic origin (Donnez et al., 1996b).

1.2.3 Deep infiltrating endometriosis (DIE)

This disease entity is also referred to as deep nodular endometriosis, recto-vaginal endometriosis or adenomyosis of the rectovaginal septum (Linda C Giudice, 2012). Deep infiltrating endometriosis has been defined as the presence of endometrial glands and stroma infiltrating more than 5 mm under the peritoneal surface (Koninckx et al., 1994, Koninckx et al., 2012, Dai et al., 2012). The use of 5 mm was postulated because from a morphology point, lesions infiltrating >5mm are reported to be the most active, aggressive and very painful (pain ranges from severe dysmenorrhoea, severe dyspareunia, chronic pelvic pain and dyschezia), forming nodules and causing pelvic distortion (Koninckx et al., 1994). This contrasts to peritoneal endometriosis whereby lesions are superficial, they may be asymptomatic or cause few symptoms, and they eventually become inactive and disappear.

DIE has been described as a different disease from peritoneal endometriosis and endometrial ovarian cysts because deep lesions have a distinct histological appearance resembling adenomyosis i.e. endometrial glands and stroma in the myometrium. The lesions are nodular in nature and contain a large amoun of fibromuscular tissue with areas containing endometrial glands and stroma (Van Kaam et al., 2008). They also have a specific anatomic distribution i.e. they are commonly found in the recto-vaginal septum, cul-de-sac, pouch of Douglas and utero-sacral ligaments. Three categories of deep lesions have been reported. Type I lesions are conical shaped lesions that are said to occur as a result of infiltration. Type II are deep lesions covered by extensive adhesions probably formed by retraction of the bowel over the lesion and Type III are

24

the largest and most severe lesions formed of a round nodule (>1cm in diameter) located under the peritoneum in the recto-vaginal septum. This type is also referred to as *adenomyosis externae* (Koninckx et al., 1994, Koninckx et al., 2012).

Aetiology of DIE remains debatable. *In situ* development i.e. metaplasia of Mullerian remnants located in the recto-vaginal septum has been proposed by some authors as the reason for the development of deep endometriosis as opposed to implantation of regurgitated menstrual debris (Nisolle and Donnez, 1997). Other authors believe that this disease develops from infiltrating superficial peritoneal lesions in the pouch of Douglas (Chapron et al., 2003, Chapron et al., 2006). Various arguments against the assumption that DIE is associated with peritoneal lesions has been reported. Thomas Cullen, one of the early pioneers of endometriosis, described adenomyoma of the round ligament, recto-vaginal septum and umbilicus in his works of 1896-1908. He reported that these deep lesions occurred as a result of direct extension of lower uterine adenomyosis (Clement, 2001). Sampson (Sampson, 1927b) also described adenoma of endometrial type which invades the cervical and uterine tissue and unites the cervix and the rectum.

1.3 Pathogenesis

The current understanding on aetiology and pathophysiology of endometriosis remains elusive. However, different theories have been proposed.

1.3.1 Transplantation theory

This theory suggests that the endometrium is replaced from the uterus to another location inside the body. The routes of dissemination involved are: retrograde menstruation, lymphatic /haematogenous dissemination and iatrogenic dissemination.

1.3.2 Retrograde menstruation

Retrograde menstruation may be defined as the backflow of menstrual fluid through the fallopian tubes to ectopic sites most commonly within the organs of the peritoneal

cavity (Kyama et al., 2003, Sampson, 1927a). This theory is the most widely recognised theory of endometriosis. There are three essential conditions that must be met in order to consider retrograde menstruation as the explanation for endometriosis (Nisolle and Donnez, 1997); endometrial cells must enter the peritoneal cavity through the fallopian tubes, the endometrial cells within the menstrual debris must be viable and able to be transplanted onto pelvic structures and the anatomic distribution of endometriosis in the pelvic cavity must be correlated with the principles of transplantation of exfoliated cells.

Reflux of endometrial cells into the peritoneal cavity during menstruation is a common physiological condition occurring in women with healthy tubes. Sampson observed menstrual blood escaping through the fimbrial ends of the fallopian tubes and endometrial tissue in the lumen of some women at the time of laparotomy scheduled during or soon after the menstrual period (Sampson, 1927b). Since then several studies have demonstrated this phenomenon. Blumenkrantz et al observed blood-stained peritoneal fluid during menses in women undergoing chronic peritoneal dialysis (Blumenkrantz et al., 1981). In these women, blood staining of the peritoneal fluid preceded vaginal bleeding for one to several days and the presence of blood was detected by observation of threads of sedimented red blood cells. Halme and colleagues found a red colour in 90% of peritoneal fluid samples of women with patent tubes at laparoscopy, suggesting the presence of blood (Halme et al., 1984) . However, only visual documentation of the colour of peritoneal fluid samples was carried out.

One study demonstrated the presence of endometrial cells in peritoneal fluid using immunohistochemistry based on application of monoclonal antibodies against various epithelial markers (Van Der Linden et al., 1995b). They compared the immunohistochemical staining properties of these fragments with those of cells present in the endometrium, menstrual effluent, peritoneal fluid and endometriotic lesions. Their study showed that peritoneal fluid contains single epithelial cells rather than endometrial tissue fragments in women with patent tubes. It is possible that endometrial epithelial cells after having left the uterine cavity are modulated in the peritoneal cavity prior to developing into an endometriotic lesion. Retrograde

menstruation has also been demonstrated in the baboon; a potential animal model for the study of endometriosis (D'hooghe et al., 1991, D'hooghe et al., 1996a, D'hooghe et al., 1996b). Prevalence, laparoscopic appearance and histology of endometriosis in the baboon resembles that of the human disease (D'hooghe et al., 1991). It was demonstrated that the prevalence of retrograde menstruation is higher in female baboons with endometriosis as compared to those without (D'hooghe et al., 1996b). This was the first study to document the prevalence of retrograde menstruation in non-human primates.

Endometrial cells must be viable and retain their capacity to adhere and proliferate. A number of adhesive molecules have been identified e.g. laminin and fibronectin (Beliard et al., 1997), integrins and cadherins (Van Der Linden et al., 1995a) that promote cell-cell and cell-matrix attachment of the endometrial cells to the peritoneal lining. The anatomic distribution of endometriosis in the pelvic cavity has been described (Jenkins et al., 1986). The authors reported that the distribution of endometriotic lesions in the abdominal cavity relates to that of tubal reflux, therefore the most frequent sites of implantation are the dependent areas of the pelvis e.g. ovaries, posterior and anterior cul-de-sac. These areas are most likely to be affected by menstrual reflux from the fallopian tube.

The theory of retrograde menstruation as the cause of endometriosis is challenged by the fact that it occurs in most women yet endometriosis only occurs in 6-10% of women in their reproductive years. It does not account for why these misplaced cells survive in women with endometriosis and not in non-diseased women. The puzzling enigma is why the pelvic environment becomes receptive to allow implantation and proliferation of endometrial cells. In this regard, it seems probable that there needs to be abnormalities of function either within the eutopic endometrium of women predisposed to endometriosis and/or defects in immune surveillance mechanisms normally responsible for the recognition and removal of endometrial fragments which find their way into the peritoneal cavity. Endometrial cells recovered from the pelvic cavity at the end of menstruation are viable and capable of proliferating. They also express the aromatase enzyme allowing them to synthesize active local oestrogen and

have the capacity to secrete angiogenic and neurogenic molecules that encourage tissue growth and potential pain generating mechanisms within the developing lesions. Other factors must therefore be involved to allow retrogradely displaced endometrial tissue to implant and develop into endometriotic lesions.

It has been hypothesised that the quantity of endometrial fragments desquamated during menstruation and deposited into the peritoneal cavity is higher in women who develop endometriotic lesions. In support of this, there is a positive correlation between the numbers of menstrual cycles and the prevalence, cumulative incidence and progression of spontaneous endometriosis on the surface of the ovaries in baboons (D'hooghe et al., 1996b). Similarly, women with short cycles and long durations of menstrual flow are more likely to develop endometriosis

1.3.3 Lymphatic/haematogenous spread (Halban's theory).

Hablan developed his theory from Sampson's original observations and hypothesised that endometrial tissue could be transferred into the myometrium or any other organ in the body through the lymphatic route (J, 1924). This theory therefore proposes that viable endometrial cells may spread from the uterus to distant sites through the lymphatic or haematogenous channels. This theory could explain the occurrence of endometriosis in unusual extra pelvic areas like the brain and lungs (Javert, 1949), but does not explain why ectopic tissue is able to become established and grow in these locations. Sampson also postulated that there may be more than one route available for the development and spread of this disease. He concluded that the invasion and dissemination of endometrial tissue employs the same channels as the invasion of cancer meaning that fragments of endometrial tissue reached other parts of the body through channels such as blood and lymph systems through metastasis (Clement, 2001).

In his paper entitled 'Metastatic or embolic endometriosis, due to menstrual dissemination of endometrial tissue into the venous circulation', Sampson describes his observations from four hysterectomy cases, three of which he performed during

menstruation and the other during the late secretory phase. He reports on the fragments of endometrium which he observed on the venous sinuses of the uterine wall, in the blood of the veins and some attached to the lining of the vessels by fibrin. These endometrial fragments were also confirmed by histology. In an earlier paper to investigate escape of foreign material into the venous circulation of the uterus, Sampson describes his observations after injecting barium into the endometrial cavity (Sampson, 1918). He reports that no barium is observed in the uterine veins when the endometrium is intact. When the endometrium is disrupted e.g. during menstruation, pregnancy or curettage the barium found its way into the uterine veins.

1.3.4 Iatrogenic transplantation

This theory assumes that endometrial cells may be transported to ectopic sites during surgery. This explanation accounts for the finding of localized endometriosis in old episiotomy sites or in caesarean section scars (Jensen and Coddington, 2010). It is proposed that biologically distinct tissue may directly attach to a site accompanied by initiation of oncogenic like cascades leading to implantation and survival (Bulun, 2009, Kao et al., 2003). The immune system should in normal circumstances be able to clear these implants, but it is not known why this is not the case. This phenomenon has also been described as auto-transplantation (Redwine, 2002) . An auto transplant is a tissue transplanted from one site to another site in the same organism either due to a pathologic process or surgical intervention.

1.3.5 *In situ* development theory

The transplantation theory cannot explain all localisations and manifestations of endometriosis (Burney and Giudice, 2012, Sourial et al., 2014) *In situ* development theory is based on the assumption that endometriosis develops *in situ* from local tissues. Development occurs from remnants of the Wolffian or Mullerian ducts or from metaplasia of peritoneal or ovarian tissue (Linda C Giudice, 2012, Nap et al., 2004).

1.3.6 Coelomic Metaplasia Theory

Embryological studies have shown that all organs in the pelvis including the endometrium originate as cells of the peritoneal (coelomic) cavity lining (Fujii, 1991). The term metaplasia refers to any type of tissue that can transform to another type. The coelomic metaplasia theory of endometriosis proposes that some of the cells of the peritoneum (abdominal wall) develop into endometrial cells instead of normal peritoneal cells or instead of the usual cells that make up the organs within the abdomen. Robert Meyer became one of the leading advocates for this theory. He believed that peritoneal inflammation stimulated the metaplastic transformation of the mesothelium to endometrial like tissue.

This theory would explain presence of endometriosis in the absence of menstruation and in unusual sites (Suginami, 1991). A case of endometriosis in a 20 year old patient with uterine agenesis (complete absence of the uterus, cervix and vagina) was reported (Mok-Lin et al., 2010). The patient underwent laparoscopy for debilitating pelvic pain and was found to have stage 1 endometriosis manifesting as red and clear lesions in the posterior cul-de-sac. A small band of flat Mullerian tissue was also found along the pelvic side wall. The endometriosis was treated but later recurred at age 25 when the patient presented again with pelvic pain. In this case, retrograde menstruation cannot explain the presence of endometriosis. Coelomic metaplasia seems like a plausible explanation for the development and recurrence of her endometriosis. Endometriosis has also been reported in a 83 year old man undergoing oestrogen treatment for prostate cancer (Martin and Hauck, 1985). The authors hypothesised that this developed from the prostate utricle which is a remnant of the uterus existing in the male.

The process by which cells differentiate into functional endometrium remains controversial. Steroid hormones and other exogenous molecules could be responsible for causing differentiation of normal mesothelial cells into endometriotic cells. One study showed that co-culture of ovarian surface epithelium and ovarian stromal cells together with 17β estradiol resulted in the ovarian surface epithelium cells forming a

lumen structure surrounded by endometrial stromal cells with an epithelial-mesenchymal structure (Matsuura et al., 1999). Immunostaining showed epithelial membrane antigen and cytokeratin in the glandular cells. This study shows that metaplasia of ovarian surface epithelium could give rise to endometriotic lesions and this process requires 17β estradiol at concentrations 10 times higher than that in the peritoneal fluid. Such a concentration is found near the ovaries and is hence a possible explanation for ovarian endometriosis.

This theory has several limitations; coelomic metaplasia should be manifested in all areas where the tissues are derived from the coelomic epithelium e.g. endometriosis should be present more often in chest wall since the the coelomic membrane contributes to cells of the thoracic cavity. If coelomic metaplasia resembles common metaplasia, the frequency of endometriosis should increase with advancing age. Peritoneal cells can easily undergo metaplastic transformation, the disease should therefore be observed more frequently in men. Additional basic and experimental data must be accumulated to confirm this theory.

1.3.7 Immune abnormalities

Endometrial lesions are frequently present in the peritoneal cavity where they are in direct contact with peritoneal fluid which bathes the pelvic cavity, uterus, fallopian tubes and ovaries. Retrograde menstruation occurs in most women, but relatively few develop endometriosis. As a result of apoptosis the body rids itself of these shed cells without eliciting a significant inflammatory reaction and it is not clear why this is not the case in women with endometriosis (Herington et al., 2011).

Several studies have shown that immune surveillance is impaired and that the innate immune system responds inadequately to displaced endometrium within the pelvic cavity of women with endometriosis (Ulukus and Arici, 2005, Herington et al., 2011, Tariverdian et al., 2007). The immune cells that play a role in the destruction of displaced cells include: macrophages, natural killer cells and cytotoxic T cells (Kaminski et al., 1995, Oosterlynck et al., 1992). A failure of these immune cells to

clear misplaced endometrium may play a central role in providing an opportunity for viable endometrial cells to attach and grow ectopically.

Activation of inflammatory responses within the peritoneal cavity may lead to local production of cytokines and chemokines that may enhance the growth of ectopic tissue by inhibiting normal apoptotic mechanisms and promoting localised angiogenesis. Inflammatory mediators are released upon onset of acute inflammation which is initiated by macrophages in the peritoneal fluid. These macrophages are responsible for extravasation of leucocytes from blood vessels into the tissue in an effort to clear ectopic endometrium. Despite this, these immune cells produce cytokines, growth factors and potent angiogenic factors which appear to support the survival and growth of ectopic tissue (Lebovic et al., 2001)

1.4 Epidemiology

Endometriosis is one of the most common benign gynaecological conditions. It is a condition in which tissue with histological structure and physiological responses of the uterine mucosa occurs in sites other than the uterus, most commonly within the pelvis (Uno et al., 2010). The physical findings are non-specific therefore clinical findings can be confused with those of pelvic inflammatory diseases, benign or malignant ovarian disease, fibroids, gastrointestinal and urinary problems.

Endometriosis is not widely and sufficiently recognised by the general public because it does not always provide a visible handicap despite its common crippling effects. Few well-conducted studies have reported data on the prevalence of endometriosis and no data is available on the incidence rates of the disease. However, from the few studies available, the disease is said to affect 6-10% of women of reproductive age in the general population. Its prevalence is 35-50% in women with chronic pelvic pain, infertility or both (Sensky and Liu, 1980, Uno et al., 2010, Giudice and Kao, 2004). It is also difficult to compare estimates of prevalence because the published studies include women with different conditions and are conducted in centres that apply different diagnostic criteria and exhibit different levels of clinical interest in endometriosis. A survey completed by 7,025 women with endometriosis reported that 65% of women were initially misdiagnosed with another condition and 46% had to see five doctors or more before they were correctly diagnosed (Mihalyi et al., 2010). (European Endometriosis Alliance). This has led to estimates of 6-11 years delay in diagnosis.

1.5 Risk factors

1.5.1 Genetic risk

Genetic and environmental factors contribute to endometriosis risk and the disease is inherited as a complex trait (Bischoff and Simpson, 2004, Kennedy, 1999, Stefansson et al., 2002). Gene markers associated with endometriosis have been reported but the results have generally not been replicated in subsequent studies. The variability between studies has also led to concerns about estimates of the genetic contribution to disease risk. Studies in Australian twins, Icelandic populations and rhesus macaques (Treloar et al., 1999, Stefansson et al., 2002, Zondervan et al., 2002) have provided strong evidence for disease heritability.

Large genome wide association studies (GWAS) have provided a powerful approach to discover genes influencing the risk of many common diseases. This method involves genotyping DNA samples with representative single nucleotide polymorphisms from across the genome and allele frequencies compared between cases and controls. A study in Japanese women (1,423 cases and 1,318 controls) reported a significant association with SNPs in the noncoding RNA CDKN2BAS on chromosome 9p21.3 (Uno et al., 2010). The International ENDOGENE study represents collaboration between research groups in the United Kingdom, USA and Australia whose main aim was to identify variants that influence susceptibility to endometriosis. This study which recruited 2,270 cases and 1,870 controls from Australia and 924 cases and 5,190 controls from the United Kingdom, identified a significant association in an intergenic region on 7p15.2. They also replicated evidence for an association near the WNT4 gene on 1p36.12, previously reported in the Japanese study (Uno et al., 2010). Data from this study provides independent evidence for a genetic contribution to disease risk, supporting results from earlier family-based studies (Stefansson et al., 2002, Treloar et al., 1999). These results also demonstrated stronger genetic loading of moderate to severe (stage III and IV) endometriosis compared to minimal to mild (stage I and II). The number of variants discovered in GWAS is strongly correlated with the experimental sample size

(Visscher et al., 2012). For each disease, doubling the study size doubles the number of genes/regions identified above a given threshold. It has been recommended that increasing the sample size for genetic studies in endometriosis will increase the number of markers and gene regions associated with disease risk (Rogers et al., 2013). Results also need to be replicated to confirm associations and to establish differences in gene expression to identify the specific genes and pathways contributing to disease risk.

1.5.2 Demographic risk factors

Endometriosis affects women of reproductive age. A study in women under 50 suggested that the frequency of endometriosis increases with age until menopause although other studies have not confirmed this (Houston et al., 1988). Age at diagnosis instead of age at onset is an assumption that due to the absence of fluctuating levels of oestrogen and progesterone, the condition is absent in girls before menarche. Endometriosis may be symptomatic after menopause if there is a resurgence of endogenous hormonal stimulation or if the use of exogenous hormones stimulates endometriosis (Ranney B, 1977). This suggests that sex steroid hormones play a role in the initiation and spread of endometriosis.

A greater frequency of endometriosis among women of higher social economic class has been reported (Arumugam and Templeton, 1990). This has been linked to their reproductive history as this groups of women tend to have fewer and later pregnancies than those from lower socio-economic groups (Williams and Pratt, 1977). It has been suggested that a lower incidence is in blacks is not because of race, but because of endemic poverty among a large portion of the black population (Chatman, 1976). However, poverty could create an apparent differential in disease occurrence across racial groups because of differential access to specialised medical care and the existence of a prejudicial diagnostic aspect based on what is 'known' about racial susceptibility to endometriosis. Differences in the incidence of endometriosis can also be attributed to the relative availability of laparoscopy for the diagnosis of pelvic conditions (Molgaard et al., 1985). This can account for the comparatively low

occurrence in certain racial and socio-economic groups, with the less affluent being constrained in their access to and utilisation of medical resources for the complete investigation of pain, infertility or both.

1.5.3 Menstrual factors

The correlation between menstrual history and risk of pelvic endometriosis is poorly understood. Some studies have suggested that women with early menarche, short and heavy menstrual cycles are at a higher risk of developing endometriosis (Cramer and Missmer, 2002, Parazzini et al., 1995, Cramer et al., 1986). This has been attributed to the fact that the potential for contamination of the pelvic cavity by refluxed menstrual debris is higher in these women hence a high risk of implantation of viable endometrial cells. It has however been noted that evidence for this is very inconsistent with interpretation of some of these findings being biased (Mangtani and Booth, 1993).

1.5.4 Oral contraceptive use

Oral contraceptives are used in the treatment of endometriosis. They do not cure the disease but instead work by down-regulating the endometrium, hence supressing menstruation, reducing pain and shrinking the ectopic implants or preventing the existing ones from growing any bigger. Information regarding the relation between oral contraceptive (OC) use and endometriosis risk remains controversial (Cramer and Missmer, 2002). A study to establish prevalence and associated risk factors in 504 women reported a lower risk among OC users (OR 0.5, 95% CI 0.2-0.90) (Sangi-Haghpeykar and Poindexter, 1995). The same has been reported in another large study where the rate of disease was lower in current/recent users as compared to those who have never used (relative risk 0.4, 95% CI 0.2-0.7) (Vessey et al., 1993). It was also reported in the same study that women who had stopped the pill for more than 2 years were at a higher risk (RR 1.8, 9.5% CI 1.0-3.1).

1.5.5 Family History

A familial risk for developing endometriosis has been proposed (Lamb et al., 1986, Moen and Magnus, 1993, Moen and Schei, 1997, Malinak et al., 1980, Simpson et al., 1980). For example high prevalence in relatives of affected women has been reported (Moen and Magnus, 1993). A wealth of genetic factors have been anticipated to be linked to increased susceptibility to endometriosis (Guidace and Kao, 2004) but common genetic markers for endometriosis have not yet been identified. Further studies however need to be conducted in order to establish the role heredity plays as a risk factor for endometriosis.

1.5.6 Environmental factors

Exposure to some environmental toxins are thought to be associated with endometriosis e.g. dioxin (2, 3, 7, 8 tetrachlorodibenzo-p-dioxin) (Rier et al., 1993). It is thought that dioxin promotes the development of endometriosis by interfering with immune-mediated mechanisms, more specifically by stimulating the production of pro-inflammatory cytokines.

1.5.7 Diet and other lifestyle factors

Alcohol intake and diet high in fat have been suggested as risk factors of endometriosis (Houston, 1984, Vessey et al., 1993, Cramer et al., 1986). An association between tobacco smoking and endometriosis has been reported in some studies (Aban et al., 2007, Chapron et al., 2010, Vidal et al., 2006). Smoke compounds lead to the disruption of oestradiol synthesis possibly reducing the risk (Vidal et al., 2006). However data is limited and further studies need to be conducted in order to establish whether diet, nutrition, alcohol and smoking could be risk factors for endometriosis.

1.6 Diagnosis

1.6.1 Surgical diagnosis

When endometriosis was first described, radical surgery was the treatment of choice. Laparotomy combined with histological examination was used to diagnose endometriosis (Linda C Giudice, 2012). Laparotomy is a highly invasive surgical procedure which involves making a large incision through the abdominal wall in order to gain access into the abdominal cavity. Later endoscopic techniques were developed which provided a less invasive way of inspecting the abdominal cavity whilst increasing understanding of the disease and improving treatment. The culdescope was the first endoscopic instrument and was used in a procedure called culdescopy. The procedure involved introducing the culdescope into the abdominal cavity via a puncture in the posterior vaginal wall. The advantage of this method is that there are no abdominal cuts or visible scars after healing. However the Inability to check the whole pelvic area and the abdominal cavity is a major disadvantage.

1.6.2 Laparoscopy

It is difficult to credit one individual with the introduction of laparoscopy. The first laparoscopic procedure is reported to have been performed in 1901 in dogs by George Kelling, a surgeon from Germany who dedicated himself to developing less invasive and non-surgical methods of diagnosis and treatment (Spaner and Warnock, 1997, Litynski, 1997). The first laparoscopy procedure in humans was performed by Hans Christian Jacobaeus on 17 patients with ascites (Litynski, 1997, Hatzinger et al., 2006). He recognised the great diagnostic and therapeutic potential of laparoscopic surgery as well as its risks and limitations e.g. the risk of damaging organs during surgery especially the bowel.

Laparoscopy is the gold standard method for the diagnosis and treatment of endometriosis (Kennedy, 2006, Kennedy et al., 2005). It is combined with the histological confirmation of the presence of endometrial glands and stroma in ectopic

sites. Laparoscopy has enabled thorough description of different types of ectopic lesions, disease classification and effective treatment of disease. In 1979, the American Fertility Society (AFS) proposed a classification of endometriosis so as to provide a standardised way of surgical recording and reporting of the disease (1979). Endometriosis was therefore classified into: stage I (minimal endometriosis), stage II (mild endometriosis), stage III (moderate endometriosis) and stage IV (severe endometriosis). Despite this it is still an invasive procedure with risks of organ damage and iatrogenic transfer of ectopic cells.

This classification, mainly based on severity of and the presence or absence of adhesions was later revised (rAFS) to allow recording of additional pathology e.g. adhesions were quantified and reported as either filmy or dense. In addition to this, it was now possible, with the rAFS classification, to predict pregnancy outcomes following treatment (1997). Some of the limitations of the rAFS classification include high observer error due to the disease manifesting in different ways and the weak relation between stage of disease and pelvic pain. There is also a weak relation between disease stage and infertility; pregnancy outcome rates do not vary with stage, therefore the classification is still not able to predict successful pregnancies following treatment with laparoscopy (Adamson et al., 1993, Adamson, 1990). A better predictive score for pregnancy outcomes after endometriosis surgery was proposed by (Adamson and Pasta, 2010). The endometriosis fertility index (EFI) is used to predict fertility after laparoscopy by providing detailed scores on ovarian and tubal functions combined with conception-related factors such as age, duration of fertility and parity history.

1.6.3 Imaging techniques

Several imaging techniques have been used to diagnose endometriosis prior to laparoscopy although they are not sensitive enough to predict all forms of the disease. Trans-vaginal ultrasound (TVU) is a readily available and inexpensive procedure adequate to detect large ovarian endometriomas but is not capable of ruling out peritoneal endometriosis, related adhesions and some deep lesions (Kennedy et al., 2005, Moore et al., 2002). Accurate staging of the disease is also not possible. It is vital that a highly trained sonographer conducts this procedure in order to achieve accurate results (Linda C Giudice, 2012, Abrao et al., 2009, Abrao et al., 2007).

Magnetic resonance imaging (MRI) has been reported as an important tool in the pre-operative diagnosis of patients with endometriosis, especially DIE and endometriosis affecting other areas away from the pelvis. This procedure has the advantage of being able to visualise the full spectrum of organ involvement. This compares to CT (computed tomography), whose resolution is not able to visualise pelvic organs well, hence its limited use in diagnosis of endometriosis (Hsu et al., 2010). Limitations of MRI include the lack of adequate resolution to identify adhesions and superficial peritoneal implants. The procedure is also very expensive compared to TVU, although it is the superior imaging method. Despite the availability of these imaging techniques, it is still not possible to detect early superficial peritoneal disease, which is the most active form of endometriosis. Early non-invasive or minimally invasive tests for the early diagnosis of endometriosis could help solve this problem.

1.7 Treatment of endometriosis

Current treatments for endometriosis include both surgical and medical therapy. Medical treatments aim at either inducing a low oestrogen environment or antagonising oestrogen action. NSAIDs and Oral contraceptives are often used as an initial approach even without definitive diagnosis. Surgical therapy can be performed alongside diagnostic surgery involving ablation which is the destruction of endometriotic tissue by electro-cautery and laser, division of scar tissue (adhesions) and removal of endometriotic cysts. In advanced endometriosis, laparoscopic surgery is done to excise visible endometrial implants, divide adhesions or surgically interrupt neural pathways for treatment of pelvic pain and is the only option for larger (>3 cm) respond poorly to medical therapy and hormonal supression doesn't influence the extent of the adhesions which are often associated with large lesions. In the case where these traditional methods of surgical treatment are ineffective or if the disease has progressed beyond the ability or desire to maiantain future fertility a hysterectomy may be performed. A review of surgical excision for endometriosis found that excision should be considered the gold standard treatment to endometriosis because it is better at controlling pain than traditional management (Garry, 2004). There is however limited data demonstrating the best method of treatment in patients with endometriosis who desire fertility.

1.7.1 Medical therapies for endometriosis

Endometriosis is an oestrogen-dependent disorder therefore current medical therapies are centered upon decreasing circulating oestrogens levels. This is usually achieved by down regulating ovarian production of steroid hormones mainly oestradiol (Linda C Giudice, 2012). Therapies for systemic hormonal supression of endometriosis include danazol; a synthetic testosterone hormone derivative, gonardotrophin releasing hormone (GnRH) analogues, progestogens, gestrinone and oral contraceptive pills. These therapies may be effective for relief of pain associated with endometriosis, but they also reduce fertility. Both danazol and GnRH analogues are associated with side effects related to hyperandrogenism e.g. hair growth and voice

deepening and hypoestrogenism e.g. hot flushes and vaginitis. The therapies are used for periods of up to 6 months, however recurrence of symptoms is common after cessation of medical therapy.

1.7.1.1 Danazol

Danazol is a 17α ethinyltestosterone that increases androgens and inhibits pituitary response by inhibiting secretion of FSH and LH .This medication induces a hyper-androgenic state which in turn causes suppresson of ovarian function, amenorrhoea and a low oestrogenic environment causing atrophy of endometriotic lesions and reduction in pelvic pain (Selak et al., 2007). In a study reporting on the effects of a danazol impregnated IUD for the management of pelvic pain in 18 women diagnosed with endometriosis and with recurrent pelvic pain, showed that pelvic pain was reduced significantly after the first 1 month with subsequent reduction in following 6 months after IUD removal (Cobellis et al., 2004). A few non-randomised trials that have demonstrated a significant decrease in pain from endometriosis after 6 months of danazol impregnated IUD therapy (Ozawa et al., 2006). This form of treatment however has substantial androgenic side effects that have been repored like hirsutism, mood changes, non-reversible voice deepening, arterial damage and liver damage (Selak et al., 2007)

1.7.1.2 Non-steroidal anti-inflammatory drugs

NSAIDs are often the first line therapy for treatment of symptoms of suspected endometriosis such as dysmennorohoea and pelvic pain. They are inexpensive, have few side effects and are available without prescription. Their mechanism of action is by inhibiting prostaglandins that are released in the pelvis and are the suscpeted cause of pain associated with endometriosis (Yap et al., 2004). NSAIDs inhibit the cycloxygenase (COX2) in the arachodonic acid pathway. The inhibition of COX2 decreases the production of prostaglandins, thus reducing (Marjoribanks et al., 2003). The most common side effects are gastro-intestinal upset with reflux disease, nausea, diarrhoea and headaches (Allen et al., 2005). A Cochrane review of NSAIDs for pain

in women with endometriosis included two randomised controlled studies comparing NSAIDs with placebo. The evidence was insufficient to show that NSAIDs are effective in treatment of pain caused by endometriosis (Allen et al., 2009).

1.7.1.3 Progestins

Progestins also act by creating a low oestrogenic environment by suppressing ovarian release of steroid hormones and thus preventing growth of endometriosis implants. The exact mechanisms by which progestins decrease endometriosis associated pain is not well understood however some postulately mechanisms include; suppression of ovulation resulting in marked endometrial decidualisation of both the eutopic and ectopic endometrium subsequently causing a decrease in active bleeding from the lesion, modulation of immune responses by suppression of interleukin 8 production in lymphocytes, reduction of TNF-α induced nuclear factor κ-B (NFκB) activaton which prevents proliferation of endometriotic stromal cells (Horie et al., 2005). Evidence for progesterone resistance in the endometrium of women with endometriosis may explain also why progestins are effective in reducing pelvic pain but ineffective in improving pregnancy rates (Bulun, 2009, Aghajanova et al., 2010). Side effects experienced by these drugs include; irregular uterine bleeding, weight gain, mood changes, bloating, fatigue, depression and nausea. Further studies aimed at understanding the precise effects of progestins on nociceptive, inflammatory and neuropathic pain in endometriosis are required to further knowledge of pathophysiology and improve the treatment of endometriosis. The combined OCP is also prescribed to women with endometriosis as a first in-line therapy and work by inducing a pseudo-pregnancy state by suppressing ovulation and causing atrophy of endometrial implants.

1.7.1.4 GnRH analogues

GnRH agonists are analogs of the hypothalamic hormone gonadotrophin releasing hormone (GnRH). GnRH is responsible for the normal function of the ovaries by stimulating the release of follicle stimulating hormone (FSH) and luteinizing hormone

(LH) from the pitutary. GnRH agonists bind to pituitary receptors resulting in a downregulation of pituitary hormone secretion which in turn down regulates the ovarian production of oestrogen. Depriving the endometriotic cells of oestrogen is thought to result in endometriotic atrophy (Kokorine et al., 1997). However, endometriotic cells are known to express aromatase ensuring their survival independent of ovarian steroids. GnRH agonist therapy is associated with side effects normally presenting during menopause e.g. hot flushes, vaginal dryness, loss of bone density (Sagsveen et al., 2003) therefore in some cases small amounts of steroid hormone can be administered indefinately in what is known as 'add back therapy' which appears to stem the severity of these side effects without significantly affecting the relief of endometriosis-associated pain.

1.7.1.5 Aromatase inhibitors

Aromatase is the key enzyme in the synthesis of oestrogens and mediates the conversion of androstenedione and testosterone to estrone and oestradiol. Aromatase inhibitors were developed to act on sex steroid-dependent neoplasms e.g. breast cancer, by suppressing the *in situ* production of oestrogen. In endometriosis, poor response to hormonal therapy may be due to the resistant nature of endometriotic cells to progestin compounds due to the under-expression of progesterone receptors, but also the over-expression of oestrogenic receptors in endometriotic cells that renders them less sensitive to progestin therapies.This enzyme is therefore a good target for inhibition of oestradiol synthesis.

Aromatase inhibiting compounds are currently being considered for the treatment of endometriosis that is unresponsive to current therapy and have shown promising preliminary results (Attar and Bulun, 2006, Nothnick, 2011). The side effects associated with this treatment are milder compared to those associated with GnRH treatment. However, long term use carries a potential risk of osteoporosis and osteoponia. Clinical trials are yet to be carried out to ascertain whether aromatase inhibitors could have a significant role in medical management of endometriosis.

1.8 Biomarkers of endometriosis

Laparoscopy is the gold standard method used to definitively diagnose endometriosis. It involves visualisation of lesions in surgery followed by histological confirmation by a pathologist (Kennedy et al., 2005). This procedure is invasive in nature, could cause complications and is very expensive. (Chapron et al., 2002, Vercellini et al., 2009). There is often a delay in diagnosis of between 7-10 years on average. The delay is partly attributed to the fact that laparoscopy is a surgical procedure and is not the primary choice for diagnosis of women presenting with suspected disease. Variability of signs and symptoms does not make it simple to make a definitive diagnosis and symptoms are often mistaken to be those of other disorders such as irritable bowel syndrome. (Rogers et al., 2009a, Husby et al., 2003, Hadfield et al., 1996). Symptoms such as dysmenorrhoea, chronic pelvic pain and dyspareunia are commonly associated with endometriosis (Ballard et al., 2008, Falcone and Lebovic, 2011). However, the predictive value of these symptoms for the diagnosis of endometriosis is limited (Eskenazi et al., 2001, Chapron et al., 2005). Imaging techniques also are not able to diagnose all forms of this disease. Thus molecular biomarkers are urgently needed for the more accurate diagnosis of endometriosis. Before establishment of diagnosis by surgery, a less invasive test would be important for initial screening of possible cases of endometriosis.

A biomarker may be defined as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention (Jain, 2010). There is a great need for biomarkers that can function as alternatives for important clinical end points in endometriosis and for the development of a non-invasive diagnostic test.

To date there are no reliable markers for the diagnosis and prognosis of endometriosis. However, efforts continue in the search for suitable markers. Several systematic literature reviews have been published on potential markers that have been identified to date (Fassbender et al., 2013, May et al., 2010, May et al., 2011). Some reported markers for the non-invasive or semi-invasive diagnosis of endometriosis are

discussed below. In some cases, these markers have been tested as part of a panel and this has shown to improve diagnostic performance. However, most markers/marker panels require appropriate independent validation.

1.8.1 Serum/plasma markers

1.8.1.1 CA125

CA125/mucin 16 is a high molecular weight mucin and antigenic determinant expressed on the surface of coelomic epithelium, including epithelium of the endocervix, endometrium, fallopian tubes, pelvic peritoneum and placental tissue. It is also found in biological fluids such as human milk, amniotic and peritoneal fluids. CA125 was initially thought to be specific for ovarian malignancies and is currently used to aid in the diagnosis of ovarian cancer and to monitor recurrence following treatment. However, CA125 has been found to be raised in various benign conditions endometriosis, pelvic inflammatory disease, liver cirrhosis, benign ovarian cysts, tubo ovarian abscesses, and fibroids.

Many studies have been carried out to evaluate the diagnostic performance of CA125 for detecting endometriosis. Some studies have reported high specificity (80-95%) but with inadequate sensitivity (rarely above 50%) (Moretuzzo et al., 1988, Barbati et al., 1994). The diagnostic performance of CA125 in endometriosis is complicated by the fact that endometriosis has varying degrees of chronicity and CA125 is mostly elevated in advanced disease. CA125 is also not a marker specific to endometriosis and therefore an elevated level cannot differentiate endometriosis from other diseases with similar indications such as fibroids or pelvic inflammatory disease.

Association of an increased serum CA125 and the presence of severe endometriosis have been known for some time with the first reports being published in the mid-1980s (Pittaway, 1989, Giudice et al., 1986). One early study (Barbieri et al., 1986) demonstrated elevated serum CA125 in patients with advanced endometriosis.

Several teams have evaluated CA125 levels in serum and peritoneal fluid but have shown conflicting results. CA125 levels in peritoneal fluid are said to be higher than in serum, but no significant difference was observed in either serum or peritoneal fluid levels between women with and without endometriosis (Moen et al., 1991, Williams et al., 1988). By using a two-step immunoradiometric assay, Barbati et al found that the levels of CA125 in peritoneal fluid are a more sensitive indicator of disease than the levels in serum (Barbati et al., 1994). However, serum levels may provide a useful indicator on the extent of disease or response to therapy (Matalliotakis et al., 1994, Pittaway and Fayez, 1986).

In a meta-analysis, Mol et al evaluated the performance of CA125 in serum of patients with laparascopically confirmed endometriosis compared to healthy controls from multiple studies (Mol et al., 1998). They also evaluated its ability to differentiate between mild disease (stage I and II) and severe disease (stage III and IV). 74% of studies included in this meta-analysis used a CA125 cut-off level of 35 IU/mL, while others used lower and higher cut-offs. ROC curve summaries were calculated for the ability of a single CA125 measurement to predict any type of endometriosis. The diagnostic performance was low with a sensitivity of 28% at 90% specificity. Increasing sensitivity to 50% resulted in a corresponding specificity of 72%. CA125 was found to be a better predictor of advanced disease (stage III and IV) with a sensitivity of 47% and specificity of 89%. A later study in 775 women found a 78% NPV (negative predictive value) using <20 IU/mL for pre-operative CA125 levels and 92.9% PPV using >30 IU/mL (Kitawaki et al., 2005)

CA125 could be a useful marker of disease recurrence although this aspect has not been extensively studied. In a longitudinal study evaluating CA125 measurements in post-operative women with endometriosis over 12 months, the authors found that women whose CA125 levels dropped to below 16 IU/ml after surgery were more successful at achieving pregnancy within one year (Pittaway et al., 1995). Rising CA125 levels however could be of great value in post-operative management of women with infertility associated with endometriosis who may benefit from a more aggressive treatment regimen and assisted reproductive procedures. However serum

CA125 on its own does not have adequate sensitivity and specificity to be used to diagnose endometriosis especially in screening for peritoneal lesions that are missed by current imaging techniques.

1.8.1.2 CA19-9

CA19-9 (Carbohydrate antigen 19-9) is a marker associated with pancreatic and colorectal cancer. In gynaecology, serum CA19-9 levels can be elevated in patients with malignant and benign ovarian tumours (Ye C et al, 1994; (Harada et al., 2002) and has been shown in endometriosis (Panidis et al., 1988). The investigators reported raised serum CA19-9 in 5 out of 8 women with endometriosis. The levels were also found to drop significantly during treatment with danazol. There was no control group that was used in this study. A drop in CA19-9 following danazol treatment was also reported in a later study (Matalliotakis et al., 1998).

Several studies have evaluated the diagnostic properties of CA19-9 in comparison to CA125 in women with and without endometriosis (Harada et al., 2002). The investigators found the mean concentrations of CA19-9 were higher in 101 women with endometriosis compared to 22 controls. There was also a positive correlation with disease stage. By using a cut-off of 37 IU/mL, 34 out of 101 (33.7%) cases had raised CA19-9. Concentrations greater than 37 IU/mL were observed in women with advanced disease (stage III and IV). Another study reported reported higher serum CA19-9 concentrations in 101 endometriosis cases compared to 78 controls with levels also higher in advanced disease (Kurdoglu et al., 2009). Using the commonly accepted cut off of 37 IU/ml the sensitivity was 89% and specificity was 52%, values similar to those of CA125. These results contradict with those from previous studies which have reported high specificities and low sensitivities for CA19-9. However, consistent with previous reports CA19-9 could only predict severe stage disease thereby limiting this marker for broad application. Most studies have been limited by small subject numbers and an inability to diagnose early stage disease.

1.8.2 Cytokines

Cytokines are glycoproteins that play a role in controlling cell proliferation, immune cell activation, motility, cell adhesion and chemotaxis. They are secreted into the extra-cellular environment by leucocytes, macrophages and other inflammatory cells. It is widely known that retrograde menstruation occurs in most women but why some end up developing endometriosis is still unknown and it has been hypothesiszed that a change in the function of the immune cells in the peritoneal environment may be a major reason (Wu and Ho, 2003, Lebovic et al., 2001).

Macrophages are major secretors of cytokines and have been shown to be increased in the peritoneal cavity of women with endometriosis inducing a local inflammatory response (Sukhikh et al., 2004, Haney et al., 1981). Changes in cytokine concentrations reflects the immune system's response to initial disease. Therefore many studies have focused on finding alterations in markers of inflammation between women with and without endometriosis to define potential biomarkers but also aid in understanding the pathogenesis of the disease.

The most studied cytokines as biomarkers of endometriosis are IL-6 and TNF-α, but the results have been conflicting (May et al., 2011). IL-6 is a pro-inflammatory cytokine involved in the activation of T cells and differentiation of B cells. Alterations in IL-6 in endometriosis have been studied by several groups (Bedaiwy et al., 2002, Bedaiwy and Falcone, 2004, Somigliana et al., 2004, Martinez et al., 2007, Seeber et al., 2008a, Othman et al., 2008, Iwabe et al., 2002). In a multi-marker panel study, increased levels of IL-6 were demonstrated in women with early stage endometriosis (Bedaiwy et al., 2002). At a threshold of 2 pg/mL, IL-6 was able to discriminate between groups with a sensitivity of 90% and specificity of 47%. The other markers tested (IL-1β, IL-8, IL-12, IL-13 and TNF-α) were not discriminatory. The authors however noted that they were unable to obtain sufficient serum to measure all the cytokines in all subjects and their findings were based on the comparison of only 20 cases and 11 controls.

49

In a prospective study to compare the diagnostic performance of IL-6 and CA125, elevated levels of IL-6 were found in women with endometriosis, but only those with minimal-mild disease yielding a sensitivity and specificity of 75% and 83% respectively, at a threshold of 5.75 pg/mL (Martinez et al., 2007). CA125 was increased 3 fold in women with late stage disease. However, the combination of both markers did not offer any additional value. These data do suggest that IL-6 may be a good marker for early stage disease where ultrasound is not as helpful and because IL-6 levels were not affected by other pelvic pathologies e.g. myomas, benign ovarian cysts. A protein array system using cytokine-specific antibody coated beads was used to quantify cytokine concentrations in samples from 68 women with endometriosis and controls (Othman et al., 2008). IL-6, MCP1 and IFN-γ were found to be higher in women with endometriosis compared to controls. At a threshold of 1.9 pg/mL, IL-6 gave the best results in discriminating endometriosis from controls. Combining IL-6, MCP1 and IFN-γ did not provide additional discriminatory power over using IL-6 alone. Combination with CA125 was not reported. In another prospective cohort study of CA125 and IL-6 no significant improvements in diagnostic accuracy were reported for this combination. In another study (Somigliana et al., 2004) classification tree analysis was applied to determine the diagnostic ability of IL-6, TNF-α, MCP1, MIF, CA125 and leptin. Combination of some of the markers improved diagnostic performance; CA125, MCP1 and leptin was able to diagnose 51% of cases with 89% accuracy, while CA125, MCP1, leptin and MIF could diagnose 48% of cases with 93% accuracy (Seeber et al., 2008a). The discrepancies observed in these studies can be attributed to various factors including the type of assay used and the different study sets used. Whilst the same ELISA kit was used in two of these studies, different results were reported (Somigliana et al., 2004, Bedaiwy et al., 2002). In addition IL-6 evaluation methods are not properly standardised and IL-6 is likely to reflect non-specific inflammatory responses that differ from case to case.

TNF-α is a monocyte-derived cytokine with pro-inflammatory and pro-angiogenic roles. An increase in serum TNF-α concentrations in endometriosis has been reported by several studies (Matalliotakis et al., 1997, Pizzo et al., 2002, Darai et al., 2003,

Xavier et al., 2006), whilst others have failed to show any significant difference (Vercellini et al., 1993, Seeber et al., 2008a, Othman et al., 2008). In one of these studies, elevated TNF-α levels were decreased after treatment with danazol (Matalliotakis et al., 1997). Significantly higher levels of TNF-α were reported in women with endometriomas compared to those with benign cysts, but were not higher in those with malignant cysts (Darai et al., 2003). High levels of serum TNF-α were shown in endometriosis patients throughout the menstrual cycle compared to controls (Xavier et al., 2006). Another study evaluating TNF-α, TGF-β, IL-8 and MCP1 in serum and peritoneal fluid reported increased TNF-α levels in early stage disease, but a significant decrease as the disease progressed (Pizzo et al., 2002). This is contrary to another report whose findings showed no difference in TNF-α levels between women with endometriosis-associated infertility to those with idiopathic infertility (Bedaiwy et al., 2002). In summary, the diagnostic potential of TNF-α is as yet unclear and more studies are required to evaluate its worth as a marker either individually or as part of a panel of other markers.

1.8.3 Cell adhesion molecules

Soluble intracellular adhesion molecule 1 (sICAM1) can be secreted from the endometrium and endometriotic implants (Vigano et al., 2001, Vigano et al., 2000). Studies of this protein as a possible biomarker of endometriosis have been conflicting. In a cross-sectional study, sICAM1 was found to be increased in serum of women with advanced stage endometriosis (Wu et al., 1998). Another study showed no difference in endometriosis patients compared to controls although when sICAM1 was evaluated by stage of disease it was reported to be increased in women with early stage disease (De Placido et al., 1998). Similarly, sICAM1 was found to be elevated in early stage disease and rose further during treatment and was maintained for 3 months after treatment (Matalliotakis et al., 2001). Decreased levels in cases with stage III-IV disease have also been reported (Barrier and Sharpe-Timms, 2002). They concluded that it could be possible that the levels are high during early stage disease, but decrease as the disease progresses.

Other studies have shown marginal increases in sICAM (Daniel et al., 2000, Somigliana et al., 2002), although women with deep endometriosis had significantly raised sICAM levels compared to those with peritoneal disease and controls. A sensitivity and specificity of 19% and 97% respectively, were reported. The use of CA125 in detecting deep lesions in the same patients yielded a sensitivity and specificity of 14% and 92%, respectively. When both markers were combined, there was an improved sensitivity of 28% at 92% specificity. Thus sICAM1 may play a role in the pathophysiology of the disease, but appears to be an unreliable non-invasive marker when used alone.

In a relatively large study, 28 plasma biomarkers were evaluated for diagnosis in 232 plasma samples from women with endometriosis and 121 without endometriosis at laparoscopy in the menstrual phase (Vodolazkaia et al., 2012). Two models of four biomarkers each were developed. Model 1 consisted of annexin V, VEGF, CA125 and glycodelin/PAEP, while in model 2, glycodelin was replaced with sICAM1. When evaluated, models yielded sensitivities and specificities of 81-90% and 68-81%, respectively. Validation in an independent set yielded a sensitivity of 82% and a specificity of 63-75%. This study demonstrates the value of using a panel of markers instead of a single biomarker to improve diagnostic performance.

1.8.4 Angiogenic factors

Angiogenesis is the formation of new blood vessels from existing vessels. The establishment of a new blood supply is essential for the survival of the endometrium attached to the peritoneum and the maintenance of endometriosis. Macrophage derived cytokines like transforming growth factor beta (TGF-β) and vascular endothelial growth factor (VEGF) have been suggested to contribute to the development of endometriosis by promoting the neovascularisation of endometrial cells attached to the peritoneum (Harada et al., 2001). Increased angiogenesis is reported to be common around peritoneal explants and increased angiogenic markers has been observed in serum and peritoneal fluid of women with endometriosis

(Mclaren et al., 1996a, Mclaren et al., 1996b, Garcia Manero et al., 2009, Taylor et al., 2002)

VEGF is the most potent angiogenic factor which has been detected in high concentrations in serum and peritoneal fluid of women with moderate to severe endometriosis. It is also secreted in endometriotic lesions, possibly as a downstream consequence of pro-inflammatory cytokine IL-1β activity (Mclaren et al., 1996a, Mclaren et al., 1996b). However, the usefulness of VEGF as a biomarker of endometriosis remains unclear. Some studies have reported increased peripheral VEGF in women with endometriosis (Xavier et al., 2006, Vodolazkaia et al., 2010, Bourlev et al., 2010), whilst others reported no significant changes (Othman et al., 2008, Pupo-Nogueira et al., 2007).

Glycodelin/PAEP is a glycoprotein expressed in reproductive tissues. It has numerous glycosylation sites and isoforms. Glycodelin A is found in amniotic fluid (Riittinen et al., 1989), the glandular epithelium of the secretory endometrium (Meola et al., 2009) and is also secreted into serum and peritoneal fluid by endometriotic lesions (Koninckx et al., 1992, Kocbek et al., 2013). Glycodelin has immunosuppressive, angiogenic and contraceptive effects. There is therefore a likelihood that it plays a role in the development of endometriosis and endometriosis-associated infertility (Fassbender et al., 2013). Few studies have evaluated glycodelin as a possible serum biomarker of endometriosis. In one study, glycodelin A levels were reported to be 10-fold higher in peritoneal fluid than serum of endometriosis patients in both the proliferative and secretory phase of the menstrual cycle (Kocbek et al., 2013). Pain frequency and intensity was directly proportional to glycodelin A concentrations. In assessing the diagnostic capabilities for ovarian endometriosis, serum glycodelin A yielded a sensitivity of 82.1% and specificity of 78.4%, compared to 79.7% sensitivity and 77.5% specificity in peritoneal fluid. Glycodelin A has been evaluated as part of a multi-marker panel for the non-invasive diagnosis of endometriosis and has shown promising results (Vodolazkaia et al., 2012). It may therefore have a potential role as a biomarker for endometriosis and further studies need to be carried out to assess its possible significance.

1.8.5 Circulating antigens and auto-antibodies

Reports on serum endometrial antigens and auto-antibodies are few and inconsistent. The inconsistency is due to different study designs, use of different cell lines as antigenic cell lines and different assays used. Despite this, efforts have been made to test the value of anti-endometrial antibodies and antigens as potential biomarkers. Anti-endometrial IgM was reported in 27% of endometriosis patients, while anti-endometrial IgG was present in 33% of the patients (Gajbhiye et al., 2008). A different study by the same group demonstrated endometrial antigens using 1D and 2D western blotting using serum of patients with endometriosis compared to healthy controls (Gajbhiye et al., 2012). Five differentially detected spots were identified by mass spectrometry as tropomyosin 3 (TPM3), stomatin-like protein 2 (SLP2) and tropomodulin 3 (TMOD3), and their auto-antibodies were found to be elevated in women with endometriosis. One of the auto-antibodies (anti-TMOD3) gave a sensitivity of 78% and specificity of 96% in detecting early stage endometriosis. Larger studies need to be carried out to evaluate these auto-antibodies as biomarkers.

1.8.6 Endometrial tissue markers

1.8.6.1 Interleukin 1 (IL-1)

Endometrial tissue cytokine expression has also been assessed with the aim of identifying candidate biomarkers. IL-1 belongs to a group of cytokines that play a role in the regulation of immune and inflammatory responses to a wide sectrum of pathophysiological processes associated with host defence and inflammation. It is also involved in normal immunological and reproductive activities that occur in the human endometrium during a normal menstrual cycle (Lawson et al., 2007, Akoum et al., 2008, Akoum et al., 2007). Several studies point to a significant role for IL-1 in endometriosis. IL-1β levels have been shown to be increased in women with endometriosis during the secretory phase (Kyama et al., 2008). IL-1RII a decoy receptor known for its ability to inhibit IL-1 functions has also been shown to be significantly altered in women with endometriosis. A significant decrease of IL-1RII

mRNA was reported in the endometrium of women with early stage disease in the proliferative and secretory phase of the menstrual cycle (Lawson et al., 2008). Another study from the same group reported that increased proteolysis causes the down-regulation of IL-1RII in the endometrium of women with endometriosis (Bellehumeur et al., 2005). This research group has conducted many studies on IL-1RII expression in endometriosis (Akoum et al., 2008, Lawson et al., 2007, Bellehumeur et al., 2005, Kharfi and Akoum, 2001, Lawson et al., 2008, Kharfi et al., 2002, Akoum et al., 2001). All their studies report on a significant decrease in IL-1RII. This decrease is proposed to result in an insufficient ability of the endometrium and ectopic implants to down regulate IL-1 function that plays a role in the abnormal inflammatory processes reported in ectopic sites. Independent validation of IL-1RII as a biomarker is therefore warranted.

1.8.6.2 Aromatase

In the human endometrium, it is known that androgens cannot be converted to oestrogen due to a of lack of the enzyme aromatase (Bulun et al., 1993). This is contrary to several reports that indicate an abnormally high expression of aromatase in eutopic and ectopic tissue of women with endometriosis (Noble et al., 1996, Maia et al., 2009, Hudelist et al., 2007, Bulun et al., 2004). Endometriosis is not associated with increased oestrogen levels in serum, therefore aromatase expression in eutopic and ectopic tissue suggest a local oestrogen production in the endometrium and endometriotic tissue of women with endometriosis. This phenomenon may promote growth and maintenance of ectopic lesions and may play a role in endometriosis-related infertility by altering the programming of eutopic endometrium.

The diagnostic usefulness of tissue aromatase has been evaluated (Kitawaki et al., 1999b, Kitawaki et al., 1999a). This group reported 91% sensitivity and 100% specificity for distinguishing-disease free women from those with endometriosis and other diseases like adenomyosis and fibroids. Other studies have reported no detectable changes in expression of aromatase in the endometrium of women with

disease (Colette et al., 2009, Velasco et al., 2006). Additional studies need to be conducted to evaluate the role of aromatase and to assess its value as a biomarker.

1.8.6.3 Nerve Fibres

High density small unmyelinated sensory nerve fibres have been shown in the functional layer of the endometrium from women with endometriosis especially in the secretory phase of the menstrual cycle (Tokushige et al., 2006b, Tokushige et al., 2006a). The detection of endometrial nerve fibres has been proposed as a semi-invasive diagnostic tool for endometriosis (Al-Jefout et al., 2009, Bokor et al., 2009). Endometrial biopsies from 99 women presenting with pelvic pain and infertility were stained for the neuronal protein PGP9.5 (Al-Jefout et al., 2009). Sensitivity and specificity was 83% and 98%. Nerve fibre density did not differ between cycle phase and women with endometriosis and pain had significantly higher nerve fibre density compared to those who were infertile with no pain. This study was however limited by lack of uniform histological confirmation of endometriosis. The other study tested the presence of sensory nerve fibres in the secretory phase of the endometrium of patients with minimal to mild endometriosis compared to controls (Bokor et al., 2009). Tissues were stained with PGP9.5, neurofilamin (NF), vasoactive intestinal peptide (VIP), substance P (SP), neuropeptide Y (NY) and calcitonin gene related peptide. The density of nerve fibres was 14 times higher in the eutopic endometrium of endometriosis patients compared to controls. A predictive model with PGP9.5, VIP and SP was able to detect early stage endometriosis with 95% sensitivity, 100% specificity and 97.5% accuracy. These results are promising, but need to be evaluated further, especially to establish whether nerve fibres could be useful markers of endometriosis and are not just related to pelvic pain.

1.8.6.4 Matrix metalloproteinase (MMPs)

MMPs are a group of enzymes important for the control of extracellular matrix turnover (Brunner K.L *et al,* 1999). They are involved in tissue remodelling and angiogenesis. MMPs are upregulated by TNF-α and IL-1, which could contribute to

56

the invasiveness of endometrial fragments in women with endometriosis (Sillem M *et al*, 2001). TNF-α may also contribute to decreased expression of endogenous tissue inhibitors of MMPs (TIMPs) (Gottschalk *et al*, 2000). The expression of MMP1 by endometriotic cells was shown to be increased as compared to eutopic endometrium of patients and controls (Hudelist G *et al*, 2005, Di Carlo 2009). Another study however reported no change in MMP1 expression (Kyama et al., 2006a). Raised MMP 3 expression has also been reported in the eutopic endometrium of patients with endometriosis versus controls (Ramon et al., 2005, Gilabert-Estelles et al., 2007).

1.9 Proteomics in endometriosis research

Proteomics is the study of the expressed proteins in a cell or tissue, also known as the proteome. This includes the expression of all protein isoforms, their chemical modifications, structures and interactions with other proteins and biomolecules (Tyers and Mann, 2003, Anderson and Anderson, 1998). Variation in protein expression in different diseases may happen by different mechanisms which often cannot be detected using genomic methods. Therefore one of the main objectives of proteomics is to determine these protein changes and to assess how these changes could affect function (Patterson and Aebersold, 2003). Areas of current research in biomedicine where proteomics is used include: study of altered protein expression at the tissue, cellular and sub-cellular level, discovery of novel biomarkers for the diagnosis and early detection of various diseases and the identification of new therapeutic targets (Hanash, 2003, Anderson and Anderson, 1998). The proteome is a very complex and dynamic entity and because of this, the impact of proteomics in medicine (as compared to genomics) has yet to be felt. Advances in proteomic technologies however are likely to improve the detail at which proteomes can be characterised and therefore provide important leads in the future.

The molecular pathways by which endometriosis occurs are currently not known. Genome analysis has been widely applied with the hope of providing additional useful information about the differences between women with and without endometriosis. Transcriptomics (gene expression profiling) which is the study of the transcripts

(mRNA) in a cell or tissue population has been used to provide information regarding the molecular basis of endometriosis (Giudice, 2003b, Giudice, 2003a). It is important to note however that changes in mRNA levels do not necessarily reflect changes to the abundance of the corresponding translated proteins. Various mechanisms of post-transcriptional regulation exist such as mRNA degradation, silencing and splicing, whilst protein translation and turnover also vary considerably. Proteomics has the potential to provide considerably more information about disease and can identify relevant targets for diagnosis and therapy (Hanash, 2003).

The main goal of proteomic approaches to date in the study of endometriosis has been the search for biomarkers (Fassbender et al., 2013). However, despite numerous studies in this area, no single biomarker or panel of markers has been deemed fit enough for accurate diagnosis for clinical use, and whilst some show promise, they require further testing and validation (Fassbender et al., 2012b, Vodolazkaia et al., 2011, Vodolazkaia et al., 2010, Bokor et al., 2009). From these studies however, a vast amount of knowledge has been acquired regarding the molecular events and pathways that lead to disease development and progression, hence providing a better understanding of this enigmatic disease.

The reason why application of proteomic technologies has not yet provided applicable tools for diagnosis of endometriosis is likely because of the heterogeneous nature of the tissue and the fact that the endometrium is constantly remodelling itself in response to fluctuations in sex steroid hormones and drug treatments. Emerging technologies for protein detection which are more sensitive, as well as better sample preparation methods, pre-fractionation, protein quantitation and improved bioinformatics are permitting a more in-depth understanding of the proteomes of tissues and biological fluids. These methods can now be applied in the search for biomarkers of endometriosis.

1.10 Proteomic technologies for biomarker discovery in endometriosis

Proteomic analysis allows comparison of proteins from any biological source and permits the identification and quantitation of multiple proteins within a single experiment (Meehan et al., 2010). Experimental reproducibility is the key to biomarker identification. Sample handling at all stages of the experiment is a crucial factor. Samples must be collected, processed and stored according to highly standardised and robust procedures. To date, the two most commonly applied technologies in the investigation of endometriosis have been two dimensional gel electrophoresis (2DE) and surface-enhanced laser desorption ionisation time-of-flight mass spectrometry (SELDI-TOF MS) (Fassbender et al., 2013, Linda C Giudice, 2012).

In 2DE, proteins are first separated by iso-electric focusing according to their native pI (iso-electric point) and then by SDS-PAGE (sodium dodecyl polyacrylamide gel electrophoresis) according to their molecular weight. Protein species are then detected in the gel by direct staining or by imaging if the proteins have first been labelled with fluorescent tags. In studies of endometriosis, 2DE has been applied to both tissue biopsies and biofluids. The first study compared protein profiles of serum and peritoneal fluid of women with and without endometriosis (Joshi et al., 1986) with the hope of establishing whether infertility in women with endometriosis was as a result of autoimmune mechanisms. No significant changes were observed, but an unidentified protein of 70 kDa was observed in 18/20 peritoneal fluid samples in the secretory phase that was not present in the proliferative phase.

In another study, 2DE and mass spectrometry were used to establish the effects of endometriosis on the proteome of eutopic endometrium comparing tissue from women with and without endometriosis (Fowler et al., 2007). Proteins found to be differentially expressed included chaperones, redox regulators, proteins involved in DNA metabolism and secreted proteins.

Protein profiling using SELDI-TOF MS is a relatively high-throughput, but low coverage technique that involves capture of proteins on protein chips which have different chromatographic surface chemistries to enrich specific protein subsets. Captured proteins are directly analysed by mass spectrometry and comparisons made between spectra to identify changes in protein abundance between samples. Some of the advantages of SELDI are that small amounts of sample can be used and it is a rapid technique. Its main limitation however is that it does not allow for the direct identification of proteins (Poliness et al., 2004, Linda C Giudice, 2012).

SELDI-TOF MS has been applied in the analysis of serum, plasma and tissue from women with and without endometriosis. The first SELDI-TOF MS study in endometriosis demonstrated several polypeptides and proteins that were expressed differentially in eutopic endometrial tissue from women with and without endometriosis (Kyama et al., 2006b). Endometrial polypeptides of 2.8-12.3 kDa were decreased in women with endometriosis whilst a 23 kDa protein subsequently identified as transgelin, was increased. A follow-up study to investigate differentially expressed proteins and peptides in endometrial tissue at the secretory phase revealed two up-regulated (90.7 kDa and 36 kDa) and two down-regulated peaks (1.9 kDa and 2.5 kDa) that together could classify minimal-mild endometriosis with 100% accuracy (Kyama et al., 2011). The up-regulated proteins were identified as T plastin and annexin V; proteins reported to play roles in cell mobility, proliferation, attachment and early invasion of endometrial cells. Five peaks (5,385 m/z, 5,425 m/z, 5,891 m/z, 6,448 m/z and 6,898 m/z) were described in another study that could discriminate endometriosis from healthy controls (Wang et al., 2010). These studies suggested that SELDI-TOF MS analysis of endometrial tissue may be a promising method for biomarker identification for the diagnosis of endometriosis.

Analysis of serum proteins using SELDI-TOF MS identified peaks that together achieved >90% specificity with 20% sensitivity (Seeber et al., 2010). Combining these with their best performing markers from a previous study (Seeber et al., 2008b), resulted in 73% of all subjects that would have been diagnosed with 94% accuracy. Another study analysed 254 plasma samples in women with and without

endometriosis during the menstrual phase (Fassbender et al., 2012b). A model was developed using five peaks: 2,058 *m/z*, 2,456 *m/z*, 14,694 *m/z* 3,883 *m/z* and 42,065 *m/z* that was able to detect ultrasound-negative endometriosis with a sensitivity of 88% and specificity of 84%. They also identified a high intensity peak (2,189 *m/z*) that was down-regulated in women with moderate to severe endometriosis compared to the controls. This peptide, identified by MALDI-TOF/TOF as a fragment of fibrinogen β chain, has been reported by the same group to be down-regulated in uterine fluid from baboons induced with endometriosis (Patentstorm.Us/Patents/7794958, Fassbender et al., 2013).

Despite these advancements, SELDI-TOF is known for its poor reproducibility and poor mass accuracy, making reliable protein identification difficult. To this end, other quantitative proteomic tools that allow for the comparison of complex protein profiles have been developed. Described below are the major approaches that will be used in this study to search for the discovery of potential biomarkers for the non-invasive diagnosis of endometriosis. These approaches will be used in combination to quantitatively characterise and identify proteins that are differentially expressed in endometrial tissue of women with and without endometriosis.

1.10.1 Two dimensional-difference gel electrophoresis (2D-DIGE)

Separation of complex protein mixtures into simpler fractions or individual constituents is normally the initial step in any proteomics experiment. For many years 2DE was the method of choice for protein separation and profiling (O'farrell, 1975, Garrels, 1979). As mentioned earlier, 2DE involves separation of proteins in two dimensions by their isoelectric points and by their molecular weight. This method however has several limitations with poor reproducibility being the most significant (Viswanathan et al., 2006, Timms and Cramer, 2008, Minden et al., 2009). It was reasoned that separating multiple samples on the same gel would partly overcome this problem, allowing a level of internal control. To facilitate this, a method of differentially labelling proteins within a mixture was needed. This was made possible in difference gel electrophoresis (DIGE) with the introduction of fluorescent dyes that

are used to label the proteins before running on a gel. The CyDyes (Cy3-NHS, Cy5-NHS and Cy2-NHS), are synthetic cyanine dyes with a reactive N-hydroxysuccinimidyl ester (NHS). The CyDyes react with primary amino groups on proteins, have the same molecular weight and preserve the charge of the target amino acid, but have specific fluorescence emission spectra . The charge and mass matching ensures that proteins from differenlly labelled samples will co-migrate to the same position during 2DE.

The principle of DIGE was originally described by Minden and colleagues (Unlu et al., 1997). Since then, this method has been extensively used to analyse protein changes in different cells, tissues and body fluids. Experimental design and sample preparation are vital in any DIGE experiment. The protocols depend on the type and number of samples to be analysed. A typical DIGE experiment involves sample lysis to extract and denature the proteins with lysis carried out using urea- and zwitterionic detergent-containing 2DE lysis buffer. Equal amounts of proteins are then differentially labelled with the fluorescent dyes (Cy3, Cy5 and Cy2) which covalently label lysine residues ε-amino groups and polypeptide N-terminal amino groups in the different samples. The labelled proteins are then mixed together in equal amounts and separated by 2DE on the same gel. The Cy2 dye is typically used to label an internal standard that is run on all gels against pairs of test samples labelled with the other two dyes (Alban et al., 2003, Gharbi et al., 2002). Resolved, labelled proteins are thereafter detected at appropriate excitation and emission wavelengths using a fluorescence scanner and signals from protein spots compared. Gels are typically post-stained after differentially expressed proteins/spots are selected and spots of interest are then picked from the gel for identification by mass spectrometry (MS).
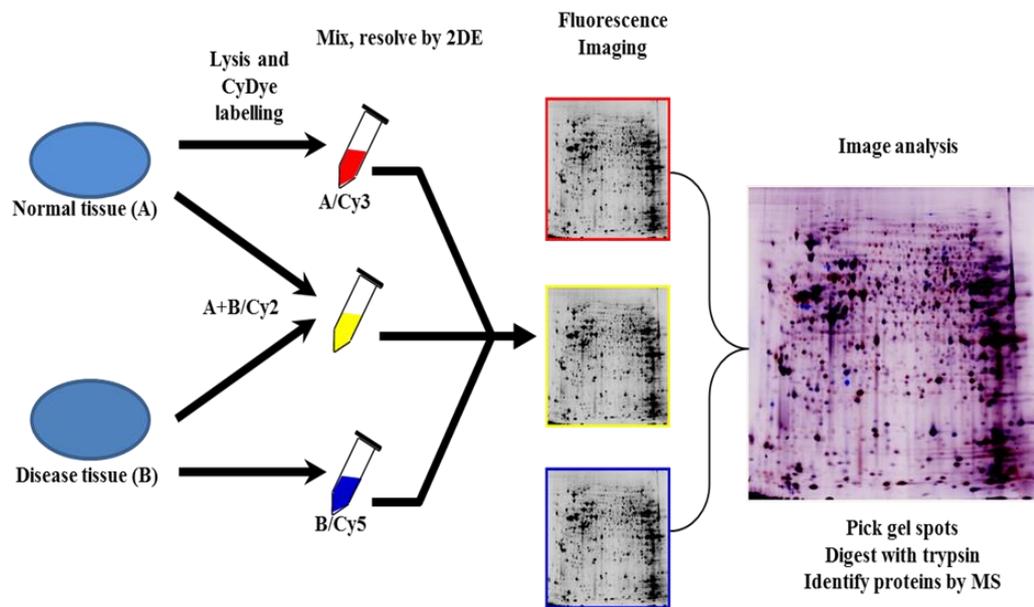
Figure 1.1 2D-Difference gel electrophoresis
Samples are labelled using three dyes; Cy2, Cy3 and Cy5. Proteins are resolved by electrophoresis in two dimensions, the gels imaged on a fluoresecence scanner and iamge analysis performed. Differentially expressed proteins are picked from gels and identified by MS**.**

Protein profiling by fluoresence 2D-DIGE has several advantages over standard 2DE using protein stains: it is more reproducible, more accurate and more sensitive with a higher throughput and a larger dynamic range of protein abundance can be analysed. Samples are subjected to the same conditions during 2D separation, hence raising the confidence with which proteins can be matched and quantified. Post-translational modifications which play a role in regulating protein function can also be detected if they affect protein pI. Introduction of the internal standard allows for normalisation of each spot across all gels enabling more accurate differentiation of biological from experimental variation (Alban et al., 2003, Diez et al., 2010). However, several problems inherent to 2DE still exist; there is poor resolution of basic, hydrophobic and large proteins. Additionally labelled proteins migrate slower than unlabelled ones causing spot misalignment in the lower molecular mass region. This may interfere with spot picking and subsequent MS identification (Gharbi et al., 2002). Other problems can include high fluoresence background, detection of non-protein signals and protein-dye bias (Timms and Cramer, 2008).

2D-DIGE has not been used extensively in endometriosis studies. 2D-DIGE and MALDI-TOF-MS were used to identify proteins with altered abundance in mid-secretory phase eutopic endometrium from women with and without endometriosis (Stephens et al., 2010). A total of 72 differentially expressed proteins spots were found with 20 of these identified by MS. Most had not been previously reported as being associated with endometriosis. Five proteins were selected for validation by immunohistochemistry and western blotting: vimentin (VIM), peroxiredoxin 6 (PDRX6), ribonuclease/angiogenin inhibitor 1 (RNH1), coronin 1A (CORO1A) and transgelin 2 (TAGLN2). RNH1, VIM and PDRX6 all showed a confirmatory decrease in patients with endometriosis, however an additional isoform of PDRX6 was over-expressed in endometriosis patient samples compared to those from healthy women. The fold-changes were compared with previously reported micro-array data (Burney et al., 2007) in an attempt to compare gene expression with protein abundance data in endometriosis patients. Most of the proteins showed no correlation with their mRNA

level in line with previous reports that protein abundance chnages are poorly correlated with changes in mRNA in endometrial tissue (Chen et al., 2009, Fassbender et al., 2012a).

1.10.2 Mass spectrometry

One of the most powerful modern analytical techniques available is mass spectrometry (MS). Over the last decade MS-based proteomics has rapidly become the analytical method of choice for identification and characterisation of proteins (Kicman et al., 2007, Yates et al., 2009). MS allows molecules to be identified by the production of ions, their subsequent separation and detection based on their mass to charge ratio, conferring a high level of specificity with sensitivity.

The development of different types of ion sources and mass analysers has had considerable impact in biomedical studies of large organic molecules (Aebersold and Mann, 2003, Mano and Goto, 2003, Kito and Ito, 2008). Ionisation techniques in the past limited research to volatile compounds or those that could be enhanced to make them volatile. Large, polar biomolecules that are involatile e.g. proteins and peptides presented a particular problem. In 1990, two ionisation techniques were able to enter fully into the biological arena. These were electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI). These approaches are known as soft ionisation techniques and they allow molecules to remain relatively intact during the ionisation process permitting accurate mass measurement and identification of polypeptides (Fenn et al., 1989, Karas and Hillenkamp, 1988, Kicman et al., 2007, Yates et al., 2009).. ESI works by ionizing the analytes from the liquid phase and is the preferred ionization method for the analysis of complex mixtures. It is often coupled to liquid-based separation tools such as chromatography. MALDI is used to analyse simpler mixtures of peptides by sublimating and ionizing the analytes from a dry, crystalline matrix via laser pulses (Aebersold and Mann, 2003, Kicman et al., 2007).

The mass analyser is interfaced with the ionization process. Ions produced are subjected to separation according to their mass to charge ratio (m/z) in the mass analyser. There are five basic types of mass analysers: ion trap (IT), quadrupole trap (QT), time of flight (TOF), orbitrap and Fourier transform ion cyclotron resonance (FT-ICR). The type of mass analyser will dictate the type and quality of experimental data obtained with reference to mass accuracy, mass resolution, mass range and sensitivity. Each mass analyser however has inherent advantages and disadvantages which should be considered before the start of any MS experiment.

The ions emitted from the mass analysers are measured by a detector. The detector is an electron multiplier or micro channel plates. When each ion hits the detector a wave of electrons is emitted resulting in amplification of the signal for improved sensitivity. To ensure that non-sample ions do not collide with sample ions, this process is performed under high vacuum. Data from the detector is then analysed to provide information regarding the m/z of the ions and their relative abundance. This generates a mass spectrum. This information is then processed through database searching to identify the molecule of interest based on its accurate mass and other information such as protease specificity and taxonomy.

Mass analysers have been hyphenated to allow fragmentation of ions and subsequent analysis of daughter ions in the second analyser. In this so called tandem MS (MS/MS), ions of interest from the first mass spectrum (MS1) are selected based on their intensity and m/z. These ions (precursor ions) are then fragmented by colliding the ions with an inert gas such as helium. This fragmentation process is commonly referred to as collision induced dissociation (CID). Another round of MS is carried out on the fragmented ions generating a second mass spectrum (MS2). Since the fragmentation can be controlled, a series of fragment ions is generated, the mass differences of which can be used to determine amino acid sequence, greatly improving on the ability to identify peptides and proteins.

1.10.3 Separation technologies

Mass spectrometry is highly dependent on powerful separation technologies that simplify complex biological samples prior to mass analysis. Multidimensional separation couples two or more different separation methods by which the analytes are first separated by one method and then by one or more independent separation methods. It is important to consider the orthogonality of the individual separation methods in which each dimension uses different (orthogonal) molecular properties of molecules as a basis for separation. Two major approaches for separation are gel-based and liquid chromatography-based.

1.10.3.1 Ion exchange chromatography (IEC)

Ion exchange chromatography is the reversible adsorption of charged molecules to immobilised ion groups on a matrix of opposite charge. Separation can be selectively achieved by adsorption and release of analytes from the matrix by applying a gradient of eluting buffer of increasing ionic strength or pH (Figure 1.2.). Fractionation of proteins and peptides by IEC depends upon differences in the charge of different proteins and peptides which depends upon the number and type of ionisable amino acid side chains.

Ion-exchange chromatography
(anion exchange)

Mix of amino acids

Flow of solution

Immobilized
cation surfaces

Binding of negatively
charged amino acids
to immobilized cation
surface

Separation of negatively
charged amino acids
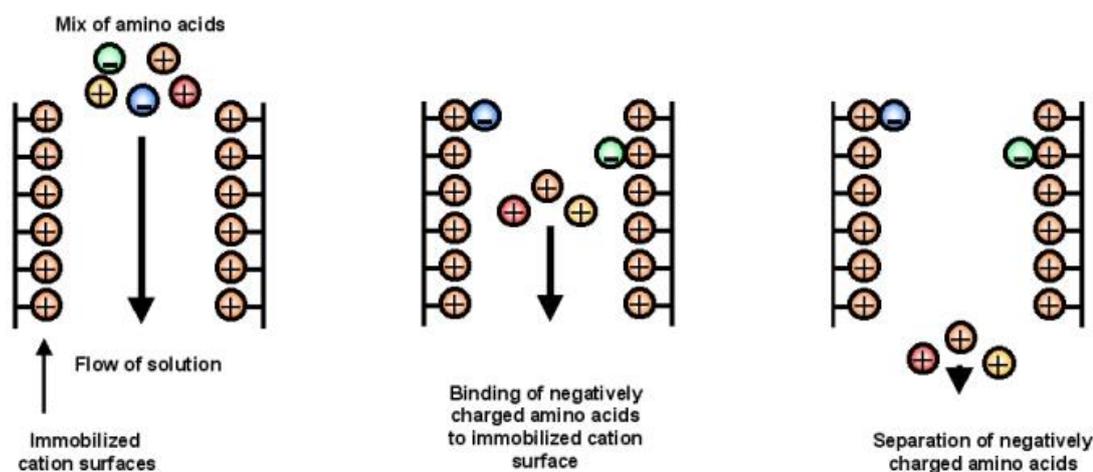
Figure 1.2 Schematic diagram showing anion exchange chromatography
The stationery phase is a positively charged resin which binds negatively charged
molecules. The strength of binding is affected by the pH and salt concentration of the buffer
and by changing either of the conditions the bound molecules can be eluted off. Increasing
the salt concentration of the buffer, the anions in the salt would displace the bound anions on
the column.

1.10.3.2 High Performance Liquid Chromatography

HPLC is a chromatographic method used to resolve molecular species on the basis of
their interactions with a column-bound stationery phase. Analyte molecules are passed
through the column and are retarded by specific chemical and physical interactions
with immobilised chromatographic supports. Bound peptides are sequentially eluted
using a graduated mobile phase. The specific point at which a peptide elutes from the
column is known as the retention time and is a unique property of the analyte
molecule. Reversed-phase liquid chromatography (RPLC) separates peptides
according to hydrophobicity using a linear gradient of organic solvent. It is usually
the final dimension of separation and can be coupled directly to a mass spectrometer
in so-called LC-MS/MS. For compatibility, low flow rates are required (200-500
nL/min), so the technique is often referred to as nano-LC-MS/MS

1.10.4 Electrospray ionisation (ESI)

ESI is a soft ionisation technique which allows the analysis of large intact biomolecules such as protein and DNA. This process involves production of ions by spraying a solution containing the analyte into an electrical field under atmospheric pressure (Fenn et al., 1989). The analyte is ionised by the source which is held at a positive and negative potential depending on the analyte of interest (Figure 1.3). Peptides are typically analysed in the positive ion mode and are dissolved in organic solvents to assist ionisation. The principle in ESI is that a spray of charged liquid droplets is produced by atomisation or nebulisation. ESI is achieved by applying a strong electrical field at the end of the capillary. As the solvent evaporates from the droplets, analyte ions are transferred into the gas phase. Heat and gas are applied to assist droplet desolvation which initiates a series of Raleigh instabilities (Coloumb fissions) which ultimately produce individual gas phase ions (Nguyen and Fenn, 2007). ESI is the method of choice for LC-MS/MS analysis due to its compatibility with reverse phase liquid chromatography (RPLC) and its ability to produce multiply charged ions. Typical solvents include acetonitrile or methanol with 1% acetic acid or 0.1% formic acid.

Electrospray Ionisation (ESI) and Ion Source Overview
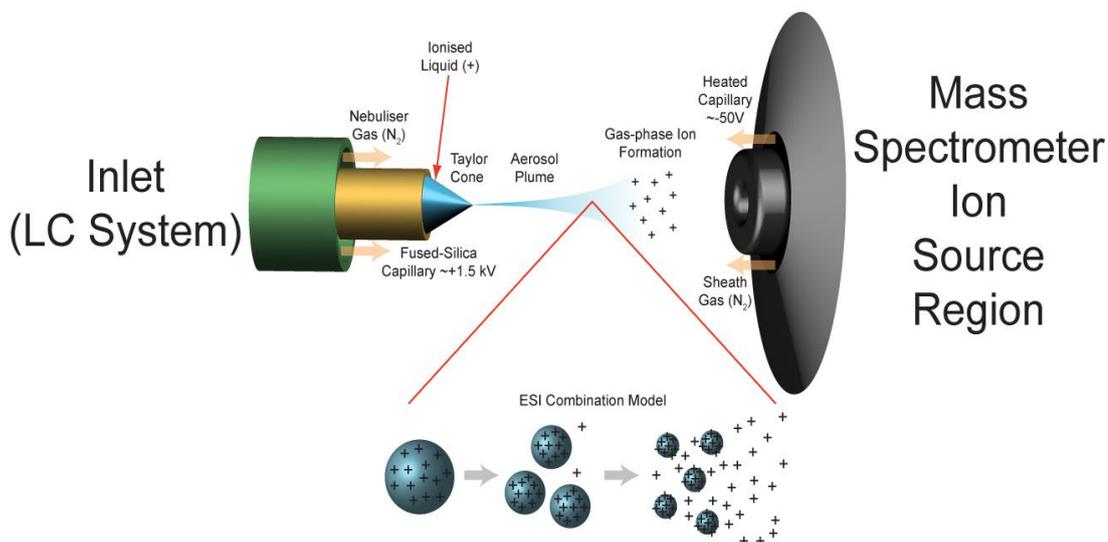
Figure 1.3 Diagram showing electro-spray ionisation.
ESI is used to generate positive or negatively charged ions from RPLC. Liquid eluting from the LC is attached to a fused silica capillary emitter which has a pulled tip of diameter less than 10 μm. The liquid is forced through this constriction and for positive ions a high voltage is applied (~1.5kV) inducing the formation of a spray of positively charged droplets from a Taylor cone generated at the tip. Solvent-free ions are produced by the combined effect of Coulomb fission and ion evaporation. Multiply charged ions are then directed towards the MS by applying an electric field.

1.10.5 Linear trap quadruple mass analyser (LTQ)

The LTQ is a high-throughput instrument which features fast scan times and high sensitivity. Typically, ions enter the analyser and are focused into the ion trap by a series of multipoles made up of four parallel cylindrical metal rods (quadrupoles). A fixed direct current voltage (DC) and an oscillating radio frequency voltage (RF) are applied to each rod, creating a continuously varying electric field along the length of the poles. Ions are sequentially attracted or repelled by the poles and focus into a concise beam which accelerates the ions towards the trap. Within the ion trap, the oscillating amplitude (RF) is ramped from low to high voltage and a complementary alternating current (AC) is applied to the horizontal (x) rods (resonance ejection

voltage). The ions oscillate back and forth in the x direction and become increasingly unstable in order of *m/z* (low to high). Unstable ions are scanned from the trap through ejection slots in the x rods. The RF amplitude at which ions become unstable is *m/z* specific and thus the instrument interprets the signal as a particular mass. Ejected ions strike a conversion dynode in the detector and release secondary particles (for positive ion mode these include negative ions and electrons). The particles are transferred to a multiplier where they create an electron cascade as they move towards the anode which results in a measurable current that is proportional to the ion intensity. This generates a mass spectrum which is the signal intensity of the ions at each value of the *m/z* scale.

1.10.6 Orbitrap mass analyser

The orbitrap mass analyser has the ability to deliver low ppm mass accuracy and extremely high resolution all within a time scale compatible with nano-LC separation. The orbitrap is an ion trap instrument without the radio frequency (RF) or magnetic field to hold ions inside. Moving ions are instead trapped in an electro-static field (Hardman and Makarov, 2003, Scigelova and Makarov, 2006). The electrostatic attraction towards the central electrode is compensated by a centrifugal force that arises from the initial tangential velocity of ions. The electrostatic field which ions experience inside the orbitrap forces them to move in complex and spiral patterns. The axial components of these oscillations is independent of initial energy, angles and positions and can be detected as an image current on the two halves of an electrode encapsulating the orbitrap. The axial component of these ion oscillations is measured by a Fourier transform resulting in accurate reading of their *m/z*. As a result, the machine has a very high resolution and mass accuracy (Hu et al., 2005, Zubarev and Makarov, 2013, Scigelova and Makarov, 2006).

1.10.7 Tandem Mass Spectrometry

Mass spectrometry can obtain primary structure information in a process known as tandem mass spectrometry (MS/MS). MS/MS is carried out in the data-dependent

71

mode whereby the most intense precursor ions from a full MS scan are isolated for fragmentation into smaller product ions for mass analysis. In the LTQ Orbitrap XL™, the orbitrap is used for high precision MS1 scans, whilst the subsequent MS2 analysis occurs in the LTQ following CID fragmentation (Figures 1.4 and 1.5). Ions generated by ESI are transferred by axial ejection into an RF only quadrupole called the C-trap. The C-trap accumulates and stores the ions before injection into the orbitrap. The development of an additional collision cell provides additional flexibility to MS/MS experiments by allowing ions to be selected in the linear ion-trap and fragmented either in the ion trap by collision induced dissociation (CID) or by higher energy collision dissociation (HCD) in an additional gas-filled collision cell. The ions encounter frequent collisions with the dampening gas thus increasing their internal vibrational energy until they fragment. The product ions are ejected from the trap and detected. CID induces fragmentation at the amino bonds along the peptide backbone. A fragment must have at least one charge for it to be detected. If the charge is retained on the N terminal, the resulting ions are referred to as a, b or c ions. If the charge is retained on the carboxy terminal, the ion type is x, y or z. By using the *m/z* of the precursor ion as a guide, the masses of the product fragments are then matched against a database that is derived from the predicted peptide and fragment ion masses of all known protein sequences.
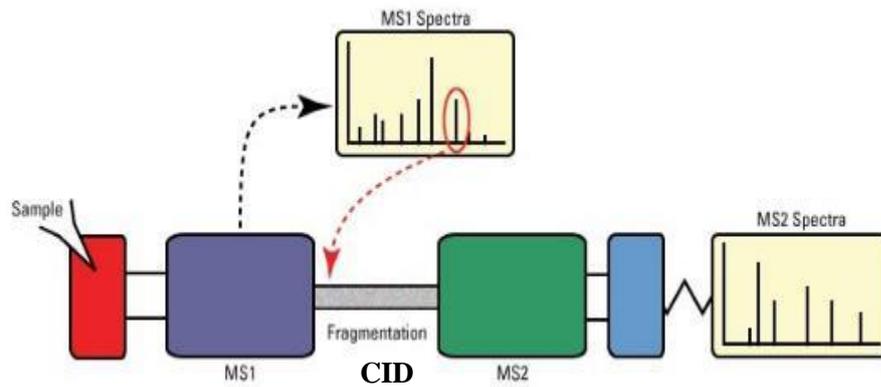
Figure 1.4 Schematic of tandem mass spectrometry (MS/MS). The sample is introduced into the mass spectrometer, ionised and then analysed by MS1. Ions from the MS1 spectra are then selectively fragmented and analysed by MS2 to give the spectra for the ion fragments.

Nano-LC system

MS operation

•Ion trap transmits all ions from ESI source to Orbitrap for high-resolution, high mass accuracy MS scan

MS/MS operation

•Automated switching to MS/MS mode
•Up to 6 of the most intense ions selected in the ion trap for fragmentation by collision induced dissociation (CID)
•Fragment ions measured in ion trap

Total ion chromatogram

MS scan (at RT 24 min)

MS/MS scan (for *m/z* 537.34)

Figure 1.5 Schematic diagram of the LTQ Orbitrap mass spectrometer – Peptides are introduced mass spectrometer by ESI . Ions are separated by their mass to charge in the first stage (MS1). These pre-cursor ions are selected and fragmented at a second MS stage (MS2). These product ions are then separated and detected.

1.10.8 MS-based profiling

Intact protein samples can be analysed directly by either 2DE, antibody-based methods or by MS using a so-called top-down approach. The commonly employed top down approaches in MS are MALDI-TOF and SELDI-TOF. The mass spectra obtained from these two methods does not provide the identity of the protein, but gives the relative abundance of the masses detected. The protein profiles are then analysed and selected masses of interest can be purified and identified. Top-down proteomics by MS/MS is not commonly used for profiling of complex protein mixtures, mainly because sufficient resolution and efficient fragmentation of large intact proteins cannot be achieved. However, the approach for characterising smaller, purified proteins has several advantages in that high sequence coverage of target proteins can be obtained, post-translational modifications can be better characterised and protein quantification is more reliable.

The bottom-up approach (also known as shotgun proteomics) is the most commonly used approach for large-scale MS-based proteomic analysis of complex mixtures.. The basis of this approach is enzymatic cleavage of proteins to peptides using a protease such as trypsin. The generation of these smaller peptides helps in ionisation and fragmentation and provides sequence information for the final identification of the peptide. Chromatographic separation is applied to reduce the complexity of peptide mixtures prior to MS/MS. Particularly, RPLC separation is often linked directly to the mass spectrometer via an ESI source such that eluting/separated peptides can be analysed 'on the fly' by MS/MS. Peptide ion and fragment ion masses are then searched against sequence databases of all predicted open reading frames which have been digested *in silico*. The search returns peptide hits based on the probability of the mass matches which is dependent upon the mass accuracy of the instrument. Peptide hits are matched to protein sequences to confirm the identy of proteins present in the sample.

Quantitative proteomics is important for understanding how global protein expression and modification ochestrates various biological processes and disease states. Bottom-

up label-free quantification is based on the ion intensity of identified peptides with peptide matching between runs and samples based on the *m/z* and retention time This has drawbacks in terms of quantitative accuracy as matching is challenging and multiple repeat runs are necessary. Quantitation of proteins is achieved by averaging or taking the median intensities of the peptides matched to a particular protein, so information may be limited for lower abundance proteins where there are fewer matching peptides. Spectral counting is another method of label-free quantification whereby the number of spectral matches for peptides from a certain protein is used as a surrogate of that protein's abundance (Liu et al., 2004). Quantitative proteomic analyses also use isotopic labelling of proteins or peptides across samples which can thereafter be differentiated and compared by MS. These labelling methods can be classified into metabolic labelling (e.g. SILAC) or enzymatic and chemical labelling (e.g. TMT, iTRAQ) and allow multiplexing, whereby mixing of differentially labelled samples reduces variability during fractionation and reduces the number of samples that need to be analysed to provide an acceptable level of quantitative accuracy.

Tandem mass tags (TMT) is the quantification method that will be used in this study. These tags are small chemical entities with identical structure and mass (isobaric) that are covalently attached to the amino groups of lysine residues and the N-termini of peptides (Thompson et al., 2003, Gygi et al., 1999, Coombs, 2011). Commercially available isobaric tags offer simultaneous analysis of up to 10 samples. In principle, each isobaric tag produces a unique reporter ion during MS/MS that are used for relative quantification of each labelled peptide. Thus in the first MS, labelled peptides are indistinguishable from each other. During the second MS, peptides are fragmented allowing each tag to produce a unique reporter ion based on isotopic labelling with a mixture of $^{12}C/^{13}C$ and $^{14}N/^{15}N$. Protein quantitation is then accomplished by comparing the intensities of the reporter ions in the MS2. It is important to note that unlabelled peptides are not quantified (Thompson et al., 2003, Bantscheff et al., 2008). Isobaric tags have enabled protein identification and quantification from various cell types, tissues and body fluids and are proving to be important tools in biomarker discovery research.

1.11 JUSTIFICATION OF RESEARCH

Endometriosis is a gynaecological disorder that affects 6-10% of women of reproductive age. It is characterised by the presence of endometrial glands and stroma in ectopic locations such as the ovaries, fallopian tubes and rectovaginal septum (Fassbender et al., 2013). Common symptoms of endometriosis include severe dysmenorrhoea, non-menstrual pelvic pain, dyspareunia, dysuria, dyschezia and infertility (Meehan et al., 2010). Pain associated with this disease is a result of inflammation in the peritoneum, presence of adhesions and innervation of endometriotic lesions. (Bulun 2009 (Ballard et al., 2006).

Endometriosis is not a life-threatening disease, but it impacts greatly on the quality of life of affected women (Chapron et al., 2003, Chapron et al., 2006). However, this impact has been poorly researched with available reports focusing on selected populations; mainly western nations; (Simoens et al., 2007) with small sample sizes, poorly selected control subjects and inadequate validation tools (Gao et al., 2006b). Despite this, it is recognised as a major cause of severe morbidity in women and impacts on their physical and emotional wellbeing. Psychologically, endometriosis and related symptoms may cause anxiety, depression and feelings of uncertainty, which in turn can interfere with a woman's perceived sense of control, handling of adverse situations and resourcefulness.

Physically, endometriosis pain can impair work-related and daily activities e.g. ability of the affected women to maintain a career (Gao et al., 2006a). It has been reported that 50% of women with endometriosis are bed-ridden several times each year, interfering with education, work and day to day living (Kjerulff et al., 1996). Loss of productivity has been estimated at 10.8 hours a week owing to reduced effectiveness while working (Nnoaham et al., 2011). Associated costs to society including those of delayed diagnosis, mistreatments and individual costs incurred when symptoms associated with this disease interfere with daily life are considerable, but poorly characterised (Gao et al., 2006a). The annual cost estimates in the US was reported as $22 billion in 2002. This was calculated from the estimated cost per patient; $1023-

$2801 per year at a prevalence of 10% among women of reproductive age (Simoens et al., 2007). Indirect costs of the disease were not calculated in this study. Despite this significant health burden the disease is still poorly researched, diagnosed and treated. Infertility/subfertility which is a major consequence of the disease causes an extra burden to the patient due to the uncertainty of ever having a family. Sexual dysfunction due to dyspareunia can disrupt the relationship between a man and a woman (Gao et al., 2006b). This comes with an extra burden of social stigma especially in settings whereby infertility is considered shameful (Somigliana et al., 2010). Efforts to assess the societal cost effects of endometriosis have been reported. The World Endometriosis Research Foundation (WERF) EndoCost study is the first prospective study from 12 centres in 10 countries to examine the direct and indirect costs of endometriosis (Simoens et al., 2011). The study measures direct health care costs like cost of medication and physician visits, direct non-health care costs and indirect costs associated with loss of productivity. The average costs of endometriosis were reported as €9579 per woman per year which equates to €3113 for direct health care costs. The inabilty to work due to symtoms is therefore twice the direct health costs. These cost estimates may be used to raise awareness of endometriosis with policy makers, health professionals and researchers inorder to emphasise the importance of early diagnosis and treatment.

The gold standard of diagnosis is laparoscopy together with histological confirmation. Laparoscopy is an invasive procedure which is expensive and bears significant risks (Kennedy et al., 2005, Bulun, 2009). Efforts aimed at early diagnosis and treatment of endometriosis have been hindered by a lack of proper methods to study and manage the disease. The mean interval between first symptom appearance and diagnosis has been reported to be 7-10 years (Hadfield et al., 1996, Dmowski, 1984, Husby et al., 2003, Ballard et al., 2006). Patients who present with severe pelvic pain that has not been relieved by pain medication or oral contraceptives and those seeking pregnancy for more than one year are the most common patients to whom laparoscopy is recommended in order to guide therapeutic interventions. The availability of a non-invasive diagnostic test would therefore be important to establish endometriosis

instead of subjecting these women to uneccessary surgery whose outcomes might be negative for endometriosis.

Development of a non-invasive test for diagnosis and follow-up has been identified as a top research priority (Rogers et al., 2009b, Rogers et al., 2013). The need for a non-invasive test for asymptomatic women (i.e. screening) is still debatable because this will mean subjecting women to unnecessary and potentially harmful procedures. It is important to note that most subfertile women with or without pelvic pain, having regular cycles, a partner with a normal sperm count and quality, and normal pelvis on ultrasound imaging, may have endometriosis (Meuleman et al., 2009). A non-invasive test would therefore be important for such women. These women could be those with early stage disease and some cases of late stage disease not picked up by imaging methods and those with pelvic adhesions and/or other pelvic pathology who on diagnosis would benefit from laparoscopic treatment.

Serum, plasma, urine, endometrial fluid, menstrual fluid, tissue biopsy and peritoneal fluid are samples that can be studied in the search for a non-invasive or minimally invasive biomarker (Vodolazkaia et al., 2012, Vodolazkaia et al., 2010, Casado-Vela et al., 2009, Kyama et al., 2007). The vital aim is therefore to develop a test in which no woman with endometriosis and/or any other pelvic pathology that would benefit from laparoscopic treatment are missed. A test with high sensitivity and specificity would be ideal for detecting or ruling out endometriosis in patients presenting with symptoms. At present no such test exists.

The WERF Endometriosis Phenome and Biobanking Harmonisation Project (EPHECT) is a global initiative involving 34 clinical/academic and 3 industrial collaborators from 16 countries with a mission to develop a consensus on standardisation and harmonisation of phenotypic surgical, clinical data and biological sample collection methods in endometriosis research mainly to address large scale, cross centre, epidemiologically robust, translational biomarker and treatment target discovery research in endometriosis (Becker et al., 2014, Fassbender et al., 2014, Rahmioglu et al., 2014). This initiative outlines detailed international guildelines for

78

standardised clinical and personal phenotyping (phenome) data to be collected from women with endometriosis and controls to improve patient disease characterisation and standard operating procedures (SOPs) for biobanking of biological samples from women with endometriosis and controls with respect to collection, transport, processing and long term storage of samples collected from these women (Becker et al., 2014).

High-throughput proteomic methods have been developed over the years and are now being used to study various diseases of the female reproductive system (Meehan et al., 2010). These approaches have the potential to identify new disease biomarkers by comparing the abundance of hundreds or thousands of proteins simultaneously across cohorts of patients and controls. This has been the focus of many proteomic studies. Endometriosis is a complex disease, therefore it is may not be possible that a single marker will have sufficient diagnostic accuracy. A marker panel however could provide better diagnostic and/or prognostic power.

Biological sampling is an important step for many biomarker studies. Different types of samples can be used depending on the type of study and subsequent downstream application. Sample quality is a critical factor in proteomic analyses to ensure reproducibility of data. Standardised techniques for sample collection, processing and storage are therefore important in any biomarker study design. Research into the human endometrium may be complex and challenging. This tissue is composed of many cell types (epithelial cells, stromal cells, fibroblasts, pre-decidual cells, leucocytes and cells of the vasculature. The endometrium is also regulated by cyclic hormones and other paracrine and autocrine factors which when combined with the individuals' genetic and environmental background may result in alterations to biological processes. Endometrial tissue is inherently heterogeneous with respect to developmental, temporal and biological composition, therefore cell types within a single tissue can be highly variable e.g. ectopic endometriotic lesions contain relatively few endometrial cells that are often dispersed along with leucocytes among the cells of their recipient surface. Biological variability may therefore arise through differences in tissue composition of the collected samples and heterogeneity of

cellular compositions due to phase of menstrual cycle. This inherent variability needs to be considered when designing studies and analysing data. Standard operating procedures for sample collection should therefore be implemented and must be highly standardised and robust to ensure proper tissue acquisition, processing, utilisation, storage and distribution. Laparoscopy together with histological dating is used to definitively diagnose endometriosis. Eutopic endometrial tissue, ectopic endometriosis tissue collected surgically and blood samples were collected for the purposes of this study. Tissue samples were used for the discovery of candidate biomarkers and selected markers were then tested in serum as potential non-invasive diagnostic markers of endometriosis.

The endometrium can be obtained in five different ways; by use of an endometrial sampling device (Pipelle®), endometrial curettage, hysteroscopy resection, post-hysterectomy excision and brushing. An endometrial sampling device is a thin plastic tube that is inserted into the uterus and used to aspirate the tissue. Curettage involves scraping 'strips' of the endometrium from the uterine lining with the use of a curette. In post-hysterectomy collection, the endometrium can be scraped off using a curette or scissors or aspirated using a Pipelle®. Endometrial brushing involves insertion of a disposable brush into the uterus that is used to collect the sample. The collection method varies with the type of study, tissue of interest and available expertise. In this study, eutopic endometrium tissue samples were collected from women with and without disease by curettage during laparoscopy or after hysterectomy. Ectopic tissue was excised during surgery.

Defining the phenotype of the study population is important to ensure that representative disease and sample types are being used. Heterogeneity resulting from improper classification of study participants may decrease both the sensitivity and power of the study. The impact of other pathological conditions affecting the endometrium must also be considered when defining phenotype. The presence of structural alterations (e.g. fibroids), cancer and immune changes may will affect the phenotype of the endometrium. Exposure to different medications e.g. those used to shrink the lesions before surgery, contraceptives that are also used to manage pain and

environmental toxins may also impact on tissue phenotype resulting in biological variability. The level of phenotypic heterogeneity between studies poses a challenge when it comes to reproducibility and replication of study findings.

Many protein extraction protocols have been developed but selecting a suitable protocol mainly depends on the nature of the starting material and on the downstream applications. Protein analysis involves a number of processing steps; homogenisation for protein extraction, denaturation, reduction of disulphide bonds, alkylation of cysteine residues, enzymatic digestion, protein fractionation/separation, analysis by MS and data analysis to identify and quantify peptides. Due to the diverse biochemical properties of cellular proteins e.g. their charge, size, hydrophobicity, susceptibility to proteolysis, ligand interactions and sub-cellular localisation, no single protein extraction method can capture the full proteome. Comprehensive, uncontaminated and representative protein populations can be difficult to extract from tissue samples partly because of the presence of structural proteins and due to contamination with blood proteins arising during tissue sampling.

1.12 Study Aims

The main aim of this study was to identify novel biomarkers for the non-invasive diagnosis of endometriosis. The hypothesis was that ectopic endometrial tissue and eutopic endometrium from women with endometriosis secrete specific proteins into the blood stream that can be used for the detection of the disease.

The specific objectives were to:

1) Apply proteomic profiling technologies to a set of well-characterised tissue samples for the discovery of candidate biomarkers of endometriosis
2) Investigate the differential expression of proteins in eutopic and ectopic endometrium of women with endometriosis compared to relevant controls
3) Verify differentially expressed proteins in individual tissue samples

4) Test these and other proteins reported in the literature as putative biomarkers in serum collected from the same women

5) Develop a model of best performing markers and validate the most promising candidates as biomarkers for the diagnosis of endometriosis in an independent cohort of women presenting with endometriosis and/or pelvic pain

1.13 Study Design

Patient tissue and serum samples used in this study were sourced from the UCLH Reproductive Medicine Unit following ethical approval (13/LO/0163) and informed consent . These samples were stratified into six clinical groups based on condition and menstrual cycle phase:

a) Eutopic tissue control secretory phase (CS) – scheduled for risk-reducing surgery, asymptomatic, no disease at laparoscopy

b) Eutopic tissue control proliferative phase (CP) - scheduled for risk-reducing surgery, asymptomatic, no disease at laparoscopy

c) Eutopic tissue pain secretory phase (PS) - PID or chronic pelvic pain, no endometriosis at laparoscopy

d) Eutopic tissue endometriosis secretory phase (ES) - diagnosed with endometriosis at laparoscopy

e) Eutopic tissue endometriosis proliferative phase (EP) - diagnosed with endometriosis at laparoscopy

f) Ectopic tissue endometriosis secretory phase (EcS) - diagnosed with endometriosis at laparoscopy.

(Eutopic tissue pain proliferative phase (PP) was not represented due to the lack of women recruited who fitted this group (see Chapter 3). Ectopic tissue proliferative phase (EcP) was also not represented due to lack of samples; some of these lesions were ablated during surgery).

The study design is depicted in Figure 6. Tissues were collected during laparoscopy and flash frozen in liquid nitrogen and stored at -80˚C. Samples were then processed as per an optimised protocol. Briefly, frozen tissue was ground whilst still frozen, lysed in denaturing buffer and then pooled into different clinical groups, as above, based on equal protein amount. Pooled tissue samples would be subjected to two previously optimised proteomic profiling strategies (peptide TMT coupled to 3D-LC-MS/MS and 2D-DIGE coupled to LC-MS/MS) in an initial discovery phase to identify potential candidate biomarkers that differ in expression between normal and diseased endometrial tissue. Both strategies would employ immunodepletion of serum proteins in an attempt to remove contaminating proteins and improve coverage. A third strategy using TMT protein labelling and separation would also be tested as a complementary profiling method. Selected candidates from the discovery work, as well as promising biomarker candidates from the literature, would then be tested in a verification phase using serum samples collected from the same women. Individual candidates and combination models would be tested for their ability to discriminate between endometriosis and control groups and to assess the effect of the menstrual cycle phase. The work would also generate data regarding differences in protein expression in eutopic tissue from women with and without endometriosis and in eutopic *versus* ectopic tissue, that may provide insights into the molecular biology of the disease.

Patient data was also collected for all subjects following informed consent. Specifically; age, ethnicity, disease history, treatment history (OCP, GnRH analogues), co-morbidities, type and severity of pain and histopathology (disease stage, stage of menstrual cycle, anatomic characteristics of disease lesions) data were collected. For all subjects serum C-reactive protein, progesterone, oestrogen and CA125 levels were measured at the Clinical Biochemistry laboratories of the UCLH NHS Foundation Trust using gold standard assays. This information would be correlated with candidate biomarker data to assess possible confounders and/or to improve diagnostic biomarker algorithms, and in the case of progesterone and oestrogen, to corroborate cycle stage.
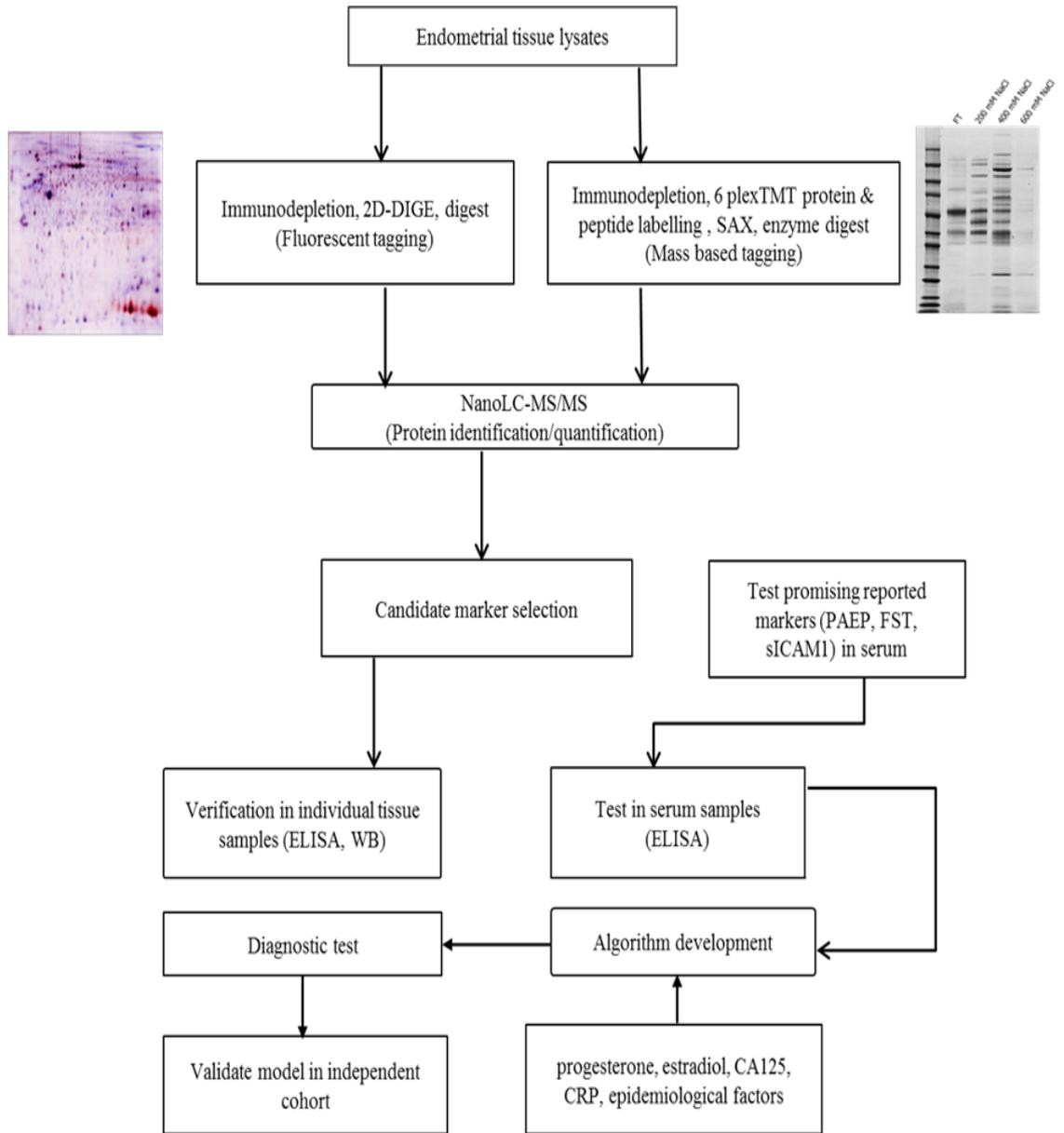
Figure 1.6. Flow chart showing the study design and workflow.

CHAPTER 2: MATERIALS AND METHODS

2.1 Introduction

This chapter explains in detail the methods used in this study including the patient recruitment process, sample selection criteria, sample processing, proteomics methods used to define potential markers, data analysis and verification of these markers in serum. Clinical specimens came from patients recruited to the UCLH Reproductive Medicine Unit (RMU) following ethical approval and informed consent. These were women with endometriosis and those without endometriosis. Endometrial tissue and blood samples were collected from these women at the time of laparoscopic surgery. Two main complementary quantitative proteomic based strategies were employed in this discovery work to define the protein expression profiles of tissue samples obtained from these women. These were 2D-DIGE linked to LC-MS/MS and protein and peptide based TMT-labelling linked to multi-dimensional separation and LC-MS/MS. A scoring system was employed to rank and select candidates that warranted further testing and verification. Verification of selected markers was evaluated using ELISA assays in serum samples taken from the same case and control women. Potential markers identified from the literature were also tested alongside markers identified from the discovery work and correlated with clinico-pathological features obtained during patient recruitment.

2.1.2 Patient recruitment

Patients and samples were sourced from the UCLH Reproductive Medicine Unit (RMU) following ethical approval and informed consent. These patients were those referred to the RMU of UCLH by their general practitioners or other clinicians for investigation of pelvic pain or for diagnosis and/or treatment of endometriosis. These patients were routinely offered laparoscopic surgery for investigation and treatment or bilateral salpingoophorectomy for those with a strong family history of breast and ovarian cancer. The patient inclusion criteria for the study were as follows: cases were defined as women diagnosed with endometriosis at laparoscopy and confirmed histologically. Controls with pain were defined as symptomatic women with pelvic pain of unknown cause or chronic pelvic inflammatory disease (PID) without surgical evidence of endometriosis. Controls without pain were regularly cycling women with no known disease at the time of surgery undergoing bilateral tubal ligation and/or prophylactic bilateral salpingoophorectomy due to a familial risk of breast and ovarian cancer and with no visual evidence of endometriosis upon laparoscopy. Only pre-menopausal women were recruited for this study. The following women were excluded from the study; women with a positive pregnancy test or unknown pregnancy status on the day of surgery, post-menopausal women, those with other benign conditions or malignant tumours (particularly fibroids and cancer were excluded as they may compromise the integrity of the endometrium), women on any hormonal medication <3 months prior to surgery and those whose surgical findings and pathological reports were inconsistent. Cycle phase was determined by a triple approach to ensure accuracy; chronologically, by histological dating and by sex steroid hormone determination. Women with unconfirmed menstrual cycle stage were also excluded.

Additional patient data was collected including age, fertility history, treatment history (oral contraceptive use, GnRH analogues) phase of menstrual cycle, pain history, histopathology findings and anatomic characteristic of disease lesions. All patient records were handled according to NHS confidentiality practices. Samples were anonymised and sequentially numbered. A lab coding system and an excel database

were developed for recording anonymised patient and sample information (see appendix).

2.1.3 Sample collection

Sample collection and processing is a critical step in any study. The adoption of SOPs for tissue collection, processing and storage and standardised phenotypic and other patient data collection are crucial to optimise sample and study quality. The collection method employed should be one that is least likely to cause alterations to the molecular composition of the tissue of interest and also one that is versatile for downstream analysis. To limit variability in tissue composition, sampling should ideally provide a homogeneous histological specimen which should be of sufficient volume for the pathologist to identify the pathology in question and to provide material for the study. Limited sample amounts, heterogeneous cellular composition, presence of abundant structural proteins, blood contamination will all impact on the outcome of subsequent analysis.

Standard operating procedures were developed for tissue and blood acquisition, processing, utilisation and storage. Human endometrial tissue biopsies were obtained by Pipelle or curettage from women of reproductive age undergoing laparoscopy as follows: eutopic endometrium from women undergoing prophylactic bilateral salpingoophorectomy without any evidence of endometriosis, eutopic endometrium from women with chronic pelvic pain of unknown cause or chronic pelvic inflammatory disease without laparoscopic evidence of endometriosis, eutopic endometrium from women undergoing laparoscopy for diagnosis and/or treatment of endometriosis and ectopic endometrial tissue from the latter group i.e. superficial lesions, endometriomas and deep infiltrating nodules were excised from the same women diagnosed and/or treated for endometriosis.

Part of each tissue obtained was fixed in 10% buffered formalin for histological examination while the remaining tissue was washed in sterile phosphate buffered saline (PBS) to remove excess blood. The tissue was then dried using lint free paper,

transferred into labelled and weighed Eppendorf tubes and snap-frozen in liquid nitrogen. The samples were transported to the lab and stored at -80°C.

Cycle stage for each patient was determined by endometrial dating by a trained pathologist without prior knowledge of sample group according to (Noyes et al. 1975). Endometriotic lesions were also confirmed histologically. Samples from patients with surgically diagnosed endometriosis, but inconclusive histological diagnosis were excluded. Samples were also excluded where there was insufficient tissue biopsy (<20 mg wet weight).

10 mL of blood was also collected from each patient before laparoscopic surgery by venepuncture into two BD vacutainer gel tubes. 5 mL of blood was sent to the clinical biochemistry lab of the UCLH NHS Foundation Trust for determination of oestradiol, progesterone, CA125 and CRP levels using standard assays, while the other half was allowed to stand at room temperature for one hour to allow clotting. This was then transported to the laboratory on wet ice and centrifuged at 3000 x g for 10 min at 4°C within 2 hrs of collection. Serum was aliquoted into labelled Eppendorf tubes and stored at -80°C.

2.1.4 Sample preparation optimisation

To enable accurate quantitation of global protein expression in biological samples, optimal sample preparation for MS-based analysis is a critical step in a proteomics workflow. A method was optimised for endometrial tissue lysate sample preparation. Test eutopic and ectopic endometrial tissue obtained by curettage from two women undergoing laparoscopy was used to optimise sample preparation methods, protein extraction and two-dimensional difference in-gel electrophoresis (2D-DIGE). Following excision during laparoscopy, tissues samples were first cleaned with sterile PBS to remove excess blood prior to freezing in liquid nitrogen and storage at -80°C. The representative samples used to optimise sample homogenisation and protein extraction for this study are shown in Table 2.1.

Table 2.1 Tissue samples used for sample preparation optimisation.

Patient ID	Phase of cycle	Tissue Type	Tissue Code	Description
001	Secretory	Ect1	001S3	Left pelvic side wall endometriosis
001	Secretory	Ect1	001S2	Pouch of Douglas endometriosis
001	Secretory	Ect1	001S1	Utero-sacral endometriosis
001	Secretory	Ect2	001E1	Left ovarian cyst
001	Secretory	EmE	001Eu	Endometrium
002	Unknown	EmE	002Eu	Endometrium
002	Unknown	Ect1	002S1	Superficial lesion
002	Unknown	Ect2	002E1	Ovarian cyst

Ect = ectopic endometrial tissue, EmE= Eutopic tissue

Frozen tissues weights were determined by subtracting the weight of the Eppendorf tube plus tissue from the weight of the Eppendorf tube taken before collection. A method was then optimised for the most efficient tissue homogenisation and protein extraction procedure from these endometrial and endometriosis tissues. The best homogenisation was achieved by grinding the tissues under liquid nitrogen prior to re-suspension in urea/CHAPS denaturing buffer. This method was useful especially in the breakdown of ectopic endometrial tissues which can be very fibrous in nature. Other methods such as continuous vortexing or repeated passage through syringe needles failed to break down the tissue, or was virtually impossible to achieve.

The ground tissue was then transferred into sterile tubes; care was taken not to leave any tissue material in the mortar. Protein extraction was then carried out using 2D lysis buffer containing 8 M urea, 4% CHAPS (w/v) and 10 mM Tris-HCL pH 8.3. Urea is a neutral chaotropic agent that is used to effectively disrupt the secondary structure of proteins by disrupting hydrogen bonds and hydrophilic interactions, enabling proteins to unfold with all ionisable groups exposed to the solution. CHAPS is a zwitterionic detergent that is used to disrupt hydrophobic interactions thereby

facilitating cell lysis, protein extraction and solubilisation. A ratio of 10 µL of lysis buffer per mg of tissue was first used to extract the proteins at room temperature.

To ensure maximum protein extraction, sonication was also performed for 20 min. The Eppendorf tubes with the samples were kept on ice during the procedure to avoid heating of the sample which causes urea breakdown and modification of proteins (carbamylation) evoking artefactual charge heterogeneity. Sonication was also efficient for separation of lipids that was observed as a fatty layer on top of the protein extract. Samples were then centrifuged to pellet insoluble material. The supernatant was separated and a Bradford protein assay performed to determine protein concentration as shown in Table 2.2 below.

Table 2.2 Amount of protein obtained after extraction

Tissue code	Protein concentration (µg/µL)	Total yield of protein (µg)
001S1	0.463	125
001S2	1.127	2073
001E1	1.888	3228
001S3	0.974	1499
001Eu	1.965	2731
002Eu	1.038	800
002E1	0.948	6325
002S1	1.353	2110

The starting extraction volume of 10µL/mg of tissue was reviewed and was reduced to 5µL/mg for subsequent experiments to ensure a sufficiently high protein concentration for optimum downstream labelling and 2D-DIGE.

2.1.5 Protein extraction of endometrial tissue lysates

Protein extraction was carried out as per the optimised protocol above. Each snap-frozen tissue sample was weighed and homogenised by grinding in liquid nitrogen into a fine powder. Ground tissue was then lysed using 2D lysis buffer (8M urea, 4% w/v CHAPS and 10mM Tris-HCL pH 8.3) at a ratio of 5 μL of lysis buffer to 1 mg of tissue at room temperature in order to extract the proteins. Sonication was carried out to ensure maximum tissue breakdown and protein extraction. Samples were then centrifuged at 3,000 g for 10 min at 4°C to pellet any insoluble material. The supernatant was separated for further protein analysis and the pellet discarded. Protein concentration was determined using a Bradford microtitre plate assay (Thermo Scientific Ltd). Standard curves were constructed using dilutions of bovine serum albumin (BSA). All samples were then normalised to the same (lowest) protein concentration using lysis buffer.

2.1.6 Quality assessment of endometrial tissue lysates

To assess the quality of tissue samples 10 μg protein from each sample was mixed with 5x sample buffer (300 mM Tris-HCL, pH 6.8, 10% SDS, 30% glycerol, 0.025% bromophenol blue and 1.3 v/v β-mercaptoethanol) and loaded onto a NuPAGE® Novex® 4-12% Bis-Tris 1.5 mm, 15-well pre-cast gel (Invitrogen). The gels were run at 70 mA for 1 hour and stained with Instant Blue gel stain (Expedion). The stained gels were imaged on a GS-800™ densitometer (BioRad)

2.1.7 Sample pooling

A sample pooling strategy was used for 2D-DIGE analysis. Protein extracted from each sample was pooled in equal amounts into six sample groups created as indicated in Table 2.3 below. There were insufficient samples in the pain-proliferative stage (PP) to warrant inclusion in the analysis.

Table 2.3 Sample groups used for 2D-DIGE analysis

Clinical group	Cycle phase	Tissue type	Pool	No. of samples per pool
Healthy controls	Secretory	Eutopic endometrium	CS	4
Healthy controls	Proliferative	Eutopic endometrium	CP	7
Pain	Secretory	Eutopic endometrium	PS	7
Endometriosis	Secretory	Ectopic endometrium	EcS	12
Endometriosis	Secretory	Eutopic endometrium	ES	19
Endometriosis	Proliferative	Eutopic endometrium	EP	9

2.2 Endometrial tissue profiling by 2D-DIGE

2.2.1 Introduction

Two-dimensional difference gel electrophoresis (2D-DIGE) was developed to overcome the inherent problem of poor reproducibility of 2DE. 2D-DIGE relies on direct labelling of lysine groups on proteins with fluorescent CyDyes before 2D separation. An important aspect of the use of 2D-DIGE technology is the ability to label 3 samples with 3 distinct CyDyes and electrophorese the samples on the same 2D gel. This reduces spot pattern variability and the number of gels that need to be run and facilitates accurate spot matching. The CyDye DIGE fluor minimal dye has an N-succinimidyl ester reactive group. It is designed to form a covalent bond with the ϵ-amino group of lysines and N-terminal residues in proteins via an amide linkage. The single positive charge of the CyDye replaces the single positive charge of the lysine at neutral and acidic pH minimising changes to the pI of the protein. The labelling reaction is dye limiting with the ratio of CyDye to protein ensuring that the dyes label approximately 1-2% of lysine residues. Therefore, each labelled protein molecule carries around one dye label and will thus have a minimal effect on the molecular weight meaning that the labelled protein will co-migrate with the unlabelled protein and run as a single spot. This minimal labelling is also performed to prevent

protein precipitation due to increased hydrophobicity with the addition of dye moieties. The ratio has been optimised to detect less abundant proteins whilst at the same time keeping the highly abundant proteins in the linear dynamic range for quantitative image analysis. The labelling reaction is quenched by addition of lysine. The use of a Cy2-labelled internal standard pool containing equal amounts of all samples and run on every gel against the Cy3- and Cy5-labelled test samples facilitates spot matching across gels and improves quantitative accuracy. The individual protein data from the controls and the diseased group samples are normalised against the Cy2 dye labelled standard.

Herein, 2D-DIGE was used as a profiling approach for discovery of markers of endometriosis from the tissue samples. A preliminary experiment on test samples showed the feasibility of the method for this sample type. In the main experiment, 2D-DIGE technology was coupled to LC-MS/MS to identify and quantify differentially expressed proteins in endometriosis and control tissue lysates that could be later tested in serum as potential biomarkers of the disease.

2.2.2 Preliminary 2D-DIGE analysis of endometrial tissue lysates

A preliminary analysis was carried out to test the feasibility of the method in profiling endometrial tissue lysates. A method was therefore optimised for lysate pooling, CyDye labelling and 2D electrophoresis using the previously described test samples (Table 2.1). Samples (100 µg total protein) were labelled using Cy3 or Cy5. An internal standard comprising equal amounts of all samples was labelled using Cy2. These samples were mixed together appropriately and separated on four 2D gels according to protein isoelectric point in the first dimension and by relative molecular weight in the second dimension (Table 2.4).

Table 2.4 Experimental design for preliminary 2D-DIGE analysis.

.

Gel	Cy3	Cy5	Cy2
1	001Eu	002Eu	Pool
2	001E1	002E1	Pool
3	001S3	002S1	Pool
4	001S2	001S1	Pool

Resolved and labelled proteins were then detected at appropriate excitation and emission wavelengths using a Typhoon 9400™ Imager and the signals compared. Automated spot detection and normalisation was performed and abundance ratios (Cy3/Cy2 and Cy5/Cy2) compared for each gel using the Differential In-gel Analysis (DIA) module of the DeCyder software. Since each sample spot map is co-detected with a standard spot map, all of the spots were compared internally to the same pooled standard.

2.2.3 Optimisation of high-abundant protein depletion in tissue lysates

Protein concentrations in biological samples may vary over 10 orders of magnitude. Low abundant protein detection is hampered by the presence of those proteins present at very high concentrations. This was a particular issue with the tissue lysates used herein which were significantly contaminated with blood. Highly abundant proteins often prevent optimal focusing, limit the loading capacity of lower abundance proteins and tend to mask considerable areas on 2D gels. Moreover, in data-dependent LC-MS/MS experiments, peptides from high abundant proteins will be preferentially sampled at the cost of reduced proteomics coverage. Immuno-depletion is one method to reduce the most abundant proteins in a sample to alleviate these problems. Depletion of high abundant serum proteins in tissue lysates was carried out using Protein Purify 12 (PP12) Human Serum Protein Immunodepletion resin (R&D) which depletes the 12 most abundant serum proteins. 500 µL of PP12 immunodepletion resin was incubated with diluted endometrial tissue lysate at two different protein amounts (200 µg and 500 µg). Samples were incubated and allowed to mix on a rotary shaker

for 30 min. Equal amounts of the mixture were then pipetted onto the upper chamber of three Spin-X Filter Units, centrifuged for 2 min, washed with PBS and centrifuged again. The filtrates were then concentrated to 25 μL using 5 kDa Molecular Weight Cut-Off (MWCO) Vivaspin Columns. Protein concentrations were then determined using the Bradford method. Depleted samples were run on a 1D gel against undepleted samples and neat serum to examine the level of depletion.

2.2.4 2D-DIGE profiling of endometrial tissue lysates

Protein extraction and sample immunodepletion of pooled samples (for 2D-DIGE) was carried out according to an optimised protocol. 1.5 mg protein from each pool was depleted using 1.5 mL of PP12 immunodepletion resin (R&D). Samples were first diluted 8-fold using PBS in 10 mL polypropylene columns (Thermo Scientific) and then incubated with the resin on a rotary shaker for 1 hr at room temperature. The flow-through was transferred to 20 mL 5 kDa MWCO filtration units and concentrated to 100 μL. To remove salt from the samples, a buffer exchange was carried out by diluting the samples to 1 mL with 10 mM Tris pH 8.5 and the samples then re-concentrated to 25 μL. Samples were diluted again to 500 μL and finally concentrated to 100 μL. Samples were then dried down in a Speed Vac and re-suspended in 2D lysis buffer (8M Urea, 4% CHAPS, 1.5 mM Tris pH 8.5). The protein concentration of each pool was determined using the Bradford method. All pools were adjusted to the same protein concentration.

2.2.5 Protein labelling with CyDyes for 2D-DIGE

80 μg protein from each pool was labelled in triplicate with NHS-cyanine dyes (Cy3, Cy5 and Cy2) (GE Healthcare) at a dye to protein ratio of 6 pmol of dye per microgram of protein on ice for 30 min in the dark. Cy2 was used to label an internal standard pool which was prepared by mixing equal amounts of proteins from each pool (Table 2.5). Reactions were quenched by adding a 20-fold molar excess of L-lysine to dye and incubating on ice for 10 min in the dark. Equal amounts of proteins from pairs of differentially labelled (Cy3 and Cy5) samples and Cy2 labelled standard were mixed

appropriately and reduced by addition of dithiothreitol (DTT) (65 mM final concentration). Carrier ampholines and pharmalyte mixture was added to a final concentration of 2% v/v and 1 µL of 2% bromophenol blue. The final volume was made up to 450 µL with 2D lysis buffer containing 65 mM DTT. Immobiline IPG strips (24 cm; pH 3-10 NL) (GE Healthcare) were rehydrated with the labelled samples overnight on a re-swelling tray overlaid with mineral oil.

Table 2.5 2D-DIGE tissue profiling: experimental design.

Gel No.	Cy3	Cy5	Cy2
1	EEuS	PEuS	Pool
2	EEcS	EEuS	Pool
3	EEuS	CEuS	Pool
4	EEuP	CEuP	Pool
5	CEuS	EEcS	Pool
6	PEuS	CEuS	Pool
7	EEuP	PEuS	Pool
8	EEcS	CEuP	Pool
9	CEuP	EEuP	Pool

Nine gels were run to compare the six different groups of pooled samples in triplicate with each gel containing the Cy2-labelled standard pool. E=endometriosis group; P=pain group; C=control group; Eu=eutopic tissue; Ec=ectopic tissue. S=secretory phase; P=proliferative phase.

2.2.6 2D gel electrophoresis and gel imaging

IEF was performed using a Multiphor II apparatus (GE Healthcare) for a total of 80 kVh at 16°C. Strips were equilibrated for 15 min in equilibration buffer (6M urea, 30% (v/v) glycerol, 50 mM Tris-HCl pH 6.8, 2% (w/v) SDS) containing 65 mM DTT and alkylated thereafter for another 15 min in the same buffer containing 240 mM iodoacetamide. For second dimension separation, equilibrated strips were transferred onto 10% SDS-PAGE gels cast between 20 x 24 cm x 1 mm low fluorescence glass plates. Gels were bonded to the inner plate at casting using bind saline (GE

Healthcare) according to the manufacturer's instructions. Strips were overlaid with 0.5 % (w/v) low melting point agarose in running buffer containing bromophenol blue. Gels were run in an Ettan Dalt 12 tank (GE Healthcare) for 16 hrs at 2.2 W per gel or until the dye front had run off the bottom of the gel. Images were acquired by scanning the gels on a Typhoon™ 9400 multi-wavelength fluorescence imager (GE Healthcare). The photomultiplier tube voltage was adjusted on each channel (Cy2, Cy3 and Cy5) for preliminary low resolution scans (1000 μm) to give maximum pixel values within 10% for each channel and below the saturation level. A final high resolution scan (100 μm) was then performed. Sequential scanning of Cy2, Cy3 and Cy5 labelled proteins was achieved using the following laser excitation/emission filters: 488/520 nm, 532/580 nm and 633/670 nm, respectively. The images were cropped in ImageQuant software (GE Healthcare) and exported into DeCyder software (GE Healthcare) for image analysis.

2.2.7 Image analysis

Images were curated and analysed using DeCyder Software V6 (GE Healthcare). This software performs the following functions: Differential In-gel Analysis (DIA) and Biological Variance Analysis (BVA). DIA assesses the differential expression of spots in the same gel by defining spot boundaries and calculating spot volume/abundance for the three channels (Cy2, Cy3 and Cy5) and performs in-gel normalization. BVA is used to match spots from multiple samples across gels, using the standard for calculating standardised abundances for all matched spots across the sample and performing statistical analysis. Differences in spot abundances between the different conditions under study were compared by applying a student t-test. Spots displaying changes in abundance were then filtered by specifying a 1.5 threshold of average fold-change with P values of <0.05. Pick lists for spots of interest were created and exported to an Ettan spot picking robot (GE Healthcare).

2.2.7 Colloidal Coomassie Blue staining

Prior to staining, gels were fixed in 35% (v/v) ethanol, 2% (v/v) phosphoric acid in distilled water. Gels were then washed for 30 min with distilled water and stained with Colloidal Coomassie Instant Blue (Expedion) overnight on a shaking platform. Stained gels were then imaged on the Typhoon 9400™ scanner using the red laser and no emission filters.

2.2.8 SYPRO-Ruby fluorescence staining

Gels were fixed overnight in 30% (v/v) methanol, 7.5% (v/v) acetic acid in distilled water. Gels were then washed for 30 min in distilled water and incubated overnight in Sypro Ruby stain (Molecular Probes) in the dark on a shaking platform. Post stained gels were washed in distilled water to remove excess stain and thereafter scanned on the Typhoon™ Scanner using the appropriate laser and emission filter.

2.2.9 Spot-picking and trypsin digestion

Automated spot picking was achieved by matching post-stained images with CyDye images using DeCyder software. A pick-list of co-ordinates was then created for spots of interest relative to a pair of reference markers fixed to the glass plates at casting. Spots of interest were then excised using an Ettan automated spot picker (GE Healthcare) with a 2 mm picking head from gels submerged in 1-2 mm of distilled water. Spots were collected in a 96 well-plate, the water drained and gel plugs stored at -20°C prior to trypsin digestion and MS analysis.

Gel plugs were transferred to MS Lo-Bind Eppendorf tubes and in-gel digestion was carried out by first washing the gel pieces three times in 200 µL of 50% acetonitrile (ACN) and then in 200 µL of 100% ACN to de-stain them. The pieces were dried in a SpeedVac for 20 min, reduced for 45 min with 200 µL of 10 mM DTT in 5 mM ammonium bicarbonate (AmBiC) pH 8.0 at 50°C. Gel pieces were then washed twice in 200 µL of 50% ACN and once in 200 µL of 100% ACN and dried in a SpeedVac. Gels were then digested overnight at 37°C with 50 ng modified trypsin (Promega) in

5 mM AmBiC. The supernatant was collected and peptides extracted twice with 200 µL of 5% trifluoroacetic acid (TFA) in 50% ACN and the supernatants pooled. Peptide extracts were vacuum dried and stored at -20°C prior to MS analysis.

2.2.10 Sample clean-up

Peptides were desalted using reversed phase C18 ZipTips™ (Merck Millipore) according to the manufacturer's instructions. Briefly, the peptides were re-suspended in 0.1% (v/v) formic acid (FA). The C18 tip was conditioned twice with elution buffer containing 50% methanol and 0.1% (v/v) TFA and equilibrated twice with washing solution containing 5% (v/v) methanol and 0.1% (v/v) TFA. Peptides were then bound onto the C18 material by carefully aspirating and dispensing the sample at least 10 times. The tip was then washed 5 times with wash solution and peptides eluted by aspirated and dispensing (10 times) with elution buffer (50% methanol and 0.1% TFA) into fresh MS Lo-Bind Eppendorf tubes. Samples were dried down in a SpeedVac prior to MS analysis.

2.2.11 Reversed-phase liquid chromatography (RPLC)

Peptides were re-suspended in 6 μL of 0.1% (v/v) FA and resolved by nano-flow capillary RPLC using an Ultimate 3000 System (Dionex Corp. Thermo Scientific Ltd). Samples were first injected onto an Acclaim PepMap 100 C18 pre-column (5 μm, 100A, 300 μm i.d x 5 mm) (Dionex Corp. Thermo Fisher Scientific Ltd) and washed for 3 min with 90% buffer A (H$_2$O + 0.1% (v/v) FA) at a flow rate of 25 µL/min. RP chromatographic separation was then performed on an Acclaim PepMap 100 C18 Nano-LC column (3 μm, 100A, 75 μm i.d x 25 cm) (Dionex Corp, Thermo Fisher Scientific Ltd) with a 60 min linear gradient of 10-50% buffer B (100% ACN + 0.1% (v/v) FA) at a flow rate of 300 nL/min.

2.2.11 Tandem Mass Spectrometry

Tandem MS/MS was performed on an LTQ-Orbitrap XL™ mass spectrometer (Thermo-Scientific Ltd) equipped with a picoview PV550 nano-electrospray ion

source (New Objective Inc.). The MS was operated in the data-dependent mode to automatically switch between MS (full ion scan) and MS/MS (fragment ion scan) acquisition. Survey full scan MS spectra (*m/z* 390-1700) were acquired in the orbitrap with a resolution of 60,000 at *m/z* 400. The most intense (top 6 ions per survey scan) were sequentially isolated for fragmentation (MS/MS) in the linear ion-trap (LTQ) by collision induced dissociation and dynamically excluded for 60 sec. To maintain accurate mass measurement, the lock mass option was enabled with polydimethylcyclosiloxane at *m/z* 455.120025 set as an internal calibrant.

2.2.12 Protein identification and quantification

Acquired mass spectra were processed using Mascot Distiller version 2.5 (Matrix Science Ltd) and searched against the Uniprot database. The following parameters and search filters were used; MS tolerance was set to +/- 10 ppm, the fragment MS/MS tolerance was set to 0.5 Da, two missed cleavages were allowed, carbamidomethylation of cysteines was set as a fixed modification, methionine oxidation, acetylation (protein N-term), deamidation (asparagine and glutamine) were set as variable modifications. MudPit scoring was enabled and peptides were required to score >20 with a Mascot significant threshold of P <0.05. Peptides were also required to be bold red and therefore were statistically the strongest assignments. At least 2 peptide matches were required for a positive protein identification. The protein identifications were then matched to specific spots in DeCyder and experimental molecular weights and pIs checked with theoretical values.

2.3 Endometrial tissue profiling using a protein based labelling and fractionation method

2.3.1 Introduction

A protein labelling and separation strategy was explored in an attempt to improve the depth of coverage in profiling the endometrial tissue proteome. Protein amino group labelling of denatured protein samples using isobaric tandem mass tags (TMT) was

coupled to multi-dimensional protein separation by strong anion exchange (SAX) chromatography and one-dimension gel electrophoresis, prior to digestion and LC-MS/MS The premise was that the heterogeneous nature of proteins would afford a better separation versus their more homogeneous peptide components and where peptides from abundant proteins would not be spread across the entire LC-MS/MS experiment, being sampled repeatedly in preference to lower abundant species. However, complete labelling of intact proteins may be challenging and blocking of trypsin sites by lysine TMT labelling would require a different digestion approach.

2.3.2 Samples used for optimisation of protein TMT labelling

Endometrial tissue samples from four patients previously collected at laparoscopy and stored at -80°C were used in this experiment (Table 2.6). Patients selected for this experiment were diagnosed with two forms of endometriosis; deep infiltrating endometriosis and ovarian endometriosis. Ectopic and eutopic endometrial tissue were collected from each patient. All patients were on GnRH treatment and were not selected for the main profiling experiments.

Table 2.6 Tissue samples used for optimisation of a protein-based profiling strategy.

Patient ID	Tissue type	Tissue description	Phase of cycle	Treatment
007	DIE	Recto-vaginal nodule	Inactive endometrium	GnRH
	Ect 2	Endometrioma		
	EmE	Endometrium		
021	DIE	Recto-vaginal nodule	Inactive endometrium	GnRH
	EmE	Endometrium		
051	DIE	Recto-vaginal nodule	Inactive endometrium	GnRH
	EmE	Endometrium		
057	Ect 2	Endometrioma	Inactive endometrium	GnRH
	EmE	Endometrium		

DIE= Deep infiltrating endometriosis; Ect = ectopic endometrial tissue; EmE= eutopic endometrium

Protein extraction was carried out according to the optimised protocol presented in the previous sections. Three pools were created by mixing equal amounts of protein from the 9 samples to create eutopic endometrial (EU), deep infiltrating endometriosis (D1) and endometrioma (E1) pools.

2.3.3 Immunodepletion and protein denaturation

Samples containing 500 µg protein were diluted with PBS (final urea concentration of 1 M) and 500 µL of PP12 immunodepletion resin was added. and incubated on a rotary shaker for 30 minutes. The flow-through was collected using Spin-X Filter Units, and then concentrated to 100 µL using 5 kDa MWCO Vivaspin filter units (GE Healthcare). Samples were diluted to 500 µL with 8M urea in 20 mM triethylammonium bicarbonate (TEAB) at pH 8.3 to denature the protein and then concentrated to 25 µL. Samples were diluted again to 500 µL as above and concentrated to 100 µL. A protein assay (Bradford) was carried out to determine protein concentration of each pool after depletion. Pools were normalised to the same concentration with 8M urea in 20 mM TEAB pH 8.3.

2.3.4 TMT protein labelling and gel-based separation

SDS (0.1% v/v final concentration) and EDTA (1 mM final concentration) were added to 100 µg protein from each of the immunodepleted pools. Samples were then reduced with 1 mM TCEP from a stock solution of 0.5 M TCEP at pH 7.0 for 1 hr at room temperature. 0.8 mg of amine-reactive Tandem Mass Tag reagents (TMT-126, TMT-127 and TMT-128) (Thermo-Scientific Ltd.) were resuspended in 41 µL of acetonitrile and used to label the pools (pool D1-TMT-126, pool E1-TMT-127, pool EU-TMT-128) for 1 hr at room temperature. Samples were then incubated with 0.25 % hydroxylamine for 30 minutes at RT to quench the reaction. TMT labelled pools were then combined in equal amounts (300 µg total protein) then mixed with 5x sample buffer and run in triplicate (100 µg per lane) on a large 10% SDS-PAGE gel at 10 mA overnight. The gel was stained with Instant Blue Coomassie stain. Each lane was cut into 48 bands and digested with Glu-C and trypsin as described below.

2.3.5 Enzymatic digestion

Excised gel pieces were washed three times in 50 % (v/v) ACN and dried in a SpeedVac. The dried pieces were then reduced in 10 mM DTT in 5 mM AmBiC, pH 8.0, for 45 minutes at 50°C and then alkylated with 50 mM iodoacetamide in 5 mM AmBic for 1 hr in the dark at room temperature. The gel pieces were then washed in 50% ACN and dried in a SpeedVac until completely dry and incubated with 100 ng Glu-C protease (Roche) at 25°C overnight. 50 ng of sequencing grade modified trypsin (Promega) in 10 µL 5 mM AmBic was then added to each of the gel pieces which were further incubated at 37°C overnight. Peptides were then extracted twice using 30 µL 50% ACN/5% TFA. The supernatant was then transferred into clean Eppendorfs and the samples dried in a SpeedVac. Dried samples were re-suspended in 10 µL of 0.1% TFA, and desalted using C18 ZipTips® (Merck Millipore) as described previously (Section 2.2.10).

2.3.6 Reverse-phase liquid chromatography tandem MS analysis

Samples were resuspended in 6 µL of 0.1% FA and analysed on an LTQ Orbitrap XL mass spectrometer (Thermo Scientific) coupled to an Ultimate 3000 nano-LC system (Dionex). The mass spectrometer was operated in the data-dependent mode to automatically switch between orbitrap MS and ion trap MS/MS acquisition. For identification of TMT labelled peptides, the three most abundant ions were selected for collision induced dissociation (CID) and then the same precursor ions triggered higher energy collision dissociation (HCD) fragmentation.

2.3.7 Optimisation of fractionation by strong anion exchange (SAX) chromatography

A method for protein fractionation by SAX before and after TMT labelling was explored as a potential separation method for improving coverage. Briefly, 150 µL of DEAE Ceramic HyperD F slurry (Pall Corporation) was prepared in a column and used to fractionate 300 µg of immunodepleted TMT labelled tissue lysate protein.

Elution buffers of increasing sodium chloride concentration in 20 mM TEAB pH 8.3 (50 mM, 100 mM, 200mM and 1M NaCl) were used to elute bound proteins from the column in a sequential fashion using 200 µL of each buffer. To resolve eluted proteins, each fraction and the flow through were run on a NuPAGE® Novex® 4-12% Bis-Tris 1.5 mm, 10 well pre-cast gel (Invitrogen). The gels were run at 70 mA for 1 hour and stained with Instant Blue gel stain (Expedeon). The stained gels were imaged on a GS-800™ densitometer (BioRad). The same fractionation was repeated without TMT labelling of proteins using 200 mM, 400 mM, 600 mM, 800 mM and 1M NaCl elution steps.

2.4 Profiling of endometrial tissues using protein TMT labelling and fractionation

Protein extraction and immunodepletion for TMT protein labelling of pooled samples was carried out according the optimised protocol (sections 2.3.3 and 2.3.4). 6-plex TMT reagents were used to label 100 µg of denatured protein from each pool as follows: pool CS-TMT126, pool CP-TMT127, pool PS-TMT128, pool EcS-TMT129, pool ES-TMT130 and pool EP-TMT131 and then combined together. TMT-labelled proteins were fractionated by strong anion exchange chromatography. A SAX column was prepared in a spin filter unit using 300 µL of DEAE ceramic HyperD F slurry (Pall Corporation). The column was centrifuged at 3,000 rpm for 2 min to remove the storage solution and then washed with 300 µL of 1 M NaCL in 20 mM TEAB pH 8.5 and centrifuged at 3,000 rpm for 2 min. The column was again washed three times with 300 µL of 200 mM TEAB pH 8.5, and then equilibrated by washing with 300 µL of 20 mM TEAB pH 8.5 followed by centrifuging at 3000 rpm for 2 min. The pooled sample was then introduced into the column and incubated with the resin on a rotary shaker for 5 min. Un-bound proteins were removed by centrifuging at 3,000 rpm for 2 min followed by washing with 300 µL of 20 mM TEAB. The two fractions were combined as the flow through (FT). Bound proteins were then sequentially eluted two times using 300 µL each of increasing salt concentration buffers of 400 mM NaCl, 600 mM NaCl and 1M NaCl. The two eluates were combined for each step. The flow through and three eluates were concentrated to 50 µL each using 5 kDa MWCO

Vivaspin filter devices (GE Healthcare). The fractions were then run on a large format hand-cast 10% SDS-PAGE gel to resolve the eluted proteins. The gel was stained using InstantBlue Coomassie stain (Expedeon) and each lane cut into 50 bands and digested with Glu-C (Roche) then trypsin (Promega) as described in section 2.3.5. Samples were desalted using C18 ZipTips as before, dried down and re-suspended in 6 μL 0.1% FA prior to LC-MS/MS analysis.

Samples were analysed in an LTQ Orbitrap XL mass spectrometer (Thermo Scientific) equipped with a picoview PV550 nano-electrospray ion source (New Objective Inc.) and coupled to an Ultimate 3000 nano-liquid chromatography system (Dionex). Samples were injected onto an Acclaim PepMap 100 C18 pre-column (5 μm, 100A, 300 μm i.d x 5 mm) (Dionex Corp. Thermo Fisher Scientific Ltd) and washed for 3 min with 10% buffer B (ACN + 0.1% (v/v) FA) at a flow rate of 25 μL/min. RP chromatographic separation was performed on an Acclaim PepMap 100 C18 Nano-LC column (3 μm, 100A, 75 μm i.d x 25 cm) (Dionex Corp, Thermo Fisher Scientific Ltd) with a 90 min linear gradient of 10-50% buffer B at a flow rate of 300 nL/min. The instrument was operated in the data-dependent mode to automatically switch between MS (full ion scan) and MS/MS (fragment ion scan) acquisition. Survey full scan MS spectra (*m/z* 390-1700) were acquired in the orbitrap with a resolution of 60,000 at *m/z*.

For identification of TMT labelled peptides, the most intense (top 3 per survey scan) were sequentially isolated for fragmentation in the linear ion-trap by CID and dynamically excluded for 60 sec. Wide band activation was enabled to account for potential neutral loss of water. The same ions were also fragmented by HCD for detection of TMT reporter ions in the orbitrap using a collision energy of 40%. To maintain accurate mass measurement, the lock mass options was enabled and the polydimethylcyclosiloxane at *m/z* 455.120025 was set as an internal calibrate.

Raw data files produced in Xcalibur software (Thermo Scientific) were processed using Mascot Distiller version 2.4 searching against the SwissProt database. For searching, taxonomy was human; the MS tolerance was set to +/-10 ppm and the

MS/MS tolerance was set to 0.5 Da. Enzyme specificity was set to trypsin and Glu-C. Two missed cleavages were allowed. TMT 6-plex modification of peptides and carbamidomethylation of cysteines were set as fixed modifications. Protein N-terminal acetylation, methionine oxidation and deamidation (N/Q) were set as variable modifications. Search result filters were as follows: only peptides with a score of >20 and below the Mascot significance threshold filter of $P<0.05$ were included. The extent of labelling was calculated by comparing the numbers of labelled versus unlabelled peptides. Protein grouping was enabled such that when a set of peptides in one protein were equal to or completely contained within the set of peptides of another protein, the two proteins were put together in a protein group. Quantitative information was calculated from reporter ion intensities taking the median value of peptide ratios to calculate protein group ratios for the following comparisons; ES/CS, ES/PS, EP/CP, EcS/ES and PS/CS.

2.5 Endometrial tissue profiling using peptide-based labelling and fractionation

2.5.1 Introduction

This work also explored a peptide TMT labelling and fractionation approach for profiling endometrial tissues for biomarker discovery. Three-dimensional separation of peptides was employed to maximise coverage by MS analysis, using SAX chromatography, off-line high pH RPLC and low pH nano-RPLC coupled to tandem mass spectrometry. The strategy was applied to the 6 tissue lysate pools described in Table 2.3.

2.5.2 Enzyme digestion and TMT labelling of immunodepleted pools

For peptide based profiling, 100 μg of each immunodepleted tissue lysate pool was resuspended in 100 mM TEAB, pH 8.5 and 0.1% SDS. Samples were reduced in 1 mM TCEP for 1 hr at 55°C and alkylated in 7.5 mM iodoacetamide for 1 hr at RT in the dark. Samples were then digested overnight at 37°C using 4 μg of modified porcine trypsin (Promega).

Samples were labelled with TMT reagents. Briefly 100 µg of each digested sample was labelled with 0.8 mg of each TMT reagent (resuspended in 41 µL of ACN) as follows: pool CS -TMT126, pool CP - TMT127, pool PS - TMT128, pool EcS - TMT129, pool ES - TMT130 and pool EP - TMT131 for 1 hr at RT. Samples were then incubated with 0.25% hydroxylamine for 30 minutes at RT to quench the reaction. The six TMT labelled samples were then combined together.

2.5.3 Sample clean-up

SDS was removed using detergent removal spin columns (Pierce). Briefly, six spin columns were prepared by first centrifuging for 1 min at 1,500 x g to remove storage buffer, followed by equilibration using 400 µL of PBS. This was done three times and the supernatants discarded. 100 µL of sample was then loaded onto each column, incubated for 2 min at RT and collected by centrifugation for 2 min at 1,500 x g into fresh Eppendorf tubes. The samples were recombined and dried down in a SpeedVac. The dried sample was then re-suspended in 1 mL acidified water containing 2% phosphoric acid. Samples were then desalted using 1cc Oasis HLB cartridges (Waters). The cartridge was first conditioned using 1 mL of 100% methanol followed by equilibration using 1 mL of distilled water. 1 mL of sample was then loaded onto the cartridge and washed using 1 mL of 5% methanol in water. Bound peptides were then eluted using 1 mL of 100% methanol. The sample was then dried down in SpeedVac and re-suspended in 300 µL of 100 mM TEAB pH 8.5 for SAX chromatography.

2.5.4 Strong anion exchange chromatography of labelled peptides

TMT-labelled peptides were fractionated by SAX chromatography. The SAX column was prepared in a spin filter unit using 300 µL of DEAE ceramic HyperD F slurry (Pall Corporation). The column was centrifuged at 3,000 rpm for 2 min to remove storage solution and then washed with 200 µL of 1 M NaCl in 100 mM TEAB pH 8.5 and centrifuged at 3,000 rpm for 2 min to remove the supernatant. The column was again washed three times with 200 µL of 200 mM TEAB pH 8.5, and then equilibrated

by washing with 100 mM TEAB pH 8.5 followed by centrifuging at 3,000 rpm for 2 min to remove the supernatant. The pooled sample was then introduced into the column and incubated with the resin on a rotary shaker for 5 min. Un-bound peptides were removed by centrifuging at 3,000 rpm for 2 min followed by washing with 200 µL of 100 mM TEAB. The two fractions were combined as the flow-through (FT). Bound peptides were then sequentially eluted using 200 µL twice of increasing salt concentration buffer; 100 mM TEAB plus 400 mM NaCl, 600 mM NaCl and 1M NaCl. The two eluates at each salt concentration were combined and desalted using 1cc Oasis HLB cartridges (Waters), lyophilised and stored at -20°C.

2.5.5 High pH reversed-phase liquid chromatography

Peptide fractions were resuspended in 42 µL of 20 mM ammonium formate pH 8.5 and separated by high pH RPLC using an Agilent 1100 series microflow pump. Briefly, 40 µL of resuspended peptides were injected onto a Poroshell 300SB-C18 (5 µm, 2.1 x 75 mm) column (Agilent). Solvents consisted of 20 mM ammonium formate pH 8.5 in water as mobile phase A and 20 mM ammonium formate pH 8.5 in 80% acetonitrile as mobile phase B. Peptide separation was accomplished using a linear gradient of 0-45% mobile phase B at a flow rate of 200 µL/min for 55 min. 30 fractions were collected for each SAX fraction (total 120 fractions). Each fraction was dried down in a SpeedVac and resuspended in 200 µL of 0.1% FA and redried prior to MS analysis.

2.5.6 LC-MS/MS

Fractions were resuspended in 6 µL 0.1% FA prior to tandem MS. Tandem MS/MS was performed on an LTQ-Orbitrap XL™ mass spectrometer (Thermo-Scientific Ltd) with the same parameters as described in section 2.4 with data-dependent acquisition of the top 3 ions by CID and HCD for optimal reporter ion measurement. The orbitrap was operated using Xcalibur software and parameters were optimised for acquisition of high quality spectra for identification and quantification of TMT-tagged peptides. Fragmentation was optimal with a HCD normalised collision energy set at 40%.

Product ions were detected in the orbitrap at a resolution of 7,500. This resulted in prominent reporter ions in HCD MS/MS spectra.

Proteome Discoverer version 2.4 was used for protein identification and quantification using Mascot for searching the SwissProt database. Search parameters were as described above, except that only trypsin was used as the enzyme and one missed cleavage was allowed. A co-isolation threshold of 25% was set in the quantification method to limit the recording of reporter ion ratios from multiple peptides. Protein groups with a ratio above 1.5 or below 0.67 in each comparison were considered to be differentially expressed. Proteins were also assigned a 'biomarker score' based on fold-change, ratio count, variability, number of unique peptide sequences, whether they were a possible serum protein contaminant and their membership of a particular expression cluster group, each of which represented the broad pattern of expression across the tissue groups; Graphical Proteomics Data Explorer (GProX) was used to cluster the proteins based on their reporter ion ratios across the different comparisons.

2.6 Gene ontology and pathway enrichment analysis

Gene ontology and pathway enrichment analysis of the changing proteins was carried out with the aim of defining potential biological and molecular networks involved in endometriosis. Gene ontology is a framework that describes gene products in terms of their molecular and biological functions and cellular localisations. GO Slim is a summarised version of the GO ontology and was used to provide a broader classification of gene product functions. Differentially expressed proteins ($\geq$1.5/$\leq$0.67-fold) were imported into WebGestalt (Wang J et al., 2013) and each clinical group analysed separately for enrichment of GO biological process, molecular function and cellular component, GO Slim terms, protein interaction networks, KEGG pathways and disease association. Significantly enriched terms were identified using a hypergeometric test with a Benjamini-Hochberg (BH) correction at a significance of P <0.05. The top 10 GO terms with the most significant P values were reported. In each comparison, the protein lists were analysed separately as up- or down-regulated proteins.

2.7 Verification of proteins identified from proteomic profiling

2.7.1 Introduction

The main goal of biomarker verification is to determine whether there is adequate and consistent evidence for the potential clinical application of a given candidate to warrant further validation studies. In this study, univariate tests of significance and a 'biomarker score' were used to filter the discovery data for the selection of proteins that displayed potential as endometriosis biomarker candidates. These candidates as well as markers identified from previous reports in the literature were verified in individual serum samples from the discovery set using commercial ELISA kits. Data analysis involved significance testing and Receiver Operating Characteristic (ROC) curve analysis of single markers in the different clinical groups irrespective of cycle phase, as well as in the different phases of the cycle. Candidates were also correlated with measurements of progesterone, oestradiol, CRP and CA125, clinico-pathological features and epidemiological data obtained in the course of the study. Multivariate tests were also conducted by incorporating multiple markers into models to assess performance. Models showing the highest diagnostic accuracies would be selected for further validation in an independent cohort.

2.7.2 Assay optimisation and preliminary analysis

Commercial ELISA kits purchased from R&D Systems were used to assay follistatin (FST), soluble intercellular adhesion molecule-1 (sICAM1), macrophage chemoattractant protein-1 (MCP1), macrophage inhibitory factor (MIF), soluble interleukin-1 receptor type 2 (IL1R2) and vascular endothelial growth factor (VEGF). The kits used to assay lumican (LUM), glycodelin (PAEP), tropomyosin beta chain (TPM2), tenascin C (TNC) and carboxypeptidase M (CPM) were purchased from Boster, Bioserve Diagnostics, Abcam and Clone Corp, respectively (Table 2.7). The assay kits were first tested for their reproducibility and technical sensitivity to ensure the particular protein of interest could be accurately detected in serum samples before using them on the full sample set. Trial runs were performed with a series of serum

dilutions coupled with various diluting reagents and incubation times to determine the most distinguishable signal on the linear portion of the standard curve. A serum pool for testing was prepared by mixing 20 µL of each serum sample from the discovery set. Different dilutions of this serum pool were optimised by testing each dilution in duplicate alongside the standards in the different assays. Each assay was calibrated by running a standard curve using a series of standards according to the manufacturer's instructions. Intra-assay CVs for the different assays are reported in the table below.

Table 2.7 ELISA assays, optimal dilutions and intra-assay CVs.

Assay	Company	Optimal dilution	Intra-assay CV%
CPM	Clone Corp	1:20	5
FST	R&D	1:2	5.6
sICAM	R&D	1:50	2
PAEP	Bioserv	1:5	13.6
IL1R-II	R&D	1:50	1.5
LUM	Boster	1:20	10.6
MCP1	R&D	1:10	5.5
MIF	R&D	1:10	4.5
TNC	Abcam	1:50	0.7
VEGF	R&D	1:2	4.6
CA125	Roche	1:1	4

2.7.4 Proseek Proximity Extension Assay (PEA) platform

PEA technology is a high-throughput immunoassay that was used to analyse a panel of 92 oncology-related protein biomarkers across a subset of discovery samples; 15 randomly selected serum samples each from the endometriosis and pain groups were assessed using the Proseek® Multiplex Oncology II panel (Olink Biosciences). Briefly, target-specific antibody pairs are linked to a single strand DNA oligonucleotide. When the two probes are in close proximity due antibody pair binding to their target, a PCR target sequence is formed by ligation of the two DNA probes. The resulting sequence is subsequently detected and quantified using real-time QPCR.

A normalised protein expression value was provided by reference to controls added to the samples.

2.8 Statistical analysis

Data analysis was carried using MS Excel 2013, Graphpad Prism software version 5.0.1 and the R statistical software environment. Where the protein of interest was undetectable in the sample, the adjusted concentration of this protein was considered to be one-half the limit of quantification of the particular assay. For baseline characteristics, the D'Agostini and Pearson omnibus test was applied to test for normality. The unpaired t-test was chosen to compare mean values for normally distributed data, whereas for non-normally distributed data, the mean values were compared by the Mann-Whitney test. For each clinical group, the data was analysed independently of cycle phase and then according to cycle phase to assess whether any of the markers were dependent on the menstrual cycle. A *P* value of <0.05 was considered significant. To determine the diagnostic performance of each biomarker, a ROC curve analysis was performed with the area under the curve (AUC) reported for each comparison (control vs endometriosis and pain vs endometriosis). Markers were also analysed for correlations to clinico-pathological parameters and to each other, reporting the Spearman correlation coefficients. It is highly likely that a panel of markers as opposed to single markers would increase the sensitivity and specificity of a non-invasive test of endometriosis. A multi-variate analysis was therefore carried out to assess the performance of combination of these biomarkers. A script in R studio was applied to combine 2 or 3 candidate markers into models, testing all combinations and reporting the best models based on AUC. The sensitivities at 80% and 90% specificity were reported.

CHAPTER 3: PATIENT RECRUITEMENT AND SAMPLE COLLECTION CHALLENGES

3.1 Introduction

Patients and samples were sourced from the UCLH Reproductive Medicine Unit (RMU) following ethical approval. Initially samples were to be collected from 80 patients who were to be selected from the total number of consenting patients who were recruited for the study. They were to be stratified as follows: 20 patients diagnosed with endometriosis, 20 patients diagnosed with chronic pelvic inflammatory disease, 20 patients diagnosed with pelvic pain of unknown cause and 20 control patients with no known disease at the time of surgery undergoing bilateral tubal ligation (BTL) or prophylactic bilateral salpingo-oophorectomy (BSO) and all in the secretory phase of the menstrual cycle. However, these criteria were later reviewed due to a lower than expected number of consenting patients. It was also not possible to obtain samples only from women in the secretory phase due to surgical scheduling and lower than expected consent rates. There were no consenting patients undergoing BTL so all healthy controls had been referred for BSO as a prophylactic measure due to a high familial risk of breast and ovarian cancer. Patients were then stratified according to menstrual cycle phase as controls with no known disease in the secretory phase (n=9) and proliferative phase (n=11). The pain group also had to be combined to include women with chronic pelvic pain of unknown cause and women with chronic pelvic inflammatory disease (n=12). All patients selected for the pain group were in the secretory phase. In the endometriosis group, eutopic endometrial tissue as well as superficial peritoneal lesions, ovarian endometriomas and deep infiltrating endometriosis ectopic tissue were collected from each patient diagnosed with endometriosis. Samples from patients in this group were stratified according to menstrual cycle phase as eutopic endometrium and the corresponding ectopic tissue in the proliferative (n=14) and secretory phases (n=28). Superficial implants and deep infiltrating endometriosis ectopic tissue samples only were used in this study. The rationale being that imaging techniques are sensitive for diagnosis of ovarian

endometriomas, but are less accurate for diagnosing endometriosis elsewhere in the pelvis.

Biomarker discovery ideally involves the analysis of a larger number of different samples. However patient recruitment and the accrual of enough tissue samples proved to be a challenge. Particularly many of the women with advanced stage endometriosis were on some form of treatment at the time of laparoscopy (mostly GnRH analogues or oral contraceptives). This meant that eutopic endometrial biopsies were compromised as the endometrium is highly down-regulated by these treatments. It is also important to note that in some instances it was not possible to obtain ectopic samples, especially from patients with subtle peritoneal disease. These endometriosis lesions were ablated and not excised during laparoscopy. Finally, to avoid too much variability in the sample set, patients diagnosed with endometriosis alongside other benign conditions such as fibroids, non-endometriosis cysts and cancer were excluded. Figure 3.1 below shows a flow chart of the recruitment and patient journey through the study.
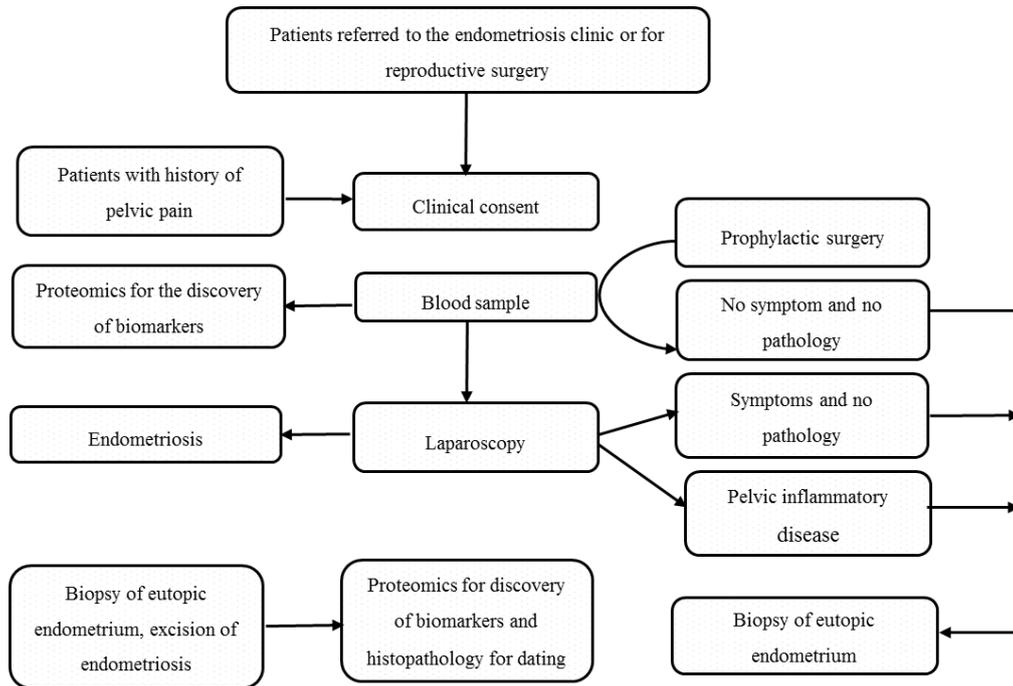
Figure 3.1 Recruitment and flow of participants through the study in the discovery phase.

In total, 122 pre-menopausal women were consented to this study. 109 women met the inclusion criteria. This set was divided into cases who were women diagnosed with endometriosis (n=62). The control group comprised of women with pelvic pain (n=24) and women with no known disease and no pain (n=23). Table 3.1 shows the characteristics of subjects included in the study and the clinical grouping. The remaining women were diagnosed with other pelvic conditions, mostly fibroids and non-endometrial ovarian cysts.

Table 3.1 Characteristics of subjects included in the study and clinical grouping

	Control	Pain	Endometriosis
Number	23	24	62
Age (years)			
Mean (SD)	37.61 (8.16)	31.83 (6.67)	35.52 (6.38)
Median (range)	36 (20-51)	32 (19-48)	35 (20-52)
Symptoms			
Subfertility	4	9	18
Pain	0	24	55
Cycle phase			
Proliferative	11	11	14
Secretory	9	12	28
Inactive	3	1	20
Treatment			
GnRH	0	1	11
OCP	0	1	6

For discovery profiling, tissue samples from patients were stratified into six groups based on condition and cycle phase with 6-20 tissue samples pooled into each group.

3.2 Sample processing optimisation

Conditions for extraction, separation and identification of the proteins extracted from endometrial and endometriosis tissue were optimised. The best extraction was achieved by combining mechanical homogenisation through grinding in liquid nitrogen followed by chemical lysis and extraction in buffered urea/CHAPS.

3.3 Quality control assessment of endometrial tissue lysates

Endometrial tissue is a highly heterogeneous tissue containing different cell types. Due to the 'bloody' nature of the post-surgical tissue specimens obtained, tissue samples were washed using PBS, dried, placed in clean Eppendorf tubes before snap-freezing in liquid nitrogen. This step was not sufficient to remove all the blood in the samples. It was therefore our aim to assess the extent of this blood 'contamination' prior to sample analysis and also to examine sample integrity. Thus 10 µg of protein

of the lysates selected for the main profiling experiments were run on 1D gels and stained to assess the extent of contamination and general protein integrity. The gel images below (Figure 3.2) illustrate the presence of albumin (blue arrows) in almost all samples and of red blood cell haemoglobin (yellow arrows). Several samples also appeared to be degraded or contained low levels of higher molecular weight proteins. Thus these samples (annotated with an asterisk in Figure 3.2) were excluded from the main profiling experiments. Additionally, the protein yields for some samples was deemed too low for subsequent downstream applications. This quality assessment thus confirmed the heterogeneity of the samples and differential contamination with blood proteins.

For proteomic profiling, a pooling approach for the individual tissue lysates was used. The advantages of pooling are that protein expression in a pool matches the mean expression of the individual samples making up the pool and reduces the number of samples for analysis. Pooling is also ideal where limited amounts of material per biological sample are available. However, individual biological variation cannot be accurately assessed and subsequent verification of changes in individual samples becomes paramount. In conclusion, patients were recruited and a set of samples accrued. These samples were assessed for their quality and the most homogeneous samples selected for pooling prior to proteomic profiling.

117
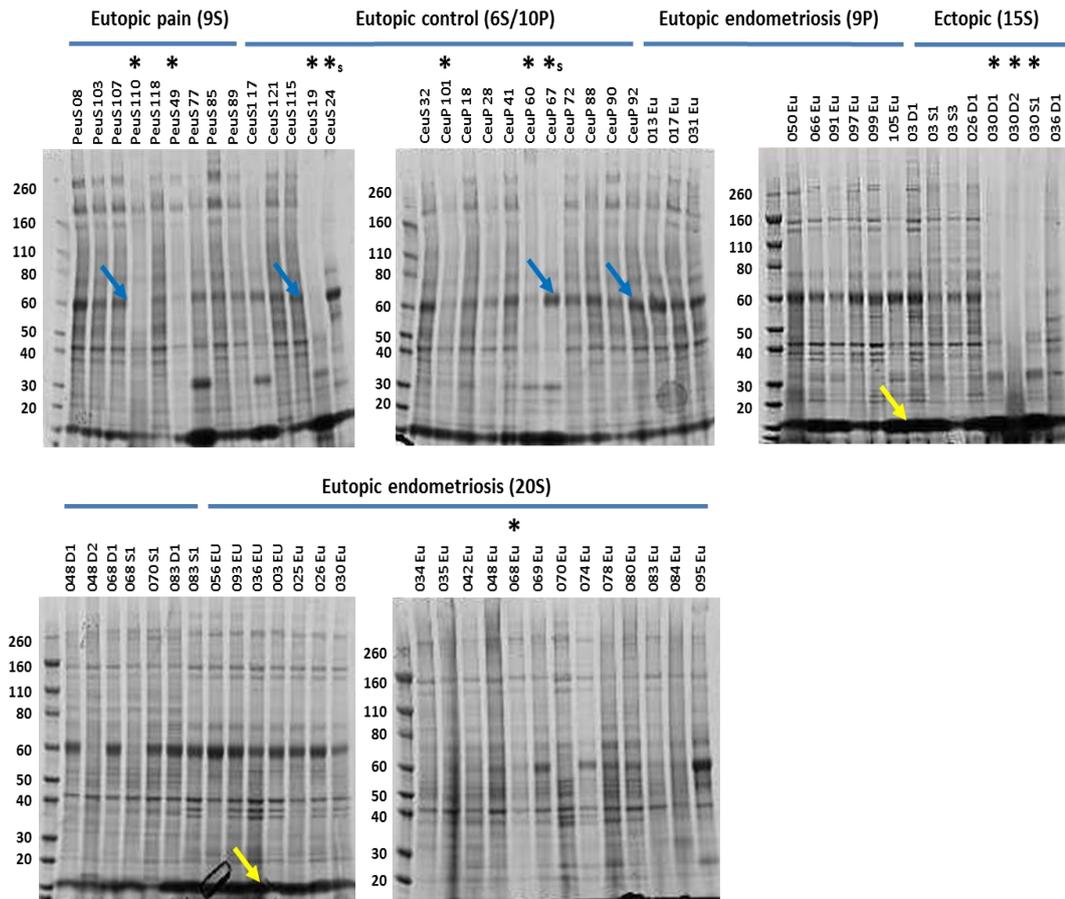
Figure 3.2 1D gel images of tissue sample lysates.

10 µg of undepleted protein from each sample was run. * indicates samples excluded from main profiling experiments due to low protein amount. Most samples were contaminated by serum album (blue arrows). Yellow arrows show haemoglobin, an indication of red blood cell contamination.

CHAPTER 4: 2D-DIGE PROFILING OF ENDOMETRIAL TISSUE FOR BIOMARKER DISCOVERY

4.1 Introduction

This chapter reports the use of two-dimensional difference in-gel electrophoresis (2D-DIGE) for the profiling of endometrial tissues collected as part of the study. The method was used to compare intact protein expression between endometriosis and control tissue samples to identify potential tissue biomarkers.

4.2 Preliminary 2D-DIGE analysis

A preliminary 2D-DIGE analysis was carried out in order to assess the feasibility of the technique in the profiling of endometrial and endometriosis tissue samples. It was possible to identify multiple protein spots from these tissue lysates as indicated by the gel images in Figure 4.1. Gels were loaded with 100 µg of differentially labelled protein representing pairs of samples. Generally, fewer protein spots were identified from ectopic tissue compared to eutopic tissue, although all samples had broadly similar patterns. The gel images produced were compared to gel images from serum samples. It was apparent that high-abundance serum proteins were present in the samples, possibly masking some lower abundance tissue-derived proteins and effectively reducing the protein load. High abundance and multiple isoforms of proteins such as albumin also complicated spot detection. It was therefore decided that an immunodepletion step should be incorporated into the main workflow in order to remove some of these highly abundant serum proteins to improve the sample load of tissue proteins and proteomic coverage.

Gel 1 (2066)

Cy3-001Eu   Cy5-002Eu

Gel 2 (1325)

Cy3-001E1   Cy5-002E1

Gel 3 (1425)

Cy3-001S3   Cy5-002S1

Gel 4 (1381)

Cy3-001S2   Cy5-001S1
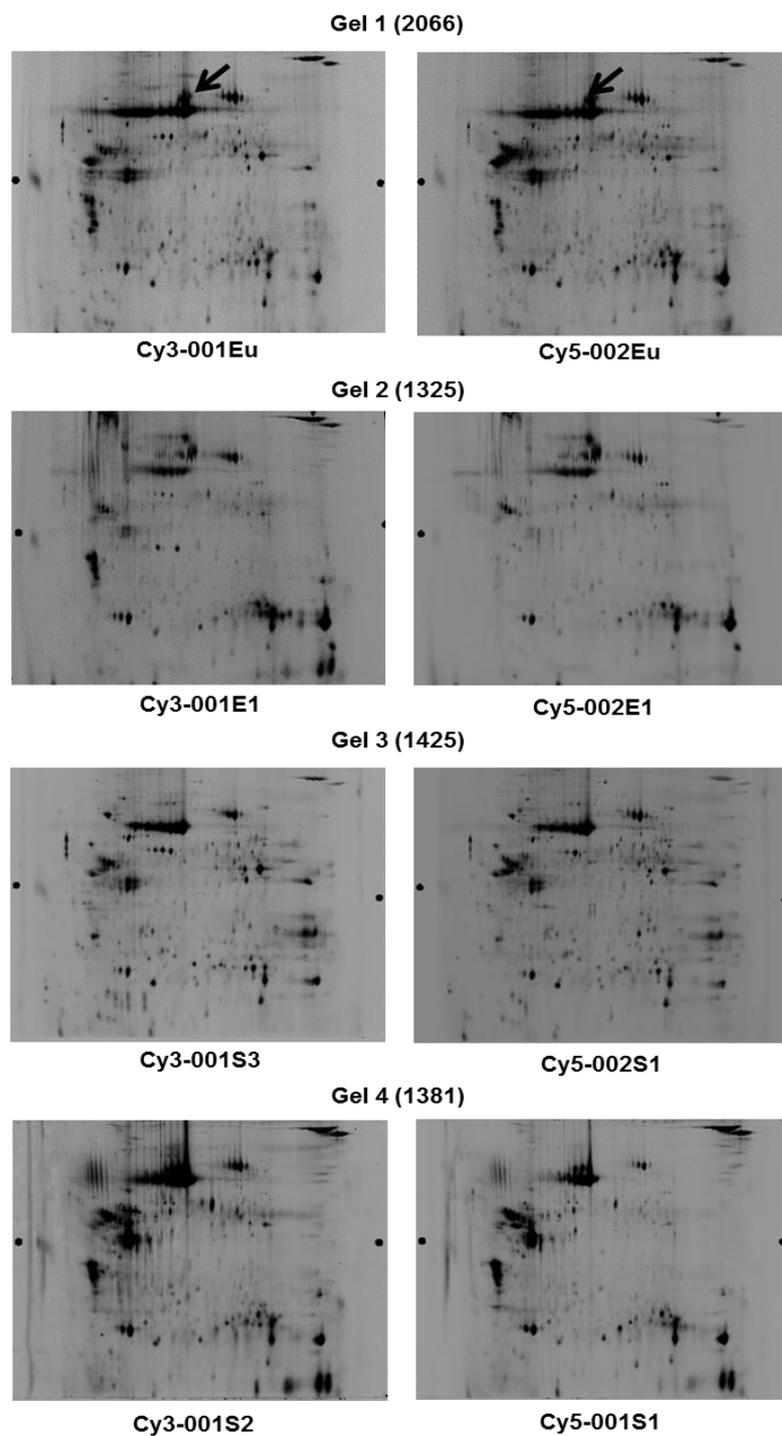
Figure 4.1 2D-DIGE images of test samples.
The number in parentheses next to the gel number reports the number of detected spots in
that gel. Arrows on gel 1 indicate the contaminating serum protein albumin

4.3 Preliminary analysis of high-abundance serum protein depletion

Sample depletion was carried out in order to remove some of the high abundant serum proteins revealed by 2D-DIGE. Curettage, which involves 'scraping' strips of endometrium, was used to collect endometrial biopsies during laparoscopy. As a consequence, most samples will have some level of blood contamination. To remove excess blood, tissue samples were washed in PBS prior to snap freezing, but this did not remove all of the blood. Protein Purify 12 (PP12) immunodepletion resin can be used to remove ~90% of the most abundant proteins from serum/plasma namely, albumin, IgG, IgA, IgM, transferrin, haptoglobin, fibrinogen, apolipoproteins AI and AII, α2-macroglobulin, α-1-antitrypsin and α1-acid glycoprotein. These proteins represent approximately 95% of the protein concentration in serum. Their removal would allow visualisation of other proteins that co-migrate or co-fractionate with the high-abundance proteins and allow increased loading of lower abundant proteins. With this in mind, 500 µL of PP12 slurry was used to deplete 200 µg and 500 µg of tissue protein extract. Flow-through depleted samples were then concentrated and 20 µg of the depleted samples were run on a 1D SDS-PAGE gel alongside 20 µg of undepleted sample and 20 µg of undepleted serum for comparison. Protein bands were visualised using colloidal Coomassie brilliant blue staining (Figure 4.2).

The image supports the notion that there are major contaminating proteins from blood in the tissue extracts, particularly serum albumin. By depleting these proteins, lower abundance (tissue-derived) proteins were now revealed when an equal load of protein was used. This 'contamination' was significant, given that the yields of proteins after depletion were 20% and 25% of the 200 µg and 500 µg starting amounts, respectively. The depletion step was thus incorporated into the main profiling workflow, in an attempt to improve proteomic coverage of tissue proteins.
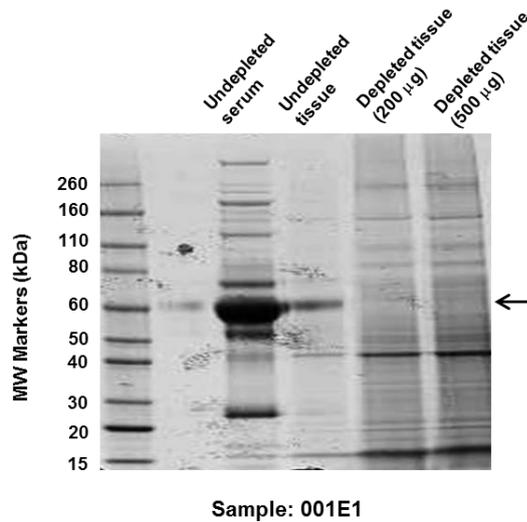
Figure 4.2 Gel image of immunodepleted tissue lysate sample.
Arrow indicates position of serum albumin

4.4 Identification of differentially expressed proteins in endometrial and endometriosis tissue

A comparison of protein expression levels in pooled tissue samples from endometriosis and controls was undertaken using 2D-DIGE. Each type of sample pool (eutopic tissue from pain, endometriosis and control groups in secretory and proliferative phases and ectopic tissue from endometriosis cases) was labelled in triplicate using CyDyes. Cy2 was used to label an internal standard comprising an equal mixture of all samples in the experiment. The standard was run on all gels thereby allowing accurate spot matching and normalisation across all gels. Protein components of these endometrial lysates were sufficiently resolved in the pH range of 4-8 (Figure 4.3). Spot detection and quantitation were performed automatically on overlaid Cy2/Cy3 and Cy2/Cy5 image pairs for each gel, followed by semi-automated gel to gel matching and statistical analysis. For a given spot, an average standardised abundance was determined as the average of the spot volume ratios between the standard and the test sample. Following inter-gel matching, the quantitative comparison of the gel images yielded 1,500 matched spots.

122
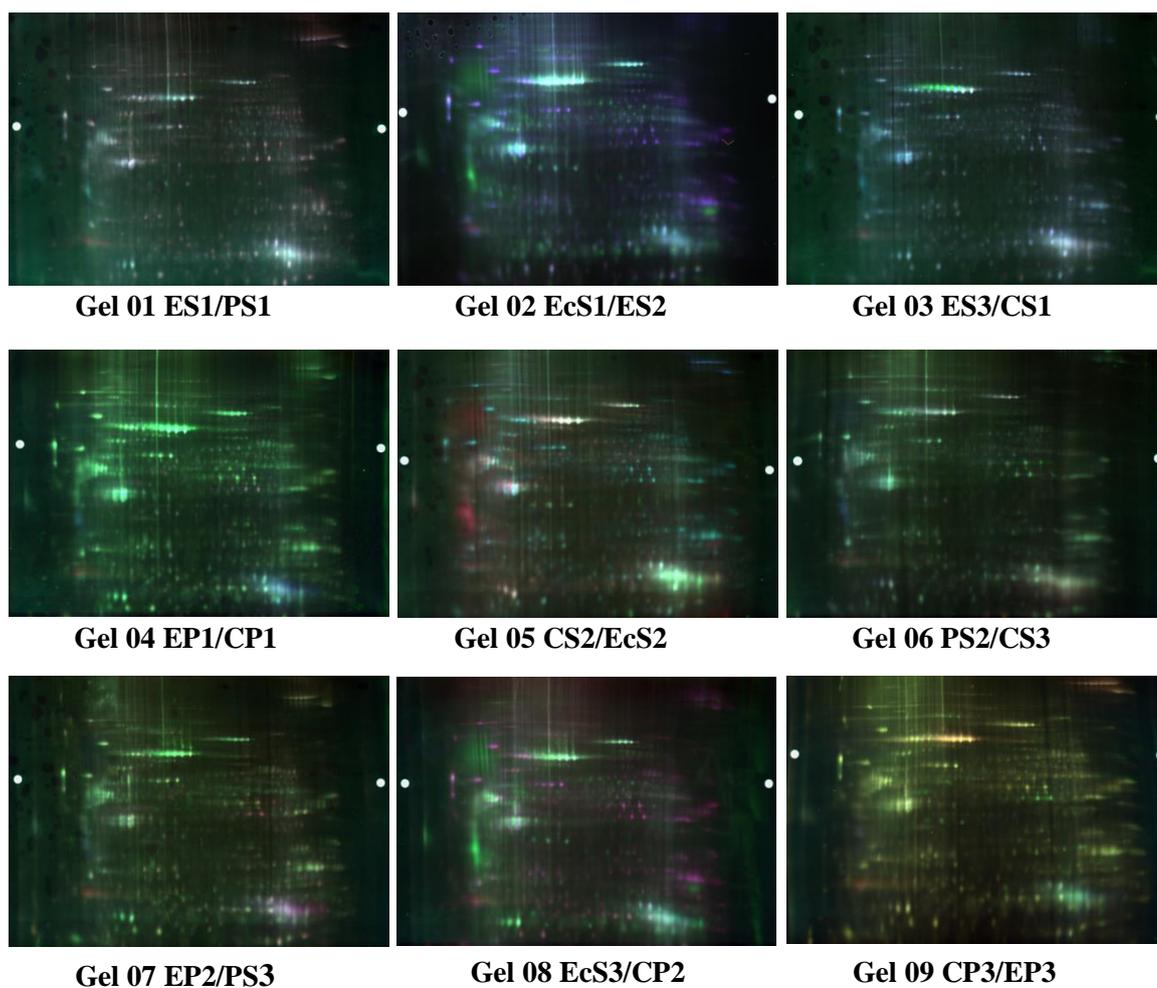
Figure 4.3 Overlaid 2D-DIGE images showing endometrial tissue lysate profiles.
80 µg of each pool was labelled in triplicate with Cy3 (red), Cy5 (blue) and Cy2 (green)
which was used to label an internal standard.

Analysis of variance was applied and data was filtered to retain spots with a significant difference $P < 0.05$ at >1.5-fold change for spots that were matched in all spot maps. This resulted in 72 differentially expressed spots of which some occurred in a series of 2-4 clusters of spots. Clusters could mean that the protein/s were present as multiple isoforms in the gels differing in their isoelectric points. Spots of interest identified through this analyses were also verified to have a 3D profile characteristic of a protein spot and features detected from non-protein sources like dust particles and background were filtered out.

Post staining of gels with colloidal Coomasie blue and matching of images to the master CyDye image (Gel 4) resulted in 52 well-defined Coomassie-detectable protein spots which were picked from gels and digested with trypsin for downstream identification by LC-MS/MS. Mass lists generated from the acquired spectra were compared against theoretically expected tryptic peptide masses in the SwissProt database searched using the Mascot search engine. The identifications were scored according to algorithms that take into account mass accuracy and peptide coverage. Following MS analysis, 613 proteins were identified with confidence from the 52 gel spots. Abundant serum proteins e.g. albumin and possible contaminants such as cytokeratins were excluded, resulting in a final list of 84 differentially expressed proteins; three low-abundance spots lacked identification. The identified proteins were also verified by comparison of their theoretical molecular weights and isoelectric points with gel positions on the master gel. A simple scoring system was also developed to prioritise candidates as possible biomarkers and was based on the proteins displaying the same direction of differential regulation between endometriosis and both control groups and in the ectopic *versus* eutopic tissue. A summary of the results with fold-changes and *P* values are presented in Table 4.1.

Table 4.1 List of proteins that were identified as significantly up/down regulated (*P* <0.05 and >1.5-fold change).

Master Spot No.	Acc No.	Protein Name	Score	Peptide Matches	Peptide Sequences	Pred Mass	Gel pI	Gel Mass	ES/CS Av. Ratio	ES/CS T-test	ES/PS Av. Ratio	ES/PS T-test	EP/CP Av. Ratio	EP/CP T-test	EcS/ES Av. Ratio	EcS/ES T-test	Biomarker Score
708	P51884	Lumican LUM	394	15	7	38747	3.74	73000	1.86	0.0051	2.25	0.0043	1.63	0.16	2.09	0.0083	4
708	P01011	Alpha-1-antichymotrypsin SERPINA3	69	2	2	47792	3.74	73000	1.86	0.0051	2.25	0.0043	1.63	0.16	2.09	0.0083	4
833	P51884	Lumican LUM	314	14	7	38747	3.87	65000	2.05	0.025	2.36	0.029	1.04	0.78	1.80	0.0064	4
833	P08670	Vimentin VIM	92	5	4	53676	3.87	65000	2.05	0.025	2.36	0.029	1.04	0.78	1.80	0.0064	4
833	O95302	Peptidyl-prolyl cis-trans isomerase FKBP9	114	5	3	63500	3.87	65000	2.05	0.025	2.36	0.029	1.04	0.78	1.80	0.0064	4
833	O75781	Paralemmin-1 PALM	97	2	2	42221	3.87	65000	2.05	0.025	2.36	0.029	1.04	0.78	1.80	0.0064	4
1363	O75874	Isocitrate dehydrogenase [NADP] cytoplasmic IDH1	474	23	11	46915	6.59	41000	-1.24	0.011	-1.97	0.00053	-2.40	0.00089	-3.39	4.60E-06	4
1463		No identification					8.38	37000	-1.58	0.0067	-1.43	0.051	-1.61	0.049	-1.90	0.0057	4
1543	P07951	Tropomyosin beta chain TPM2	1066	37	12	32945	3.5	33000	3.17	0.001	2.13	0.0039	1.10	0.11	4.65	0.0014	4
1543	P08670	Vimentin VIM	376	15	10	53676	3.5	33000	3.17	0.001	2.13	0.0039	1.10	0.11	4.65	0.0014	4
1548	P07951	Tropomyosin beta chain TPM2	885	33	10	32945	3.77	30000	4.15	1.00E+00	2.36	0.0057	1.15	0.19	4.15	0.00022	4
1548	P08670	Vimentin VIM	268	15	11	53676	3.77	30000	4.15	2.10E-06	2.36	0.0057	1.15	0.19	4.15	0.00022	4
1548	P51858	Hepatoma-derived growth factor HDGF	77	3	2	26886	3.77	30000	4.15	2.00E+00	2.36	0.0057	1.15	0.19	4.15	0.00022	4
1689	Q32P51	Heterogeneous nuclear ribonucleoprotein A1-like 2 HNRNPA1L2	113	2	2	34375	9.56	24000	-1.43	0.046	-1.48	0.031	-1.34	0.03	-5.71	0.0002	4
1981	P02647	Apolipoprotein A-I APOA1	1190	40	16	30759	4.2	15000	1.70	0.065	3.07	0.0065	1.43	0.067	1.59	0.13	4
1981	P12111	Collagen alpha-3(VI)chain COL6A3	206	9	7	345167	4.2	15000	1.70	0.065	3.07	0.0065	1.43	0.067	1.59	0.13	4
1981	P60709	Actin, cytoplasmic 1 ACTB	244	7	6	42052	4.2	15000	1.70	0.065	3.07	0.0065	1.43	0.067	1.59	0.13	4
1981	P04792	Heat shock protein beta-1 HSPB1	101	6	4	22826	4.2	15000	1.70	0.065	3.07	0.0065	1.43	0.067	1.59	0.13	4
492	P06396	Gelsolin GSN	195	10	4	86043	5.17	88000	-1.02	0.69	-1.56	0.0052	-1.92	0.00025	-2.92	0.0043	3
492	P21333	Filamin-A FLNA	149	6	6	283301	5.17	88000	-1.02	0.69	-1.56	0.0052	-1.92	0.00025	-2.92	0.0043	3
492	P07900	Heat shock protein HSP 90-alpha HSP90AA1	148	5	5	85006	5.17	88000	-1.02	0.69	-1.56	0.0052	-1.92	0.00025	-2.92	0.0043	3
492	O75369	Filamin-B FLNB	105	4	4	280157	5.17	88000	-1.02	0.69	-1.56	0.0052	-1.92	0.00025	-2.92	0.0043	3
497	P07900	Heat shock protein HSP 90-alpha HSP90AA1	103	4	3	85006	5.7	90000	-1.04	0.67	-1.63	0.016	-1.55	0.00037	-2.17	0.013	3

604	P11021	78 kDa glucose-regulated protein HSPA5	3062	99	34	72402	4.09	80000	-1.20	0.061	-1.74	0.0027	-1.62	0.0022	-9.47	0.00075	3
604	P51884	Lumican LUM	101	4	2	38747	4.09	80000	-1.20	0.061	-1.74	0.0027	-1.62	0.0022	-9.47	0.00075	3
604	P14923	Junction plakoglobin JUP	79	3	3	82434	4.09	80000	-1.20	0.061	-1.74	0.0027	-1.62	0.0022	-9.47	0.00075	3
883	O60701	UDP-glucose 6-dehydrogenase UGDH	753	29	14	55674	7.52	65000	-1.53	0.00022	-1.53	0.00011	1.02	0.84	-1.46	0.21	3
883	P14866	Heterogeneous nuclear ribonucleoprotein L HNRNPL	575	20	6	64720	7.52	65000	-1.53	0.00022	-1.53	0.00011	1.02	0.84	-1.46	0.21	3
883	P04040	Catalase CAT	242	13	8	59947	7.52	65000	-1.53	0.00022	-1.53	0.00011	1.02	0.84	-1.46	0.21	3
883	P30038	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial ALDH4A1	91	4	3	62137	7.52	65000	-1.53	0.00022	-1.53	0.00011	1.02	0.84	-1.46	0.21	3
962	P78371	T-complex protein 1 subunit beta CCT2	169	3	2	57794	5.9	59000	1.25	0.013	1.52	0.01	1.55	0.022	2.56	0.003	3
1149	P31943	Heterogeneous nuclear ribonucleoprotein H HNRNPH1	827	28	10	49484	5.34	51000	-1.21	0.26	-1.64	0.033	-1.77	1.60E-05	-2.08	0.036	3
1149	P05091	Aldehyde dehydrogenase, mitochondrial ALDH2	105	5	4	56859	5.34	51000	-1.21	0.26	-1.64	0.033	-1.77	1.00E+00	-2.08	0.036	3
1210	P61158	Actin-related protein 3 ACTR3	119	4	4	47797	5.8	50000	1.02	0.99	-1.73	0.018	-4.03	0.0047	-1.93	0.027	3
1210	P36957	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial DLST	139	4	4	49067	5.8	50000	1.02	0.99	-1.73	0.018	-4.03	0.0047	-1.93	0.027	3
1210	P35998	26S protease regulatory subunit 7 PSMC2	66	3	3	49002	5.8	50000	1.02	0.99	-1.73	0.018	-4.03	0.0047	-1.93	0.027	3
1210	Q9BQE3	Tubulin alpha-1C chain TUBA1C	47	2	2	50548	5.8	50000	1.02	0.99	-1.73	0.018	-4.03	0.0047	-1.93	0.027	3
1212	Q5VTE0	Putative elongation factor 1-alpha-like 3 EEF1A1P5 PE=5	365	17	9	50495	9	46000	-1.14	0.26	-1.76	0.0083	-2.11	0.0061	-8.86	0.00022	3
1368	O75874	Isocitrate dehydrogenase [NADP] cytoplasmic IDH1	916	39	17	46915	7.5	42000	-1.33	0.0027	-2.15	9.00E-05	-3.09	4.90E-05	-4.11	2.10E-05	3
1370	P00558	Phosphoglycerate kinase 1 PGK1	1016	38	19	44985	7.7	42000	-1.47	2.80E-05	-1.57	1.80E-05	1.07	0.69	-1.61	1.00E-01	3
1370	P62333	26S protease regulatory subunit 10B PSMC6	73	3	3	44430	7.7	42000	-0.47	1.00E+00	-0.57	1.00E+00	1.07	0.69	-1.61	1.00E-01	3
1371	Q9UBG3	Cornulin CRNN	205	8	4	53730	7.24	41000	-1.33	0.0032	-2.33	0.00018	-4.01	1.00E+00	-3.98	0.00016	3
1371	P04083	Annexin A1 ANXA1	314	7	6	38918	7.24	41000	-1.33	0.0032	-2.33	0.00018	-4.01	2.10E-05	-3.98	0.00016	3
1666	P04083	Annexin A1 ANXA1	1136	34	16	38918	6.87	25000	-1.05	0.24	-1.72	7.70E-05	-1.69	1.80E-05	-2.97	5.50E-08	3
1666	Q9H9H4	Vacuolar protein sorting-associated protein 37B VPS37B	75	3	2	31345	6.87	25000	-1.05	0.24	-1.72	2.00E+00	0.31	2.00E+00	-0.97	5.50E-08	3
1666	O00151	PDZ and LIM domain protein 1 PDLIM1	52	2	2	36505	6.87	25000	-1.05	0.24	-1.72	1.00E+00	-0.69	1.00E+00	-1.97	5.50E-08	3
1997	P01834	Ig kappa chain C region IGKC	104	4	2	11773	4.2	15000	1.10	0.14	1.85	0.00037	1.44	0.0019	2.15	0.0011	3
1997	P01593	Ig kappa chain V-I region AG	118	3	2	12099	4.2	15000	1.10	0.14	1.85	0.00037	1.44	0.0019	2.15	0.0011	3
403	P55072	Transitional endoplasmic reticulum ATPase VCP	175	6	3	89950	4	100000	-1.22	0.027	-1.57	0.0065	-1.15	0.042	-2.75	1.50E-05	2

404	P55072	Transitional endoplasmic reticulum ATPase VCP	1177	36	16	89950	4.34	95000	-1.17	0.00077	-1.68	0.0054	-1.20	0.018	-2.16	0.0082	2
745	P54652	Heat shock-related 70 kDa protein 2 HSPA2	603	25	13	70263	5	67000	1.64	0.037	1.52	0.18	1.21	0.25	-1.18	0.37	2
745	P38646	Stress-70 protein, mitochondrial HSPA9	105	4	3	73920	5	67000	1.64	0.037	1.52	0.18	1.21	0.25	-1.18	0.37	2
754	P02545	Prelamin-A/C LMNA	927	38	19	74380	7	65000	-1.12	0.23	-1.60	0.016	-1.25	0.044	-1.90	0.0047	2
754	Q9NSD9	Phenylalanine--tRNA ligase beta subunit FARSB	173	8	6	66701	7	65000	-1.12	0.23	-1.60	0.016	-1.25	0.044	-1.90	0.0047	2
754	Q9Y3Z3	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1	94	6	6	72896	7	65000	-1.12	0.23	-1.60	0.016	-1.25	0.044	-1.90	0.0047	2
754	P51888	Prolargin PRELP	167	6	4	44181	7	65000	-1.12	0.23	-1.60	0.016	-1.25	0.044	-1.90	0.0047	2
767	Q02413	Desmoglein-1 DSG1	91	4	4	114702	7.31	70000	-1.12	0.068	-1.80	0.026	-1.39	0.062	-1.92	0.00056	2
767	P02545	Prelamin-A/C LMNA	101	3	3	74380	7.31	70000	-1.12	0.068	-1.80	0.026	-1.39	0.062	-1.92	0.00056	2
864	P15924	Desmoplakin DSP	604	20	16	334021	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
864	P10809	60 kDa heat shock protein, mitochondrial HSPD1	866	17	9	61187	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
864	P07339	Cathepsin D CTSD	291	8	7	45037	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
864	P07355	Annexin A2 ANXA2	154	4	4	38808	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
864	Q13867	Bleomycin hydrolase BLMH	76	3	2	53155	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
864	Q96QA5	Gasdermin-A GSDMA	78	3	2	49619	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
864	P04040	Catalase CAT	70	2	2	59947	4.43	65000	-1.11	0.28	-1.53	0.02	-1.37	0.0016	-2.45	0.028	2
899	P04040	Catalase CAT	1337	42	16	59947	7.64	61000	-1.64	0.0017	-1.82	0.0013	1.11	0.11	-1.37	0.13	2
899	P14618	Pyruvate kinase PKM	506	15	8	58470	7.64	61000	-1.64	0.0017	-1.82	0.0013	1.11	0.11	-1.37	0.13	2
899	Q99832	T-complex protein 1 subunit eta CCT7	45	2	2	59842	7.64	61000	-1.64	0.0017	-1.82	0.0013	1.11	0.11	-1.37	0.13	2
899	O60701	UDP-glucose 6-dehydrogenase UGDH	82	2	2	55674	7.64	61000	-1.64	0.0017	-1.82	0.0013	1.11	0.11	-1.37	0.13	2
998	P50995	Annexin A11 ANXA11	186	6	5	54697	8.04	57000	-1.32	0.046	-1.89	0.0072	-1.28	0.12	-2.43	0.04	2
998	P00390	Glutathione reductase, mitochondrial GSR	132	5	4	56791	8.04	57000	-1.32	0.046	-1.89	0.0072	-1.28	0.12	-2.43	0.04	2
998	P34897	Serine hydroxymethyltransferase, mitochondrial SHMT2	59	2	2	56414	8.04	57000	-1.32	0.046	-1.89	0.0072	-1.28	0.12	-2.43	0.04	2
1150	P25705	ATP synthase subunit alpha, mitochondrial ATP5A1	402	14	7	59828	8.47	51000	-1.57	0.041	1.04	0.91	-1.25	0.029	-1.80	0.024	2
1150	P14618	Pyruvate kinase PKM	39	2	2	58470	8.47	51000	-1.57	0.041	1.04	0.91	-1.25	0.029	-1.80	0.024	2
1213	Q5VTE0	Putative elongation factor 1-alpha-like 3 EEF1A1P5	256	9	5	50495	9.41	48000	-1.11	0.28	-1.68	0.011	-1.34	0.034	-9.97	8.00E-05	2
1234	P17661	Desmin DES	428	15	11	53560	4.29	47000	-1.54	0.017	1.69	0.036	1.65	0.042	2.34	0.009	2
1234	P60709	Actin, cytoplasmic 1 ACTB	552	11	6	42052	4.29	47000	-1.54	0.017	1.69	0.036	1.65	0.042	2.34	0.009	2
1234	P01009	Alpha-1-antitrypsin SERPINA1	449	9	5	46878	4.29	47000	-1.54	0.017	1.69	0.036	1.65	0.042	2.34	0.009	2

1234	P07437	Tubulin beta chain TUBB	205	7	6	50095	4.29	47000	-1.54	0.017	1.69	0.036	1.65	0.042	2.34	0.009	2
1234	O60664	Perilipin-3 PLIN3	119	3	2	47217	4.29	47000	-1.54	0.017	1.69	0.036	1.65	0.042	2.34	0.009	2
1234	Q8TBC4	NEDD8-activating enzyme E1 catalytic subunit UBA3	50	2	2	52504	4.29	47000	-1.54	0.017	1.69	0.036	1.65	0.042	2.34	0.009	2
1270	P06733	Alpha-enolase ENO1	2245	68	20	47481	7.14	46000	-1.20	0.04	-1.79	8.90E-05	-1.18	0.083	-2.40	4.20E-04	2
1270	Q9NVA2	Septin-11 SEPT11	205	7	4	49652	7.14	46000	-1.20	0.04	-1.79	1.00E+00	-1.18	0.083	-2.40	4.20E-04	2
1319	P49411	Elongation factor Tu, mitochondrial TUFM	1156	30	18	49852	7.14	43000	-1.27	0.011	-1.55	8.80E-05	-1.21	0.047	-2.92	6.50E-04	2
1319	P06733	Alpha-enolase ENO1	741	14	8	47481	7.14	43000	-1.27	0.011	-1.55	1.00E+00	-1.21	0.047	-2.92	6.50E-04	2
1319	Q02413	Desmoglein-1 DSG1	216	7	7	114702	7.14	43000	-1.27	0.011	-1.55	2.00E+00	-1.21	0.047	-2.92	6.50E-04	2
1319	P15924	Desmoplakin DSP	67	5	5	334021	7.14	43000	-1.27	0.011	-1.55	3.00E+00	-1.21	0.047	-2.92	6.50E-04	2
1319	P14923	Junction plakoglobin JUP	132	3	3	82434	7.14	43000	-1.27	0.011	-1.55	4.00E+00	-1.21	0.047	-2.92	6.50E-04	2
1319	Q08554	Desmocollin-1 DSC1	64	2	2	101406	7.14	43000	-1.27	0.011	-1.55	5.00E+00	-1.21	0.047	-2.92	6.50E-04	2
1323	P60709	Actin, cytoplasmic 1 ACTB	242	8	5	42052	4.8	42000	-1.17	0.16	-1.71	0.0068	-1.88	0.0024	-1.06	0.65	2
1323	P14923	Junction plakoglobin JUP	61	3	3	82434	4.8	42000	-1.17	0.16	-1.71	0.0068	-1.88	0.0024	-1.06	0.65	2
1323	P12277	Creatine kinase B-type CKB	231	3	2	42902	4.8	42000	-1.17	0.16	-1.71	0.0068	-1.88	0.0024	-1.06	0.65	2
1323	Q14240	Eukaryotic initiation factor 4A-II EIF4A2	64	3	2	46601	4.8	42000	-1.17	0.16	-1.71	0.0068	-1.88	0.0024	-1.06	0.65	2
1351	P00558	Phosphoglycerate kinase 1 PGK1	1922	58	19	44985	8.46	41000	-1.24	0.01	-1.77	0.011	-1.18	0.082	-4.95	0.00021	2
1351	P50454	Serpin H1 SERPINH1	241	9	4	46525	8.46	41000	-1.24	0.01	-1.77	0.011	-1.18	0.082	-4.95	0.00021	2
1351	P22695	Cytochrome b-c1 complex subunit 2, mitochondrial UQCRC2	62	3	2	48584	8.46	41000	-1.24	0.01	-1.77	0.011	-1.18	0.082	-4.95	0.00021	2
1515	P02675	Fibrinogen beta chain FGB	546	23	10	56577	4.75	33000	-1.35	0.052	2.26	0.013	-1.46	0.01	1.52	0.048	2
1515	P60709	Actin, cytoplasmic 1 ACTB	288	9	7	42052	4.75	33000	-1.35	0.052	2.26	0.013	-1.46	0.01	1.52	0.048	2
1675	P02647	Apolipoprotein A-I APOA1	1592	49	19	30759	4.3	28000	-1.21	0.064	-1.56	0.0025	-1.19	0.022	-2.24	0.0096	2
1675	Q15181	Inorganic pyrophosphatase PPA1	222	4	3	33095	4.3	28000	-1.21	0.064	-1.56	0.0025	-1.19	0.022	-2.24	0.0096	2
1693	P04406	Glyceraldehyde-3-phosphate dehydrogenase GAPDH	1331	39	15	36201	7.5	28000	-1.21	0.005	-1.50	0.00069	-1.15	0.065	-2.93	6.70E-05	2
1693	P40926	Malate dehydrogenase, mitochondrial MDH2	695	22	11	35937	7.5	28000	-1.21	0.005	-1.50	0.00069	-1.15	0.065	-2.93	6.70E-05	2
1693	P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 HNRNPA2B1	400	17	9	37464	7.5	28000	-1.21	0.005	-1.50	0.00069	-1.15	0.065	-2.93	6.70E-05	2
1850	P00915	Carbonic anhydrase 1 CA1	1838	49	10	28909	7.41	18000	-1.58	0.00092	-1.43	0.031	-1.51	0.00073	1.24	0.085	2
1850	P25789	Proteasome subunit alpha type-4 PSMA4	66	3	2	29750	7.41	18000	-1.58	0.00092	-1.43	0.031	-1.51	0.00073	1.24	0.085	2
1850	Q13126	S-methyl-5~-thioadenosine phosphorylase MTAP	126	3	2	31729	7.41	18000	-1.58	0.00092	-1.43	0.031	-1.51	0.00073	1.24	0.085	2
1905		No identification					7.91	17000	-1.63	0.025	-1.01	0.91	-2.32	0.0059	1.23	0.31	2
1924	P68871	Hemagloblin subuint beta HBB	1024	29	9	16102	8.25	16000	-1.71	0.014	1.00	0.92	-1.60	0.06	1.41	0.011	2

1924	P69905	Hemagloblin subuint alpha HBA1	290	13	5	15305	8.25	16000	-1.71	0.014	1.00	0.92	-1.60	0.06	1.41	0.011	2
1924	P00915	Carbonic anhydrase 1 CA1	96	4	2	28909	8.25	16000	-1.71	0.014	1.00	0.92	-1.60	0.06	1.41	0.011	2
1929	P68871	Hemagloblin subuint beta HBB	1564	41	11	16102	8.05	16000	-1.92	0.00051	-1.15	0.13	-1.85	0.029	1.44	0.0068	2
1929	P69905	Hemagloblin subuint alpha HBA1	471	18	7	15305	8.05	16000	-1.92	0.00051	-1.15	0.13	-1.85	0.029	1.44	0.0068	2
1929	P00915	Carbonic anhydrase 1 CA1	58	4	3	28909	8.05	16000	-1.92	0.00051	-1.15	0.13	-1.85	0.029	1.44	0.0068	2
1932	P68871	Hemagloblin subuint beta HBB	558	19	9	16102	9	15000	-1.86	0.0019	-1.20	0.052	-2.18	0.025	1.27	0.095	2
1932	P69905	Hemagloblin subuint alpha HBA1	204	7	4	15305	9	15000	-1.86	0.0019	-1.20	0.052	-2.18	0.025	1.27	0.095	2
1942	P68871	Hemagloblin subuint beta HBB	849	26	10	16102	7.64	16000	-1.62	0.0011	-1.05	0.56	-1.99	0.014	1.06	0.61	2
1942	P69905	Hemagloblin subuint alpha HBA1	250	8	5	15305	7.64	16000	-1.62	0.0011	-1.05	0.56	-1.99	0.014	1.06	0.61	2
1947	P68871	Hemagloblin subuint beta HBB	621	17	8	16102	8.21	16000	-1.84	0.0011	-1.08	0.27	-1.74	0.012	1.09	0.58	2
1947	P69905	Hemagloblin subuint alpha HBA1	445	17	7	15305	8.21	16000	-1.84	0.0011	-1.08	0.27	-1.74	0.012	1.09	0.58	2
1948	P69905	Hemagloblin subuint alpha HBA1	245	9	5	15305	8.8	16000	-2.03	0.015	-1.12	0.39	-1.65	0.19	1.44	0.08	2
1968	P68871	Hemagloblin subuint beta HBB	461	14	6	16102	7.54	16000	-1.52	0.0025	1.10	0.36	-1.79	0.057	-1.10	0.52	2
1968	P69905	Hemagloblin subuint alpha HBA1	140	4	3	15305	7.54	16000	-1.52	0.0025	1.10	0.36	-1.79	0.057	-1.10	0.52	2
2005	P04792	Heat shock protein beta-1 HSPB1	913	28	12	22826	6.88	34000	1.52	0.0036	-1.06	0.75	-1.29	0.07	1.92	0.0011	2
2167	P08559	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial PDHA1	66	3	2	43952	4.99	15000	-1.81	0.05	1.21	0.12	1.26	0.42	-1.46	0.0018	2
2167	P06733	Alpha-enolase ENO1	52	2	2	47481	4.99	15000	-1.81	0.05	1.21	0.12	1.26	0.42	-1.46	0.0018	2
2175	P04792	Heat shock protein beta-1 HSPB1	813	29	12	22826	5.11	15000	1.67	0.0025	-1.39	0.23	-1.02	0.93	2.08	0.01	2
2179		No identification					4	73000	1.55	0.23	2.42	0.038	1.22	0.68	3.17	0.00013	2

Numerous spots were matched to more than one protein. Shading indicates altered expression (green increased; red decreased). Proteins of interest are shaded yellow.

4.5 Discussion

Since multiple proteins were identified in many of the spots, it was difficult to assign which protein was responsible for the differences in expression. To address this, proteins were selected based on abundance, the number of matched peptides and matching of the theoretical and experimental molecular weight and isoelectric points. A total of 84 differentially expressed proteins were identified from the profiling. The differentially expressed proteins comprised cytoskeletal proteins, metabolic enzymes, extracellular matrix proteins, muscle proteins, serum proteins, haemoglobin and those involved in protein folding.

The proteins of interest were ranked based on a score, which was higher if the protein showed the same directionality of altered expression between endometriosis and both controls groups, and did not change between the control groups (PS/CS). Thus, a protein was assigned a score of 4 if its expression displayed a significant increase ($P \leq 0.05$, ≤ 1.5-fold change) in the four comparisons ES/CS, ES/PS, EP/CP and EcS/ES. These proteins included; lumican, tropomyosin β chain and heterogeneous nuclear ribonucleoprotein A-1-like 2.

Expression of lumican was higher in the endometriosis group compared to both the control and pain groups (ES/CS av. ratio=1.86, P=0.0051; ES/PS av. ratio=2.25, P=0.0043). Expression was also higher in the ectopic endometrial tissue compared to the eutopic endometrium in the secretory phase (EcS/ES av. ratio=2.09, P=0.008). Expression also differed between phases, with higher differential expression seen in the secretory phase. Lumican was also identified from several spots.

Lumican is an extracellular matrix protein that is expressed in different tissues e.g. muscle, cartilage and cornea. It is a member of the small leucine rich proteoglycan (SLRP) family which comprises of 17 genes that share structural homologies e.g. cysteine residues, leucine rich repeats and at least one glycosaminoglycan chain. Members of this family have been reported to play a role in cell migration and proliferation during embryonic development, tissue repair and tumour growth. *In*

vitro, SLRPs have been shown to regulate collagen fibrillogenesis (Ezura et al., 2000), a process important in development, tissue repair and metastasis. Altered expression of lumican has been demonstrated in several cancer types such as breast, colorectal, neuroendocrine, uterine, cervical, pancreatic, and lung (Koninger et al., 2004, Leygue et al., 1998, Lu et al., 2002, Shinji et al., 2005), with conflicting data on its role in promoting or abrogating tumour progression.

The exact molecular function of lumican is poorly understood. It has been shown to down-regulate the proteolytic activities of MMP14 and MMP9 associated with endothelial cell membranes, thereby exerting an angiostatic effect (Brezillon et al., 2013). In the endometrium, lumican is present in the ECM of pre-menopausal fertile women. Its expression has been reported to increase during the transition of the endometrium from the late proliferative to early secretory phase. The level of its expression decreases in menopausal women and in other pathological states. Its presence in the endometrium has been demonstrated in mice where it is spatially and temporally modulated in the pre-implantation period (San Martin *et al*., 2003). The role if any of lumican in endometriosis is unclear. The expression pattern in tissues shown from this profiling may suggest that lumican is involved in the proliferation, invasion and/or adhesion of ectopic endometrial cells to neighbouring peritoneal organs. Its usefulness as a biomarker of endometriosis is yet to be established.

Tropomyosins are actin binding proteins that form the major structural constituent of microfilaments. High molecular weight tropomyosins (TPM1 and TPM2) are thought to play a role in stabilising the organisation of actin filaments, which in turn play a role in muscle contraction and the maintenance of cell shape, cell motility, cell-cell and cell-matrix interactions (Shah et al., 2001). The loss of tropomyosin expression reported in tumour cells may prevent proper assembly of microfilaments and thereby contribute to the invasive and metastatic properties of cancer cells (Shah et al., 2001). The role of tropomyosins in endometriosis is not well studied. Herein, TPM2 expression was highest in ectopic endometrial tissue compared to eutopic tissue (av. ratio=4.65, P=0.001) and was increased in endometriosis compared to control and pain groups in the secretory phase, suggesting differential expression across the cycle.

Its differential expression possibly points to an involvement in cellular structural change occurring in the endometrium allowing refluxed endometrial cells to invade and adhere to peritoneal organs, eventually leading to disease development. Its potential as a biomarker of endometriosis has not been evaluated.

Heterogeneous nuclear ribonucleoprotein A1-like 2 belongs to the A/B subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs); RNA binding proteins that are associated with pre-mRNA in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. These proteins may have a primary role in the expression of specific myometrial protein species in parturition (Pollard et al., 2000). In this experiment, HNRNPA12 was down-regulated in both endometriosis compared to healthy and pain controls in the secretory phase (av. ratio= -1.43, P=0.04, av.ratio= -1.48, P=0.03 respectively). A significant down-regulation of this protein was also observed in the ectopic tissue compared to eutopic endometrium in the secretory phase (av. ratio=-5.71, P=0.0002). The potential role of this protein in endometriosis is unknown.

4.6 Conclusion

Gel-based proteomics is a well-established technique for global protein separation and quantification for profiling biological and clinical samples. Development of the 2D-DIGE technique was a milestone in increasing the reproducibility of 2D-PAGE for more accurate quantification. 2D-DIGE was used successfully herein to profile endometrial and endometriosis tissue samples with the aim of identifying potential tissue biomarkers. One of the main challenges experienced in this analysis was contamination of the tissue samples with blood and serum proteins. Highly abundant serum proteins were identified in the preliminary analysis and were effectively lowering the load of tissue-derived proteins and could possibly obscure lower abundance proteins. To address this problem, an immunodepletion step was incorporated into the main workflow. Although this was not sufficient to remove all of these proteins (as evidenced by the presence of significant levels of serum albumin and other proteins targeted by the depletion), the complexity of the spot patterns was

improved. The identification of haemoglobin and several serum proteins as being differentially expressed, further highlights this contamination issue. This will have undoubtedly increase the false discovery rate and complicate candidate selection for verification. Another major challenge experienced was the presence of multiple cytokeratins and cytokeratin-related proteins. This may be due to contamination during sample processing, although one cannot rule out the possibility that these are *bona fide* endogenous proteins.

The heterogeneous cellular composition of the tissues was evidenced by the presence of high abundance structural proteins known to be expressed by different cell types and these may also mask lower abundance proteins during the analysis. A possible explanation is the fibrous nature of some of the ectopic tissue samples which were nodular in nature and contained large amounts of fibro-muscular tissue. Another possible explanation for the differential expression of these proteins could be the use of ectopic tissue samples from different sites. This presents a challenge in that while some lesions may have clearly defined endometrial tissue, including glands and stroma, other lesions could only be clusters of a few cells and a single gland. This results in variability of the cellular composition of these different endometriosis lesion types thereby compromising the quality of the analysis. Whilst the pooling approach would average out some of this heterogeneity, it is possible that outlier samples may skew the data leading to a high false discovery rate in terms of identifying potential biomarkers.

CHAPTER 5: MS-BASED PROTEOMIC PROFILING OF ENDOMETRIAL TISSUE LYSATES

5.1 Introduction

The aim of the work presented in this chapter was to explore protein Tandem Mass Tag (TMT) labelling and separation as a potential method improving quantitative coverage in profiling endometrial and endometriosis tissue biopsies for biomarker discovery. The fractionation of intact proteins by molecular weight and charge potentially offers better separation than peptide-based methods due to the greater degree of heterogeneity of protein species versus their peptide components. Additionally multiple peptides from abundant proteins will be preferentially sampled across the entire chromatographic space in data-dependent mode at the expense of lower abundant peptides, thus reducing coverage of the proteome. However, the high level of TMT labelling required may be hard to achieve for proteins versus peptides due to secondary structure constraints. Lysine labelling will also block some tryptic sites so additional proteases should be used. Profiling with protein separation is also likely to be complementary to peptide-based profiling approaches when datasets are combined. A peptide-based labelling and separation approach that has previously shown high proteomic coverage in this lab was also used to profile the six immunodepleted tissue pools that were first digested with trypsin prior to TMT labelling and separation. A three dimensional separation approach was used to maximise the number of peptides for analysis by mass spectrometry. First dimension separation of peptides was achieved by SAX liquid chromatography, with the six collected fractions further separated in a second dimension by off-line high pH reverse-phase liquid chromatography prior to low pH nano-RPLC-MS/MS.

5.2 Optimisation of protein TMT labelling and SAX fractionation

Four steps were incorprated into this workflow namely; immunodepletion of abundant serum proteins from the tissue lysates, protein denaturation and labelling using TMT reagents for relative quantification, SAX chromatography and SDS-PAGE for

separation, followed bydigestion and LC-MS/MS analysis. Tissue homogenisation and protein extraction was achieved as previously described (Chapter 2). Optimisation of protein labelling was carried out using pools of equal amounts of protein from each of three types of tissue lysate, namely; pool D1 (deep infiltrating endometriosis), pool E1 (ovarian endometriosis) and pool EU (eutopic endometrium). As well as optimising the strategy, this would allow the determination of any differences in protein levels in the different forms of endometriosis.

Immunodepletion was carried out on pooled samples to improve coverage as previously described. Briefly, PP12 immunodepletion resin was used to deplete 500 µg total protein from each pool giving a final protein yield of ~20%. Test samples were fractionated using SAX and 1D-SDS-PAGE prior to digestion. Different concentrations of elution buffers were tested. In the first test, pooled lysate was fractionated after immunodepletion without TMT labelling and proteins eluted from the SAX resin using 50 mM, 100 mM, 200 mM and 1M NaCl in 20 mM TEAB (Figure 5.1A).

Generally most proteins were eluted from the resin by 200 mM NaCl, however retention was still apparent. For this reoson, the elution steps were expanded at the higher salt concentrations in a second experiment where TMT labelling was incorporated. This generated somewhat improved separation although it was apparent that the more prominent protein bands were present across multiple fractions (Fig 5.1 B). This also suggested that labelling affects the charge and hence retention of proteins.
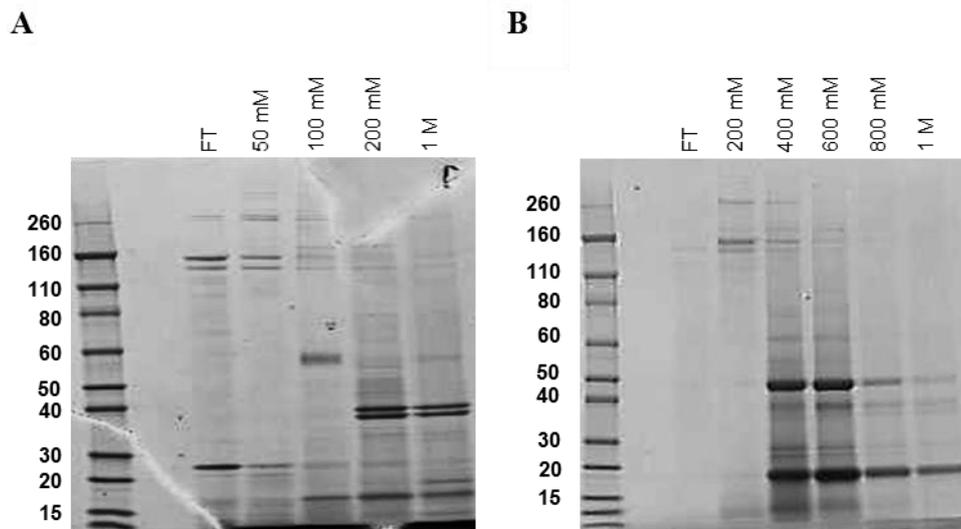
Figure 5.1 Representative 1D gel images of SAX protein fractionation.

Bands were cut from the 400 mM and 600 mM lanes on the second gel, digested sequentially with Glu-C then trypsin and analysed by LC-MS/MS. TMT labelling efficiency was determined to be ~95% with a total of 365 proteins identified. However, only 217 (60%) of the proteins had quantitative information. This was due to the identification of proteins using matched peptides that lacked a lysine residue and hence were not labelled with a TMT tag. Proteins displaying a median reporter ion ratio of ≥1.5-fold were considered to be differentially regulated when comparing the pairs of tissue types; 108 proteins were differentially expressed in the comparison between ectopic ovarian tissue and eutopic tissue, whilst 50 proteins were differentially expressed between the DIE tissue compared to eutopic tissue (Fig. 5.2)

136
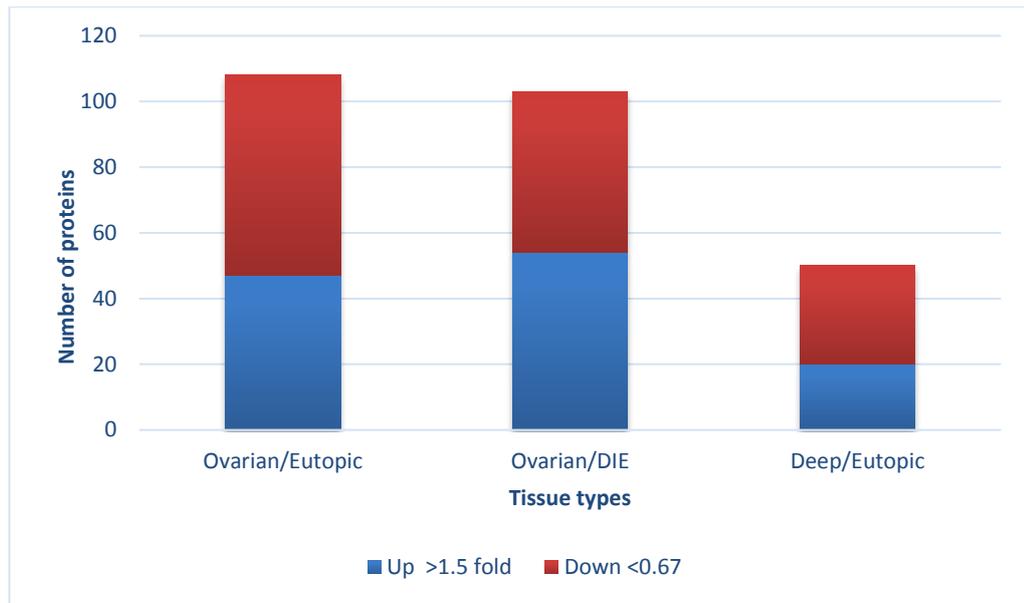
Figure 5.2 Graphical representation of differentially expressed proteins.The number of proteins displaying a fold-change of >1.5 for the different tissue comparisons are shown.

The experiment was repeated on the full set of sample pools which were labelled with 6-plex TMT tags for comparison of the CS, PS, ES, EcS, CP and EP groups using the higher NaCl concentrations for SAX elution. The fractions were run on a 1D-SDS-PAGE gel (Fig 5.3). There were no clear protein bands seen from the gel indicating considerable protein loss. This loss might have occurred during concentration of the samples after immunodepletion and/or during denaturation, labelling and clean-up, and may have been exacerbated by poor SAX fractionation.
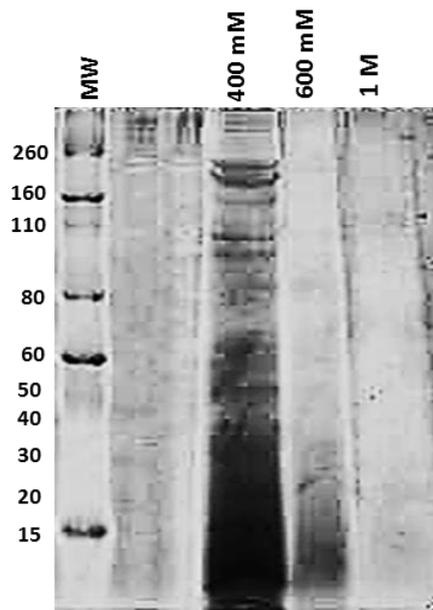
Figure 5.3 1D gel image of SAX fractions in the main protein profiling experiment.

Despite the protein loss, 100 gel slices were excised from lanes 400 mM and 600 mM of the gel and digested with Glu-C and trypsin and analysed by LC-MS/MS. A total of 187 proteins were identified of which 109 (58%) proteins had quantitative information. This was somewhat lower coverage compared to the optimisation experiments because of the protein loss. Table 5.1 shows proteins identified and quantified across the different comparisons.. A scoring system was used to rank the proteins to aid in selection of proteins of interest. Proteins that showed the same directionality of altered expression (>1.5-fold) between endometriosis and both controls groups and between ectopic and eutopic tissue would be assigned a total score of 4, although no protein achieved this score.

Table 5.1 Quantified proteins from TMT-protein profiling of endometrial tissues.

Accession	Description	Protein Score	# Proteins	# Unique Peptides	# Peptides	# PSMs	Ratio 130/126 ES/CS	130/126 Count	130/126 Variability [%]	Ratio 130/128 ES/PS	130/128 Count	130/128 Variability [%]	Ratio 131/127 EP/CP	127/131 Count	127/131 Variability [%]	Ratio 129/130 EcS/ES	129/130 Count	129/130 Variability [%]	Biomarker Score
P05546	Heparin cofactor 2 SERPIND1	54.22	1	1	1	2	1.748	1		1.197	1		4.444	1		3.394	1		3
P09493	Tropomyosin alpha-1 chain TPM1	64.01	1	2	2	3	1.684	2	120.3	2.733	2	271.6	0.954	2	41.3	5.118	2	131.4	3
P30740	Leukocyte elastase inhibitor SERPINB1	697.04	1	7	7	25	1.542	21	30.3	1.636	21	14	1.706	21	28.5	0.679	21	42.5	3
P07355	Annexin A2 ANXA2	75.78	2	2	2	3	0.892	3	15	0.388	3	52.9	0.521	3	11.9	0.652	3	65.8	3
Q9Y342	Plasmolipin PLLP	31.48	1	1	1	1	0.826	1		0.46	1		0.43	1		0.177	1		3
P07384	Calpain-1 catalytic subunit CAPN1	64.34	1	2	2	2	0.663	1		0.573	1		0.695	1		0.601	1		3
P53618	Coatomer subunit beta COPB1	25.64	1	1	1	1	0.494	1		0.392	1		0.799	1		0.325	1		3
P06396	Gelsolin GSN	66.1	1	1	1	1	0.358	1		0.43	1		0.175	1		1.089	1		3
P02679	Fibrinogen gamma chain FGG	39.38	1	1	1	1	6.887	1		5.607	1		0.435	1		0.951	1		2
P08729	Keratin, type II cytoskeletal 7 KRT7	1602.97	5	8	12	41	6.179	11	62.5	2.185	11	42.5	1.497	11	69.3	0.133	11	90.7	2
P02458	Collagen alpha-1(II) chain COL2A1	116.95	1	1	1	6	4.74	5	22	0.592	5	8.4	1.073	5	39.2	2.864	5	22.7	2
P13647	Keratin, type II cytoskeletal 5 KRT5	1110.76	3	5	9	51	4.709	4	72.5	4.247	4	110	1.195	4	17.1	0.197	4	88.4	2
P35908	Keratin, type II cytoskeletal 2 KRT2	19495.96	3	21	28	320	4.696	7	108.7	5.332	7	109.7	1.2	7	7.5	0.203	7	129.1	2
P02538	Keratin, type II cytoskeletal 6A KRT6A	1737.7	3	2	10	70	4.696	7	108.7	5.332	7	109.7	1.2	7	7.5	0.203	7	129.1	2
P48668	Keratin, type II cytoskeletal 6C KRT6C	1705.22	3	2	10	70	4.696	7	108.7	5.332	7	109.7	1.2	7	7.5	0.203	7	129.1	2
Q03135	Caveolin-1 CAV1	30.59	1	1	1	1	1.916	1		1.247	1		1.307	1		4.173	1		2
Q99715	Collagen alpha-1(XII) chain COL12A1	509.02	1	7	7	16	1.693	2	110.5	1.046	2	0.1	3.484	2	302.2	1.073	2	12	2
P30041	Peroxiredoxin-6 PRDX6	128.49	1	1	1	4	1.693	4	12.7	0.967	4	8.6	1.745	4	8.9	1.08	4	34.5	2
P61026	Ras-related protein Rab-10 RAB10	62.03	1	1	1	2	1.352	1		0.91	1		0.491	1		0.266	1		2
P21333	Filamin-A FLNA	264.64	1	2	2	6	1.212	3	28.4	0.685	3	33.1	1.828	3	25.7	2.546	3	7.1	2
P00738	Haptoglobin HP	168.08	1	4	4	7	1.082	6	6	1.176	6	14.1	1.686	6	26.4	3.881	6	24.5	2
Q14315	Filamin-C FLNC	162.96	1	4	4	5	1.066	1		0.339	1		9.174	1		4.701	1		2
P08758	Annexin A5 ANXA5	373.11	1	4	4	11	0.897	11	26.7	0.886	11	20.4	0.508	11	39.9	0.475	11	32	2
Q06830	Peroxiredoxin-1 PRDX1	37.72	1	1	1	1	0.534	1		0.646	1		0.687	1		0.542	1		2
P04264	Keratin, type II cytoskeletal 1 KRT1	33659.78	2	41	44	1000	0.524	2	9.4	2.113	2	111.4	0.603	2	21.8	0.669	2	98	2
P01833	Polymeric immunoglobulin receptor PIGR	587.34	1	4	4	23	0.452	18	21.4	2.452	18	30.2	0.451	18	46.7	0.286	18	46.4	2
P01009	Alpha-1-antitrypsin SERPINA1	608.07	1	9	9	23	0.318	7	100.8	1.039	7	17.9	1.647	7	1.2	2.977	7	126.2	2
P19652	Alpha-1-acid glycoprotein 2 ORM2	54.64	1	1	2	2	0.218	1		1.629	1		0.357	1		4.093	1		2
Q04695	Keratin, type I cytoskeletal 17 KRT17	307.13	10	1	7	12	6.296	1		1.075	1		0.861	1		0.214	1		1
P08779	Keratin, type I cytoskeletal 16 KRT16	1241.61	14	1	9	68	4.1	4	28	0.946	4	7.8	0.811	4	30.7	0.147	4	132.3	1

P08727	Keratin, type I cytoskeletal 19 KRT19	2225.31	10	7	17	127	3.382	37	72.3	0.805	37	35.3	0.864	37	27	0.248	37	59.8	1
P27797	Calreticulin CALR	146.6	1	2	2	6	2.669	4	29.8	1.267	4	31	1.238	4	4	0.364	4	31.5	1
P05787	Keratin, type II cytoskeletal 8 KRT8	5450.61	5	25	32	188	2.636	127	35.7	0.815	127	23.8	0.87	127	27.9	0.192	126	93.3	1
Q6KB66	Keratin, type II cytoskeletal 80 KRT80	489.22	2	1	5	23	2.48	15	77.5	0.738	15	27.2	0.98	15	36.7	0.349	15	151.8	1
P02545	Prelamin-A/C LMNA	616.83	1	14	14	30	2.466	18	46.9	1.255	18	22.3	1.015	18	32.6	1.011	18	23.5	1
P09012	U1 small nuclear ribonucleoprotein A SNRPA	158.8	1	3	3	5	2.37	1		1.186	1		0.829	1		0.418	1		1
P05783	Keratin, type I cytoskeletal 18 KRT18	5206.97	9	27	29	174	2.272	41	95.8	0.972	41	43.4	0.572	41	31.6	0.383	41	137.7	1
P00488	Coagulation factor XIII A chain F13A1	139.56	1	1	1	5	2.158	3	48.1	1.203	3	5.4	0.503	3	4.6	0.989	3	24.5	1
P67936	Tropomyosin alpha-4 chain TPM4	22.49	1	1	1	1	2.001	1		1.061	1		1.225	1		0.574	1		1
P62995	Transformer-2 protein homolog beta TRA2B	551.69	1	3	3	13	1.999	11	31.7	1.136	11	17.8	1.314	11	23.9	0.409	11	80.2	1
P08670	Vimentin VIM	117.17	7	3	4	8	1.93	7	93.7	1.154	7	48	0.955	7	28.5	0.539	7	145.8	1
Q16629	Serine/arginine-rich splicing factor 7 SRSF7	141.89	1	3	3	5	1.914	3	0.5	1.191	3	9.3	1.366	3	34.5	0.562	3	223.5	1
P35237	Serpin B6 SERPINB6	808.17	1	8	8	27	1.66	11	29	0.985	11	22.2	1.404	11	22.8	0.96	11	37.6	1
Q9NVD7	Alpha-parvin PARVA	657.83	1	3	3	13	1.435	3	20.4	0.944	3	0.3	0.887	3	4.3	1.636	3	18	1
P01008	Antithrombin-III SERPINC1	3745.15	1	17	17	89	1.414	21	42.2	1.435	21	47.3	1.247	21	46.9	1.605	21	50.5	1
O00264	Membrane-associated progesterone receptor component 1 PGRMC1	529.48	1	3	3	16	1.372	16	9.3	1.746	16	20	0.864	16	10.5	0.308	16	39.2	1
P51884	Lumican LUM	113.54	1	1	1	6	1.333	5	63.2	0.537	5	52.7	1.01	5	9.1	18.962	5	53.1	1
P34741	Syndecan-2 SDC2	42.95	1	1	1	1	1.322	1		3.383	1		1.443	1		0.194	1		1
P11142	Heat shock cognate 71 kDa protein HSPA8	2026.68	1	6	6	84	1.191	56	42.8	0.824	56	29.4	1.919	56	34.6	0.676	56	77.4	1
Q15043	Zinc transporter ZIP14 SLC39A14	119.81	1	2	2	4	1.132	1		0.422	1		0.113	1		0.272	1		1
P68871	Hemoglobin subunit beta HBB	1770.45	5	7	7	93	1.129	23	36.1	1.578	23	40	0.816	23	27.1	0.886	23	33.9	1
P08107	Heat shock 70 kDa protein 1A/1B HSPA1A	888.11	2	2	2	167	1.019	1		0.61	1		1.374	1		0.812	1		1
P02787	Serotransferrin TF	567.05	1	2	2	12	0.982	1		0.592	1		0.978	1		0.882	1		1
P02792	Ferritin light chain FTL	2602.39	1	12	12	85	0.936	18	15.9	1.382	18	12	1.684	18	12.4	3.358	18	21.7	1
Q14974	Importin subunit beta-1 KPNB1	327.89	1	2	2	7	0.931	1		1.131	1		0.573	1		1.164	1		1
P48637	Glutathione synthetase GSS	234.04	1	1	1	7	0.927	7	3.5	1.021	7	3.5	2.198	7	101.4	1.164	7	3.4	1
O15258	Protein RER1	37.81	1	1	1	1	0.921	1		1.038	1		1.506	1		1.171	1		1
P05141	ADP/ATP translocase 2 SLC25A5	52.75	1	1	1	1	0.918	1		0.347	1		0.88	1		1.102	1		1
O15260	Surfeit locus protein 4 SURF4	338.68	1	4	4	12	0.906	3	1.6	0.648	3	10.4	1.179	3	173.5	1.165	3	4.5	1
P07942	Laminin subunit beta-1	283.15	2	3	3	10	0.901	8	46.6	0.499	8	10.8	1.368	8	38.1	1.016	8	59	1

Accession	Protein																		
O60884	DnaJ homolog subfamily A member 2 DNAJA2	40.2	1	1	1	1	0.899	1		0.47	1		1.096	1		1.157	1		1
P35579	Myosin-9 MYH9	321.49	1	3	3	12	0.892	7	7.5	0.48	7	31.7	1.174	7	5.5	0.735	7	44.8	1
O75400	Pre-mRNA-processing factor 40 homolog A PRPF40A	449.21	1	2	2	18	0.886	18	23.4	0.597	18	30.1	0.747	18	50.5	0.703	17	77.1	1
Q01130	Serine/arginine-rich splicing factor 2 SRSF2	157.17	1	2	2	6	0.861	6	55.4	0.652	6	17.6	0.857	6	11.4	0.701	3	89.3	1
P17612	cAMP-dependent protein kinase catalytic subunit alpha PRKACA	187.51	2	2	2	5	0.855	4	62.9	0.644	4	57	0.91	4	3.1	0.668	4	24.1	1
Q00610	Clathrin heavy chain 1 CLTC	59.17	1	1	1	3	0.724	2	38.3	0.613	2	105	0.687	2	8.1	0.722	2	80.5	1
Q86VP6	Cullin-associated NEDD8-dissociated protein 1 CAND1	84.6	1	1	1	1	0.67	1		0.446	1		0.71	1		1.39	1		1
P01011	Alpha-1-antichymotrypsin SERPINA3	7285.6	1	20	20	215	0.614	101	64.9	1.076	101	27	1.247	101	19.4	1.367	101	34.7	1
P01024	Complement C3	653.1	1	4	4	13	0.609	2	116.1	0.561	2	40.6	0.789	2	16.9	2.592	2	79.3	1
P43652	Afamin AFM	153.57	1	1	1	7	0.367	4	36.9	1.064	4	8.9	0.855	4	60.3	1.852	4	41.2	1
P02675	Fibrinogen beta chain FGB	44.97	1	1	1	1	1.444	1		1.015	1		0.942	1		1.228	1		0
P63267	Actin, gamma-enteric smooth muscle ACTG2	5945.74	5	2	9	224	1.41	134	49.4	1.045	134	22.2	0.826	134	32.4	1.164	134	27.3	0
P02790	Hemopexin HPX	120.54	1	2	2	4	1.357	3	3.1	1.268	3	3.9	1.067	3	7.2	2.572	3	20.8	0
P14866	Heterogeneous nuclear ribonucleoprotein L HNRNPL	220.81	1	1	1	6	1.321	6	28.3	1.078	6	48.8	0.943	6	23.5	0.514	6	86	0
P11021	78 kDa glucose-regulated protein HSPA5	390.59	1	4	4	9	1.318	4	17.2	1.055	4	11.2	1.148	4	36.9	0.303	4	46.7	0
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2 HNRNPC	307.02	1	1	1	9	1.279	7	25.8	0.729	7	12.4	0.95	7	27.4	0.249	7	12.6	0
P60709	Actin, cytoplasmic 1 ACTB	9252.18	7	7	13	271	1.258	167	47.8	1.045	167	18.7	0.842	167	38.5	1.164	167	22.2	0
P13611	Versican core protein VCAN	72.25	1	1	1	2	1.215	2	8.8	1.238	2	2.3	0.99	2	19.8	0.787	2	4.4	0
P27824	Calnexin CANX	4600.94	1	4	4	149	1.162	122	34.3	0.969	122	33.9	1.247	122	35	0.433	122	46.4	0
P14625	Endoplasmin HSP90B1	31.49	1	1	1	1	1.084	1		0.697	1		1.05	1		0.168	1		0
Q9BYX7	Putative beta-actin-like protein 3 POTEKP	855.11	1	1	2	33	1.071	12	31.8	0.967	12	27.6	0.742	12	30.7	1.313	12	29.6	0
P0CG48	Polyubiquitin-C UBC	330.11	4	4	4	19	1.038	14	15.4	0.869	14	18.2	0.721	14	23	0.718	14	27.3	0
P09211	Glutathione S-transferase P GSTP1	664.79	1	3	3	20	1.032	10	23	0.888	10	30.7	1.038	10	38.9	0.724	10	40.5	0
P07900	Heat shock protein HSP 90-alpha HSP90AA1	221.63	3	2	4	7	0.946	2	4.5	0.708	2	12.1	0.876	2	22.9	0.778	2	65.6	0
Q15084	Protein disulfide-isomerase A6 PDIA6	671.87	1	2	2	30	0.934	11	9.2	1.098	11	15	1.042	11	91.2	1.164	11	7.7	0
P46940	Ras GTPase-activating-like protein IQGAP1	258.82	1	3	3	6	0.931	4	4.5	1.031	4	21.9	1.242	4	0.5	1.127	4	20.4	0
P02751	Fibronectin FN1	1978.06	1	11	11	59	0.93	7	6.4	0.884	7	36.2	0.824	7	16.2	1.322	7	19	0
P14384	Carboxypeptidase M CPM	38.45	1	1	1	1	0.929	1		1.092	1		0.745	1		1.164	1		0
P08697	Alpha-2-antiplasmin SERPINF2	285.26	1	3	3	8	0.928	1		0.672	1		1.255	1		1.141	1		0

Accession	Description	Score	#Prot	#Unique	#Pep	#PSM	Ratio	Count	Var	Ratio	Count	Var	Ratio	Count	Var	Ratio	Count	Var	Score
P02730	Band 3 anion transport protein SLC4A1	693.08	1	4	4	40	0.927	3	0	0.836	3	2.2	1.25	3	0	1.154	3	0.1	0
P08236	Beta-glucuronidase GUSB	4745.09	1	21	21	109	0.927	16	0.2	1.045	16	6	1.247	16	15.3	1.164	16	0.8	0
P08962	CD63 antigen CD63	48.71	1	1	1	1	0.927	1		1.045	1		1.247	1		1.164	1		0
P02452	Collagen alpha-1(I) chain COL1A1	40	1	1	1	1	0.927	1		1.045	1		1.247	1		1.164	1		0
P39656	DDOST	53.76	1	1	1	1	0.927	1		1.045	1		1.247	1		1.164	1		0
P07099	Epoxide hydrolase 1 EPHX1	112.01	1	1	1	3	0.927	1		1.045	1		1.247	1		1.164	1		0
Q13228	Selenium-binding protein 1 SELENBP1	883.46	1	10	10	40	0.927	22	29.1	0.929	22	31.3	1.271	22	8.1	1.893	22	42.4	0
Q13813	Spectrin alpha chain, non-erythrocytic 1 SPTAN1	845.54	1	7	8	37	0.927	10	4.3	1.045	10	4.9	1.316	10	7.9	1.164	10	0.1	0
P29401	Transketolase TKT	92.73	1	1	1	4	0.927	3	37.7	0.928	3	16.7	0.709	3	18.8	1.151	3	1.7	0
P24821	Tenascin TNC	702.7	1	5	5	22	0.915	9	11.1	1.111	9	12.5	1.261	9	16	1.205	9	5.1	0
P50395	Rab GDP dissociation inhibitor beta GDI2	2045.57	2	5	5	60	0.912	51	24.1	0.694	51	66.8	0.827	51	47.1	1.119	51	11	0
P12814	Alpha-actinin-1 ACTN1	537.49	4	6	6	26	0.908	14	28.1	0.922	14	54.4	0.763	14	47.5	1.102	14	43.1	0
P02794	Ferritin heavy chain FTH1	697.43	1	9	9	26	0.902	3	19.3	1.275	3	26.3	1.043	3	45.5	1.692	3	15.7	0
P13667	Protein disulfide-isomerase A4 PDIA4	128.17	1	2	2	6	0.899	4	21	0.819	4	13.7	0.706	4	35.5	0.247	4	19.8	0
P05155	Plasma protease C1 inhibitor SERPING1	45.87	1	1	1	1	0.895	1		1.241	1		1.287	1		2.35	1		0
Q05707	Collagen alpha-1(XIV) chain COL14A1	2474.99	1	11	11	79	0.85	31	13.3	0.974	31	18.5	0.767	31	66.7	3.781	31	73.7	0
P68104	Elongation factor 1-alpha 1 EEF1A1	203.01	3	3	3	11	0.837	8	30.3	0.705	8	14.2	0.876	8	4.6	0.491	8	145	0
P02768	Serum albumin ALB	51389.16	1	24	24	1529	0.837	167	20	0.982	167	32.4	1.247	167	45.8	2.04	167	61.4	0
P26641	Elongation factor 1-gamma EEF1G	228.61	1	3	3	9	0.77	8	53.3	0.768	8	33.1	0.805	8	16.5	0.421	8	216.2	0
P54725	UV excision repair protein RAD23 homolog A	105.1	1	2	2	4	0.754	1		0.754	1		0.855	1		1.07	1		0

Data was generated using Proteome Discoverer version 2.4. Protein changes are shown as median peptide reporter ion ratios for peptides matching that protein. Up-regulation in the different comparisons is displayed in shades of red and down-regulation is displayed in shades of blue. The protein score, numbers of proteins in the protein group, unique peptides, matched peptides, peptide spectral matches, reporter ion count, variability and 'biomarker score' are shown.

The proteins identified were enriched for classical serum proteins (n=18) and cytokeratins (n=12) reflecting contamination of the samples with blood and also perhaps contaminatio n from sample handling/processing. Extracelluar matrix proteins, cytoskeletal proteins, chaperones, mRNA processing poroteins and several muscle-specific proteins were also represented, the latter indicative of the presence of myometrial cells in the samples. When comparing endometriosis and controls in the secretory phase, 28 proteins were identified as being overexpressed in, whilst 11 were under-expressed (Table 5.2). In the comparison of the endometriosis and pain groups, 14 proteins were over-expressed, whilst 23 were under-expressed.12 and 13 proteins were similarly differentially regulated in proliferative phase endometriosis *versus* control samples . There were six non-serum proteins with a highest biomarker score of 3; tropomyosin alpha-1 chain (TPM1), annexin A2 (ANXA2), plasmolipin (PLLP), calpain-1 catalytic subunit (CAPN1), coatomer subunit beta (COPB1) and gelsolin (GSN).

Table 5.2 Summary of numbers of differentially expressed proteins in each comparison

	Over-expression	Under-expression	Total
ES vs. CS	28	11	39
ES vs. PS	14	23	37
EP vs. CP	12	13	25
EcS vs. ES	21	36	57

TPM1 was up-regulated in endometriosis compared to pain in the secretory phase**.** This protein was also observed in the 2D-DIGE profiling, showing similar changes across the groups. Differential lumican expression was also found, although its altered expression in endometriosis *versus* pain was somewhat different between the experiments. Other proteins of interest included plasmolipin (PLLP) and membrane-associated progesterone receptor component 1 (PGRMC1).

5.2.1 Profiling of endometrium tissues using a TMT peptide-based 3D-LC-MS/MS strategy

Profiling of pooled samples in the six clinical groups using a 3D-LC-MS/MS workflow with TMT peptide labelling identified a much more respectable 1,581 proteins groups, of which 1,433 (91%) had quantitative information across all six clinical groups. The full list of protein IDs, peptide information and fold-changes across the six groups can be found in Additional Information. There were a considerable number of protein changes identified (Table 5.3 and Figure 5.4). Ineed the average overall percentage of proteins changing in expression greater than 1.5-fold was 28% and as high as 49%. For example, 474 proteins were differentially expressed at >1.5-fold between the secretory phase endometriosis and control groups with 232 proteins over-expressed and 242 under-expressed in this comparison. The largest differences were seen when comparing secretory phase ectopic and eutopic tissue with 313 proteins over-expressed and 386 under-expressed. This may be attributed to greater sample heterogeneity for the ectopic tissue samples and also variable levels of blood contamination. The lowest number of changes (14%) was observed in the proliferative phase endometriosis *versus* control group, possibly indicating that these samples are more homogeneous.

Table 5.3 Summary of number of differentially expressed proteins (>1.5-fold) in each tissue comparison

	Over-expressed	Under-expressed	Total
ES vs. CS	232	242	474
ES vs. PS	298	97	395
EP vs. CP	82	120	202
EcS vs. ES	313	386	699
PS vs. CS	191	337	528
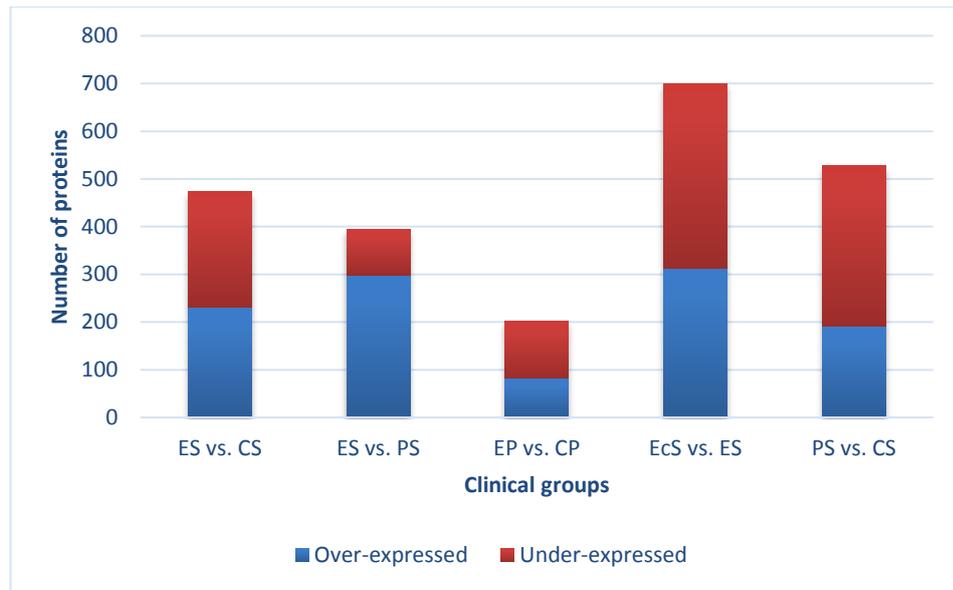
Figure 5.4 Graphical display of the number of differentially expressed proteins in each comparison (>1.5 fold)

A scoring system was devised to rank the proteins to aid in selection of biomarker candidates for verification. This involved summing of individual scores for each expression ratio, reporter ion count, variability and number of unique peptide sequences for each protein. An expression pattern score was added to this by reference to membership of one of five expression cluster groups, with a higher score given to those proteins differentially expressed in the endometriosis group versus both control groups, irrespective of cycle stage (Figure 5.5). The top scoring proteins are shown in Table 5.4.
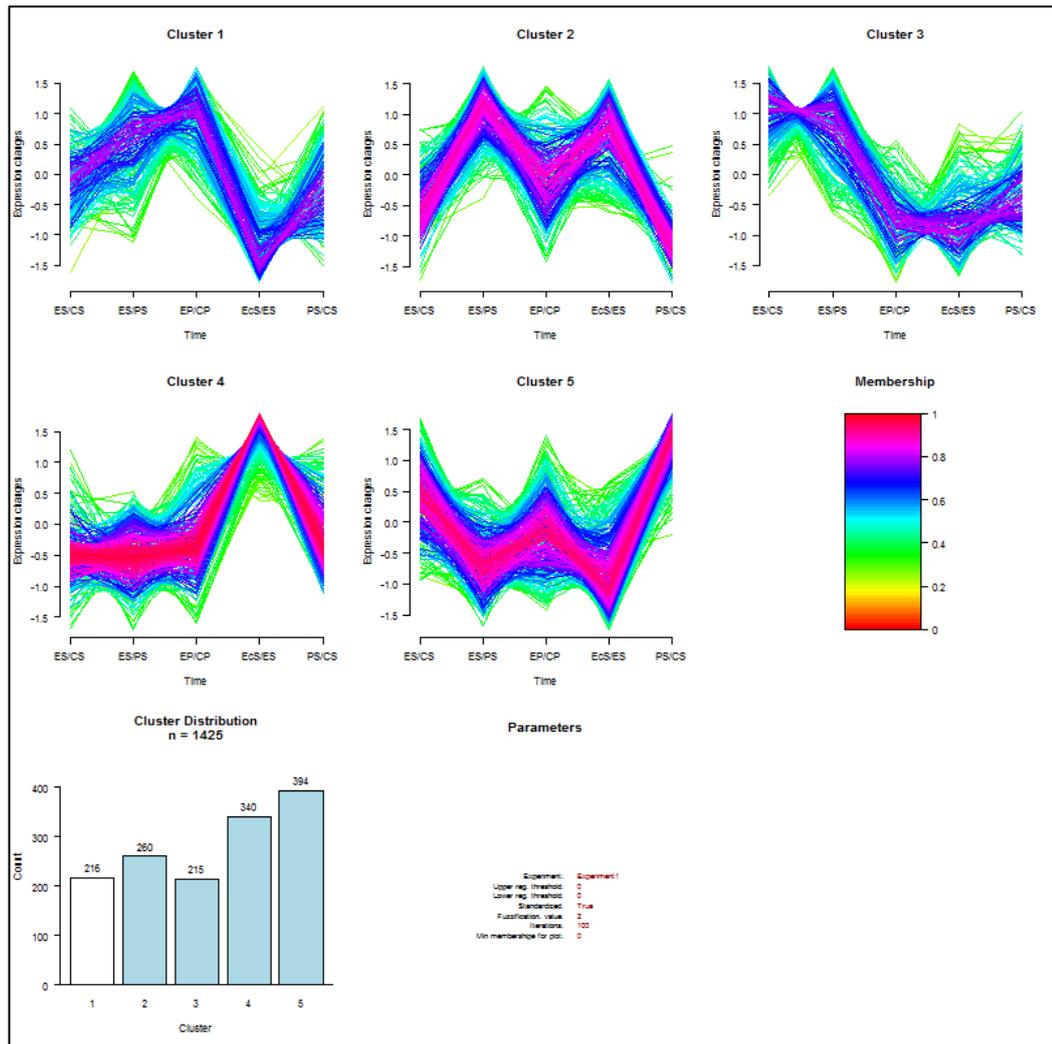
Cluster	Score	Name
Cluster 1	8	cycle-independent; up-regulated proteins; ectopic down
Cluster 2	1	pain-specific; down-regulated
Cluster 3	6	cycle-specific
Cluster 4	4	cycle-independent; low fold-changes; ectopic up
Cluster 5	2	pain specific; up-regulated

Figure 5.5 Protein expression pattern clustering.
GPRoX software was used to generate proteinexpression pattern clusters based on abundance ratios across the groups. Clusters were given a cluster score (lower panel) which was incorprated into the biomarker score.

Accession	Description	Biomarker Score	Classical serum protein	RBC Protein	POI	Σ# Proteins in group	Score	ΣCoverage	# AAs	MW [kDa]	calc. pI	Σ# Unique Peptides	Σ# Peptides	Σ# PSMs	Ratio 130/126 ES/CS	Ratio 130/128 ES/PS	Ratio 131/127 EP/CP	Ratio 129/130 EcS/ES	Ratio 128/126 PS/CS
P02042	Hemoglobin subunit delta HBD - [HBD_HUMAN]	24.2		1		2	53600	92.52	147	16.0	8.05	7	14	3664	3.357	3.512	0.921	0.936	1.005
P14384	Carboxypeptidase M CPM - [CBPM_HUMAN]	24			1	1	154	6.32	443	50.5	7.36	2	2	6	1.619	2.525	2.451	0.319	0.633
P02549	Spectrin alpha chain, erythrocytic 1 SPTA1 - [SPTA1_HUMAN]	24		1		1	1342	13.93	2419	279.8	5.05	22	22	38	2.907	3.041	0.938	0.633	0.871
P02671	Fibrinogen alpha chain FGA - [FIBA_HUMAN]	23.8	1			1	3116	30.83	866	94.9	6.01	25	26	103	8.515	13.491	0.920	0.862	0.809
P02730	Band 3 anion transport protein SLC4A1 - [B3AT_HUMAN]	23.6		1		1	1223	14.93	911	101.7	5.19	9	9	33	2.639	3.066	0.921	0.533	1.074
P02675	Fibrinogen beta chain FGB - [FIBB_HUMAN]	23.4	1			1	1208	46.03	491	55.9	8.27	16	16	62	5.924	9.261	0.855	1.264	0.737
P32119	Peroxiredoxin-2 PRDX2 - [PRDX2_HUMAN]	23.2				1	2422	52.53	198	21.9	5.97	14	15	112	2.649	3.043	1.032	0.815	0.847
P69905	Hemoglobin subunit alpha HBA1 - [HBA_HUMAN]	23.2		1		2	111145	78.17	142	15.2	8.68	15	15	4226	2.924	3.171	1.001	0.895	1.127
P19013	Keratin, type II cytoskeletal 4 KRT4 - [K2C4_HUMAN]	23				7	678	19.66	534	57.2	6.61	5	11	28	2.931	8.667	2.032	0.317	0.298
P68871	Hemoglobin subunit beta HBB - [HBB_HUMAN]	23		1		2	142146	84.35	147	16.0	7.28	9	16	6548	2.853	3.499	0.942	0.892	0.937
P11277	Spectrin beta chain, erythrocytic SPTB - [SPTB1_HUMAN]	23		1		2	1333	11.00	2137	246.3	5.27	17	19	42	2.560	3.382	0.911	0.837	0.830
P26599	Polypyrimidine tract-binding protein 1 PTBP1 - [PTBP1_HUMAN]	22.6			1	3	264	23.92	531	57.2	9.17	7	8	11	2.646	2.771	0.953	0.377	0.787
P02679	Fibrinogen gamma chain FGG - [FIBG_HUMAN]	22.4	1			1	1117	32.89	453	51.5	5.62	11	11	50	7.880	13.703	0.727	0.563	0.459
P04040	Catalase CAT - [CATA_HUMAN]	22.2				1	1220	29.79	527	59.7	7.39	12	12	38	2.772	2.800	1.043	1.150	1.176
P00918	Carbonic anhydrase 2 CA2 - [CAH2_HUMAN]	22.2				1	836	36.92	260	29.2	7.40	8	9	32	2.485	3.316	0.819	1.243	0.800
P00491	Purine nucleoside phosphorylase PNP - [PNPH_HUMAN]	21.4				1	166	10.73	289	32.1	6.95	3	3	6	3.715	3.221	0.942	0.573	1.195
P16157	Ankyrin-1 ANK1 - [ANK1_HUMAN]	21.2		1		1	301	4.25	1881	206.1	6.01	5	5	10	3.054	3.352	0.945	0.866	0.871
P29373	Cellular retinoic acid-binding protein 2 CRABP2 - [RABP2_HUMAN]	21			1	1	83	21.74	138	15.7	5.40	2	2	3	2.078	1.612	2.544	1.436	1.382
Q14764	Major vault protein MVP - [MVP_HUMAN]	21			1	1	151	4.03	893	99.3	5.48	3	3	7	0.143	0.158	0.425	10.978	0.894
P69891	Hemoglobin subunit gamma-1 HBG1 - [HBG1_HUMAN]	20.8		1		2	17502	64.63	147	16.1	7.20	1	10	1756	3.547	3.546	0.905	1.098	1.007
P69892	Hemoglobin subunit gamma-2 HBG2 - [HBG2_HUMAN]	20.8		1		2	17579	64.63	147	16.1	7.20	1	10	1759	3.546	3.545	0.905	1.099	1.007
O75884	Putative hydrolase RBBP9 - [RBBP9_HUMAN]	20.4			1	1	78	17.20	186	21.0	6.20	2	2	2	0.305	0.293	1.006	4.505	1.028
Q01995	Transgelin TAGLN - [TAGL_HUMAN]	20.2			1	2	2119	69.15	201	22.6	8.84	17	17	95	1.105	1.324	2.171	18.832	1.240
Q01105	Protein SET - [SET_HUMAN]	20.2				2	937	33.79	290	33.5	4.32	6	6	29	0.914	0.919	1.084	0.330	0.881
A8MVG2	Putative selection and upkeep of intraepithelial T-cells protein 1 homolog SKINTL - [SKIT1_HUMAN]	20.2				1	1125	3.21	218	25.4	5.87	1	1	138	4.314	6.142	0.829	0.700	0.773
P00738	Haptoglobin HP - [HPT_HUMAN]	20	1			1	714	35.22	406	45.2	6.58	15	15	33	1.459	2.619	1.699	3.321	0.577
P00915	Carbonic anhydrase 1 CA1 - [CAH1_HUMAN]	20				1	5011	57.47	261	28.9	7.12	14	14	263	2.086	2.902	0.933	1.285	0.627
Q00796	Sorbitol dehydrogenase SORD - [DHSO_HUMAN]	20				1	181	11.48	357	38.3	7.97	4	4	9	0.538	0.752	0.543	0.299	0.695
O94788	Retinal dehydrogenase 2 ALDH1A2 - [AL1A2_HUMAN]	19.4				3	301	23.94	518	56.7	6.05	9	10	12	0.838	0.898	1.542	0.456	0.903
Q13308	Inactive tyrosine-protein kinase 7 PTK7 - [PTK7_HUMAN]	19.2				1	362	2.71	1070	118.3	7.09	3	3	13	0.510	0.677	1.497	0.293	0.807
P16949	Stathmin STMN1 - [STMN1_HUMAN]	19.2				3	293	41.61	149	17.3	5.97	7	7	16	1.044	0.926	1.411	0.471	0.994
P10909	Clusterin CLU - [CLUS_HUMAN]	19.2	1			1	649	13.14	449	52.5	6.27	7	7	22	0.281	1.831	1.179	4.014	0.146
Q6SPF0	Atherin SAMD1 - [SAMD1_HUMAN]	19			1	1	25	2.23	538	56.0	7.58	1	1	1	3.178	3.470	1.757	0.566	0.903
P06703	Protein S100-A6 S100A6 - [S10A6_HUMAN]	19			1	1	762	55.56	90	10.2	5.48	5	5	47	1.379	1.653	0.795	3.364	0.774
P0DJI8	Serum amyloid A-1 protein SAA1 - [SAA1_HUMAN]	19	1			2	77	11.48	122	13.5	6.79	1	1	2	9.667	10.125	0.594	1.832	0.942
Q13263	Transcription intermediary factor 1-beta TRIM28 - [TIF1B_HUMAN]	18.6			1	1	320	5.99	835	88.5	5.77	7	7	19	0.841	0.800	1.135	0.497	1.086
P09466	Glycodelin PAEP - [PAEP_HUMAN]	18.6			1	1	698	12.78	180	20.6	5.57	2	2	18	0.622	0.863	0.950	0.210	0.705

Accession	Protein	Score					Coverage		MW	pI								
P04843	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 RPN1 - [RPN1_HUMAN]	18.6			1	1752	30.31	607	68.5	6.38	15	16	70	0.893	1.171	0.936	0.396	0.865
P11171	Protein 4.1 EPB41 - [41_HUMAN]	18.6	1		1	273	6.48	864	97.0	5.58	4	4	7	2.388	3.231	0.833	0.823	0.764
P02751	Fibronectin FN1 - [FINC_HUMAN]	18.4			1	251	4.07	2386	262.5	5.71	8	8	9	2.706	1.288	0.630	2.104	2.005
P07741	Adenine phosphoribosyltransferase APRT - [APT_HUMAN]	18.2			1	729	32.78	180	19.6	6.02	4	5	31	2.260	1.612	0.921	1.436	1.258
P07305	Histone H1.0 H1F0 - [H10_HUMAN]	18.2			1	120	8.76	194	20.9	10.84	2	2	6	3.084	1.811	0.828	2.121	1.675
Q96KP4	Cytosolic non-specific dipeptidase CNDP2 - [CNDP2_HUMAN]	18.2			1	1519	48.84	475	52.8	5.97	16	16	43	0.724	1.078	0.535	0.333	0.710
P0CG48	Polyubiquitin-C UBC - [UBC_HUMAN]	18.2			4	681	44.67	685	77.0	7.66	4	4	23	0.282	1.407	0.430	4.231	0.230
P06401	Progesterone receptor PGR - [PRGR_HUMAN]	18		1	1	37	1.39	933	98.9	6.49	1	1	1	1.415	1.762	2.420	0.321	0.792
P49959	Double-strand break repair protein MRE11A - [MRE11_HUMAN]	18			1	42	1.84	708	80.5	5.90	1	1	2	22.475	11.335	1.253	0.161	1.956
Q92598	Heat shock protein 105 kDa HSPH1 - [HS105_HUMAN]	18			1	26	0.82	858	96.8	5.39	1	1	1	5.432	3.687	1.231	0.390	1.453
P21964	Catechol O-methyltransferase COMT - [COMT_HUMAN]	18			1	31	2.95	271	30.0	5.47	1	1	1	2.446	3.077	0.955	0.326	0.784
Q8IZ73	RNA pseudouridylate synthase domain-containing protein 2 RPUSD2 - [RUSD2_HUMAN]	18			1	35	1.28	545	61.3	7.17	1	1	3	24.809	15.889	0.571	0.367	1.540
P00167	Cytochrome b5 CYB5A - [CYB5_HUMAN]	18			1	692	55.97	134	15.3	4.96	7	7	29	0.835	0.880	0.205	0.979	0.948
P14678	Small nuclear ribonucleoprotein-associated proteins B and B' SNRPB - [RSMB_HUMAN]	17.8			2	83	9.58	240	24.6	11.19	3	3	4	0.779	1.058	1.350	0.072	0.677
Q14980	Nuclear mitotic apparatus protein 1 NUMA1 - [NUMA1_HUMAN]	17.8			1	311	6.05	2115	238.1	5.78	10	11	15	0.930	0.743	1.327	0.452	1.063
P05387	60S acidic ribosomal protein P2 RPLP2 - [RLA2_HUMAN]	17.8			1	219	46.09	115	11.7	4.54	4	5	6	0.292	1.764	1.118	0.268	0.204
P14625	Endoplasmin HSP90B1 - [ENPL_HUMAN]	17.8			2	2266	37.24	803	92.4	4.84	27	29	86	0.805	1.027	1.071	0.482	0.768
P50454	Serpin H1 SERPINH1 - [SERPH_HUMAN]	17.6	1		1	2078	42.58	418	46.4	8.69	17	17	82	0.955	0.963	1.539	0.678	0.904
P17661	Desmin DES - [DESM_HUMAN]	17.6			4	4279	67.23	470	53.5	5.27	24	30	216	0.904	0.930	1.381	3.056	1.063
P59665	Neutrophil defensin 1 DEFA1 - [DEF1_HUMAN]	17.6			2	137	20.21	94	10.2	6.99	3	3	11	0.325	2.476	1.315	4.989	0.133
P02787	Serotransferrin TF - [TRFE_HUMAN]	17.6	1		1	2077	49.86	698	77.0	7.12	29	29	75	0.778	1.501	1.184	3.285	0.503
P09651	Heterogeneous nuclear ribonucleoprotein A1 HNRNPA1 - [ROA1_HUMAN]	17.6			2	2159	38.44	372	38.7	9.13	9	13	68	0.468	1.240	1.127	0.316	0.303
P49756	RNA-binding protein 25 RBM25 - [RBM25_HUMAN]	17.6			1	156	2.37	843	100.1	6.32	3	3	12	2.260	1.599	1.122	1.436	1.382
Q00610	Clathrin heavy chain 1 CLTC - [CLH1_HUMAN]	17.6			2	1215	17.79	1675	191.5	5.69	22	22	55	0.628	0.801	0.994	0.545	0.842
Q05682	Caldesmon CALD1 - [CALD1_HUMAN]	17.4			1	677	7.06	793	93.2	5.66	6	6	25	1.162	0.950	1.260	7.685	1.223
Q14257	Reticulocalbin-2 RCN2 - [RCN2_HUMAN]	17.4			1	539	27.76	317	36.9	4.40	6	6	14	1.054	0.889	1.200	0.488	0.990
P02545	Prelamin-A/C LMNA - [LMNA_HUMAN]	17.4			1	3831	50.30	664	74.1	7.02	39	41	174	0.918	0.801	1.052	3.330	1.148
P01011	Alpha-1-antichymotrypsin SERPINA3 - [AACT_HUMAN]	17.4	1		1	933	15.37	423	47.6	5.52	7	7	34	0.731	1.056	1.013	3.257	0.716
Q7KZ85	Transcription elongation factor SPT6 SUPT6H - [SPT6H_HUMAN]	17.4			1	53	0.98	1726	198.9	4.91	1	2	6	2.143	3.101	0.960	0.843	0.682
P23284	Peptidyl-prolyl cis-trans isomerase B PPIB - [PPIB_HUMAN]	17.4			1	692	53.70	216	23.7	9.41	13	13	33	0.837	0.929	0.925	0.565	0.821
Q9HC84	Mucin-5B MUC5B - [MUC5B_HUMAN]	17.2		1	1	471	3.31	5762	596.0	6.64	12	12	18	0.193	3.853	0.386	0.650	0.088
P46783	40S ribosomal protein S10 RPS10 - [RS10_HUMAN]	17.2			1	74	10.91	165	18.9	10.15	2	2	3	0.516	1.987	1.734	0.241	0.263
P27635	60S ribosomal protein L10 RPL10 - [RL10_HUMAN]	17.2			2	125	11.68	214	24.6	10.08	3	3	4	1.097	1.136	1.441	0.242	0.952
P13667	Protein disulfide-isomerase A4 PDIA4 - [PDIA4_HUMAN]	17.2			1	776	26.36	645	72.9	5.07	16	16	31	0.976	0.969	1.191	0.543	0.982

Accession	Description	Score				Count													
Q13247	Serine/arginine-rich splicing factor 6 SRSF6 - [SRSF6_HUMAN]	17.2				3	242	18.02	344	39.6	11.43	6	6	12	0.882	0.930	0.966	0.446	0.982
P27797	Calreticulin CALR - [CALR_HUMAN]	17.2				1	1868	29.02	417	48.1	4.44	11	11	71	1.036	1.223	0.954	0.540	0.877
P04921	Glycophorin-C GYPC - [GLPC_HUMAN]	17.2		1		1	366	20.31	128	13.8	4.84	1	1	6	2.281	2.069	0.929	1.555	0.983
P54819	Adenylate kinase 2, mitochondrial AK2 - [KAD2_HUMAN]	17.2				1	515	24.27	239	26.5	7.81	5	5	13	0.902	1.066	0.901	0.378	0.743
Q9NR12	PDZ and LIM domain protein 7 PDLIM7 - [PDLI7_HUMAN]	17			1	1	56	2.41	457	49.8	8.41	1	1	1	0.286	0.390	1.499	60.688	0.724
P51884	Lumican LUM - [LUM_HUMAN]	17			1	1	1340	41.72	338	38.4	6.61	13	13	51	1.021	0.952	1.216	14.892	1.033
P60660	Myosin light polypeptide 6 MYL6 - [MYL6_HUMAN]	17			1	4	2726	88.74	151	16.9	4.65	12	12	92	1.378	0.878	1.148	5.003	1.382
P85037	Forkhead box protein K1 FOXK1 - [FOXK1_HUMAN]	17				1	31	1.23	733	75.4	9.32	1	1	2	0.462	2.505	5.066	0.477	0.182
P12111	Collagen alpha-3(VI) chain COL6A3 - [CO6A3_HUMAN]	17				1	687	5.19	3177	343.5	6.68	12	12	27	1.933	1.254	1.209	18.478	1.597
P36955	Pigment epithelium-derived factor SERPINF1 - [PEDF_HUMAN]	17	1			1	506	12.44	418	46.3	6.38	4	4	14	2.224	1.587	1.171	1.458	1.381
P60174	Triosephosphate isomerase TPI1 - [TPIS_HUMAN]	17				1	2705	80.07	286	30.8	5.92	22	22	83	0.853	1.052	1.069	0.635	0.775
P27105	Erythrocyte band 7 integral membrane protein STOM - [STOM_HUMAN]	17		1		1	138	10.42	288	31.7	7.88	2	2	5	3.050	3.058	1.033	0.941	0.984
Q9BT73	Proteasome assembly chaperone 3 PSMG3 - [PSMG3_HUMAN]	17				1	35	8.20	122	13.1	7.88	1	1	1	2.647	3.041	0.989	1.794	0.858
P49411	Elongation factor Tu, mitochondrial TUFM - [EFTU_HUMAN]	17				1	1069	28.32	452	49.5	7.61	10	10	29	0.948	0.903	0.981	0.622	1.002
A8MYA2	Uncharacterized protein CXorf49 - [CX049_HUMAN]	17				1	4395	2.33	514	54.4	9.03	1	1	627	2.260	1.612	0.921	1.436	1.382
P16452	Erythrocyte membrane protein band 4.2 EPB42 - [EPB42_HUMAN]	17		1		1	228	2.75	691	77.0	8.09	1	1	15	2.722	2.844	0.827	0.573	0.944
P05023	Sodium/potassium-transporting ATPase subunit alpha-1 ATP1A1 - [AT1A1_HUMAN]	17				6	424	8.31	1023	112.8	5.49	8	8	17	0.779	0.658	0.600	0.565	0.716
P02766	Transthyretin TTR - [TTHY_HUMAN]	16.8	1			1	1022	69.39	147	15.9	5.76	8	8	23	0.974	1.938	1.293	4.226	0.496
O60701	UDP-glucose 6-dehydrogenase UGDH - [UGDH_HUMAN]	16.8				1	305	21.05	494	55.0	7.12	7	7	10	0.911	1.047	1.191	0.426	0.716
P09012	U1 small nuclear ribonucleoprotein A SNRPA - [SNRPA_HUMAN]	16.8				2	102	9.22	282	31.3	9.83	3	3	5	0.792	1.132	1.174	0.305	0.669
P06748	Nucleophosmin NPM1 - [NPM_HUMAN]	16.8				1	995	47.28	294	32.6	4.78	11	11	51	0.821	0.962	1.022	0.524	0.834
P09525	Annexin A4 ANXA4 - [ANXA4_HUMAN]	16.8				1	2615	63.32	319	35.9	6.13	21	22	104	0.856	1.037	0.881	0.642	0.716
Q15293	Reticulocalbin-1 RCN1 - [RCN1_HUMAN]	16.8				1	390	23.87	331	38.9	5.00	7	7	18	0.774	0.975	0.753	0.538	0.689
P07951	Tropomyosin beta chain TPM2 - [TPM2_HUMAN]	16.6			1	1	4277	49.65	284	32.8	4.70	6	25	178	0.988	1.032	1.239	9.162	0.913
P06702	Protein S100-A9 S100A9 - [S10A9_HUMAN]	16.6				1	286	34.21	114	13.2	6.13	3	3	7	0.406	7.751	2.632	7.255	0.061
P09455	Retinol-binding protein 1 RBP1 - [RET1_HUMAN]	16.6	1			1	290	48.15	135	15.8	5.11	6	6	8	0.425	1.697	1.731	0.749	0.287
Q99729	Heterogeneous nuclear ribonucleoprotein A/B HNRNPAB - [ROAA_HUMAN]	16.6				1	58	2.41	332	36.2	8.21	1	1	2	0.656	1.233	1.594	0.078	0.525
P08865	40S ribosomal protein SA RPSA - [RSSA_HUMAN]	16.6				1	363	29.49	295	32.8	4.87	7	7	15	1.022	1.106	1.363	0.691	1.055
P09493	Tropomyosin alpha-1 chain TPM1 - [TPM1_HUMAN]	16.6				1	4361	47.18	284	32.7	4.74	8	25	186	1.004	1.007	1.245	10.119	0.964
P20774	Mimecan OGN - [MIME_HUMAN]	16.6				1	731	31.88	298	33.9	5.63	9	9	28	0.831	0.798	1.212	20.121	0.905
P01009	Alpha-1-antitrypsin SERPINA1 - [A1AT_HUMAN]	16.6	1			2	1713	33.01	418	46.7	5.59	14	14	75	0.438	1.878	1.144	4.483	0.202
O00264	Membrane-associated progesterone receptor component 1 PGRMC1 - [PGRC1_HUMAN]	16.4			1	1	191	15.90	195	21.7	4.70	4	4	8	1.013	1.035	1.561	0.609	1.228
P24821	Tenascin TNC - [TENA_HUMAN]	16.4			1	1	83	3.13	2201	240.7	4.89	5	5	5	1.817	1.682	1.502	1.151	1.022

Accession	Protein																
Q14978	Nucleolar and coiled-body phosphoprotein 1 NOLC1 - [NOLC1_HUMAN]	16.4		1	120	2.15	699	73.6	9.47	1	1	3	1.122	0.908	2.007	0.337	1.219
P67936	Tropomyosin alpha-4 chain TPM4 - [TPM4_HUMAN]	16.4		2	4288	77.02	248	28.5	4.69	12	30	201	1.014	1.023	1.230	7.301	0.918
Q6NZI2	Polymerase I and transcript release factor PTRF - [PTRF_HUMAN]	16.4		1	334	17.95	390	43.4	5.60	7	8	14	1.313	0.690	1.210	10.697	1.467
P08729	Keratin, type II cytoskeletal 7 KRT7 - [K2C7_HUMAN]	16.4		6	1531	41.36	469	51.4	5.48	14	20	105	0.887	1.008	1.113	2.509	1.000
P21796	Voltage-dependent anion-selective channel protein 1 VDAC1 - [VDAC1_HUMAN]	16.4		2	514	36.75	283	30.8	8.54	8	8	14	1.155	1.067	0.719	0.384	1.031
P36542	ATP synthase subunit gamma, mitochondrial ATP5C1 - [ATPG_HUMAN]	16.2		1	175	15.10	298	33.0	9.22	4	4	9	0.917	1.791	1.261	0.611	0.521
P08133	Annexin A6 ANXA6 - [ANXA6_HUMAN]	16.2		1	2095	38.63	673	75.8	5.60	22	23	75	1.057	0.852	1.191	2.891	1.105
P10809	60 kDa heat shock protein, mitochondrial HSPD1 - [CH60_HUMAN]	16.2		1	2447	56.20	573	61.0	5.87	30	30	93	1.161	0.993	1.104	0.657	0.999
Q14697	Neutral alpha-glucosidase AB GANAB - [GANAB_HUMAN]	16.2		1	512	16.31	944	106.8	6.14	13	13	23	0.685	1.131	1.018	0.603	0.597
P51991	Heterogeneous nuclear ribonucleoprotein A3 HNRNPA3 - [ROA3_HUMAN]	16.2		1	899	27.25	378	39.6	9.01	8	11	30	0.410	1.027	1.018	0.531	0.452
P53999	Activated RNA polymerase II transcriptional coactivator p15 SUB1 - [TCP4_HUMAN]	16.2		1	245	15.75	127	14.4	9.60	2	2	7	0.385	1.386	0.910	0.319	0.256
P52907	F-actin-capping protein subunit alpha-1 CAPZA1 - [CAZA1_HUMAN]	16.2		1	145	18.53	286	32.9	5.69	4	4	7	3.031	1.078	0.885	1.459	1.280
Q02818	Nucleobindin-1 NUCB1 - [NUCB1_HUMAN]	16.2		1	483	25.38	461	53.8	5.25	9	9	18	0.793	0.943	0.861	0.530	0.711
Q02252	Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial ALDH6A1 - [MMSA_HUMAN]	16.2		1	239	8.22	535	57.8	8.50	5	5	11	0.664	0.840	0.548	0.382	0.741
Q6PF15	Kelch-like protein 35 KLHL35 - [KLH35_HUMAN]	16		1	44	2.57	583	62.9	7.83	1	2	4	29.628	4.085	3.745	0.823	7.155
P21333	Filamin-A FLNA - [FLNA_HUMAN]	16		1	2103	23.31	2647	280.6	6.06	41	44	77	1.227	0.967	1.509	2.216	1.357
P51888	Prolargin PRELP - [PRELP_HUMAN]	16		1	84	6.81	382	43.8	9.38	3	3	5	0.504	0.471	1.466	33.466	1.145
P22105	Tenascin-X TNXB - [TENX_HUMAN]	16		2	85	2.12	4289	464.0	5.34	3	3	3	1.548	1.570	1.155	2.143	0.973
P30043	Flavin reductase (NADPH) BLVRB - [BLVRB_HUMAN]	16		1	897	50.00	206	22.1	7.65	7	7	34	1.620	2.564	1.051	1.372	0.710
P07738	Bisphosphoglycerate mutase BPGM - [PMGE_HUMAN]	16		1	54	4.63	259	30.0	6.54	1	1	2	3.572	3.046	0.925	0.847	1.157
P62195	26S protease regulatory subunit 8 PSMC5 - [PRS8_HUMAN]	16		1	240	10.34	406	45.6	7.55	2	3	7	2.260	1.612	0.921	1.436	1.382
P49748	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial ACADVL - [ACADV_HUMAN]	16		1	100	5.80	655	70.3	8.75	3	3	4	2.260	1.612	0.921	1.436	1.382
Q05707	Collagen alpha-1(XIV) chain COL14A1 - [COEA1_HUMAN]	16		1	1210	13.36	1796	193.4	5.30	18	18	49	1.294	0.834	0.898	10.120	1.426
P78527	DNA-dependent protein kinase catalytic subunit PRKDC - [PRKDC_HUMAN]	16		1	30	0.19	4128	468.8	7.12	1	1	1	0.313	0.510	0.733	0.279	0.605
P10412	Histone H1.4 HIST1H1E - [H14_HUMAN]	16		3	431	32.88	219	21.9	11.03	2	10	24	2.349	1.835	0.509	1.323	1.559
P16401	Histone H1.5 HIST1H1B - [H15_HUMAN]	16		1	172	11.06	226	22.6	10.92	2	4	11	2.625	1.714	0.508	0.986	1.506
P0C0S5	Histone H2A.Z H2AFZ - [H2AZ_HUMAN]	15.8		2	237	31.25	128	13.5	10.58	2	4	14	0.642	1.584	1.217	0.744	1.025
P02647	Apolipoprotein A-I APOA1 - [APOA1_HUMAN]	15.8	1	1	1131	46.44	267	30.8	5.76	15	16	57	0.349	2.125	1.169	3.265	0.150
Q9Y490	Talin-1 TLN1 - [TLN1_HUMAN]	15.8		1	2009	14.25	2541	269.6	6.07	23	30	64	1.351	1.474	1.142	1.038	0.949
P04792	Heat shock protein beta-1 HSPB1 - [HSPB1_HUMAN]	15.8		1	1495	43.41	205	22.8	6.40	8	8	69	1.298	0.775	1.077	4.369	1.586
P49419	Alpha-aminoadipic semialdehyde dehydrogenase ALDH7A1 - [AL7A1_HUMAN]	15.8		1	98	8.35	539	58.5	7.99	2	2	3	2.197	1.568	1.042	1.436	1.382
Q13838	Spliceosome RNA helicase DDX39B - [DX39B_HUMAN]	15.8		2	543	36.68	428	49.0	5.67	11	11	24	0.717	0.884	1.023	0.611	0.647

Accession	Description																
Q13151	Heterogeneous nuclear ribonucleoprotein A0 HNRNPA0 - [ROA0_HUMAN]	15.8		1	158	16.39	305	30.8	9.29	3	5	7	0.731	1.743	0.960	0.136	0.413
P01859	Ig gamma-2 chain C region IGHG2 - [IGHG2_HUMAN]	15.8	1	1	976	31.60	326	35.9	7.59	5	10	77	1.195	1.620	0.777	2.720	0.729
P07355	Annexin A2 ANXA2 - [ANXA2_HUMAN]	15.8		2	6078	69.62	339	38.6	7.75	26	26	208	0.754	0.731	0.655	2.435	1.127
P62191	26S protease regulatory subunit 4 PSMC1 - [PRS4_HUMAN]	15.8		1	98	9.77	440	49.2	6.21	2	3	3	2.369	1.891	0.425	0.728	1.236
Q9UNM6	26S proteasome non-ATPase regulatory subunit 13 PSMD13 - [PSD13_HUMAN]	15.8		1	127	9.04	376	42.9	5.81	2	2	2	1.164	4.359	0.342	0.216	0.263
Q8N684	Cleavage and polyadenylation specificity factor subunit 7 CPSF7 - [CPSF7_HUMAN]	15.6		1	41	5.31	471	52.0	8.00	1	1	2	0.488	2.108	3.279	0.777	0.228
Q01081	Splicing factor U2AF 35 kDa subunit U2AF1 - [U2AF1_HUMAN]	15.6		2	64	12.50	240	27.9	8.81	3	3	3	0.645	0.768	2.660	0.939	0.828
Q32P28	Prolyl 3-hydroxylase 1 LEPRE1 - [P3H1_HUMAN]	15.6		1	139	4.62	736	83.3	5.14	3	3	4	1.399	1.023	1.415	0.547	1.260
P16104	Histone H2AX H2AFX - [H2AX_HUMAN]	15.6		2	345	27.27	143	15.1	10.74	2	5	27	0.671	1.935	1.277	0.661	0.830
P31939	Bifunctional purine biosynthesis protein PURH ATIC - [PUR9_HUMAN]	15.6		1	304	18.41	592	64.6	6.71	7	7	9	1.134	1.012	1.144	0.914	1.006
P01042	Kininogen-1 KNG1 - [KNG1_HUMAN]	15.6		1	246	9.78	644	71.9	6.81	8	8	9	0.700	1.596	1.134	2.627	0.498
P50395	Rab GDP dissociation inhibitor beta GDI2 - [GDIB_HUMAN]	15.6		1	870	43.82	445	50.6	6.47	17	17	35	0.887	1.064	1.132	0.860	0.806
Q13595	Transformer-2 protein homolog alpha TRA2A - [TRA2A_HUMAN]	15.6		1	191	10.28	282	32.7	11.27	2	3	7	0.734	1.031	1.108	0.195	0.740
Q12805	EGF-containing fibulin-like extracellular matrix protein 1 EFEMP1 - [FBLN3_HUMAN]	15.6		1	62	6.69	493	54.6	5.07	2	2	3	2.222	1.596	0.921	1.974	1.373
P13804	Electron transfer flavoprotein subunit alpha, mitochondrial ETFA - [ETFA_HUMAN]	15.6		1	148	12.01	333	35.1	8.38	3	3	10	0.657	1.011	0.864	0.526	0.729
P23526	Adenosylhomocysteinase AHCY - [SAHH_HUMAN]	15.6		1	104	7.18	432	47.7	6.34	2	3	5	0.682	0.976	0.658	0.358	0.689
Q9ULC5	Long-chain-fatty-acid--CoA ligase 5 ACSL5 - [ACSL5_HUMAN]	15.6		1	318	13.03	683	75.9	6.92	7	7	11	0.605	0.802	0.413	0.712	0.772
P06454	Prothymosin alpha PTMA - [PTMA_HUMAN]	15.4		1	6865	35.14	111	12.2	3.78	5	5	310	0.900	1.008	1.270	0.485	1.109
P13010	X-ray repair cross-complementing protein 5 XRCC5 - [XRCC5_HUMAN]	15.4		1	617	21.58	732	82.7	5.81	11	11	26	0.522	1.231	1.134	0.508	0.487
P18206	Vinculin VCL - [VINC_HUMAN]	15.4		1	1911	31.48	1134	123.7	5.66	29	30	73	1.404	0.904	1.131	2.429	1.425
Q14498	RNA-binding protein 39 RBM39 - [RBM39_HUMAN]	15.4		1	374	8.11	530	59.3	10.10	3	3	9	1.032	0.849	1.111	0.459	1.068
P21291	Cysteine and glycine-rich protein 1 CSRP1 - [CSRP1_HUMAN]	15.4		1	489	44.56	193	20.6	8.57	6	6	17	0.912	1.039	1.080	21.887	0.735
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 HNRNPA2B1 - [ROA2_HUMAN]	15.4		1	3204	58.36	353	37.4	8.95	17	20	127	0.681	1.069	1.023	0.608	0.531
P02452	Collagen alpha-1(I) chain COL1A1 - [CO1A1_HUMAN]	15.4		1	830	9.02	1464	138.9	5.80	11	11	45	1.124	0.428	0.904	5.938	2.536
P04083	Annexin A1 ANXA1 - [ANXA1_HUMAN]	15.4		1	2360	53.47	346	38.7	7.02	17	17	73	1.096	1.213	0.658	1.887	0.898
O75874	Isocitrate dehydrogenase [NADP] cytoplasmic IDH1 - [IDHC_HUMAN]	15.4		1	1237	52.17	414	46.6	7.01	16	18	42	0.828	0.761	0.235	0.423	1.048
P12004	Proliferating cell nuclear antigen PCNA - [PCNA_HUMAN]	15.2		1	107	16.09	261	28.8	4.69	4	4	7	0.954	0.709	1.790	0.414	1.326
P09382	Galectin-1 LGALS1 - [LEG1_HUMAN]	15.2		1	1621	65.93	135	14.7	5.50	9	9	66	0.664	0.862	1.398	1.653	0.928
Q9BVK6	Transmembrane emp24 domain-containing protein 9 TMED9 - [TMED9_HUMAN]	15.2		1	88	17.02	235	27.3	8.02	1	3	5	0.409	2.208	1.238	0.226	0.201
Q9NYF8	Bcl-2-associated transcription factor 1 BCLAF1 - [BCLF1_HUMAN]	15.2		1	75	3.80	920	106.1	9.98	4	4	5	0.816	1.105	1.232	0.343	0.807
P35749	Myosin-11 MYH11 - [MYH11_HUMAN]	15.2		1	1059	13.13	1972	227.2	5.50	14	20	40	0.986	1.006	1.228	3.765	1.372
P05455	Lupus La protein SSB - [LA_HUMAN]	15.2		1	266	10.29	408	46.8	7.12	4	4	8	0.591	1.193	0.989	0.412	0.522
P17858	ATP-dependent 6-phosphofructokinase, liver type PFKL - [PFKAL_HUMAN]	15.2		1	154	5.90	780	85.0	7.50	3	4	5	0.609	1.716	0.842	0.422	0.462

Accession	Protein	Score		#									R1	R2	R3	R4	R5
P67809	Nuclease-sensitive element-binding protein 1 YBX1 - [YBOX1_HUMAN]	15.2		3	261	11.42	324	35.9	9.88	2	2	3	0.622	0.771	0.738	0.367	0.729
P06396	Gelsolin GSN - [GELS_HUMAN]	15.2		1	1356	37.98	782	85.6	6.28	22	23	45	0.882	0.836	0.582	1.793	1.099
P16402	Histone H1.3 HIST1H1D - [H13_HUMAN]	15.2		3	415	32.58	221	22.3	11.02	2	10	22	2.365	1.851	0.547	1.035	1.506
Q27J81	Inverted formin-2 INF2 - [INF2_HUMAN]	15	1	1	34	1.20	1249	135.5	5.38	1	1	1	1.720	1.565	1.893	1.464	1.084
O95810	Serum deprivation-response protein SDPR - [SDPR_HUMAN]	15	1	1	195	2.12	425	47.1	5.21	1	1	9	1.557	1.373	1.007	6.635	0.872
O60888	Protein CutA CUTA - [CUTA_HUMAN]	15		1	115	15.08	179	19.1	5.50	1	1	1	1.104	1.535	2.324	1.506	0.710
O14773	Tripeptidyl-peptidase 1 TPP1 - [TPP1_HUMAN]	15		1	59	3.73	563	61.2	6.48	2	2	3	0.725	0.888	1.614	0.660	0.789
P09497	Clathrin light chain B CLTB - [CLCB_HUMAN]	15		1	49	9.61	229	25.2	4.64	3	3	3	0.821	1.557	1.234	0.292	0.520
P62318	Small nuclear ribonucleoprotein Sm D3 SNRPD3 - [SMD3_HUMAN]	15		1	69	7.14	126	13.9	10.32	1	1	2	1.189	1.325	1.187	0.268	0.885
P68032	Actin, alpha cardiac muscle 1 ACTC1 - [ACTC_HUMAN]	15		4	12184	61.54	377	42.0	5.39	7	20	431	1.000	0.746	1.154	2.721	1.382
Q9NZN4	EH domain-containing protein 2 EHD2 - [EHD2_HUMAN]	15		1	184	6.08	543	61.1	6.46	3	3	7	0.847	0.699	1.047	7.219	1.154
Q8N2N9	Ankyrin repeat domain-containing protein 36B ANKRD36B - [AN36B_HUMAN]	15		3	28	0.96	1353	153.5	8.85	1	1	2	2.181	3.061	0.979	0.542	0.703
Q9Y6C9	Mitochondrial carrier homolog 2 MTCH2 - [MTCH2_HUMAN]	15		1	163	7.92	303	33.3	7.97	1	1	5	2.260	1.612	0.921	1.436	1.382
Q6XQN6	Nicotinate phosphoribosyltransferase NAPRT1 - [PNCB_HUMAN]	15		1	62	6.88	538	57.5	5.68	2	2	3	3.682	2.629	0.908	0.866	1.381
Q09666	Neuroblast differentiation-associated protein AHNAK - [AHNK_HUMAN]	15		1	1937	26.60	5890	628.7	6.15	50	54	75	0.825	0.717	0.795	2.339	1.109
P61019	Ras-related protein Rab-2A RAB2A - [RAB2A_HUMAN]	15		2	46	9.91	212	23.5	6.54	2	2	2	5.393	2.011	0.783	0.292	2.645
Q9Y6Q1	Calpain-6 CAPN6 - [CAN6_HUMAN]	15		1	29	1.56	641	74.5	7.05	1	1	1	2.162	3.493	0.620	0.661	0.611
A6NDG6	Phosphoglycolate phosphatase PGP - [PGP_HUMAN]	15		1	38	5.92	321	34.0	6.14	1	1	1	3.560	2.633	0.520	0.900	1.334
O43252	Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase 1 PAPSS1 - [PAPS1_HUMAN]	15		1	152	9.94	624	70.8	6.86	5	5	6	0.595	1.194	0.460	0.372	0.529
Q15746	Myosin light chain kinase, smooth muscle MYLK - [MYLK_HUMAN]	15	1	1	263	2.04	1914	210.6	6.15	4	4	9	0.795	0.688	1.454	7.900	1.297

Table 5.4 Summary of top-scoring proteins displaying differential expression between clinical groups. Data was generated using Proteome Discoverer v2.4. Protein changes are shown as peptide reporter ion ratios for peptides matching that protein. Differential expression for each comparison is represented on a continuous colour scale from red (high) to blue (low). A 'biomarker score' combining scores for ratio, count, variability, unique peptide matches and expression pattern is presented for each protein, along with an indication of whether it is a classical serum or red blood cell protein. Yellow accession numbers represent proteins of interest that were considered for verification. The full list is provided as additional information

Protein groups were also annotated as possible blood contaminants that would not be considered for verification; 77 proteins were identified as classical serum proteins and 14 as red blood cell-specific proteins with several of these scoring highly. Identification of serum proteins that had been targeted for depletion suggested incomplete immunodepletion. The presence of these blood proteins may have arisen as a result of contamination during sampling and/or due to vascularisation of the tissue, especially the eutopic tissues. The presence of abundant structural cellular proteins was also noticeable and may reduce coverage of less-abundant proteins. For selection of proteins of interest, weight was given to candidates whose expressions differed significantly between endometriosis and both control groups and in both cycle phases. Selection was also weighted based on prior knowledge of the function of these gene products, those known to be secreted proteins and those for which commercial detection reagents were available for verification.

5.2.2 Proteins of interest

Twenty two proteins were identified as proteins of interest. Carboxypeptidase M (CPM) was the highest scoring non-blood contaminant protein. It was increased in the comparison between endometriosis and controls in both the proliferative and secretory phases (ratio=2.45 and 1.62, respectively). CPM expression was also increased in the endometriosis *versus* pain group in the secretory phase (ratio=2.53), although was lower in ectopic compared to eutopic tissue (ratio=0.32). CPM is an extracellular peptidase attached to the outer membrane by a glycosyl-phosphoinositol anchor. Its function is to cleave C-terminal lysines and arginines from peptides and proteins as part of their processing into mature forms. A role in inflammmation is supported by its regulated expression on macrophages and by the ability of CPM to inactivate anaphylatoxins and alter the receptor specificity for bradykinin. CPM also modulates the activity of chemokines CXCL12 and CCL1 providing further evidence for CPMs involvement in the inflammatory response, but also stem cell mobilisation, monocyte recruitment and cancer. Its increased expression in endometriosis supports a role in inflammation, although its lowered expression in ectopic lesions is somewhat at odds

with this. Lumican was also identified as a candidate biomarker from the profiling. Its expression was highest in ectopic tissue compared to eutopic endometrium in the secretory phase (ratio=14.89) and similar to that seen by 2D-DIGE profiling (Chapter 4), Tropomyosin beta chain (TPM2) expression was also elevated in ectopic tissue compared to eutopic tissue (ratio=9.16) as seen by the 2D-DIGE profiling, although its expression changed little in the other comparisons.

Progesterone receptor (PGR) was identified from this profiling. It was over-expressed in endometriosis compared to controls and the pain group in the secretory phase (ratio=1.45 and 1.76, respectively) and also between endometriosis and controls in the proliferative phase (ratio=2.42). However, PGR expression was lower in ectopic tissue compared to eutopic tissue in the secretory phase (ratio=0.32). The effects of progesterone are mediated via intracellular progesterone receptors that are expressed from a single gene as two protein isoforms, PGRA and PGRB. From the MS data it was not clear which isoform had been identified.

Impaired gene expression in the endometrium of patients with endometriosis was reported to occur throughout the menstrual cycle(Burney et al., 2007) (Burney et al 2007). Some of the dysregulated genes that were reported (FOXO1A, MIG6 and CYP26A1) are known targets of progesterone and their overall pattern of abberrant expression suggested a prolongation of the proliferative phenotype even after ovulation. An incomplete transition of the endometrium from the proliferative phase to the secretory phase is a common characteristic in endometriosis attributed to progesterone resistance. The mechanisms underlying this resistance are poorly understood and may be caused by altered PGR expression, reduced release of progesterone by the corpus luteum or the action of endometrium-specific transcriptional repressors of the PGR. Whether the higher expression of PGR observed herein in eutopic endometrium from endometriosis patients is involved in or a response to progesterone resistance is not clear. However, its lowered expression in ectopic *versus* eutopic tissue does suggest a direct involvement in progesterone resistance that may promote the progression of endometriosis.

Membrane-associated progesterone receptor component 1 (PGRMC1) was also identified as a protein of interest from this data set. Similar to PGR, the protein was over-expressed in proliferative eutopic tissue in endometiosis *versus* control (ratio=1.56) and under-expressed in ectopic tissue (ratio=0.61). PGRMC1 and 2 are porrly characterised but may mediate non-genomic progesterone signalling and act as regulators of steroid hormone synthesis. PGRMC1 is expressed and regulated in the human endometrium in a cycle-dependent manner, but also in endometriotic lesions with the lowest expression reported in the secretory phase (Kao et al 2002, Chen et al 2009, Bunch et al 2014). This may also contribute to the progesterone resistance seen in endometriosis. PGRMC1 is also over-expressed in some cancers (Neubauer et al 2008), suggesting that these receptors are active during rapid growth and proliferation. The results presented herein are in partial agreement with the literature in that it was down-regulated in endometriotic lesions, however this down-regulation was not observed in secretory phase eutopic endometrium. Its up-regulation in the proliferative phase eutopic tissue from endometriosis patients matched that of PGR, suggesting that it supports progesterone-dependent proliferation.

Glycodelin (PAEP) was another protein found to be under-expressed in secretory phase ectopic tissue compared to eutopic tissue (ratio=0.21). This lowered expression may be explained by the fact that PAEP is a progesterone-regulated gene and thus matches the down-regulation of PGR in ectopic tissue. In the normal endometrium, PAEP is down-regulated in the proliferative phase and then highly expressed during the mid-secretory, peaking during the implantation window (Meola et al.,2009). Depending on its glycosylation state, PAEP has roles in contraception, immunosuppression, angiogenesis and apoptosis.

Tenascin C (TNC) was up-regulated in endometriosis compared to controls in both the proliferative and secretory phases (ratios=1.82 and 1.50 respectively). It was also up-regulated in endometriosis compared to pain in the secretory phase (ratio=1.68). Tenascin C is an extracellular matrix glycoprotein that plays a role in cell differentiation, proliferation, invasion and axonal guidance. The role of TNC in the endometrium and endometriosis is not well understood, although its expression has

been demonstrated in human endometrium and is reported to be regulated across the menstrual cycle, with highest expression during the proliferative phase (Sasano et al., 1993, Harrington et al., 1999). The data presented here suggests that this regulated expression may be lost in endometriosis. Expression of TNC has been reported to be upregulated in ectopic lesions compared to eutopic endometrium throughout the cycle (Tan, 2008 #324). Whilst an increased expression of TNC in ectopic tissue was not observed here, its increased expression in the endometrium of cases may drive the disease process by promoting the proliferation and invasion of endometrial cells at ectopic sites.

Transgelin was moderately increased in endometriosis compared to pain in the secretory phase (ratio=1.32) and *versus* controls in the proliferative phase (ratio=2.17), with a high fold-change observed in ectopic *versus* eutopic tissue (ratio=18.83). Transgelin is a smooth muscle actin-binding protein that has been reported to play a potential role in the early invasion of endometrial cells into the mesothelium after initial attachment to the peritoneal wall. One study showed a significant upregulation of transgelin gene expression in ectopic tissue compared to eutopic endometrium (Dos Santos Hidalgo et al., 2011), in agreement with the present findings. Regulation of transgelin expression is thought to be controlled by transforming growth factor beta (TGF-β) in smooth muscle, and thus links transgelin function to cellular proliferation, motility, apoptosis, differentiation, immunologic response and tumourigenesis.

5.2.3 Gene ontology (GO) and enrichment analysis

To try and gain insight into the functional consequences of the altered protein expression, a functional enrichment analysis was undertaken. GO term enrichment for all differentially expressed proteins (>1.5-fold) proved rather ambigous since all terms relating to each gene product are automatically included in the analysis. Several general parent terms for the same processes and containing the same matched genes were therefore recorded as enriched and may have masked more specific annotations. Splitting the genes lists into up-regulated and down-regulated gene products generated

slightly more specific terms with slightly less multiplicity, but the results were still somewhat ambigous. GO enrichment analysis of up-regulated proteins mostly revealed enrichment of the biological processes RNA metabolism, mRNA splicing, translation and gene expression (Table 5.5). This included the over-representation of ribosomal proteins involved in these processes. This may suggest an up-regulation of protein translation in proliferating endometrial cells in endometriosis. Notably however, proteins involved in translational termination were also up-regulated in endometriosis.

KEGG pathway mapping revealed enrichment of genes involved in metabolic pathways, the proteosome and focal adhesion/extra-cellular matrix interactions, although both up- and down-regulated gene products were enriched and pathways were not common to all comparisons. Genes enriched in the proteosome pathway were proteosome components PSMB5, PSMD2, PSMD1, PSMD12, PSMD13, PSMA2, PSMC3, PSMB7, PSMC5, PSMC1 and PSMA4. The proteosome regulates protein turnover in proccesses such as regulation of the cell cycle, cell differentiation, signal transduction, antigen processing, stress signalling, inflammatory responses and apoptosis. The proteosome has not been studied extensively in endometriosis. However the few studies that have examined proteosomal degradation have been aimed at understanding the pathophysiology of the disease (Gonzalez-Ramos et al., 2010) and identifying the proteosome as a potential drug target to block the action of NF-κB (Celik et al., 2008, Gonzalez-Ramos et al., 2008). NF-κB binds to target DNA in the nucleus and upregulates the transcription of genes involved in inflammation, cell growth and apoptosis (Celik et al., 2008). Dysregulation of protein proteosomal expression may potentially contribute to the development of ectopic endometrial lesions due to reduced sensitivity to apoptosis. Furthermore dysregulation of proteosomal function in implants may induce inflammation due to the consequent activation of NF-κB. Thus, the activation of NF-κB via altered proteosome function could play a potential role in the pathophysiology of endometriosis by inducing an inflammatory response and the growth of endometriotic lesions.

Biological Process	ES vs CS		EP vs CP		ES v PS		EcS v ES	
	Upregulated	Downregulated	Upregulated	Downregulated	Upregulated	Downregulated	Upregulated	Downregulated
Cellular macromolecular complex subunit organization	5.34E-13	3.16E-05						
mRNA metabolic process	5.34E-13	2.20E-05	0.0099		4.85E-06			1.10E-25
Cellular protein complex disassembly	3.46E-10	3.34E-05						
Protein complex disassembly	4.90E-10							
Translational termination	6.69E-09		0.0044					
Cellular aromatic/cyclic compound metabolic process			0.0092					
Cellular nitrogen compound metabolic process			0.0139				1.92E-10	
Nucleic acid metabolic process					4.85E-06			
Cellular catabolic process					3.42E-05			5.02E-10
Response to wounding							8.51E-14	
Actin filament-based process/cytoskeleton organisation							8.51E-14	
Acute inflammatory response			0.0197	2.58E-06	2.63E-09		1.57E-12	
KEGG Pathways								
Metabolic pathways	7.26E-11	4.41E-07				1.40E-05		2.75E-20
Ribosome	4.72E-15							3.29E-16
Focal adhesion	2.21E-05				3.74E-05		1.48E-18	6.72E-32
Regulation of actin cytoskeleton		3.77E-06	4.82E-07				2.48E-11	
ECM-receptor interaction					0.0001	1.40E-05	2.92E-09	
Spliceosome	9.27E-06	9.31E-09			0.0005			
Proteasome	3.55E-05	4.29E-05			1.36E-10			
Oxidative phosphorylation	5.93E-05							4.13E-07
Disease associations								
Carcinoma		9.60E-07		2.33E-07		8.97E-05		
Neoplasm Invasiveness		7.87E-06					3.61E-08	5.76E-10
Cancer or viral infections		6.40E-06				0.0007		
Anemia, Hemolytic	2.15E-13				3.69E-12		1.50E-09	
Adhesion		5.10E-08				1.38E-06	1.39E-18	
Hematologic Diseases	1.30E-10							
Protein Deficiency	5.19E-08				2.13E-11			2.18E-10
Cardiovascular Diseases					5.04E-11		5.54E-10	
Huntington's disease		1.10E-06			3.87E-15			2.11E-12

Table 5.5 GO enrichment analysis for altered (≥1.5 fold-change) proteins. Enrichment analyis was performed using WebGestalt. Each clinical group was analysed separately for enrichment of biological process, molecular function, cellular component, GO slim, protein interaction networks, KEGG pathways and disease association. Significantly enriched terms were identified using a hypergeometric test with a Benjamini-Hochberg (BH) correction. The corrected *P* values are shown.

Biological processes and pathways involved in cell-matrix interactions and focal adhesions were also enriched in the up-regulated proteins groups. Several genes involved in focal adhesion formation and regulation (CRKL, VTN, ITGA2, RAP1B, TNXB, PRKCA, TNC, LAMA5, MAPK3) and the extracellular matrix (HSPG2, FN1, VTN, LAMA5, TNC, ITGA2, TNXB, COL6A1) were enriched. Focal adhesions are specialised structures formed between cells and the extracellular matrix (ECM) through transmembrane integrin complexes that link to the actin cytoskeleton and signalling proteins (Figures 5.6 and 5.7). Cell-matrix adhesions play essential roles in involving cell motility, proliferation, differentiation, survival and regulation of gene expression, serving a significant role in tissue and organ morphogenesis and structure.
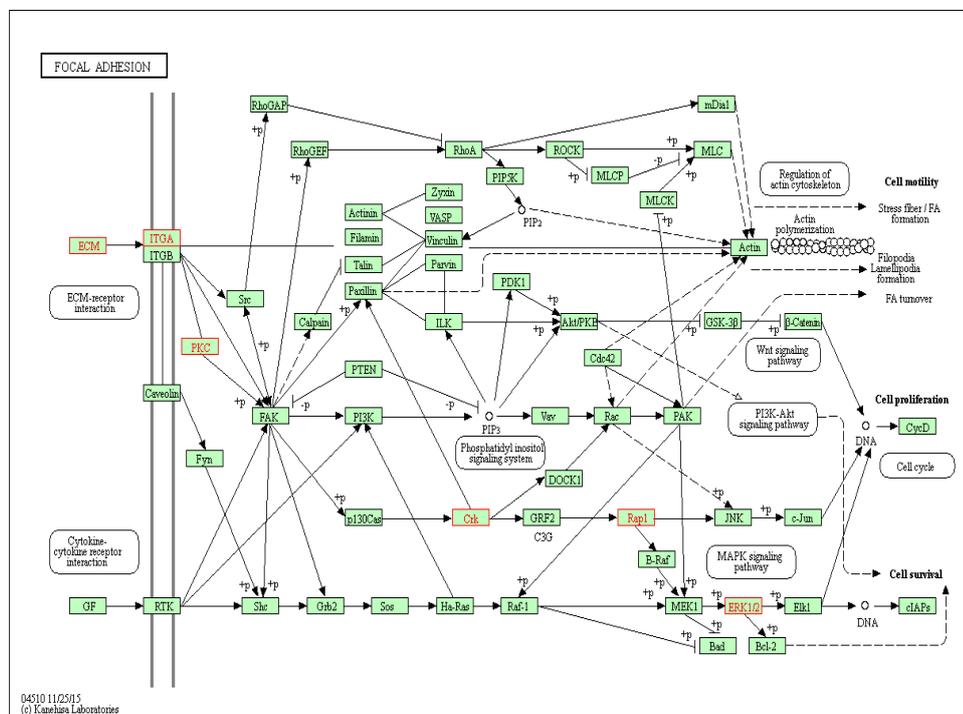
Figure 5.6 Differentially expressed genes within the focal adhesion pathway.
Red colour indicates some of the significantly enriched genes from this dataset. Green colour indicates other genes involved in this pathway.
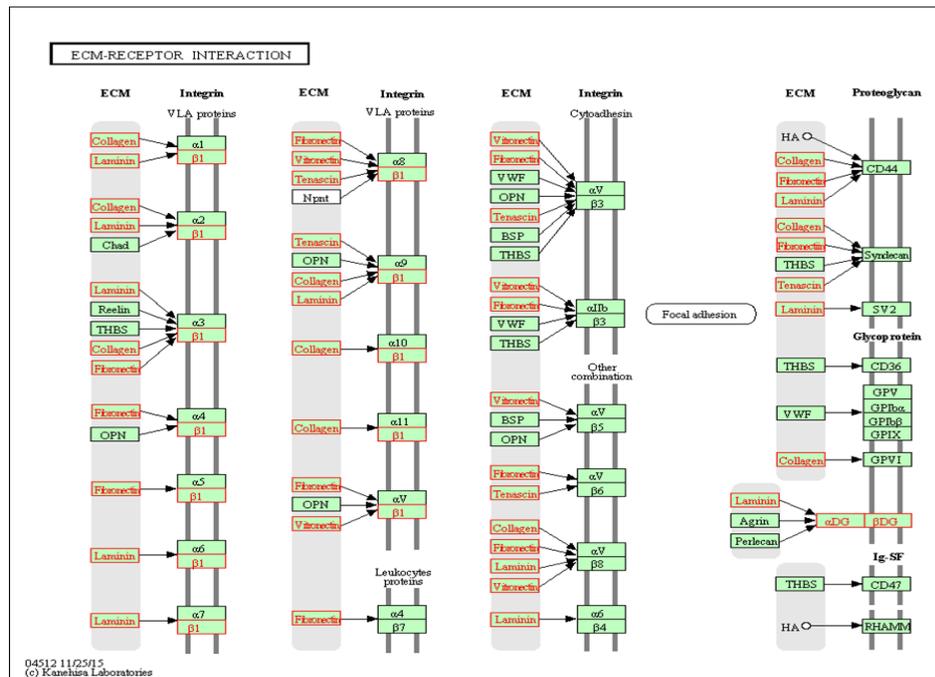
Figure 5.7 differentially expressed genes within the ECM. Red colour indicates significantly enriched genes from the dataset. Green colour indicates other genes involved in cell-ECM interactions.

Integrin signalling is dependent upon the non-receptor tyrosine kinase activities of the FAK (focal adhesion kinase) and SRC proteins as well as adaptor protein functions that initiate downstream signalling events. These signalling events culminate in re-organisation of the actin cytoskeleton, a pre-requisite for changes in cell shape, motility and invasion. Similar morphological alterations and modulations of gene expression are initiated by the binding of growth factors to their respective receptors emphasising the significant cross-talk between adhesion and growth factor-mediated signalling. The adhesion of retrograde shed endometrial cells to the ECM is one of the vital stages in the implantation of ectopic endometrial cells. The data presented here suggest some of the adhesion and ECM proteins that may be involved in the implantation process and establishment of endometriotic lesions. Studies have suggested that FAK signalling plays an important role in mediating increased cell migration in endometriosis induced by ovarian steroid hormones (Mu and Ma, 2015). Despite such reports, the mechanisms involved in ectopic cell attachment to the peritoneum and other ectopic areas remains poorly defined and further studies are

160

required to establish the molecular determinants involved and whether these proteins would be good markers of disease development and progression.

Functional analysis also revealed genes involved in cancer and neoplastic invasiveness. Endometriosis is a benign condition which shares some features with cancer such as local and distant invasion, attachment and damage to affected tissues. Women with the disease are known to be at an increased risk of developing epithelial ovarian cancer, although the molecular mechanisms involved are not clear (Worley et al., 2013). Further work should be carried out to ascertain the relevance of these genes in the pathogenesis of endometriosis and their potential role as therapeutic targets.

5.3 Conclusions

For the protein-based profiling there was a significant difference in coverage between the optimisation and main experiments. This was attributed to considerable loss of protein during the processing steps. A significant number of proteins identified comprised of cytokeratins, serum proteins and muscle proteins. This may be attributed to differential contamination due to sampling and may also be due to poor immunodepletion. However, protein labelling has some advantages over peptide labelling. Separation methods based on molecular weight e.g. SDS-PAGE can be applied to intact proteins as a fractionation step and differentially labelled samples can be mixed earlier in the workflow reducing technical variation. TMT labelling of lysines at the protein level would also compromise tryptic cleavage, and although sequential digestion with Glu-C and trypsin would alleviate this issue to generate shorter peptides more amenable to MS analysis, the number of labelled peptides identified was consequently reduced, thus compromising quantitative coverage.

Due to time constraints and the lack of material remaining, this experiment could not be repeated. It is therefore difficult to conclude whether this protein-based profiling approach is complementary to peptide-based approaches or offers improved

proteomic coverage. Because of the protein loss, possible contamination and poor labelling efficiency in the protein labelling experiment, selection of meaningful candidates for verification was not possible.

There was better coverage afforded by the peptide-based profiling approach with 1,431 proteins quantified across the six clinical groups. Blood protein contamination was still a serious issue and will exacerbate sample heterogeneity leading to a high false discovery rate. The present data suggested dysregulation of progesterone signalling that may contribute to the progesterone resistance associated with endometriosis. PGR and PGRMC1 were found to be dysregulated, as was the downstream target PAEP. The role of these changes in promoting progesterone resistance in endometriosis warrant further investigation. Differentially expressed proteins involved in regulating translation and cell-matrix interactions were also enriched and their roles in promoting endometriosis also warrant further investigation. Proteins of interest identified from this data that warrant verification include; TNC, CPM, TPM2, LUM, PAEP, PGR and PGRMC1.

CHAPTER 6: CANDIDATE VERIFICATION IN SERUM

6.1 Introduction

Proteins of interest identified from the discovery profiling were selected for verification as potential biomarkers of endometriosis. The present work also aimed to validate several candidate markers reported in the literature. These literature candidates were sICAM1, MCP1, MIF, IL1R2, VEGF and FST, whilst LUM, TPM2, CPM, PAEP and TNC were selected from the discovery profiling work. One of the original aims was to verify the expression of candidates in the individual tissue lysates by western blotting. However, this was not possible as there was insufficient sample remaining for some cases and controls. Additionally, verification by western blotting in over 50 samples would have been laborious and only semi-quantitative at best. Also, since the overall aim of the project was to develop a non-invasive diagnostic test, it was decided that the candidates should be further tested in serum samples collected from the same and additional patients using commercial ELISA kits. Assays were first tested for reproducibility and sensitivity using a test pool of all samples. Candidate serum measurements were correlated with measurements of progesterone, oestradiol, CRP and CA125, clinico-pathological features and epidemiological data, particularly phase of the menstrual cycle. The data was incorporated into multi-marker models to assess if the candidates could complement one another and improve classification performance. A commercial multiplex platform was also employed to potentially identify additional candidates, this using a subset of the serum samples.

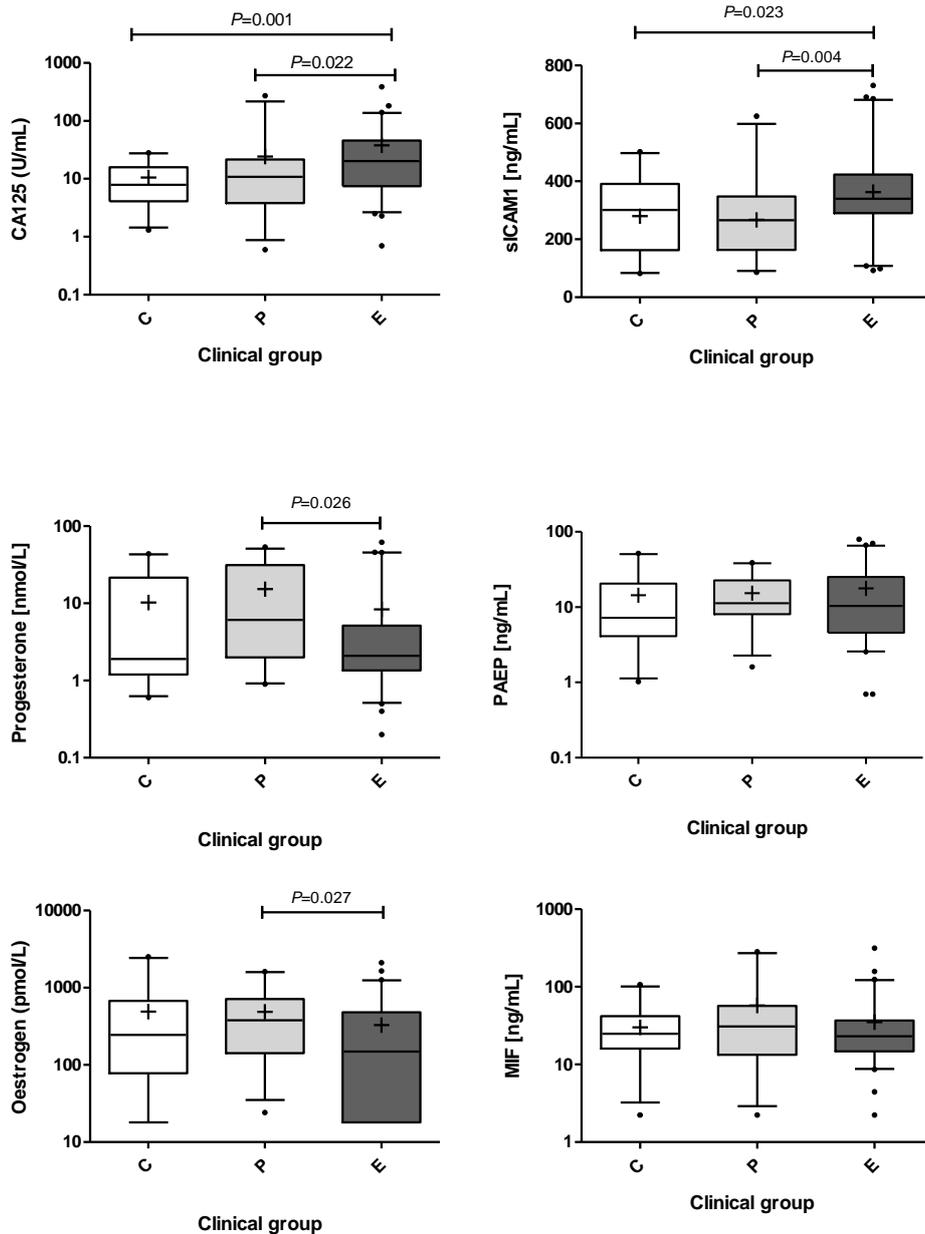## 6.2 Single marker analysis

In total, 13 candidates (sICAM1, MCP1, MIF, IL1R2, VEGF, FST, PAEP, LUM, TPM2, CPM, TNC, CA125 and CRP) were measured in a set of 109 discovery set serum samples (Control=23, Pain=24 and Endometriosis=62) from the same patients from which tissue samples had been analysed and additional case controls that had been excluded from the tissue analysis. Oestrogen and progesterone measurements were also analysed. The differences in concentrations of single markers was first

assessed to determine any differences between the clinical groups (Table 6.1). As previously reported, there was a significant difference in the levels of CA125 between healthy controls and endometriosis (P=0.001) and pain and endometriosis (P=0.022) (Figure 6.1). The concentration of sICAM1 in serum also displayed a significant difference between endometriosis and healthy controls (P=0.022) and between endometriosis and pain controls (P=0.004) with no difference between the two control groups (P=0.541). Oestrogen and progesterone levels were also significantly changing between the pain and endometriosis groups (P=0.025 and P=0.025, respectively), but not between the endometriosis and healthy controls. VEGF, IL1R2, MIF, MCP, CPM, TNC, LUM, PAEP, CRP and FST failed to show any significant differences between the clinical groups, whilst the TPM2 assay failed to give a signal on the standard curve for even undiluted serum.

Table 6.1 Median concentrations and ranges of individual marker candidates in serum and *P* values for comparison of different clinical groups.

						E v C	E v P	C v P
Candidate biomarker	Units	Control	Pain	Endometriosis	Phase	*P* value	*P* value	*P* value
sICAM1	ng/mL	301.4 (162.2-391.3)	265.7 (162.8-347.9)	339.7 (290.3-423.5)	All	**0.023**	**0.004**	0.541
IL1R2	ng/mL	12.71 (10.51-14)	11.37 (9.66-16.88)	12.69 (10.14-15.06)	All	0.738	0.598	0.477
MCP1	ng/mL	0.32 (0.22-0.41)	0.25 (0.17-0.43)	0.26 (0.21-0.37)	All	0.238	0.608	0.481
MIF	ng/mL	24.81 (15.99-41.88)	30.87 (13.31-56.62)	23.09 (14.74-36.81)	All	0.654	0.362	0.442
VEGF	ng/mL	301.4 (162.2-391.3)	265.7 (162.8-347.9)	339.7 (290-423.5)	All	0.105	0.968	0.279
FST	ng/mL	0.82 (0.59-1.25)	0.73 (0.56-1.24)	0.67 (0.5-0.9)	All	0.185	0.295	0.919
PAEP	ng/mL	7.22 (4.10-20.61)	11.23 (8.04-22.73)	10.35 (4.57-25.26)	All	0.367	0.631	0.196
LUM	ng/mL	47.05 (31.64-62.06)	39.47 (13.53-63.7)	46.1 (34.86-61.1)	All	0.862	0.498	0.461
CA125	U/mL	7.9 (4.1-15.9)	10.85 (3.85-21.7)	20.4 (7.45-45.95)	All	**0.001**	**0.022**	0.5301
Oestrogen	pmol/L	245 (78-673)	379 (142-716)	148.5 (18-478.8)	All	0.238	**0.027**	0.424
Progesterone	nmol/L	1.9 (1.2-21.55)	6.1 (2-31.4)	2.1 (1.35-5.15)	All	0.966	**0.026**	0.115
CRP	mg/L	0.75 (0.6-1.88)	0.6 (0.6-2.27)	0.7 (0.6-2.7)	All	0.914	0.544	0.665
CPM	ng/mL	4.31 (0.92-25.30)	3.83 (1.16-9.69)	5.0 (2.0-121.0)	All	0.629	0.968	0.663
TNC	ng/mL	73.24 (26.54-88.33)	40.31 (21.83-63.35)	45.24 (34.86-65.13)	All	0.369	0.359	0.279

The student t-test or Mann-Whitney test was used. *P* values <0.05 were considered significant and are shown in yellow.

Figure 6.1 Graphical representation of single marker analysis in the three clinical groups. (C= healthy controls (n=24), P= Pain (n=23) and E=endometriosis (n=62)). All graphs were generated using GraphPad Prism Version 5.01. A student t-test was applied where data was normally distributed and a Mann Whitney test where data was not normally distributed. A *P* value of <0.05 was considered significant.

The data was also analysed to determine whether these candidate markers differed across the menstrual cycle. CA125 was significantly differentially expressed between both healthy and pain control groups and endometriosis in the secretory phase (P=0.006 and 0.03) but not the proliferative phase (Figure 6.3).
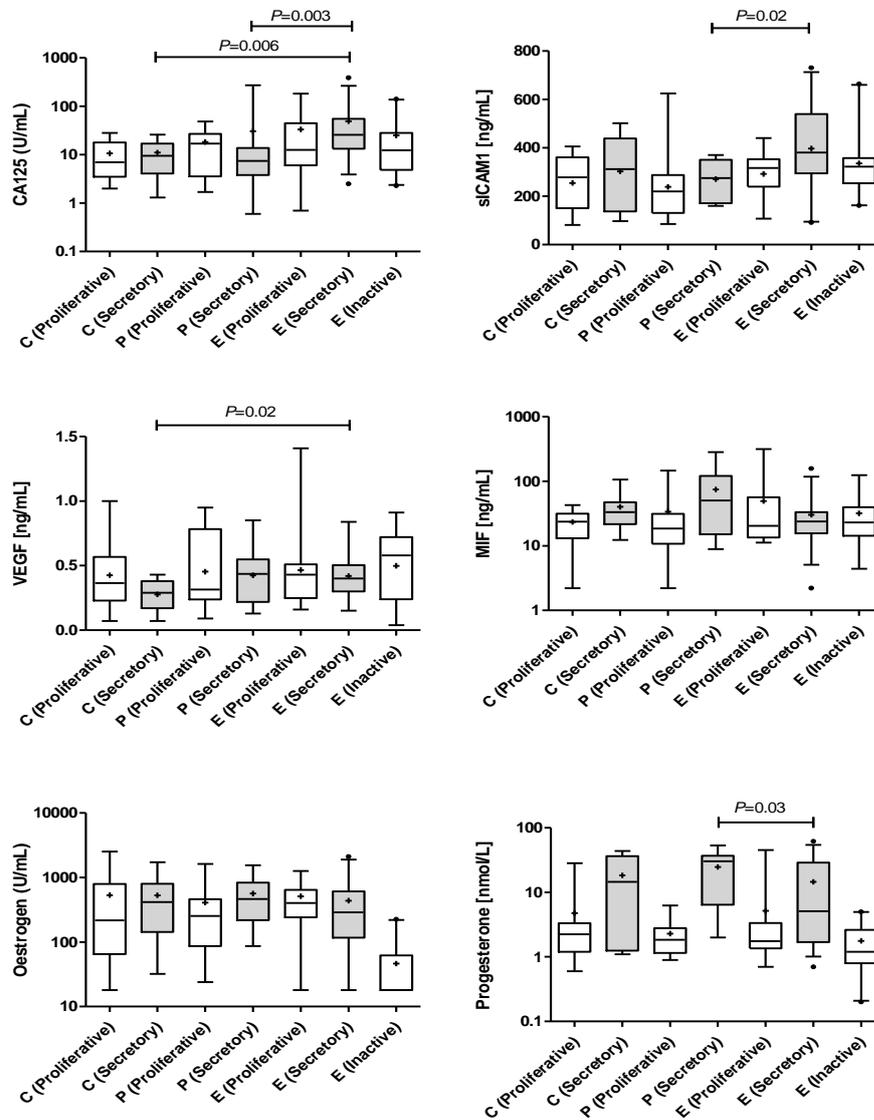
Figure 6.2 Graphical representation of single marker analysis by cycle phase. Graphs were generated using GraphPad Prism Version 5.0.1. A student t-test was applied where data was normally distributed and Mann Whitney where data was not normally distributed. A P value of <0.05 was considered statistically significant.

166

sICAM1 was also significantly elevated in the secretory phase endometriosis *versus* pain groups (*P*=0.023), whilst VEGF was elevated in secretory phase endometriosis *versus* healthy controls (*P*=0.023). As expected progesterone was significantly higher in the secretory phase than proliferative phase, but showed a non-significant down-regulation in the endometriosis group. No differences were found for any of the other candidates suggesting that they would not make useful diagnostic markers.

6.3 Receiver operator characteristic (ROC) curve analysis

ROC curve analysis was performed to determine the diagnostic performance of each biomarker separately. The area under the curves (AUCs) and sensitivities at 90% specificity were generated using GraphPad Prism software (Tables 6.2 and 6.3).

Table 6.2 Area under the ROC curve (AUC) analysis for single candidates.

Candidate	Phase	E v C (AUC)	*P* value	E v P (AUC)	*P* value
sICAM1	All	0.628	0.080	0.703	0.004
IL1R2	All	0.528	0.717	0.501	0.988
MCP1	All	0.593	0.230	0.537	0.664
MIF	All	0.566	0.397	0.677	0.039
VEGF	All	0.619	0.105	0.503	0.965
FST	All	0.598	0.183	0.577	0.293
PAEP	All	0.566	0.365	0.534	0.628
LUM	All	0.513	0.859	0.594	0.238
CA125	All	0.724	0.002	0.660	0.064
Oestrogen	All	0.586	0.239	0.657	0.026
Progesterone	All	0.504	0.962	0.659	0.025
CRP	All	0.508	0.914	0.541	0.566
CPM	All	0.537	0.625	0.504	0.963
TNC	All	0.568	0.366	0.574	0.356

AUCs with significant *P* values (<0.05) are highlighted in yellow.

Table 6.3 Sensitivities of individual candidates at 90% specificity.

Candidate	E v C	E v P
sICAM1	18%	34%
IL1R2	21%	10%
MCP1	5%	12%
MIF	15%	5%
VEGF	10%	3%
FST	23%	10%
PAEP	10%	27%
LUM	11%	13%
CA125	47%	31%
Oestrogen	26%	34%
Progesterone	8%	18%
CRP	11%	12%
CPM	12%	8%
TNC	12%	5%

The highest AUC was observed for CA125 (AUC=0.724; *P*=0.002) for discriminating between endometriosis and controls (Table 6.2). At a cut-off of >23 IU/mL the sensitivity of CA125 was 47% at 90% specificity for discriminating between endometriosis and healthy controls and 31% sensitivity at a cut off of >38 IU/mL for discriminating between endometriosis and pain controls (Table 6.3). CA125 has been investigated extensively as a potential peripheral biomarker of endometriosis, although is reported to lack diagnostic accuracy as a single biomarker. The data presented here supports this notion, with studies rarely reporting sensitivities above 50% at high specificity (80-95%). The diagnostic performance of CA125 in endometriosis is confounded by the fact that endometriosis has varying degrees of chronicity and is mostly elevated in advanced disease (see Chapter 1). CA125 is also not specific for endometriosis and does not efficiently differentiate endometriosis from other diseases e.g. fibroids, PID and ovarian cancer.

In discriminating between endometriosis and controls, the sensitivity of sICAM1 was 18% (>450 ng/mL), and in discriminating between endometriosis and pain was 34% (>373 ng/mL) at 90% specificity. sICAM1 is secreted from the endometrium and endometrioitic implants. Studies have been conflicting about the usefulness of

sICAM1 in diagnosing endometriosis. The data presented here does not support its use as a single marker for the diagnosis of endometriosis. MIF was able to discriminate between endometriosis and pain (AUC=0.677; P=0.039). MIF is a pro-inflammatory cytokine involved in T cell activation, cell growth, apoptosis and angiogenesis. It has been reported to have higher expression in advanced stage endometriosis (Morin et al 2005). However inflammatory cytokines are unlikely to make useful markers for differential diagnosis of endometriosis and non-endometriotic pathologies presenting with similar symptoms e.g. pelvic pain and infertility where inflammatory pathways are activated.

Combining markers could afford higher diagnostic accuracy by complementing one another to overcome the variable aetiology and disease manifestations of endometriosis. With this in mind, the diagnostic potential of combinations of the candidate markers was tested. Due to the relatively small sample size of this study, only combinations of two or three markers per comparison could be assessed to avoid overfitting. Logistic regression was used to combine marker candidates and their diagnostic performances tested using R software. The AUC values of these models were also cross-validated using R software. Table 6.4 and 6.5 displays the performance and cross-validation of the best models for discriminating between endometriosis and both control groups, reporting sensitivity at >90% and >80% specificity.

A model combining CA125 and sICAM1 was able to discriminate between endometriosis and controls with 55% sensitivity at 91% specificity (AUC=0.781); an improvement of 8% over using CA125 alone. By reducing the specificity to 83%, the CA125 sICAM1 model gave a sensitivity of 66% with no benefit of adding a third candidate. In discriminating between endometriosis and pain, the best combination was CA125 LUM TNC (AUC=0.739; sensitivity=63%, specificity=83%), although increasing the specificity to 92% reduced the sensitivity to 47%. At 63% sensitivity and 81% specificity, the best model that discriminated between endometriosis and both control groups (healthy and pain groups) was CA125 ICAM FST LUM.

Table 6.4 Combined marker model performances at different specificities

E vs C	AUC	Sensitivity	Specificity
Marker combination at ≥90% specificity			
CA125 ICAM	0.781	55	91
CA125 ICAM FST	0.796	53	91
Marker combination at ≥80% specificity			
CA125 ICAM	0.781	66	83
CA125 ICAM FST	0.796	66	83
E vs P	**AUC**	**Sensitivity**	**Specificity**
Marker combination at ≥90% specificity			
ICAM TNC	0.718	31	92
ICAM LUM TNC	0.739	47	92
Marker combination at ≥80% specificity			
ICAM TNC	0.718	55	83
ICAM LUM TNC	0.739	63	83
Endometriosis vs All Controls (C and P)	**AUC**	**Sensitivity**	**Specificity**
Marker combination at ≥90% specificity			
CA125 ICAM FST LUM	0.756	30	92
CA125 ICAM FST	0.750	30	92
CA125 ICAM	0.736	37	92
Marker combination at ≥80% specificity			
CA125 ICAM FST LUM	0.756	63	81
CA125 ICAM FST	0.750	55	83
CA125 ICAM	0.736	53	83

Cross-validation of the AUC values of the above models showed that there was some loss in performance in the models especially in discriminating between endometriosis and pain. There was 5% and 16% loss in sensitivity at 92% and 83% specificity respectively in the ICAM TNC model (AUC=0.66) and 10% and 19% loss in sensitivity of the ICAM LUM TNC model (AUC= 0.667) in discriminating between endometriosis and pain. There was also loss in performance in the CA125 ICAM FST LUM model (AUC= 0.691) for discriminating between endometriosis and both healthy and pain controls with 4% and 15% loss in sensitivity at 91% and 81% specificity. Performance of models CA125 ICAM (AUC=0.744) and CA125 ICAM FST (AUC=0.75) for discriminating between endometriosis and healthy controls did

170

not change much. However, there was a 5% decrease in sensitivity at 91 and 83% specificity.

Table 6.5. Cross validated results of combined marker models

| E vs C	AUC	Sensitivity	Specificity
Marker combination at ≥90% specificity			
CA125 ICAM	0.744	52	91
CA125 ICAM FST	0.75	52	91
Marker combination at ≥80% specificity			
CA125 ICAM	0.744	61	83
CA125 ICAM FST	0.75	60	83
E vs P	**AUC**	**Sensitivity**	**Specificity**
Marker combination at ≥90% specificity			
ICAM TNC	0.66	26	92
ICAM LUM TNC	0.667	37	92
Marker combination at ≥80% specificity			
ICAM TNC	0.66	39	83
ICAM LUM TNC	0.667	44	83
Endometriosis vs All Controls (C and P)	**AUC**	**Sensitivity**	**Specificity**
Marker combination at ≥90% specificity			
CA125 ICAM FST LUM	0.691	26	91
CA125 ICAM FST	0.703	21	91
Marker combination at ≥80% specificity			
CA125 ICAM FST LUM	0.691	48	81
CA125 ICAM FST	0.75	47	85

6.4 Proseek analysis

30 serum samples (15 endometriosis and 15 pain group) were randomly selected and analysed on the Proseek Oncology II platform (Olink Bioscience) which uses a sensitive qPCR proximity extension assay. The protein analysis was reported as normalised protein expression values (NPX). All assay characteristics including detection limits and measurements of assay performance and validations are available from the manufacturer's web page (http://www.olink.com/proseekmultiplex/ oncology). Table 6.5 displays the univariate analysis of the 17 analytes (of 92) that displayed significant difference between the endometriosis and pain groups. 16 of the

candidates were down-regulated in the endometriosis compared to pain group, with only MUC16/CA125 showing an elevation. This perhaps suggested some bias in the sample set. Significant changes were only apparent in the secretory phase. Data for the most significantly changing proteins are shown in Figure 6.3.

Table 6.6. List of significant proteins from the Proseek Oncology II analysis.

Marker Candidate	Endo vs. Pain (all stages)	ES vs. PS	EP vs. PP	Regulation (E vs P)	Expression	Biological function
DKN1A	0.044	0.014	ns	down		
GPNMB	0.025	0.021	ns	down		
CA9	0.030	0.027	ns	down		
CD27	0.004	0.012	ns	down	T-cell specific	apoptosis, immune function
ESM1	0.004	0.005	ns	down	endothelial cells	angiogenesis
LYN	0.006	0.011	ns	down	widely expressed	immune response, signalling
MUC16	0.047			up		
WFDC2	0.010	0.028	ns	down		
ANXA1	0.016			down		
MSLN	0.042	0.038	ns	down		
TFP1-2	0.042	ns	ns	down		
S1004	0.030	0.050	ns	down		
TGFa	0.008	0.010	ns	down	keratinocytes, tumour cell lines	EGFR ligand, proliferation
FADD	0.034	0.040	ns	down		
METAP2	0.047	0.008	ns	down		
S100A11	0.031	0.050	ns	down		
EPHA2	0.032	0.029	ns	down		

Yellow shading shows the most significantly altered proteins.

172
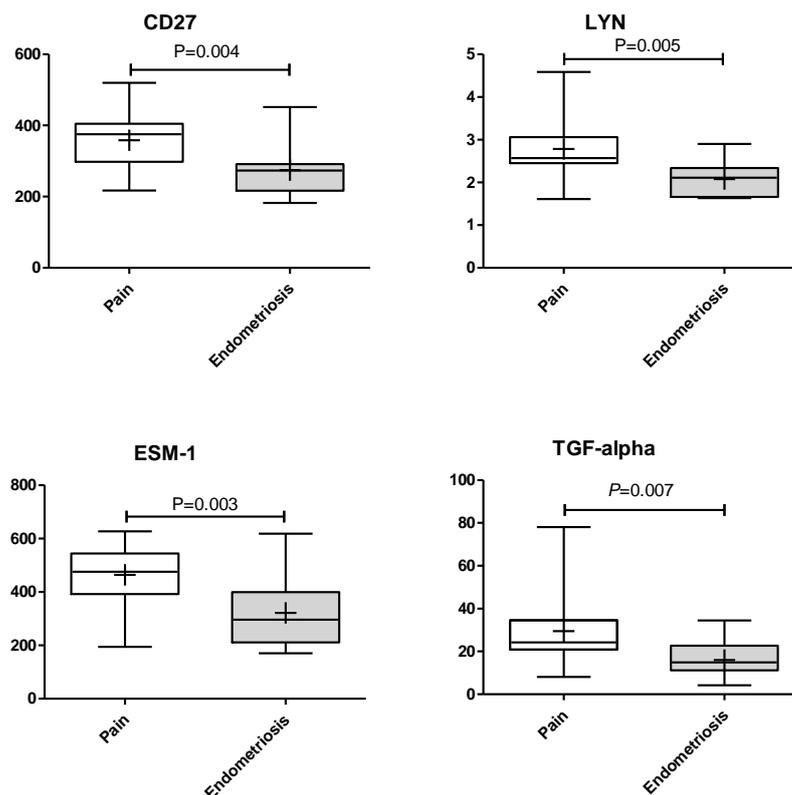
Figure 6.3 Graphical representation of most significant candidates from the Proseek®· Oncology II panel analysis. A student t-test was applied where data was normally distributed and a Mann Whitney test where data was not normally distributed. A *P* value of <0.05 was considered to be significant.

6.5 Conclusions

Selected markers from the discovery profiling and those reported in the literature were verified using ELISA in serum samples collected from all study participants. ROC curve analyses were performed to assess their diagnostic performance either as single markers or as combinations of markers for discriminating between endometriosis and controls. From this data, CA125 displayed the highest performance as a single marker, but lacked sensitivity to be considered for clinical use. The best marker combination in discriminating between endometriosis and healthy controls was CA125, ICAM at 61% sensitivity and 83% specificity. Adding FST to this combination did not improve

173

sensitivity. Combination of ICAM, LUM, and TNC gave the best performance in discriminating endometriosis and pain controls at 63% sensitivity and 83% specificity. CA125, ICAM, FST, LUM displayed the best performance in discriminating endometriosis and both pain and healthy controls. However, there was a loss in sensitivity of these models after cross-validation of these models. Whilst the use of a panel of markers improved on the sensitivity, this may still be too low for clinical use. Future work should involve validation of these models in independent samples with a more detailed analysis by cycle phase. Several candidates from the Proseek analysis also warrant verification. Notably however, none of the markers would be significant if adjusted for multiple testing.

CHAPTER 7: DISCUSSION AND FUTURE WORK

7.1 Discussion

Diagnosis of endometriosis using a non-invasive method is a major priority in endometriosis research (Rogers et al., 2013). The current gold standard is surgery together with histological confirmation. The symptoms of endometriosis are neither sensitive nor specific for the condition, therefore diagnosis based on symptoms alone is unreliable. Thus, there is a 5-11 year delay in diagnosis from the onset of symptoms to definitive diagnosis. Laparoscopic diagnosis also carries the potential risks of bladder, bowel and major blood vessel damage, infection and adhesion formation. The procedure is also expensive and adds significant economic burden to the health service as well as to the patient. A clinically reliable test would have significant impact on reducing health care and individual costs by reducing the latency period of time to diagnosis. Early diagnosis could also mean early treatment and improved outcomes.

The main aim of this study was to identify potential biomarkers for the non-invasive diagnosis of endometriosis by proteomic profiling of a set of well-characterised tissue samples obtained from women with endometriosis and relevant control groups without endometriosis. Patient and control selection criteria and conditions for sampling are of particular importance in biomarker studies and account for the majority of biological and technical variability. Endometriosis is an oestrogen-dependent disease therefore only pre-menopausal women were used in this study. Selection of appropriate controls is of critical importance. On one hand it is important to ascertain molecular diffferences between women with endometriosis and those without, but it is also important to establish differences in women with endometriosis and those presenting with similar symptoms, especially pelvic pain and/or infertility. For this group of patients, a non-invasive test would be useful to rule out endometriosis. To address this, two groups of control patients were therefore selected for this study; a healthy control group and a pain control group. The healthy control group were women with no known disease at laparoscopy, while the pain group

comprised of women with chronic PID and chronic pelvic pain of unknown cause who would benefit from a non-invasive test to rule out endometriosis.

Endometriosis is phenotypically heterogeneous with each form having a different presentation and possibly aetiology. This heterogeneity results in sample variability due to differences in tissue composition of the samples, especially ectopic tissue, as was apparent from the profiling in this study. The changing cellular composition of the endometrium, especially in response to a changing hormonal environment is a major challenge that has to be considered while designing endometriosis biomarker studies. An ideal biomarker for endometriosis would maintain a high sensitivity and specificity regardless of when during the menstrual cycle a sample was obtained. However, influence of the menstrual cycle poses a significant challenge in endometrial-based biomarker development especially for proteins whose expression are regulated by oestrogen or progesterone. Samples were matched according to the stage of the menstrual cycle using a triple approach (documented chronologically, confirmed histologically and by serum measurement of sex hormones), that would allow determination of cycle stage-specific changes and isolating those due to disease. This was important in maintaining biological homogeneity across the clinical groups. However, this meant that sample numbers in some clinical groups were compromised e.g. healthy controls in the secretory phase (n=4) and no representation of pain controls in the proliferative phase.

Oestrogen-dependent diseases such as fibroids, endometrial polyps and cancer that have been observed to exist with endometriosis in affected women may also confound the interpretation of molecular profiles derived from endometrial tissue. These conditions pose a challenge for biomarker discovery, especially from endometrial tissue, by making it difficult to delineate a biomarker that is unique to endometriosis. To address this issue, patients who were diagnosed with benign or malignant conditions affecting the endometrium were excluded from the study. Exposure to medications used to treat endometriosis causes down-regulation of the endometrium, therefore these patients were also excluded from the study, thus reducing the sample numbers in the endometriosis groups.

Proteomics is increasingly being applied to the human endometrium. Despite the challenges that come with analysis of the proteome, advances in MS instrumentation has enhanced it as a powerful tool in biomarker discovery research. Proteomics studies on endometriosis have mostly focused on identifying differences between eutopic endometrium in women with and without endometriosis, mostly for the purpose of gaining a better understanding of the pathophysiology of the disease. However, efforts have been made to search for biomarkers for non-invasive diagnosis of the disease, but despite this, no biomarkers have yet been validated for clinical use.

This study was divided into a discovery phase using untargeted proteomic strategies to identify candidate biomarkers from endometrial and endometriosis tissues collected in the course of this study, and a candidate selection and verification phase using serum samples from the same women. The aim was to identify secreted proteins that could be validated as potential markers for non-invasive differential diagnosis. In the discovery phase, two main proteomic strategies were employed namely 2D-DIGE and protein and peptide TMT labelling coupled to multi-dimensional fractionation and LC-MS/MS. These profiling methods provided the largest protein coverage of all techniques employed to date on a phenotypically well-characterised cohort.

Due to the heterogeneous nature of endometrial tissues there were some challenges experienced in proteomic profiling of these tissue samples. Presence of high abundance cellular structural proteins was observed, possibly masking proteins of low abundance. Blood contamination was a challenge as evidenced by the identification of red blood cell proteins from the profiling. This contamination may have arisen as a result of contamination during sampling or due to vascularisation of the tissue, especially the eutopic tissues samples. Additionally, there was also presence of high abundance serum proteins that would lower proteomic coverage. To address this issue, an immunodepletion step for the tissue lysates was incorporated. Whilst this undoubtedly improved the protein load and coverage, several proteins targeted during this step were identified by MS as differentially expressed, suggesting incomplete depletion. Due to protein loss, there was poor coverage in the protein labelling experiment that was possibly confounded by contamination, incomplete

immunodepletion and sub-stoichiometric TMT labelling. This experiment could not be repeated and therefore it was difficult to conclude whether this approach would offer improved and complementary coverage of the endometrial proteome. These challenges made selection of potential candidates from this profiling somewhat redundant, although some candidates common to the 2D-DIGE profiling were observed. The peptide profiling method was much more successful, providing quantitative coverage of over 1,400 proteins. Whilst the issues of sample heterogeneity and blood contamination were still an issue and would raise the number of false discoveries, the strategy did provide a list of potential biomarker candidates that warranted verification. The profiling also provided some insight into the biology of endometriosis most notably progesterone resistance that is associated with endometriosis.

Proteins identified from the discovery profiling that had potential as biomarkers were PAEP, LUM, CPM, TPM2 and TNC, and were selected for testing in serum using ELISA. These proteins have been previously identified in some studies ((Hood et al., 2015, Tan et al., 2008, Harrington et al., 1999, Meola et al., 2009) whose aim was to gain molecular understanding on the human endometrium during the menstrual cycle and to understand aetiology, however their usefulness as diagnostic biomarkers has yet to be assessed. Promising markers identified from the literature were CA125, MIF, VEGF, sICAM1, IL1R2, MCP1 and FST and these were also tested alongside the other candidates in the discovery set serum. Candidate measurements were also incorporated into multi-marker models to assess whether this would improve diagnostic performance. The most studied serum marker for endometriosis is CA125. However, its widespread presence in tissues derived from the coloemic epithelium impairs its usefulness as a stand-alone diagnostic test. Of the marker candidates tested herein, CA125 was the best performing single marker, however its sensitivity in discriminating endometriosis from healthy and pain controls was rather modest. The other candidate markers were essentially ruled out as single markers, failing to provide accurate discrimination of the endometriosis and control groups. This failure could also be attributed to lack of secretion of these markers from tissue to serum. Another

reason could be that the pooling of tissue samples presents an average level of expression across the groups to mask the effects of individual outliers and increase the FDR

Combining several markers improved the sensitivity, with a logistic regression model combining CA125, sICAM1, FST and LUM giving the highest performance with a sensitivity of 63% and specificity of 81%. This indicates that adding these other markers complements CA125 in detecting cases. To avoid over-fitting and reporting overly optimistic diagnostic performance of this model, this model was cross-validated resulting in decrease in sensitivity; 26% and 48% at 91% and 81% specificity. It is difficult to conclude whether this model warrants future independent validation in a further prospective study to assess if it will be suitable for translation into clinical practice. Biomarker data could also be tested against outcome measures to establish if any of the candidate markers have prognostic potential.

7.2 Future work

Challenges arising due to the heterogeneous nature of endometrial tissue and its sampling should be addressed in any future discovery work. The cell types within a single tissue sample are often highly variable e.g. ectopic endometrial lesions may contain few endometrial cells alongside other structural components. One approach to address this problem would be the microdissection of tissue samples, although this would require considerable input from a trained pathologist. Laser capture microdissection could be used to obtain purer and relevant endometrial cell populations prior to proteomic analyses. Future research could also look into menstrual tissue and material obtained from the peritoneal cavity at laparoscopy performed at the time of menstruation. Differences in retrograde-shed menstrual material between women with and without endometriosis might reveal unique proteins with diagnostic potential. One could also use the serum samples for profiling to potentially identify blood-borne biomarkers directly. However, this is also likely to be

challenging. The dynamic range of protein expression in serum may be as high as ten orders of magnitude, with 99% of the protein content represented by the 22 most abundant proteins. Immunodepletion would be applied as used herein, although deep coverage (sub-ng/mL) would still be difficult to achieve and multi-dimensional fractionation would need to be applied. Since MS instrument time is a limiting factor, the running of multiple fractions from single samples would not be possible, even with multiplex mass tagging. Thus, a pooling approach would need to be employed with the caveat that individual protein expression information is lost, outliers cannot be identified and the false discovery rate is increased. In turn, this would necessitate more verification testing.

Future work should involve the validation of the derived multi-marker model (CA125, sICAM1, FST and LUM) in a larger independent cohort of sera collected prospectively from women presenting with pelvic pain and no laparoscopic evidence of endometriosis and those diagnosed with endometriosis. The profiling work presented herein also identified numerous other proteins of interest that may have biomarker potential and hence warrant further verification. In future work, these would be tested as serum markers alone and in combination. ELISA is the ideal method for quantitative analysis of proteins in serum and is the gold standard in clinical diagnostics. However, a proportion of the candidates lack reliable ELISA tests for verification. Indeed, the availability of such reagents was applied here as a filtering criteria when selecting putative markers for testing. Due to the costs involved and time required in developing ELISAs, the majority of candidates identified from discovery profiling are left untested and potentially useful biomarkers may be missed. Priority should therefore be given to the development of low-cost, high-throughput, multiplex assays for protein quantification that do not rely on antibodies. Multiple reaction monitoring (MRM) assays based on MS are a favourable alternative to immunoassays. Such assays are highly-specific, cost-effective and can be more rapidly developed in comparison to ELISA assays. MRM also offers better multiplexing capabilities, allowing simultaneous quantification of numerous proteins within a single run. Future work would thus involve the development of MRM. This would first involve detection

of unique proteotypic peptides for each candidate (n≥3; ideally already observed in the discovery phase profiling) by high-accuracy mass monitoring of trypsinised crude or immunodepleted serum. Peptides would be subjected to optimised fragmentation on a triple quad instrument to confirm identification and predominant fragment ions chosen as transitions. Absolute quantification would be achieved by comparing parent ion intensities with spiked standards using ^{13}C isotopically-labelled synthetic peptides. Assays would be optimised for minimal sample processing and multiplexed where possible.

The work also identified putative tissue markers and some of these warrant verification by immunohistochemical staining. These may have applications as semi-invasive biomarkers for the staging of endometriosis and/or for its prognosis. One protein of particular interest was the progesterone receptor (PGR). Endometriosis is associated with progesterone resistance. PGR expression possibly suggests perturbed progesterone signalling may be at play. Further studies are required to establish the role of altered PGR (and PGRMC1) expression in endometriosis, whether this affects downstream progesterone-regulated genes such as PAEP, as suggested from the present data and how this might be involved in progesterone resistance. Future work should also involve a re-analysis of the profiling data, specifically to identify menstrual cycle stage-specific differences in the healthy controls (CS *versus* CP), and thus the effect of endometriosis and pain on these changes. This would add to the functional analysis and could potentially identify novel sex-steroid hormone-dependent protein changes. The functional enrichment analysis of the profiling data suggested differences in the focal adhesion, ECM and actin re-arrangement pathways in endometriosis. Future research using cell-based and animal models should explore the role of the altered proteins in promoting the proliferation, invasion and establishment of endometriotic lesion.

REFERENCES

1979. Classification of endometriosis. The American Fertility Society. *Fertil Steril,* 32**,** 633-4.

1997. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertil Steril,* 67**,** 817-21.

Aban, M., Ertunc, D., Tok, E. C., Tamer, L., Arslan, M. & Dilek, S. 2007. Modulating interaction of glutathione-S-transferase polymorphisms with smoking in endometriosis. *J Reprod Med,* 52**,** 715-21.

Abrao, M. S., Goncalves, M. O., Ajossa, S., Melis, G. B. & Guerriero, S. 2009. The sonographic diagnosis of deep endometriosis. *J Ultrasound Med,* 28**,** 408-9; author reply 409-10.

Abrao, M. S., Goncalves, M. O., Dias, J. A., Jr., Podgaec, S., Chamie, L. P. & Blasbalg, R. 2007. Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. *Hum Reprod,* 22**,** 3092-7.

Adamson, G. D. 1990. Diagnosis and clinical presentation of endometriosis. *Am J Obstet Gynecol,* 162**,** 568-9.

Adamson, G. D., Hurd, S. J., Pasta, D. J. & Rodriguez, B. D. 1993. Laparoscopic endometriosis treatment: is it better? *Fertil Steril,* 59**,** 35-44.

Adamson, G. D. & Pasta, D. J. 2010. Endometriosis fertility index: the new, validated endometriosis staging system. *Fertil Steril,* 94**,** 1609-15.

Aebersold, R. & Mann, M. 2003. Mass spectrometry-based proteomics. *Nature,* 422**,** 198-207.

Aghajanova, L., Velarde, M. C. & Giudice, L. C. 2010. Altered gene expression profiling in endometrium: evidence for progesterone resistance. *Semin Reprod Med,* 28**,** 51-8.

Akoum, A., Al-Akoum, M., Lemay, A., Maheux, R. & Leboeuf, M. 2008. Imbalance in the peritoneal levels of interleukin 1 and its decoy inhibitory receptor type II in endometriosis women with infertility and pelvic pain. *Fertility and Sterility,* 89**,** 1618-24.

Akoum, A., Jolicoeur, C., Kharfi, A. & Aube, M. 2001. Decreased expression of the decoy interleukin-1 receptor type II in human endometriosis. *Am J Pathol,* 158**,** 481-9.

Akoum, A., Lawson, C., Herrmann-Lavoie, C. & Maheux, R. 2007. Imbalance in the expression of the activating type I and the inhibitory type II interleukin 1 receptors in endometriosis. *Hum Reprod,* 22**,** 1464-73.

Al-Jefout, M., Dezarnaulds, G., Cooper, M., Tokushige, N., Luscombe, G. M., Markham, R. & Fraser, I. S. 2009. Diagnosis of endometriosis by detection of nerve fibres in an endometrial biopsy: a double blind study. *Hum Reprod,* 24**,** 3019-24.

Alban, A., David, S. O., Bjorkesten, L., Andersson, C., Sloge, E., Lewis, S. & Currie, I. 2003. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics,* 3**,** 36-44.

Allen, C., Hopewell, S. & Prentice, A. 2005. Non-steroidal anti-inflammatory drugs for pain in women with endometriosis. *Cochrane Database Syst Rev***,** Cd004753.

Allen, C., Hopewell, S., Prentice, A. & Gregory, D. 2009. Nonsteroidal anti-inflammatory drugs for pain in women with endometriosis. *Cochrane Database Syst Rev***,** Cd004753.

Anderson, N. L. & Anderson, N. G. 1998. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis,* 19**,** 1853-61.

Arumugam, K. & Templeton, A. A. 1990. Endometriosis and social class. *Med J Aust,* 153**,** 567.

Attar, E. & Bulun, S. E. 2006. Aromatase inhibitors: the next generation of therapeutics for endometriosis? *Fertil Steril,* 85**,** 1307-18.

Ballard, K., Lowton, K. & Wright, J. 2006. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. *Fertil Steril,* 86**,** 1296-301.

Ballard, K. D., Seaman, H. E., De Vries, C. S. & Wright, J. T. 2008. Can symptomatology help in the diagnosis of endometriosis? Findings from a national case-control study--Part 1. *BJOG,* 115**,** 1382-91.

Bantscheff, M., Boesche, M., Eberhard, D., Matthieson, T., Sweetman, G. & Kuster, B. 2008. Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Mol Cell Proteomics,* 7**,** 1702-13.

Barbati, A., Cosmi, E. V., Spaziani, R., Ventura, R. & Montanino, G. 1994. Serum and peritoneal fluid CA-125 levels in patients with endometriosis. *Fertil Steril,* 61**,** 438-42.

Barbieri, R. L., Niloff, J. M., Bast, R. C., Jr., Scaetzl, E., Kistner, R. W. & Knapp, R. C. 1986. Elevated serum concentrations of CA-125 in patients with advanced endometriosis. *Fertil Steril,* 45**,** 630-4.

Barrier, B. F. & Sharpe-Timms, K. L. 2002. Expression of soluble adhesion molecules in sera of women with stage III and IV endometriosis. *J Soc Gynecol Investig,* 9**,** 98-101.

Becker, C. M., Laufer, M. R., Stratton, P., Hummelshoj, L., Missmer, S. A., Zondervan, K. T., Adamson, G. D., Adamson, G. D., Allaire, C., Anchan, R., Becker, C. M., Bedaiwy, M. A., Buck Louis, G. M., Calhaz-Jorge, C., Chwalisz, K., D'hooghe, T. M., Fassbender, A., Faustmann, T., Fazleabas, A. T., Flores, I., Forman, A., Fraser, I., Giudice, L. C., Gotte, M., Gregersen, P., Guo, S. W., Harada, T., Hartwell, D., Horne, A. W., Hull, M. L., Hummelshoj, L., Ibrahim, M. G., Kiesel, L., Laufer, M. R., Machens, K., Mechsner, S., Missmer, S. A., Montgomery, G. W., Nap, A., Nyegaard, M., Osteen, K. G., Petta, C. A., Rahmioglu, N., Renner, S. P., Riedlinger, J., Roehrich, S., Rogers, P. A., Rombauts, L., Salumets, A., Saridogan, E., Seckin, T., Stratton, P., Sharpe-Timms, K. L., Tworoger, S., Vigano, P., Vincent, K., Vitonis, A. F., Wienhues-Thelen, U. H., Yeung Jr, P. P., Yong, P. & Zondervan, K. T. 2014. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: I. Surgical phenotype data collection in endometriosis research. *Fertility and Sterility,* 102**,** 1213-1222.

Bedaiwy, M. A. & Falcone, T. 2004. Laboratory testing for endometriosis. *Clin Chim Acta,* 340**,** 41-56.

Bedaiwy, M. A., Falcone, T., Sharma, R. K., Goldberg, J. M., Attaran, M., Nelson, D. R. & Agarwal, A. 2002. Prediction of endometriosis with serum and peritoneal fluid markers: a prospective controlled trial. *Hum Reprod,* 17**,** 426-31.

Beliard, A., Donnez, J., Nisolle, M. & Foidart, J. M. 1997. Localization of laminin, fibronectin, E-cadherin, and integrins in endometrium and endometriosis. *Fertil Steril,* 67**,** 266-72.

Bellehumeur, C., Collette, T., Maheux, R., Mailloux, J., Villeneuve, M. & Akoum, A. 2005. Increased soluble interleukin-1 receptor type II proteolysis in the endometrium of women with endometriosis. *Hum Reprod,* 20**,** 1177-84.

Benagiano, G. & Brosens, I. 1991. The history of endometriosis: identifying the disease. *Hum Reprod,* 6**,** 963-8.

Benagiano, G. & Brosens, I. 2011. Who identified endometriosis? *Fertil Steril,* 95**,** 13-6.

Bischoff, F. & Simpson, J. L. 2004. Genetics of endometriosis: heritability and candidate genes. *Best Pract Res Clin Obstet Gynaecol,* 18**,** 219-32.

Blumenkrantz, M. J., Gallagher, N., Bashore, R. A. & Tenckhoff, H. 1981. Retrograde menstruation in women undergoing chronic peritoneal dialysis. *Obstet Gynecol,* 57**,** 667-70.

Bokor, A., Kyama, C. M., Vercruysse, L., Fassbender, A., Gevaert, O., Vodolazkaia, A., De Moor, B., Fulop, V. & D'hooghe, T. 2009. Density of small diameter sensory nerve fibres in endometrium: a semi-invasive diagnostic test for minimal to mild endometriosis. *Human Reproduction,* 24**,** 3025-32.

Bourlev, V., Iljasova, N., Adamyan, L., Larsson, A. & Olovsson, M. 2010. Signs of reduced angiogenic activity after surgical removal of deeply infiltrating endometriosis. *Fertil Steril,* 94**,** 52-7.

Brezillon, S., Pietraszek, K., Maquart, F. X. & Wegrowski, Y. 2013. Lumican effects in the control of tumour progression and their links with metalloproteinases and integrins. *Febs j,* 280**,** 2369-81.

Brosens, I. & Steeno, O. 2000. A compass for understanding endometriosis. *Fertil Steril,* 73**,** 179-80; author reply 180-1.

Brosens, I. A. 1994. Is mild endometriosis a progressive disease? *Hum Reprod,* 9**,** 2209-11.

Brosens, I. A., Puttemans, P. J. & Deprest, J. 1994. The endoscopic localization of endometrial implants in the ovarian chocolate cyst. *Fertil Steril,* 61**,** 1034-8.

Bulun, S. E. 2009. Endometriosis. *N Engl J Med,* 360**,** 268-79.

Bulun, S. E., Fang, Z., Imir, G., Gurates, B., Tamura, M., Yilmaz, B., Langoi, D., Amin, S., Yang, S. & Deb, S. 2004. Aromatase and endometriosis. *Semin Reprod Med,* 22**,** 45-50.

Bulun, S. E., Mahendroo, M. S. & Simpson, E. R. 1993. Polymerase chain reaction amplification fails to detect aromatase cytochrome P450 transcripts in normal human endometrium or decidua. *J Clin Endocrinol Metab,* 76**,** 1458-63.

Bumpass, L. L., Rindfuss, R. R. & Janosik, R. B. 1978. Age and marital status at first birth and the pace of subsequent fertility. *Demography,* 15**,** 75-86.

Burney, R. O. & Giudice, L. C. 2012. Pathogenesis and Pathophysiology of Endometriosis. *Fertility and sterility,* 98**,** 10.1016/j.fertnstert.2012.06.029.

Burney, R. O., Talbi, S., Hamilton, A. E., Vo, K. C., Nyegaard, M., Nezhat, C. R., Lessey, B. A. & Giudice, L. C. 2007. Gene expression analysis of endometrium reveals progesterone resistance and candidate susceptibility genes in women with endometriosis. *Endocrinology,* 148**,** 3814-26.

Casado-Vela, J., Rodriguez-Suarez, E., Iloro, I., Ametzazurra, A., Alkorta, N., Garcia-Velasco, J. A., Matorras, R., Prieto, B., Gonzalez, S., Nagore, D., Simon, L. & Elortza, F. 2009. Comprehensive proteomic analysis of human endometrial fluid aspirate. *J Proteome Res,* 8**,** 4622-32.

Celik, O., Hascalik, S., Elter, K., Tagluk, M. E., Gurates, B. & Aydin, N. E. 2008. Combating endometriosis by blocking proteasome and nuclear factor-kappaB pathways. *Hum Reprod,* 23**,** 2458-65.

Chapron, C., Barakat, H., Fritel, X., Dubuisson, J. B., Breart, G. & Fauconnier, A. 2005. Presurgical diagnosis of posterior deep infiltrating endometriosis based on a standardized questionnaire. *Hum Reprod,* 20**,** 507-13.

Chapron, C., Chopin, N., Borghese, B., Foulot, H., Dousset, B., Vacher-Lavenu, M. C., Vieira, M., Hasan, W. & Bricou, A. 2006. Deeply infiltrating endometriosis: pathogenetic implications of the anatomical distribution. *Hum Reprod,* 21**,** 1839-45.

Chapron, C., Fauconnier, A., Goffinet, F., Breart, G. & Dubuisson, J. B. 2002. Laparoscopic surgery is not inherently dangerous for patients presenting with benign gynaecologic pathology. Results of a meta-analysis. *Hum Reprod,* 17**,** 1334-42.

Chapron, C., Fauconnier, A., Vieira, M., Barakat, H., Dousset, B., Pansini, V., Vacher-Lavenu, M. C. & Dubuisson, J. B. 2003. Anatomical distribution of deeply infiltrating endometriosis: surgical implications and proposition for a classification. *Hum Reprod,* 18**,** 157-61.

Chapron, C., Souza, C., De Ziegler, D., Lafay-Pillet, M. C., Ngo, C., Bijaoui, G., Goffinet, F. & Borghese, B. 2010. Smoking habits of 411 women with histologically proven endometriosis and 567 unaffected women. *Fertil Steril,* 94**,** 2353-5.

Chatman, D. L. 1976. Endometriosis in the black woman. *Am J Obstet Gynecol,* 125**,** 987-9.

Chen, J. I., Hannan, N. J., Mak, Y., Nicholls, P. K., Zhang, J., Rainczuk, A., Stanton, P. G., Robertson, D. M., Salamonsen, L. A. & Stephens, A. N. 2009. Proteomic characterization of midproliferative and midsecretory human endometrium. *J Proteome Res,* 8**,** 2032-44.

Clement, P. B. 2001. History of gynecological pathology. IX. Dr. John Albertson Sampson. 1921. *Int J Gynecol Pathol,* 20**,** 86-101.

Cobellis, L., Razzi, S., Fava, A., Severi, F. M., Igarashi, M. & Petraglia, F. 2004. A danazol-loaded intrauterine device decreases dysmenorrhea, pelvic pain, and dyspareunia associated with endometriosis. *Fertil Steril,* 82**,** 239-40.

Colette, S., Lousse, J. C., Defrere, S., Curaba, M., Heilier, J. F., Van Langendonckt, A., Mestdagt, M., Foidart, J. M., Loumaye, E. & Donnez, J. 2009. Absence of aromatase protein and mRNA expression in endometriosis. *Hum Reprod,* 24**,** 2133-41.

Coombs, K. M. 2011. Quantitative proteomics of complex mixtures. *Expert Rev Proteomics,* 8**,** 659-77.

Cramer, D. W. & Missmer, S. A. 2002. The epidemiology of endometriosis. *Ann N Y Acad Sci,* 955**,** 11-22; discussion 34-6, 396-406.

Cramer, D. W., Wilson, E., Stillman, R. J., Berger, M. J., Belisle, S., Schiff, I., Albrecht, B., Gibson, M., Stadel, B. V. & Schoenbaum, S. C. 1986. The relation of endometriosis to menstrual characteristics, smoking, and exercise. *JAMA,* 255**,** 1904-8.

D'hooghe, T. M., Bambra, C. S., Cornillie, F. J., Isahakia, M. & Koninckx, P. R. 1991. Prevalence and laparoscopic appearance of spontaneous endometriosis in the baboon (Papio anubis, Papio cynocephalus). *Biol Reprod,* 45**,** 411-6.

D'hooghe, T. M., Bambra, C. S., Raeymaekers, B. M. & Koninckx, P. R. 1996a. Development of spontaneous endometriosis in baboons. *Obstet Gynecol,* 88**,** 462-6.

D'hooghe, T. M., Bambra, C. S., Raeymaekers, B. M. & Koninckx, P. R. 1996b. Increased prevalence and recurrence of retrograde menstruation in baboons with spontaneous endometriosis. *Hum Reprod,* 11**,** 2022-5.

Dai, Y., Leng, J. H., Lang, J. H., Li, X. Y. & Zhang, J. J. 2012. Anatomical distribution of pelvic deep infiltrating endometriosis and its relationship with pain symptoms. *Chin Med J (Engl),* 125**,** 209-13.

Daniel, Y., Geva, E., Amit, A., Eshed-Englender, T., Baram, A., Fait, G. & Lessing, J. B. 2000. Do soluble cell adhesion molecules play a role in endometriosis? *Am J Reprod Immunol,* 43**,** 160-6.

Darai, E., Detchev, R., Hugol, D. & Quang, N. T. 2003. Serum and cyst fluid levels of interleukin (IL) -6, IL-8 and tumour necrosis factor-alpha in women with

endometriomas and benign and malignant cystic ovarian tumours. *Hum Reprod,* 18**,** 1681-5.

De Placido, G., Alviggi, C., Di Palma, G., Carravetta, C., Matarese, G., Landino, G. & Racioppi, L. 1998. Serum concentrations of soluble human leukocyte class I antigens and of the soluble intercellular adhesion molecule-1 in endometriosis: relationship with stage and non-pigmented peritoneal lesions. *Hum Reprod,* 13**,** 3206-10.

Diez, R., Herbstreith, M., Osorio, C. & Alzate, O. 2010. 2-D Fluorescence Difference Gel Electrophoresis (DIGE) in Neuroproteomics. *In:* Alzate, O. (ed.) *Neuroproteomics.* Boca Raton (FL).

Dmowski, W. P. 1984. Pitfalls in clinical, laparoscopic and histologic diagnosis of endometriosis. *Acta Obstet Gynecol Scand Suppl,* 123**,** 61-6.

Donnez, J., Nisolle, M., Casanas-Roux, F., Brion, P. & Da Costa Ferreira, N. 1996a. Stereometric evaluation of peritoneal endometriosis and endometriotic nodules of the rectovaginal septum. *Hum Reprod,* 11**,** 224-8.

Donnez, J., Nisolle, M., Gillet, N., Smets, M., Bassil, S. & Casanasroux, F. 1996b. Large ovarian endometriomas. *Human Reproduction,* 11**,** 641-646.

Donnez, J., Nisolle, M., Smoes, P., Gillet, N., Beguin, S. & Casanas-Roux, F. 1996c. Peritoneal endometriosis and "endometriotic" nodules of the rectovaginal septum are two different entities. *Fertil Steril,* 66**,** 362-8.

Dos Santos Hidalgo, G., Meola, J., Rosa, E. S. J. C., Paro De Paz, C. C. & Ferriani, R. A. 2011. TAGLN expression is deregulated in endometriosis and may be involved in cell invasion, migration, and differentiation. *Fertil Steril,* 96**,** 700-3.

Eskenazi, B., Warner, M., Bonsignore, L., Olive, D., Samuels, S. & Vercellini, P. 2001. Validation study of nonsurgical diagnosis of endometriosis. *Fertil Steril,* 76**,** 929-35.

Ezura, Y., Chakravarti, S., Oldberg, Å., Chervoneva, I. & Birk, D. E. 2000. Differential Expression of Lumican and Fibromodulin Regulate Collagen Fibrillogenesis in Developing Mouse Tendons. *The Journal of Cell Biology,* 151**,** 779-788.

Falcone, T. & Lebovic, D. I. 2011. Clinical management of endometriosis. *Obstet Gynecol,* 118**,** 691-705.

Fassbender, A., Rahmioglu, N., Vitonis, A. F., Viganò, P., Giudice, L. C., D'hooghe, T. M., Hummelshoj, L., Adamson, G. D., Becker, C. M., Missmer, S. A.,

Zondervan, K. T., Adamson, G. D., Allaire, C., Anchan, R., Becker, C. M., Bedaiwy, M. A., Buck Louis, G. M., Calhaz-Jorge, C., Chwalisz, K., D'hooghe, T. M., Fassbender, A., Faustmann, T., Fazleabas, A. T., Flores, I., Forman, A., Fraser, I., Giudice, L. C., Gotte, M., Gregersen, P., Guo, S. W., Harada, T., Hartwell, D., Horne, A. W., Hull, M. L., Hummelshoj, L., Ibrahim, M. G., Kiesel, L., Laufer, M. R., Machens, K., Mechsner, S., Missmer, S. A., Montgomery, G. W., Nap, A., Nyegaard, M., Osteen, K. G., Petta, C. A., Rahmioglu, N., Renner, S. P., Riedlinger, J., Roehrich, S., Rogers, P. A., Rombauts, L., Salumets, A., Saridogan, E., Seckin, T., Stratton, P., Sharpe-Timms, K. L., Tworoger, S., Vigano, P., Vincent, K., Vitonis, A. F., Wienhues-Thelen, U. H., Yeung Jr, P. P., Yong, P. & Zondervan, K. T. 2014. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonisation Project: IV. Tissue collection, processing, and storage in endometriosis research. *Fertility and Sterility,* 102**,** 1244-1253.

Fassbender, A., Verbeeck, N., Bornigen, D., Kyama, C. M., Bokor, A., Vodolazkaia, A., Peeraer, K., Tomassetti, C., Meuleman, C., Gevaert, O., Van De Plas, R., Ojeda, F., De Moor, B., Moreau, Y., Waelkens, E. & D'hooghe, T. M. 2012a. Combined mRNA microarray and proteomic analysis of eutopic endometrium of women with and without endometriosis. *Hum Reprod,* 27**,** 2020-9.

Fassbender, A., Vodolazkaia, A., Saunders, P., Lebovic, D., Waelkens, E., De Moor, B. & D'hooghe, T. 2013. Biomarkers of endometriosis. *Fertil Steril,* 99**,** 1135-45.

Fassbender, A., Waelkens, E., Verbeeck, N., Kyama, C. M., Bokor, A., Vodolazkaia, A., Van De Plas, R., Meuleman, C., Peeraer, K., Tomassetti, C., Gevaert, O., Ojeda, F., De Moor, B. & D'hooghe, T. 2012b. Proteomics Analysis of Plasma for Early Diagnosis of Endometriosis. *Obstetrics and Gynecology,* 119**,** 276-285.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science,* 246**,** 64-71.

Fowler, P. A., Tattum, J., Bhattacharya, S., Klonisch, T., Hombach-Klonisch, S., Gazvani, R., Lea, R. G., Miller, I., Simpson, W. G. & Cash, P. 2007. An investigation of the effects of endometriosis on the proteome of human eutopic endometrium: a heterogeneous tissue with a complex disease. *Proteomics,* 7**,** 130-42.

Fujii, S. 1991. Secondary mullerian system and endometriosis. *Am J Obstet Gynecol,* 165**,** 219-25.

Gajbhiye, R., Sonawani, A., Khan, S., Suryawanshi, A., Kadam, S., Warty, N., Raut, V. & Khole, V. 2012. Identification and validation of novel serum markers for early diagnosis of endometriosis. *Human Reproduction,* 27**,** 408-417.

Gajbhiye, R., Suryawanshi, A., Khan, S., Meherji, P., Warty, N., Raut, V., Chehna, N. & Khole, V. 2008. Multiple endometrial antigens are targeted in autoimmune endometriosis. *Reproductive Biomedicine Online,* 16**,** 817-24.

Gao, X., Outley, J., Botteman, M., Spalding, J., Simon, J. A. & Pashos, C. L. 2006a. Economic burden of endometriosis. *Fertil Steril,* 86**,** 1561-72.

Gao, X., Yeh, Y. C., Outley, J., Simon, J., Botteman, M. & Spalding, J. 2006b. Health-related quality of life burden of women with endometriosis: a literature review. *Curr Med Res Opin,* 22**,** 1787-97.

Garcia Manero, M., Olartecoechea, B., Auba, M., Alcazar, J. L. & Lopez, G. 2009. [Angiogenesis and endometriosis]. *Rev Med Univ Navarra,* 53**,** 8-13.

Garrels, J. I. 1979. Two dimensional gel electrophoresis and computer analysis of proteins synthesized by clonal cell lines. *J Biol Chem,* 254**,** 7961-77.

Garry, R. 2004. The effectiveness of laparoscopic excision of endometriosis. *Curr Opin Obstet Gynecol,* 16**,** 299-303.

Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M. J., Cramer, R., Waterfield, M. D. & Timms, J. F. 2002. Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. *Mol Cell Proteomics,* 1**,** 91-8.

Gilabert-Estelles, J., Ramon, L. A., Espana, F., Gilabert, J., Vila, V., Reganon, E., Castello, R., Chirivella, M. & Estelles, A. 2007. Expression of angiogenic factors in endometriosis: relationship to fibrinolytic and metalloproteinase systems. *Hum Reprod,* 22**,** 2120-7.

Giudice, L. C. 2003a. Elucidating endometrial function in the post-genomic era. *Hum Reprod Update,* 9**,** 223-35.

Giudice, L. C. 2003b. Genomics' role in understanding the pathogenesis of endometriosis. *Semin Reprod Med,* 21**,** 119-24.

Giudice, L. C., Jacobs, A., Pineda, J., Bell, C. E. & Lippmann, L. 1986. Serum levels of CA-125 in patients with endometriosis: a preliminary report. *Fertil Steril,* 45**,** 876-8.

Giudice, L. C. & Kao, L. C. 2004. Endometriosis. *Lancet,* 364**,** 1789-99.

Gonzalez-Ramos, R., Van Langendonckt, A., Defrere, S., Lousse, J. C., Colette, S., Devoto, L. & Donnez, J. 2010. Involvement of the nuclear factor-kappaB pathway in the pathogenesis of endometriosis. *Fertil Steril,* 94**,** 1985-94.

Gonzalez-Ramos, R., Van Langendonckt, A., Defrere, S., Lousse, J. C., Mettlen, M., Guillet, A. & Donnez, J. 2008. Agents blocking the nuclear factor-kappaB pathway are effective inhibitors of endometriosis in an in vivo experimental model. *Gynecol Obstet Invest,* 65**,** 174-86.

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H. & Aebersold, R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol,* 17**,** 994-9.

Hadfield, R., Mardon, H., Barlow, D. & Kennedy, S. 1996. Delay in the diagnosis of endometriosis: a survey of women from the USA and the UK. *Hum Reprod,* 11**,** 878-80.

Halme, J., Hammond, M. G., Hulka, J. F., Raj, S. G. & Talbert, L. M. 1984. Retrograde menstruation in healthy women and in patients with endometriosis. *Obstet Gynecol,* 64**,** 151-4.

Hanash, S. 2003. Disease proteomics. *Nature,* 422**,** 226-32.

Haney, A. F., Muscato, J. J. & Weinberg, J. B. 1981. Peritoneal fluid cell populations in infertility patients. *Fertil Steril,* 35**,** 696-8.

Harada, T., Iwabe, T. & Terakawa, N. 2001. Role of cytokines in endometriosis. *Fertil Steril,* 76**,** 1-10.

Harada, T., Kubota, T. & Aso, T. 2002. Usefulness of CA19-9 versus CA125 for the diagnosis of endometriosis. *Fertil Steril,* 78**,** 733-9.

Hardman, M. & Makarov, A. A. 2003. Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal Chem,* 75**,** 1699-705.

Harrington, D. J., Lessey, B. A., Rai, V., Bergqvist, A., Kennedy, S., Manek, S., Barlow, D. H. & Mardon, H. J. 1999. Tenascin is differentially expressed in endometrium and endometriosis. *J Pathol,* 187**,** 242-8.

Hatzinger, M., Kwon, S. T., Langbein, S., Kamp, S., Hacker, A. & Alken, P. 2006. Hans Christian Jacobaeus: Inventor of human laparoscopy and thoracoscopy. *J Endourol,* 20**,** 848-50.

Herington, J. L., Bruner-Tran, K. L., Lucas, J. A. & Osteen, K. G. 2011. Immune interactions in endometriosis. *Expert Rev Clin Immunol,* 7**,** 611-26.

Hood, B. L., Liu, B., Alkhas, A., Shoji, Y., Challa, R., Wang, G., Ferguson, S., Oliver, J., Mitchell, D., Bateman, N. W., Zahn, C. M., Hamilton, C. A., Payson, M., Lessey, B., Fazleabas, A. T., Maxwell, G. L., Conrads, T. P. & Risinger, J. I. 2015. Proteomics of the human endometrial glandular epithelium and stroma from the proliferative and secretory phases of the menstrual cycle. *Biol Reprod,* 92**,** 106.

Horie, S., Harada, T., Mitsunari, M., Taniguchi, F., Iwabe, T. & Terakawa, N. 2005. Progesterone and progestational compounds attenuate tumor necrosis factor alpha-induced interleukin-8 production via nuclear factor kappa B inactivation in endometriotic stromal cells. *Fertil Steril,* 83**,** 1530-5.

Hoshiai, H., Ishikawa, M., Sawatari, Y., Noda, K. & Fukaya, T. 1993. Laparoscopic evaluation of the onset and progression of endometriosis. *Am J Obstet Gynecol,* 169**,** 714-9.

Houston, D. E. 1984. Evidence for the risk of pelvic endometriosis by age, race and socioeconomic status. *Epidemiol Rev,* 6**,** 167-91.

Houston, D. E., Noller, K. L., Melton, L. J., 3rd & Selwyn, B. J. 1988. The epidemiology of pelvic endometriosis. *Clin Obstet Gynecol,* 31**,** 787-800.

Hsu, A. L., Khachikyan, I. & Stratton, P. 2010. Invasive and noninvasive methods for the diagnosis of endometriosis. *Clin Obstet Gynecol,* 53**,** 413-9.

Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M. & Graham Cooks, R. 2005. The Orbitrap: a new mass spectrometer. *J Mass Spectrom,* 40**,** 430-43.

Hudelist, G., Czerwenka, K., Keckstein, J., Haas, C., Fink-Retter, A., Gschwantler-Kaulich, D., Kubista, E. & Singer, C. F. 2007. Expression of aromatase and estrogen sulfotransferase in eutopic and ectopic endometrium: evidence for unbalanced estradiol production in endometriosis. *Reproductive Sciences,* 14**,** 798-805.

Hudelist, G., Keckstein, J. & Wright, J. T. 2009. The migrating adenomyoma: past views on the etiology of adenomyosis and endometriosis. *Fertil Steril,* 92**,** 1536-43.

Hughesdon, P. E. 1957. The structure of endometrial cysts of the ovary. *J Obstet Gynaecol Br Emp,* 64**,** 481-7.

Husby, G. K., Haugen, R. S. & Moen, M. H. 2003. Diagnostic delay in women with pain and endometriosis. *Acta Obstet Gynecol Scand,* 82**,** 649-53.

Iwabe, T., Harada, T. & Terakawa, N. 2002. Role of cytokines in endometriosis-associated infertility. *Gynecol Obstet Invest,* 53 Suppl 1**,** 19-25.

J, H. 1924. Hysteroadenosis metastica (die lymphogene genes dersog. adeno-fibromatosis hetertopica). *Wein Klin Wochenschr,* 37**,** 1205.

Jain, K. K. 2010. *The Handbook of Biomarkers*, Springer Science.

Javert, C. T. 1949. Pathogenesis of endometriosis based on endometrial homeoplasia, direct extension, exfoliation and implantation, lymphatic and hematogenous metastasis, including five case reports of endometrial tissue in pelvic lymph nodes. *Cancer,* 2**,** 399-410.

Jenkins, S., Olive, D. L. & Haney, A. F. 1986. Endometriosis: pathogenetic implications of the anatomic distribution. *Obstet Gynecol,* 67**,** 335-8.

Jensen, J. R. & Coddington, C. C., 3rd 2010. Evolving spectrum: the pathogenesis of endometriosis. *Clin Obstet Gynecol,* 53**,** 379-88.

Joshi, S. G., Zamah, N. M., Raikar, R. S., Buttram, V. C., Jr., Henriques, E. S. & Gordon, M. 1986. Serum and peritoneal fluid proteins in women with and without endometriosis. *Fertil Steril,* 46**,** 1077-82.

Kaminski, K., Kotarski, J. & Gogacz, M. 1995. Characterization of cellular components in the peritoneal fluid and in the endometrial tissue of women with endometriosis. *Arch Immunol Ther Exp (Warsz),* 43**,** 247-52.

Kao, L. C., Germeyer, A., Tulac, S., Lobo, S., Yang, J. P., Taylor, R. N., Osteen, K., Lessey, B. A. & Giudice, L. C. 2003. Expression profiling of endometrium from women with endometriosis reveals candidate genes for disease-based implantation failure and infertility. *Endocrinology,* 144**,** 2870-81.

Karas, M. & Hillenkamp, F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem,* 60**,** 2299-301.

Kennedy, S. 1999. The genetics of endometriosis. *Eur J Obstet Gynecol Reprod Biol,* 82**,** 129-33.

Kennedy, S. 2006. Should a diagnosis of endometriosis be sought in all symptomatic women? *Fertil Steril,* 86**,** 1312-3; discussion 1317.

Kennedy, S., Bergqvist, A., Chapron, C., D'hooghe, T., Dunselman, G., Greb, R., Hummelshoj, L., Prentice, A. & Saridogan, E. 2005. ESHRE guideline for the diagnosis and treatment of endometriosis. *Hum Reprod,* 20**,** 2698-704.

Kennedy, S. H., Cederholm-Williams, S. A. & Barlow, D. H. 1992. The effect of injecting endometriotic 'chocolate' cyst fluid into the peritoneal cavity of mice. *Hum Reprod,* 7**,** 1329.

Kharfi, A. & Akoum, A. 2001. Correlation between decreased type-II interleukin-1 receptor and increased monocyte chemotactic protein-1 expression in the endometrium of women with endometriosis. *Am J Reprod Immunol,* 45**,** 193-9.

Kharfi, A., Boucher, A. & Akoum, A. 2002. Abnormal interleukin-1 receptor type II gene expression in the endometrium of women with endometriosis. *Biol Reprod,* 66**,** 401-6.

Kicman, A. T., Parkin, M. C. & Iles, R. K. 2007. An introduction to mass spectrometry based proteomics-detection and characterization of gonadotropins and related molecules. *Mol Cell Endocrinol,* 260-262**,** 212-27.

Kitawaki, J., Ishihara, H., Koshiba, H., Kiyomizu, M., Teramoto, M., Kitaoka, Y. & Honjo, H. 2005. Usefulness and limits of CA-125 in diagnosis of endometriosis without associated ovarian endometriomas. *Hum Reprod,* 20**,** 1999-2003.

Kitawaki, J., Kusuki, I., Koshiba, H., Tsukamoto, K., Fushiki, S. & Honjo, H. 1999a. Detection of aromatase cytochrome P-450 in endometrial biopsy specimens as a diagnostic test for endometriosis. *Fertil Steril,* 72**,** 1100-6.

Kitawaki, J., Kusuki, I., Koshiba, H., Tsukamoto, K. & Honjo, H. 1999b. Expression of aromatase cytochrome P450 in eutopic endometrium and its application as a diagnostic test for endometriosis. *Gynecol Obstet Invest,* 48 Suppl 1**,** 21-8.

Kito, K. & Ito, T. 2008. Mass spectrometry-based approaches toward absolute quantitative proteomics. *Curr Genomics,* 9**,** 263-74.

Kjerulff, K. H., Erickson, B. A. & Langenberg, P. W. 1996. Chronic gynecological conditions reported by US women: findings from the National Health Interview Survey, 1984 to 1992. *Am J Public Health,* 86**,** 195-9.

Knapp, V. J. 1999. How old is endometriosis? Late 17th- and 18th-century European descriptions of the disease. *Fertil Steril,* 72**,** 10-4.

Kocbek, V., Vouk, K., Mueller, M. D., Rizner, T. L. & Bersinger, N. A. 2013. Elevated glycodelin-A concentrations in serum and peritoneal fluid of women with ovarian endometriosis. *Gynecological Endocrinology,* 29**,** 455-9.

Kokorine, I., Nisolle, M., Donnez, J., Eeckhout, Y., Courtoy, P. J. & Marbaix, E. 1997. Expression of interstitial collagenase (matrix metalloproteinase-1) is related to the activity of human endometriotic lesions. *Fertil Steril,* 68**,** 246-51.

Koninckx, P. R., Oosterlynck, D., D'hooghe, T. & Meuleman, C. 1994. Deeply infiltrating endometriosis is a disease whereas mild endometriosis could be considered a non-disease. *Ann N Y Acad Sci,* 734**,** 333-41.

Koninckx, P. R., Riittinen, L., Seppala, M. & Cornillie, F. J. 1992. CA-125 and placental protein 14 concentrations in plasma and peritoneal fluid of women with deeply infiltrating pelvic endometriosis. *Fertil Steril,* 57**,** 523-30.

Koninckx, P. R., Ussia, A., Adamyan, L., Wattiez, A. & Donnez, J. 2012. Deep endometriosis: definition, diagnosis, and treatment. *Fertil Steril,* 98**,** 564-71.

Koninger, J., Giese, T., Di Mola, F. F., Wente, M. N., Esposito, I., Bachem, M. G., Giese, N. A., Buchler, M. W. & Friess, H. 2004. Pancreatic tumor cells influence the composition of the extracellular matrix. *Biochem Biophys Res Commun,* 322**,** 943-9.

Kurdoglu, Z., Gursoy, R., Kurdoglu, M., Erdem, M., Erdem, O. & Erdem, A. 2009. Comparison of the clinical value of CA 19-9 versus CA 125 for the diagnosis of endometriosis. *Fertility and Sterility,* 92**,** 1761-3.

Kyama, C. M., Debrock, S., Mwenda, J. M. & D'hooghe, T. M. 2003. Potential involvement of the immune system in the development of endometriosis. *Reprod Biol Endocrinol,* 1**,** 123.

Kyama, C. M., Mihalyi, A., Gevaert, O., Simsa, P., Waelkens, E., De Plas, R. V., Mwenda, J. M., Meuleman, C. & D'hooghe, T. M. 2007. Evaluation of endometrial biomarkers for semi-invasive diagnosis of endometriosis. *Human Reproduction,* 22**,** I107-I107.

Kyama, C. M., Mihalyi, A., Gevaert, O., Waelkens, E., Simsa, P., Van De Plas, R., Meuleman, C., De Moor, B. & D'hooghe, T. M. 2011. Evaluation of endometrial biomarkers for semi-invasive diagnosis of endometriosis. *Fertil Steril,* 95**,** 1338-U173.

Kyama, C. M., Overbergh, L., Debrock, S., Valckx, D., Vander Perre, S., Meuleman, C., Mihalyi, A., Mwenda, J. M., Mathieu, C. & D'hooghe, T. M. 2006a. Increased peritoneal and endometrial gene expression of biologically relevant cytokines and growth factors during the menstrual phase in women with endometriosis. *Fertil Steril,* 85**,** 1667-75.

Kyama, C. M., Overbergh, L., Mihalyi, A., Meuleman, C., Mwenda, J. M., Mathieu, C. & D'hooghe, T. M. 2008. Endometrial and peritoneal expression of aromatase, cytokines, and adhesion factors in women with endometriosis. *Fertility and Sterility,* 89**,** 301-10.

Kyama, C. M., T'jampens, D., Mihalyi, A., Simsa, P., Debrock, S., Waelkens, E., Landuyt, B., Meuleman, C., Fulop, V., Mwenda, J. M. & D'hooghe, T. M. 2006b. ProteinChip technology is a useful method in the pathogenesis and diagnosis of endometriosis: a preliminary study. *Fertil Steril,* 86**,** 203-9.

Lamb, K., Hoffmann, R. G. & Nichols, T. R. 1986. Family trait analysis: a case-control study of 43 women with endometriosis and their best friends. *Am J Obstet Gynecol,* 154**,** 596-601.

Lawson, C., Al-Akoum, M., Maheux, R. & Akoum, A. 2007. Increased expression of interleukin-1 receptor type 1 in active endometriotic lesions. *Reproduction,* 133**,** 265-74.

Lawson, C., Bourcier, N., Al-Akoum, M., Maheux, R., Naud, F. & Akoum, A. 2008. Abnormal interleukin 1 receptor types I and II gene expression in eutopic and ectopic endometrial tissues of women with endometriosis. *J Reprod Immunol,* 77**,** 75-84.

Lebovic, D. I., Mueller, M. D. & Taylor, R. N. 2001. Immunobiology of endometriosis. *Fertil Steril,* 75**,** 1-10.

Leygue, E., Snell, L., Dotzlaw, H., Hole, K., Hiller-Hitchcock, T., Roughley, P. J., Watson, P. H. & Murphy, L. C. 1998. Expression of lumican in human breast carcinoma. *Cancer Res,* 58**,** 1348-52.

Linda C Giudice, J. L. H. E., David L. Healey 2012. *Endometriosis. Science and Practice,* United Kingdom, Wiley-Blackwell.

Litynski, G. S. 1997. Laparoscopy--the early attempts: spotlighting Georg Kelling and Hans Christian Jacobaeus. *JSLS,* 1**,** 83-5.

Liu, H., Sadygov, R. G. & Yates, J. R., 3rd 2004. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem,* 76**,** 4193-201.

Lu, Y. P., Ishiwata, T., Kawahara, K., Watanabe, M., Naito, Z., Moriyama, Y., Sugisaki, Y. & Asano, G. 2002. Expression of lumican in human colorectal cancer cells. *Pathology International,* 52**,** 519-526.

Maia, H., Jr., Casoy, J. & Valente Filho, J. 2009. Is aromatase expression in the endometrium the cause of endometriosis and related infertility? *Gynecological Endocrinology,* 25**,** 253-7.

Malinak, L. R., Buttram, V. C., Jr., Elias, S. & Simpson, J. L. 1980. Heritage aspects of endometriosis. II. Clinical characteristics of familial endometriosis. *Am J Obstet Gynecol,* 137**,** 332-7.

196

Mangtani, P. & Booth, M. 1993. Epidemiology of endometriosis. *J Epidemiol Community Health,* 47**,** 84-8.

Mano, N. & Goto, J. 2003. Biomedical and biological mass spectrometry. *Anal Sci,* 19**,** 3-14.

Marjoribanks, J., Proctor, M. L. & Farquhar, C. 2003. Nonsteroidal anti-inflammatory drugs for primary dysmenorrhoea. *Cochrane Database Syst Rev***,** Cd001751.

Martin, J. D., Jr. & Hauck, A. E. 1985. Endometriosis in the male. *Am Surg,* 51**,** 426-30.

Martinez, S., Garrido, N., Coperias, J. L., Pardo, F., Desco, J., Garcia-Velasco, J. A., Simon, C. & Pellicer, A. 2007. Serum interleukin-6 levels are elevated in women with minimal-mild endometriosis. *Hum Reprod,* 22**,** 836-42.

Matalliotakis, I., Makrigiannakis, A., Karkavitsas, N., Psaroudakis, E., Froudarakis, G. & Koumantakis, E. 1994. Use of CA-125 in the diagnosis and management of endometriosis: influence of treatment with danazol. *Int J Fertil Menopausal Stud,* 39**,** 100-4.

Matalliotakis, I., Neonaki, M., Zolindaki, A., Hassan, E., Georgoulias, V. & Koumantakis, E. 1997. Changes in immunologic variables (TNF-a, sCD8 and sCD4) during danazol treatment in patients with endometriosis. *Int J Fertil Womens Med,* 42**,** 211-4.

Matalliotakis, I., Panidis, D., Vlassis, G., Neonaki, M., Goumenou, A. & Koumantakis, E. 1998. Unexpected increase of the CA 19-9 tumour marker in patients with endometriosis. *Eur J Gynaecol Oncol,* 19**,** 498-500.

Matalliotakis, I. M., Vassiliadis, S., Goumenou, A. G., Athanassakis, I., Koumantakis, G. E., Neonaki, M. A. & Koumantakis, E. E. 2001. Soluble ICAM-1 levels in the serum of endometriotic patients appear to be independent of medical treatment. *J Reprod Immunol,* 51**,** 9-19.

Matsuura, K., Ohtake, H., Katabuchi, H. & Okamura, H. 1999. Coelomic metaplasia theory of endometriosis: evidence from in vivo studies and an in vitro experimental model. *Gynecol Obstet Invest,* 47 Suppl 1**,** 18-20; discussion 20-2.

May, K. E., Conduit-Hulbert, S. A., Villar, J., Kirtley, S., Kennedy, S. H. & Becker, C. M. 2010. Peripheral biomarkers of endometriosis: a systematic review. *Human Reproduction Update,* 16**,** 651-74.

May, K. E., Villar, J., Kirtley, S., Kennedy, S. H. & Becker, C. M. 2011. Endometrial alterations in endometriosis: a systematic review of putative biomarkers. *Human Reproduction Update,* 17**,** 637-53.

Mclaren, J., Prentice, A., Charnock-Jones, D. S., Millican, S. A., Muller, K. H., Sharkey, A. M. & Smith, S. K. 1996a. Vascular endothelial growth factor is produced by peritoneal fluid macrophages in endometriosis and is regulated by ovarian steroids. *J Clin Invest,* 98**,** 482-9.

Mclaren, J., Prentice, A., Charnock-Jones, D. S. & Smith, S. K. 1996b. Vascular endothelial growth factor (VEGF) concentrations are elevated in peritoneal fluid of women with endometriosis. *Hum Reprod,* 11**,** 220-3.

Meehan, K. L., Rainczuk, A., Salamonsen, L. A. & Stephens, A. N. 2010. Proteomics and the search for biomarkers of female reproductive diseases. *Reproduction,* 140**,** 505-519.

Meola, J., Dentillo, D. B., Rosa E Silva, J. C., Ferriani, R. A., Veiga, L. C., Paro De Paz, C. C., Giuliatti, S. & Martelli, L. 2009. Glycodelin expression in the endometrium of healthy women and in the eutopic and ectopic endometrium of women with endometriosis. *Fertil Steril,* 91**,** 1676-80.

Meuleman, C., D'hoore, A., Van Cleynenbreugel, B., Beks, N. & D'hooghe, T. 2009. Outcome after multidisciplinary CO2 laser laparoscopic excision of deep infiltrating colorectal endometriosis. *Reprod Biomed Online,* 18**,** 282-9.

Mihalyi, A., Gevaert, O., Kyama, C. M., Simsa, P., Pochet, N., De Smet, F., De Moor, B., Meuleman, C., Billen, J., Blanckaert, N., Vodolazkaia, A., Fulop, V. & D'hooghe, T. M. 2010. Non-invasive diagnosis of endometriosis based on a combined analysis of six plasma biomarkers. *Human Reproduction,* 25**,** 654-64.

Minden, J. S., Dowd, S. R., Meyer, H. E. & Stuhler, K. 2009. Difference gel electrophoresis. *Electrophoresis,* 30 Suppl 1**,** S156-61.

Moen, M. H., Hagen, B. & Onsrud, M. 1991. CA 125 in peritoneal fluid from patients with endometriosis. *Hum Reprod,* 6**,** 1400-3.

Moen, M. H. & Magnus, P. 1993. The familial risk of endometriosis. *Acta Obstet Gynecol Scand,* 72**,** 560-4.

Moen, M. H. & Schei, B. 1997. Epidemiology of endometriosis in a Norwegian county. *Acta Obstet Gynecol Scand,* 76**,** 559-62.

Mok-Lin, E. Y., Wolfberg, A., Hollinquist, H. & Laufer, M. R. 2010. Endometriosis in a patient with Mayer-Rokitansky-Kuster-Hauser syndrome and complete

uterine agenesis: evidence to support the theory of coelomic metaplasia. *J Pediatr Adolesc Gynecol,* 23**,** e35-7.

Mol, B. W., Bayram, N., Lijmer, J. G., Wiegerinck, M. A., Bongers, M. Y., Van Der Veen, F. & Bossuyt, P. M. 1998. The performance of CA-125 measurement in the detection of endometriosis: a meta-analysis. *Fertil Steril,* 70**,** 1101-8.

Molgaard, C. A., Golbeck, A. L. & Gresham, L. 1985. Current concepts in endometriosis. *West J Med,* 143**,** 42-6.

Moore, J., Copley, S., Morris, J., Lindsell, D., Golding, S. & Kennedy, S. 2002. A systematic review of the accuracy of ultrasound in the diagnosis of endometriosis. *Ultrasound Obstet Gynecol,* 20**,** 630-4.

Moretuzzo, R. W., Dilauro, S., Jenison, E., Chen, S. L., Reindollar, R. H. & Mcdonough, P. G. 1988. Serum and peritoneal lavage fluid CA-125 levels in endometriosis. *Fertil Steril,* 50**,** 430-3.

Mu, L. & Ma, Y.-Y. 2015. Expression of focal adhesion kinase in endometrial stromal cells of women with endometriosis was adjusted by ovarian steroid hormones. *International Journal of Clinical and Experimental Pathology,* 8**,** 1810-1815.

Nap, A. W., Groothuis, P. G., Demir, A. Y., Evers, J. L. & Dunselman, G. A. 2004. Pathogenesis of endometriosis. *Best Pract Res Clin Obstet Gynaecol,* 18**,** 233-44.

Nguyen, S. & Fenn, J. B. 2007. Gas-phase ions of solute species from charged droplets of solutions. *Proc Natl Acad Sci U S A,* 104**,** 1111-7.

Nisolle, M. & Donnez, J. 1997. Peritoneal endometriosis, ovarian endometriosis, and adenomyotic nodules of the rectovaginal septum are three different entities. *Fertil Steril,* 68**,** 585-96.

Nnoaham, K. E., Hummelshoj, L., Webster, P., D'hooghe, T., De Cicco Nardone, F., De Cicco Nardone, C., Jenkinson, C., Kennedy, S. H. & Zondervan, K. T. 2011. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertil Steril,* 96**,** 366-373 e8.

Noble, L. S., Simpson, E. R., Johns, A. & Bulun, S. E. 1996. Aromatase expression in endometriosis. *J Clin Endocrinol Metab,* 81**,** 174-9.

Nothnick, W. B. 2011. The emerging use of aromatase inhibitors for endometriosis treatment. *Reproductive Biology and Endocrinology : RB&E,* 9**,** 87-87.

O'farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem,* 250**,** 4007-21.

Oosterlynck, D. J., Meuleman, C., Waer, M., Vandeputte, M. & Koninckx, P. R. 1992. The natural killer activity of peritoneal fluid lymphocytes is decreased in women with endometriosis. *Fertil Steril,* 58**,** 290-5.

Othman, E. E. D. R., Homung, D., Salem, H. T., Khalifa, E. A., El-Metwally, T. H. & Al-Hendy, A. 2008. Serum cytokines as biomarkers for nonsurgical prediction of endometriosis. *European Journal of Obstetrics Gynecology and Reproductive Biology,* 137**,** 240-246.

Ozawa, Y., Murakami, T., Terada, Y., Yaegashi, N., Okamura, K., Kuriyama, S. & Tsuji, I. 2006. Management of the pain associated with endometriosis: an update of the painful problems. *Tohoku J Exp Med,* 210**,** 175-88.

Panidis, D., Vlassis, G., Matalliotakis, J., Skiadopoulos, S. & Kalogeropoulos, A. 1988. Serum levels of the oncofetal antigens CA-125, CA 19-9 and CA 15-3 in patients with endometriosis. *J Endocrinol Invest,* 11**,** 801-4.

Parazzini, F., Ferraroni, M., Fedele, L., Bocciolone, L., Rubessa, S. & Riccardi, A. 1995. Pelvic endometriosis: reproductive and menstrual risk factors at different stages in Lombardy, northern Italy. *J Epidemiol Community Health,* 49**,** 61-4.

Patentstorm.Us/Patents/7794958, W.

Patterson, S. D. & Aebersold, R. H. 2003. Proteomics: the first decade and beyond. *Nat Genet,* 33 Suppl**,** 311-23.

Pittaway, D. E. 1989. CA-125 in women with endometriosis. *Obstet Gynecol Clin North Am,* 16**,** 237-52.

Pittaway, D. E. & Fayez, J. A. 1986. The use of CA-125 in the diagnosis and management of endometriosis. *Fertil Steril,* 46**,** 790-5.

Pittaway, D. E., Rondinone, D., Miller, K. A. & Barnes, K. 1995. Clinical evaluation of CA-125 concentrations as a prognostic factor for pregnancy in infertile women with surgically treated endometriosis. *Fertil Steril,* 64**,** 321-4.

Pizzo, A., Salmeri, F. M., Ardita, F. V., Sofo, V., Tripepi, M. & Marsico, S. 2002. Behaviour of cytokine levels in serum and peritoneal fluid of women with endometriosis. *Gynecol Obstet Invest,* 54**,** 82-7.

Poliness, A. E., Healey, M. G., Brennecke, S. R. & Moses, E. K. 2004. Proteomic approaches in endometriosis research. *Proteomics,* 4**,** 1897-1902.

Pollard, A. J., Sparey, C., Robson, S. C., Krainer, A. R. & Europe-Finner, G. N. 2000. Spatio-temporal expression of the trans-acting splicing factors SF2/ASF and

heterogeneous ribonuclear proteins A1/A1B in the myometrium of the pregnant human uterus: a molecular mechanism for regulating regional protein isoform expression in vivo. *J Clin Endocrinol Metab,* 85**,** 1928-36.

Pupo-Nogueira, A., De Oliveira, R. M., Petta, C. A., Podgaec, S., Dias, J. A., Jr. & Abrao, M. S. 2007. Vascular endothelial growth factor concentrations in the serum and peritoneal fluid of women with endometriosis. *Int J Gynaecol Obstet,* 99**,** 33-7.

Rahmioglu, N., Fassbender, A., Vitonis, A. F., Tworoger, S. S., Hummelshoj, L., D'hooghe, T. M., Adamson, G. D., Giudice, L. C., Becker, C. M., Zondervan, K. T., Missmer, S. A., Adamson, G. D., Allaire, C., Anchan, R., Becker, C. M., Bedaiwy, M. A., Buck Louis, G. M., Calhaz-Jorge, C., Chwalisz, K., D'hooghe, T. M., Fassbender, A., Faustmann, T., Fazleabas, A. T., Flores, I., Forman, A., Fraser, I., Giudice, L. C., Gotte, M., Gregersen, P., Guo, S. W., Harada, T., Hartwell, D., Horne, A. W., Hull, M. L., Hummelshoj, L., Ibrahim, M. G., Kiesel, L., Laufer, M. R., Machens, K., Mechsner, S., Missmer, S. A., Montgomery, G. W., Nap, A., Nyegaard, M., Osteen, K. G., Petta, C. A., Rahmioglu, N., Renner, S. P., Riedlinger, J., Roehrich, S., Rogers, P. A., Rombauts, L., Salumets, A., Saridogan, E., Seckin, T., Stratton, P., Sharpe-Timms, K. L., Tworoger, S., Vigano, P., Vincent, K., Vitonis, A. F., Wienhues-Thelen, U. H., Yeung Jr, P. P., Yong, P. & Zondervan, K. T. 2014. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking Harmonization Project: III. Fluid biospecimen collection, processing, and storage in endometriosis research. *Fertility and Sterility,* 102**,** 1233-1243.

Ramon, L., Gilabert-Estelles, J., Castello, R., Gilabert, J., Espana, F., Romeu, A., Chirivella, M., Aznar, J. & Estelles, A. 2005. mRNA analysis of several components of the plasminogen activator and matrix metalloproteinase systems in endometriosis using a real-time quantitative RT-PCR assay. *Hum Reprod,* 20**,** 272-8.

Redwine, D. B. 2002. Was Sampson wrong? *Fertil Steril,* 78**,** 686-93.

Rier, S. E., Martin, D. C., Bowman, R. E., Dmowski, W. P. & Becker, J. L. 1993. Endometriosis in rhesus monkeys (Macaca mulatta) following chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Fundam Appl Toxicol,* 21**,** 433-41.

Riittinen, L., Julkunen, M., Seppala, M., Koistinen, R. & Huhtala, M. L. 1989. Purification and characterization of endometrial protein PP14 from mid-trimester amniotic fluid. *Clin Chim Acta,* 184**,** 19-29.

Rogers, P. A., D'hooghe, T. M., Fazleabas, A., Gargett, C. E., Giudice, L. C., Montgomery, G. W., Rombauts, L., Salamonsen, L. A. & Zondervan, K. T.

2009a. Priorities for endometriosis research: recommendations from an international consensus workshop. *Reproductive Sciences,* 16**,** 335-46.

Rogers, P. A., D'hooghe, T. M., Fazleabas, A., Gargett, C. E., Giudice, L. C., Montgomery, G. W., Rombauts, L., Salamonsen, L. A. & Zondervan, K. T. 2009b. Priorities for endometriosis research: recommendations from an international consensus workshop. *Reprod Sci,* 16**,** 335-46.

Rogers, P. A., D'hooghe, T. M., Fazleabas, A., Giudice, L. C., Montgomery, G. W., Petraglia, F. & Taylor, R. N. 2013. Defining Future Directions for Endometriosis Research: Workshop Report From the 2011 World Congress of Endometriosis in Montpellier, France. *Reprod Sci*.

Sagsveen, M., Farmer, J. E., Prentice, A. & Breeze, A. 2003. Gonadotrophin-releasing hormone analogues for endometriosis: bone mineral density. *Cochrane Database Syst Rev***,** Cd001297.

Sampson, J. A. 1918. The escape of foreign material from the uterine cavity into the uterine veins. *American Journal of Obstetrics,* 78**,** 161-175.

Sampson, J. A. 1927a. Metastatic or Embolic Endometriosis, due to the Menstrual Dissemination of Endometrial Tissue into the Venous Circulation. *Am J Pathol,* 3**,** 93-110 43.

Sampson, J. A. 1927b. Peritoneal endometriosis due to menstrual dissemination of endometrial tissue into the peritoneal cavity. *American Journal of Obstetrics and Gynaecology,* 14**,** 422-469.

San Martin, S., Soto-Suazo, M., De Oliveira, S. F., Aplin, J. D., Abrahamsohn, P. & Zorn, T. M. 2003. Small leucine-rich proteoglycans (SLRPs) in uterine tissues during pregnancy in mice. *Reproduction,* 125**,** 585-95.

Sangi-Haghpeykar, H. & Poindexter, A. N., 3rd 1995. Epidemiology of endometriosis among parous women. *Obstet Gynecol,* 85**,** 983-92.

Sasano, H., Nagura, H., Watanabe, K., Ito, K., Tsuiki, A., Sato, S., Yajima, A., Kusakabe, M. & Sakakura, T. 1993. Tenascin expression in normal and abnormal human endometrium. *Mod Pathol,* 6**,** 323-6.

Scigelova, M. & Makarov, A. 2006. Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics,* 6 Suppl 2**,** 16-21.

Seeber, B., Sammel, M. D., Fan, X., Gerton, G. L., Shaunik, A., Chittams, J. & Barnhart, K. T. 2008a. Panel of markers can accurately predict endometriosis in a subset of patients. *Fertility and Sterility,* 89**,** 1073-81.

Seeber, B., Sammel, M. D., Fan, X., Gerton, G. L., Shaunik, A., Chittams, J. & Barnhart, K. T. 2008b. Panel of markers can accurately predict endometriosis in a subset of patients. *Fertil Steril,* 89**,** 1073-81.

Seeber, B., Sammel, M. D., Fan, X. J., Gerton, G. L., Shaunik, A., Chittams, J. & Barnhart, K. T. 2010. Proteomic analysis of serum yields six candidate proteins that are differentially regulated in a subset of women with endometriosis. *Fertil Steril,* 93**,** 2137-2144.

Selak, V., Farquhar, C., Prentice, A. & Singla, A. 2007. Danazol for pelvic pain associated with endometriosis. *Cochrane Database Syst Rev***,** Cd000068.

Sensky, T. E. & Liu, D. T. 1980. Endometriosis: associations with menorrhagia, infertility and oral contraceptives. *Int J Gynaecol Obstet,* 17**,** 573-6.

Shah, V., Bharadwaj, S., Kaibuchi, K. & Prasad, G. L. 2001. Cytoskeletal organization in tropomyosin-mediated reversion of ras-transformation: Evidence for Rho kinase pathway. *Oncogene,* 20**,** 2112-21.

Shinji, S., Tajiri, T., Ishiwata, T., Seya, T., Tanaka, N. & Naito, Z. 2005. Different expression levels of lumican in human carcinoid tumor and neuroendocrine cell carcinoma. *Int J Oncol,* 26**,** 873-80.

Simoens, S., Hummelshoj, L. & D'hooghe, T. 2007. Endometriosis: cost estimates and methodological perspective. *Hum Reprod Update,* 13**,** 395-404.

Simoens, S., Hummelshoj, L., Dunselman, G., Brandes, I., Dirksen, C. & D'hooghe, T. 2011. Endometriosis cost assessment (the EndoCost study): a cost-of-illness study protocol. *Gynecol Obstet Invest,* 71**,** 170-6.

Simpson, J. L., Elias, S., Malinak, L. R. & Buttram, V. C., Jr. 1980. Heritable aspects of endometriosis. I. Genetic studies. *Am J Obstet Gynecol,* 137**,** 327-31.

Somigliana, E., Vercellini, P., Vigano, P., Benaglia, L., Crosignani, P. G. & Fedele, L. 2010. Non-invasive diagnosis of endometriosis: the goal or own goal? *Hum Reprod,* 25**,** 1863-8.

Somigliana, E., Vigano, P., Candiani, M., Felicetta, I., Di Blasio, A. M. & Vignali, M. 2002. Use of serum-soluble intercellular adhesion molecule-1 as a new marker of endometriosis. *Fertil Steril,* 77**,** 1028-31.

Somigliana, E., Vigano, P., Tirelli, A. S., Felicetta, I., Torresani, E., Vignali, M. & Di Blasio, A. M. 2004. Use of the concomitant serum dosage of CA 125, CA 19-9 and interleukin-6 to detect the presence of endometriosis. Results from a series of reproductive age women undergoing laparoscopic surgery for benign gynaecological conditions. *Hum Reprod,* 19**,** 1871-6.

Sourial, S., Tempest, N. & Hapangama, D. K. 2014. Theories on the pathogenesis of endometriosis. *Int J Reprod Med,* 2014**,** 179515.

Spaner, S. J. & Warnock, G. L. 1997. A brief history of endoscopy, laparoscopy, and laparoscopic surgery. *J Laparoendosc Adv Surg Tech A,* 7**,** 369-73.

Spuijbroek, M. D., Dunselman, G. A., Menheere, P. P. & Evers, J. L. 1992. Early endometriosis invades the extracellular matrix. *Fertil Steril,* 58**,** 929-33.

Stefansson, H., Geirsson, R. T., Steinthorsdottir, V., Jonsson, H., Manolescu, A., Kong, A., Ingadottir, G., Gulcher, J. & Stefansson, K. 2002. Genetic factors contribute to the risk of developing endometriosis. *Hum Reprod,* 17**,** 555-9.

Stephens, A. N., Hannan, N. J., Rainczuk, A., Meehan, K. L., Chen, J., Nicholls, P. K., Rombauts, L. J., Stanton, P. G., Robertson, D. M. & Salamonsen, L. A. 2010. Post-translational modifications and protein-specific isoforms in endometriosis revealed by 2D DIGE. *J Proteome Res,* 9**,** 2438-49.

Suginami, H. 1991. A reappraisal of the coelomic metaplasia theory by reviewing endometriosis occurring in unusual sites and instances. *Am J Obstet Gynecol,* 165**,** 214-8.

Sukhikh, G. T., Sotnikova, N. Y., Antsiferova, Y. S., Posiseeva, L. V., Veryasov, V. N. & Van'ko, L. V. 2004. Cytokine production by immunocompetent cells of peritoneal fluid in women with external genital endometriosis. *Bull Exp Biol Med,* 137**,** 568-71.

Tan, O., Ornek, T., Seval, Y., Sati, L. & Arici, A. 2008. Tenascin is highly expressed in endometriosis and its expression is upregulated by estrogen. *Fertility and Sterility,* 89**,** 1082-1089.

Tariverdian, N., Theoharides, T. C., Siedentopf, F., Gutierrez, G., Jeschke, U., Rabinovich, G. A., Blois, S. M. & Arck, P. C. 2007. Neuroendocrine-immune disequilibrium and endometriosis: an interdisciplinary approach. *Semin Immunopathol,* 29**,** 193-210.

Taylor, R. N., Lebovic, D. I. & Mueller, M. D. 2002. Angiogenic factors in endometriosis. *Ann N Y Acad Sci,* 955**,** 89-100; discussion 118, 396-406.

Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K. & Hamon, C. 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem,* 75**,** 1895-904.

Timms, J. F. & Cramer, R. 2008. Difference gel electrophoresis. *Proteomics,* 8**,** 4886-97.

Tokushige, N., Markham, R., Russell, P. & Fraser, I. S. 2006a. High density of small nerve fibres in the functional layer of the endometrium in women with endometriosis. *Hum Reprod,* 21**,** 782-7.

Tokushige, N., Markham, R., Russell, P. & Fraser, I. S. 2006b. Nerve fibres in peritoneal endometriosis. *Hum Reprod,* 21**,** 3001-7.

Treloar, S. A., O'connor, D. T., O'connor, V. M. & Martin, N. G. 1999. Genetic influences on endometriosis in an Australian twin sample. sueT@qimr.edu.au. *Fertil Steril,* 71**,** 701-10.

Tyers, M. & Mann, M. 2003. From genomics to proteomics. *Nature,* 422**,** 193-7.

Ulukus, M. & Arici, A. 2005. Immunology of endometriosis. *Minerva Ginecol,* 57**,** 237-48.

Unlu, M., Morgan, M. E. & Minden, J. S. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis,* 18**,** 2071-7.

Uno, S., Zembutsu, H., Hirasawa, A., Takahashi, A., Kubo, M., Akahane, T., Aoki, D., Kamatani, N., Hirata, K. & Nakamura, Y. 2010. A genome-wide association study identifies genetic variants in the CDKN2BAS locus associated with endometriosis in Japanese. *Nat Genet,* 42**,** 707-10.

Van Der Linden, P. J., De Goeij, A. F., Dunselman, G. A., Erkens, H. W. & Evers, J. L. 1995a. Expression of cadherins and integrins in human endometrium throughout the menstrual cycle. *Fertil Steril,* 63**,** 1210-6.

Van Der Linden, P. J., Dunselman, G. A., De Goeij, A. F., Van Der Linden, E. P., Evers, J. L. & Ramaekers, F. C. 1995b. Epithelial cells in peritoneal fluid--of endometrial origin? *Am J Obstet Gynecol,* 173**,** 566-70.

Van Kaam, K. J., Schouten, J. P., Nap, A. W., Dunselman, G. A. & Groothuis, P. G. 2008. Fibromuscular differentiation in deeply infiltrating endometriosis is a reaction of resident fibroblasts to the presence of ectopic endometrium. *Hum Reprod,* 23**,** 2692-700.

Velasco, I., Rueda, J. & Acien, P. 2006. Aromatase expression in endometriotic tissues and cell cultures of patients with endometriosis. *Molecular Human Reproduction,* 12**,** 377-81.

Vercellini, P., De Benedetti, F., Rossi, E., Colombo, A., Trespidi, L. & Crosignani, P. G. 1993. Tumor necrosis factor in plasma and peritoneal fluid of women with and without endometriosis. *Gynecol Obstet Invest,* 36**,** 39-41.

Vercellini, P., Somigliana, E., Vigano, P., Abbiati, A., Barbara, G. & Crosignani, P. G. 2009. Surgery for endometriosis-associated infertility: a pragmatic approach. *Hum Reprod,* 24**,** 254-69.

Vessey, M. P., Villardmackintosh, L. & Painter, R. 1993. Epidemiology of Endometriosis in Women Attending Family-Planning Clinics. *British Medical Journal,* 306**,** 182-184.

Vidal, J. D., Vandevoort, C. A., Marcus, C. B., Lazarewicz, N. R. & Conley, A. J. 2006. In vitro exposure to environmental tobacco smoke induces CYP1B1 expression in human luteinized granulosa cells. *Reprod Toxicol,* 22**,** 731-7.

Vigano, P., Somigliana, E., Di Blasio, A. M., Cozzolino, S., Candiani, M. & Vignali, M. 2001. Suppression of natural killer cell function and production of soluble ICAM-1: endometrial stroma versus melanoma. *Am J Reprod Immunol,* 46**,** 342-8.

Vigano, P., Somigliana, E., Gaffuri, B., Santorsola, R., Busacca, M. & Vignali, M. 2000. Endometrial release of soluble intercellular adhesion molecule 1 and endometriosis: relationship to the extent of the disease. *Obstet Gynecol,* 95**,** 115-8.

Visscher, P. M., Brown, M. A., Mccarthy, M. I. & Yang, J. 2012. Five years of GWAS discovery. *Am J Hum Genet,* 90**,** 7-24.

Viswanathan, S., Unlu, M. & Minden, J. S. 2006. Two-dimensional difference gel electrophoresis. *Nat Protoc,* 1**,** 1351-8.

Vodolazkaia, A., El-Aalamat, Y., Popovic, D., Mihalyi, A., Bossuyt, X., Kyama, C. M., Fassbender, A., Bokor, A., Schols, D., Huskens, D., Meuleman, C., Peeraer, K., Tomassetti, C., Gevaert, O., Waelkens, E., Kasran, A., De Moor, B. & D'hooghe, T. M. 2012. Evaluation of a panel of 28 biomarkers for the non-invasive diagnosis of endometriosis. *Human Reproduction,* 27**,** 2698-2711.

Vodolazkaia, A., Fassbender, A., Kyama, C. M., Bokor, A., Clerinx, P., Gevaert, O., Schols, D., Huskens, D., Meuleman, C., Peeraer, K., Tomassetti, C., De Moor, B. & D'hooghe, T. M. 2010. Evaluation of 28 different plasma biomarkers for potential non-invasive diagnosis of endometriosis. *Human Reproduction,* 25**,** I9-I10.

Vodolazkaia, A., Fassbender, A., Kyama, C. M., Bokor, A., Schols, D., Huskens, D., Meuleman, C., Peeraer, K., Tomassetti, C. & D'hooghe, T. M. 2011. Evaluation of VEGF, sVEGFR-1, sVEGFR-2 and PLGF plasma levels: potential biomarkers in endometriosis? *Human Reproduction,* 26**,** I215-I216.

Wang, L., Zheng, W., Ding, X. Y., Yu, J. K., Jiang, W. Z. & Zhang, S. Z. 2010. Identification biomarkers of eutopic endometrium in endometriosis using artificial neural networks and protein fingerprinting. *Fertil Steril,* 93**,** 2460-2.

Wiegerinck, M. A., Van Dop, P. A. & Brosens, I. A. 1993. The staging of peritoneal endometriosis by the type of active lesion in addition to the revised American Fertility Society classification. *Fertil Steril,* 60**,** 461-4.

Williams, R. S., Rao, C. V. & Yussman, M. A. 1988. Interference in the measurement of CA-125 in peritoneal fluid. *Fertil Steril,* 49**,** 547-50.

Williams, T. J. & Pratt, J. H. 1977. Endometriosis in 1,000 consecutive celiotomies: incidence and management. *Am J Obstet Gynecol,* 129**,** 245-50.

Worley, M. J., Welch, W. R., Berkowitz, R. S. & Ng, S. W. 2013. Endometriosis-associated ovarian cancer: a review of pathogenesis. *Int J Mol Sci,* 14**,** 5367-79.

Wu, M. H., Yang, B. C., Hsu, C. C., Lee, Y. C. & Huang, K. E. 1998. The expression of soluble intercellular adhesion molecule-1 in endometriosis. *Fertil Steril,* 70**,** 1139-42.

Wu, M. Y. & Ho, H. N. 2003. The role of cytokines in endometriosis. *Am J Reprod Immunol,* 49**,** 285-96.

Xavier, P., Belo, L., Beires, J., Rebelo, I., Martinez-De-Oliveira, J., Lunet, N. & Barros, H. 2006. Serum levels of VEGF and TNF-alpha and their association with C-reactive protein in patients with endometriosis. *Arch Gynecol Obstet,* 273**,** 227-31.

Yap, C., Furness, S. & Farquhar, C. 2004. Pre and post operative medical therapy for endometriosis surgery. *Cochrane Database Syst Rev***,** Cd003678.

Yates, J. R., Ruse, C. I. & Nakorchevsky, A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng,* 11**,** 49-79.

Zondervan, K., Cardon, L., Desrosiers, R., Hyde, D., Kemnitz, J., Mansfield, K., Roberts, J., Scheffler, J., Weeks, D. E. & Kennedy, S. 2002. The genetic epidemiology of spontaneous endometriosis in the rhesus monkey. *Ann N Y Acad Sci,* 955**,** 233-8; discussion 293-5, 396-406.

Zubarev, R. A. & Makarov, A. 2013. Orbitrap mass spectrometry. *Anal Chem,* 85**,** 5288-96.