Endometriosis biomarkers: discovery experiment and contextualization

Proteomic profiling of human stromal and epithelial endometriotic cell lines compared to normal endometrial cell lines, and characterization of estradiol and lipoxin A4 treatment.

Thesis supervisors: Dr. Geraldine Canny and Dr. Manfredo Quadroni

External Experts: Prof. Jean-Charles Sanchez and Dr. Patricia Palagi

Jonathan Sobel
23/12/2011
# Table of contents

**Index of Figures** .................................................................................................................. 3  
**Index of Tables** ...................................................................................................................... 4  
**Abstract** .................................................................................................................................. 5  
**Introduction** ............................................................................................................................. 6  
Pathogenesis of Endometriosis ........................................................................................................ 6  
Role of estrogens, estrogen receptors and estrogen production .................................................... 8  
Currently used and potential therapies ......................................................................................... 11  
Lipoxins: description, production and potential targets ................................................................. 13  
Cell lines used in this study ........................................................................................................... 15  
**Proteomics** ............................................................................................................................... 16  
Nano LC-MS/MS orbitrap .............................................................................................................. 16  
Quantitative proteomics ............................................................................................................... 17  
Biomarker validation ..................................................................................................................... 18  
Bioinformatics ............................................................................................................................. 19  
Data analysis ............................................................................................................................... 19  
Contextualization......................................................................................................................... 20  
Aim of the project .......................................................................................................................... 20  
**Materials and methods** ........................................................................................................... 21  
Experimental design ................................................................................................................... 21  
Main questions ............................................................................................................................. 21  
General work flow ....................................................................................................................... 22  
Cell culture ................................................................................................................................... 23  
Homogeneization ......................................................................................................................... 23  
Protein quantification .................................................................................................................... 23  
Separation and fractionation .......................................................................................................... 23  
In gel digestion ............................................................................................................................. 23  
 Extraction ..................................................................................................................................... 24  
LC-MS orbitrap .............................................................................................................................. 24  
Protein identification and validation .............................................................................................. 25  
Statistical analysis ......................................................................................................................... 25  
Enrichment analysis ....................................................................................................................... 26  
Data visualization .......................................................................................................................... 26  
**Results** .................................................................................................................................... 27
Mascot identification and Scaffold validation ................................................................. 27
Quality control .............................................................................................................. 29
Differences between endometriotic and normal endometrial cell lines ....................... 30
Effects of treatment .................................................................................................... 33
Known Biomarkers ..................................................................................................... 41
Contextualization ....................................................................................................... 44
Synergetic or suppressive effect .................................................................................. 48
Discussion .................................................................................................................. 51
Effects of treatments ................................................................................................. 51
Contextualization ....................................................................................................... 53
Conclusions and perspectives .................................................................................... 54
Acknowledgments ..................................................................................................... 55
Bibliography ................................................................................................................ 55
Annexes ...................................................................................................................... 63
Protein index ................................................................................................................ 63

Index of Figures
Figure 1: (A) retrograde menstruation, (B) the formation of the early lesions, (C) the late endometriosis lesions. ....................................................................................... 6
Figure 2 : Estrogen receptors (ER) ........................................................................... 8
Figure 3 : Estradiol biosynthesis ............................................................................... 9
Figure 4 : Sources of estradiol in endometriosis ....................................................... 10
Figure 5 : Pathways involved in proliferation and inflammation in endometriosis .......... 10
Figure 6 : Lipid mediators and the inflammatory response ........................................ 13
Figure 7 : Lipoxin-mediated actions on neutrophils and macrophages ...................... 13
Figure 8 : Role of lipoxins in the resolution of inflammation Taken directly from .......... 14
Figure 9 : LTQ-Orbitrap mass spectrometer ................................................................ 16
Figure 10 : Quantitative proteomics strategies ......................................................... 17
Figure 11 : Biomarker development process .............................................................. 18
Figure 12 : Experimental design ............................................................................... 21
Figure 13 : General work flow .................................................................................. 22
Figure 14 : Sequence coverage of COMT in the hTERT EEC LXA₄ sample and identified spectrum .... 27
Figure 15 : Quality control: A) Dendrogram of all cell lines and treatments, B) correlation matrix .... 29
Figure 16 : A) Volcano plot of 12Z control versus hTERT EEC control, B) Volcano plot of 22B control versus stromal control, C) Box plot of log ratio on protein expression data on 12Z control versus hTERT EEC control and 22B control versus stromal control ........................................................................... 30
Figure 17 : Pathway enrichment analysis of proteins differentially expressed between endometriotic and non endometriotic cell lines ..................................................................................... 31
Figure 18 : Boxplot and heatmap of treatment versus control log ratio expression ........ 34
Figure 19: volcano plots treatments vs ctrl for all cell lines, names of protein significantly
differentially expressed are displayed in red (p-value < 0.01) according to adjusted p-value with
Bonferroni (n=3) and Benjamini Hochberg correction for multiple comparisons. .......................... 35
Figure 20: Pathway enrichment analysis of proteins highly impacted by the different treatments ... 36
Figure 21: Impact of treatment on the potential biomarkers of endometriosis in the different cell
lines ........................................................................................................................................... 37
Figure 22: Protein expression in all cell lines and for all treatment of genes retrieved from the ERE
database and responsive to estradiol in hTERT EEC cells (log ratio > log (2) (A) or log ratio < -log(2)
(B)). ........................................................................................................................................... 39
Figure 23: Heatmap of known biomarkers ...................................................................................... 41
Figure 24: Pathway annotation enrichment analysis of known biomarkers .................................. 43
Figure 25: Biomarker protein-protein interaction network .............................................................. 44
Figure 26: Contextualization of potential biomarkers: A) Epithelial cells, B) Stromal cells .......... 46
Figure 27: Contextualization of known biomarkers: A) Epithelial cells, B) Stromal cells .......... 47
Figure 28: Suppressive or synergic effect and linear model: A) 3D plot of all protein in the different
cell lines in function of the effect of E2, LXA4 and both, B) residues histogram and density distribution
................................................................................................................................................ 49

Index of Tables
Table 1: Contribution of valid biomarkers to clinical research ...................................................... 18
Table 2: Arrangement of the spectral count data in a two-way table .............................................. 25
Table 3: Mascot identification statistics ......................................................................................... 27
Table 4: Spectral count, sequence coverage and identification probability of COMT in each sample 28
Table 5: Uniprot ID of proteins significantly differentially expressed between endometriotic and non
endometriotic cell lines according to the Fischer exact test with Benjamini Hochberg correction ..... 31
Table 6: Uniprot ID of proteins significantly differentially expressed between treated and non treated
cell lines according to Fischer exact test (p-value<0.01) adjusted with Benjamini Hochberg method
and Bonferroni method .................................................................................................................. 36
Table 7: Proteins (UniProt ID) from ERE regulated genes responsive to estradiol in hTERT EEC cells 38
Table 8: Statistical analysis of estrogen responsive genes .............................................................. 40
Table 9: Protein network GO Term analysis ..................................................................................... 45
Table 10: Suppressive or synergic effect of estradiol and lipoxin A4 treatment on specific proteins in
each cell line .................................................................................................................................... 50
Abstract

Endometriosis is a common disease, affecting 10% of women of reproductive age for which there is no cure. The current treatment of endometriosis mainly involves surgery along with medical therapies including oral contraceptive pills, to suppress the resurgence of the disease. The diagnosis is only possible with laparoscopy, thus there is an unmet need for novel diagnostic biomarkers. This pathology is characterized by the presence of endometrial tissue outside the uterus (termed lesions), inflammation and proliferation. Endometriotic lesions are principally composed of two cell types, stromal cells and epithelial cells. The hormone 17-beta estradiol (E2) promotes lesion growth and is thought to play a major role in endometriosis pathology. Lipoxins are anti-inflammatory lipids, which mediate a variety of protective effects in inflammatory contexts, their role in endometrial tissue is largely unknown. As Lipoxin A4 (LXA4) decreases lesion size in a mouse model of endometriosis, it is of interest to further characterize the effect of this lipid on endometrial cells.

Employing a LC-MS Orbitrap shotgun proteomics approach, we compared the normal cell lines (hTERT EEC endometrial epithelial cell line and an endometrial stromal cell line) to the endometriotic cell lines (12Z epithelial cell line and the 22B stromal cell line) to discover new biomarkers and biological processes implicated in the disease, and evaluated the effect of treatment with E2, LXA4 and E2 in combination on each cell line.

We then contextualized the data using different enrichment annotation tools such as DAVID and Biocompendium and visualization tools such as Cytoscape.

Results: 1139 proteins were detected and quantified with unweighted spectral count in the different samples. These studies revealed 31 potential endometriosis biomarkers, which were significantly differentially expressed between normal and endometriotic cell lines, according to fisher exact test corrected with Benjamini Hochberg method. The expression of 24 known biomarkers was altered by the above-named treatments and 10 other proteins were significantly impacted by at least one treatment.

To assess the estrogen responsiveness of the different cell lines the expression of genes known to be E2-regulated was investigated and we observed that hTERT EEC cells were the most E2 responsive cells. We also evaluated whether the combination of both treatments exerted a suppressive or synergic effect on specific proteins in each cell line using a mathematical model. These data provide novel insights into signaling pathways activated in normal and endometriotic endometrial epithelial and stromal cells.
Introduction

Pathogenesis of Endometriosis

Endometriosis affects roughly 5-10% of the women of reproductive age (1). This gynecological inflammatory disease is estrogen dependant and is characterized by the presence of endometrium-like tissue outside the uterine cavity, mostly on the pelvic peritoneum and ovaries. The clinical symptoms are severe dysmenorrhoea, dyspareunia, dysuria, chronic pelvic pain and infertility, which are affecting patient’s quality of life (2, 3). Endometriosis associated infertility is characterized by altered folliculogenesis, reduced preovulatory steroidogenesis of granulosa cells and decreased fimbrial ovum capture capability. The main causes are still unclear but many hypotheses have been made regarding environmental causes, defects of the immune system, genetic and epigenetics causes, and diverse anatomical or biochemical uterine aberrations (4). The most well accepted theory is that pertaining to retrograde menstruation described by Sampson, which could explain how the lesions are formed, although this occurs in almost all women. This process is depicted in figure 1 (5).

![Figure 1: (A) retrograde menstruation, (B) the formation of the early lesions, (C) the late endometriosis lesions. The retrograde flux of endometrial tissue fragments into the peritoneal cavity through the fallopian tube. This endometrial tissue adheres to the peritoneal wall, the ovaries or other organs and proliferates, leading to inflammation, bleeding in the cavity and pain.](image)

A second hypothesis is based on coelomicmetaplasia, whereby the genesis of endometriotic lesions within the peritoneal cavity is due to the differentiation of mesothelial cells into endometriallike tissue (6). A third hypothesis proposes that veins or lymphatic vessels can transport menstrual tissue from the endometrial cavity to distant sites (5). Others speculate that circulating blood cells originating from bone marrow can differentiate into endometriotic tissue at various anatomic locations (7). Molecular defects or immunologic abnormalities likely underlie the implantation of refluxed endometrium onto the peritoneal surface. In eutopic endometrium, many molecular abnormalities are due to the activation of oncogenic pathways (8) or biosynthetic cascades favoring increased production of estrogen, cytokines, prostaglandins, and metalloproteinases (9-11). Implant survival is probably caused by a failure of the immune system to clear implants from the peritoneal surface (12).

There are three distinct clinical forms of endometriosis, notably peritoneal endometriosis, which is characterized by endometriotic implants on the pelvic peritoneal surface and ovaries, endometriomas, ovarian cysts lined by endometrioid mucosa, and finally the rectovaginal endometriotic nodule, a solid mass comprised of endometriotic tissue blended with adipose and fibromuscular tissue, localized between the rectum and the vagina (13).

Endometriotic lesions are principally composed of two cell types, stromal cells and epithelial cells, and are associated with chronic bleeding, inflammation, chronic pain and an increased risk of infertility (14). The inflammation process causes pain, reduces the function of uterine tubes,
decreases receptivity of the endometrium and increases the difficulty for the development of the oocyte and embryo (15). Endometriotic lesions physically blocking the fallopian tubes leads to infertility (14).

A link between ovarian endometriosis and distinct types of ovarian cancer is suggested by epidemiologic and laboratory data (16).

Ovulatory cycles have a drastic effect on the development and the persistence of endometriosis (17). The fact that endometriosis appears after menarche and vanishes after menopause clearly shows the link between endometriosis and ovulatory cycles (18). However, in rare cases rectovaginal nodules are still found in postmenopausal women, indicating that persistence is independent of ovarian estrogen secretion (19).

Endometriotic and eutopic endometrial tissue are quite different in terms of estrogen, prostaglandin and cytokines production (20). Candidate genes related to implantation failure, infertility, and progesterone resistance have been found by comparing endometrial tissue from a disease-free group and endometriosis patients (21). Inflammation, which is a known feature of endometriosis, is associated with an increase of prostaglandins, metalloproteinases, cytokines, and chemokines, notably interleukin-1β, interleukin-6 and tumor necrosis factor. These proteins might enhance the adhesion of endometrial-tissue fragments onto peritoneal surfaces, and proteolytic membrane metalloproteinases may then promote implantation. Granulocytes, natural killer cells and macrophages are attracted by monocyte chemoattractant protein 1, interleukin-8, and RANTES. The accumulation of these immune cells, cytokines and chemokines in endometriotic lesions is regulated through positive feedback loops (22).

Progesterone resistance is mainly due to loss of progesterone signaling in the endometrium, this is commonly observed in women with the disease. The level of progesterone receptors in endometriotic stromal cells are significantly decreased, leading to a loss of paracrine signaling. These women no longer respond to progestin therapy. The starting point involves a deficient methylation of estrogen receptor beta (ERbeta) promoter resulting in over expression of this receptor in stromal endometriotic cells. It is probable that the relative levels of ERalpha and ERbeta dictates estradiol regulation and progesterone receptor expression in endometrial tissue (23).
Role of estrogens, estrogen receptors and estrogen production

In endometriosis, inflammatory and immune responses, angiogenesis and apoptosis are involved in the survival of endometriotic tissue. These pathologic processes are driven by estrogen, prosataglandins and progesterone resistance (23).

The steroid hormone 17β-Estradiol (E2) is a naturally produced predominant sex hormone and the most potent estrogen. Estradiol has an important role in reproductive and sexual functioning (24) and affects other organs such as bones or brain (25, 26).

The link between high estrogen and prostaglandin levels and the induction of progesterone resistance has become a subject of considerable interest (27). The therapeutic targeting of aromatase in the estrogen biosynthetic pathway, of cyclooxygenase-2 (COX-2) in the prostaglandin pathway or of the progesterone receptor reduces pelvic pain. These three pathways are involved in the over expression of the nuclear receptors steroidogenic factor 1 (SF1) and estrogen receptor β (28).

STAR (steroidogenic acute regulatory protein) mediates the first step of estrogen biosynthesis, notably the entry of cytosolic cholesterol into the mitochondrion. Subsequent steps are driven by five proteins catalyzing six enzymatic steps (side-chain cleavage enzyme, 3β-hydroxysteroid dehydrogenase 2, 17-hydroxylase–17-20-lyase, aromatase, and 17β-hydroxysteroid dehydrogenase 1) converting cholesterol to biologically active estradiol. The key step, biotransformation of C19 steroids to estrogens, is catalyzed by aromatase. Estradiol biosynthesis is presented in figure 3. (30)
Estrogens are produced at three sites (figure 4). Estradiol secreted by the ovaries is delivered to endometriotic tissue through blood vessels. Follicular rupture leads to a leakage of a large quantity of estradiol directly onto pelvic lesions. Another major source of estrogen are adipose tissue and skin, where circulating androstenedione is converted to estrone by aromatase, and subsequently transformed into estradiol. It can then travel through blood circulation to endometriotic lesions. Cholesterol can be also converted locally into estradiol, because endometriotic tissue expresses the complete enzyme set for estrogen biosynthesis (30).
In summary, estrogens are responsible for survival and proliferation of endometriotic tissue. Prostaglandins and cytokines mediate pain, inflammation and infertility (28). Prostaglandin E2 recruits SF1 to the promoter of steroidogenic genes, which allows estradiol biosynthesis in endometriotic stromal cells. The inhibition of estradiol biosynthesis through aromatase targeting leads to a reduction of endometriotic lesions and pelvic pain (31).

---

**Figure 4:** Sources of estradiol in endometriosis Taken directly from (28).

**Figure 5:** Pathways involved in proliferation and inflammation in endometriosis (28)
Currently used and potential therapies

The usual treatment of endometriosis is surgical removal of endometriotic tissue combined with medical suppression of ovulation to target pain. Endometriosis associated infertility needs surgery as well, and eventually the use of assisted reproductive technology. Peritoneal implants can be removed by laser, electric current or resection, while ovarian endometriomas and rectovaginal nodules can only be treated by classical surgery (32).

Aromatase inhibitors are used as therapeutics for the management of endometriosis. Anastrozole is a non-steroidal aromatase inhibitor developed to treat breast cancer. Anastrozole reduces pelvic pain as well as lesion size. Anastrozole combined with oral contraceptives may be a good therapy for women who do not tolerate the classical treatments (33). Letrozole is also a non-steroidal competitive aromatase inhibitor which is similar to Anastrozole and is equally effective in reducing pelvic pain associated with endometriosis (34). These aromatase inhibitors have shown to be compatible with fertility restoration and assisted reproduction (35).

Nowadays, approved endometriosis pharmacotherapy mainly involves gonadotropin releasing hormone (GnRH) agonists, but they are associated with side effects including post-menopausal-like state with bone loss, which limits their use to six month (36).

Oral contraceptives, progestins or GnRH analogues, reduce pain and pelvic disease by the suppression of ovulation. This renders pregnancy impossible, because these hormone treatments inhibit ovulation and interfere with cyclic remodeling of the endometrium. GnRH antagonists block the GnRH receptor directly, leading to a down-regulation of pituary gland function and the suppression of ovarian steroid production (37). The GnRH antagonist has a higher efficiency than the GnRH agonist to reduce endometriosis symptoms (38).

Many non hormonal treatments can be used as therapeutics for endometriosis but unfortunately few large, well-conducted trials have been performed. For instance TNF-alpha inhibitors, such as TNF binding protein 1 (TBP1), and GnRH agonists, reduce the development of endometriosis in baboons (39). Matrix metalloprotease (MMP) inhibitors could be used as well, because they inhibit the release of pro inflammatory mediators. Progesterone inhibits MMPs and induces the expression of MMP inhibitors like Tissue inhibitor of metalloproteinases (TIMPs) (40). Anti-angiogenic agents, which inhibit vascularization, can be used to decrease the establishment and maintenance of new blood vessels in endometriotic lesions. The reduction of vascular endothelial growth factor (VEGF) levels can also be induced by anti-angiogenic agents or GnRH agonists (41). Indeed VEGF levels are elevated in the peritoneal fluid of endometriosis patient. Other factors influencing angiogenesis, such as IL-1, IL-6, TNF-α, TGF β and MMP2 are potential therapeutic targets. Anti-angiogenic agents like endostatin, TNP-470, celecoxib and rosaglitazone caused a reduction in lesion size in mouse models of endometriosis (42).

As the immune system is involved in the development of endometriosis, immunomodulators can be used as a therapy. Pentoxifylline is a multi-site immunomodulator, which reduces the inflammatory action of TNF-α and IL-1β on granulocytes. This immunomodulator does not influence ovulation, and could be used to treat infertility associated with endometriosis (32).
Oxidative stress is likely involved in impaired infertility in endometriosis patient, and the use of anti-oxidants could attenuate disease progression (43). Pro-oxidant and pro-inflammatory factors such as haemoglobin, heme and iron are released by erythrocytes into the peritoneal cavity and are implicated in the formation of reactive oxygen species. The stress response can be monitored by measuring levels of endothelial nitric oxide synthase, heat shock protein 27 and 70 (HSP27, HSP70) in the ectopic endometrium, and lysophosphatidilcholine concentration in the peritoneal fluid. An impaired expression of proteins participating in the defense against oxidative stress, such as manganese and copper/zinc superoxide dismutases and glutathione peroxidase, has been demonstrated in patient’s endometrium. The injection of superoxide dismutase and catalase into the peritoneal cavity prevented the formation of intraperitoneal adhesions in a rabbit model of endometriosis (44).

Lipoxins and resolvins are anti-inflammatory and pro-resolving lipid mediators, which may represent a potential therapeutic for endometriosis. These mediators will be discussed in more detail in the next section.
**Lipoxins: description, production and potential targets**

Lipoxins are **eicosanoids** which exert anti-inflammatory and pro-resolution effects in many inflammatory disorders, including periodontitis, arthritis, nephritis, tuberculosis, cystic fibrosis and toxoplasmosis (45-47). Lipoxins inhibit leukotriene and natural killer cell function, leukocyte migration, TNF-induced chemokine production, NFκB translocation, chemokine receptor and adhesion molecule expression and pathogen-triggered IL-12 production (48, 49). Lipoxins were traditionally thought to signal via two receptors, a G-coupled protein receptor known as FPR2/ALX and AhR, a nuclear receptor (50-52). Lipoxin A₄ has recently been shown to signal via estrogen receptor alpha in endometrial epithelial cells (53). The intracellular mechanisms of lipoxin action are not yet fully understood. The suppressor of cytokine signaling (SOCS) protein family may play an important role in lipoxin mediated-actions. These proteins regulate signal transduction pathways triggered by cytokines and hormone receptors (54). SOCS-1 and SOCS-3 are induced by IL-10 and modulate the IFN-gamma receptor downstream cascade. Lipoxin analogues induce renal SOCS-2 expression (55).

![Diagram](image1.png)

**Figure 6 : Lipid mediators and the inflammatory response (56)**

Figure 6 shows the development and resolution of acute inflammation. Microbial invaders, tissue injury and surgical trauma may activate the release and formation of arachidonate-derived prostaglandins, which regulate early events in the inflammatory response. Mediators involved in inflammation such as prostaglandins and leukotrienes are probably the same as those involved in chronic inflammation or the resolution of inflammation. Specialized chemical mediators participate actively in leukocyte responses required for resolution (56).

![Diagram](image2.png)

**Figure 7 : Lipoxin-mediated actions on neutrophils and macrophages (56)**
Polymorphonuclear neutrophils (PMNs) start to produce lipoxins and stop leukotriene biosynthesis. At inflammatory sites, prostaglandins initiate a number of inflammatory responses. The transcriptional regulation of 15-lipoxygenase (15-LO) in human neutrophils allows the temporal dissociation and transcellular production of lipoxins. Decreased PMNs in the exudate is also associated with a switch of the families of lipid mediators generated from eicosanoids to resolvins as well as protectins. Then apoptotic PMNs are phagocytosed by macrophages, which clear them from the site of inflammation, restoring tissue homeostasis (57).

Lipoxins have a dual action on PMNs and monocytes. Initial chemoattractants recruit PMNs that migrate from postcapillary venules. PMN recruitment is amplified by the production of a 5-LO-pathway product, Leukotriene B4 (LTB4), a potent chemoattractant. During ongoing inflammation, platelet-leukocyte interactions stimulate the formation of lipoxin A4 (LXA4) and lipoxin B4 (LXB4), represented in the Figure 8 by a red ×, by blocking the further recruitment of PMNs from the postcapillary venules. Lipoxins stimulate monocyte recruitment and increase the uptake of apoptotic PMN by macrophages. Lipoxins reduce dendritic cell motility and reduce IL-12, as well as regulating T cell-produced cytokines and MMPs from fibroblasts leading to the resolution of inflammation in vivo (58).

Figure 8: Role of lipoxins in the resolution of inflammation Taken directly from (56)
**Cell lines used in this study**

The four cell lines used are the hTERT EEC immortalized human endometrium epithelial cell line, 12Z human epithelial endometriotic cell line, human endometrium stromal cell line, and 22B human endometriotic cell line.

**hTERT EEC cells:**

Immortalized human primary endometrial epithelial cells (EECs) were isolated from a normal proliferative phase endometrium by stably transfecting the catalytic subunit (hTERT) of the human telomerase complex. hTERT-EECs have a polarized, non-invasive epithelial phenotype with apical microvilli and production of a basal lamina when grown on a three-dimensional collagen–fibroblast lattice. Employing atomic force microscopy, living hTERT-EECs were shown to produce extracellular matrix (ECM) components and ECM secretion is known to be modified by estrogen and progesterone (P4). hTERT-EECs express inducible and functional endogenous estrogen receptor-alpha (ER-alpha) as demonstrated by estrogen response element reporter assays and induction of P4 receptor (PR). Thus, hTERT-EECs may be regarded as a novel *in vitro* model to investigate the role of human EECs in steroid hormone-dependent normal physiology (59).

**Normal stromal cells:**

Primary stromal cells isolated from the endometrium in the proliferative phase, and were previously characterized by immunocytochemical analysis. These cells were provided by Dr.Nick Bersinger’s lab, Bern.

**22B and 12Z cells:**

Gene expression profiling and functional characterization of human immortalized endometriotic epithelial and stromal cells revealed that several genes associated with estrogen and progesterone biosynthesis and signaling, cell cycle regulation, extracellular matrix degradation, angiogenesis, cell growth, cytokines production and prostaglandin E2 biosynthesis, transport, and signaling, were expressed in these cell lines. These cells show a migrating and invading potential and produce a high quantity of prostaglandin E2. Human endometriotic cell lines are an ideal model to study molecular and cellular aspect of endometriosis (60).
Proteomics

Nano LC-MS/MS orbitrap

Shotgun analysis, a non-targeted approach, is widely used in proteomics studies. This type of analysis needs high-performance instrumentation, such as the LTQ-Orbitrap mass spectrometer, which belongs to the high resolution Fourier transform mass spectrometer class (FTMS). This type of mass spectrometer has various application including metabolomic or clinical proteomics (61). The Orbitrap abilities include the detection with a high mass accuracy of less than 2 parts per million using internal standards and 5 ppm with external calibration. Its resolving power exceeds 100'000 FWHM at 400 m/z, with a wide dynamic range, high sensitivity and a fast duty cycle. This MS device can analyze complex peptide mixtures in great depth and with high confidence. LTQ-Orbitrap offers various fragmentation methods including collision induced dissociation (CID), pulsed-Q-dissociation (PQD) and higher energy collision dissociation (HCD).

Figure 9: LTQ-Orbitrap mass spectrometer

Tandem MS analysis involves the detection of a precursor ion by a first MS scan and fragments generated by CID are detected in data-dependent acquisition (62). The combined information from MS and MS/MS scans is used to search tandem spectra against a protein sequence database using search algorithms which will be discussed in the following bioinformatics section.

Reversed-phase high performance liquid chromatography is very popular because it can be easily coupled with electrospray and nano-electrospray ionization techniques to analyze complex peptide mixtures derived from protein samples (63). Sample complexity is usually reduced prior to LC-MS analyses by separation methods at the protein or peptide level, such as polyacrylamide gel electrophoresis, isoelectric focusing, 2D-gel or strong cation exchange chromatography (SCX) (64).
Quantitative proteomics

Mass spectrometry allows qualitative and quantitative analysis of complex mixtures with label or label-free methods. Label strategies are more accurate, but unfortunately more expensive, and the sample preparation is more difficult (65). Quantitative proteomic analysis strategies are briefly described in this section.

Spectral counting and peak intensity based approach are used as label free MS quantification in proteomic analysis (66). The number of peptide MS/MS spectra serves as an estimator of the peptide relative abundance for spectral counting, but it is less accurate and reproducible than evaluating peak intensity. The intensity of MS peaks can also serve as an estimator of the peptide relative concentration in a sample, but is computationally more difficult. Spectral counting sensitivity can be increased using mass analyzer with high MS/MS sampling rate, such as linear ion trap, and improvement in the accuracy of quantification by ion intensities can be obtained using a high resolution and high accuracy mass analyzers, such as Orbitrap instruments. In this study we used spectral counting as a label free method for protein quantification.

Several stable isotope labeling approaches have been developed during the past decades. These include Isotope-Coded Affinity Tag (ICAT) (67), Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) (68), 15N/ 14N metabolic labeling (69), 18O/ 16O enzymatic labeling (70), Isotope Coded Protein Labeling (ICPL) (71), Tandem Mass Tags (TMT), Isobaric Tags for Relative and Absolute Quantification (iTRAQ) among others (72). Unfortunately, most labeling-based quantification approaches have potential limitations, including increased time and complexity of sample preparation, high reagent cost, incomplete labeling and the requirement for specific quantification software (73). In contrast, TMT and iTRAQ allow the comparison of multiple (up to 8) samples in the same experiment. Isotope stable labeling approaches can be used in vivo or in vitro on peptides or proteins.
Biomarker validation

Biological markers have been defined as cellular, biochemical or molecular alterations that are measurable in biological media such as human tissue, cells or fluids. This definition is extended to include biological characteristics, which can be objectively measured, and used to monitor biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention. Biomarkers include tools and technologies that help to predict as well as to understand the cause, the diagnosis, the progression, the regression or the effect of the treatment on a specific disease. Their uses are summarized in table 1. They are used to reduce the misclassification of disease and exposure, opening a window to potential mechanisms related to disease pathogenesis (74, 75).

**Table 1 : Contribution of valid biomarkers to clinical research, taken directly from (75)**

| Delineation of events between exposure and disease |
| Establishment of dose-response |
| Identification of early events in the natural history |
| Identification of mechanisms by which exposure and disease are related |
| Reduction in misclassification of exposure or risk factor and disease |
| Establishment of variability and effect modification |
| Enhanced individual and group risk assessments |

The phases involved in biomarker development are discovery, characterization and clinical qualification or validation (76).

**Figure 11 : Biomarker development process (76)**

In this study we used cell lines to identify novel biomarkers and follow known and candidate biomarkers to monitor the treatment effect of estrogen and lipoxin A4 on normal and endometriotic epithelial and stromal endometrial cells. These biomarkers can be used to develop a diagnosis test if they are detectable in urine or blood.
Bioinformatics

Data analysis
In recent years MS based analysis has become a powerful technology to carry out large-scale protein analysis. The combination of innovative experimental strategies and advances in computational methods enable the global study of cellular proteomes. Manual mapping of complex proteomics data to biological process is impossible and computer-aided data analysis is essential for further progress in the field (77). Bioinformatics is defined as a means for functional analysis and data mining of data sets leading to biologically interpretable results and insights. Bioinformatics covers large-scale data analysis, peptide identification and quantification, databases, tools and software development and usage.

In a typical proteomic MS-based workflow, the starting point is a protein sample usually digested with trypsin. Peptides resulting from the digestion of the protein mixture are then separated using chromatography and loaded directly in the mass spectrometer. The peptides are isolated and fragmented by the mass spectrometer and a second stage of mass analysis is used to acquire a MS/MS spectrum. For each spectrum software is used to determine which peptide sequence, in a database of in silico digested proteins, gives the best match. Several algorithms have been developed to score the matches between the spectrum and the peptide sequence (78). The search engine of choice is MASCOT (79). The result of the identification is then validated using Scaffold, which attempts to increase the confidence in protein identification reports, through the use of several statistical methods. This includes the re-implementation of the Peptide Prophet algorithm, which converts the MASCOT score into probabilities of peptide identification. Scaffold also allows the combination of the score from different search engines to increase sensitivity and accuracy of peptide probability (80).

After protein identification, R statistical software can be used to analyze the expression of the protein according to the peptide quantitative data (here spectral counting). Between two conditions the log ratio is computed to estimate if the expression of a specific protein is changing. At that point several statistical tests can be used depending on the number of replicates and conditions. The Fischer’s exact test, G-test and AC test can be used when the number of replicates is limited to one or two, whereas the t-test is adapted with three or more replicates (81). Adjusted p-values with Bonferroni correction or Benjamini & Hochberg are used to reduce the over prediction from multiple testing leading to a higher false positive rate. Clustering methods as hierarchical clustering or k-means clustering are then useful to identify expression patterns, which can be enriched with existing functional annotations (78). Many tools have been developed to provide Gene ontology (GO), pathway (KEGG, Biocarta, Panther, Reactome), disease, protein family and domain, protein-protein interactions, transcription factor or tissue specific annotations. David annotation tool (82), biocompendium or CTD (83) are often used to perform enrichment analysis. Another interesting computational method called Gene Set Enrichment Analysis (GSEA), determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states according to the expression profile (84). One last commercial tool developed by Metacore, called GeneGO, is useful for statistical and enrichment analysis including pathway map layouts and protein network visualization (85).
Contextualization

To understand the role of a specific protein, it is important to look at the targeted protein in the context of a cellular system. Each interaction can provide a valuable clue to the function of the protein. Many resources allow the retrieval of protein-protein, protein-small molecules and DNA/RNA-protein molecular interactions, based on high throughput assays, such as yeast two hybrid system or tandem affinity purification. Important efforts are made by the PSIMEx consortium to unify the molecular interactions format with a standard format and to allow the query of the different databases (IntAct DIP, MINT, MatrixDB, MPIDB, I2D, Innate DB, and Molecular Connections) with a single tool known as PSICQUIC (86, 87). These databases are used to reconstitute the protein network and identify cluster of proteins, protein hub, and protein complex, to understand their function with enrichment analysis, such as GO term or pathway. GeneMANIA is based on the principle guilt by association and it can also be used to build networks from a list of genes or proteins (88). These manually curated protein networks or pathways can be used to visualize the expression level of each detected protein in their context using visualization and analysis tools (Plug-in), such as GeneMAPP, Cytoscape (89), Celldesigner (90) or GeneGO (85). These networks are also employed to develop models by system biologists to understand the kinetics of these cellular processes.

Aim of the project

Using a proteomic approach, the aims of this project are to:

1) Monitor existing biomarkers of endometriosis.
2) Find new candidate biomarkers.
3) Understand molecular processes involved in the disease.
4) Compare normal cell lines versus endometriotic cell lines.
5) Understand the role of epithelial cells versus stromal cells.
6) Understand which signaling pathways are activated or inhibited upon lipoxin A4 and Estradiol treatment.
7) Investigate a suppressive or synergic effect of the combination of both treatments.
**Materials and methods**

**Experimental design**

Four cell lines (stromal non endometriotic, 22B stromal endometriotic, hTERT epithelial non endometriotic and epithelial 12Z endometriotic) treated with lipoxin A₄, estradiol, estradiol + lipoxin A₄ in combination and vehicle control were used. The experimental design is presented below in the figure 12.

**Main questions**

The selected experimental design is useful for answering several fundamental questions. This experiment will give new insights into the role of stromal cells compared to epithelial cells in normal and endometriotic situations. We will be able to compare the normal cell lines versus the endometriotic cell lines to discover new biomarkers and processes implicated in the disease, and to evaluate the effect of E₂, lipoxin A₄ and lipoxin A₄ with E₂ on each cell line. Suppressive or Synergic effect of the combination of both treatments on specific protein, known biomarkers, new candidate biomarkers in each cell line, will be evaluated.

![Experimental design](image-url)
General work flow

Sample preparation
- Homogenisation, quantification
- Separation, fractionation
- Digestion
- Extraction

LC-MS orbitrap
- Reverse phase liquid chromatography
- Tandem MS
- High resolution LTQ-Orbitrap

Peptide or protein identification
- Mascot search engine for peptide identification
- Scaffold software for validation

Statistical analysis R
- Quality control, correlation, dendrogram
- Log ratio and fisher's exact test, volcano plot, box plot
- Clustering K-means/hierarchical, heat map

Functional analysis
- Previously annotated biomarkers (NCBI)
- Enrichment analysis: Gene Ontology analysis, pathway analysis, other annotation analysis (disease, drug interaction, protein interaction, chemical interaction, transcription factor, estrogen regulated genes ERE)

Network analysis
- Data visualisation: R, Cytoscape
- Topology analysis, modeling

Figure 13: General work flow
**Cell culture**

6 \( \times 10^5 \) **12Z** epithelial endometriotic cells passage 65, 5 \( \times 10^5 \) **hTERT** epithelial non endometriotic cells passage 48, 1.1 \( \times 10^5 \) **2B** stromal endometriotic cells passage 22 and 3.75 \( \times 10^5 \) **Stromal** non endometriotic cells passage 3 were plated in DMEM F12 medium containing 10% FBS and 1% Penicillin Streptomycin in 100mm dishes for each treatment (ctrl, E2, LXA$_4$, LXA$_4$ + E2). The following day the cells were washed with pre warmed PBS, and were cultured for 24h in white DMEM with 10% charcoal striped FBS, 1% Penicillin Streptomycin and 1% L-glutamine. The cells were treated with 100nM LXA$_4$ and/or 10 nM E2 or vehicle for 24h.

**Homogeneization**

The medium was discarded and cells were scraped into 2ml PBS. The cells were then transferred into a falcon tube and centrifuged at 1500g for 3 min. The pellet was washed twice with PBS and resuspended in 100ul FASP buffer (4% SDS, 100mM Tris/HCl ph 7.5 , 100mM DTT). Samples were sonicated twice for 20s and then heated at 95°C for 5 min to solubilize the proteins before being allowed to cool down to room temperature for 5 min. Samples were then centrifuged for 10 min at 13000 rpm and the supernatants were transferred into clean eppendorf tubes and stored at -20°C.

**Protein quantification**

A 12% polyacrylamide gel was prepared for each cell line (10ml Acrylamide 30%, 8.6ml ddWater, 6.25ml 1.5M Tris pH 8.8, 125 ul APS 10% 12ul TEMED ; Stacking :1.3ml Acrylamide 30%, 6.1ml ddWater, 2.5 ml 0.5M Tris pH 6.8, 50 ul APS 10%, 10 ul TEMED). 7ul of molecular weight and 2 ul of each sample with 20 ul SDS buffer 2X were loaded on the gel, together with 5ug, 10 ug and 15ug of standard Jurkat cell extract, for calculating the approximate concentration of proteins in the samples. The gel was run at 80V for 30 min and 120V for 1h. After migration, the gel was fixed 10 min in a solution of 10% acetic acid and 50% EtOH. The gel was then washed in ddwater for 10min. The gel was stained for 2h using Candiano Coomassie Staining. The gel was then scanned and the concentration of total proteins in each sample was calculated by densitometry by linear regression relative to the standard, using the total OD for individual lanes.

**Separation and fractionation**

A second 12% polyacrylamide gel was prepared for each cell line. 7ul molecular weight and 15ug of total protein for each sample was loaded onto a gel, each sample being separated by an empty lane. The gel was migrated on 2cm length only and then stained using Candiano Coomassie Staining (as described above). Each sample lane was cut into 6 fractions and transferred into 1.5ml clean Eppendorf tubes.

**In gel digestion**

The gel pieces were washed with 200 ul of 50 mM ammonium bicarbonate with 30% acetonitrile and incubated 20 min at room temperature on a shaker. The supernatant was then discarded and 150 ul of 10mM DTT in 25mM ammonium bicarbonate was added to each tube and incubated for 45 min at 56 °C with occasional shaking. The tubes were then cooled to room temperature and the liquid was discarded. 150 ul of 100mM iodoacetamide in 50mM ammonium bicarbonate was added to each tube, and incubated for 45 min at room temperature in the dark with occasional shaking. The liquid was removed, and another wash with 200ul of 50mM ammonium bicarbonate was performed for 10 min incubation at room temperature on shaker. The liquid was again removed and the gel was dehydrated with 100ul acetonitrile 100% for 10 min on a shaker. An additional incubation step was
made with 200 ul of 50 mM ammonium bicarbonate with 30% acetonitrile for 20 min at room temperature on a shaker, followed by a second dehydratation step with 100ul acetonitrile 100%. Then gel pieces were dried and 100ul of trypsin solution (6.25ng/ul) was added into the tubes, which were then kept for 30 min on ice. The digestion was carried out overnight at 37°C.

**Extraction**
The sample solution was transferred into a clean tube using a loading tip kept for the whole extraction process (one different tip for each tube). 100ul of 10% formic acid was added to the gel pieces, and incubated 20min at room temperature on a shaker. The liquid was then transferred into the corresponding tube. Two further extractions of gel pieces using 100ul of 100% acetonitrile and an incubation for 20min at room temperature on a shaker were performed. The solutions were then pooled with previous extracts. The tubes were frozen in liquid nitrogen and dried in a speed vacuum device.

**LC-MS orbitrap**
The LC-MS/MS analyses were carried out by the Protein Analysis Facility (PAF) of the University of Lausanne. Samples were analyzed on a hybrid linear trap LTQ-Orbitrap XL mass spectrometer (Thermo Fisher, Bremen, Germany) interfaced via a TriVersa Nanomate (Advion Biosciences, Norwich, UK) to a Agilent 1100 nano HPLC system (Agilent Technologies, Waldbronn, Germany). Solvents used for the mobile phase were 95:5 H₂O:acetonitrile (v/v) with 0.1 % formic acid (A) and 5:95 H₂O:acetonitrile (v/v) with 0.1 % formic acid (B).

Peptides were loaded onto a trapping microcolumn ZORBAX 300SB C18 (5 mm x 300 μm ID, 5 μm, Agilent) in H₂O:acetonitrile 97:3 (v/v) + 0.1 % formic acid at a flow rate of 10 μl/min. After 5 min, they were back-flush eluted and separated on a reversed-phase nanocolumn ZORBAX 300SB C18 column (75 μm ID x 15 cm, 3.5 μm, Agilent) at a flow rate of 300 nl/min with a gradient from 5 to 85 % acetonitrile in 0.1% formic acid: 5 min at 0 % of solvent B, from 0 to 25 % of B in 35 min, 25 to 50 % B in 15 min, 50 to 90 % in 5 min, 90 % B during 10 min, 90 to 0 % in 5 min and 15 min at 0 % (total time: 90 min).

For spraying, a 400 nozzle ESI Chip (Advion Biosciences) with a voltage of 1.65 kV was used, and the mass spectrometer capillary transfer temperature was set at 200°C. In data-dependent acquisition controlled by Xcalibur 2.0.7 software (Thermo Fisher), the six most intense precursor ions detected in the full MS survey performed in the Orbitrap (range 350-1500 m/z, resolution 60000 at m/z 400) were selected and fragmented. MS/MS was triggered by a minimum signal threshold of 10000 counts, carried out at relative collision energy of 35 % and with an isolation width of 4.0 amu. Only precursors with a charge higher than one were selected for CID fragmentation and fragment ions were analyzed in the LTQ linear trap. The m/z of fragmented precursors was then dynamically excluded, with a tolerance of 0.01 amu, from any selection for a period of 120 s. From raw files, MS/MS spectra were de-isotoped and exported as mgf files (Mascot Generic File, text format) using MascotDistiller 2.3.2 (Matrix Science, London, UK).
Protein identification and validation

DATABASE SEARCHING: All MS/MS samples were analyzed using Mascot 2.3 (Matrix Science, London, UK). Mascot was set up to search the Swiss-Prot database (www.uniprot.org) restricted to human taxonomy (database release used was 2011_03, 20'234 sequences after taxonomy filter) and a custom-built database containing the sequences of usual contaminants (enzymes, keratins). Trypsin (cleavage at K,R, not before P) was used as the enzyme definition. Mascot was searched with a fragment ion mass tolerance of 0.50 Da and a parent ion tolerance of 10 ppm. Iodoacetamide derivative of cysteine was specified in Mascot as a fixed modification. N-terminal acetylation of protein, deamidation of asparagine and glutamine, and oxidation of methionine were specified as variable modifications.

CRITERIA FOR PROTEIN IDENTIFICATION: Scaffold (version Scaffold_3, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identification. Peptide identifications were accepted if they could be established at greater than 95.0 % probability as specified by the Peptide Prophet algorithm (91). Protein identifications were accepted if they could be established at greater than 99.0 % probability and contained at least one identified peptide. Protein probabilities were assigned using the Protein Prophet algorithm (92). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

Statistical analysis

As we did not use biological replicates (one sample per condition), we decided to use the Fischer's exact test to determine the significantly differentially expressed proteins (81). This test doesn’t necessitate data normalization and assumes that the row total and the column total are fixed in a two-way table (see table 2). Therefore, any entry of this table completely determines the others. The Fischer’s exact test assumes a hyper geometric distribution for the expected spectral counting.

<table>
<thead>
<tr>
<th>Table 2 : Arrangement of the spectral count data in a two-way table</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Spectral count for target protein X</td>
</tr>
<tr>
<td>Spectral count for any other protein</td>
</tr>
<tr>
<td>Total spectral count across all proteins</td>
</tr>
</tbody>
</table>

\[
p = \frac{\binom{x_1+x_2}{x_2} \binom{y_1+y_2}{y_2}}{\binom{n}{y_2}}
\]

Hyper geometric distribution

Given a target protein X, the spectral counts from a pairwise experiment are arranged in a two way table. The total spectral count for a specific condition defines the sampling number for this experiment. Then X is differentially expressed under the two conditions if the difference between these two conditions is statistically significant. As the Fisher’s exact test does not require replicates, the spectral count data is pooled to mimic replicated experiments. Consequently, we are comparing the spectral count in condition 1 and 2 for one targeted protein to the sum of spectral counts in a sample in condition 1 and 2. As more than 1000 proteins are tested simultaneously in our study we use the Benjamini and Horchberg correction for multiple testing adjustment to control the false discovery rate. A protein with a p-value of p is considered differentially expressed if
\[ p \leq \left( \frac{j}{t} \right)^q \]

where \( t \) is the total number of proteins tested and \( j \) is the rank of the targeted protein in the list of \( t \) proteins sorted in increasing order by their p-value, and \( q \) is the desired false discovery rate. Another correction can be done for the multiple testing between the different conditions (treatments versus control). The Bonferroni correction is the most widely used and simply involves the division of \( \alpha \) (significance level) by the number of comparisons. All calculations were performed using R statistical computing software (http://www.r-project.org/).

The un-weighted spectral count was used as a quantitative value. This spectral count includes spectra which are not specific to a unique protein. This means that spectra belonging to several proteins are counted for each protein. This can produce artifacts with exaggerated p-value for some proteins. The log ratios were calculated with the following formula:

\[ \text{logratio} = \log \left( \frac{\text{Spectral count}_{\text{cond1}} + 1}{\text{Spectral count}_{\text{cond2}} + 1} \right) \]

This calculation avoids NaN (not a number) and Inf (infinite) due to null spectral count in the denominator, but add an error for the proteins of low spectral count.

**Enrichment analysis**

We wished to monitor known biomarkers of endometriosis, and specific proteins which are implicated in the biological process underlying the disease, such as inflammation, adhesion, proliferation, apoptosis and estrogen metabolism. In order to find known endometriosis biomarkers, three main sources were used: the EAGLi text mining tool (Engine for question-Answering in Genomics Literature, http://eagl.unige.ch/EAGLi/), the NCBI annotated gene list for endometriosis, and a gene list provided by the Mucosal Immunity research group. The significantly differentially expressed genes according to Fisher’s exact test were annotated using the DAVID tool (The Database for Annotation, Visualization and Integrated Discovery, http://david.abcc.ncifcrf.gov/) for gene ontology and pathway annotations, and the Biocompendium tool for transcription factor binding profile and pathway annotations (http://biocompendium.embl.de/). A database for estrogen responsive element (ERE) from Genome-Wide Identification of High-Affinity in Human (http://www.mapageweb.umontreal.ca/maders/eredatabase/) was also used for the enrichment analyses (93).

**Data visualization**

Finally, the data were explored and visualized using several tools and software. R statistical computing software provides many tools for assessing the quality of the data such as dendrograms and correlation heat maps (for all detected proteins). Data exploration was performed using hierarchical clustered heat maps of log ratio expression values from different conditions (e.g. treatment vs ctrl), box plots of log ratio values, histograms, and volcano plots (log ratio versus – (log p-values)). Protein networks for interesting candidates were then built with PSICQUIC or GENEMANIA to understand how these proteins were working in their context (87, 88). These protein networks were visualized and analyzed using Cytoscape (94), including Reactome and KEGG pathways of interest.
Results

Mascot identification and Scaffold validation

In order to obtain reliable results, stringent parameters for protein identification were set. The protein identification probability was set above 99% and the peptide identification probability was set above 95% based Peptide Prophet algorithms \(^{(91)}\). We wanted at least 3 spectra assigned to one protein in at least one sample to consider the protein for quantitative comparisons. Table 3 shows the statistics relevant to the identification for each sample.

Table 3: Mascot identification statistics

<table>
<thead>
<tr>
<th>Category</th>
<th>Dn Sample</th>
<th>PLS Sample</th>
<th>#Prot</th>
<th>#IDs</th>
<th>#Spec</th>
<th>%Ids</th>
</tr>
</thead>
<tbody>
<tr>
<td>122-E2</td>
<td>4H4-E2</td>
<td>mascot_danem_merge (F00 5063)</td>
<td>924</td>
<td>4717</td>
<td>1503</td>
<td>28%</td>
</tr>
<tr>
<td>122-LXaE2</td>
<td>H17-LXaE2</td>
<td>mascot_danem_merge (F00 5063)</td>
<td>525</td>
<td>4975</td>
<td>17545</td>
<td>28%</td>
</tr>
<tr>
<td>122-LXaH2</td>
<td>4H4-LXaH2</td>
<td>mascot_danem_merge (F00 5063)</td>
<td>938</td>
<td>5697</td>
<td>18329</td>
<td>30%</td>
</tr>
<tr>
<td>122-cfl</td>
<td>4H4-cfl</td>
<td>mascot_danem_merge (F00 5063)</td>
<td>944</td>
<td>5204</td>
<td>19124</td>
<td>25%</td>
</tr>
<tr>
<td>22B-Clfl</td>
<td>2B-Clfl</td>
<td>mascot_danem_merge (F00 5715)</td>
<td>1006</td>
<td>6330</td>
<td>20796</td>
<td>32%</td>
</tr>
<tr>
<td>22B-E2</td>
<td>2B-E2</td>
<td>mascot_danem_merge (F00 5715)</td>
<td>998</td>
<td>6493</td>
<td>18128</td>
<td>38%</td>
</tr>
<tr>
<td>22E-LXaA</td>
<td>2B-LXaA</td>
<td>mascot_danem_merge (F00 2715)</td>
<td>964</td>
<td>5699</td>
<td>17942</td>
<td>35%</td>
</tr>
</tbody>
</table>

The total number of identified proteins irrespective of the occurrence across every sample is 1139 proteins with 91144 spectra analyzed and approximately 30% of spectra were positively identified.

Catechol O-methyltransferase identification

Catechol O-methyltransferase (COMT), an enzyme involved steroid hormone metabolism, has three different identified peptides in samples generated from LXA4-treated hTERT EEC cells, with sequence coverage of 26% and 6 spectra assigned. The identification of a spectrum is based on the tandem
MS/MS analysis. It provides a peak list with M/Z ratio which usually represents b (represented in red) or y ions (represented in blue). The difference of mass between two consecutive peaks gives the mass of a single amino acid. The peak pattern thus contains sequence information allowing the identification of the peptide. Each correctly assigned spectrum is then counted giving the spectral count value for each protein in each sample.

Table 4: Spectral count, sequence coverage and identification probability of COMT in each sample

<table>
<thead>
<tr>
<th>Sequence Coverage</th>
<th>Protein</th>
<th>Category</th>
<th>Bio Sample</th>
<th>MS/MS Sa...</th>
<th>P</th>
<th>%Spec</th>
<th>#Repl</th>
<th>#Unq</th>
<th>#Spec</th>
<th>%Cov</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>Strain-ClF</td>
<td>Strain-ClF</td>
<td>100%</td>
<td>0.037%</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>11%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>Strain-E2</td>
<td>Strain-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>13%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>Strain-LxA4</td>
<td>Strain-LxA4</td>
<td>100%</td>
<td>0.037%</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>13%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>Strain-LxA4</td>
<td>Strain-LxA4</td>
<td>100%</td>
<td>0.037%</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>13%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>22A-E2</td>
<td>22A-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>22A-LX4-E2</td>
<td>22A-LX4-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>INTER-E1</td>
<td>INTER-E1</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>INTER-E1</td>
<td>INTER-E1</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>INTER-E2</td>
<td>INTER-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>INTER-E2</td>
<td>INTER-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>22A-E2</td>
<td>22A-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catechol O-methyltransferase...</td>
<td>22A-LX4-E2</td>
<td>22A-LX4-E2</td>
<td>100%</td>
<td>0.037%</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>30 kDa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All statistical analysis and protein quantitation are based on the unweighted spectral count raw data, not normalized, allowing by the usage of the fisher exact test, as described above.
Quality control

Figure 15: Quality control: A) Dendrogram of all cell lines and treatments, B) correlation matrix

For quality control, a dendrogram, produced by hierarchical clustering (with average linkage on Euclidian distance), was used to compare each sample and to determine if all the samples of the
same cell line clustered together based on protein quantitation data. The correlation matrix was another possibility to compare relationships between samples. Both analyses showed that the samples from the same cell line were highly correlated or close in the hierarchical tree analysis. And interestingly both endometriotic cell lines were more closely related to each other than the normal cell lines.

**Differences between endometriotic and normal endometrial cell lines**

![Figure 16: A) Volcano plot of 12Z control versus hTERT EEC control, B) Volcano plot of 22B control versus stromal control, C) Box plot of log ratio on protein expression data on 12Z control versus hTERT EEC control and 22B control versus stromal control](image)

Figure 16 A represents the data analysis of (unstimulated) 12Z endometriotic cell line control versus hTERT EEC non endometriotic cell line control, where the log ratio of 12Z versus hTERT for each protein was computed on the x-axis and the p-value from the fisher exact test corrected with the Benjamini Hochberg method was computed on the y-axis. Significantly differentially expressed proteins are shown in red (significance level at $10^{-4}$ due to Benjamini Hochberg correction), which are potential biomarkers for endometriosis. The volcano plot in figure 16 B, compares the 22B stromal endometriotic cell line with the stromal non endometriotic cell line, to identify potential biomarkers for endometriosis. The box plots in figure 17 C show the log ratio of each protein for both comparisons. We can see that the variance of the 12Z versus hTERT EEC is larger than the variance of 22B versus normal stromal cells.

Using the Fisher-Snedecore test, we observed that the difference between the variance of the log ratio of 12Z epithelial endometriotic cell line versus hTERT epithelial non endometriotic cells and the variance of the log ratio of 22B stromal endometriotic cell line versus stromal non endometriotic cell line was significant at 99% confidence level ($p$-value = 0.001228). This suggests that the differences between the two epithelial cell lines are more striking than the differences between the two stromal
cell lines. Epithelial cells may play a more important role in the disease process, but this assumption needs further investigation.

Uniprot identifiers for the 31 significantly differentially expressed proteins, for the two volcano plot presented above, are listed in table 5.

| Table 5: Uniprot ID of proteins significantly differentially expressed between endometriotic and non endometriotic cell lines according to the Fischer exact test with Benjamini Hochberg correction |
|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| **up regulation** | **down regulation** |
| **12Z ctrl vs hTERT ctrl** | FLNC_HUMAN, VIME_HUMAN, MYH9_HUMAN, ACTG_HUMAN, TBB3_HUMAN, HSP72_HUMAN, FLNC_HUMAN, TBB6_HUMAN, FLNA_HUMAN, LOH8_HUMAN, ALLA3_HUMAN, TBA1C_HUMAN, ACTC_HUMAN, TBA1B_HUMAN, GSTP1_HUMAN | K2CB_HUMAN, EPIPL_HUMAN, H4_HUMAN, KIC18_HUMAN, 4F2_HUMAN, KIC14_HUMAN, H2A1D_HUMAN, PROF1_HUMAN, EF1A2_HUMAN, H33_HUMAN, FAS_HUMAN, PIEC_HUMAN, COF1_HUMAN, LAT1_HUMAN |
| **22B ctrl vs stromal ctrl** | TBA1B_HUMAN, K2CB_HUMAN, KIC18_HUMAN, SPTB2_HUMAN, SPTA2_HUMAN, FAS_HUMAN, AL1A3_HUMAN, ENOA_HUMAN, K1C16_HUMAN, GCN1L_HUMAN | VIME_HUMAN, KIC10_HUMAN, ACTG_HUMAN, COG3A3_HUMAN, FAS_HUMAN, K22E_HUMAN, ACTC_HUMAN |

As we wished to investigate the pathway annotations related to the potential biomarkers discovered, we used Biocompendium to generate the next figure.

**Figure 17**: Pathway enrichment analysis of proteins differentially expressed between endometriotic and non endometriotic cell lines.
Figure 17 represents all the pathway annotations (in red) for the potential biomarkers (in green) discovered using both volcano plot analyses. This time gene names are used to represent the corresponding protein. Some of these potential biomarkers are implicated in known biological processes underlying the disease. For example, the protein actin gamma 1 (ACTG1) related to adherence junction and leukocyte trans-endothelial migration is significantly differentially expressed. Fatty acid synthase (FASN) is an important enzyme in long chain fatty acid synthesis and fatty acid synthesis has been implicated in the pathology of endometriosis (95). The gap junction related proteins are also detected with tubulins such as TUBB3, TUBB6, TUBA1C and TUBA1B. The spectrin alpha chain (SPTAN1), the myosin heavy chain 9(MYH9) and the protein actin gamma 1 (ACTG1) are tight junction related proteins. We can see that Mitogen-activated protein kinase (MAPK) signaling pathway related proteins might be important biomarkers, notably the heat shock-related 70 kDa protein 2 (HSPA2), filamin-C (FLNC) and filamin-A (FLNA). Some links with metabolic pathways have been identified through L-lactate dehydrogenase B chain (LDHB), enolase 1 (ENO1), Collagen alpha-3(VI) chain (COL6A3) and aldehyde dehydrogenase family 1 member A3 (ALDH1A3). The actin cytoskeleton regulation pathway is playing a role in the disease through profilin-1 (PFN1), myosin-9 (MYH9), coflin-1 (CFL1) and fibronectin (FN1).

Some of the potential biomarkers, such as vimentin (VIME) or keratin, type II cytoskeletal 2 epidermal (K22E), don’t have a pathway annotation, but could conceivably play an important role in the disease process; further gene ontology analysis could clarify this.
**Effects of treatment**

Each treatment is compared with the vehicle control sample for all four cell lines. The log ratio of the treatment spectral count versus the control spectral count for the 1139 proteins was computed. The figure 18 A represents the heat map with a hierarchical clustering of each protein pattern under different conditions. Each protein with a similar pattern across all the conditions is regrouped allowing the visualization of clusters with similar responses to different treatments. These clusters are very useful to understand how lipoxin A₄ and estradiol E2 act on the different cell lines. The up regulated proteins are represented in red and the down regulated proteins are represented in green. The brown color represents the proteins whose expression remains unchanged. Figure 18 B shows the box plot of log ratios for each condition (treatment versus control in each cell line). The red box represents 12Z endometriotic epithelial cells, the green box represents 22B stromal endometriotic cells, the blue box represents hTERT EEC epithelial non endometriotic cells and the brown box represents stromal non endometriotic cells.
Proteins which were significantly differentially expressed for each treatment were determined using the Fisher exact test corrected with Benjamini Hochberg method for multiple testing on the proteins and Bonferroni method for multiple testing on the different conditions (figure 19 and table 6). Unfortunately, with these stringent corrections, only a low number of proteins were significantly expressed upon the different treatments. In particular, the Benjamini Hochberg correction reduced the number of significantly differentially expressed proteins because it drastically decreased the p-value significance ($10^{-6}$ for a FDR of 0.01).
Interestingly none of the proteins were significantly differentially expressed in the stromal non endometriotic cells, indicating a lack of E2 responsiveness. Responses vary across cell lines but the overall response is not very dramatic.

The list of significantly differentially expressed proteins is given in the table below.
Table 6: Uniprot ID of proteins significantly differentially expressed between treated and non treated cell lines according to Fischer exact test (p-value<0.01) adjusted with Benjamini Hochberg method and Bonferroni method

<table>
<thead>
<tr>
<th></th>
<th>up regulation</th>
<th>down regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>12Z E2 vs 12Z ctrl</td>
<td></td>
<td>HSP72_HUMAN</td>
</tr>
<tr>
<td>12Z LXA4 vs 12Z ctrl</td>
<td></td>
<td>HSP72_HUMAN, G3P_HUMAN</td>
</tr>
<tr>
<td>12Z LXA4 E2 vs 12Z ctrl</td>
<td></td>
<td>HSP72_HUMAN</td>
</tr>
<tr>
<td>22B E2 vs 22B ctrl</td>
<td></td>
<td>K1C16_HUMAN, K2C6C_HUMAN, K1C14_HUMAN, K2C5_HUMAN</td>
</tr>
<tr>
<td>22B LXA4 vs 22B ctrl</td>
<td></td>
<td>K2C6C_HUMAN, K1C14_HUMAN, K2C5_HUMAN</td>
</tr>
<tr>
<td>22B LXA4 E2 vs 22B ctrl</td>
<td></td>
<td>K1C16_HUMAN, K2C6C_HUMAN, K1C14_HUMAN, K2C5_HUMAN</td>
</tr>
<tr>
<td>hTERT E2 vs hTERT ctrl</td>
<td>PPIA_HUMAN, TBB3_HUMAN</td>
<td>TBB2B_HUMAN</td>
</tr>
<tr>
<td>hTERT LXA4 vs hTERT ctrl</td>
<td>PPIA_HUMAN, TBB3_HUMAN</td>
<td>TBB2B_HUMAN</td>
</tr>
<tr>
<td>hTERT LXA4 E2 vs hTERT ctrl</td>
<td>PPIA_HUMAN</td>
<td>PLEC_HUMAN</td>
</tr>
<tr>
<td>stromal E2 vs stromal ctrl</td>
<td>PPIA_HUMAN</td>
<td>PLEC_HUMAN</td>
</tr>
<tr>
<td>stromal LXA4 vs stromal ctrl</td>
<td>PPIA_HUMAN</td>
<td>PLEC_HUMAN</td>
</tr>
</tbody>
</table>

The 22B endometriotic stromal cells showed a statistically significantly down regulation of keratin proteins. This could have a possible link with proliferation, but unfortunately cannot be distinguished from a contamination by keratins during sample preparation.

Figure 20: Pathway enrichment analysis of proteins highly impacted by the different treatments

Here we find that heat shock-related 70 kDa protein 2 (HSPA2) is annotated for MAPK signaling pathway and antigen processing, tubulins TUBB3 and TUBB2B for gap junction and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is linked with metabolic pathway. Some proteins such as peptidyl-prolyl cis-trans isomerase A (PPIA) or plectin (PLEC) do not have a pathway annotation but have been annotated in Gene Ontology which allows an insight into their potential roles.
Figure 21: Impact of treatment on the potential biomarkers of endometriosis in the different cell lines

Potential endometriosis biomarkers discovered were monitored upon treatment in each cell line. Figure 21 shows the heatmap of the potential biomarkers of endometriosis. The tubulins TUB6, TUBB3, TUBBA1B and TUBBA1C are up regulated by the different treatments in the hTERT EEC non endometriotic cell lines and are slightly down regulated in the other cell lines. Heat shock-related 70 kDa protein 2 (HSPA2) was decreased by the different treatments in 12Z epithelial endometriotic cell line. Lipoxin A₄ caused an increase in Filamin-C (FLNC) expression in the hTERT EEC cells whereas it was slightly decreased in the other conditions. Histone H4 (HIST1H4E) was up regulated by the different treatments in 12Z epithelial endometriotic cell lines and down regulated in the other conditions, especially in the 22B stromal endometriotic cell line treated with lipoxin A₄. Translational activator GCN1 (GCN1L1) was enhanced in the stromal non endometriotic cell line and strongly positively impacted by the estradiol treatment. This protein was down regulated in the other conditions and strongly reduced in 22B stromal cell line treated with lipoxin A₄. Actin, cytoplasmic 2 (ACTC1) was highly augmented in hTERT EEC epithelial non endometriotic cell line treated with estradiol. Enolase 1 (ENO1) was a strong positive response in the hTERT EEC epithelial non endometriotic cell line with a decrease by estradiol and increase by lipoxin A₄. The vimentin (VIM) was increased by lipoxin A₄ in hTERT EEC. Aldehyde dehydrogenase family 1 member A3 (ALDH1A3) was strongly down regulated by estradiol in 12Z epithelial endometriotic cell lines and Glutathione S-transferase P (GSTP1) expression was strongly augmented in the same condition but was decreased by lipoxin A₄ treatment. L-lactate dehydrogenase B chain (LDHB) was strongly reduced in 22B stromal endometriotic cell lines by...
lipoxin A₄ treatment and increased in the stromal non endometriotic cell line by estradiol treatment. Fatty acid synthase (FASN) showed a strong up regulation in the stromal non endometriotic cell line treated with estradiol. 4F2 cell-surface antigen heavy chain (SLC3A2) was slightly augmented by the lipoxin A₄ in 12Z epithelial endometriotic cell line and in stromal non endometriotic cell line. Keratins KRT1 8 and KRT2 showed a similar response with a light up regulation in 12Z upon all treatments and a down regulation by estradiol in 22B stromal endometriotic cells. KRT2 was not impacted by lipoxin A₄ in 22B stromal endometriotic cell line. Interestingly, Cofilin-1 (CFL1) showed a strong rise with the combination of the two treatments in the 22B stromal endometriotic cell line and a reduction in hTERT EEC epithelial non endometriotic cell line upon estradiol E2 treatment. Myosin-9 (MYH9) exhibited a similar pattern to CFL1 in the 22B stromal endometriotic cell line but in hTERT EEC cells it was down regulated in the same manner by each treatment. The large neutral amino acid transporter small subunit 1 (SLC7A5) was markedly attenuated by lipoxin A₄ in hTERT EEC epithelial endometriotic cell line. Profilin-1(PFN1) was strongly reduced by the lipoxin A₄ treatment in 22B stromal endometriotic cells.

Filamin-A (FLN1), Plectin (PLEC), Epipilakin (EPPK1), Fibronectin FN1, Spectrin beta chain (SPTB2) and Spectrin alpha chain (SPTA2) had a similar pattern across all conditions and they showed a probable synergic response in 22B stromal endometriotic cell upon lipoxin A₄ and estradiol treatment.

As hTERT EEC epithelial non endometriotic cells are known to be estrogen responsive (59), we compared the response for genes containing estrogen responsive elements (ERE) in their promoter and highly up or down regulated in the hTERT EEC cell line treated with estradiol. We selected a subset list of 36 genes containing an ERE and searched for corresponding 18 up-regulated (fold change > 2) or 18 down-regulated (fold change < 0.5) proteins (Table 7).

<table>
<thead>
<tr>
<th>Negatively regulated proteins from ERE genes in hTERT cells</th>
<th>Positively regulated proteins from ERE genes in hTERT cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDIR1_HUMAN CH10_HUMAN LASP1_HUMAN CPT1A_HUMAN ACSL3_HUMAN COPE_HUMAN ODO2_HUMAN PDL1_HUMAN DEC_R_HUMAN VIGLN_HUMAN TIMK2_HUMAN LASP1_HUMAN TOP2B_HUMAN SPTB2_HUMAN CPT1A_HUMAN CALC_HUMAN SRP68_HUMAN CPT1A_HUMAN SYW_HUMAN FA49B_HUMAN VIGLN_HUMAN HNRPL_HUMAN ODO2_HUMAN ODO1_HUMAN PELP1_HUMAN ACSL3_HUMAN</td>
<td>ARF4_HUMAN LEG1_HUMAN SSBP_HUMAN ARPC2_HUMAN SRR1_HUMAN ARPC2_HUMAN RFBM_HUMAN CALM_HUMAN FKBP4_HUMAN MYL6_HUMAN ARF1_HUMAN DOPD_HUMAN SSBP_HUMAN ATPD_HUMAN PHL2_HUMAN SSBP_HUMAN COMT_HUMAN ARF1_HUMAN RM43_HUMAN PTH2_HUMAN SSBP_HUMAN RS12_HUMAN SDOS_HUMAN ARF1_HUMAN LEG1_HUMAN COMT_HUMAN MGN_HUMAN CALM_HUMAN SDOS_HUMAN MGN_HUMAN PTH2_HUMAN</td>
</tr>
</tbody>
</table>

The fold change of these proteins for each condition in each cell line was calculated and is displayed in figure 22. The 12Z epithelial endometriotic cell line is represented in red, the 22B stromal endometriotic cell line is in green, the hTERT epithelial non endometriotic cell line is in blue and the stromal non endometriotic cell line is presented in brown.
Figure 22: Protein expression in all cell lines and for all treatment of genes retrieved from the ERE database and responsive to estradiol in hTERT EEC cells (log ratio > log (2) (A) or log ratio < -log(2) (B)).

Figure 22 A shows the down-regulated proteins from ERE genes in the hTERT EEC treated with estradiol E2 versus hTERT EEC control. The responses in the other conditions and cell lines were quite different. Firstly, the 12Z epithelial endometriotic cell line seemed to be less affected by the treatment than hTERT EEC cells. Secondly, both stromal cell lines were mostly not affected by the treatment. Thirdly, lipoxin A4 alone seemed to have a similar effect to estradiol on those genes in a lower proportion, but a slightly antagonistic effect with a combined treatment.

Figure 22 B shows the responses of the proteins from ERE genes which were up-regulated in hTERT EEC cells treated with estradiol. In this case most of those genes were not present in 12Z epithelial endometriotic cell line, and again both stromal cell lines were slightly affected by the treatments.

The statistical analysis is presented below (Table 8).
Table 8: Statistical analysis of estrogen responsive genes

<table>
<thead>
<tr>
<th>treatments</th>
<th>cell line</th>
<th>t-test p val</th>
<th>var test p val</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2 vs LXA4 E2</td>
<td>12z</td>
<td>7.70E-01</td>
<td>5.98E-01</td>
</tr>
<tr>
<td>E2 vs LXA4 E2</td>
<td>12z</td>
<td>1.40E-02</td>
<td>5.95E-01</td>
</tr>
<tr>
<td>LXA4 vs LXA4 E2</td>
<td>12z</td>
<td>4.22E-02</td>
<td>2.91E-01</td>
</tr>
<tr>
<td>E2 vs LXA4</td>
<td>22b</td>
<td>3.99E-01</td>
<td>3.93E-01</td>
</tr>
<tr>
<td>E2 vs LXA4 E2</td>
<td>22b</td>
<td>5.99E-01</td>
<td>6.86E-04</td>
</tr>
<tr>
<td>LXA4 vs LXA4 E2</td>
<td>22b</td>
<td>8.78E-02</td>
<td>9.01E-03</td>
</tr>
<tr>
<td>E2 vs LXA4 hTERT</td>
<td></td>
<td>1.17E-02</td>
<td>3.89E-05</td>
</tr>
<tr>
<td>E2 vs LXA4 E2 hTERT</td>
<td></td>
<td>2.10E-03</td>
<td>1.00E+00</td>
</tr>
<tr>
<td>LXA4 vs LXA4 E2 hTERT</td>
<td></td>
<td>2.12E-01</td>
<td>6.36E-02</td>
</tr>
<tr>
<td>E2 vs LXA4 str</td>
<td></td>
<td>5.13E-01</td>
<td>9.64E-01</td>
</tr>
<tr>
<td>E2 vs LXA4 E2 str</td>
<td></td>
<td>3.41E-01</td>
<td>9.31E-01</td>
</tr>
<tr>
<td>LXA4 vs LXA4 E2 str</td>
<td></td>
<td>7.56E-01</td>
<td>8.95E-01</td>
</tr>
<tr>
<td>E2 12z vs h tert</td>
<td></td>
<td>3.69E-06</td>
<td>1.78E-08</td>
</tr>
<tr>
<td>LXA4 12 z vs hTERT</td>
<td></td>
<td>1.66E-03</td>
<td>1.75E-02</td>
</tr>
<tr>
<td>LXA4E2 12z vs hTERT</td>
<td></td>
<td>6.62E-05</td>
<td>6.07E-01</td>
</tr>
<tr>
<td>E2 22b vs str</td>
<td></td>
<td>5.10E-01</td>
<td>1.17E-01</td>
</tr>
<tr>
<td>LXA4 22b vs str</td>
<td></td>
<td>2.61E-01</td>
<td>4.43E-01</td>
</tr>
<tr>
<td>LXA4E2 22b vs str</td>
<td></td>
<td>3.92E-01</td>
<td>4.46E-02</td>
</tr>
<tr>
<td>E2 12z vs 22b</td>
<td></td>
<td>4.16E-03</td>
<td>6.60E-01</td>
</tr>
<tr>
<td>LXA4 12z vs 22b</td>
<td></td>
<td>5.85E-02</td>
<td>3.47E-01</td>
</tr>
<tr>
<td>LXA4E2 12z vs 22b</td>
<td></td>
<td>1.17E-01</td>
<td>1.24E-02</td>
</tr>
<tr>
<td>E2 hTERT vs str</td>
<td></td>
<td>1.84E-14</td>
<td>2.11E-06</td>
</tr>
<tr>
<td>LXA4 hTERT vs str</td>
<td></td>
<td>8.79E-11</td>
<td>4.78E-01</td>
</tr>
<tr>
<td>LXA4E2 hTERT vs str</td>
<td></td>
<td>3.98E-06</td>
<td>3.03E-01</td>
</tr>
</tbody>
</table>

The fisher-snedecore test was used to estimate whether the variance was significantly different between two log ratio comparisons, as 22B cell line treated with estradiol compared to control versus 22B cell line treated with lipoxin A<sub>4</sub> and estradiol compared to control. We only considered the list of estrogen responsive genes up regulated or down regulated in the hTERT EEC treated with estradiol compared to control. We also used the student t-test for the comparison of the means. Apparently there was a probable suppressive effect in 22B treated with lipoxin A<sub>4</sub> and estradiol compared either to lipoxin A<sub>4</sub> or estradiol alone. Interestingly, the negatively regulated estrogen responsive genes didn’t respond in the same way in hTERT EEC treated with lipoxin A<sub>4</sub> or with estradiol. There was an important difference between hTERT EEC and 12Z cells response in term of means in the negatively and positively estrogen regulated genes for each treatment. The same observation was applicable to hTERT EEC cells compared to the normal stromal cells. The variance was significantly impacted by estradiol treatment in hTERT EEC cells compared to 12Z cells and in hTERT EEC cells compared to stromal cells for the negatively estrogen regulated genes, whereas the combination of both treatment significantly impacted the variance in the same comparisons in positively estrogen regulated genes.
Known Biomarkers

We wished to determine if we could detect some known biomarkers of endometriosis. We used the EAELi text mining tool, the NCBI annotated genes for endometriosis and a list of interesting proteins form the mucosal immunity lab to build a list of known biomarkers. The intersection between the list of detected proteins and the list of known biomarkers was computed to find the 24 known biomarkers detected among the detected proteins in our experiment. These biomarkers were monitored in each conditions and the log ratio of treatment versus control are presented in the figure 23.

Solute carrier family 2, facilitated glucose transporter member 1 (SLC2A1) expression was increased by the combination of both treatments versus control in the 22B stromal endometriotic cell line, but it was slightly decreased in the other conditions. Fibronectin (FN1) exhibed a similar pattern with SLC2A1 with a small difference in 22B treated either with lipoxin A₄ or with estradiol, where its expression was slightly augmented. Mitogen-activated protein kinase 1(MAPK1) was enhanced in hTERT EEC non endometriotic cells by the combination of both treatments whereas it was reduced in all other conditions except in 12Z epithelial endometriotic cells, where it didn’t change. HLA class I histocompatibility antigen, A-2 alpha chain (HLA-A) was up regulated by the combination of both treatments in 12Z epithelial endometriotic cell lines and slightly up regulated by lipoxin A₄ alone.
HLA-A was diminished in all other cell lines and treatments. The apoptosis regulator BAX was increased in hTERT EEC treated with the combination of lipoxin A₄ and estradiol and with the estradiol alone, but it was slightly decreased by lipoxin A₄ in the same cell line and in all other conditions except in stromal non endometriotic cell line treated with estradiol where the reduction was stronger. Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCHL1) was down regulated in 12Z epithelial endometriotic cells treated with estradiol and it was strongly up regulated in stromal non endometriotic cell line treated with the combination of both treatments. Interestingly, intercellular adhesion molecule 1 (ICAM1) was reduced in 12Z endometriotic cell lines by both treatments alone, whereas it was strongly increased by the combination of lipoxin A₄ and estradiol. ICAM1 was down regulated in the other conditions. Peroxiredoxin-6 (PRDX6) expression was down regulated in 12Z epithelial endometriotic cell line treated with lipoxin A₄ and slightly enhanced in 22B stromal endometriotic cell line treated with estradiol. Glutathione S-transferase P (GSTP1) was up regulated in 12Z and 22B endometriotic cell lines treated with estradiol and it was down regulated in other conditions, especially in 12Z treated with lipoxin A₄ or lipoxin A₄ with estradiol and in stromal non endometriotic cell line treated with the combination of lipoxin A₄ and estradiol. Cellular tumor antigen p53 (P53) showed a marked increase in 12Z cells treated with estradiol and with the combination of both treatments, but it slightly decreased in all other conditions. Neural cell adhesion molecule L1 (L1CAM) was strongly down regulated in 12Z epithelial endometriotic cell line treated with lipoxin A₄ and the combination of both treatments. Actin (ACTN1) was down regulated in hTERT epithelial non endometriotic cell line treated with estradiol. ACTN1 decreased in hTERT EEC epithelial non endometriotic cell line by estradiol and was increased in the stromal non endometriotic cell line by lipoxin A₄. CD44 antigen (CD44) was down regulated by estradiol on 22B stromal endometriotic cell line and it was slightly up regulated by lipoxin A₄ in the same cell line. We saw that Transgelin-2 expression (TAGLN2) was diminished by lipoxin A₄ in 22B stromal endometriotic cell line. Vimentin expression (VIM) had a marked augmentation in hTERT EEC epithelial non endometriotic cell line treated with lipoxin A₄, but it was slightly decreased in the other conditions. Cofilin-1 (CFL) showed a down regulation in hTERT EEC epithelial endometriotic cell line treated with estradiol and Macrophage migration inhibitory factor (MIF) was decreased in 22B stromal endometriotic cell line treated with lipoxin A₄. Low molecular weight phosphotyrosine protein phosphatase (ACP1) was as well down regulated in 22B stromal endometriotic cell line treated with lipoxin A₄. Stathmin (STMN1) levels were increased in hTERT EEC epithelial non endometriotic cells by lipoxin A₄ treatment. Epidermal growth factor receptor (EGFR) was markedly down regulated by the lipoxin A₄ treatment in 22B stromal endometriotic cell line. Catenin beta-1(CTNNB1), Integrin alpha-V (ITGAV) and Annexin A1 (ANXA1) showed a similar pattern and were decreased by the three different treatments in hTERT EEC epithelial non endometriotic cell line with notably a strong down regulation of CTNNB1 in the estradiol treated sample. There was a slightly declination of ANXA1 in 12Z epithelial endometriotic cell line treated with lipoxin A₄. Finally, Catechol O-methyltransferase (COMT) was marked attenuated in 12Z epithelial endometriotic cell line treated with estradiol or lipoxin A₄ and in stromal non endometriotic cells treated with the combination of lipoxin A₄ and estradiol.

We then wanted to identify the different pathways related to the known biomarkers to understand their potential role in the disease process. The enrichment analysis provides much information on the potential target of estradiol and lipoxin A₄ in terms of pathway or process.
As can be seen in figure 24, the known biomarkers for endometriosis have numerous pathway annotations (as evaluated using Biocompendium) because they are much more studied than the potential biomarkers discovered in this work. Mitogen-activated protein kinase 1 (MAPK1) is the most studied and it has links with some relevant pathway like vascular endothelial growth factor (VEGF) signaling, transforming growth factor (TGF) beta signaling and of course MAPK signaling pathway, where stathmin (STMN1) is playing a role. MAPK1 is related to epidermal growth factor receptor (ErbB) signaling pathway, Gap junction, gonadotropin-releasing hormone (GnRH) signaling and inflammation with B cells and T cells receptor signaling and mammalian target of rapamycin (mTOR) signaling. Cofilin-1 (CFL1) is associated with actin cytoskeleton regulation. HLA class I histocompatibility antigen (HLA-A) and neural cell adhesion molecule L1 (L1CAM) and Intercellular adhesion molecule (ICAM) are linked to adhesion process. Epidermal growth factor receptor (EGFR) is related to proliferation through ErbB signaling pathway and apoptosis regulator (BAX) is a key protein for apoptosis. Cellular tumor antigen p53 (TP53) is connected to cell cycle and cancer, which have similarity in some aspect with endometriosis. Catenin beta-1 (CTNNB1) is related to adhesion and leukocyte transendothelial migration as well as actin (ACTN1). Glutathione S-transferase P (GSTP1) is linked to cancer, glutathione and drug metabolism and CD44 is associated with extra cellular matrix receptor and hematopoietic cell lineage. Solute carrier family 2, facilitated glucose transporter member 1 (SLC2A1) is linked to adipocytokine signaling and cancer. Catechol O-methyltransferase COMT is related to steroid hormone biosynthesis and metabolism. Macrophage migration inhibitory factor (MIF) and Peroxiredoxin-6 (PRDX6) are implicated in metabolism and immune response. Low molecular weight phosphotyrosine protein phosphatase (ACP1) is linked with cancer and metabolism. These pathway links give a very interesting view of the role of each biomarker.
**Contextualization**

The contextualization will be discussed in detail in the discussion section.

In the figure 25 how the known biomarkers interact with the potential biomarkers and the proteins highly affected by the different treatments (treatment biomarkers) is delineated. We built the protein-protein interaction network employing Cytoscape software, using interactome data (GeneMANIA).

![Figure 25: Biomarker protein-protein interaction network](image)

Red nodes represent the known biomarkers, yellow nodes represent the potential biomarkers and turquoise nodes represent the proteins which are highly affected by the treatment. Orange represents proteins which are known biomarkers and potential biomarkers discovered by the statistical analysis. Potential biomarkers are represented in green, whose expressions were highly affected by one of the treatments. The grey nodes represent proteins which were not directly related to this network, and are involved in other biological processes. Interestingly potential biomarkers, known biomarkers and treatment biomarkers have a lot of physical-interactions. Several potential biomarkers were highly affected by one of the treatments, notably PLEC, TUBB3, TUBB2B and KRT14, KRT16.

The table 9 shows the Gene ontology analysis of the protein network above.
GeneMANIA only considers annotations in GO terms with between 10 and 300 annotations in the organism of interest. The GO categories and Q-values from a FDR corrected hypergeometric test for enrichment were reported, along with coverage ratios for the number of annotated genes in the displayed network compared to the number of genes with that annotation in the genome. The Q-values were estimated using the Benjamini-Hochberg procedure. Categories are displayed up to a Q-value cutoff of 0.1.

The potential biomarkers, whose expression were statistically differentially expressed in normal versus endometriotic cell lines, and the known biomarkers in the stromal cell lines and in the epithelial cell lines, are contextualized in the following figures. The pathway annotation from bio compendium and the treatment effect are mapped onto the images. The yellow color represents the biomarkers which are up regulated in endometriotic case whereas the blue color represents the biomarkers which are up regulated in normal cases (figure 26-27).
Figure 26: Contextualization of potential biomarkers: A) Epithelial cells, B) Stromal cells
Figure 27: Contextualization of known biomarkers: A) Epithelial cells, B) Stromal cells
**Synergic or suppressive effect**

To determine whether the combination of lipoxin A₄ treatment and estradiol treatment leads to a suppressive or a synergic effect on the expression of specific proteins, we develop a mathematical model. With the data of the log ratio of the treatment versus the control for each cell line we could compute the effect of lipoxin A₄ treatment alone, the effect of estradiol alone and the effect of the combination of both treatments. In the case of no interaction between treatments, the effect of both treatments should be equal to the sum of the effect of lipoxin A₄ treatment and the effect of estradiol E₂ treatment. This gave the equation of the linear model below.

**Linear model:**

\[
\log_{10} \frac{LXA₄E₂}{Ctrl} \sim \log_{10} \frac{LXA₄}{Ctrl} + \log_{10} \frac{E₂}{Ctrl}
\]

Mathematically the synergic effect means that the effect of both treatments was greater than the sum of the effect of each treatment.

**Synergic effect**

\[
\log_{10} \frac{LXA₄E₂}{Ctrl} > \log_{10} \frac{LXA₄}{Ctrl} + \log_{10} \frac{E₂}{Ctrl}
\]

The suppressive effect was defined in the same way, but in this case the effect of both treatments was lower than the sum of each treatment effect.

**Suppressive effect**

\[
\log_{10} \frac{LXA₄E₂}{Ctrl} < \log_{10} \frac{LXA₄}{Ctrl} + \log_{10} \frac{E₂}{Ctrl}
\]

Finally we compared the expression values of proteins using the linear model to find those which showed either a suppressive or a synergic effect. We used a 99% confidence interval around the linear model to classify the proteins. If their residue was outside of the 99% confidence interval and negative or positive, they were considered to show a suppressive or a synergic effect respectively.

The residue, which represents the distance to the linear model for a single protein, is estimated with the following equation:

\[
\text{residue} = \log_{10} \frac{LXA₄E₂}{Ctrl} - \left( \log_{10} \frac{LXA₄}{Ctrl} + \log_{10} \frac{E₂}{Ctrl} \right)
\]

The resulting calculation for the synergic and the suppressive case are given by the following equations.

**Synergic case**

\[
\log_{10} \frac{LXA₄E₂}{Ctrl} > \log_{10} \frac{LXA₄}{Ctrl} + \log_{10} \frac{E₂}{Ctrl} + \Delta
\]

**Suppressive case**

\[
\log_{10} \frac{LXA₄E₂}{Ctrl} < \log_{10} \frac{LXA₄}{Ctrl} + \log_{10} \frac{E₂}{Ctrl} - \Delta
\]

With

\[
\Delta = \Phi^{-1}(99\%) \times \sigma \quad \sigma \text{ is the standard deviation of residues.}
\]

\(\Delta\) represents the 99% confidence interval, assuming that the residues follow a normal distribution.
We constructed a 3D space where the x-axis represents the effect of estradiol treatment alone, the y-axis represents the effect of lipoxin A4 treatment alone and the z-axis represents the effect of the combination of both treatments. The proteins present in 12Z epithelial endometriotic cell line are represented in red, the proteins present in 22B stromal endometriotic cell line are in green, the proteins present in hTERT EEC epithelial non endometriotic cell line are in blue and the proteins present in stromal non endometriotic cell line are presented in brown. The black plane represents the linear model and both of the red planes represent the limit of the 99% confidence interval. The proteins which are above or below these planes are respectively in synergic case or suppressive case. The histogram and the density curve of the residue show that we can approximate the distribution by a normal distribution.

The table below gives the uniprot identifiers of the proteins which are either in synergic case or in suppressive case in each cell lines. Interestingly we find some of our potential biomarkers, known biomarkers and treatment biomarkers in these lists (shown in red).

Figure 28: Suppressive or synergic effect and linear model: A) 3D plot of all protein in the different cell lines in function of the effect of E2, LXA4 and both, B) residues histogram and density distribution.
Table 10: Suppressive or synergic effect of estradiol and lipoxin A₄ treatment on specific proteins in each cell line

<table>
<thead>
<tr>
<th>12Z</th>
<th>22B</th>
<th>hTERT</th>
<th>stromal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Suppressive effect</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K1C16_HUMAN</td>
<td>PSA4_HUMAN</td>
<td>TBB3_HUMAN</td>
<td>PPT1_HUMAN</td>
</tr>
<tr>
<td>K1C14_HUMAN</td>
<td>VAMP7_HUMAN</td>
<td>PPIA_HUMAN</td>
<td>LYRIC_HUMAN</td>
</tr>
<tr>
<td>K2C6C_HUMAN</td>
<td>EIF3L_HUMAN</td>
<td>TBB6_HUMAN</td>
<td>GCN1c_HUMAN</td>
</tr>
<tr>
<td>SRF53_HUMAN</td>
<td>TSN_HUMAN</td>
<td>NDBK_HUMAN</td>
<td>PEP1_HUMAN</td>
</tr>
<tr>
<td>LYP1A_HUMAN</td>
<td></td>
<td>KCY_HUMAN</td>
<td>XP05_HUMAN</td>
</tr>
<tr>
<td>E12BA_HUMAN</td>
<td></td>
<td>RM43_HUMAN</td>
<td>NIT2_HUMAN</td>
</tr>
<tr>
<td>RPAB1_HUMAN</td>
<td></td>
<td>COMT_HUMAN</td>
<td>SNUT2_HUMAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RAN_HUMAN</td>
<td>THOC4_HUMAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MYOF_HUMAN</td>
<td>NU205_HUMAN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DIMP_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBB27_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RIR2_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DOPD_HUMAN</td>
<td></td>
</tr>
<tr>
<td><strong>Synergic effect</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHD4_HUMAN</td>
<td>SC24C_HUMAN</td>
<td>CNA1_HUMAN</td>
<td>EHD1_HUMAN</td>
</tr>
<tr>
<td>ZYX_HUMAN</td>
<td>STLML2_HUMAN</td>
<td>CM2_HUMAN</td>
<td>ABCE1_HUMAN</td>
</tr>
<tr>
<td>ERG11_HUMAN</td>
<td>NATIO_HUMAN</td>
<td>SEPT2_HUMAN</td>
<td>RHEB_HUMAN</td>
</tr>
<tr>
<td>DPP3_HUMAN</td>
<td>NIPS1_HUMAN</td>
<td>STAT1_HUMAN</td>
<td>ACAD9_HUMAN</td>
</tr>
<tr>
<td>PLOD3_HUMAN</td>
<td>CO3A1_HUMAN</td>
<td>NIBL1_HUMAN</td>
<td>RPAB1_HUMAN</td>
</tr>
<tr>
<td>LMO7_HUMAN</td>
<td>NU155_HUMAN</td>
<td>SYEP_HUMAN</td>
<td>CDK1_HUMAN</td>
</tr>
<tr>
<td>SMHD1_HUMAN</td>
<td>LAMC1_HUMAN</td>
<td>ARL3_HUMAN</td>
<td>PAPS2_HUMAN</td>
</tr>
<tr>
<td>SPTB2_HUMAN</td>
<td>ADDA_HUMAN</td>
<td>SYMC_HUMAN</td>
<td>RFC2_HUMAN</td>
</tr>
<tr>
<td>PAPS2_HUMAN</td>
<td>TOP2B_HUMAN</td>
<td>IPYR_HUMAN</td>
<td>PLAK_HUMAN</td>
</tr>
<tr>
<td>IF4G1_HUMAN</td>
<td>HP1B3_HUMAN</td>
<td>CN0T1_HUMAN</td>
<td>UGPA_HUMAN</td>
</tr>
<tr>
<td>LASC1_HUMAN</td>
<td>RBM14_HUMAN</td>
<td>NU160_HUMAN</td>
<td>DESP_HUMAN</td>
</tr>
<tr>
<td>MCM5_HUMAN</td>
<td>GBF1_HUMAN</td>
<td>TWF1_HUMAN</td>
<td></td>
</tr>
<tr>
<td>USO1_HUMAN</td>
<td>NOP2_HUMAN</td>
<td>ODO2_HUMAN</td>
<td></td>
</tr>
<tr>
<td>UBQL2_HUMAN</td>
<td>EGRF_HUMAN</td>
<td>DNB_HUMAN</td>
<td></td>
</tr>
<tr>
<td>RBP2_HUMAN</td>
<td>NCKP1_HUMAN</td>
<td>SNUT2_HUMAN</td>
<td></td>
</tr>
<tr>
<td>COMT_HUMAN</td>
<td>IPYR_HUMAN</td>
<td>CPT1A_HUMAN</td>
<td></td>
</tr>
<tr>
<td>LAMC1_HUMAN</td>
<td>NRDC_HUMAN</td>
<td>PDA4_HUMAN</td>
<td></td>
</tr>
<tr>
<td>ACSL3_HUMAN</td>
<td>SYEP_HUMAN</td>
<td>XRC5_HUMAN</td>
<td></td>
</tr>
<tr>
<td>RRBP1_HUMAN</td>
<td>AP2A2_HUMAN</td>
<td>SPTA2_HUMAN</td>
<td></td>
</tr>
<tr>
<td>FUS_HUMAN</td>
<td>K2C5_HUMAN</td>
<td>ALDOC_HUMAN</td>
<td></td>
</tr>
<tr>
<td>VATA_HUMAN</td>
<td>K1C14_HUMAN</td>
<td>C1QBP_HUMAN</td>
<td></td>
</tr>
<tr>
<td>FPS5_HUMAN</td>
<td>K2C6c_HUMAN</td>
<td>NAT10_HUMAN</td>
<td></td>
</tr>
<tr>
<td>LAP2A_HUMAN</td>
<td>KINH_HUMAN</td>
<td>DSRA_HUMAN</td>
<td></td>
</tr>
<tr>
<td>RMB14_HUMAN</td>
<td></td>
<td>VACT1_HUMAN</td>
<td></td>
</tr>
<tr>
<td>ACA_HUMAN</td>
<td></td>
<td>FA4B_HUMAN</td>
<td></td>
</tr>
<tr>
<td>HSP72_HUMAN</td>
<td></td>
<td>PIGS_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD166_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AFG32_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FKB3_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RBM39_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ANKA3_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GLRK3_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SYM_HUMAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TBB2B_HUMAN</td>
<td></td>
</tr>
</tbody>
</table>
Discussion

MS based proteomics has several drawbacks for protein detection and identification. Peptide detection is dependent on abundance, solubility and ionization. For example insoluble membrane proteins are very difficult to detect (96, 97). The identification is linked with the spectrum quality in term of peak intensity, post-translational modification either natural or due to the sample preparation, and the quality of the database. Frame shift, incorrect open reading frames (ORF), or poor quality expressed sequence tag (EST) can affect the quality of the in silico digested database for peptide identification. There are many identification algorithm and validation software which have been developed to overcome some of these difficulties (98). Despite all these efforts we have detected only 1139 proteins and identified approximately 30% of the spectra produced by the LC-LTQ-Orbitrap-MS/MS, with a high confidence. Some sample preparation modifications might allow the increase the detection of some proteins of low abundance, such as using an increased number of fractions.

The quality control of the experiment shows that all samples for same cell line cluster together in the denrogram or are highly correlated in the correlation heatmap. This means that the experiment was well done, that the MS analysis was reproducible, and that the treatments did not have a major effect. Another interesting outcome of the quality control is that the two endometriotic cell lines are closer to each other in terms of profiling than the two normal cell lines. This tendency was also observed in a gene expression profiling study of human immortalized endometriotic epithelial and stromal cells (60).

The statistical analysis using the fisher exact test corrected with Benjamini-Hochberg method from endometriotic versus normal cell lines revealed 31 potential biomarkers, among the 1139 proteins detected. The box plot of log ratio between endometriotic and normal cell lines exert a higher variance for the epithelial cell line (p-value =0.001228) suggesting that epithelial cells are more impacted in the disease than the stromal cells, which is coherent with a gene expression profiling study (60) and the protein characterization study of the 12Z endometriotic epithelial cell line (99).

Effects of treatments

The treatment effect was evaluated comparing control sample versus each treatment for each cell line using the Fischer exact test corrected with Benjamini-Hochberg method and Bonferroni with comparisons. This analysis revealed 10 proteins, whose expression was highly modulated by at least one of the treatments. The treatment effect on known biomarkers and potential biomarkers were not statistically significant, but it showed interesting patterns which are described in the contextualization part.

Heat shock-related 70 kDa protein 2 (HSP72) was strongly down regulated in 12Z endometriotic epithelial cell line treated by lipoxin A4 and estradiol. This protein is implicated in the biotic stimulus response, defined as any process that results in a change in state or activity of a cell or an organism in terms of movement, secretion, enzyme production or gene expression. HSP72 is related to anti apoptotic effects in lung carcinoma (100). HSP72 stimulates neutrophil phagocytic ability (101) and was significantly associated with the presence of tumor infiltration, lymph node, and remote metastasis (102). HSP72 stimulates innate immunity through toll like receptor 4 and promotes cell
survival (103). This protein has been associated with fat oxidation and insulin resistance in skeletal muscle (104). This attenuated HSP72 expression could lead to a decrease of cell survival and infiltration of the 12Z endometriotic cell line. Lipoxin A₄ caused a reduction of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in the 12Z endometriotic cell line. The glycolytic enzyme GAPDH is a sensor for oxidative stress and promotes angiogenesis under hypoxic conditions in human chondrosarcoma-derived cell line (105). Consequently, lipoxin A₄ might exert an anti-angiogenic action in 12Z cells.

Both treatments attenuate the expression of several keratins (KRT6B, KRT14, KRT5, and KRT16) in 22B stromal endometriotic cells. A decrease of KRT14 has been associated with increased susceptibility to TNF-alpha induced apoptosis in keratinocytes (106). KRT14 and KRT5 have been used as a marker for immunohistochemical breast cancer study (107), which demonstrate some common points between cancer and endometriosis. KRT16 has been associated with keratinocyte migration and differentiation (108). It could therefore be speculated that estradiol and lipoxin A₄ can reduce migration in 22B stromal endometriotic cells.

Both estradiol and lipoxin A₄ treatments have increased peptidyl-prolyl cis-trans isomerase A (PPIA) and slightly Tubulin beta-3 chain (TBB3) in the hTERT EEC epithelial non endometriotic cells. PPIA known as Cyclophilin A, encodes a member of the peptidyl-prolyl cis-trans isomerase (PPIase) family. PPIases catalyze the cis-trans isomerization of proline imidic peptide bonds in oligopeptides and accelerate the folding of proteins. The encoded protein is a cyclosporin binding-protein and may play a role in cyclosporin A-mediated immunosuppression. PPIA induce the leukocyte chemotaxis and contributes to the pathogenesis of inflammation-mediated diseases (109). PPIA increased cell proliferation in liver fluke-associated cholangiocarcinoma cell lines (110). Therefore PPIA might promote inflammation and cell proliferation of hTERT EEC epithelial non endometriotic cells. TUBB3 is involved in Gap junction, melanogenesis and infection mechanisms according to (biocompendium) kegg pathway annotation. TUBB3 is induced by hypoxia and is implicated in drug resistance mechanisms (111, 112). Interestingly, the combination of lipoxin A₄ and estradiol decreased plectin (PLEC) expression in hTERT EEC cells. Plectin is essential for anchoring keratin filament cytoskeleton to the extracellular matrix via hemidesmosomal integrins in basal keratinocytes (113). The down regulation of plectin could contribute to prevent the adhesion of epithelial cells during retrograde menstruation.

As endometriosis is related to estrogen signaling (23), we investigated the response of the different cell lines to estrogen and lipoxin A₄ on a small group of estrogen responsive genes, highly modulated in hTERT EEC cells. These cells are E2 responsive (114). We observed that the response of stomal cell lines was very different from the epithelial cell lines and that the endometriotic epithelial cell line was less responsive to estradiol then the normal epithelial cell line. This is most likely related to the presence and the abundance of Estrogen receptors ERα, ERβ and GPER30 (60) or the transcription factor SF-1. From western blotting carried out in the laboratory we know that ER expression in the other cells lines is much lower than that in hTERT EEC cells. Interestingly, the response of the different cell lines to lipoxin A₄ is similar to those evoked by estradiol, which is expected as lipoxin A₄ has recently been shown to be a ERα agonist (53).
Contextualization
Please refer to figures 25-27 in the results section.

We wished to investigate the physical interactions of the potential biomarkers (figure 25), the treatment biomarkers and the known biomarkers discovered using NCBI annotated genes for endometriosis, EAGLi text mining tool and a list of proteins given by members of the mucosal immunity lab. A compact network including known biomarkers, potential biomarkers and treatment biomarkers was built with the interactome data (87) and visualized using cytoscape (94). Many potential biomarkers are reported to interact with the known biomarkers, indicating that these proteins are implicated in the same biological process. 8 biomarkers were not connected to the dense network. These biomarkers are ENSG00000206503 (HLA-A), HIST1H2AD, ALDH1A3, HSPA2, TUBB6, COMT, COL6A3 and STMN1. ENSG00000206503, HSPA2 and STMN1 are annotated with response for biotic stimulus as biological process gene ontology. COMT is quite interesting because it is up regulated in epithelial cells (12Z versus hTERT EEC) and down regulated in stromal cells (22B versus stromal). This protein is involved in the steroid hormone bio-synthesis pathway, and was related to cell proliferation (115). The COL6A3, Collagen alpha-3(VI) chain was associated with inflammation in adipose tissue of obese patients (116). COL6A3 was down regulated in 22B endometriotic stromal cells compared to normal stromal cells (figure 26), which is surprising, because the inflammation should be more apparent in the endometriotic cell line, but inflammation is also likely implicated in the renewal of the endometrium during menstruation. Stathmin (STMN1), a ubiquitous cytosolic phosphoprotein proposed to function as an intracellular relay integrating regulatory signals of the cellular environment, is involved in the regulation of the microtubule filament system by destabilizing microtubules. It prevents assembly and promotes disassembly of microtubules. STMN1 was linked with cell adhesion and migration in a hepatocellular carcinoma (HCC) cell line (117). STMN1 was down regulated in 22B stromal endometriotic cells compared to stromal normal cells. Aldehyde dehydrogenase family 1 member A3 (ALDH1A3), implicated in Retinoic acid biosynthesis (118), was up regulated in both endometriotic cell lines compared to normal cell line. Retinoic acid is implicated in growth and development (differentiation). ALDH1A3 was down regulated in 12Z epithelial endometriotic cell line by all of the treatments, whereas it was up regulated in 22B stromal endometriotic cell line.

Interestingly most of the potential biomarkers were up regulated in endometriotic cell lines (figure 27), apart from PFN1, HIST1H4A HIST1H2AD in epithelial cells and COL6A3 and FN1 in stromal cells. PFN1 is implicated in actin cytoskeletal regulation and it was down regulated by the different treatments in hTERT EEC non endometriotic cell line. Profilin 1 (PFN1) is down regulated in breast cancer, and was associated with a reduction of both cell-cell and cell-matrix adhesions with a concomitant increase in motility and dramatic scattering of normal human mammary epithelial cells (119). Fibronectin (FN1) was increased in 12Z epithelial cell compared with hTERT EEC epithelial non endometriotic cell, and decreased in 22B compared with normal stromal cells. This suggests that epithelial cells are more active in term of adhesion. All of the treatments are acting against the expression of FN1, which is up regulated in 12Z control versus hTERT control and down regulated in 22B versus stromal (down regulation by any of the treatments in 12Z as and up regulation in 22B).

Fatty acid synthase (FASN) was slightly increased in the stromal endometriotic cell line compared to normal stromal cells, and slightly attenuated in the epithelial endometriotic cell line compared to normal cell line. A down regulation of FASN leads to apoptosis in melanoma cells (120). FASN is a
potential therapeutic target in cancer (121), and was decreased by estradiol and lipoxin A₄ in epithelial endometriotic cells.

Several proteins, notably HSPA2, MAPK1, FLNC and FLNA, implicated in the MAPK signaling pathway are up regulated in 12Z epithelial endometriotic cell line compared to the hTERT EEC normal epithelial cell line (figure 27-28), indicative of potential higher proliferation rates of 12Z epithelial endometriotic cells. HSPA2 is highly down regulated by at least one treatment, which might be due to the usage of unweighted spectral counting, whereas FLNA, FLNB and FLNC are slightly down regulated by estradiol.

Looking carefully at the known biomarkers (figure 28), we can infer the role of stromal and epithelial cells in the disease process. Some of these biomarkers are regulated in a very different manner in stromal cells compared to epithelial cells. We can list CD44, COMT, FN1, SLC2A1, MIF, CFL1, GSTP1 and CTNNB1 in this group. The epithelial endometriotic cells might have a predominant role in glucose transport through SLC2A1, in extra-cellular matrix interactions via CD44 and FN1, in drug metabolism and oxidative stress response through GSTP1, and in steroid hormone metabolism involving COMT. In contrast, stromal endometriotic cells assert a predominant role in phenylalanine and tyrosine metabolism and inflammation via MIF, and a predominant structural role via CFL1 (implicated in cytoskeleton regulation) and CTNNB1 (implicated in cancer, Wnt pathway, infection and tight junction).

On the other hand some known biomarkers exhibit a similar expression pattern in both cases, like PRDX6, ACP1, HLA-A, L1CAM, ICAM1, EGFR, TP53 and BAX. These known biomarkers are slightly up regulated in both endometriotic cell lines and are implicated in adhesion for L1CAM, ICAM and ACP1, proliferation through EGFR, BAX has been implicated in apoptosis and cell cycle in the case of TP53.

Finally we investigated the suppressive or synergic effect of estradiol combined with lipoxin A₄. Unexpectedly, our results indicate that the combination of the treatments has a suppressive effect on keratins KRT6B, KRT14, and KRT16 in 12Z epithelial endometriotic cells whereas it has a synergic effect on 22B stromal endometriotic cells. Interestingly, the combination of both treatments exerts a synergic effect on SPTB2, COMT and HSP72 in 12Z endometriotic epithelial cells and a synergic effect on EGFR in 22B endometriotic stromal cells. This model might be a useful tool for further study of suppressive/synergic effects. It would be interesting to develop this model for multiple treatments and statistical analysis for mass spectrometry data.

**Conclusions and perspectives**

This project based on LC-orbitrap MS/MS discovery experiment and bioinformatics provides new insights into estradiol and lipoxin A₄ signaling in endometriosis. We detected known biomarkers, as well as potential biomarkers and treatment biomarkers to investigate the molecular processes underlying the disease. This information was consistent with the literature for the known biomarkers monitored. These known and new candidate biomarkers could potentially be used to develop new non-invasive diagnostic tests, if their detection is possible in blood or urine. Many signaling pathways were previously discovered but not heretofore linked in this context. However, this data needs further investigation and validation. Mouse experiments are currently being performed to study the effect of lipoxin A₄ on endometriosis in vivo. In parallel, human samples (blood, peritoneal fluid, and
biopsies) are being used. Experiments on 1Z and hTERT EEC cell lines, involving quantitative mass spectrometry by ITRAQ analysis, are in progress.

Other types of mass spectrometry studies would provide additional information on the disease, like metabolomics or lipidomics studies. Indeed, lipid mediators are implicated in the inflammatory/resolution process, which is an important aspect of the disease, and small molecules are essential in hormone signaling, a driving force for disease development and progression.

Acknowledgments

This project was amazingly exciting, due to the diversity of topics implicated. I had the opportunity to work in wet lab for sample preparation, to improve my data analysis and statistical skills using the mass spectrometry data and to discover powerful bioinformatics tools and databases. I’m very grateful to Dr. Geraldine Canny and Dr. Manfredo Quadroni for giving me the opportunity to work on such a fascinating project. Thanks to Dr. Patrice Waridel for his help and advice during this project. Many thanks to Dr. Ilaria Gori, Msc. Rajesh Kumar and Msc. Chiara Pellegrini for their help in cell culture and their expertise on the disease. A special thanks to Dr. Jachen Barblan and Alexandra Potts for their help in sample preparation and Mass spectrometry expertise. Thanks to Dr. Patricia Palagi for giving me the opportunity to be part of the Bioinformatics and proteomics Master program at the University of Geneva. Thanks also to Prof. Ioannis Xenarios, and Anne Niknejad, Dr. Frederic Schütz, Dr. Rico Ruedel for their bioinformatics and statistics advice. I want to express my gratitude to my family for supporting me during my studies.

Bibliography


53. Russell, R., Gori, I., Pellegrini, C., Kumar, R., Achtari, C., and Canny, G. O. Lipoxin A4 is a novel estrogen receptor modulator. FASEB J


65. Wang, H., Alvarez, S., and Hicks, L. M. Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two <i>Chlamydomonas reinhardtii</i> strains of interest for biofuels engineering. *J Proteome Res*


68. Wang, S., Tian, R., Li, L., Figeys, D., and Wang, L. An enhanced chemically defined SILAC culture system for quantitative proteomics study of human embryonic stem cells. *Proteomics* 11, 4040-4046


72. Pichler, P., Kocher, T., Holzmann, J., Mazanek, M., Taus, T., Ammerer, G., and Mechtler, K. Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap. *Anal Chem* 82, 6549-6558


105. Kondo, S., Kubota, S., Mukudai, Y., Nishida, T., Yoshihama, Y., Shirotta, T., Shintani, S., and Takigawa, M. Binding of glyceraldehyde-3-phosphate dehydrogenase to the cis-acting element of structure-anchored repression in ccn2 mRNA. *Biochem Biophys Res Commun* 405, 382-387


---

**Annexes**

**Protein index**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Synonyms</th>
<th>GO annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC7A5</td>
<td>solute carrier family 7 (cationic amino acid transporter, y+ system), member 5 pseudogene 1 [Source:HGNC Symbol;Acc:29458]</td>
<td>4F2LC;8140;D16S469E;E16;ENSG00000103257;ENSP00000261622;LAT1_HUMAN;LAT1_HUMAN;MPE16;NM_003486;NP_003477;Q01650;QBMH63;SLC7A5;hiLAT1;</td>
<td>leukocyte migration</td>
</tr>
<tr>
<td>SLC2A1</td>
<td>solute carrier family 2 (facilitated glucose transporter), member 1 [Source:HGNC Symbol;Acc:11005]</td>
<td>6513;DYT17;DYT18;ENSG00000117394;ENSP000361578;ENSP00000361579;ENSP00000380214;ENSP00000394591;ENSP00000395521;ENSP00000416293;GLUT;GLUT1;GLUT1D5;GTR1_HUMAN;MIC5;NM_001040649;NM_001143963;NM_007099;NP_001035739;NP_00291;NP_009030;NM_024080;P24666;PPAC_HUMAN;</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>ACP1</td>
<td>acid phosphatase 1, soluble [Source:HGNC Symbol;Acc:122]</td>
<td>52;ACP1;ENSG00000143727;ENSP00000272065;ENSP00000272067;ENSP00000384184;ENSP00000384307;ENSP00000385404;ENSP00000389681;ENSP00000408373;ENSP00000408596;ENSP00000410331;ENSP00000411121;HAAP;MGC111030;MGC141895;MGC141896;NM_006516;NP_006507;P11166;PPAC_HUMAN;</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>L1CAM</td>
<td>L1 cell adhesion molecule [Source:HGNC Symbol;Acc:6470]</td>
<td>3897;CAML1;CD171;ENSG00000198910;ENSP00000354572;ENSP00000355380;ENSP00000359072;ENSP00000359074;ENSP00000359075;ENSP00000359077;ENSP00000384902;ENSP00000392524;ENSP00000396079;ENSP00000397972;ENSP00000402407;HAAS;HAS1;L1CAM;L1CAM_HUMAN;MGC111030;NM_001143963;NCAM-1;NM_000425;NM_001137435;NP_000416;NP_001137435;NP_0076493;P24666;PPAC_HUMAN;</td>
<td>axon guidance; leukocyte migration</td>
</tr>
<tr>
<td>KRT6C</td>
<td>keratin 6C [Source:HGNC Symbol;Acc:20406]</td>
<td>286887;ENSG000000170465;ENSP00000252250;ENSP00000405595;K2C6C_HUMAN;K2E;KRT6C;KRT6E;MGC163455</td>
<td>MGC163457;MGC163457;NM_173086;NP_001035739;P32004;SPG1;</td>
</tr>
<tr>
<td>KRT2</td>
<td>keratin 2 [Source:HGNC Symbol;Acc:6439]</td>
<td>3849;C2K2E;ENSG00000172867;ENSP00000310861;K2C2E_HUMAN;K2E;KRT2;KRT2A;KRT2B;KRT2E;KRT2E;MGC116966;MGC116968;NM_000423;NP_000414;P435908;</td>
<td>epidermis development; intermediate filament; intermediate filament cytoskeleton; structural constituent of cytoskeleton</td>
</tr>
</tbody>
</table>
CD44 (CD44_HUMAN)
CD44 molecule (Indian blood group) [Source:HGNC Symbol;Acc:1681]
960;CD44;CD44_HUMAN;CDW44;CSPG8;ECMR-
III;ENSG00000026508;ENSP000000263398;ENSP0-
0000278385;ENSP000000278386;ENSP000000279-
452;ENSP00000309732;ENSP00000353280;ENSP0-
0000389830;ENSP00000391008;ENSP00000393-
233;ENSP00000395953;ENSP00000398099;ENSP-
P00000398632;ENSP00000403990;ENSP000004-
0447;ENSP00000414567;ENSP00000415159;ENSP-
00000415529;HCELL;HUTCH-I;IN;MC56;MDU2;
MDU3;MGC10468;MIC4;NM_000601;P16070;Pgp1;
ERK1 and ERK2 cascade; cell-matrix adhesion; cell-
substrate adhesion; peptidyl-tyrosine modification; peptidyl-tyrosine phosphorylation; positive regulation of intracellular protein kinase cascade; regulation of ERK1 and ERK2 cascade; regulation of peptidyl-tyrosine phosphorylation

BAX (BAX_HUMAN)
BCL2-associated X protein [Source:HGNC Symbol;Acc:959]
581;BAX;BAX_HUMAN;BCL2L4;ENSG000000870-
88;ENSP00000263262;ENSP00000293288;ENSP0-
0000346461;ENSP000037574;ENSP0000389971;ENSP0000-
000426328;NM_004324;NM_138761;NM_138764;
NM_138763;NM_138764;NP_004315;NP_620116;
NP_620119;NP_027882;Q07812;cellular component
disassembly; cellular component disassembly at
cellular level; cellular component disassembly involved in apoptosis

ITGAV (ITA-
V_HUMAN)
integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51) [Source:HGNC Symbol;Acc:6150]
3685;CD51;DKFZp686A08142;ENSG0000001384-
4;ENSP00000036402;ENSP0000389442;ENSP0000-
0042491;ITAV_HUMAN;ITGAV;MSK8;NM_001144999;
NM_001145000;NP_001138471;NP_001138472;
No significantly enriched
annotations

UCHL1 (UCHL1_HUMAN)
ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase) [Source:HGNC Symbol;Acc:12513]
7345;ENSG000000154277;ENSP000000284440;
ENSP000000293288;ENSP0000346461;ENSP0000348871;ENSP000037574;
ENSP0000389971;ENSP0000426328;ENSP0000426895;
NM_004181;NP_001138471;NP_001138472;P06756;
VNRA;axon guidance; cell-matrix adhesion; cell-substrate adhesion; leucocyte migration

MIF (MIF_HUMAN)
macrophage migration inhibitory factor (glycosylation-inhibiting factor) [Source:HGNC Symbol;Acc:7097]
4282;ENSG000000240972;ENSP00000215754;Gli-
F;MIF;MIF_HUMAN;MMIF;NM_002415;NP_00046-
06;P14174;ERK1 and ERK2 cascade; fibroblast proliferation; peptidyl-tyrosine modification; peptidyl-
tyrosine phosphorylation; positive regulation of intracellular protein kinase cascade; regulation of ERK1 and ERK2 cascade; regulation of fibroblast proliferation; regulation of peptidyl-
tyrosine phosphorylation

ICAM1 (ICAM1_HUMAN)
intercellular adhesion molecule 1 [Source:HGNC Symbol;Acc:5344]
3383,BB2;CD54;ENSG00000090339;ENSP000002-
64832;ENSP000003977754;ENSP00000397889;EN-
SP00000413124;ICAM1;ICAM1_HUMAN;NM_00-
2021;NP_000192;P05362;P3.58;
leukocyte migration

KRT16 (K1C16_HUMAN)
keratin 16 [Source:HGNC Symbol;Acc:6423]
3868;CK16;ENSG00000018632;ENSP0000003016-
5;FNEPPK;K16;K1C16_HUMAN;K1CP;KRT16;KRT1-
6A;NM_005557;NP_005548;P08779;
epidermis development; intermediate filament; intermediate filament cytoskeleton; structural constituent of cytoskeleton

LDHB (LDHB_HUMAN)
lactate dehydrogenase B [Source:HGNC Symbol;Acc:6541]
3945;ENSG00000011716;ENSP00000229319;ENS-
P0000037983;ENSP00000379386;ENSP00000398-
015;LDH-H;LDHB;LDHB_HUMAN;NM_001174097;NM_002291;NP_001167568;NP_002291;P07195;TRG-5;
No significantly enriched annotations
<table>
<thead>
<tr>
<th>Protein</th>
<th>Source</th>
<th>Gene Symbol</th>
<th>Gene Acc</th>
<th>Candidate Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN1</td>
<td>FINC_HUMAN</td>
<td>fibronectin 1</td>
<td>3778</td>
<td>cell-substrate adhesion; exocytosis; leukocyte migration; platelet activation; platelet degranulation</td>
</tr>
<tr>
<td>PROX6</td>
<td>PRDX6_HUMAN</td>
<td>peroxiredoxin 6</td>
<td>3675</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>TAGLN2</td>
<td>TAGL2_HUMAN</td>
<td>transgelin 2</td>
<td>11554</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>KRT10</td>
<td>K1C10_HUMAN</td>
<td>keratin 10</td>
<td>6413</td>
<td>epidermis development; intermediate filament cytoskeleton</td>
</tr>
<tr>
<td>SLC3A2</td>
<td>4F2_HUMAN</td>
<td>solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2</td>
<td>11026</td>
<td>leukocyte migration</td>
</tr>
<tr>
<td>KRT14</td>
<td>K1C14_HUMAN</td>
<td>keratin 14</td>
<td>6416</td>
<td>cell junction assembly; cell-substrate junction assembly; epidermis development; hemidesmosome assembly; intermediate filament cytoskeleton; structural constituent of cytoskeleton</td>
</tr>
<tr>
<td>PFN1</td>
<td>PROF1_HUMAN</td>
<td>profilin 1</td>
<td>8881</td>
<td>cytoskeletal protein binding; exocytosis; platelet activation; platelet degranulation</td>
</tr>
<tr>
<td>KRT5</td>
<td>K2CS_HUMAN</td>
<td>keratin 5</td>
<td>6442</td>
<td>cell junction assembly; cell junction organization; cell-substrate junction assembly; epidermis development; hemidesmosome assembly; intermediate filament cytoskeleton; structural constituent of cytoskeleton</td>
</tr>
<tr>
<td>FLNC</td>
<td>FLNC_HUMAN</td>
<td>filamin C, gamma</td>
<td>3756</td>
<td>cell junction assembly; cell junction organization</td>
</tr>
<tr>
<td>TPS3</td>
<td>P53_HUMAN</td>
<td>tumor protein p53</td>
<td>11998</td>
<td>peptidyl-tyrosine modification; peptidyl-tyrosine phosphorylation; phosphatase binding; regulation of peptidyl-tyrosine phosphorylation; response to biotic stimulus</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTP1_HUMAN</td>
<td>glutathione S-transferase pi 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K2C8_HUMAN</td>
<td>keratin 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC107016.2_HUMAN</td>
<td>Keratin, type I cytoskeletal 18 (Cytokeratin-18) (K18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUBB2B_HUMAN</td>
<td>tubulin, beta 2B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENO1_HUMAN</td>
<td>enolase 1, (alpha)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEF1A2_HUMAN</td>
<td>eukaryotic translation elongation factor 1 alpha 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUBA1B_HUMAN</td>
<td>tubulin, alpha 1b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANXA1_HUMAN</td>
<td>annexin A1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTN1_HUMAN</td>
<td>actinin, alpha 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFL1_HUMAN</td>
<td>coflin 1 (non-muscle)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTC1_HUMAN</td>
<td>actin, alpha, cardiac muscle 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTG1_HUMAN</td>
<td>actin, gamma 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source:** HGNC Symbol; Acc: [4638]
<table>
<thead>
<tr>
<th>Gene</th>
<th>Symbol</th>
<th>Description</th>
<th>Enriched Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPIA</td>
<td>PPIA_HUMAN</td>
<td>Peptidylprolyl isomerase A (cyclophilin A)</td>
<td>Exocytosis; leukocyte migration; platelet activation; platelet degranulation; protein folding</td>
</tr>
<tr>
<td>MAPK1</td>
<td>MK01_HUMAN</td>
<td>Mitogen-activated protein kinase 1</td>
<td>ERK1 and ERK2 cascade; axon guidance; phosphatase binding; platelet activation</td>
</tr>
<tr>
<td>EGFR</td>
<td>EGFR_HUMAN</td>
<td>Epidermal growth factor receptor</td>
<td>Axon guidance; cytoskeletal protein binding; peptidyl-tyrosine modification; peptidyl-tyrosine phosphorylation; phosphatase binding; positive regulation of intracellular protein kinase cascade; regulation of peptidyl-tyrosine phosphorylation</td>
</tr>
<tr>
<td>PLEC</td>
<td>PLEC_HUMAN</td>
<td>Plectin</td>
<td>Cell junction assembly; cell junction organization; cellular component disassembly; cellular component disassembly at cellular level; cellular component disassembly involved in apoptosis; hemidesmosome assembly; intermediate filament cytokinin</td>
</tr>
<tr>
<td>TUBA1C</td>
<td>TBA1C_HUMAN</td>
<td>Tubulin, alpha 1c</td>
<td>'de novo' posttranslational protein folding; 'de novo' protein folding; protein folding</td>
</tr>
<tr>
<td>TUBB3</td>
<td>TBB3_HUMAN</td>
<td>Melanocortin 1 receptor (alpha melanocyte stimulating hormone receptor)</td>
<td>'de novo' posttranslational protein folding; 'de novo' protein folding; axon guidance; positive regulation of intracellular protein kinase cascade; protein folding</td>
</tr>
<tr>
<td>FASN</td>
<td>FAS_HUMAN</td>
<td>Fatty acid synthase</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>FLNA</td>
<td>FLNA_HUMAN</td>
<td>Filamin A</td>
<td>Cell junction assembly; cell junction organization; cytoskeletal protein binding; exocytosis; platelet activation; platelet degranulation; positive regulation of intracellular protein kinase cascade</td>
</tr>
<tr>
<td>SPTBN1</td>
<td>SPTB2_HUMAN</td>
<td>Spectrin, beta, non-erythrocytic 1</td>
<td>Axon guidance</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Description</td>
<td>Enriched Annotations</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>CTNNB1</td>
<td>Catenin (cadherin-associated protein), beta 1, 88kDa</td>
<td>cell junction assembly; cell junction organization; cell-matrix adhesion; cell-substrate adhesion; cellular component disassembly; cellular component disassembly at cellular level; cellular component disassembly involved in apoptosis; epidermis development; fibroblast proliferation; phosphatase binding; positive regulation of intracellular protein kinase cascade; regulation of fibroblast proliferation</td>
<td></td>
</tr>
<tr>
<td>VIM</td>
<td>Vimentin</td>
<td>cellular component disassembly; cellular component disassembly at cellular level; cellular component disassembly involved in apoptosis; intermediate filament; intermediate filament cytoskeleton; structural constituent of cytoskeleton</td>
<td></td>
</tr>
<tr>
<td>SPTAN1</td>
<td>Spectrin, alpha, non-erythrocytic 1 (alpha-fodrin)</td>
<td>axon guidance; cellular component disassembly at cellular level; cellular component disassembly involved in apoptosis; structural constituent of cytoskeleton</td>
<td></td>
</tr>
<tr>
<td>GCN1L</td>
<td>GCN1 general control of amino-acid synthesis 1-like 1 (yeast)</td>
<td>No significantly enriched annotations</td>
<td></td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>No significantly enriched annotations</td>
<td></td>
</tr>
<tr>
<td>MYH9</td>
<td>Myosin, heavy chain 9, non-muscle</td>
<td>axon guidance; cytoskeletal protein binding; leukocyte migration</td>
<td></td>
</tr>
<tr>
<td>COL6A3</td>
<td>Collagen, type VI, alpha 3</td>
<td>axon guidance</td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Description</td>
<td>Enriched Annotations</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>COMT (COMT_HUMAN)</td>
<td>catechol-O-methyltransferase</td>
<td>1312;COMT;COMT_HUMAN;ENSG00000093010;ENSP00000207636;ENSP00000354511;ENSP00000383966;ENSP00000384654;ENSP00000385150;ENSP00000385831;ENSP00000385917;ENSP00000387695;ENSP00000403958;ENSP00000410938;ENSP00000416778;NM_000754;NM_001135161;NM_001135162;NM_007310;NP_000745;NP_001128633;NP_001128634;NP_009294;P21964;</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>STMN1 (STMN1_HUMAN)</td>
<td>stathmin 1</td>
<td>3925;C1orf215;ENSG000000117632;ENSP00000035031;ENSP000000363409;ENSP000000363410;ENSP000000382632;ENSP000000382633;ENSP00000387858;ENSP00000407576;ENSP00000410452;FLJ32206;LAP18;Lag;MGC138869;MGC138870;NM_001145454;NM_005563;NM_003430;NP_01138926;NP_005554;NP_918944;NP_981946;OP18;P16949;PP19;PR22;STMN1;STMN1_HUMAN;</td>
<td>cellular component disassembly; cellular component disassembly at cellular level; cytoskeletal protein binding; response to biotic stimulus</td>
</tr>
<tr>
<td>ENSG00000206503 (1A03_HUMAN)</td>
<td>HLA class I histocompatibility antigen, A-11 alpha chain Precursor (MHC class I antigen A*11)</td>
<td>1A03_HUMAN;1A11_HUMAN;1A30_HUMAN;3105;ENSG00000206503;ENSP00000348012;ENSP00000365998;ENSP00000366002;ENSP00000366005;ENSP00000379873;FLJ26655;HLAA;HLA-A;HLA-A1;NM_002116;NP_002107;P13746;P16188;</td>
<td>response to biotic stimulus</td>
</tr>
<tr>
<td>ALDH1A3 (ALDH1A3_HUMAN)</td>
<td>aldehyde dehydrogenase 1 family, member A3</td>
<td>220;ALDH1A3_HUMAN;ALDH1A3;ALDH1A6;ALDH6;ENSG000000184254;ENSP00000332256;ENSP00000343924;ENSP00000389573;ENSP00000395353;NM_000693;NP_000684;P47895;RALDH3;</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>HSPA2 (HSP72_HUMAN)</td>
<td>heat shock 70kDa protein 2</td>
<td>3306;ENSG00000126803;ENSP00000242707;ENSP00000378199;HSP70-3;HSP72_HUMAN;HSPA2;NM_021799;NP_068814;P54562;</td>
<td>response to biotic stimulus</td>
</tr>
<tr>
<td>HIST1H2AD (H2A1D_HUMAN)</td>
<td>histone cluster 1, H2ad</td>
<td>3013;ENSG000000196866;ENSP000000341049;H2A.3;H2A/g/H2A1D_HUMAN;H2AF7;HIST1H2AD;NM_021805;NP_066409;P20671;</td>
<td>No significantly enriched annotations</td>
</tr>
<tr>
<td>TUBB6 (TUBB6_HUMAN)</td>
<td>tubulin, beta 6</td>
<td>84617;ENSG00000176614;ENSP00000231697;ENSP00000389084;HsT1601;MGC132410;MGC4083;NM_032525;NP_0115914;Q9BUF5;TUBB6_HUMAN;TUBB-5;TUBB6;</td>
<td>'de novo' posttranslational protein folding; 'de novo' protein folding; protein folding</td>
</tr>
</tbody>
</table>