

# Identify and characterise stem cell population of the human endometrium

A thesis submitted for the degree of Masters (by research) of Science by  
Marina Ellie Afami

Lancaster, 2014



Shoot for the moon, even if you miss you will land among the stars.

## ***Acknowledgments***

For this project I would like to thank my supervisors Dr Nigel Fullwood and Professor Frank Martin for giving me this great opportunity to work with them. Thank you for all the help, the guidance and all the knowledge you have provided me. It has not been an easy way and all your patience and understanding is much appreciated.

I would like to also thank the support staff at Lancaster Environment Centre and the colleagues I have worked with during this course especially Debra Hurst, Holly Butler, Becky Strong, Kelly Heys, Alana Mitchel and Georgios Theophilou. Thank you for all your patience, time and help. I am thankful for all the moral support and encouragement my special friends, Eleni Chrysostomou, Maria Michael, Victoria Eugenia Martinez Miguel, Kirsten Snidjers and Antonis Antoniou, gave me in the whole duration of this project. I wish all the best to everyone.

I would like to dedicate this work in the loving memory of close family members who passed away in the last few years: my grandmother Ellie Afami, a sweet and honest person who suffered from Gullain-Barre syndrome and this was one of the reasons which made me follow this certain path of my studies, my grandmother Eleni Tsiaklis, who was an inspiration for whatever I have done so far and her wise words will always be my guidance in life, my uncle Vasos Hadjiyerou, who was an intellectual teacher always willing to give the best of advice, my aunty Maria Tsiaklis, who despite her chronic lung disease was grateful for how life had treated her and never lost the power of her beliefs, and my uncle Kyriakos Christofi, who was in love with the beauties of nature both on earth and underwater.

Last but not least, would like to thank my parents, Andreas Afamis and Christa Afami, and my sisters, Eleni Afami and Rena Afami, for their amazing support and for all the values in life they have taught me so far. They will be the greatest motivation to do my best and achieve my goals and I will always want to make them proud for the person I will become.

I appreciate each and everyone.

## ***Declaration***

I declare that this thesis is my own work and has not been submitted, in part or in whole for the award of a higher degree at this or any other institution.

# Contents

Title page.....	I
Acknowledgments.....	III
Declaration.....	IV
Table of Contents.....	V
Abstract.....	VIII

---

## Chapter 1 - General Introduction.....1

### 1.1 Setting the scene.....2

1.1.1 The female reproductive system.....	2
1.1.2 The menstrual cycle.....	5
1.1.3 Gynaecological diseases.....	8

### 1.2 Stem cells.....11

1.2.1 Biology of stem cells.....	11
1.2.2 Evidence of stem/progenitor cells in human endometrium.....	14
1.2.3 Role of endometrial stem cells in endometrial disorders.....	20
1.2.4 Regenerative medicine.....	24

### 1.3 Background on applied methods.....29

1.3.1 Scanning electron microscopy.....	29
1.3.2 Biospectroscopy.....	31
1.3.3 Spectral pre-processing.....	33
1.3.4 Computational analysis.....	34

### 1.4 Aims and objectives.....36

---

## Chapter 2 - Materials and Methods.....37

### 2.1 Samples.....38

### 2.2 Scanning Electron Microscopy.....38

<b>2.3 Biospectroscopy.....</b>	<b>43</b>
<b>2.4 Light Microscopy.....</b>	<b>46</b>

---

**Chapter 3 - Results.....48**

<b>3.1 Scanning Electron Microscopy.....</b>	<b>49</b>
3.1.1 Scanning Elctron Photomicrographs.....	49
3.1.2 Image Analysis.....	74
<b>3.2 FTIR spectroscopy.....</b>	<b>82</b>
3.2.1 Basal Cells Vs Basal Cells.....	85
3.2.2 Luminal Cells Vs Luminal Cells.....	113
3.2.3 Stromal Cells Vs Stromal Cells.....	141
3.2.4 Basal Cells Vs Luminal Cells Vs Stromal Cells.....	169
<b>3.3 Raman spectroscopy.....</b>	<b>190</b>
3.3.1 Basal Cells Vs Basal Cells.....	192
3.3.2 Luminal Cells Vs Luminal Cells.....	221
3.3.3 Stromal Cells Vs Stromal Cells.....	249
3.3.4 Basal Cells Vs Luminal Cells Vs Stromal Cells.....	277
<b>3.4 Light Microscopy.....</b>	<b>299</b>

---

**Chapter 4 - Discussion.....307**

<b>4.1 Scanning Electron Microscopy.....</b>	<b>308</b>
4.1.1 Scanning Electron Photomicrographs.....	308
4.1.2 Image Analysis.....	315
<b>4.2 FTIR spectroscopy.....</b>	<b>317</b>
4.2.1 Basal Cells.....	317
4.2.2 Luminal Cells.....	319
4.2.3 Stromal Cells.....	320
4.2.4 Basal Cells Vs Luminal Cells Vs Stromal Cells.....	321
4.2.5 Applications of Infrared Spectroscopy.....	325

<b>4.3 Raman spectroscopy.....</b>	<b>326</b>
4.3.1 Basal Cells.....	326
4.3.2 Luminal Cells.....	327
4.3.3 Stromal Cells.....	328
4.3.4 Basal Cells Vs Luminal Cells Vs Stromal Cells.....	329
4.3.5 Applications of Infrared Spectroscopy.....	331
<b>4.4 Conclusion.....</b>	<b>332</b>

---

<b>References.....</b>	<b>335</b>
------------------------	------------

## Abstract

The human endometrium is a highly regenerative tissue that experiences functional and structural changes during each menstrual cycle in order to provide a favourable environment for implantation of the embryo. Underpinning this cycle must be a stem cell population. The aim of this project was to characterise this stem cell population *in situ* in post-menopausal endometrium. To this end, Scanning electron microscopy, image analysis and biospectroscopy techniques, Fourier transform Infrared spectroscopy and Raman spectroscopy, were employed.

In a menstruating endometrium, the functional layer is sloughed off during each cycle whilst the basal layer, which contains the potential for endometrial regeneration remains intact. Cyclical and functional changes of the endometrium are regulated by the production of oestrogen and progesterone during the ovarian cycle. Post-menopausal endometrium is thinner and is primarily composed of the basal layer. It is inactive and undergoes apoptotic changes however it retains its regenerative capacity to respond to exogenous hormones. The current evidence supports the hypothesis that epithelial stem/progenitor cells are located in the endometrial crypts, which probably reach the basal layer and migrate to the functional layer to regenerate the epithelial lining of the endometrial glands whilst stromal stem/progenitor cells are located near blood vessels in the functional and basal layer and are responsible for restoring the lost stroma surrounding the endometrial glands.

Results from Scanning electron microscopy revealed the different architecture of the endometrial surface in the samples examined. Variations were noticed in cellular morphology as well as in formation of ciliated cells and pinopode-like structures. It was also observed that cells surrounding the endometrial crypts had a more elongated shape relative to the cells away from the crypts which exhibited a more spherical shape.

To the best of our knowledge, no prior studies have investigated endometrial stem cell populations by means of biospectroscopy techniques. Spectra were collected from epithelial cells at the base and



the lumen of endometrial glands as well as from cells in the surrounding stroma. Pre-processing and subsequent multivariate analysis were applied on the derived spectra to examine segregation between the three populations of cells as well as to characterise their biochemical composition in the hope of identifying spectral regions that discriminate the cell populations and thus support the theory about the putative location of endometrial stem/progenitor cells. Unfortunately, we were not successful in achieving our objectives; however a degree of separation was observed between the different cell types with epithelial cells at the lumen of the glands being the most dissimilar whilst cells at the base of the glands and the surrounding stroma were more similar with respect to each other. Our work though can be used to develop new approaches to further study and provide more insights about stem cell populations in the human endometrium.



# CHAPTER 1

## General Introduction

## **1.1 Setting the scene**

### 1.1.1 The female reproductive system

The main function of the female reproductive system is to produce ova which upon fertilisation will form an embryo. Secretion of hormones provides a favourable environment for embryo implantation and its subsequent growth into a fetus. The system undergoes several changes to support pregnancy but after parturition the original conditions are established. Despite the pregnancy period, structural as well as functional changes occur during the life of a female i.e. from childhood to puberty to menopause under the influence of hormones. Most of these changes take place in the menstrual cycle during the reproductive years of a female.

Three main structural units constitute the female reproductive system based on their function; the ovaries, the genital tract (fallopian tubes, uterus vagina) and the breasts (Young et al., 2006). The ovaries are the site of oogenesis and produce oestrogen and progesterone in a cyclical manner under the influence of luteinising hormone (LH) and follicle stimulating hormone (FSH) via a feedback mechanism (Young et al., 2006). The fallopian tubes are the site of ova release and fertilisation. The uterus is the site of embryo implantation and experiences the most alterations due to ovarian hormones and to support pregnancy. At birth, the baby passes out of the body through the vagina. The breasts develop during puberty and are the site of lactation during pregnancy.

#### *Uterus*

The uterus is a flattened pear-shaped muscular organ and it is anatomically divided into the body, including the fundus at its upper part, and the cervix which opens into the vagina (Fig.1.1). The endometrium is the mucosal lining of the uterine wall whilst the myometrium constitutes the outer thick muscular layer. The endometrium undergoes cyclical changes during the menstrual cycle to

establish conditions for embryo implantation in which case both layers grow and differentiate to support pregnancy (Ross et al., 1995). After menopause the layers become atrophic and thinner.

### *The endometrium*

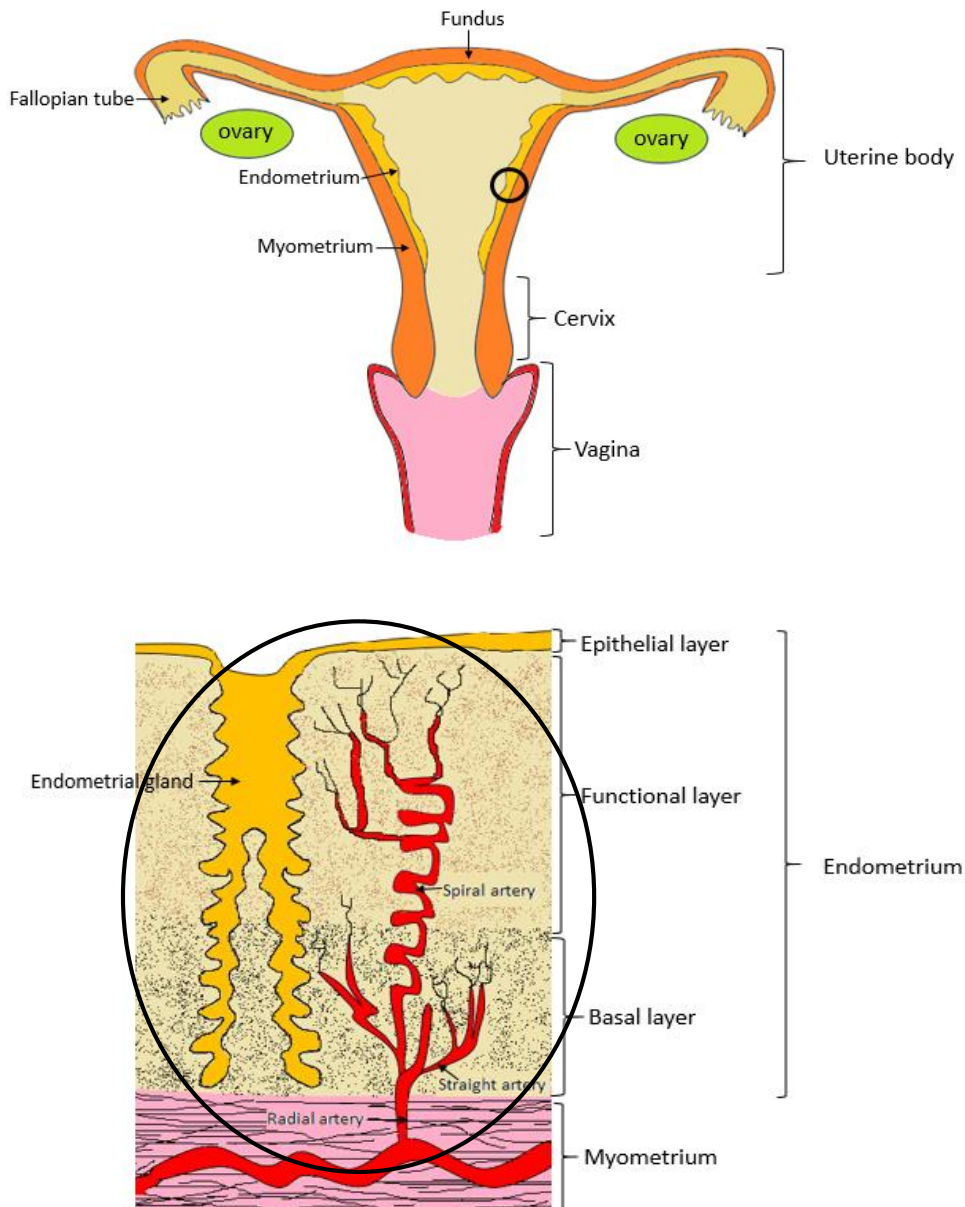
The endometrium is separated into two main layers; the stratum basale/basal layer and the stratum functionale/ functional layer (Fig.1.1.1.1). The latter is sloughed off in each menstrual cycle whilst the first remains intact and provides the tools for regeneration of the lost tissue. During the reproductive years of a female the thickness of the endometrium varies in thickness from 1 to 6 mm (Ross et al., 1995) due to cyclical functional and structural changes. Simple columnar epithelial cells line the endometrial surface along with some secretory and ciliated cells. Mesenchymal cells are the main cell type making the connective tissue of the endometrial stroma (Spencer et al., 2005). Invagination of the stroma results in the formation of endometrial glands.

In postmenopausal women the endometrium experiences profound changes as well; it becomes atrophic and thin consisting only of the basal layer, the glands are made up of cuboidal or low columnar epithelial cells and become inactive, the stroma loses its mitotic activity and contains more collagen fibres (Young et al., 2006).

The radial artery in the myometrium branches into the basal layer as a straight artery which becomes highly coiled and spiral as it enters the functional layer. The arterioles of the spiral arteries form a rich capillary bed at the endometrial surface (Ross et al., 1995). Spiral arteries are sensitive to fluctuating levels of oestrogen and progesterone during the menstrual cycle so they also proliferate and degenerate resulting into bleeding. So the menstrual flow contains blood from the ruptured vasculature along with the shed stromal and epithelial cells from the functional layer.

The endometrium of the cervix though is a lot different from the endometrium of the uterine body. No tissue is lost during the menstrual cycle but the ovarian hormones influence secretion of mucus from the tall columnar epithelial cells in such a way as to provide optimal conditions for sperm

migration and subsequent fertilisation of the egg (Ross et al., 1995). Whereas glycogen secretions in the rest of the endometrium are responsible for the nutrition of the implanted embryo until the placenta is formed (Young et al., 2006).



**Figure 1.1:** Schematic drawing of the uterus and a magnification of the uterine wall showing the myometrium and the structure of the endometrium (reproduced from Embryology.ch).

### 1.1.2 The menstrual cycle

The cyclical morphological and functional changes of the endometrial glands and stroma happen every 28 days, on average, during the reproductive years of a female. The cycle is orchestrated with follicular development since it is regulated by the ovarian hormones. According to the observed changes the cycle is divided into three successive phases; menstruation, proliferation and secretion.

The menstrual phase starts on the first day of menstruation and lasts for about five days. If neither fertilisation nor implantation take place, the corpus luteum regresses and consequently there is a rapid fall in levels of progesterone and oestrogen so the endometrial thickness cannot be maintained. The walls of the spiral arteries contract and consequently the blood flow to the functional layer decreases. As a result it becomes ischemic, secretion from the endometrial glands ceases and the stroma becomes less edematous (Ross et al., 1995). Eventually the vasculature ruptures and the menstrual flow contains blood and the sloughed off and desquamated stromal and epithelial cells. The remnants at the surface of the undisturbed basal layer will restore the lost tissue. This depends on the new cycle of follicular development whereby the secreted oestrogen stimulates the proliferation and migration of endometrial cells (Young et al., 2006).

The proliferative phase starts and ends with ovulation (mid-cycle) due to the effects of the increasingly oestrogen production during the development of ovarian follicles (Young et al., 2006). The rapid proliferation of epithelial, stromal and endothelial cells re-establish the glands, connective tissue and vasculature respectively. By the end of this phase, which is just after ovulation, the glands start producing glycogen and the endometrium is about 3 mm thick.

Events occurring in the secretory phase are primarily under the regulation of progesterone which is produced by the active corpus luteum at the end of ovulation. Epithelial cells get hypertrophic, instead of undergoing mitosis, the glands enlarge and are full with secretory products. Spiral arteries also grow to reach the endometrial surface. Both oestrogen and progesterone trigger the decidualisation of

stromal cells which get rich in glycogen and fat. By the end of this phase the endometrium is edematous and at its maximum thickness providing the necessary conditions for implantation.

Restoration of the epithelial surface at the wound areas begins at the same time with menstrual breakdown (Henriet et al., 2009; Maybin et al., 2009) and it only takes about 48 hours for the whole process of re-epithelisation to finish (Ludwig et al., 1990; Okulicz et al., 1998).

#### *Hormonal regulation in menstruating endometrium*

Fluctuating levels of oestrogen and progesterone affect primarily regeneration processes in the functional layer rather the basal layer since the latter is not that sensitive to hormonal changes (Brenner et al., 2003; Slayden et al., 2004). In the proliferative phase, once re-epithelisation is completed, increasing levels of oestrogen trigger the rapid proliferation of epithelial and stromal cells for glandular and stroma regeneration via activation of oestrogen receptor- $\alpha$  (ER $\alpha$ ) and progesterone receptor (PR) (Padykula et al., 1989). The vasculature is also regenerated to ensure oxygen and nutrient supply to the growing tissue (Gargett and Rogers, 2001). Interaction of oestrogen with oestrogen ER $\alpha$  on stromal cells and the subsequent production of growth factors result in the indirect regulation of proliferation of endometrial epithelial cells (Cooke et al., 1997; Kurita et al., 2005).

Wnt signalling pathway, which itself is regulated by oestrogen, is also involved in regulating proliferation of epithelial cells (van der Horst et al., 2011). This can be supported by the differential expression of Wnt signalling molecules between epithelial cells in the post- and pre-menopausal endometrium (Nguyen et al., 2012). Also the expression of the negative regulator AXIN2 of this pathway (Cong and Varmus, 2004) in the epithelium of the basal layer in post- and pre-menopausal endometrium suggests that Wnt signalling restricts its proliferative effect within the functional layer (Nguyen et al., 2012).

With the formation of corpus luteum after ovulation and the subsequent release of progesterone, differentiation is mediated with the beginning of the secretory phase. Production of progesterone



negatively regulates responses of the functional layer, but not of the basal layer, to oestrogen via suppressing expression of ER $\alpha$  and PR and thus promoting differentiation (Jabbour et al., 2006). However, PR of stromal cells are still activated and cells continue to proliferate and start forming decidualised cells (Gargett et al., 2012). In the meantime, differentiation of endometrial glands results in glycogen accumulation preparing the endometrium for implantation (Spencer et al., 2005). If implantation does not take place, decidualized stromal cells undergo apoptosis due to cessation of progesterone production as the corpus luteum regresses (Gargett et al., 2012). A sequence of events results in menstruation and the subsequent shedding of the functionalis and the process of endometrial regeneration starts once again (Salamonsen, 2003).

#### *Postmenopausal endometrium and response to hormones*

The loss of oestrogen-sensitive cells by apoptosis in the functional layer and the low circulating levels of oestrogen account for the atrophic and inactive status of postmenopausal endometrium (Gargett et al., 2012). However it has the potential to regenerate and gain its normal thickness and function under the influence of hormonal therapies (Ferenczy et al., 1991; Ettinger et al., 1997) and be able to support pregnancy via IVF (Paulson et al., 2002). This response of postmenopausal endometrium to exogenous sex steroid hormones is brought about by the expression of oestrogen receptors (Sauer et al., 1993; Klaassens et al., 2006). Activation of these receptors triggers proliferation and differentiation of epithelial, stromal and endothelial cells to re-establish the endometrium (Gargett et al., 2012). This property of postmenopausal endometrium can be supported by the fact that genes expressed in epithelial cells are similar to genes expressed in the pre-menopausal (Nguyen et al., 2012) and menstrual (Gaide Chevronnay et al., 2009) endometrial epithelium in the basal layer.

### 1.1.3 Gynaecological diseases

#### *Endometrial conditions*

Endometriosis, endometrial carcinoma, endometrial hyperplasia, adenomyosis, fibroids and cysts are the most common pathologies of the endometrium that lead to its abnormal growth or spread and may result in infertility, chronic pelvic pain, abnormal menstrual cycle and vaginal bleeding in postmenopausal females along with other symptoms.

Endometriosis is defined as the abnormal growth of endometrial tissue at areas outside the uterus such as the fallopian tubes, vagina and intestines. These endometrial 'implants' are responsive to ovarian hormones and get thicker during the menstrual cycle and bleed. However, there is no route available to release the blood so these areas get irritated and sore and eventually scar tissue or cysts are formed. Endometrial implants block the fallopian tubes and decrease the chances of conception and successful pregnancy. Available therapies include medication to decrease inflammation or the stimulation by steroid hormones and the surgical removal of the implants. Endometriosis affects millions of women. Some of the risk factors that make women prone to develop endometriosis are family history, never been pregnant, frequent and long periods and starting menstruating at an early age.

Endometrial hyperplasia is defined as the increased proliferation of endometrial glands relative to the stroma. This proliferation results in glands varying in size and shape. According to the architecture and cytological presentation of endometrial glands, hyperplasia is classified into simple, complex and atypical with the latter being more prone to develop into malignancy. The main cause of hyperplasia is the prolonged oestrogen stimulation due to increased oestrogen production or being on oestrogen replacement therapy, obesity and anovulation.

Endometrial carcinoma is the most common invasive cancer of the female genital tract and accounts for 7% of all invasive cancers in women. It occurs more commonly in peri & postmenopausal women

with a peak incidence age of 55 to 65 years old. According to the pathology endometrial carcinomas are classified into type I (the most common) and type II which can be further classified into adenocarcinoma, adenosquamous carcinoma, serous carcinoma and clear cells carcinoma according to their histological appearance. Along with the risk factors of endometrial hyperplasia, infertility, endometrial atrophy and family history increase the chances of cancer development.

In adenomyosis endometrial glands and/or stroma from the basal layer invade the myometrium. It results in the expansion of the uterine wall and upon examination several small cysts are observed. The main cause is unknown but it is associated with smooth muscle hyperplasia (Gargett, 2007). Adenomyosis affects 15-20% of women and is more frequently seen in women with endometriosis.

### *Infertility*

Infertility is a complex and multidimensional health issue that affects 8-10% of couples globally. It is defined as the inability of a woman, aged over than 35 years old, to get pregnant whilst having normal sexual intercourse during a period of 6 months to a year without any means of birth control (Roupa et al., 2009). The problem could be in either the male or the female however it is discovered only after marriage and in some cases the causes may be unknown.

Concerning the females, both the incident and causes of infertility are associated with geographical differences (Roupa et al., 2009). As mentioned above, endometrial conditions have a negative effect on female fertility. Other health related causes include problems with the fallopian tubes, disorders of the menstrual cycle, ovarian failure, vagina and uterus problems. Lifestyle factors considered as causes of infertility include smoking, obesity, sexual disorders, stress and deciding to have a child at an older age. The latter can be explained by the fact that women have high career goals and use contraceptive pills to delay pregnancy until they are settled in the professional arena. The consequences of giving birth at an older age are the decreased reproductive capacity whilst the chances of having a miscarriage and chromosomal abnormalities increase.

The problem of infertility has been addressed by the development of assisted reproductive technology (ART). In the case of male infertility the method of intracytoplasmic sperm injection is employed whilst in vitro fertilisation is used to overcome male and female infertility. However not all couples use the approach of ART to treat their problem because they do not like the idea or because they cannot afford it. Many successful pregnancies have been achieved via ART however it can have adverse outcomes for women and infants.

Assisted reproduction attempts to get pregnant may fail as well especially if the endometrium, the main tissue that supports pregnancy, does not function properly. Surrogacy is an option to overcome this problem or it is often for couples to consider adoption. Research in medicine and science is ongoing and has given many solutions to health problems and the research on stem cells for regenerative medicine is evolving. What if the engineering of endometrial tissue was a solution not only for infertility but also for the proliferative conditions mentioned above? Several studies have been carried out that provide evidence for the existence of stem cells in the human endometrium however the location of this stem cell population and its niche is yet to be defined.

## 1.2 Stem cells

### 1.2.1 Biology of stem cells

Stem cells exist as soon as a zygote is formed and at this stage they are known as embryonic stem cells. When a female gamete is fertilised a zygote is formed whose cells are totipotent which means that they can produce all the cell types in the embryo and the extraembryonic tissues (Eckfeldt et al., 2005). Division and maturation of the zygote result in the formation of an embryo. Human embryonic stem cells are found in the inner cell mass of the blastocyst and since they are more mature than the cells of the zygote they are pluripotent with the ability to differentiate into all derivatives of the germ layers (ectoderm, mesoderm and endoderm) and also trophoectoderm (Gage, 2000; Trounson, 2006). However this differentiation potential gets restricted during the process of embryonic development (Eckfeldt et al., 2005). Despite their pluripotency, another property that distinguishes embryonic stem cells is their ability to propagate themselves indefinitely especially when under defined conditions.

Multipotential somatic stem cells, also known as adult stem cells, are the progeny of endometrial stem cells and reside in many adult tissues and organs. They are in an undifferentiated state until triggered to differentiate into the types of cells making up that tissue. Along with the differentiation potential, self-renewability and high proliferative potential are functional characteristics specific for stem cells (Morrison et al., 1997; Weissman, 2002).

Stem cells undergo asymmetric cell division whereby they produce either identical daughter cells to maintain the stem cell pool within tissues, or progenitor cells (Gargett, 2007). Proliferation of progenitor cells produces transit amplifying cells which further proliferate and differentiate into terminally differentiated functional cells which can no longer proliferate (Fuchs et al., 2004; Chan et al., 2004).

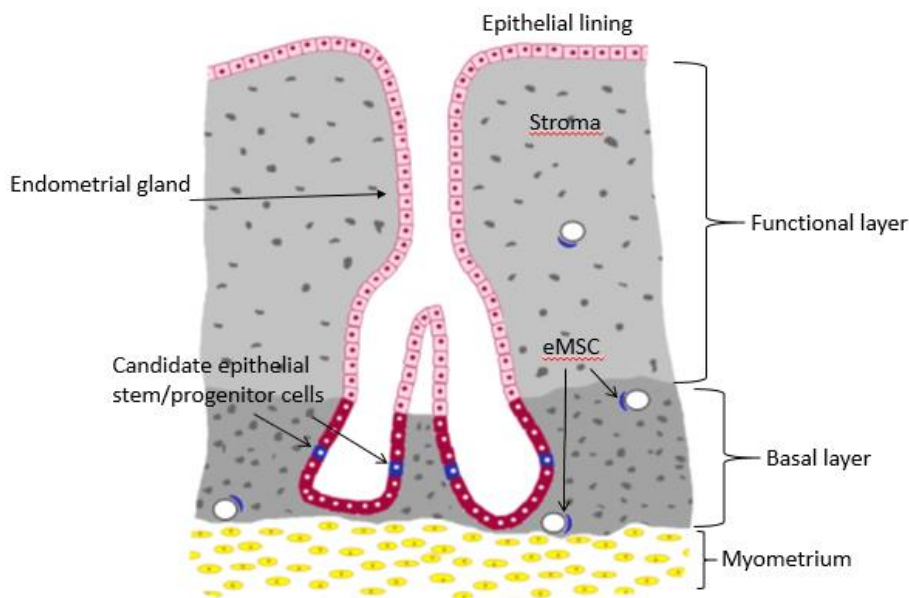
The functions of adult stem cells are regulated by their microenvironment niche (Schofield, 1978). Within this niche, stem cells are in close communication with each other, the surrounding differentiated tissue and the extracellular matrix (Li and Xie, 2005). Molecules secreted by niche cells initiate signalling pathways whereby resident stem cells either remain in a dormant state or undergo cellular division (Fuchs et al., 2004) or in the case of injury and apoptosis they are induced to proliferate and differentiate to replace the lost tissue (Moore and Lemischka, 2006). Thus a balance is achieved between self-renewal and differentiation which is necessary to maintain tissue homeostasis and avoid tumour growth (Shostak, 2006).

Sometimes the need to replace the lost cells can be sensed by stem cells of other lineages whose location allows them to circulate within the body. Plasticity is the term given to the ability of adult stem cells to adapt to changes of their present extracellular environment and transdifferentiate into cells of lineages other than the tissue they normally reside and consequently acquiring all the characteristics of the new cell type (Gargett, 2007; Tosh and Slack, 2002; Wagers and Weissman, 2004).

An important breakthrough in stem cell research was the achievement to generically reprogramme differentiated human somatic cells and induce characteristics similar to those of embryonic stem cells so these cells are called as induced pluripotent stem cells (Takahashi et al., 2007). These cells are modified to express genes and transcription factors (Oct3/4, Sox2, c-Myc, and Klf4) which are responsible for the distinguishing properties of endometrial stem cells. However induced pluripotent stem cells are rather pluripotent.

The human body has several tissues that undergo continuous and rapid turnover such as the skin and the intestine. Their ability to regenerate is due to the presence of an adult stem cell population whose niche has been identified and characterised; stem cells of the skin are located at the bulge of the hair follicle and at the interfollicular regions of the epidermis (Morris et al., 2004; Ito et al., 2005) whilst stem cells in the intestine are located just above the Paneth cells at the loop of the microvilli (Sancho

et al., 2004). A stem cell population underpins the mechanisms of structural and functional changes taking place in endometrial regeneration during each menstrual cycle. Definitive identification of this pool and niche which give rise to progenitor/stem cells is yet to be achieved. It has been long hypothesised and generally accepted that adult stem/progenitor cells reside in the basal layer since it remains intact during the menstrual cycle. Specifically the loop and base of endometrial glands in the basal layer are thought to be the location of epithelial stem cells whilst endometrial mesenchymal stem/stromal cells (eMSC) are located near blood vessels (Fig. 1.2). Stem cells persist in the atrophic postmenopausal endometrium (Gargett, 2007) since it is mainly comprised by the basal layer. Examination of menstrual debris remaining in the uterine cavity, provided evidence for the presence of stem cells in the functional layer and their contribution to endometrial regeneration (Maruyama et al., 2010) based on markers specific for endometrial stem/progenitor cells (Schwab and Gargett, 2007). In addition, stem cells from the bone marrow can sometimes participate in endometrial regeneration (Du and Taylor, 2010).



**Figure 1.2:** Schematic drawing of the hypothesised location of candidate stem/progenitor cells in human endometrium. Epithelial progenitor cells are located at the base of glands in the basal layer and endometrial MSC-like cells are located near blood vessels in the functional and basal layer (reproduced from Gargett and Masuda, 2010).

### 1.2.2 Evidence for stem/progenitor cells in human endometrium

#### *Colony forming unit (CFU) activity*

Functional studies measuring the ability of endometrial cells to produce clones provided the first evidence for the existence of stem/progenitor cells in human endometrium (Gargett, 2004; Schwab et al., 2005). Single cell suspensions of epithelial (EpCAM<sup>+</sup>) and stromal (EpCAM<sup>-</sup>) cells were freshly isolated and purified from hysterectomised endometrial tissue of menstruating women (Chan et al., 2004). Both types of cells, 0.22% and 1.25% of epithelial and stromal cells respectively, produced large and small CFUs (Chan et al., 2004). Clonogenic activity of stromal and epithelial cells was also observed when isolated from post- and peri-menopausal women and no differences in their frequency was found when compared with proliferative and secretory endometrium (Gargett, 2006; Schwab et al., 2005). Even though large colonies were rare, they showed a significant capacity to self-renew *in vitro* and a high proliferative potential whereas in the small loose colonies these two functional properties were limited (Gargett et al., 2009) and displayed no differentiation (Gargett et al., 2012). When cultured in appropriate conditions, large CFUs of epithelial cells differentiated into cytokeratin<sup>+</sup> gland-like structures and large CFUs of stromal cells displayed a multipotency by differentiating into adipocytes, osteoblasts, chondrocytes and smooth muscle cells (Gargett et al., 2012). It was also observed that large stromal CFUs expressed surface markers specific for the phenotype of mesenchymal stem cells (MSC) (Dominici et al., 2006). Plasticity of endometrial stromal cells was displayed by their ability to differentiate into lineages of mesodermal, ectodermal and endodermal origin (Wolff et al., 2010; Santamaria et al., 2011).

Based on the expansion and differentiation of the colonies it can be concluded that a population of epithelial and stromal progenitor cells residing in the basal layer give rise to large CFUs whereas transit amplifying cells with limited proliferative capacity residing in the functional layer give rise to small CFUs (Gargett and Masuda, 2010). The presence of eMSC in human endometrium accounts for



properties observed in stromal colonies and have similar multipotent properties with MSC found in bone marrow and adipose tissue (Gargett et al., 2009). Also these eMSC are responsible for stromal regeneration since they were absent from the myometrium and fallopian tubes (Gargett and Masuda, 2010).

#### *Side population (SP) cells*

Some adult stem cells can be identified by their ability to efflux the DNA-binding dye Hoechst 33342 via the ATP-binding cassette transporter G2 (ABCG2) which is considered to be a marker for this discrete population of stem cells (Gargett et al., 2012; Challen and Little, 2006). Somatic stem cells showing such activity have been isolated from several adult tissues (Preffer et al., 2002; Smalley and Clarke, 2005) using dual wavelength flow cytometry (Goodell et al., 1997) and they are characterised as SP cells since this ability is not conserved in all adult stem cells.

A very small percentage of freshly isolated human endometrial cells were identified as SP cells (Masuda et al., 2010) which when cultured displayed little clonogenic activity so it was suggested that they were in quiescence and thus inactive (Tsuji et al., 2008). It was also observed that they expressed intermediate levels of telomerase along with markers specific for the undifferentiated state (Cervello et al., 2010). SP cells were also identified in cultured human endometrial cells in similar percentages but in this case SP cells displayed an enhanced proliferative capacity to form colonies since they were primarily in G1, M or S phase of the cell cycle (Tsuji et al., 2008). Phenotypic analysis revealed typical characteristics of MSC in SP populations of endometrial epithelial and stromal cells (Cervello et al., 2010). Immunofluorescence staining illustrated the expression of ABCG2 in the basal and functional layers of human cyclic endometrium (Zhou et al., 2001); specifically the expression was detected at the vascularised areas since it was co-localised with CD13<sup>+</sup> cells (Masuda et al., 2010). It was also observed that SP cells occurred more frequently in proliferative (Masuda et al., 2010; Tsuji et al., 2008)

and menstrual (Kato et al., 2007) human endometrium despite that their numbers varied amongst individuals.

Identification of SP cells in human endometrium provides evidence for the existence of progenitor/stem cells which appear to be in close association with the vascular niche (Gargett, 2007) but their exact nature is yet to be defined (Gargett and Masuda, 2010).

### *Markers of stem/progenitor cells*

At the moment identification of specific markers for epithelial cells in human endometrium is yet to be achieved. On-going studies assess the functions of candidate adult stem cells, such as clonogenic activity, self-renewal, ability to differentiate and proliferative potential, for their verification (Gargett and Ye, 2012). However, there are known markers which are specific for endometrial MSC.

Endometrial stromal cells with clonogenic activity expressed CD146, a characteristic surface phenotype of MSC, along with the platelet derived growth factor receptor  $\beta$  (PDGF-R $\beta$ ) (Schwab and Gargett, 2007) which is important for regulating proliferation and differentiation. These cells showed their multipotent nature by their ability to differentiate into mesodermal lineages, such as adipocytes, smooth muscle cells, chondrocytes and osteocytes. It was also found that they occupied perivascular areas in the basal and functional layer (Schwab and Gargett, 2007). These observations were confirmed in another study by application of gene profiling on endometrial stromal cells whereby expression of pericyte markers as well as expression of genes associated with angiogenesis were identified (Spitzer et al., 2012). Genes responsible for maintaining self-renewal and multipotency in MSC were upregulated i.e. hypoxia response and signalling pathway genes. In the same study it was suggested that CD146<sup>+</sup>PDGF-R $\beta$ <sup>+</sup> cells are more likely to differentiate into stromal fibroblasts since they clustered with endometrial fibroblasts but could be distinguished from endothelial cells (Spitzer et al., 2012). Other MSC surface markers expressed by CD146<sup>+</sup>PDGFR $\beta$ <sup>+</sup> eMSC were CD29, CD44, CD73, CD90 and CD 105 (Dominici et al., 2006) but no expression of hematopoietic (CD34) and endothelial

(CD34 and CD31) stem cell markers was detected (Dimitrov et al., 2008). Endometrial and MSC cells can be identified by expressing MSCA-1, a surface marker for MSC in the bone marrow. In this case though, a second marker is needed to distinguish the two cell types (Sobiesiak et al., 2010). Because of their identified location in both layers of the endometrium, it can be said with certainty that mesenchyme cells in the stroma are shed in menstrual blood and it is very likely to be transported to neighbouring sites where they consequently induce growth of endometrial tissue which is what defines endometriosis (Gargett and Masuda, 2010).

Application of antibodies to examine the expression of perivascular markers in endometrial tissue and cell suspensions revealed W5C5 as a novel marker for MSC in human endometrium (Gargett et al., 2012). Endometrial cells expressing W5C5 were isolated and showed similar differentiation properties and behaviour to CD146<sup>+</sup>PDGF-R $\beta$ <sup>+</sup> cells during *in vitro* studies (Masuda et al., 2012). Transplantation of these cells in immunocompromised mice resulted in the generation of human endometrial tissue (Schuring et al., 2011). Mesenchymal stromal cells have been found in the functional layer of endometrial biopsies (Schuring et al., 2011) which could serve as a source of W5C5<sup>+</sup> cells with the potential to be used in cell therapies for tissue regeneration (Gargett et al., 2012).

Pluripotency in embryonic stem cells is maintained by expression of the transcription factor Oct-4 which was found to be expressed in adult stem cells too (Tai et al., 2005). Expression of this marker was observed in human endometrium without identifying though, neither the cell types nor their location (Matthai et al., 2006). In a cohort study hysterectomised endometrial tissue taken during the follicular or luteal phase of menstruating women expressed OCT-4 mRNA however no differences in expression patterns were detected during the cycle (Bentz et al., 2010). Identification of other adult stem cell markers in endometrial tissue includes bcl-2, c-kit and CD34 (Cho et al., 2000) but they were expressed by many endometrial cells other than clonogenic or side population cells so they could not be considered as definitive markers (Chan et al., 2004; Kato et al., 2007).

A known epithelial progenitor cell marker is Musashi-1, an RNA-binding protein which regulates self-renewal (Gotte et al., 2008). Musashi-1 positive cells were found in endometrial glands and stroma. In addition, during the proliferative phase of the menstrual cycle their numbers increased in the stroma and glands of the basalis in comparison to the functionalis. Secretory endometrium had less cells expressing Musashi-1 which is an indication that they probably function as stem/progenitor cells during the proliferative phase. Significant high levels of Musashi-1 were detected in cells from endometriotic and endometrial carcinoma tissue (Gotte et al., 2008).

#### *Mesenchymal cells in menstrual blood*

The presence of endometrial MSC in menstrual blood was examined in culture studies whereby menstrual blood cells were grown in similar way as MSC from bone marrow (Hida et al., 2008). Cultured cells looked like fibroblasts and had a high proliferative potential when *in vitro* (Patel et al., 2008). They displayed stable karyotype and telomerase activity as well as expression of characteristic eMSC surface phenotype (Meng et al., 2009). They were able to differentiate into various types of cells of mesodermal (Cui et al., 2007) and ectodermal (Patel et al., 2008) origin and also cardiac and skeletal muscle cells (Cui et al., 2007). Expression of immune-modulatory genes (Murphy et al., 2008) was found in menstrual blood cells and expressed only MHC-class I genes suggesting that they have low immunogenicity (Caplan, 2009). These immune properties have also been observed in MSC of bone marrow and adipose tissue. The fact that eMSC are found in menstrual blood is also an indication of their possible role in endometrial growth at ectopic sites (Sasson and Taylor, 2008).

#### *Epithelial stem/progenitor cells in the basal layer*

Application of the HUMARA assay for the X-linked androgen receptor gene (Tanaka et al., 2003) demonstrated the monoclonality of individual glands. The results from this assay showed the shared clonality between adjacent glands and concluded that they were derived from the same epithelial

progenitor cells whose location though could not be precisely indicated. Monoclonal endometrial glands have been identified in normal human endometrium by using PTEN immunostaining (Mutter et al., 2000). *PTEN gene*, a tumour suppressor gene that responds to oestrogen (Mutter et al., 2001), was not expressed in some glands either because of mutation or deletion or both (Mutter et al., 2000) and clones of these glands were found in the basal layer and proliferation of their cells during successive menstrual cycles resulted in the formation of PTEN-null glands in the functional layer. In endometrial hyperplasia the number of PTEN-null glands increases since unopposed oestrogen stimulation results in continuous proliferation of cells and no expression of the gene (Mutter et al., 2001).

Methylation patterns of genes have been used as epigenetic markers to indicate the presence of epithelial stem/progenitor cells in individual glands of human endometrium (Kim et al., 2005). Mitotic activity of adult stem cells resulted in the inheritance of these markers by their daughter cells whereas markers in TA or more mature cells, residing in the functional layer, were lost as the cells were shed in menstrual blood (Figueira et al., 2011). Studies on methylation patterns in endometrial glands revealed that the number of epigenetic markers reflected the rounds of stem/progenitor cell divisions (Ro and Rannala, 2001). Kim et al. observed that the degree of gene methylation increased with age but remained constant during menopause inferring the presence and activity of endometrial epithelial progenitor cells in the basal layer in menstruating and inactive endometrium.

Also the fact that epithelial cells were not detected in cultures of menstrual blood is an indication that their progenitors are located in the basal layer which remains intact during menstruation (Musina et al., 2008).

### 1.2.3 Role of endometrial stem cells in endometrial disorders

Adult stem cells in all regenerative tissues are responsible for their homeostasis i.e. they ensure that cell loss is balanced with cell turnover and thus maintain normal function of the tissue. Endometriosis, adenomyosis, endometrial hyperplasia and endometrial cancer are gynaecological diseases characterised by abnormal proliferation of endometrial cells. It is believed that initiation and progression of these diseases is due to abnormal functioning of either the endometrial stem/progenitor cells or their niche or both (Gargett, 2007).

#### *Endometriosis*

Ectopic growth of endometrial glands and stroma is the typical characteristic of endometriosis which can be described as a chronic benign disease. The growth usually occurs in pelvic organs, the peritoneum and the rectovaginal pouch resulting in symptoms such as pain, infertility and inflammation (Giudice and Kao, 2004). The mechanism underlying the formation of endometrial implants involves the attachment of inappropriately shed endometrial cells on the peritoneal cavity to achieve invasion into the mesothelium and the cells then survive and proliferate (Jensen and Coddington, 2010). The pathogenesis of endometriosis is not clearly understood; a few hypotheses have been developed to explain the origin of endometriotic implants. The most accepted theory states that retrograde menstruation results in deposition of menstrual debris in the peritoneal cavity establishing endometrial implants (Sampson, 1927b). The coelomic metaplasia theory suggests that endometriosis results from induced metaplastic changes in the cells lining the visceral and abdominal peritoneum (Gruenwald, 1942). It is also believed that remnant fetal mullerian cells can establish ectopic growth of endometrial tissue (von Recklinghausen, 1986). The last theory sets the lymphatic system as a mean for the spread of endometrial cells to ectopic sites (Sampson, 1927a).

Studies have provided evidence for the presence of endometrial stem/progenitor cells in endometriotic lesions and thus suggesting their possible role in the pathogenesis of the disease. Cultures of stromal cells from endometriotic tissue were found to have MSC-like cells which expressed a typical MSC surface phenotype and displayed multipotent differentiation potential (Kao et al., 2011). Ovarian endometrioma was found to have epithelial and stromal cells with clonogenic activity; CFU of both types of cells were able to self-renew whilst stromal CFU showed multipotency (Chan et al., 2011). Some endometriotic lesions were found to contain monoclonal epithelial cells whilst other lesions had polyclonal stromal cells leading to the conclusion that a single cell can initiate the formation of the lesion which can get contaminated by cells from other sources, such as the bone marrow, or stem cells from shed endometrial fragments may be established within the lesion (Gargett and Masuda, 2010).

Further examinations are needed to define the direct role of endometrial stem/progenitor cells or bone marrow stem cells in endometriosis. It was suggested though that stem cells in the basalis are taken up by the blood flow during menstruation and reach the peritoneal cavity via the fallopian tube where they eventually develop endometriotic implants (Sasson and Taylor, 2008). Abnormal functioning of endometrial stem cells is a possible cause for ectopic growth but also an abnormal peritoneum may provide favourable conditions for normal stem cells to implant and establish endometrial growth (Figueira et al., 2011).

### *Endometrial cancer*

Cancer cells vary in their ability to proliferate, differentiate, initiate tumors, express different markers and genes, and vary in their lifespan too (Gargett et al., 2012). Genetic or epigenetic changes constitute the transformation of normal adult stem cells, progenitor or mature cells into cancer stem cells (CSC) which are able to self-renew (Jordan, 2006) as well as produce more differentiated daughter tumour cells (Visvader and Lindeman, 2008). As tumour development progresses though, CSC may

acquire more mutations which will eventually result in changes in CSC population and its phenotype (Nguyen et al., 2012).

The underlying cause of endometrial adenocarcinoma is the loss of control over proliferation and regeneration of endometrial epithelial cells. Type I adenocarcinoma develops in pre- and perimenopausal women and is oestrogen dependent whereas type II affects post-menopausal women and is oestrogen independent (Di Cristofano and Ellenson, 2007). Also these two types of endometrial adenocarcinoma result from mutations in different genes; mutations in Pten, K-ras and  $\beta$ -catenin genes are associated with type I whilst mutations in p53 and HER-2/neu account for type II (Di Cristofano and Ellenson, 2007).

A subpopulation of CSC was found in human endometrial cancer cell lines (Friel et al., 2008). Some freshly isolated endometrial cancer cells were able to self-renew *in vitro* and *in vivo* and had clonogenic activity (Hubbard et al., 2009). Expression of genes responsible for self-renewal in stem cells, such as *BMI-1*, *NANOG* and *SOX2*, was detected in clonogenic endometrial cancer cells (Hubbard et al., 2009).

It is believed that SP cells also have a role in tumour development as they have been identified in endometrial cancer and endometrial cancer cell lines (Kato et al., 2010). In one type of cell line, SP cells were in quiescence but could self-renew and showed chemoresistance (Friel et al., 2008) whilst in other cell lines, used as a model of type II adenocarcinoma, SP cells had properties of self-renewal and were also able to produce clones and large tumours *in vivo* (Kato et al., 2010).

Two studies used CD133 as a marker of CSC in human endometrial cancer tissue. Cells expressing CD133 formed clones more efficiently, displayed higher proliferative potential and chemoresistance. Either expressing CD133 or not, cells failed to initiate tumours *in vivo* in contrast to CD133<sup>+</sup> and CD133<sup>-</sup> cells from endometrial cancer cell xenografts which were equally able to form tumours (Rutella et al., 2009). In the other study though, it was observed that tumours were initiated more readily by CD133<sup>+</sup> cells (Friel et al., 2010). The use of CD133 as a marker for the presence of CSC in endometrial tissues



is not agreed (Mather, 2012) so it requires further research but their role in tumor initiation and progression cannot be doubted.

### *Adenomyosis*

It is very possible that invasion of the myometrium in adenomyosis is caused by the abnormal behaviour of endometrial stem cells or even the cells constituting their niche microenvironment (Gargett, 2007). An abnormal orientation of the niche may result in differentiated cells establishing layers within the myometrium instead of moving towards the functional layer. Niche cells may lose regulation over differentiation potential of endometrial stem cells which may cause their excessive differentiation into smooth muscle cells and this can account for myometrial hyperplasia which is a feature of adenomyosis (Gargett, 2007). In terms of pathophysiology, it was suggested that injury to the endometrial myometrial junction is caused by the contractions of myometrium for sperm transport (Leyendecker et al., 2009). This results in production of oestrogen at the damaged area which in turn activates endometrial stem cells in the basal layer as a response to tissue injury. Stem cells proliferate and invade the myometrium to repair the damaged tissue (Gargett and Masuda, 2010). The presence of stromal cells in adenomyotic tissue was proven by their ability to differentiate into lineages of mesodermal origin and express markers for MSC surface phenotype when cultured (Chen et al., 2010). But no other conclusions were drawn concerning their clonogenicity and ability to self-renew. Application of gene profiling on these cells though showed their differences with stromal cells from normal endometrium and that they over expressed COX-2 which has a key role in repairing the injured myometrial tissue and local release of oestrogen (Leyendecker et al., 2009).

### *Inadequate endometrial regeneration*

In conditions such as Asherman's syndrome or in endometrial areas covered with scar tissue the endometrium is thin and dysfunctional possibly because stem/progenitor cells are either reduced in

number or have reduced function or both (Gargett and Ye, 2012). The endometrium cannot support implantation (Gargett and Healy, 2011) so some of the symptoms associated with these conditions are infertility and unsuccessful pregnancy (Panayiotides et al., 2009). Intrauterine adhesions which establish Asherman's syndrome are a consequence of trauma to the basal and the myometrium caused by miscarriage, elective abortion, curettage or pelvic surgeries (Yu et al., 2008). Release of inflammatory mediators as a response to infection and inflammation inhibits the regeneration of the traumatised endometrial tissue by damaging stem cells or their niche and instead induces the formation of fibrotic tissue (Gargett et al., 2012). The epithelial surface of the endometrium is re-established without any scar tissue during the menstrual cycle and after parturition (Salamonsen, 2003). A major trauma at the basal layer and the underlying myometrium can lead to the damage or loss of resident stem cells so the endometrium is thin, atrophic and inactive (Yu et al., 2008). The MSC in the functional layer cannot compensate for this loss so endometrial regeneration fails and the intrauterine adhesions get covered with fibrotic tissue and eventually a scar is formed (Schenker and Margalioth, 1982).

#### 1.2.4 Regenerative medicine

##### *Adult stem cell therapies*

The ability of stem cells to self-renew and differentiate into various types of cells has been of great interest in regenerative medicine and tissue engineering over the years for the development of adult stem cell therapies. Even though embryonic stem cells have a wide range of differentiation potential, their use in regenerative medicine is limited due to several ethical issues so multipotent mesenchymal stem cells derived from adult tissues have been used as an alternative. Several clinical cases have reported the use of bone-marrow derived mesenchymal cells to treat bone loss (Vacanti et al., 2001), fracture non-union (Bajada et al., 2007), osteogenesis imperfecta (Horwitz et al., 1999; 2001), knee

osteoarthritis by generation of cartilage (Wakitani et al., 2002), acute spinal cord injury whereby improvement of sensory and motor function was observed (Park et al., 2005), myocardial infarction whereby implantation of stem cells resulted in significant reduction of myocardial infarct size and better systolic function (Janssens et al., 2006). The stem cells in the corneal limbus have been well characterised and their application in stem cell therapies was identified quite early. Pallegriani et al., (1997) cultivated corneal progenitor cells taken from the patients to produce autologous corneal epithelial sheets which were then grafted on the damaged eye and resulted in restoration the corneal-limbal surface.

#### *Application of endometrial stem cells*

Endometrial MSC have shown similar properties to MSC derived from bone marrow and adipose tissue such as differentiation into adipocytes, osteoblasts and chondrocytes (Gargett et al., 2007; Kao et al., 2011), self-renewal and colony-forming ability. MSC from bone marrow and adipose tissue have been used in regenerative medicine for tissue repair (Caplan, 2009) since they promote angiogenesis and proliferation of stem cells, prevent fibrosis and inflammatory responses and have low immunogenicity (Caplan, 2007). Hysterectomy endometrial tissue (Massassa and Taylor, 2012), endometrial biopsies and menstrual blood (Patel et al., 2008) are an available source for eMSC whose application in tissue regeneration has been of great interest recently and their potential has been examined in animal and human studies.

The regenerative potential of eMSC from menstrual blood was first seen in a mouse model of Duchenne muscular dystrophy whereby transplantation of the cells into atrophied skeletal muscle fibres contributed to muscle repair (Cui et al., 2007). Cell fusion and *in situ* differentiation were postulated as mechanisms for repair but it also appeared that angiogenesis had a role in the reparative effect since the transplanted eMSC homed to peri-muscle fibre regions which are vascularised (Cui et al., 2007). The role of angiogenesis in tissue repair was reported in a mouse study where menstrual

blood MSCs improves critical limb ischaemia induced by femoral artery ligation (Murphy et al., 2008). In a tissue engineering approach, EGFP-labelled menstrual blood MSCs were grafted in a murine model of myocardial infarction whereby grafted cells were able to differentiate into striated muscle cells expressing troponin and actinin and there was a significant reduction in infarct area when compared to control animals treated MSC from bone marrow (Hida et al., 2008). In a murine model of encephalomyelitis, endometrial stem cells were intraperitoneally delivered and whereby their anti-inflammatory effect was observed (Peron et al., 2012). In another murine model of intracranial glioma, a decrease in tumor size was observed after the intravenous administration of endometrial stem cells probably because of their anti-angiogenic effects (Han et al., 2009). It was reported that eMSC can promote survival and regeneration of neural cells by exerting a trophic effect as it was observed in oxygen-deprived primary neuronal cell cultures and in a murine model of ischaemic stroke whereby histological and behavioural improvement was observed after injection of eMSC (Borlongan et al., 2010).

Application of tissue engineering to treat pelvic organ prolapse, a disorder resulting from vaginal birth injury, by incorporating eMSC into scaffolds has been tested *in vivo* on an animal skin wound repair model and the observations were enhanced neovascularization, tissue integration, reduced inflammation and deposition of collagen fibres (Ulrich et al., 2012; Edwards et al., 2013). The scaffolds were constructed using artificial meshes which displayed distensibility after they were transplanted and enhanced integration within the tissue due to eMSC preventing tissue responses to foreign bodies. In another approach to engineer the urinary bladder wall, endometrial stem cells were incorporated into fabricated nanofibrous silk-collagen fibres and subsequently were able to differentiate into smooth muscle cells (Shoae-Hassani et al., 2013).

In the case of human studies, clinical-grade MSC from menstrual blood were administered via intravenous and intrathecal routes in multiple sclerosis patients during a phase I clinical trial and neither immunological reactions neither adverse side effects were reported after a year (Zhong et al.,

2009). Similarly no side effects were reported during a period of 3 years after the intramuscular injections of endometrial stem cells in a patient with Duchenne muscular dystrophy (Ichim et al., 2010). On the contrary the muscles got stronger and the patient had less respiratory infections. Endometrial stem cells were administered via intravenous routes in a patient suffering from congestive heart failure and after a year of follow up the ventricular pumping was improved (Ichim et al., 2010).

There are always limitations though about using endometrial stem cells for regenerative therapeutics. Endometrial stem cells can be isolated using magnetic beads based on the expression of W5C5 marker however W5C5 is not considered as a definitive marker specific for MSC. Also there is a risk of vaginal infection by the use of the menstrual cup for collection of menstrual blood to derive MSC. More importantly larger animal studies need to be introduced in order to assess and get a better understanding of the mechanisms underlying the regenerative potential of endometrial stem cells before being further applied in human studies (Verdi et al., 2014).

#### *Endometrial tissue generation*

Human embryonic stem cells (hESC) are pluripotent and their differentiation can be manipulated in a clinical setting for tissue regeneration (Murry and Keller, 2008). The use though of human embryos to derive pluripotent stem cells raises some ethical issues so research has been focused on using induced pluripotent stem (iPS) cells instead. In addition, the use of iPS cells also eliminates the immunological issues posed by hESC (Gargett et al., 2012).

A model of endometrial epithelial cells of human endometrium was generated from hESC (Ye et al., 2011). At first the aim was to mimic the development of the Mullerian duct and this was achieved by inducing mesodermal differentiation in embryoid bodies which were then combined with mesenchyme cells from the uterus of a neonatal mouse. The recombinant tissue produced was then transplanted *in vivo* whereby further differentiation resulted in expression of characteristic Mullerian

duct markers and formation of ciliated cells was also observed (Ye et al., 2012). The function of the hESC-derived epithelium was assessed by exposure to exogenous oestrogen which induced the proliferation of cells and production of glycodelin A. generation of a functional hormonally responsive epithelium tissue resembling adult human endometrial epithelium was achieved emphasising the role of endometrial MSC in regulating differentiation of hESC. This strategy can be applied *in vitro* for the generation of an adult human endometrial epithelium using the patient's own iPS cells and endometrial stromal cells to produce Mullerian derivatives (Gargett and Ye, 2012) which could potentially be transplanted into patients for the *in vivo* regeneration of endometrial tissue.

## 1.3 Background on applied methods

### 1.3.1 Scanning electron microscopy

#### *Endometrial architecture during the menstrual cycle*

Ludwig and Spornitz (1991) were amongst the firsts to visualise the changes in the microarchitecture of human endometrium under the Scanning electron microscope (SEM). They examined normal human endometrium hysterectomised from women with mostly cervical pathologies and they ordered the events of changes on the endometrial surface according to the day of the menstrual cycle of the patients.

They observed that on the first day of the cycle, during which there is on-going bleeding, epithelial cells are already lost from some areas of the endometrial surface and open capillary vessels run in parallel to the wound. Desquamation of the functionalis results in the appearance of stumps, from the remaining glands in the basalis, on the stroma on the second and third day of the menstrual cycle. Epithelial cells at the margins of the stumps were seen to grow horizontally towards the surface but formation of the new glands of the functionalis requires grow in a vertical direction so that the glands have an appropriate length. Soon the glandular stumps adopt a cone shape. Newly synthesised epithelial cells come out from the stumps at the wound areas whilst the surrounding tissue is covered by a fibrin network and trapped red blood cells.

Rapid proliferation of epithelial cells continuous during the fourth day of the cycle; they grow in a spiral pattern and fuse to produce the new epithelial lining. They manage to cover most of the denuded surface as local fibrinolytic activity breaks down the fibrin mesh which is not seen after day four. It is quite common that the rapid epithelial proliferation from the remaining glands results in the

formation of micropolyps next to the glandular openings. The few though that are able to survive grow larger and become vascularised.

By days five and six the endometrial wound is covered by the new epithelial lining, the stromal tissue starts to grow and the endometrium gets thicker. Stromal cells grow at a higher rate than glandular cells, as a result the endometrial surface appears to be uneven and glandular openings are depressed. By the middle of the cycle the surface becomes smooth and this appearance persists for a while. Probably the endometrium now provides favourable conditions for implantation to take place.

Ciliated cells seen during the first days of the menstrual cycle are those which already existed within the surviving glandular tissue. By the beginning of the second week new ciliated cells appear between epithelial cells. Depending on the location and maturation stage within the endometrium, ciliogenesis occurs at different rates and the ciliated cells vary in type i.e. ciliated cells at the endometrial surface are of different type than the ciliated cells found at the glandular epithelium of the basalis. By midcycle though, ciliated cells reach full development.

White blood cells and macrophages migrate to the stroma when menstrual breakdown is about to begin, that is around day 22. They serve to clear away secretions and cellular debris filling the glandular opening. If implantation does not take place, epithelial cells at the surface lining lose their integrity and clefts appear on the surface. Just before menstruation on day 28, apical defects are noticed on the membrane of some cells. On the other hand, ciliated cells remain unchanged; some though may have flat microvilli or even lack microvilli.

The vasculature breaks down and bleeding commences, as a result white and red blood cells accumulate on the endometrial surface. Local fibrinolytic, lysosomal and macrophage activity increases to clear away the fibrin mesh and cellular debris as well as to create a non-coagulable menstrual flow. The triggering factors at this point are considered to be paracrine and endocrinological. By definition, the onset of menstrual bleeding is taken as the first day of the menstrual cycle.



### 1.3.2 Biospectroscopy

In the recent years, biomedical studies turned their focus on the application of vibrational spectroscopic techniques which have shown their potential as very important diagnostic tools for the detection of malignancy and cancer (Dukor et al., 2002; Choo-Smith et al., 2002; Parker, 2005). These techniques are very simple in the context of sample preparation, are easily repeated and most importantly they are non-destructive to the tissues (Zanyar et al., 2007; 2008). Spectroscopy results in a molecular fingerprint for the biochemical composition of a tissue. Vibrational spectral bands provide direct information about the types of bonds and functional groups present within a molecule. In diseased tissues, the molecules experience structural and conformational changes which consequently result in changes on the spectral bands and thus enabling the optical detection of these molecular changes (Short et al., 2006; Sebag et al., 1993). Application of biospectroscopy imaging techniques has also allowed for cancer detection (Zeng et al., 2004; Dekker and Fockens, 2005; Demos et al., 2006). Of great interest is the use of spectroscopy for the characterisation (Adam et al., 2007; and isolation (Patel et al., 2012) of stem cells in human tissues. In this project Fourier Transform Infrared (FTIR) spectroscopy and Raman spectroscopy have been employed in an attempt to locate the stem cell population in endometrial tissue.

#### *FTIR Spectroscopy*

For a molecule to be IR active its chemical bonds must experience a change in electric dipole moment (Stuart, 2004). These vibrational changes allow biomolecules in cells to absorb in the mid IR-region ( $\lambda=2.5\text{-}25\mu\text{m}$ ) (Kelly et al., 2009; Walsh et al., 2007) and thus they are quantitatively measurable (Griffiths and Haseth, 2007) and can be used for comparisons between samples. A spectrum of wavenumbers is generated with absorbance intensities according to the biochemical composition of the cell (Llabjani et al., 2011). In FTIR spectroscopy this information is found between the spectral

region 1800-900  $\text{cm}^{-1}$ ; the so called; 'biochemical cell fingerprint' (Martin et al., 2010). Certain regions of wavenumbers and their assignments which are of particular importance during interrogation of a biological sample with IR spectroscopy include 1650  $\text{cm}^{-1}$  for Amide I, 1550  $\text{cm}^{-1}$  for Amide II, 1425  $\text{cm}^{-1}$  for protein, 1260  $\text{cm}^{-1}$  for Amide III, 1225  $\text{cm}^{-1}$  for asymmetric phosphate stretching vibrations ( $V_{as} \text{PO}_2^-$ ), 1155  $\text{cm}^{-1}$  for carbohydrate, 1080  $\text{cm}^{-1}$  for symmetric phosphate stretching vibrations ( $V_s \text{PO}_2^-$ ) and 970  $\text{cm}^{-1}$  for protein phosphorylation (Patel et al., 2011). Despite FTIR spectroscopy, other IR techniques of equivalent importance which are employed in research are attenuated total reflection fourier transform IR (ATR-FTIR), photothermal microscopy, synchrotron FTIR and pressure-tuning FTIR.

Biomedical studies have different aims and objectives depending on what is being investigated, however it has always been the case that application of FTIR analysis gave significant results. FTIR spectroscopy can be used as a diagnostic tool for cervical cancer screening (Wood et al., 1996), to observe maturation and differentiation of cells allowing detection of abnormalities (Chiriboga et al., 1998; 1998), to detect biomarkers that are either common between normal and diseased tissues (i.e. cancer, malignant or benign tissues) or enable the types of tissues to be distinguished (Mordechai et al., 2004). Also, the application of more than one IR technique in a study can indicate which of those techniques is more suitable to obtain meaningful results of good quality (Wong et al., 1995). Parameters within the system can be set according to the experimental design and the type of tissue or cells being investigated.

### *Raman spectroscopy*

For a molecule to be considered as 'Raman-active' there should be a change in the molecular polarisation (Patel et al., 2011). Collision of an incident photon, from a laser beam, with a molecule results in a transfer of energy between the two. Chemical bonds within the molecule get excited but then they relax into a different vibrational state causing the scattered photon be at a lower or higher

energy level than the incident photon. These changes in energy levels results in Raman scattering and frequency shift (Singh and Reiser, 1993; Lin-Vien et al., 1991). The IR spectra is generated based on the difference between the incident and vibrational frequencies of the molecular bonds in the sample (Martin et al., 2010). The 'biochemical cell fingerprint' in Raman spectroscopy, including frequency shifts assigned for important cellular biomolecules, is between the region 1800-900  $\text{cm}^{-1}$ . Important regions of frequency shifts and their assignments include 1660  $\text{cm}^{-1}$  for Amide I, 1575  $\text{cm}^{-1}$  for nucleic acids (DNA/RNA), 1460  $\text{cm}^{-1}$  for lipid and protein, 1265  $\text{cm}^{-1}$  for Amide III, 1001  $\text{cm}^{-1}$  for phenylalanine, 787  $\text{cm}^{-1}$  for DNA and 530  $\text{cm}^{-1}$  for protein (Patel et al., 2011). Raman spectroscopy is a complementary technique to IR spectroscopy because it allows spectral acquisition from aqueous samples (Kelly et al., 2011). Other Raman spectroscopic techniques are near-infrared (NIR) Raman spectroscopy, surface-enhanced Raman spectroscopy (SERS) and resonance Raman spectroscopy. The latter two allow for rapid and greater signal-to-noise ratio spectral acquisitions by enhancing the inelastic scattering of photons which can be relatively weak (Wood et al., 2009; Kneipp et al., 2002).

Similarly to FTIR spectroscopy, Raman spectroscopy has been employed in several biomedical studies to characterise tissues producing significant outcomes and conclusions as well. This technique can be used as a method for the *in vivo* detection of cervical precancer based on squamous dysplasia (Uttinger et al., 2001), classification of cancer in epithelial tissues (Stone et al., 2004; 2002), diagnosis of lung cancer (Huang et al., 2003), discriminating skin lesions and thus allow for detection and classification of skin cancer (Sigurdsson et al., 2004), diagnosis of basal cell carcinoma (Gniadecka et al., 1997). Raman spectroscopy can also be used to derive images from biological tissues which provide significant information as well for tissue analysis (Patel et al., 2011).

### 1.3.3 Spectral pre-processing

There are several pre-processing techniques to be applied and the choice of method depends on the instrument being used. It is important to mention that the choice in turn affects the results of

Multivariate analysis. Many issues may arise during spectral acquisition or sample preparation which are corrected by pre-processing of raw spectra so as to make sure that the resulted spectra is a proper representation of the biochemical composition of the sample. Common issues that may appear on the acquired raw spectra have to do with sloped or oscillatory baseline effects due to the dispersion of IR light (i.e. Mie scattering) (Basson et al., 2009;2010), correction of noise, and variations in absorbance intensity due to variations in sample thickness, concentration or contact between the sample being examined and the instrument. The latter is a main concern when using ATR spectroscopy whereby the sample comes in contact with the ATR crystal.

To approach these issues, usually the first step is to cut the spectra to the region of interest, i.e. for FTIR spectroscopy that is the region between 1800-900  $\text{cm}^{-1}$  and for Raman spectroscopy is between 1750-500  $\text{cm}^{-1}$ . The most common technique used to remove sloped or oscillatory baseline is the rubber-band baseline correction. Removal of discrepancies in absorbance intensities is achieved by applying normalisation to a consistent peak; for example in FTIR spectroscopy normalisation to the Amide I peak is often applied. Whereas in Raman spectroscopy, or in cases when the spectra does not contain a consistent peak, vector normalisation is applied instead (Kelly et al., 2009). These normalisation techniques account for noise-reduction as well (Kelly et al., 2009).

#### 1.3.4 Computational analysis

Multivariate analysis is employed after pre-processing of spectra datasets in order to allow observation of similarities and differences between the classes within the dataset and to identify the peaks responsible for these observations (Martin et al., 2010). There are several techniques of data analysis that can be used. The multivariate techniques applied in this project were Principal Component Analysis (PCA-LDA) and Linear Discriminant Analysis (LDA).

Both techniques transform the spectra dataset into linear combinations by forming new variables or factors based on the original variables, which in this case are the absorbance intensities (Martin et al., 2010). The linear combinations are represented by a loading vector but these two techniques use different approaches to generate their respective loading vectors (Martin et al., 2010).

### *PCA*

PCA is an unsupervised method that does not take into account the assigned classes. It reduces the dataset into a few factors (10-20), called principal components (PCs), which are uncorrelated with each other (Duda et al., 2001) and capture as much as 99% of the variance within the dataset (Fearn, 2002).

Application of PCA allows for data to be visualised in a reduced dimensional space of one, two or three dimensions whilst capturing most of the information from the dataset (Martin et al., 2010). However, in this concept, heterogeneity in the dataset and thus the observed results account for variations within the classes rather than variations between the classes which cannot be distinguished by PCA (Martin et al., 2010). So further analysis is required for a more valid interpretation of the results.

### *LDA*

On the other hand, LDA is a supervised technique that maximises variation between classes in relation to variation within classes (Duda et al., 2001). The issue in this approach is the over fitting especially when the data sets are small, but this can be avoided by applying LDA after PCA. PCA reduces the number of variables into a few factors which are used as an input for the application of LDA (Martin et al., 2010). So in this concept, the analysis method is termed as PCA-LDA and allows for segregation of classes to be visualised in a dimensional space of one, two or three dimensions.

### *Visualisation*

Scores plots are scatter plots whereby the first two or three factors allow for similarities and differences between classes to be visualised at the axes (Martin et al., 2010). Whereas, loadings or cluster vectors allow for peaks responsible for segregation or similarities to be identified by visualising loading vectors as a function of wavenumbers (Martin et al., 2010).

#### **1.4 Aims and objectives**

Post-menopausal endometrium has not been investigated a lot. In this project, Scanning electron microscopy was used to investigate the architecture of post-menopausal endometrium in different endometrial biopsies with attention to the cellular composition at areas where crypts of endometrial glands were located. Biospectroscopy techniques were employed to investigate the biochemical composition of epithelial cells of the glandular elements as well as of stromal cells in the surrounding connective tissue in an attempt to identify and characterise the location of endometrial stem/progenitor cells and thus get a more clear understanding about their role in the structure and function of post-menopausal endometrium.

# CHAPTER 2

## Materials and Methods

## 2.1 Samples

For all the experiments, endometrial samples were provided by the Royal Preston Hospital. An informed consent was obtained from the patients, who had undergone hysterectomy, for the experimental use of their endometrial biopsies. All experimental procedures were ethically approved by the Research Ethics Committee (REC).

Endometrial specimens were provided from female patients who experienced postmenopausal bleeding because of either endometrial tumours or endometriosis. Hysterectomised endometrial samples were used for Scanning Electron Microscopy (SEM) investigation whilst for biospectroscopy examination, using Fourier Transform Infrared (FTIR) spectrometer and Raman spectrometer, paraffin embedded blocks of the samples were provided. See table 2.1 for detailed information on patients and the samples.

Sample	Diagnosis	Examination	Blocks supplied
ECNT 68	Endometrial cancer	SEM, Biospectroscopy, H&E	H09-12890-9 (tumour) H09-12890-11 (normal endometrium)
ECNTN 50	Endometriosis	SEM, Biospectroscopy, H&E	H09-12292-1 (normal endometrium)
ECNTN 49	Normal	SEM, Biospectroscopy, H&E	H09-11708-4 (normal endometrium)
ECNT 65	G1 endometrial cancer	SEM	
ECNTN 51	Endometriosis	SEM	
ECNT 64	G1 endometrial cancer	SEM, Biospectroscopy, H&E	H09-9102-1 (tumour) H09-9102-6 (normal endometrium)
ECNT 70	Clear cell endometrial carcinoma	SEM, Biospectroscopy, H&E	H09-13558-A5 (tumour) H09-13558-A8 (normal endometrium)
ECNT 69	G3 endometrial cancer	SEM	

**Table 2.1:** Information provided by the Preston Hospital about the endometrial biopsies examined by SEM and the paraffin embedded blocks used in biospectroscopy and H&E staining.



## **2.2 Scanning Electron Microscopy**

### 2.2.1 Sample preparation

All the steps were carried out in the fume hood.

Prior to processing, samples were kept in 10% formalin for fixation and stored at 4°C for preservation. Before SEM examination, samples were cut into smaller pieces using a blade (Agar Scientific Ltd/ Stansted, Essex, UK) and were then fixed in 2.5% glutaraldehyde in Phosphate Buffered Saline (PBS) (Sigma® Life Science, sigma-aldrich.com/ St. Louis, MO, USA) for an overnight wash at 4°C. This solution was prepared by adding 10ml of 25% glutaraldehyde EM (TAAB Laboratories Equipment LTD/ Aldermaston, Berks, England) in 90ml of PBS.

The overnight wash was followed by three washes in PBS, each lasting five minutes. Samples were then post-fixed in Osmium tetroxide (OsO<sub>4</sub>) 2% solution (Agar Scientific Ltd/ Stansted, Essex, UK) for 90 minutes. The lids of the vials were discarded in a plastic bag and replaced by new ones. Three more washes of 10 minutes in PBS were then carried out. A series of ethanol washes was performed to dehydrate the samples; single washes of 30 minutes in 50%, 70%, 80% and 90% ethanol followed by two washes of 40 minutes in 100% ethanol. Hexamethyldisilazane (HMDS) (Sigma® Life Science, sigma-aldrich.com/ St. Louis, MO, USA) was used for sample drying during two washes of 30 minutes each. A new volume of HMDS was then added for an overnight wash of the samples in glass vials whose tops were removed.

### 2.2.2 Sputter coating

The pieces of the samples were attached on the adhesive disc on an aluminium SEM stub with conductive carbon cement. The SEM stubs were then placed in Edwards S150A sputter coater to be coated with gold.

The vacuum chamber of the coater was evacuated by opening the air admit valve. Once an ideal gas pressure of  $10^{-1}$  mbar (1mbar=100Pa) was reached, indicated by the pirani gauge, Argon gas was introduced in the chamber at a pressure of 4psi (pound per square inch). The pressure was then increased by turning the gas admit valve anticlockwise and a low current of 40mA but with high voltage was then applied in the chamber for four minutes. This ionised Argon gas and turned it into plasma which could be seen as a purple glow. The gas admit valve and the argon cylinder were closed and the chamber was again evacuated.

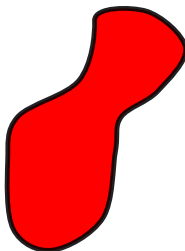
### 2.2.3 Examination

A digital Scanning Electron Microscope (JEOL JSM 5600, Herts, UK) was used for examination of the endometrial samples. Digital images were taken from different areas and at different magnifications allowing for cells to be distinguished.

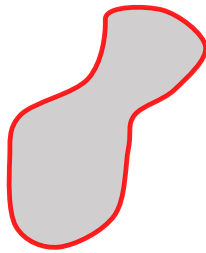
### 2.2.4 Image Analysis

The computer software ITEM (Universal ITEM Imaging Platform, Soft Imaging System) was used for image analysis. Images were manually calibrated and the parameters of the cells of interest were manually defined as well (Fig. 2.1).

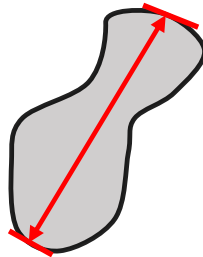
- Area: the area of the particle is (number of pixels of the particle) times (calibration factors in X and Y direction).



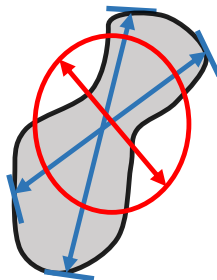
- Perimeter: the sum of the pixel distances along the closed boundary.



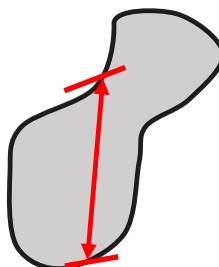
- Diameter max: to evaluate the maximum diameter of a particle, the diameter for different evaluation axes will be determined. The evaluation axis is varied in  $1^\circ$  steps and the maximum diameter at each angle is determined.



- Diameter mean: the arithmetic mean of all diameters of a particle (for angles in the range  $0^\circ$  through  $179^\circ$  with step width  $1^\circ$ ).

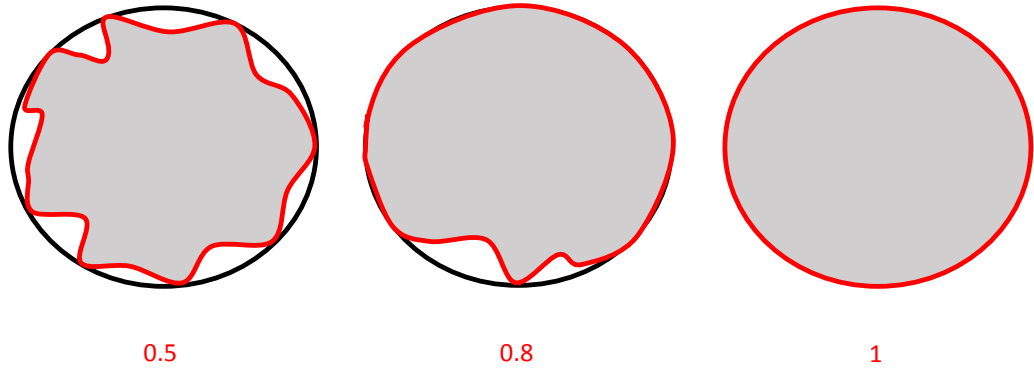


- Diameter min: the minimum diameter of a particle (for angles in the range  $0^\circ$  through  $179^\circ$  with step width  $1^\circ$ ).

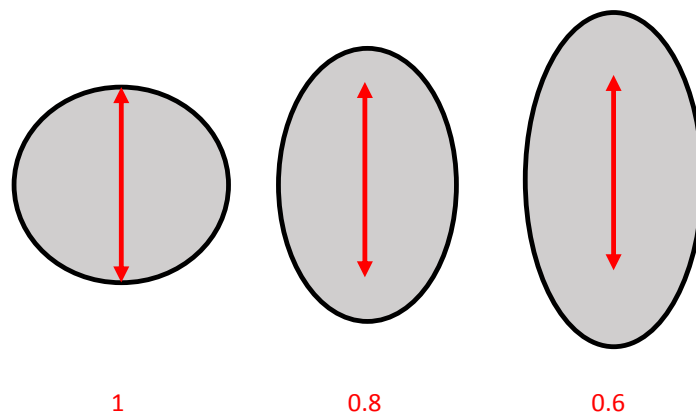


- Shape Factor: the shape factor provides information about the “roundness” of the particle.

For a spherical particle the shape factor is 1, for all other particles it is smaller than 1.



- Sphericity: describes the sphericity or “roundness” of the particle by using central moments.



- Aspect ratio: the maximum ratio of width and height if a bounding rectangle for the particle.

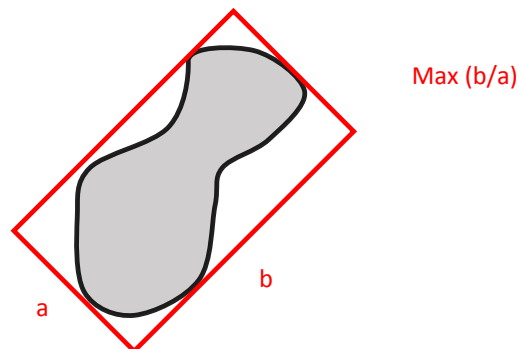


Figure 2.1: The analysed parameters and their defined measurements.

Microsoft Office Excel 2013 was used to represent the parameters of the counted cells on bar charts and allow for comparisons and further data analysis. GraphPad Prism 4.0 Software was used to perform the statistical analysis (unpaired t-test) which allowed the testing of the significant difference between the counted cells in the concept of the measured parameters, whereby a *p-value* of <0.05 was considered to be statistically significant.

## **2.3 Biospectroscopy**

### 2.3.1 Sample preparation

Paraffin embedded blocks of endometrial tissues were prepared by the Royal Preston Hospital. A microtome (Surgipath Medical Industries. INC.) was used to cut tissue sections of 10 $\mu$ m thickness from the paraffin embedded blocks of endometrial tissues. Histochoice Clearing Agent (Sigma<sup>®</sup> Life Science, sigma-aldrich.com/ St. Louis, MO, USA) was used to clean the microtome and the bench area prior to cutting of the tissues. The sections were then floated in water at 40°C in a paraffin section mounting bath to be then placed onto low-E reflective microscope slides (Kevley Technologies/ Chesterland, OH, USA) and were then left in air to dry. Dewaxing of tissue sections was achieved through three sequential washes of five minutes each in fresh xylene (Fisher Scientific/ Loughborough, UK) and were then washed and cleared acetone (Fisher Scientific/ Loughborough, UK) for five minutes to ensure removal xylene and thus avoiding tissue contamination. The washes took place in the fume hood. Slides were left again in air to dry and were later placed in petri dishes to be stored in a dessicator, to avoid sample contamination from moisture, until they were examined.

The same eight tissue sections (H09-12890-9/11, H09-9102-1/6, H09-13558-A5/8, H09-12292-1, H09-11708-4) were used for both types of biospectroscopy examination; FTIR and Raman microspectroscopy.

### 2.3.2 Experimental design

Spectra were acquired from five, randomly chosen, different glandular elements of each tissue. Specifically a point map was used to select 10 individual points for spectral acquisition from the luminal and basal cells of each gland and its surrounding connective tissue/stromal cells (a total of 30 spectra per glandular area and 150 spectra per tissue section).

### 2.3.3 FTIR microspectroscopy

Spectral acquisition was recorded using a Thermo Nicolet 6700 FTIR spectrometer coupled to a Nicolet Continuum microscope (Thermo Fisher Scientific/ Waltham, MA, USA) equipped with a KBr beam splitter and a mercury cadmium telluride detector. Liquid nitrogen was used to cool down the spectrometer before, in the middle and at the end of the experiment. The Thermo OMNIC software was used to operate the spectrometer for spectral acquisition in reflection mode and for the conversion of spectra to absorbance. Spectra were collected at a resolution of  $4\text{ cm}^{-1}$  and co-added for 256 scans. The Atlus microscope application was used to visualise the glandular elements through a x15 Replachromat objective lens at a  $40 \times 40\ \mu\text{m}$  aperture size in order to achieve a good signal-to-noise ratio. Background spectra were acquired from a region of no sample, after recording each set of 10 spectral points, for the automatic removal of atmospheric spectra from the sample spectra.

### 2.3.4 Raman microspectroscopy

Spectral acquisition was recorded using an InVia Renishaw Raman spectrometer (Renishaw PLC, Gloucestershire, UK) configured to a 785 nm laser diode operating at 1200 lines per millimetre ( $1\text{ cm}^{-1}$  spectral resolution) diffraction grating whereby Raman signals were detected by a Master Renishaw Pelletier cooled charge-coupled detector camera. Using a white light camera mounted on a Leica microscope, glandular elements and locations for spectral acquisition were visualised through a x50 objective lens. Spectral acquisition was set up using Renishaw Wire 3.1 software whereby an extended

grating scan type and a spectrum range of 500-2000 (Raman shift/  $\text{cm}^{-1}$ ) were selected. Locations for spectral acquisition were exposed to Raman signals for 25 s at 100% laser power and acquisitions were repeated twice.

Whilst the laser power was switched on, a Renishaw silicon calibration source was used to calibrate the system before spectral recording. The ideal wavenumber shifts for the centre of the recorded spectra should be 520.5 waves/cm and for the width it should be 4.1-4.2 waves/cm. The height should be as high as possible. When the recorded spectra exceeded the variation range of 0.1-0.2 waves/cm, calibration process was performed again.

#### 2.3.5 Data pre-processing

Raw spectra were pre-processed using MATLAB R2013b software (TheMathsWorks, Nattick, MA, USA) and the IRootLab software (<http://code.google.com/p/irootlab/>) was applied as a toolbox for data pre-processing and analysis. FTIR spectra were cut between the spectral range 1800-900  $\text{cm}^{-1}$  and were then pre-processed using second order differentiation followed by vector normalisation and wavelet de-noising. Cosmic rays were removed from Raman spectra using Renishaw Wire 3.1. Raman spectra were cut to the region 1750-500  $\text{cm}^{-1}$  and were then rubber band baseline corrected followed by vector normalisation and wavelet de-noising. Rubber band baseline correction eliminates slopes while differentiation methods resolve the overlapped bands as well. Vector normalisation accounts for confounding factors such as variations in sample thickness.

#### 2.3.6 Computational analysis

MATLAB was also used for the application of multivariate analysis to obtain spectral differences between the segregated classes. Principal Component Analysis (PCA), an unsupervised classification technique, was firstly used to reduce the dataset into linear variables, called PCs, and allowing for variance between and within the classes. PCA Pareto charts were plotted to derive the number of PCs

for which PCA was applied. The PCs were chosen according to when the graph started to plateau whilst capturing a very high variance of about 99%. The maximum number though of the chosen PCs was 10. Linear Discriminant Analysis (LDA), a supervised classification technique, was applied to the output of PCA so the whole process is called PCA-LDA. LDA maximises inter-category variance whilst minimizing intra-category variance of the specified classes and thus allowing for optimal class segregation. Segregation of classes, was visualised on score plots whilst the wavenumbers accounting for the segregation were shown on loading plots and cluster vector plots.

Statistical analysis, unpaired t-test and one-way analysis of variance (ANOVA) test, were performed using GraphPad Prism 4.0 Software to determine the significance of segregation. A  $p$  value  $<0.05$  suggested significant difference.

## **2.4 Light Microscopy**

### 2.4.1 Sample preparation

All steps were carried out in the fume hood.

Parallel tissue sections of 4 $\mu$ m in thickness were cut from the paraffin embedded blocks of the samples used in biospectroscopy. They were bathed on microslides (Chance Propper Ltd/ West Midlands, UK) and left in air to dry. Tissue sections were then heated in a 60°C oven for 15 minutes and then de-waxed in xylene for five minutes. Tissues were rehydrated through two minutes washes in 100% and 95% alcohol and a minute wash in 70% alcohol. They were then washed in distilled water for a minute and then stained in Gill 3 haematoxylin (Thermo Scientific/ Waltham, MA, USA) for four minutes. A one minute wash in running tap water was followed by and repeated after differentiation in 1% acid alcohol for three seconds. Bluing reagent Thermo Scientific/ Waltham, MA, USA) was added in the water and left for a minute after which tissues were washed again in running tap water for a minute.



Tissues were left for a minute in 70% alcohol to partially dehydrate and were then stained in Eosin Y alcoholic (Thermo Scientific/ Waltham, MA, USA) for a minute. Dehydration was completed in 95% alcohol for two minutes followed by a two minutes and a minute wash in 100% alcohol. Sections were finally cleared in xylene for five minutes and were mounted with DPX mountant (Sigma® Life Science, sigma-aldrich.com/ St. Louis, MO, USA) and covered with a coverslip.

Images of the tissue sections were taken using the Live video tool on Renishaw Wire 3.1 software of the Raman spectrometer. Glandular areas of interest were visualised through the x5 objective lens. These images were used for histological comparison with observations from biospectroscopy analysis.

# CHAPTER 3

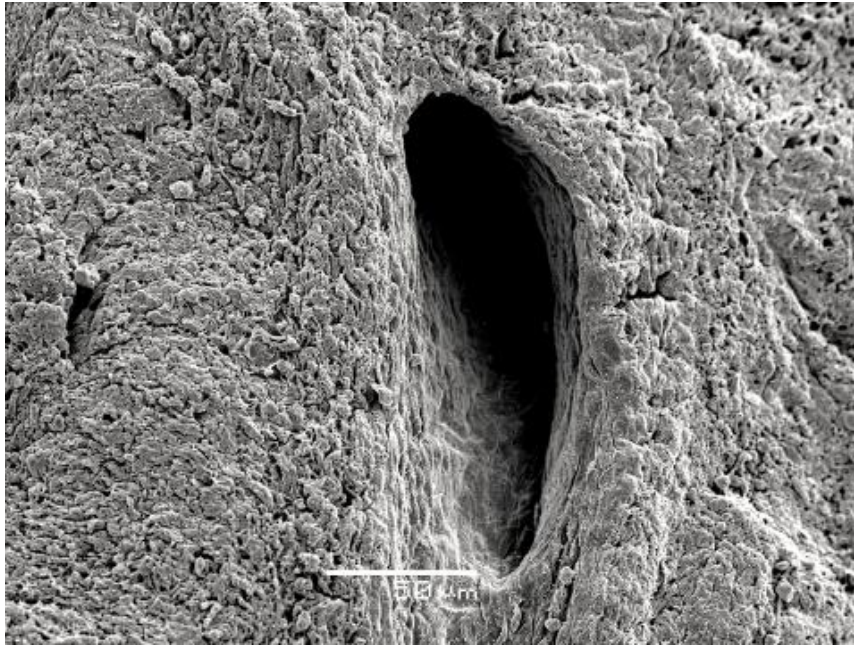
## Results

### **3.1 Scanning Electron Microscopy**

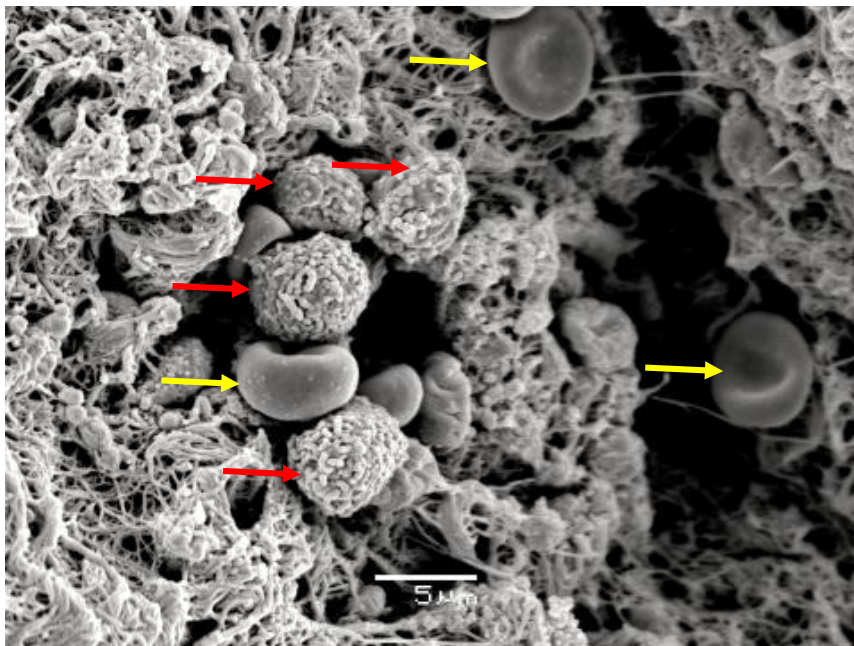
#### 3.1.1 Scanning Electron Photomicrographs

The ultrastructure of human postmenopausal endometrium was examined using SEM (Fig. 3.1-3.38). A menstruating endometrium experiences structural and functional changes during the reproductive years of a female. In postmenopause though, the endometrium becomes atrophic and inactive but it can gain back its regenerative capacity when exposed to exogenous steroid hormones. Most of the related studies so far, characterised menstruating endometrium or endometrium of postmenopausal females taking hormonal therapies. In this study endometrial biopsies taken from patients who experienced postmenopausal bleeding because of endometrial cancer or endometriosis were examined (Table 2.1), however no information was given if the patients had received or were receiving hormonal therapy. Also the provided biopsies were taken from non-diseased sites of the endometrium. The endometrium is a complex tissue and knowledge so far is limited so further research is required to establish a more clear understanding and characterisation.

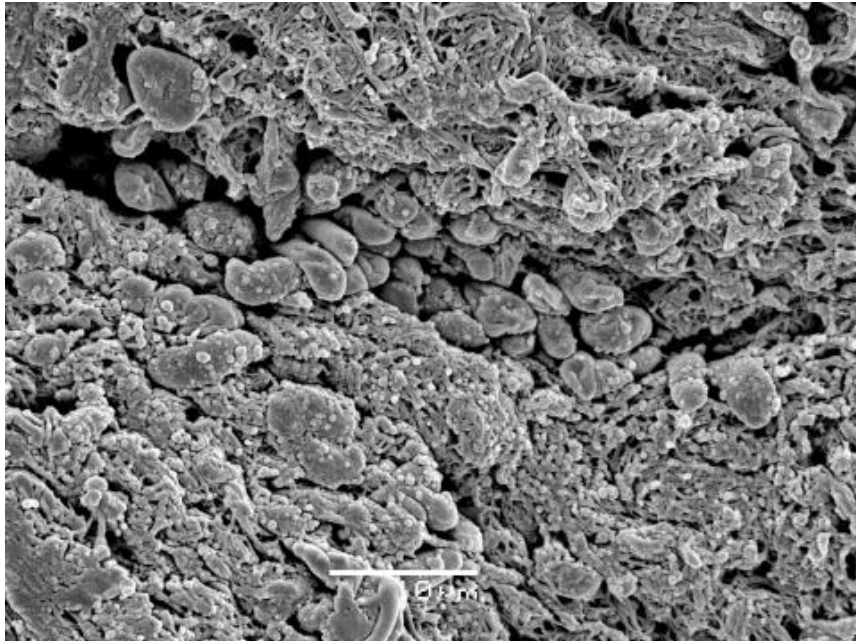
Representative photomicrographs of examined samples are illustrated below.



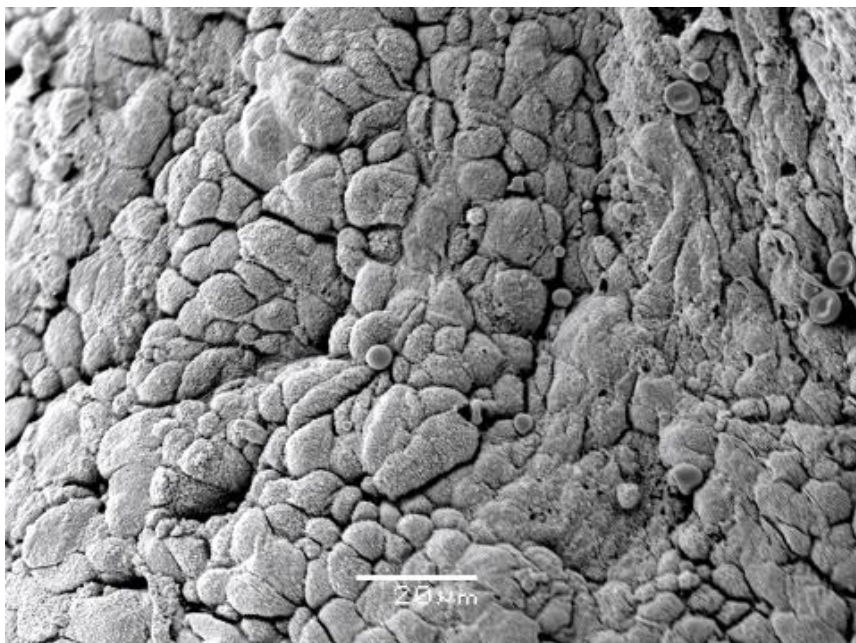
**Figure 3.1:** Scanning electron micrograph of human postmenopausal endometrium. A remaining stumps of an endometrial gland forming a crypt on the surface. The surrounding tissue is mainly composed of fibrin network. Scale bar: 50  $\mu\text{m}$ .



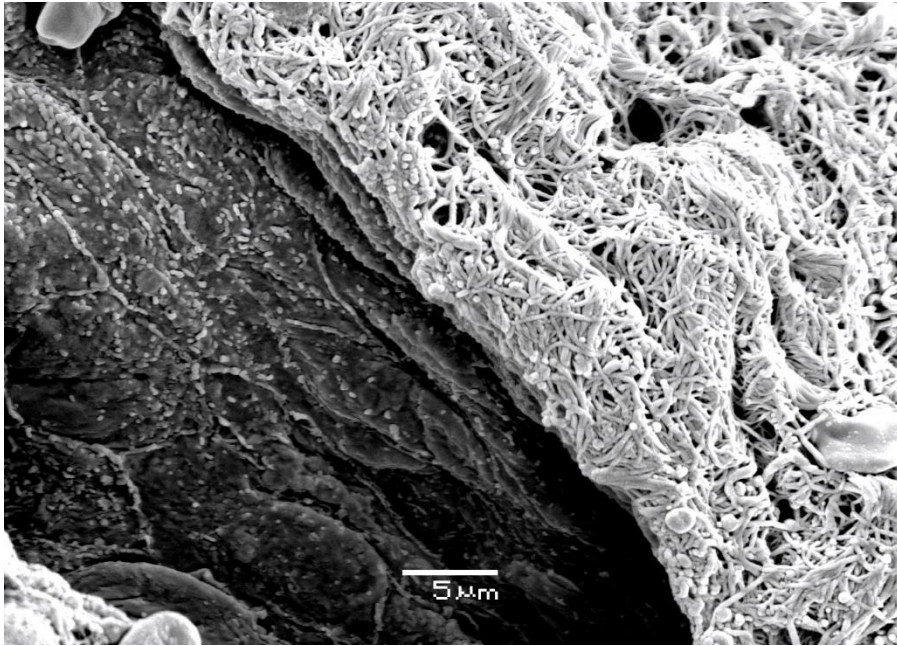
**Figure 3.2:** Scanning electron micrograph of human postmenopausal endometrium. Epithelial cells within the fibrin mesh present a flat pattern of microvilli (red arrows). A few red blood cells can also be distinguished (yellow arrows). Scale bar: 5  $\mu\text{m}$ .



**Figure 3.3:** Scanning electron micrograph of human postmenopausal endometrium. The endometrial surface is mainly covered by fibrin whilst epithelial cells can be identified. In some cells the microvilli are flat whereas in other cells they are absent. Scale bar: 10  $\mu\text{m}$ .

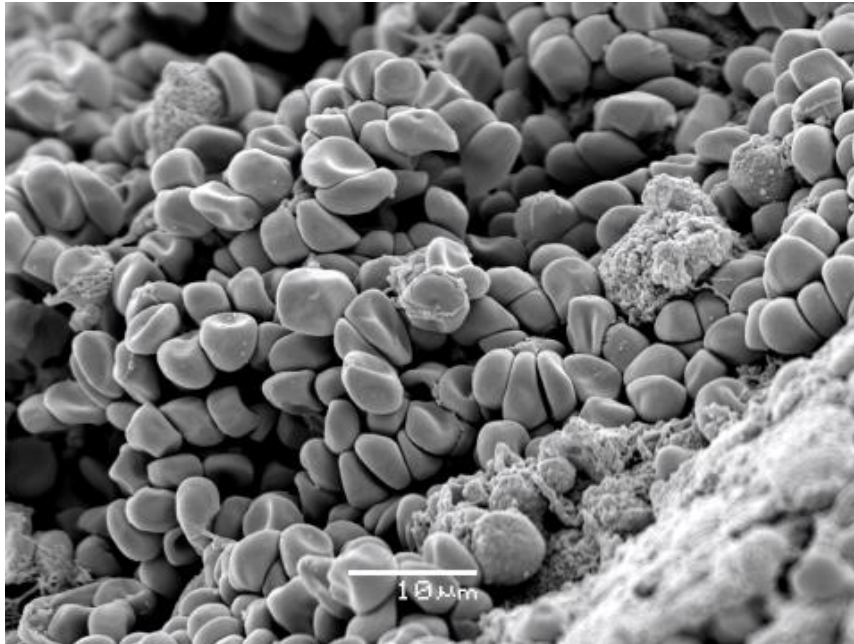


**Figure 3.4:** Scanning electron micrograph of human postmenopausal endometrium. Epithelial cells of various sizes. A few red blood cells can be seen on the right of the image. Scale bar: 10  $\mu\text{m}$ .

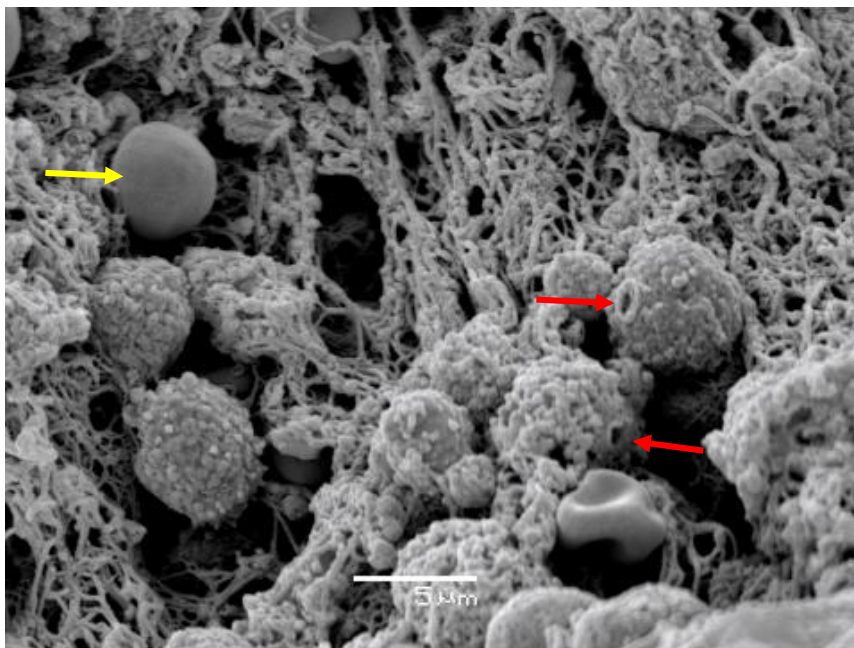


**Figure 3.5:** Scanning electron micrograph of human postmenopausal endometrium. The image was taken at an endometrial crypt which on the outside is surrounded by a fibrin mesh whilst the inside is covered with epithelial surface. However borders of epithelial cells are indiscernible. Scale bar: 5  $\mu\text{m}$ .

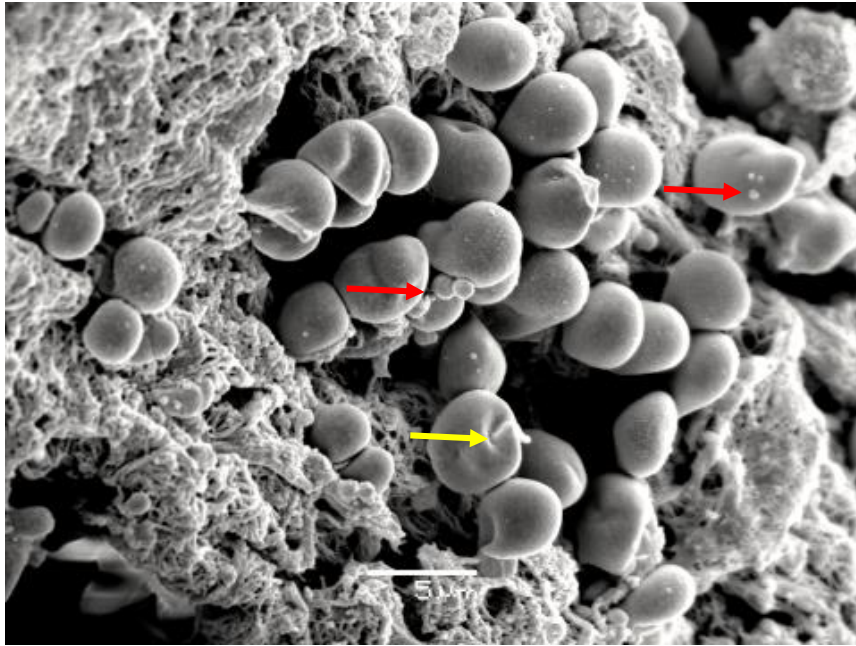
All the above images (Figures 3.1- 3.5) were taken from non-diseased endometrial sites from the same patient who was diagnosed with grade 1 endometrial cancer (sample ECNT 64m see Table 2.1).



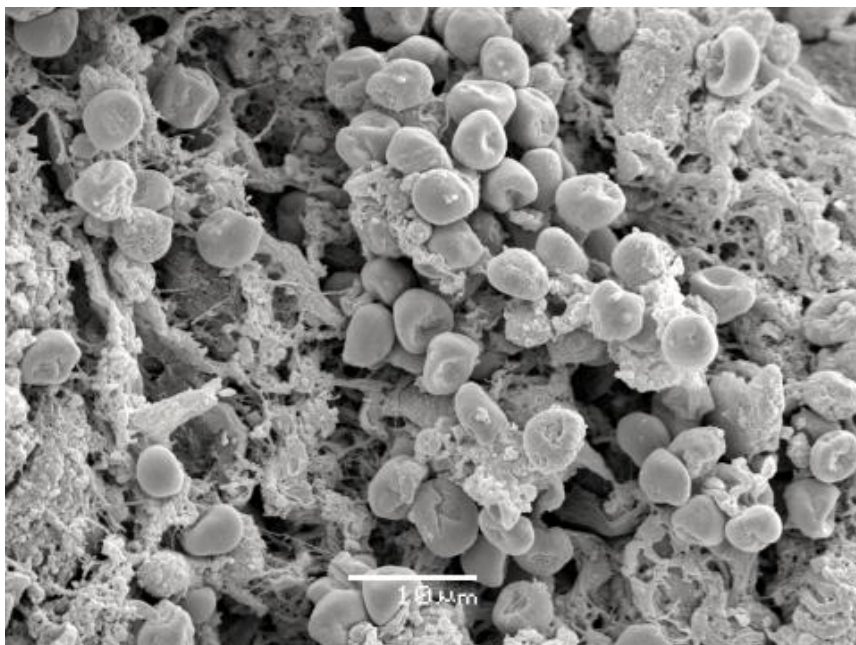
**Figure 3.6:** Scanning electron micrograph of human postmenopausal endometrium. Pinopode-like structures of the same maturation state. They appear to have a more rounded shape rather than biconcave like red blood cells. Scale bar: 10  $\mu\text{m}$ .



**Figure 3.7:** Scanning electron micrograph of human postmenopausal endometrium. Rounded epithelial cells covered with low microvilli. Apical defects can be noticeable (red arrows). Pinopode-like structures arising from the fibrin mesh (yellow arrow). Scale bar: 5  $\mu\text{m}$ .

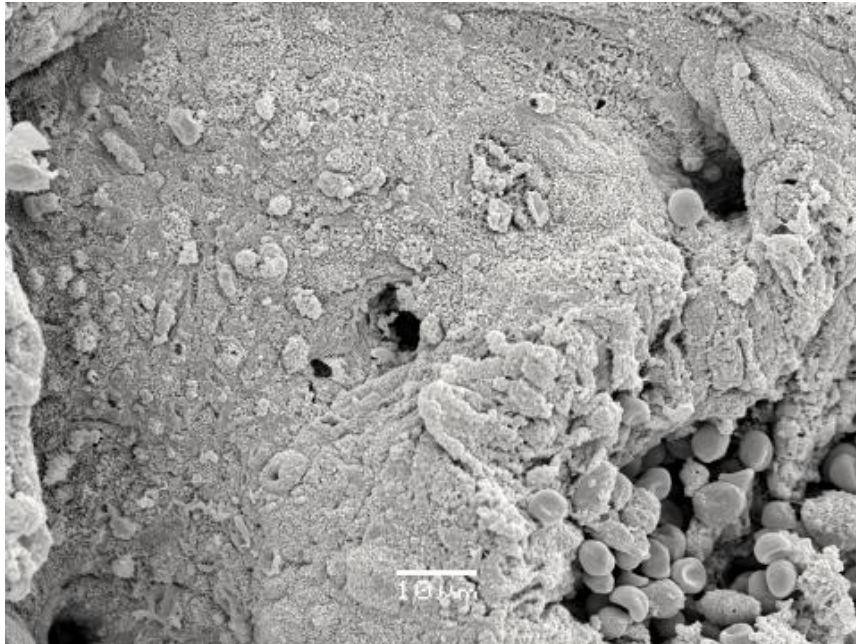


**Figure 3.8:** Scanning electron micrograph of human postmenopausal endometrium. Pinopode-like structures some of which present secretory droplets (red arrows). One is also presenting a cilia-like projection (yellow arrow). Scale bar: 5  $\mu\text{m}$ .



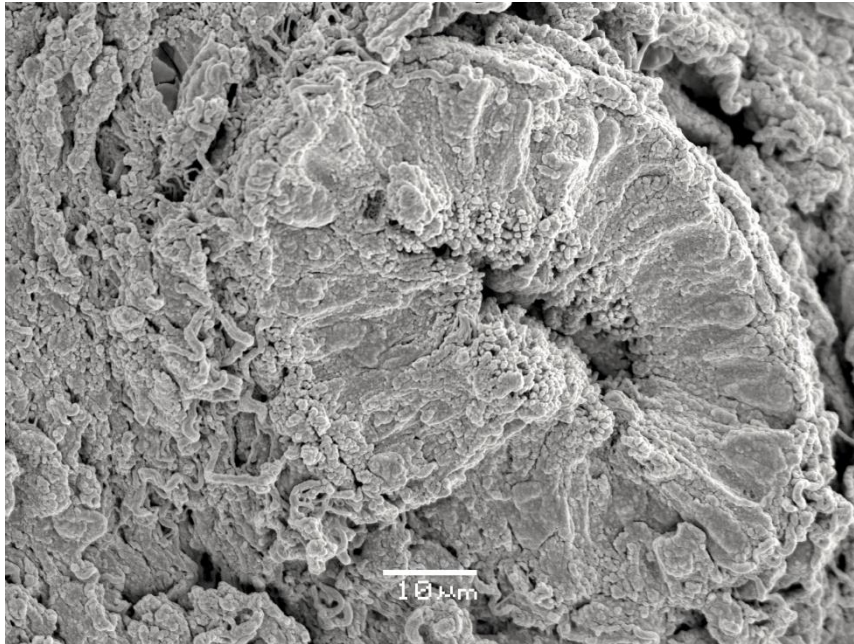
**Figure 3.9:** Scanning electron micrograph of human postmenopausal endometrium. Endometrial surface is covered with pinopode-like structures. Most of the pinopodes appear to be fully developed whilst a few present features of regression. Scale bar: 10  $\mu\text{m}$ .



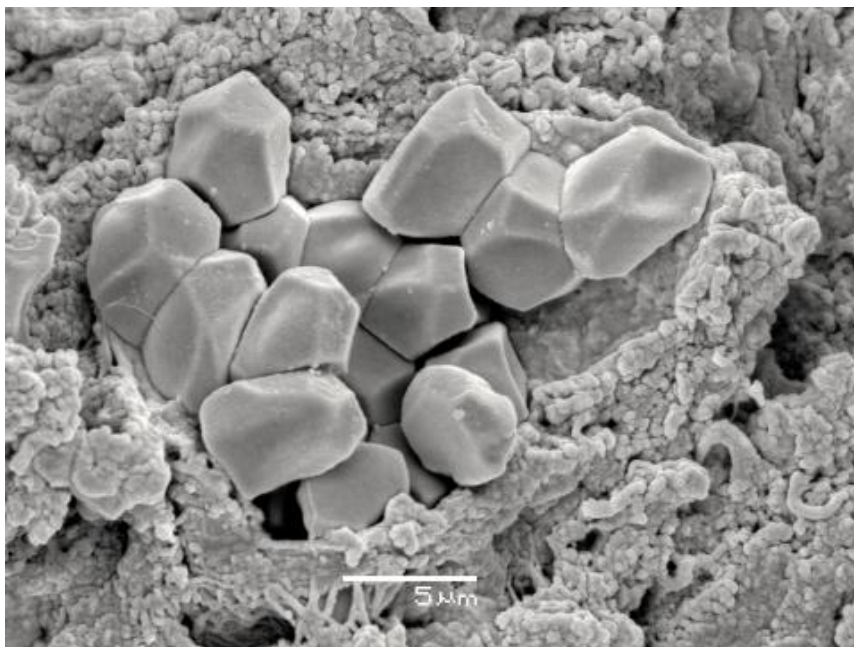


**Figure 3.10:** Scanning electron micrograph of human postmenopausal endometrium. Pinopode-like structures can be seen at the bottom right hand corner. Endometrial surface is covered by flat cells without cell borders. A few glandular stumps protrude from the stroma. Focal disruptions can be noticed at the epithelial surface. Scale bar: 10  $\mu\text{m}$ .

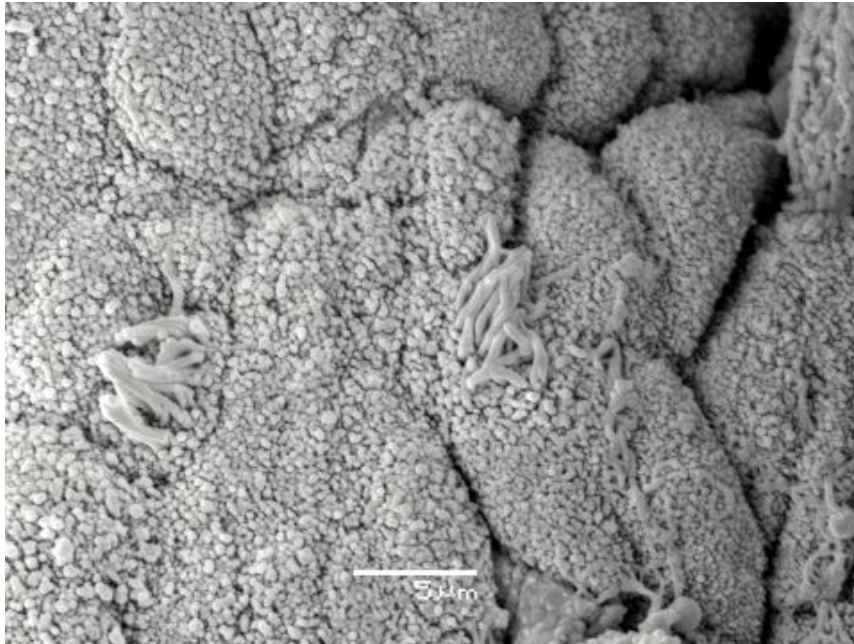
All the above images (Figures: 3.6-3.10) were taken from non-diseased endometrial sites from the same patient who was diagnosed with endometrial cancer (sample ECNT 65, see Table 2.1).



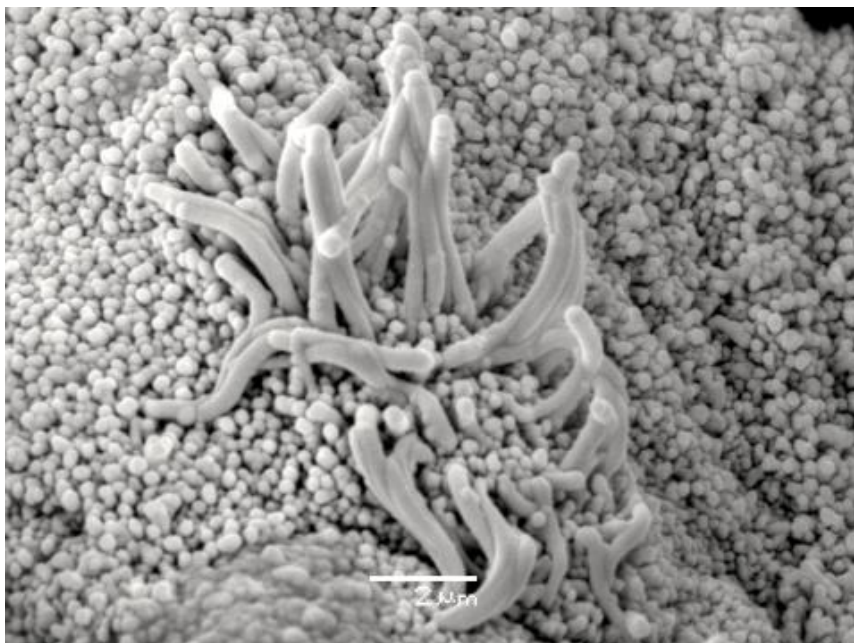
**Figure 3.11:** Scanning electron micrograph of human postmenopausal endometrium. An endometrial gland at the site where the specimen was cut. Borders of epithelial cells are not clearly distinguished, however cells appear to be tall presenting low microvilli at the lumen of the gland. The surrounding stroma is covered by fibrin. Scale bar: 10 μm.



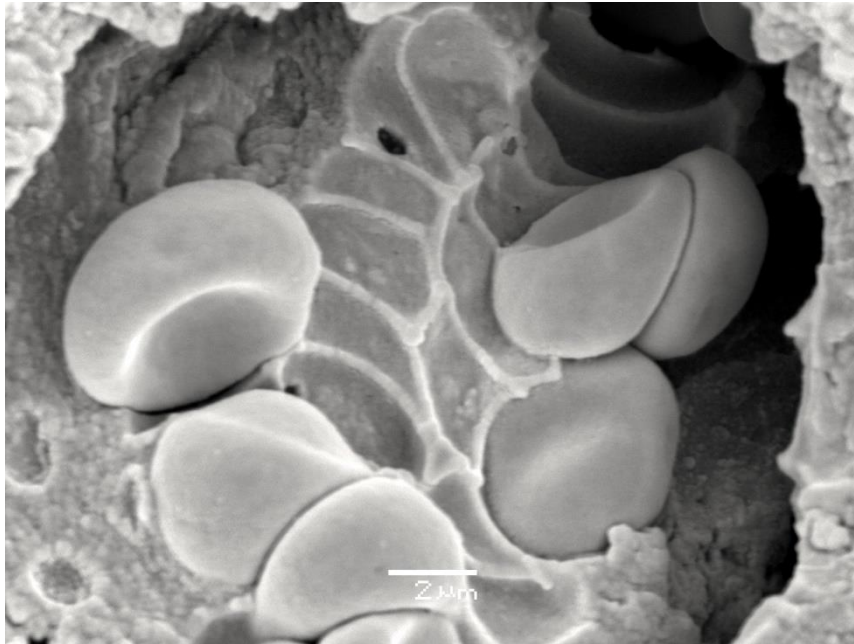
**Figure 3.12:** Scanning electron micrograph of human postmenopausal endometrium. The illustrated cells most probably are red blood cells squashed into strange shapes since endometrial cells have not been reported to have such morphology. Scale bar: 5 μm.



**Figure 3.13:** Scanning electron micrograph of human postmenopausal endometrium. Ciliated cells are created within the lining of the epithelial surface. Epithelial cells have numerous microvilli. The borders of some epithelial cells can be defined. Scale bar: 5  $\mu\text{m}$ .

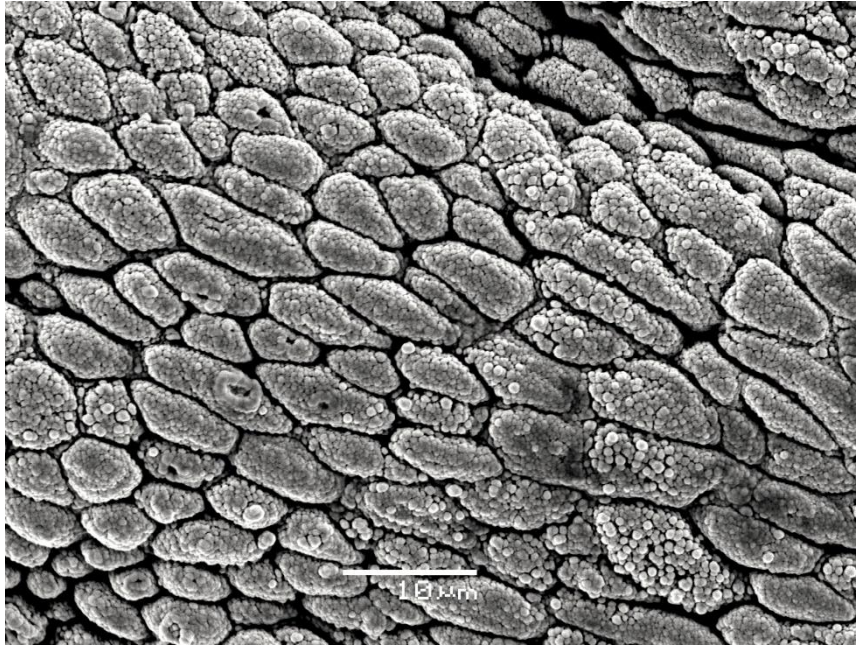


**Figure 3.14:** Scanning electron micrograph of human postmenopausal endometrium. A fully developed ciliated cell found at the epithelial surface. Scale bar: 2  $\mu\text{m}$ .

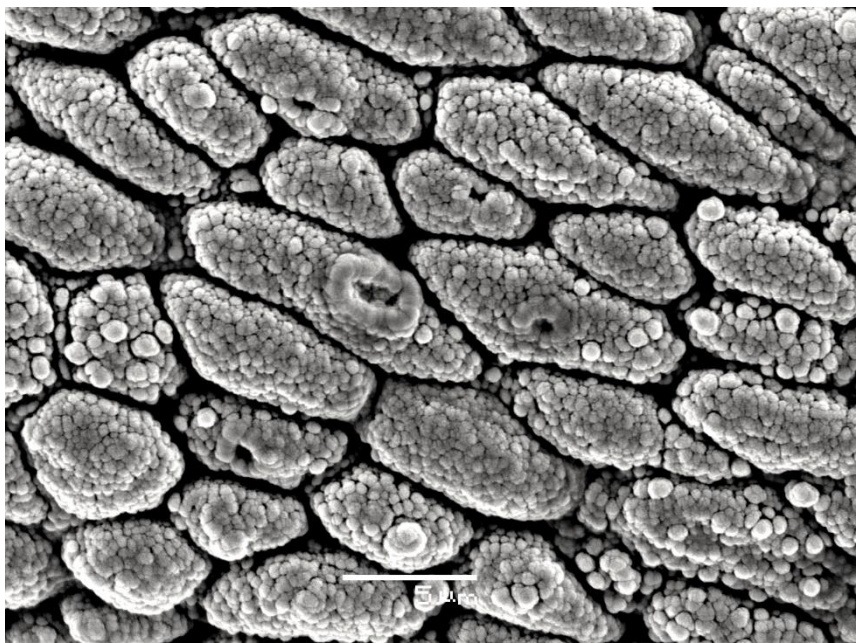


**Figure 3.15:** Scanning electron micrograph of human postmenopausal endometrium. The biconcave shape presented by the cells suggests that they are red blood cells. The endometrial surface with which they are in contact has retained their shape. Scale bar: 2  $\mu\text{m}$ .

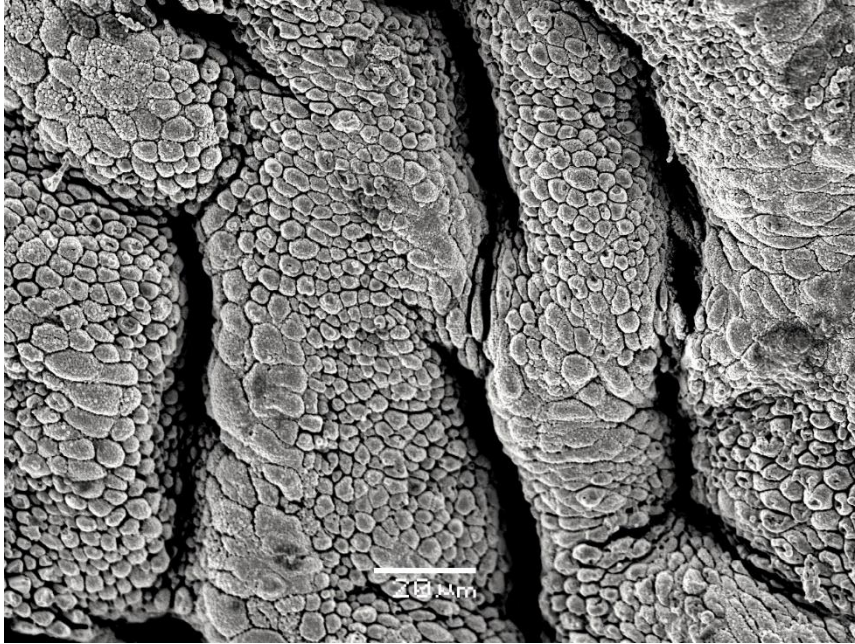
All the above images (Figures 3.11-3.15) were taken from non-diseased endometrial sites from the same patient who was diagnosed with endometrial cancer (sample ECNT 68, see Table 2.1).



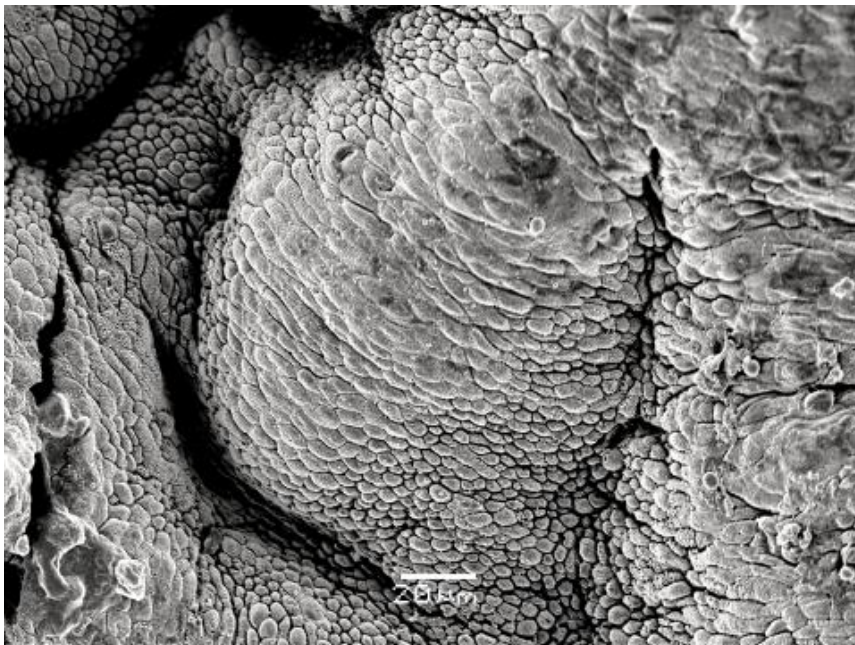
**Figure 3.16:** Scanning electron micrograph of human postmenopausal endometrium. Epithelial cells of various sizes. Their position relative to each other and their shape result in endometrial surface having a cobblestone-like appearance. Scale bar: 10  $\mu\text{m}$ .



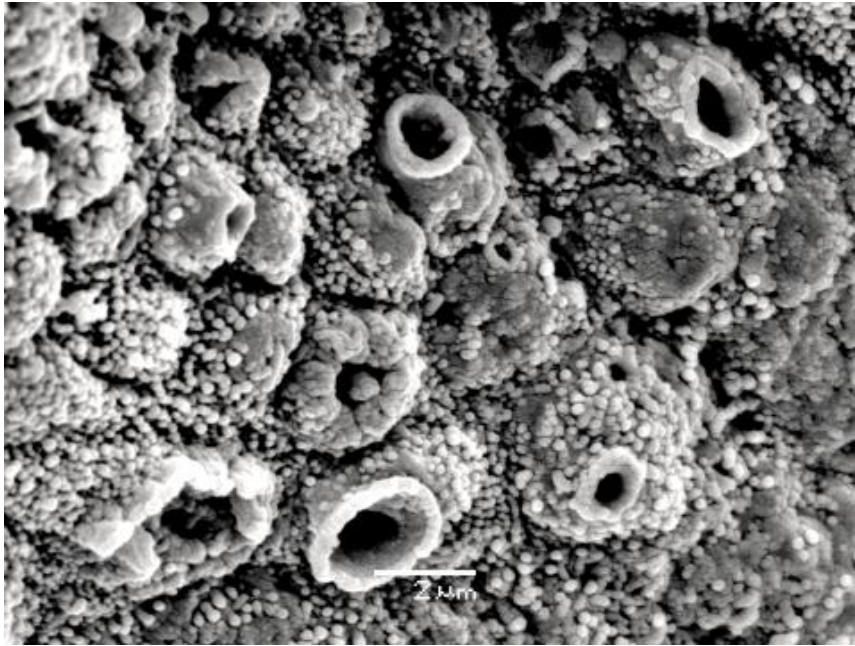
**Figure 3.17:** Scanning electron micrograph of human postmenopausal endometrium. A higher magnification of Figure 3.16. Epithelial cells are covered with low microvilli and a few present apical defects. A space can be noticed separating adjacent cells. Scale bar: 5  $\mu\text{m}$ .



**Figure 3.18:** Scanning electron micrograph of human postmenopausal endometrium. Endometrial surface has a cobblestone appearance. Invaginations are also presented. Epithelial cells are pleomorphic with obvious variations in size and shape. Scale bar: 20  $\mu\text{m}$ .

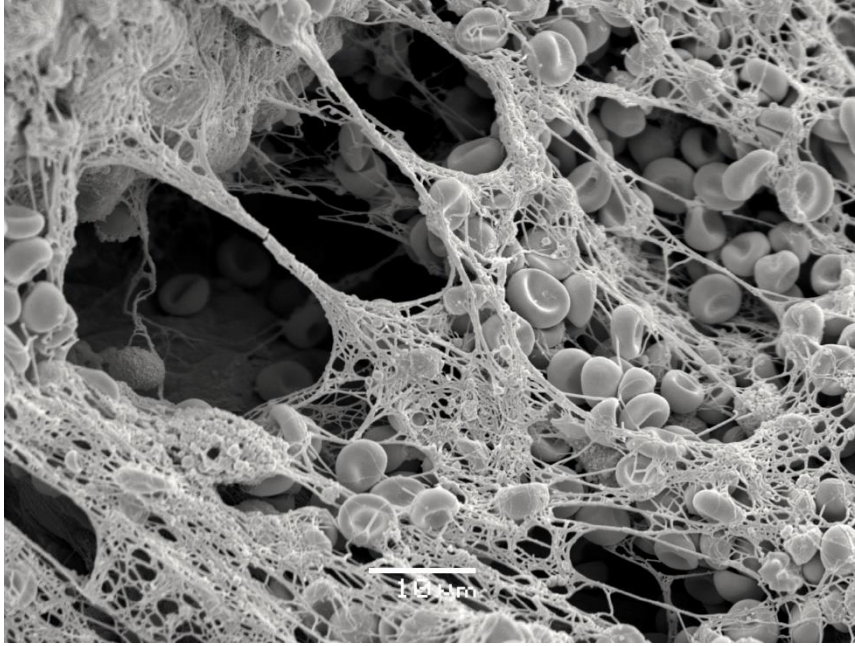


**Figure 3.19:** Scanning electron micrograph of human postmenopausal endometrium. Endometrial cells at the centre of the image appear to be more elongated and fuse with adjacent cells thus cellular borders are lost. Scale bar: 20  $\mu\text{m}$ .

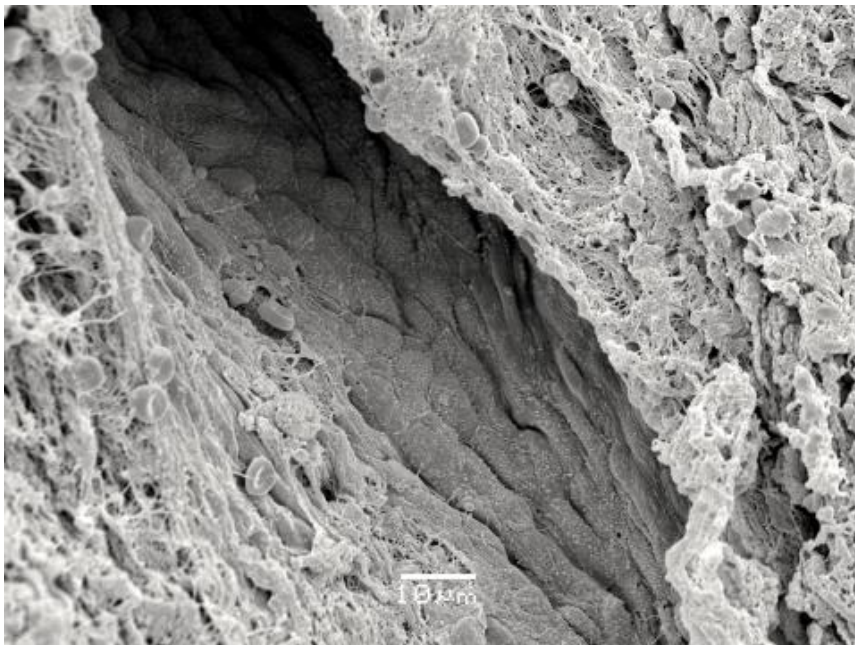


**Figure 3.20:** Scanning electron micrograph of human postmenopausal endometrium. Break-down of the membrane of epithelial cells. Apical defects vary in size among cells. Scale bar: 2  $\mu$ m.

All the above images (Figures 3.16-3.20) were taken from non-diseased endometrial sites from the same patient who was diagnosed with grade 3 endometrial cancer (sample ECNT 69, see Table 2.1).

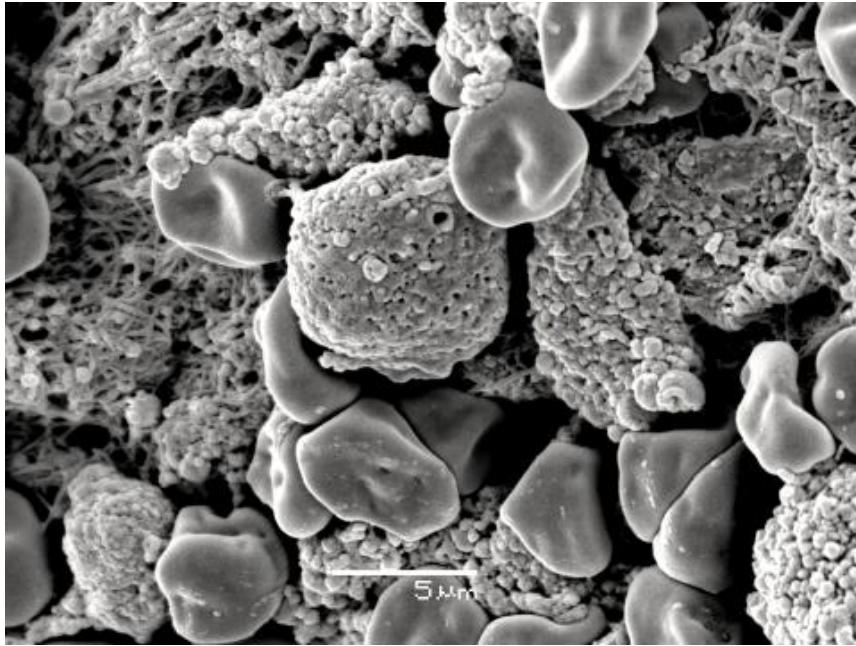


**Figure 3.21:** Scanning electron micrograph of human postmenopausal endometrium. Red blood cells trapped within fibrin mesh. Scale bar: 10 μm.



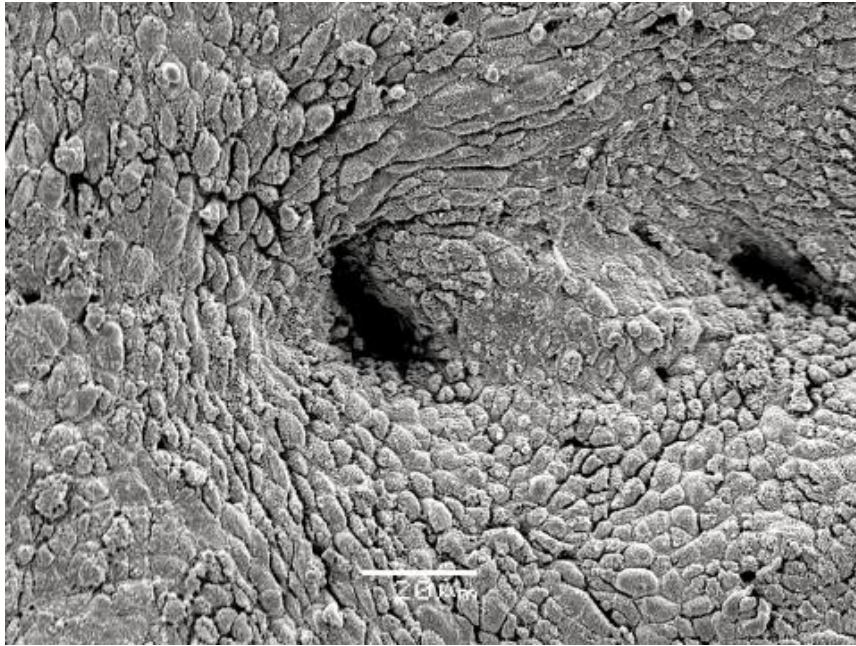
**Figure 3.22:** Scanning electron micrograph of human postmenopausal endometrium. The interior of the endometrial crypt is lined with fused epithelial cells whilst the outside is covered by a fibrinous stroma. Scale bar: 10 μm.



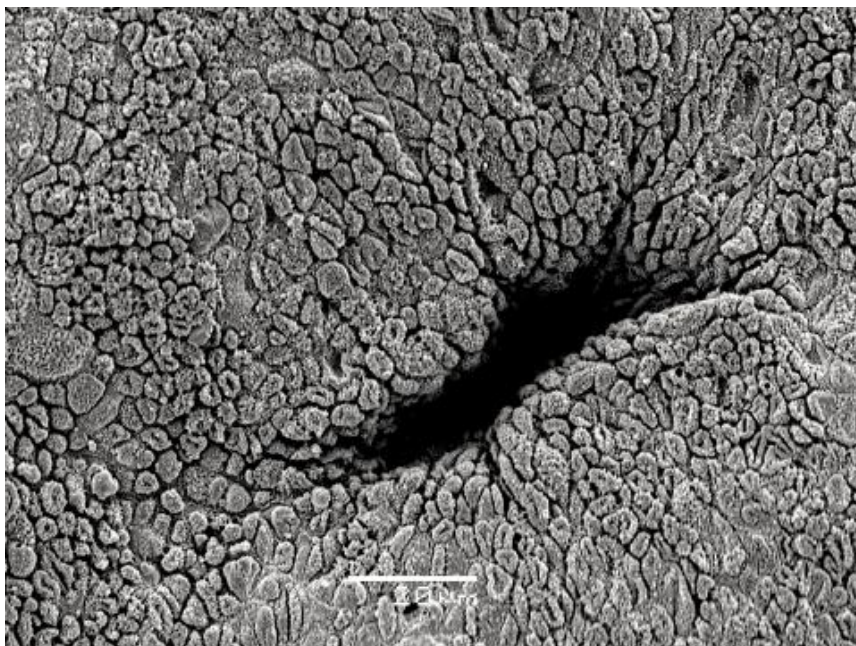


**Figure 3.23:** Scanning electron micrograph of human postmenopausal endometrium. Pinopode-like structures presenting signs of regression. Scale bar: 10  $\mu\text{m}$ .

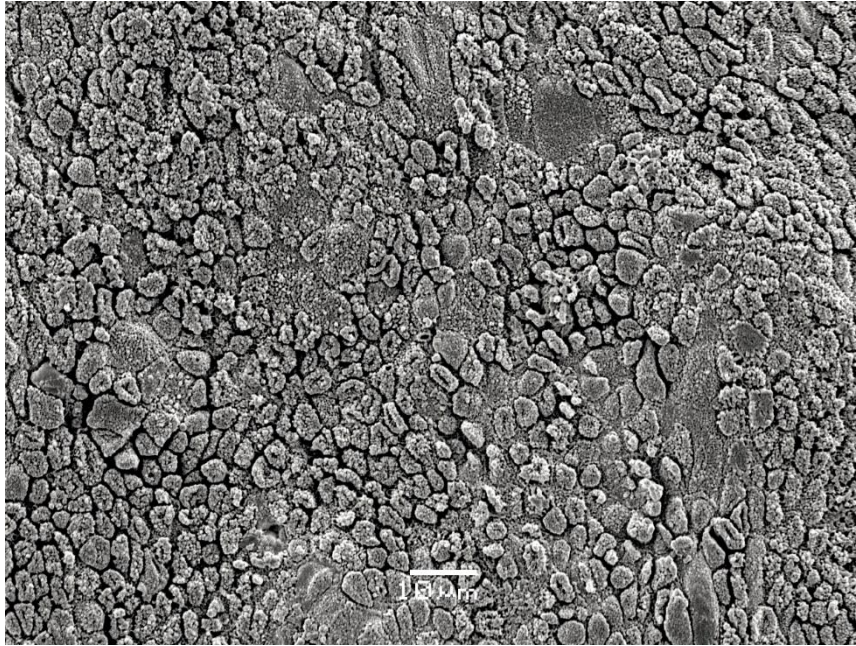
All the above images (Figures 3.21-3.23) were taken from non-diseased endometrial sites from the same patient who was diagnosed with endometrial cancer (sample ECNT 70, see Table 2.1).



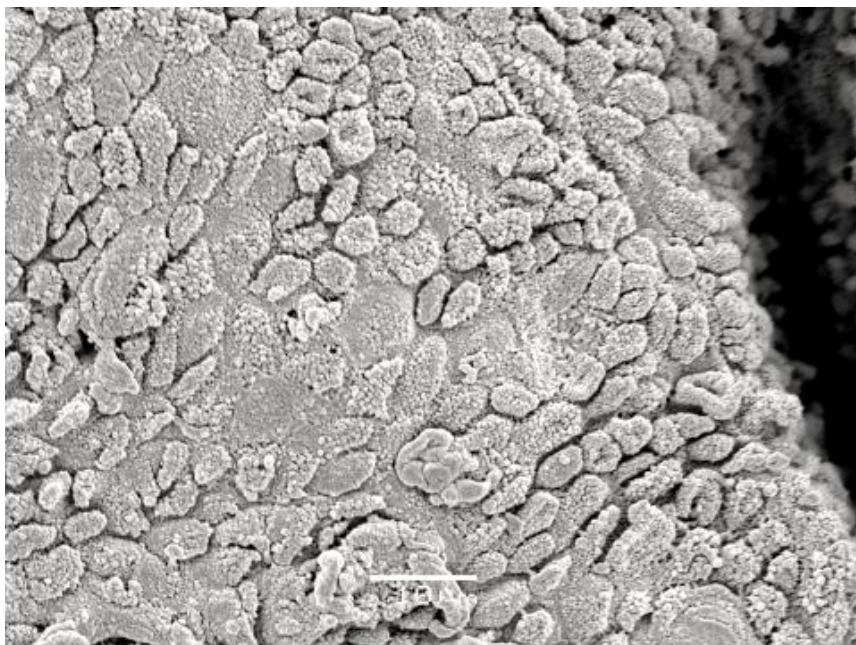
**Figure 3.24:** Scanning electron micrograph of human postmenopausal endometrium. An endometrial gland forming a crypt on the surface. Cells of various sizes and shapes surrounding the crypt. Some show a spiral orientation. Scale bar: 20  $\mu\text{m}$ .



**Figure 3.25:** Scanning electron micrograph of human postmenopausal endometrium. An endometrial crypt from which epithelial cells seem to emerge. Cells present depressions on their surface rather than apical defects. Scale bar: 20  $\mu\text{m}$ .

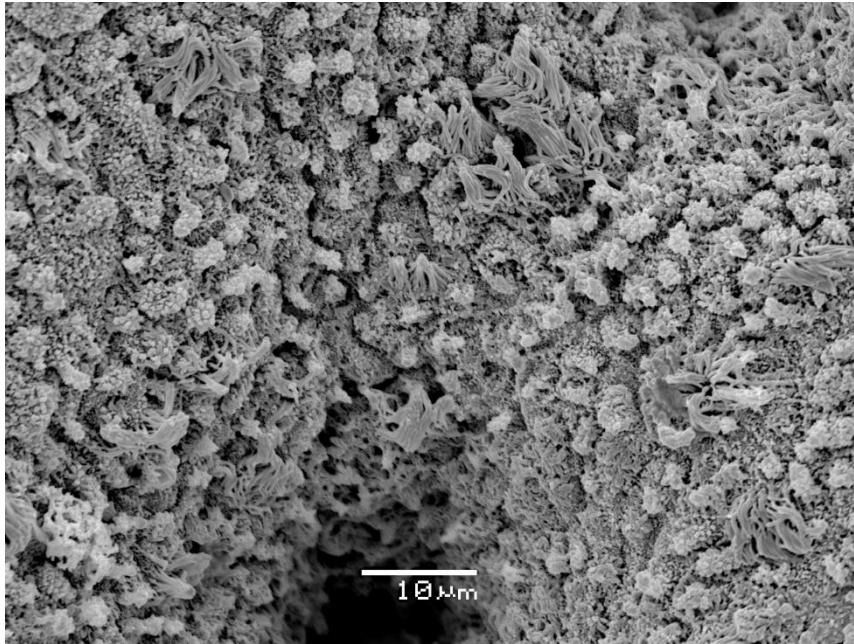


**Figure 3.26:** Scanning electron micrograph of human postmenopausal endometrium. A representative image of the endometrial surface of the specimen. Cell flattening is presented rather than cell bulging like previously observed. A few fields are scant from cells whilst some cells have depressions on their surface. No signs of fibrinous stroma were observed. Scale bar: 10  $\mu$ m.

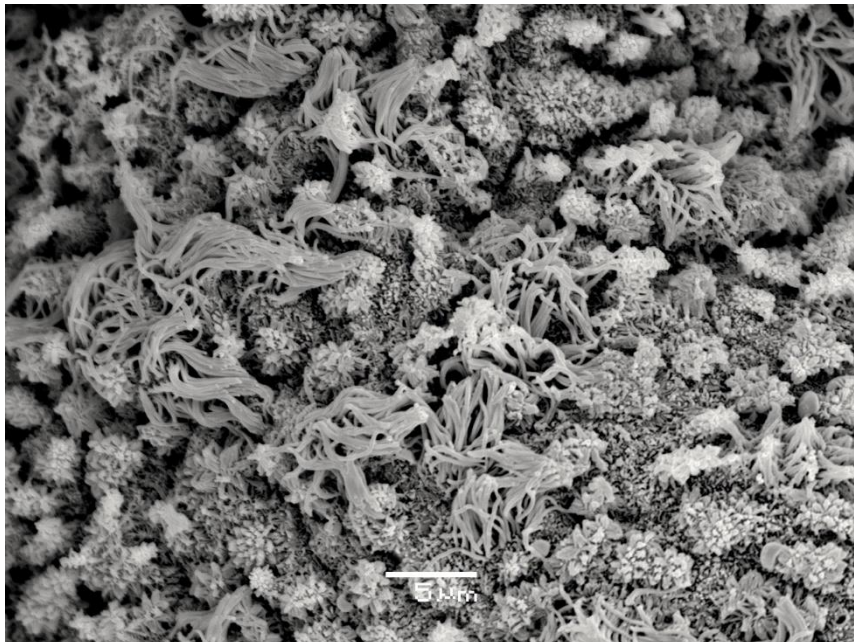


**Figure 3.27:** Scanning electron micrograph of human postmenopausal endometrium. A closer look at the side of the endometrial crypt. Certain fields are interspersed. Even though blurry, cells can be seen at the depths of the crypt. Scale bar: 10  $\mu$ m.

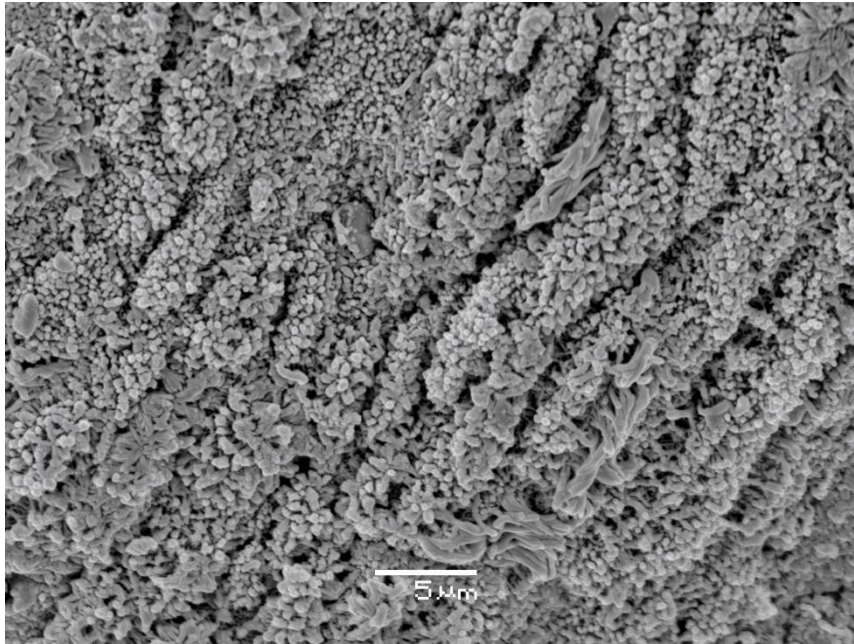
All the above images (Figures 3.24-3.27) were taken from normal post-menopausal endometrium (sample ECNTN 49, see Table 2.1).



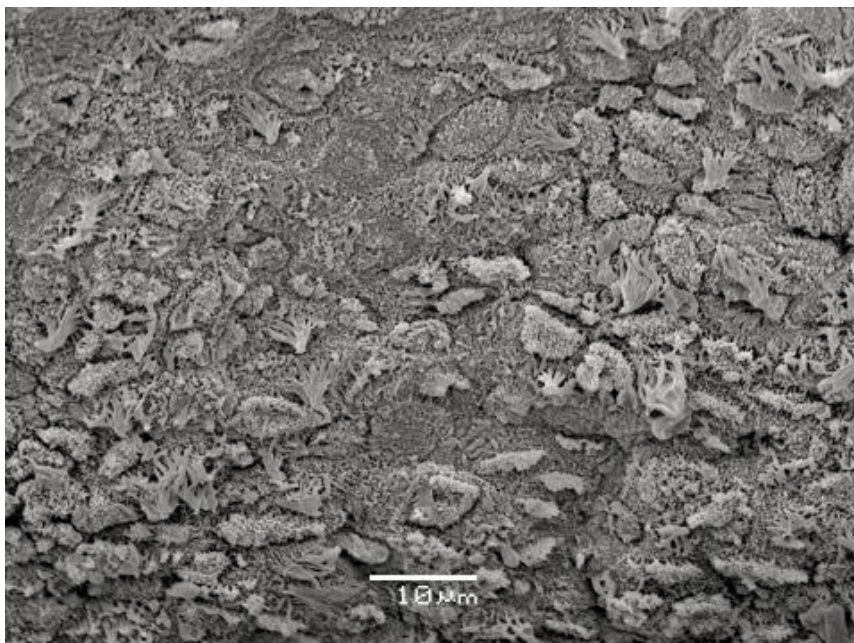
**Figure 3.28:** Scanning electron micrograph of human postmenopausal endometrium. Numerous ciliated cells at the luminal surface of the glandular crypt. Scale bar: 10  $\mu\text{m}$ .



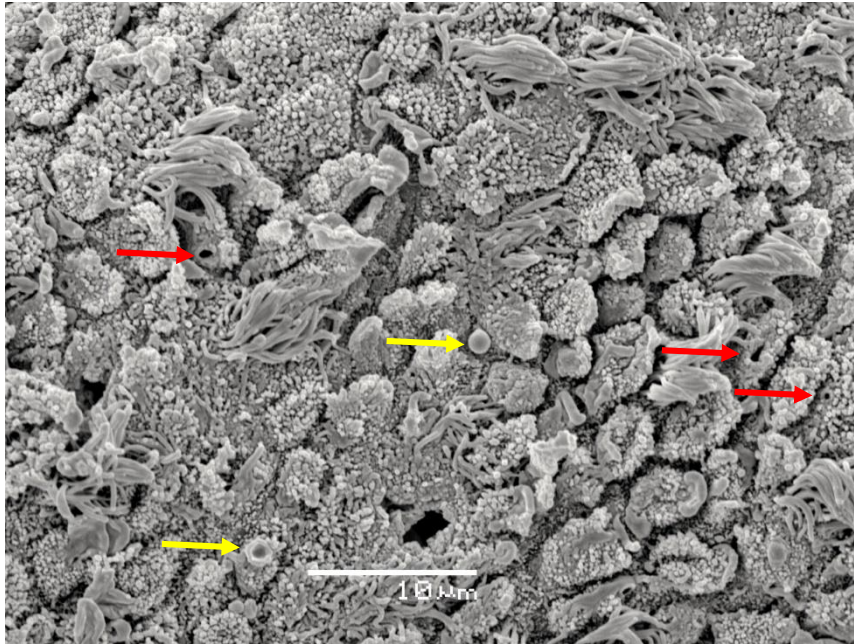
**Figure 3.29:** Scanning electron micrograph of human postmenopausal endometrium. Ciliated cells arise within the epithelial surface. Cilia vary in their length whilst epithelial cells are covered by a dense microvillus net. Scale bar: 5  $\mu\text{m}$ .



**Figure 3.30:** Scanning electron micrograph of human postmenopausal endometrium. Shorter cilia than previously observed. Epithelial cells seem to be elongated however their cell borders are not clearly identified. Scale bar: 5  $\mu\text{m}$ .

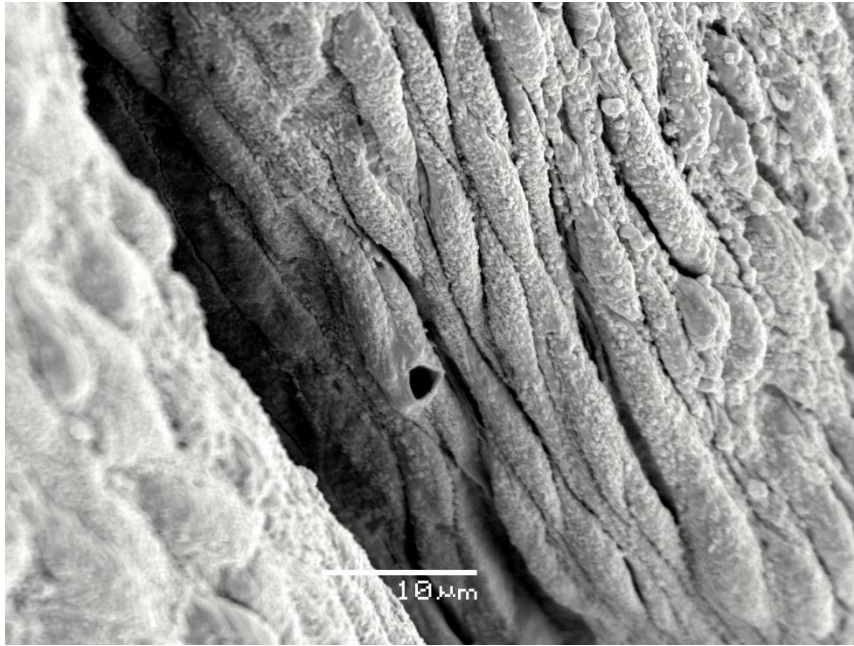


**Figure 3.31:** Scanning electron micrograph of human postmenopausal endometrium. Ciliated cells are rather spread and their cilia seem to regress. Epithelial cells are flat and scattered and vary in morphology. Scale bar: 10  $\mu\text{m}$ .

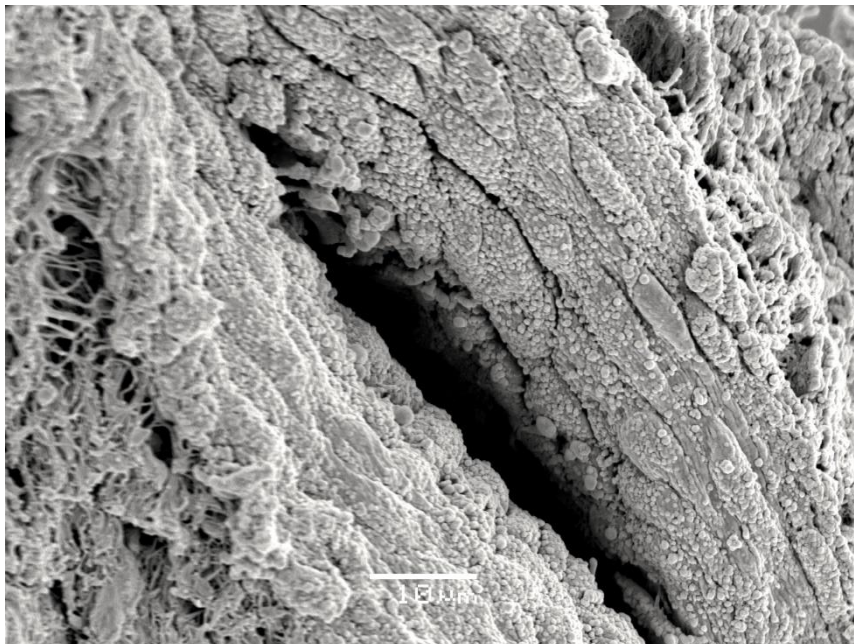


**Figure 3.32:** Scanning electron micrograph of human postmenopausal endometrium. Epithelial surface interspersed with ciliated cells. Surface of epithelial cells is covered by microvilli whilst a few present apical defects (red arrows). Secretory droplets-like can also be seen (yellow arrows). Scale bar: 10  $\mu\text{m}$ .

All the above images (Figures 3.28-3.32) were taken from normal post-menopausal endometrium of a patient diagnosed with endometriosis (sample ECNTN 50, see Table 2.1).

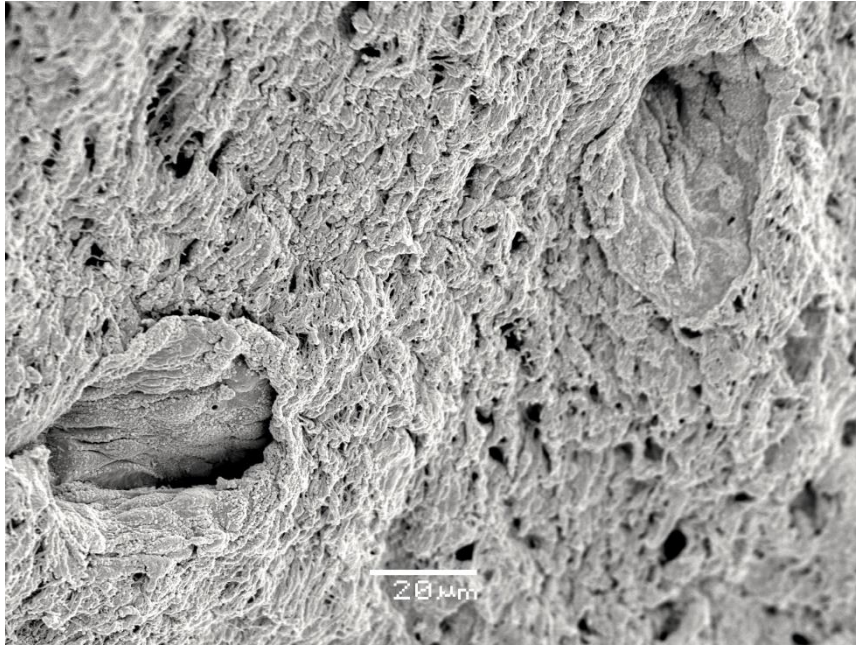


**Figure 3.33:** Scanning electron micrograph of human postmenopausal endometrium. Epithelial cells at the margin of a glandular opening. Cells are elongated and fuse with each other. Defects on the cell surface can also be noticed. Scale bar: 10  $\mu\text{m}$ .

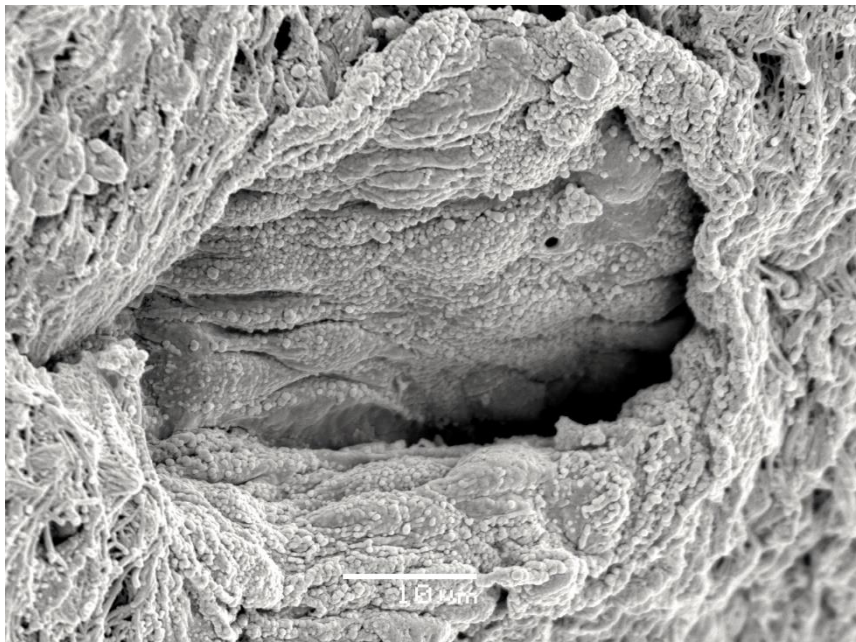


**Figure 3.34:** Scanning electron micrograph of human postmenopausal endometrium. Epithelial cells at the margins of a glandular opening which is surrounded by a fibrinous stroma. Fusiform cells display a spiral growth. Scale bar: 10  $\mu\text{m}$ .

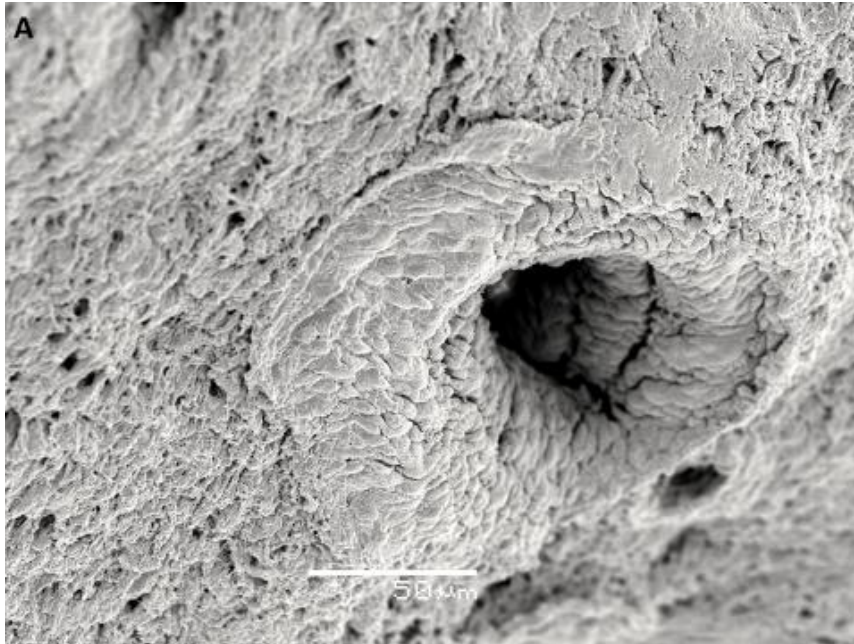




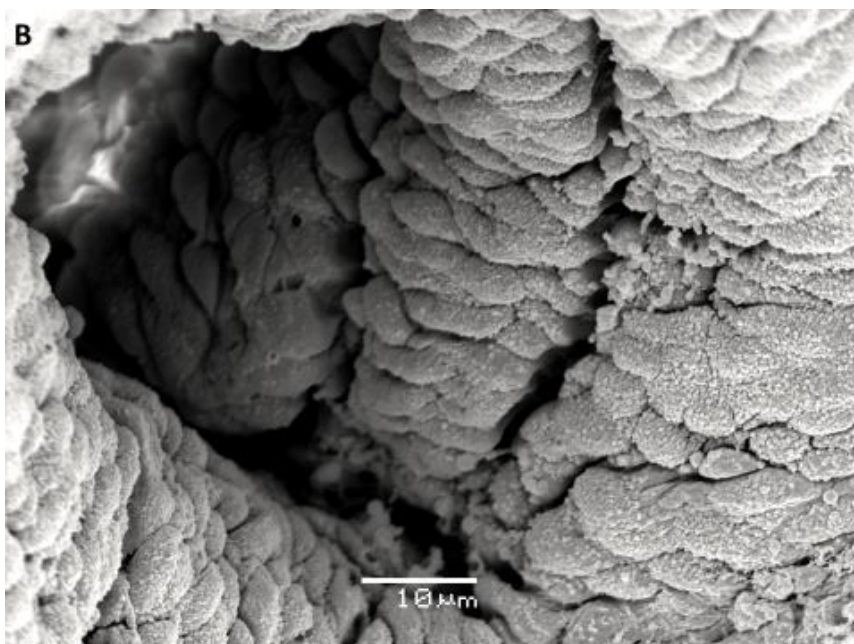
**Figure 3.35:** Scanning electron micrograph of human postmenopausal endometrium. Two glandular stumps on the endometrial surface surrounded by fibrin mesh. Scale bar: 20  $\mu\text{m}$ .

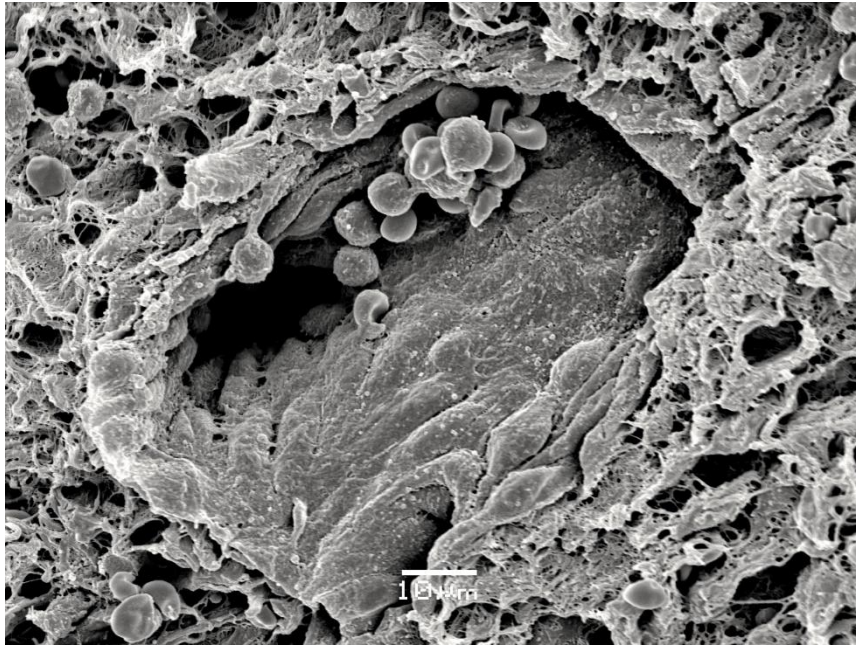


**Figure 3.36:** Scanning electron micrograph of human postmenopausal endometrium. Higher magnification of glandular opening in Figure 3.1.1.35. Fused epithelial cells line the internal surface of the opening one of which presents an apical defect. Scale bar: 10  $\mu\text{m}$ .



**Figure 3.37:** Scanning electron micrograph of human postmenopausal endometrium. (A) An endometrial gland opens through the fibrinous surface. Scale bar: 50  $\mu\text{m}$ . (B) At higher magnification layers of epithelial cells can be seen at the margins of the glandular opening. Borders of some adjacent cells are vaguely identified. Scale bar: 10  $\mu\text{m}$ .





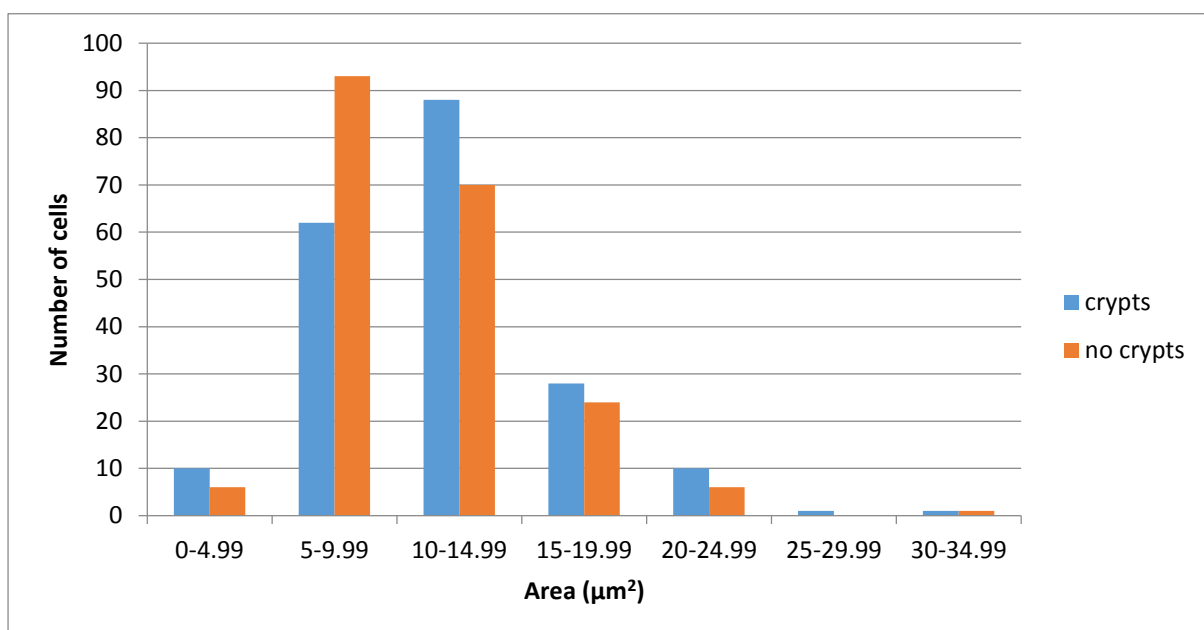
**Figure 3.38:** Scanning electron micrograph of human postmenopausal endometrium. Opening of an endometrial gland from which red blood cells and macrophages come out. Scale bar: 10  $\mu\text{m}$ .

All the above images (Figures 3.33-3.38) were taken from normal post-menopausal endometrium of a patient diagnosed with endometriosis (sample ECNTN 51, see Table 2.1).

### 3.1.2 Image Analysis

The images obtained during SEM examination, specifically the images taken from sample ECNTN 49, were used to assess the ultrastructure of cells around from the crypts of endometrial glands and cells away from the crypts. Figures 3.39-3.46 illustrate the results obtained from analysis of the parameters (Area, Perimeter, Diameter Max, Diameter Mean, Diameter Min, Shape Factor, Sphericity, Aspect Ratio) of the two groups of cells and the statistical analysis (Unpaired t-test)

**A**



**B**

<b>P value</b>	0.1302
<b>Mean ± SEM of crypts (n=200)</b>	11.68 ± 0.3259
<b>Mean ± SEM of no crypts (n=200)</b>	11.00 ± 0.3086

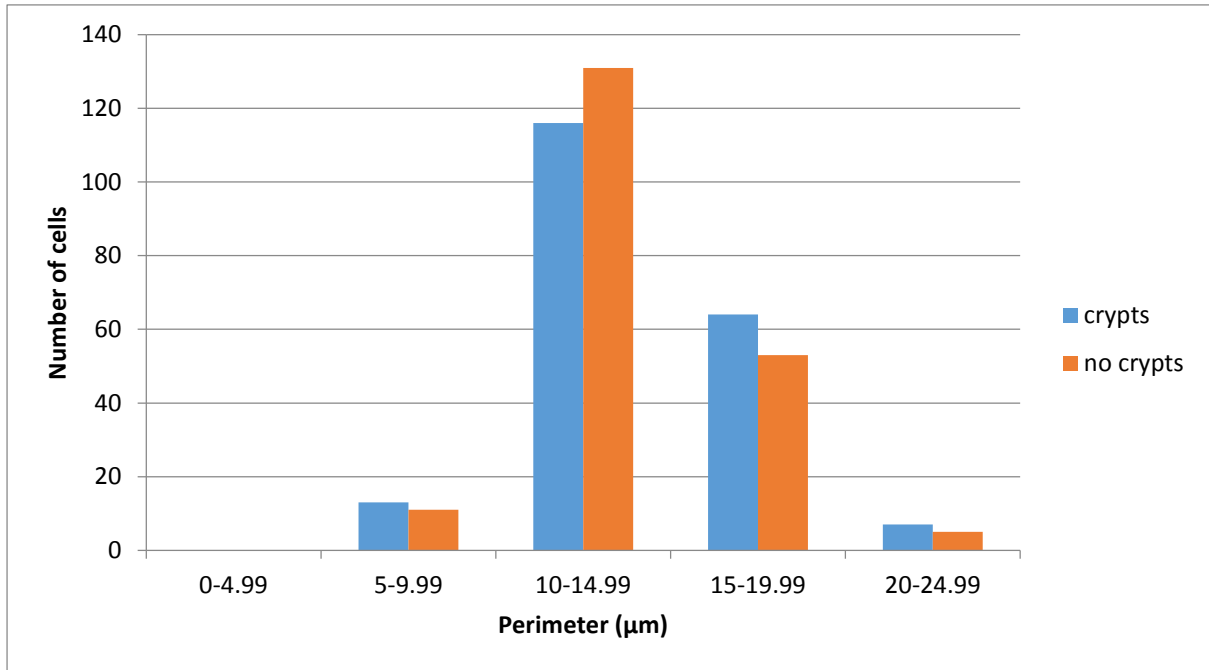
**Figure 3.39: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured area. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured from each group was 200.

The obtained results suggested that groups of cells were not significantly different from each other.

The mean area of cells around from the crypts was higher than the mean of cells away from the crypts.

Neither group maintained higher number of cells through the ranges of measured area.

**A**



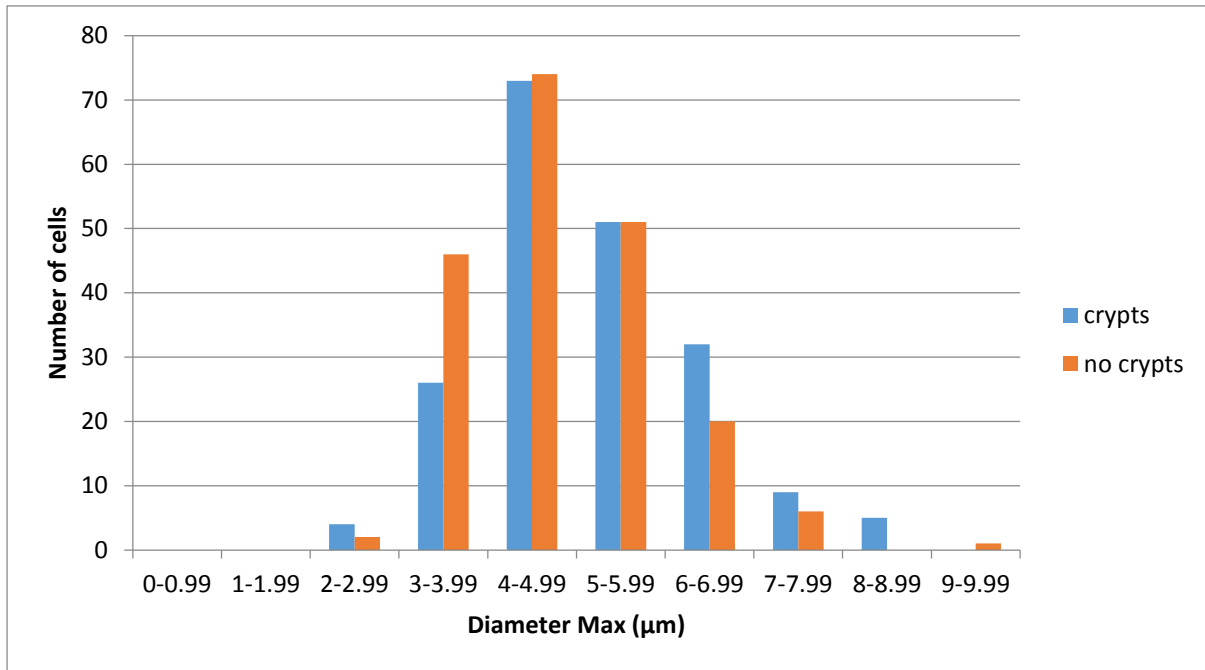
**B**

<b>P value</b>	0.0194
<b>Mean ± SEM of crypts (n=200)</b>	14.30 ± 0.2058
<b>Mean ± SEM of no crypts (n=200)</b>	13.63 ± 0.1955

**Figure 3.40: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured perimeter. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were significantly different from each other. The mean perimeter of cells around from the crypts was higher than the mean of cells away from the crypts. Number of cells away from the crypts was higher only for the region 10-14.99 μm.

**A**



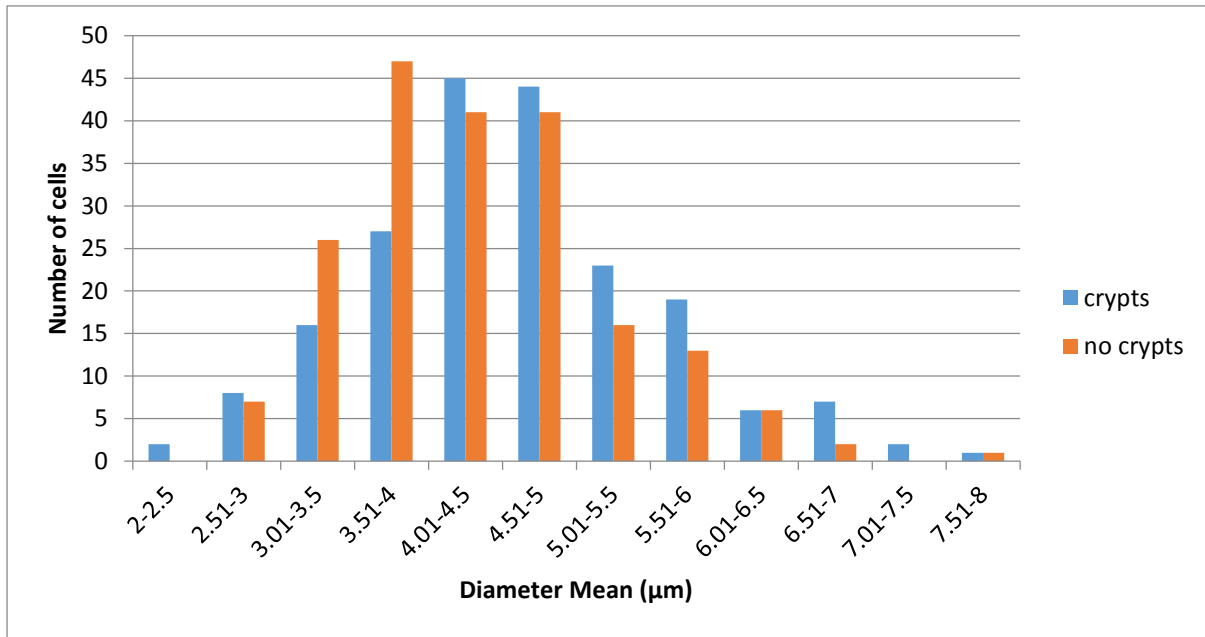
**B**

<b>P value</b>	0.0025
<b>Mean ± SEM of crypts (n=200)</b>	5.182 ± 0.08580
<b>Mean ± SEM of no crypts (n=200)</b>	4.838 ± 0.07388

**Figure 3.41: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured diameter max. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were significantly different from each other. The mean diameter max of cells around from the crypts was higher than the mean of cells away from the crypts. Neither group maintained higher number of cells through the ranges of measured diameter max.

**A**



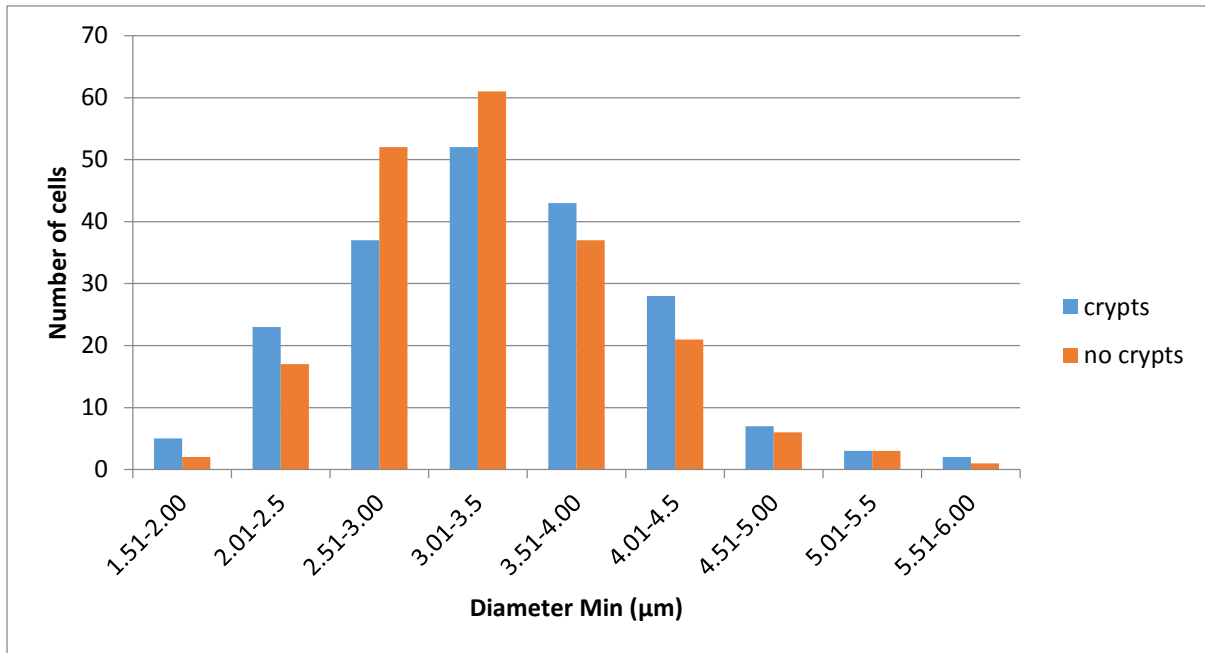
**B**

<b>P value</b>	0.0038
<b>Mean ± SEM of crypts (n=200)</b>	4.599 ± 0.07082
<b>Mean ± SEM of no crypts (n=200)</b>	4.326 ± 0.06181

**Figure 3.42: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured diameter mean. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were significantly different from each other. The mean diameter mean of cells around from the crypts was higher than the mean of cells away from the crypts. Neither group maintained higher number of cells through the ranges of measured diameter mean.

**A**



**B**

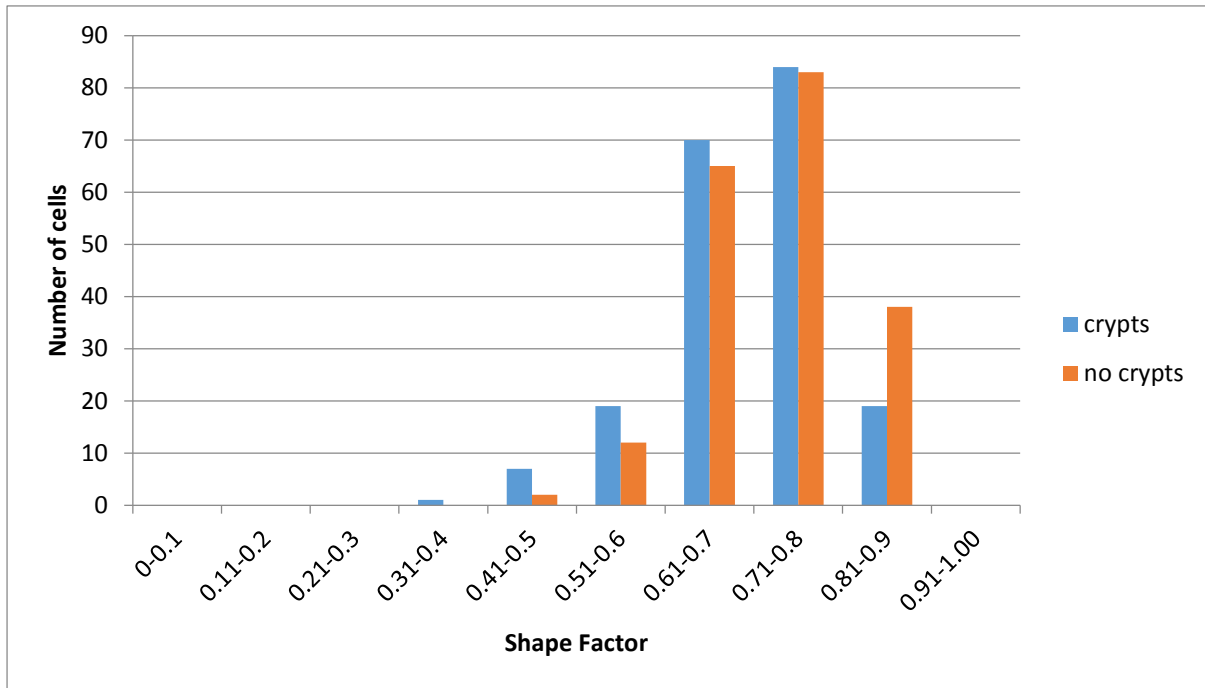
<b>P value</b>	0.5891
<b>Mean ± SEM of crypts (n=200)</b>	3.356 ± 0.05314
<b>Mean ± SEM of no crypts (n=200)</b>	3.318 ± 0.04715

**Figure 3.43: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured diameter min. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were not significantly different from each other. The mean diameter min of cells around from the crypts was slightly higher than the mean of cells away from the crypts. Neither group maintained higher number of cells through the ranges of measured diameter min.



**A**



**B**

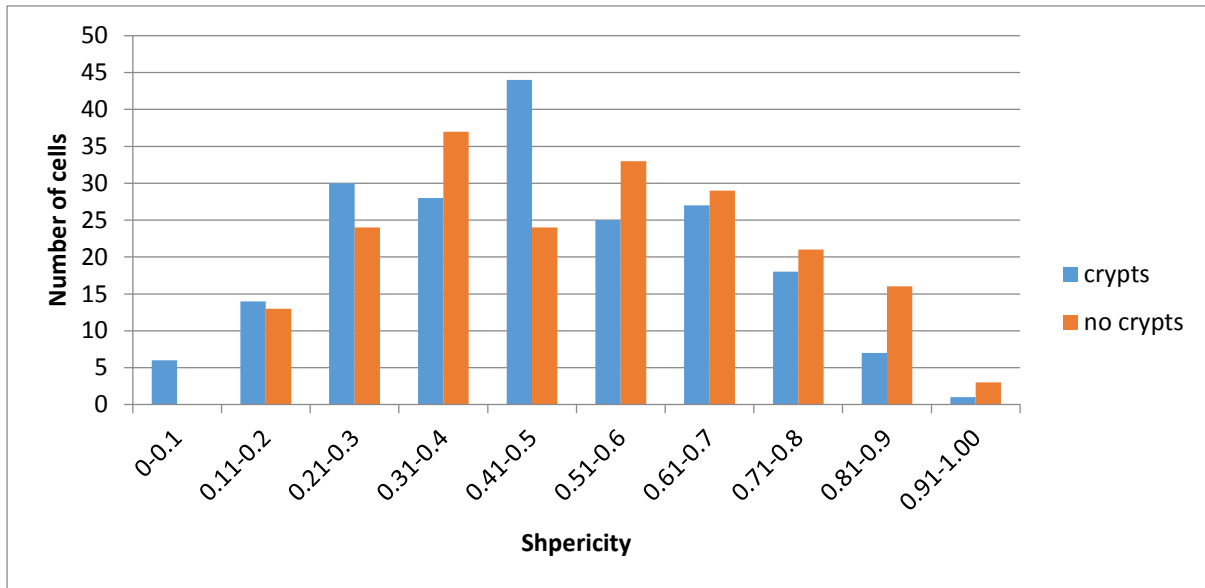
<b>P value</b>	0.0631
<b>Mean ± SEM of crypts (n=200)</b>	0.7065 ± 0.008933
<b>Mean ± SEM of no crypts (n=200)</b>	0.7265 ± 0.005908

**Figure 3.44: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured shape factor. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were not significantly different from each other.

The mean shape factor of cells away from the crypts was higher than the mean of cells around from the crypts. Number of cells away from the crypts was higher only for the region 0.81-0.9.

**A**



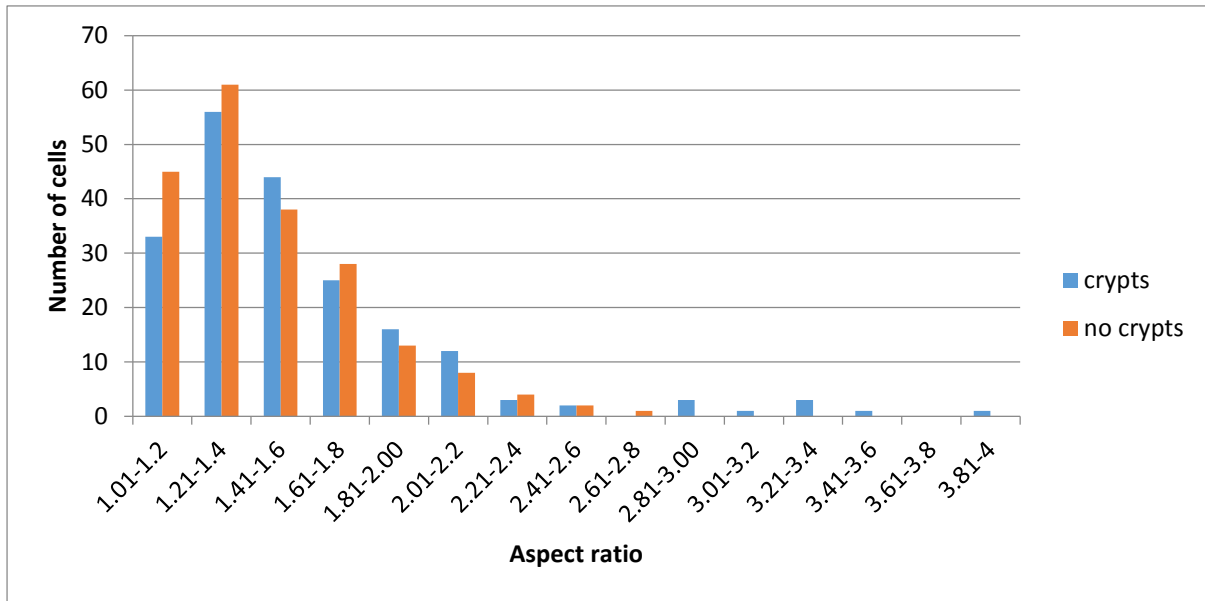
**B**

<b>P value</b>	0.0121
<b>Mean ± SEM of crypts (n=200)</b>	0.4606 ± 0.01393
<b>Mean ± SEM of no crypts (n=200)</b>	0.5111 ± 0.01440

**Figure 3.45: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured sphericity. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were significantly different from each other. The mean sphericity of cells away from the crypts was higher than the mean of cells around from the crypts. Neither group maintained higher number of cells through the ranges of measured sphericity.

**A**



**B**

<b>P value</b>	0.0049
<b>Mean ± SEM of crypts (n=200)</b>	1.580 ± 0.03357
<b>Mean ± SEM of no crypts (n=200)</b>	1.465 ± 0.02288

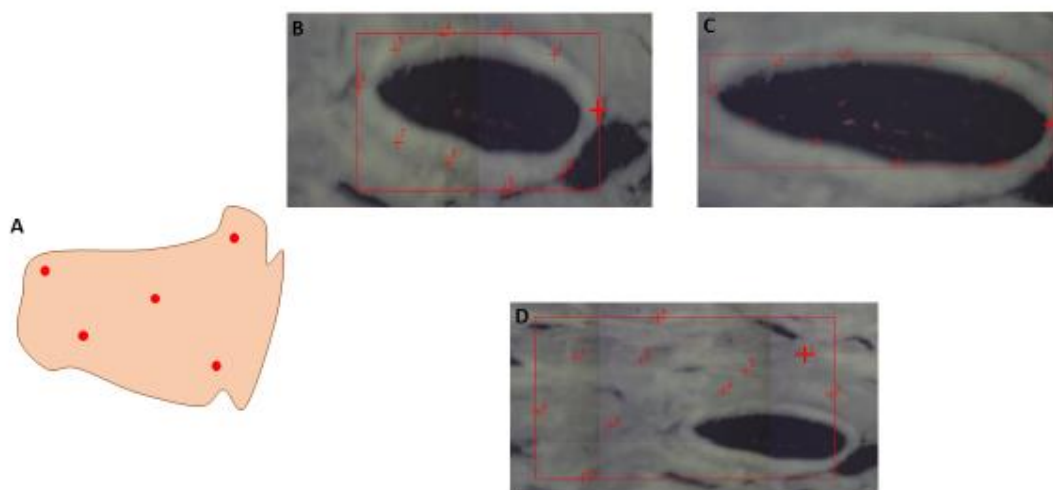
**Figure 3.46: (A)** Number of cells surrounding the crypts (blue) and away from the crypts (orange) for each range of measured aspect ratio. **(B)** Results of statistical analysis by employment of Unpaired t-test. The total number of cells measured in each group was 200.

The obtained results suggested that groups of cells were significantly different from each other. The mean aspect ratio of cells around from the crypts was higher than the mean of cells away from the crypts. Neither group maintained higher number of cells through the ranges of measured area however both groups show a decrease in cell numbers after the range 1.21-1.4 until 2.61-2.8.

### 3.2 FTIR spectroscopy

Tissue samples were interrogated using FTIR spectroscopy in the hope of identifying potential biomarkers for basal, luminal and stromal cells by analysing their biochemical composition. Diseased tissue sections, H09-12890-9, H09-9102-1 and H09-13558-A5, were taken from cancerous lesions of the endometrium. Their corresponding non-diseased tissues, H09-12890-11, H09-9102-6 and H09-13558-A8, were taken from normal sites of the endometrium away from the cancerous lesions. The interrogated non-diseased tissue samples which did not have corresponding diseased samples were H09-12292-1 and H09-11708-4. The first was taken from a patient with endometriosis but for the latter sample no information was provided about the health status of the patient. More information about the samples can be found in Table 2.1.

Five randomly selected glandular areas from each sample were exposed to FTIR spectroscopy whereby spectra was collected from the basal, luminal cells of the glands and stromal cells surrounding the glands (Fig. 3.47). A x15 Replachromat objective lens was used to visualise and select the glandular elements. Spectra were collected from 10 points corresponding to the location of basal epithelial cells (hypothesised location of epithelial stem/progenitor cells) in a gland (Fig. 3.47 D), 10 points from locations of luminal epithelial cells (Fig. 3.47 C), which is adjacent to the lumen of the gland, and 10 points from stromal cells from the surrounding connective tissue (Fig. 3.47 E). A total of 30 spectral points were selected in each glandular element and a total of 150 spectral points per tissue sample.



**Figure 3.47: Experimental design – an example of a tissue sample.** (A): shape of interrogated tissue section indicating the locations of the 5 different glandular areas, (B): 10 selected points for spectral acquisition from basal cells, (C): 10 selected points for spectral acquisition from luminal cells and (D): 10 selected points for spectral acquisition from surrounding stromal cells.

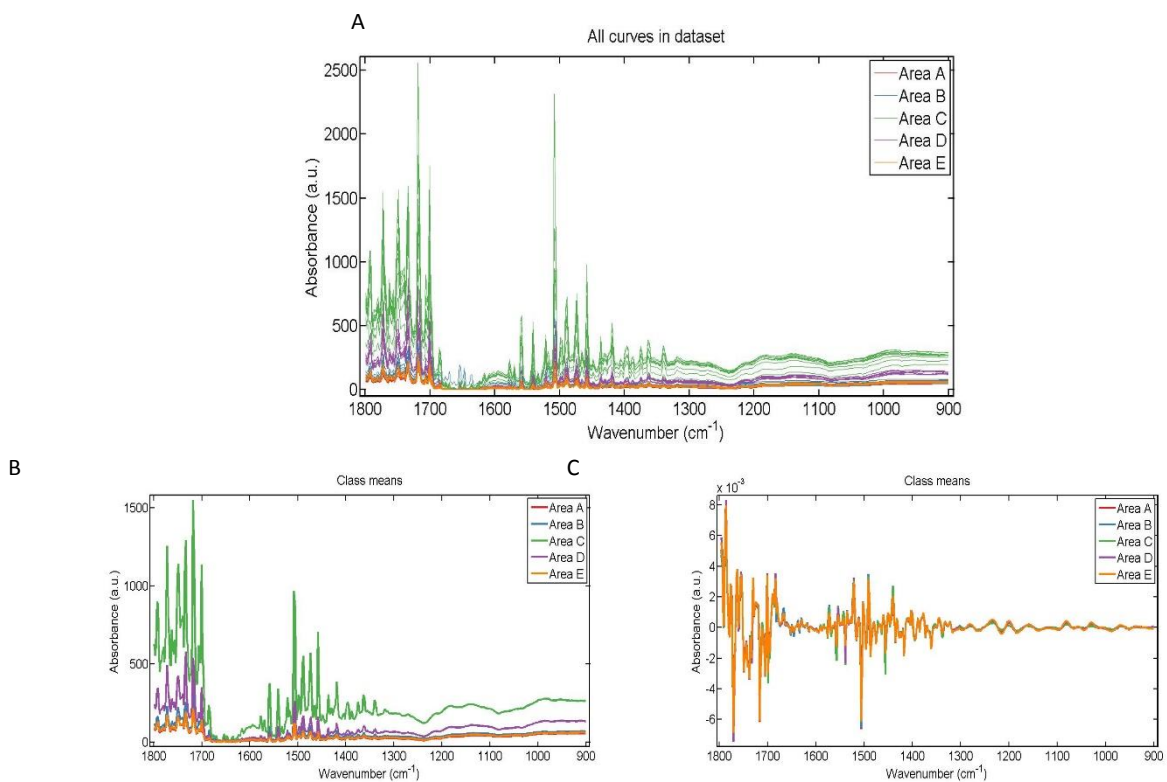
The raw spectral fingerprint acquired from the cells was pre-processed by application of second-order differentiation followed by vector normalisation and wavelet-denoising to remove any issues that arose during spectral collection (Fig. 3.48). Pre-processing was followed by multivariate analysis via employment of PCA-LDA to allow segregation of classes which was visualised on scores plots of one-dimensional (1D), two-dimensional (2D) or three-dimensional (3D) space, depending on the number of classes. In each case generation of loading plots revealed five major discriminant wavenumbers responsible for variations and thus segregation of classes. Cluster vector plots were presented in a different way illustrating the expression levels of wavenumbers in the classes. Classes were represented by different shapes whose sizes correlated to expression levels i.e. the bigger the size of the shape the higher the expression.

The aim of data analysis was to identify potential biomarkers for basal, luminal and stromal cells that would enable their characterisation based on their location within the glandular elements and eventually assess their 'stemness' based on the hypothesised location of endometrial stem/progenitor cells. In this concept basal cells were expected to have more endometrial-like nature. For

wavenumbers and/or spectral regions to present a potential as biomarkers, they should be shared in common among classes being compared.

Statistical analysis was also employed to assess significance of inter-class variations. One-way ANOVA or unpaired t-test were performed, depending on number of classes being analysed at each section.

On a first approach, the biochemical composition of basal, luminal and stromal cells in tissue samples was analysed individually. Then their spectra was compared and analysed. Unfortunately the spectra acquired during FTIR spectroscopy did not have the expected shape of an FTIR spectra, but due to limited time available the obtained data was analysed.



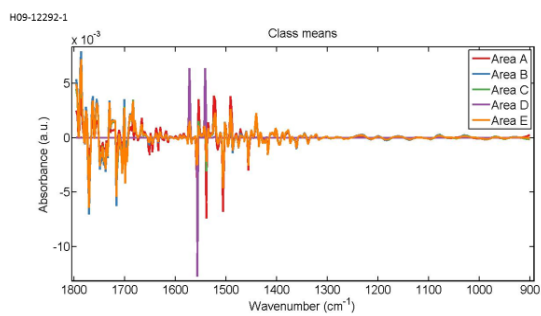
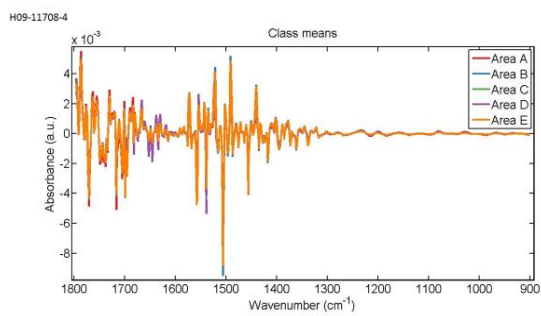
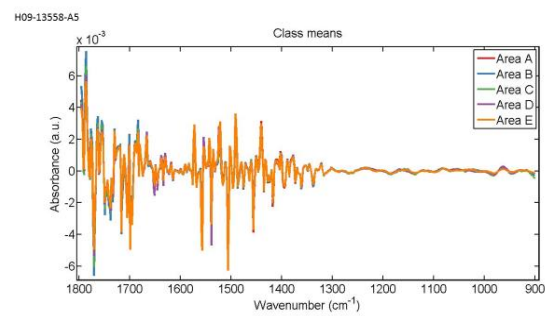
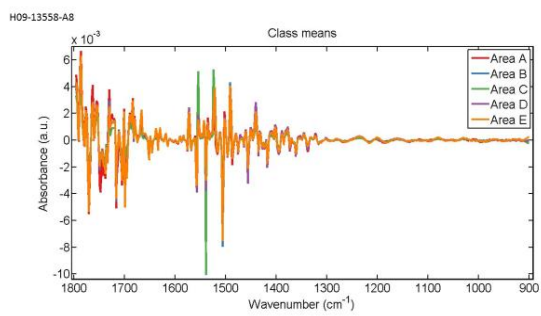
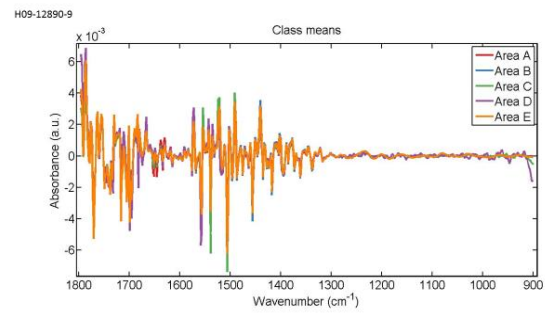
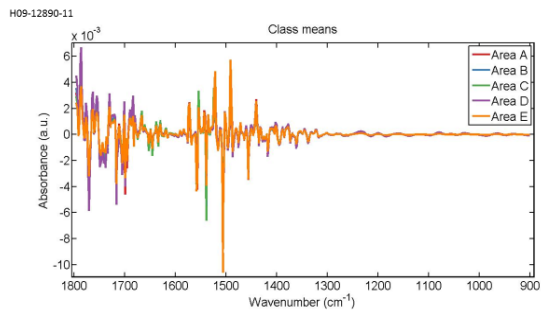
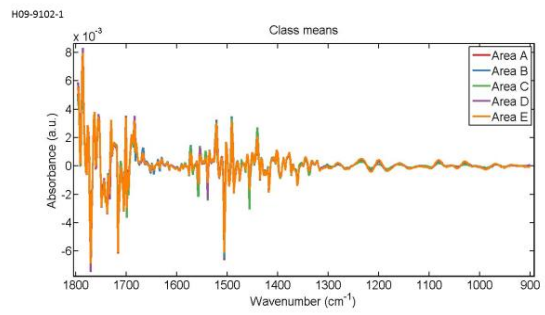
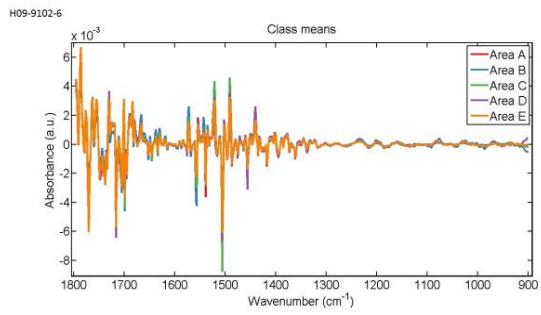
**Figure 3.48:** An example of (A) raw spectra obtained during interrogation of a tissue sample by FTIR spectroscopy and class means (B) before and (C) after pre-processing.

### 3.2.1 Basal Cells Vs Basal Cells

#### *In individual tissue samples*

Basal cells from all glandular areas in an individual tissue were compared. Basal cells from all tissues exhibited a very similar shape in mean spectra displaying variations in peak intensities (Fig. 3.49). No peaks were detected in the spectral region 900-1300  $\text{cm}^{-1}$ . Also shape of spectra was relatively consistent among tissues. No specific observations were made that would correlate corresponding diseased and non-diseased tissues.

Statistical analysis (Fig. 3.50) indicated that in most tissues at least three areas were not significantly different from the tissue as a whole. Whereas when individual areas within a tissue were compared with each other, they were significantly different in most of the cases.



**Figure 3.49:** Mean FTIR spectra for comparison of spectra acquired from basal cells between the five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.



## H09-9102-6

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.01
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-12890-11

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P < 0.01
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P > 0.05
Area B vs Area E	P < 0.01
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

## H09-13558-A8

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001
Area B vs Area C	P < 0.01
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.01
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

## H09-11708-4

Parameters	P value
LDA1 vs Area A	P < 0.05
LDA1 vs Area B	P < 0.01
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P < 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P > 0.05
Area A vs Area D	P > 0.05
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-12292-1

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-9102-1

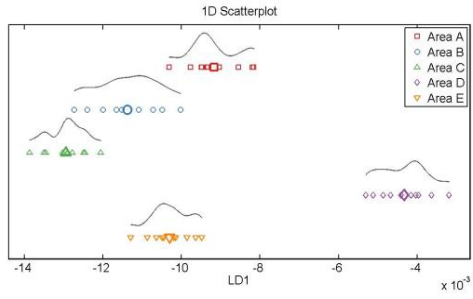
Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P < 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.01
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.001

H09-12890-9		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Area A	P > 0.05	LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05	LDA1 vs Area B	P < 0.01
LDA1 vs Area C	P > 0.05	LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001	LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05	LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05	Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05	Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001	Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05	Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05	Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.001	Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05	Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001	Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05	Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.001	Area D vs Area E	P > 0.05

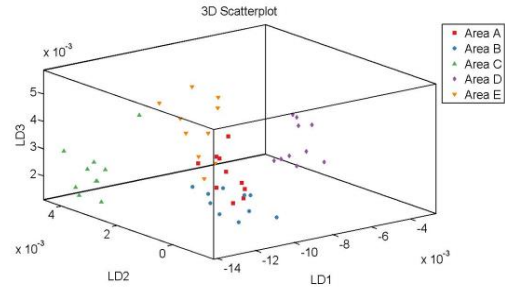
**Figure 3.50:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal in five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Segregation of spectra acquired from the glandular elements was visualised in scores plots of 1D and 3D space. Figure 3.51 illustrates scores plots of non-diseased tissue samples and figure 3.52 illustrates scores plots of diseased tissue samples. No consistency amongst samples was observed; the spectra from an individual area would overlap with the spectra of at least two other areas classes or no overlapping was observed. In the 3D scores plot spectral points of individual areas would either have a compact or spread arrangement. No observations were made that would correlate corresponding non-diseased and diseased tissue sections.

H09-9102-6

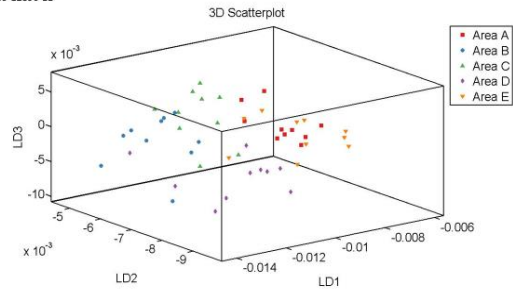
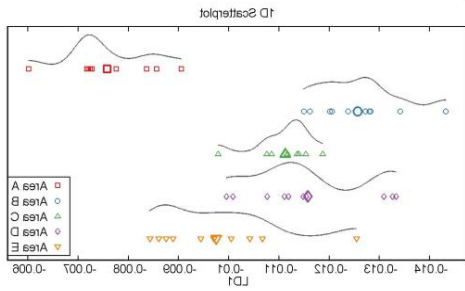


H09-9102-6

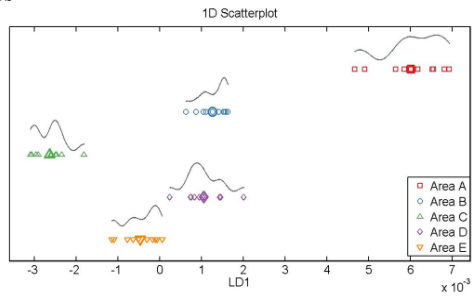


LI-02851-60H

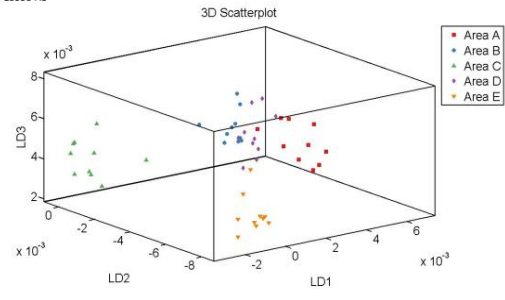
H09-12890-11



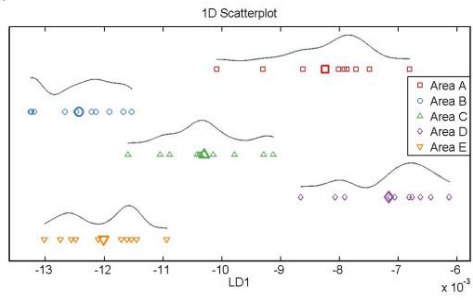
H09-13558-A8



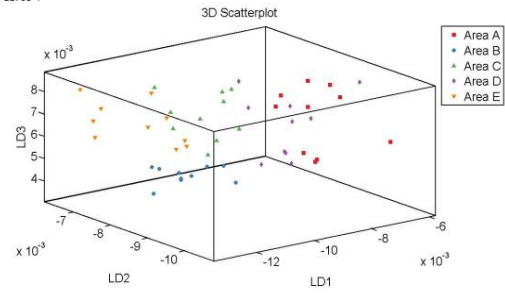
H09-13558-A8

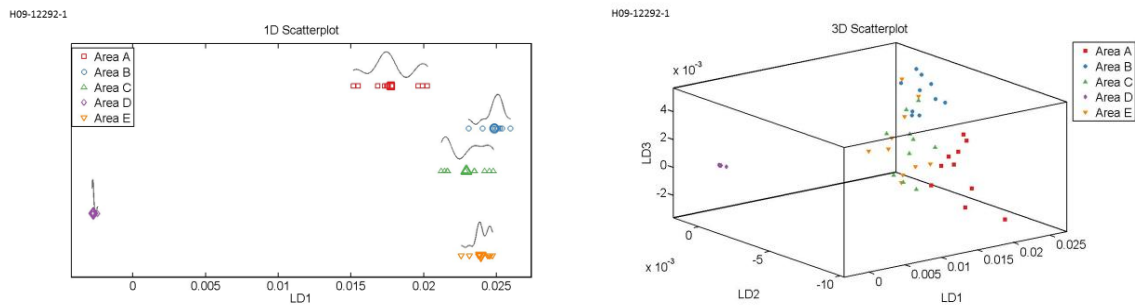


H09-11708-4

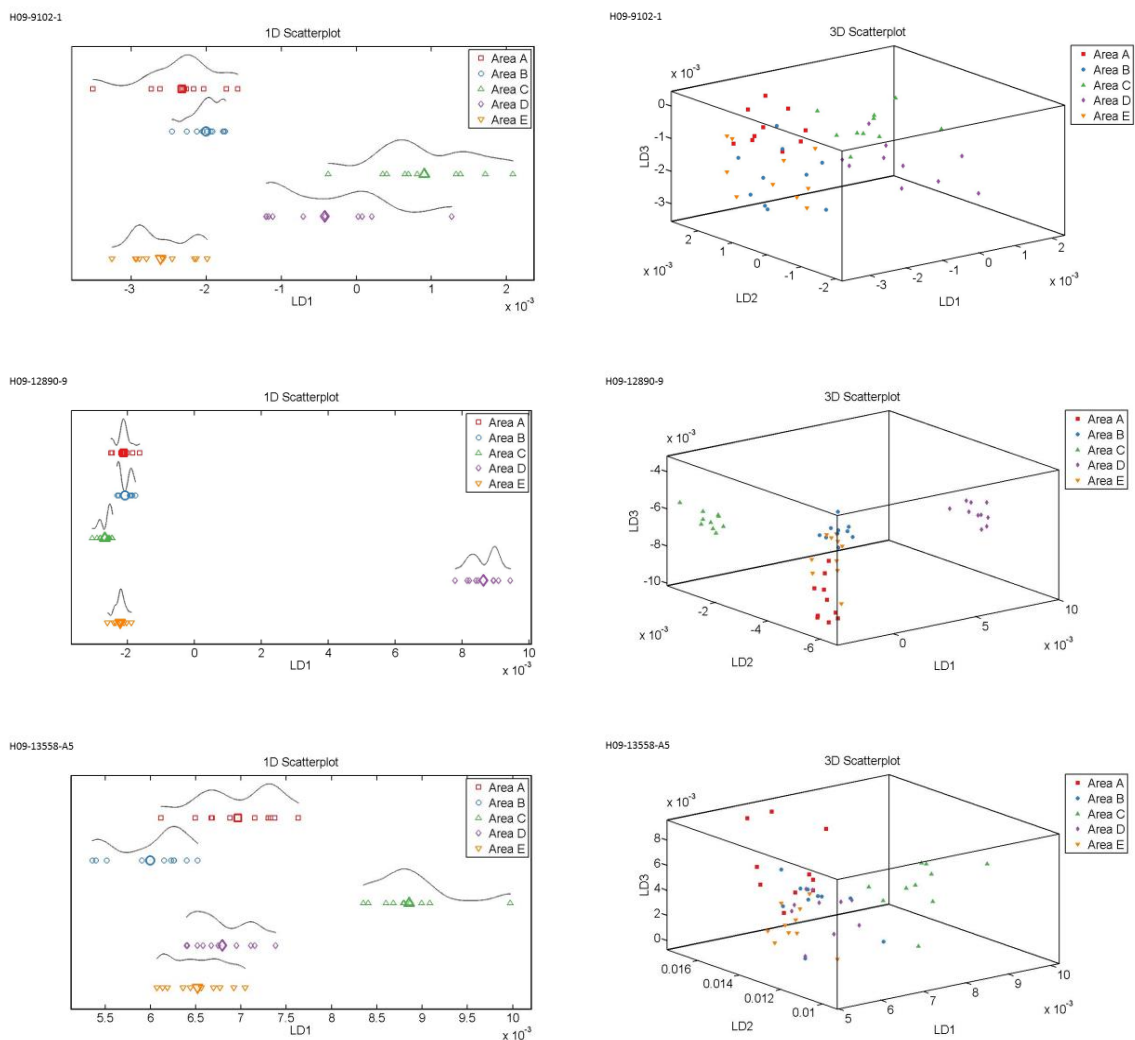


H09-11708-4



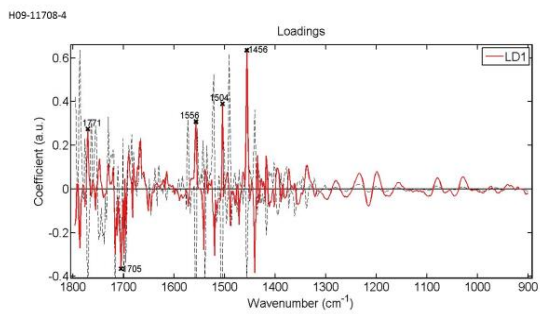
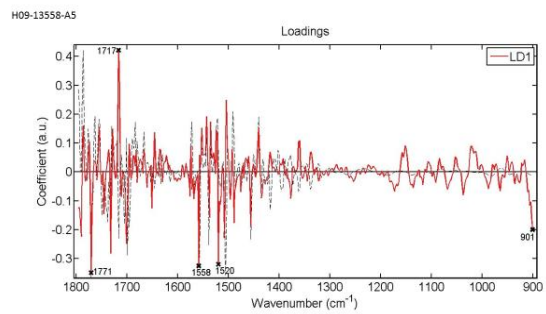
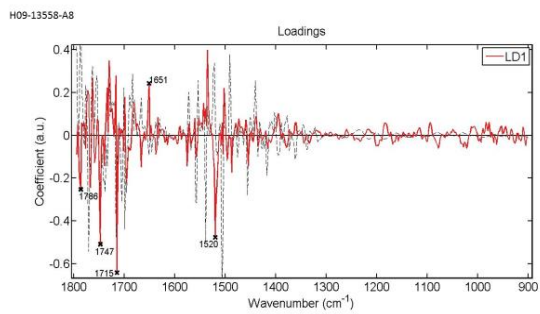
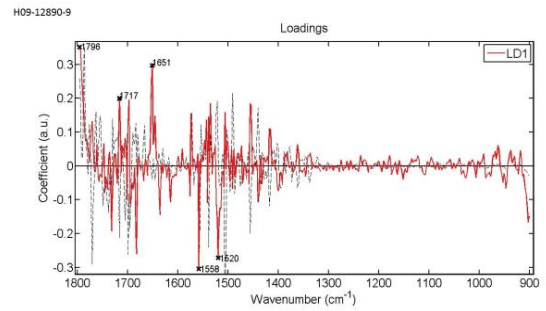
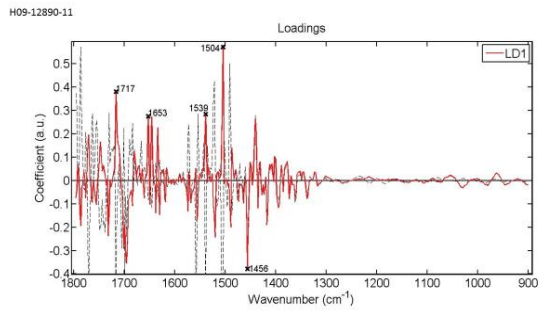
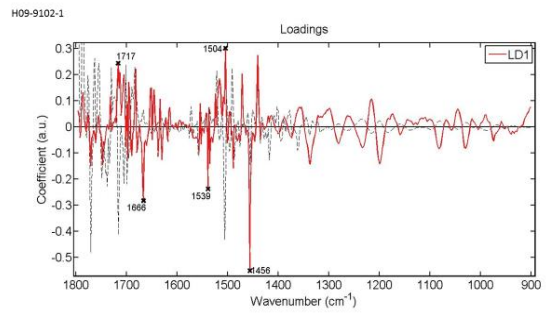
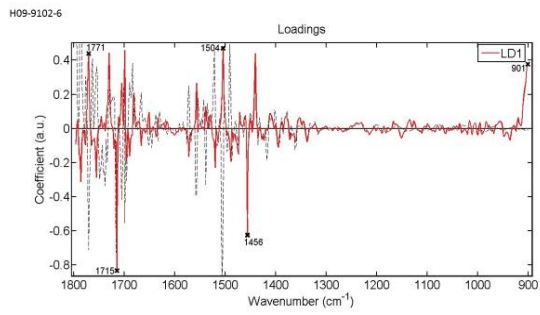


**Figure 3.51:** 1D scores plots and their corresponding 3D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from basal cells from five different glandular elements.

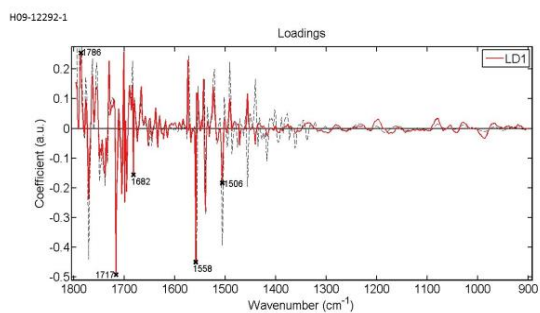


**Figure 3.52:** 1D scores plots and their corresponding 3D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from basal cells from five different glandular elements.

Figure 3.53 illustrates the loading plot for each tissue sample revealing the discriminant wavenumbers responsible for variations between the spectra acquired from basal cells at glandular areas. In sample H09-9102-6 the discriminant wavenumbers were 1771  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1504  $\text{cm}^{-1}$  (phenyl rings), 1456  $\text{cm}^{-1}$  (lipids and proteins) and 901  $\text{cm}^{-1}$ . In sample H09-12890-11 the wavenumbers were 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1653  $\text{cm}^{-1}$  (amide I), 1539  $\text{cm}^{-1}$  (amide II), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). For sample H09-13558-A8 the identified wavenumbers were 1786  $\text{cm}^{-1}$ , 1747  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$  and 1520  $\text{cm}^{-1}$  (amide II). In sample H09-11708-4 the wavenumbers were 1771  $\text{cm}^{-1}$ , 1705  $\text{cm}^{-1}$  (lipids), 1556  $\text{cm}^{-1}$ , 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). The wavenumbers identified in sample H09-12292-1 were 1786  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1682  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1506  $\text{cm}^{-1}$ . In the diseased tissue H09-9102-1 the wavenumbers were 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1666  $\text{cm}^{-1}$  (amide I), 1539  $\text{cm}^{-1}$  (amide II), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-12890-9 the observed wavenumbers were 1796  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1651  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1520  $\text{cm}^{-1}$  (amide II). In sample H09-13558-A5 the wavenumbers were 1771  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1558  $\text{cm}^{-1}$ , 1520  $\text{cm}^{-1}$  (amide II) and 901  $\text{cm}^{-1}$ .

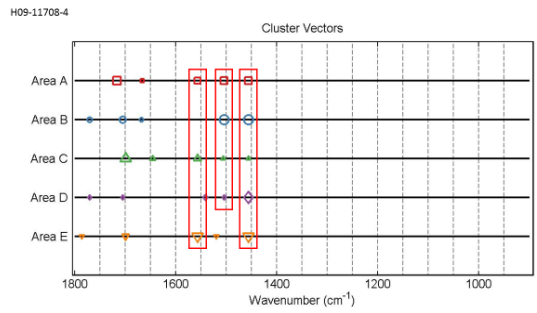
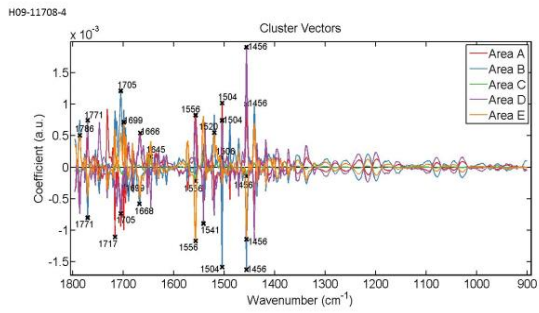
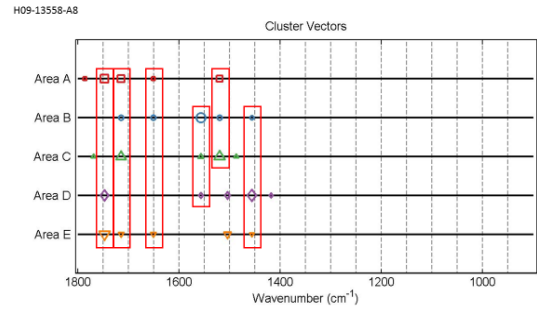
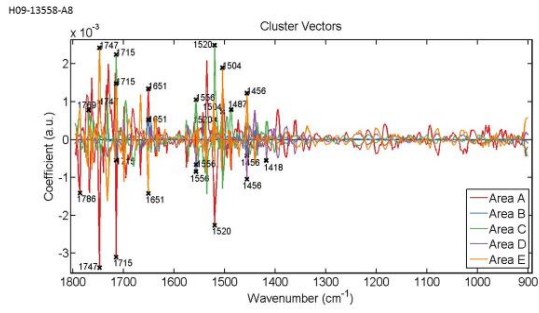
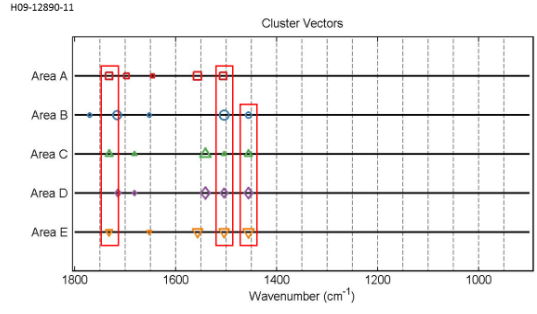
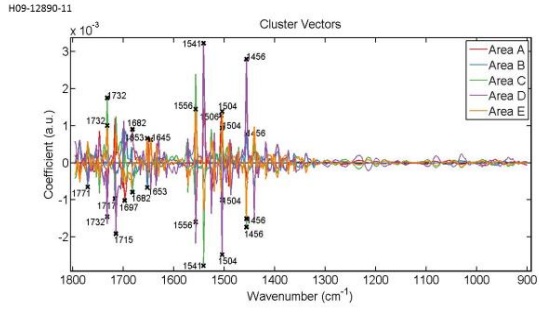
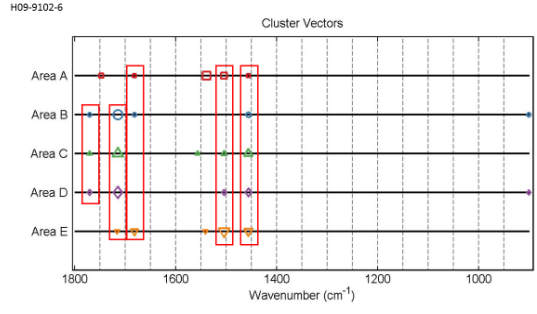
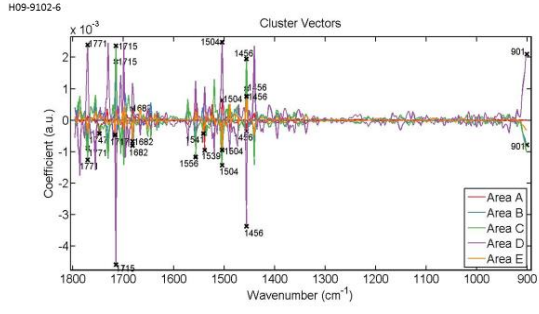


**Figure 3.53:** Loadings plots showing wavenumbers that discriminate basal cells between five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

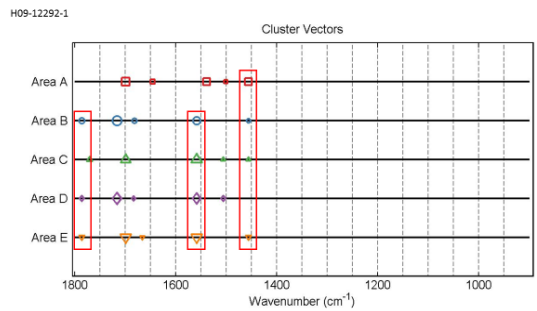
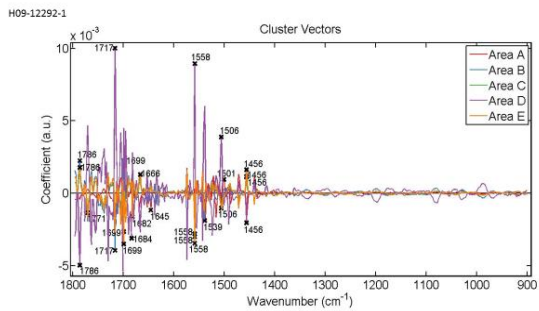


Cluster vector plots of non-diseased (Fig. 3.54) and diseased (Fig. 3.55) tissue samples were analysed in a different way in the hope of identifying possible biomarkers for basal cells, based on the occurrence of wavenumbers and their expression levels in the glandular areas. Looking at individual samples, for a wavenumber and/or spectral region to display a potential as biomarker it should be commonly shared between three or more glandular areas and also have a smooth vertical alignment on the plots.

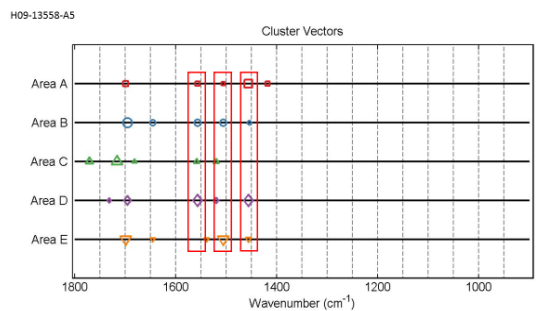
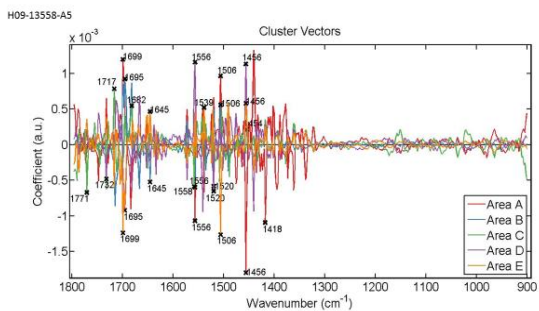
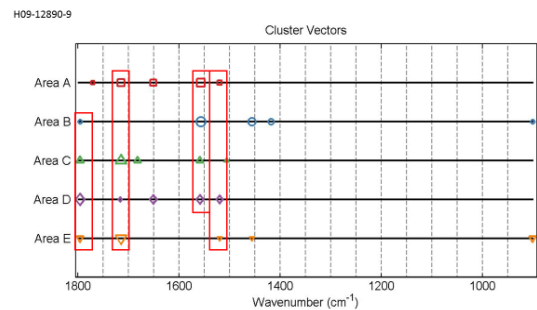
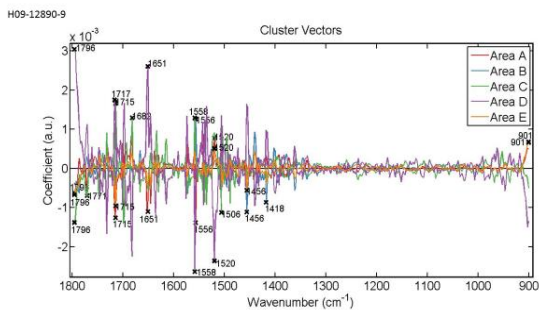
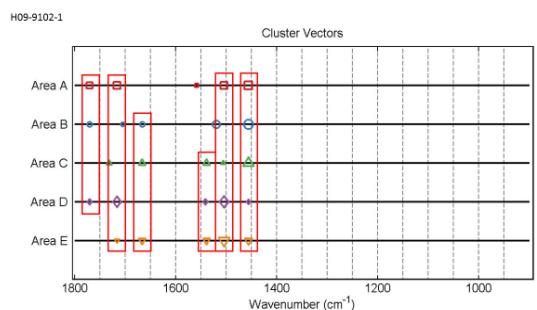
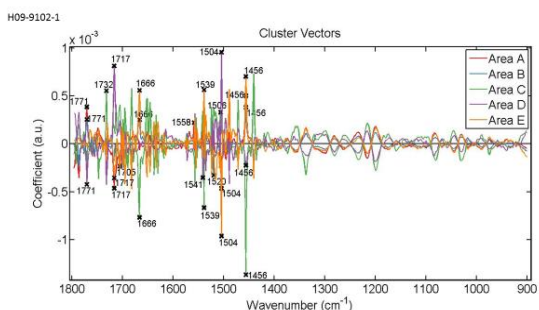
In most plots, the wavenumbers displaying common occurrence were the same as the discriminant wavenumbers observed in loading plots. In sample H09-9102-6 the wavenumbers for which three areas were common were  $1771\text{ cm}^{-1}$  and  $1682\text{ cm}^{-1}$  whilst four areas were common for  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base) with  $1715\text{ cm}^{-1}$ ,  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1456\text{ cm}^{-1}$  (lipids and proteins). Common occurrence of  $1732\text{ cm}^{-1}$  (lipids),  $1506\text{ cm}^{-1}$  with  $1504\text{ cm}^{-1}$  (phenyl rings), and  $1456\text{ cm}^{-1}$  (lipids and proteins) was observed in three, five and four areas respectively in sample H09-12890-11. In sample H09-13558-A8 three glandular areas were common for  $1747\text{ cm}^{-1}$ ,  $1651\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$ ,  $1520\text{ cm}^{-1}$  (amide II) and  $1456\text{ cm}^{-1}$  (lipids and proteins) and four areas were common for  $1715\text{ cm}^{-1}$ . In sample H09-11708-4 common occurrence was displayed by three, four and five areas for  $1556\text{ cm}^{-1}$ ,  $1506\text{ cm}^{-1}$  with  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1456\text{ cm}^{-1}$  (lipids and proteins) respectively. The wavenumber  $1786\text{ cm}^{-1}$  was observed in three areas of sample H09-12292-1 and four areas were common for  $1558\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins). Three areas in sample H09-9102-1 were common for  $1771\text{ cm}^{-1}$ ,  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base),  $1663\text{ cm}^{-1}$  and  $1541\text{ cm}^{-1}$  (amide II) with  $1539\text{ cm}^{-1}$  (amide II), four areas were common for  $1506\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings) whilst all five areas were common for  $1456\text{ cm}^{-1}$  (lipids and proteins). In sample H09-12890-9 the wavenumbers  $1796\text{ cm}^{-1}$ ,  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base) with  $1715\text{ cm}^{-1}$ , and  $1558\text{ cm}^{-1}$  with  $1556\text{ cm}^{-1}$  were commonly shared among four areas whilst three areas were common for  $1520\text{ cm}^{-1}$  (amide II). In sample H09-13558-A5 four areas were common for  $1558\text{ cm}^{-1}$  with  $1556\text{ cm}^{-1}$ , and  $1456\text{ cm}^{-1}$  (lipids and proteins) with  $1454\text{ cm}^{-1}$  (asymmetric methyl deformation) whilst  $1506\text{ cm}^{-1}$  was commonly shared between three areas.







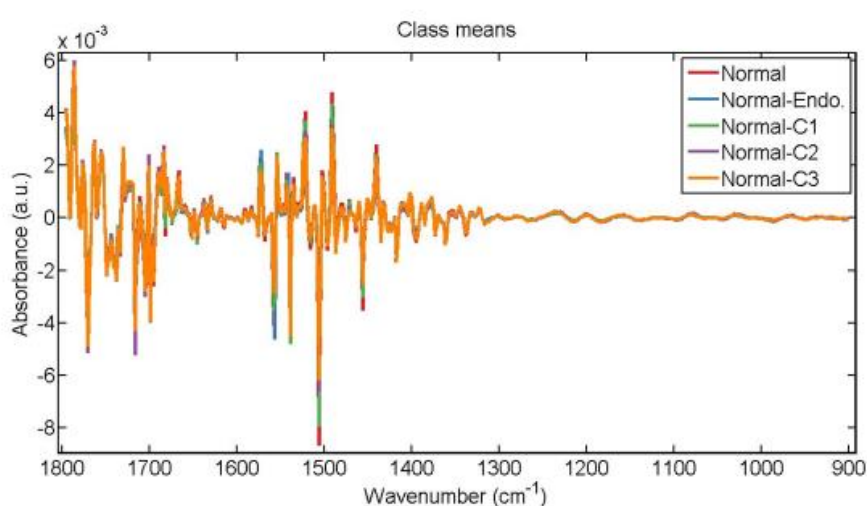
**Figure 3.54:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells, in five glandular areas in individual non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1). The red rectangles highlight the wavenumbers shared in common among glandular areas.



**Figure 3.55:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells, in five glandular areas in individual diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight the wavenumbers shared in common among glandular areas.

*Between all normal tissue samples*

All the non-diseased tissue samples were compared between them for the spectra interrogated from their basal cells. Classes of tissue samples were labelled according to the information given about the samples (see Table 2.1); Normal (H09-11708-4), Normal-Endo. (H09-12292-1), Normal-C1 (H09-12890-11), Normal-C2 (H09-9102-6) and Normal-C3 (H09-13558-A8). Sample H09-11708-4 was diagnosed as normal whilst the other samples were taken from non-diseased sites of the endometrium from patients diagnosed with endometriosis (H09-12292-1) and endometrial cancer (H09-12890-11, H09-9102-6, H09-13558-A8). In this section tissues will be referred according to their class labels to associate plots with observations. All classes representing individual tissues, exhibited a very similar shape in mean spectra with most evident variations in absorbance intensities occurring at  $\sim 1572\text{ cm}^{-1}$ ,  $\sim 1556\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$ ,  $\sim 1491\text{ cm}^{-1}$  and  $\sim 1456\text{ cm}^{-1}$  (Fig. 3.56). No peaks were detected in the spectral region  $900\text{--}1300\text{ cm}^{-1}$ . Statistical analysis (Fig. 3.57) indicated that classes were significantly different for the spectra acquired from basal cells but Normal-C1 was not significantly different from Normal and Normal-C2. The latter was also not significantly different from Normal-C3.

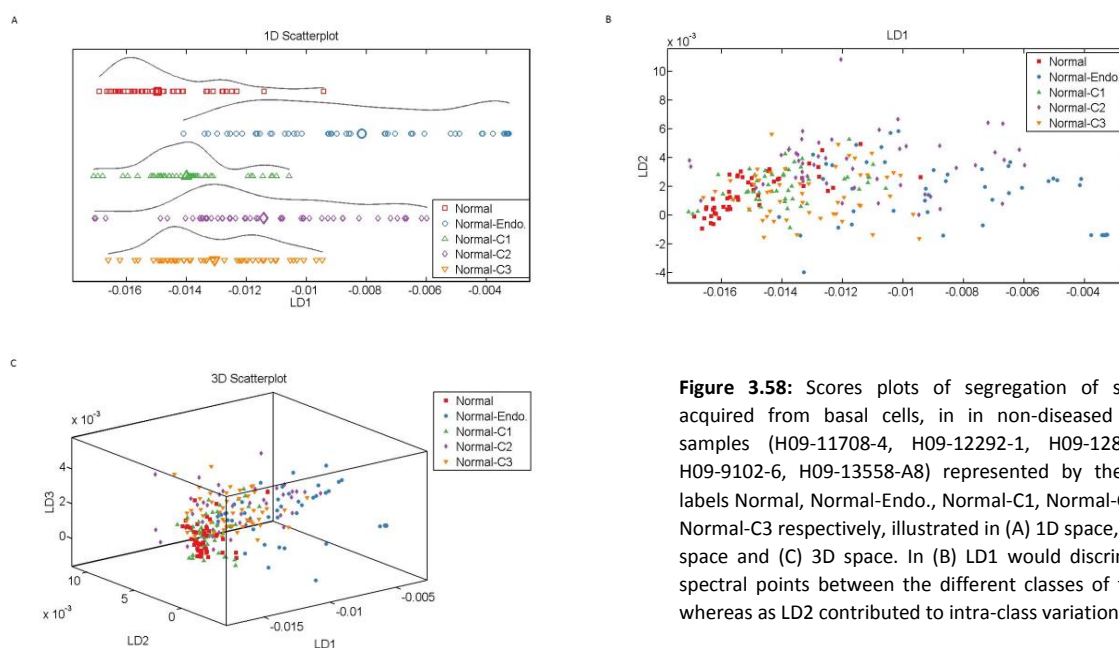


**Figure 3.56:** Mean FTIR spectra for comparison of spectra acquired from basal cells in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Parameters	P value
LDA1 vs Normal	P < 0.001
LDA1 vs Normal-Endo.	P < 0.001
LDA1 vs Normal-C1	P < 0.01
LDA1 vs Normal-C2	P > 0.05
LDA1 vs Normal-C3	P > 0.05
Normal vs Normal-Endo.	P < 0.001
Normal vs Normal-C1	P > 0.05
Normal vs Normal-C2	P < 0.001
Normal vs Normal-C3	P < 0.05
Normal-Endo. vs Normal-C1	P < 0.001
Normal-Endo. vs Normal-C2	P < 0.001
Normal-Endo. vs Normal-C3	P < 0.001
Normal-C1 vs Normal-C2	P < 0.001
Normal-C1 vs Normal-C3	P > 0.05
Normal-C2 vs Normal-C3	P > 0.05

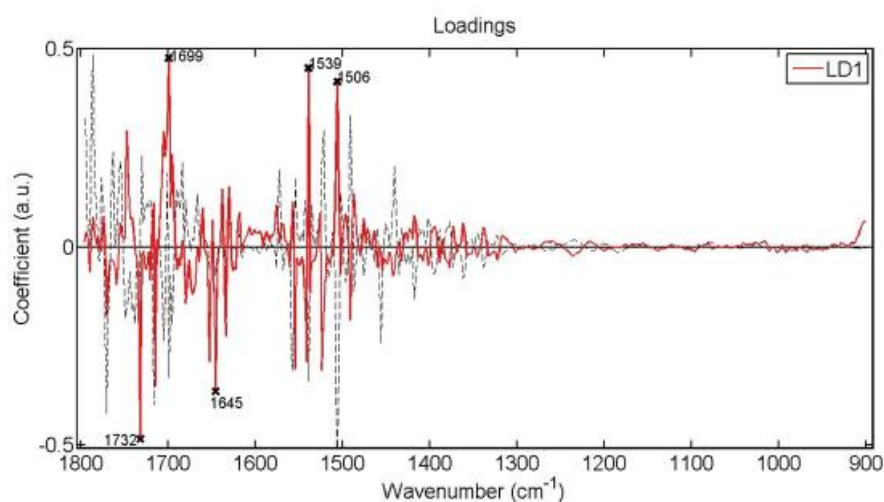
**Figure 3.57:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal cells in all five glandular elements between non-diseased tissue sections H09-11708-4, H09-9102-1, H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D, 2D and 3D space (Fig. 3.58). All classes would overlap with each other either throughout their whole spectra or spectral regions. The spectral points of class Normal had a relatively more compact arrangement which signified the least intra-class variation. Whereas spectral points of classes Normal-Endo. and Normal-C2 displayed a relatively spread arrangement and thus more intra-class variation.



**Figure 3.58:** Scores plots of segregation of spectra acquired from basal cells, in in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively, illustrated in (A) 1D space, (B) 2D space and (C) 3D space. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were  $1732\text{ cm}^{-1}$  (lipids),  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1645\text{ cm}^{-1}$ ,  $1539\text{ cm}^{-1}$  (amide II) and  $1506\text{ cm}^{-1}$  (Fig. 3.59).

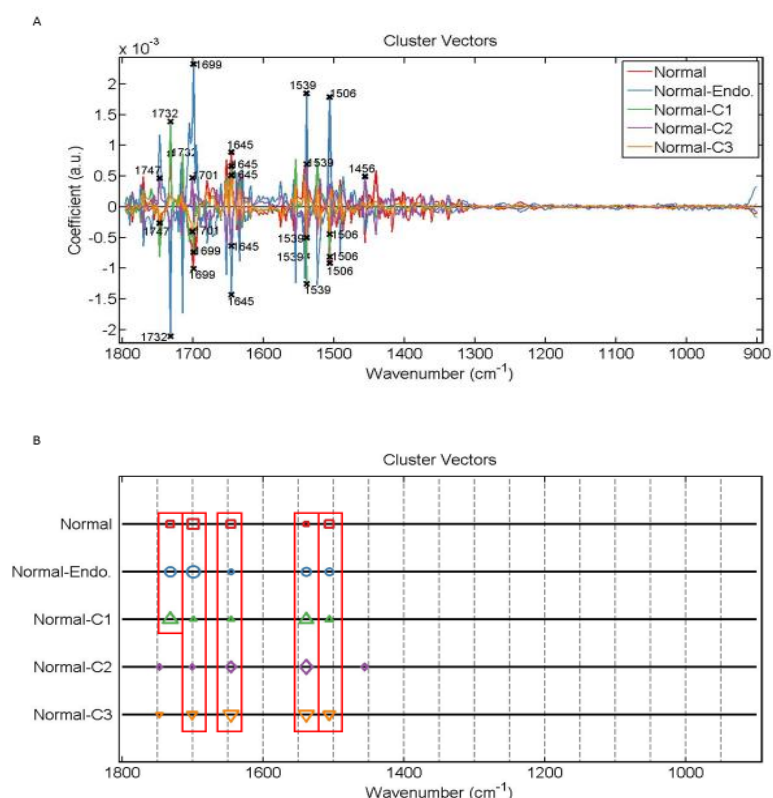


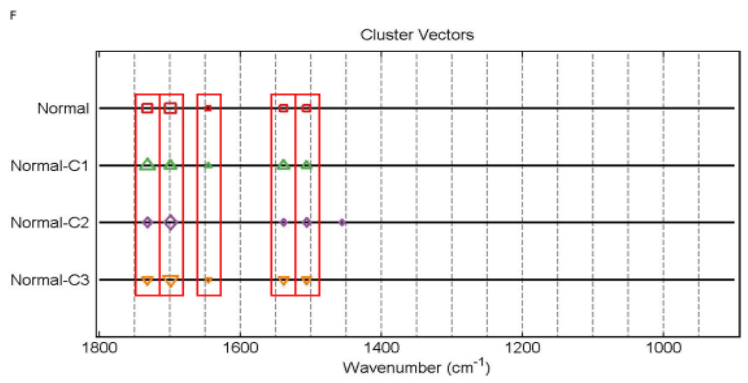
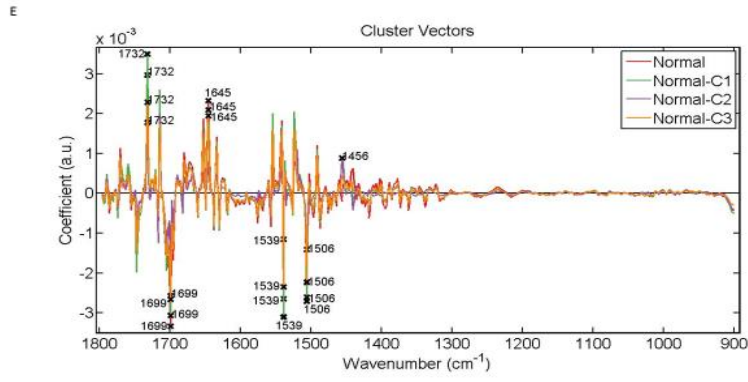
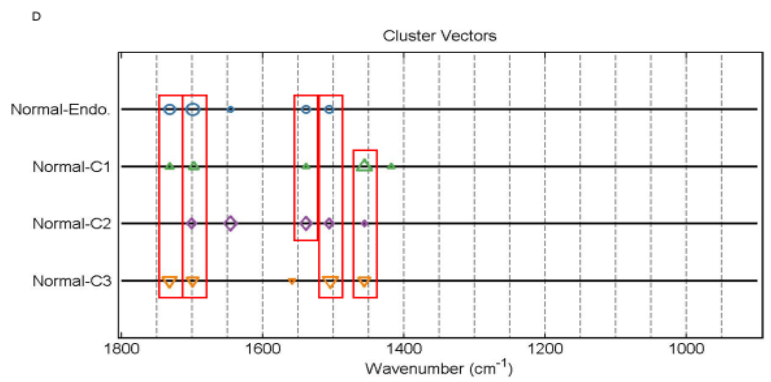
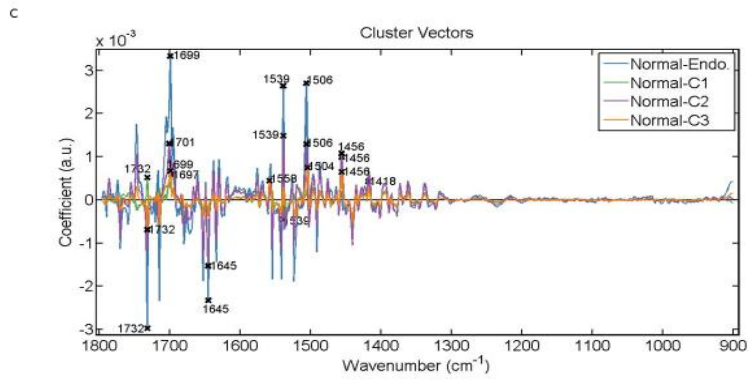
**Figure 3.59:** Loadings plots showing wavenumbers that discriminate basal cells in non-diseased tissue sections (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8). The red line is a pseudospectrum and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots (Fig. 3.60) were presented and analysed in a different way in order to identify potential biomarkers for basal cells by observing the occurrence and expression levels of wavenumbers and/or spectral regions between the tissues by having a class as a reference origin or when no class was used as a reference origin. Wavenumbers would display common occurrence if they were observed in at least three classes representing the tissues as long as they had a smooth vertical alignment in the plots.

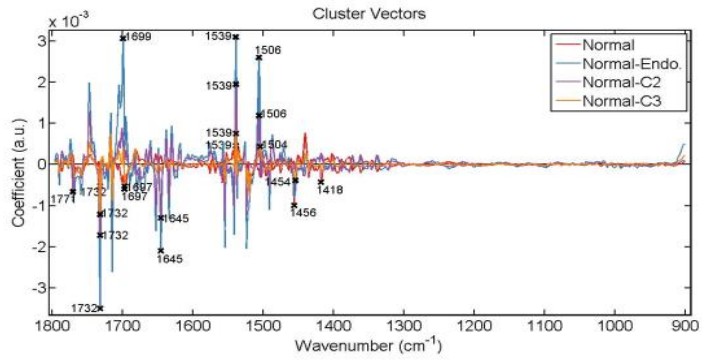
Having no class as a reference, three tissues were common for  $1732\text{ cm}^{-1}$  (lipids) and all five tissues were common for  $1701\text{ cm}^{-1}$ ,  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1645\text{ cm}^{-1}$ ,  $1539\text{ cm}^{-1}$  (amide II) whilst four tissues were common only for  $1506\text{ cm}^{-1}$ . Using class Normal as a reference, common occurrence of  $1732\text{ cm}^{-1}$  (lipids),  $1539\text{ cm}^{-1}$  (amide II),  $1506\text{ cm}^{-1}$  with  $1504\text{ cm}^{-1}$  (phenyl rings), and  $1456\text{ cm}^{-1}$  (lipids and proteins) was observed amongst three tissues whilst four tissues were common for  $1701\text{ cm}^{-1}$ ,

1699  $\text{cm}^{-1}$  (guanine/thymine) and 1697  $\text{cm}^{-1}$ . Having class Normal-Endo. as a reference, four tissues were common for 1732  $\text{cm}^{-1}$  (lipids), 1699  $\text{cm}^{-1}$  (guanine/thymine), 1539  $\text{cm}^{-1}$  (amide II) and 1506  $\text{cm}^{-1}$  whilst 1645  $\text{cm}^{-1}$  was commonly shared between three tissues. Having class Normal-C1 as reference the wavenumbers 1732  $\text{cm}^{-1}$  (lipids) and 1539  $\text{cm}^{-1}$  (amide II) were commonly shared by four tissues and 1699  $\text{cm}^{-1}$  (guanine/thymine), 1697  $\text{cm}^{-1}$ , 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings) were commonly shared by three tissues. Having class Normal-C2 as reference, common occurrence of 1701  $\text{cm}^{-1}$ , 1699  $\text{cm}^{-1}$  (guanine/thymine) and 1645  $\text{cm}^{-1}$  was observed in three tissues and four tissues were found common for 1539  $\text{cm}^{-1}$  (amide II) and 1506  $\text{cm}^{-1}$ . Having class Normal-C3 as a reference, three tissues were common for 1732  $\text{cm}^{-1}$  (lipids) and 1539  $\text{cm}^{-1}$  (amide II) whilst four tissues were common for 1701  $\text{cm}^{-1}$ , 1699  $\text{cm}^{-1}$  (guanine/thymine), 1697  $\text{cm}^{-1}$  and 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings).

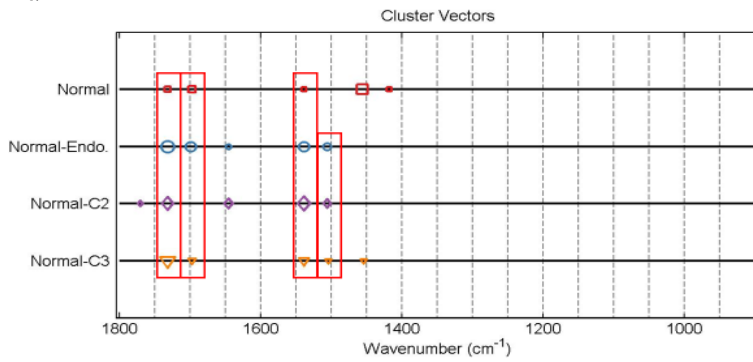




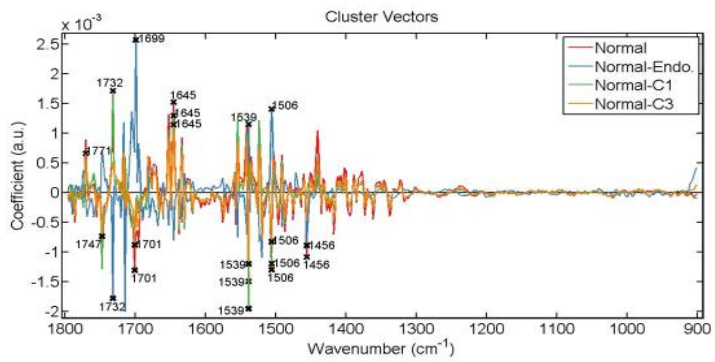
G



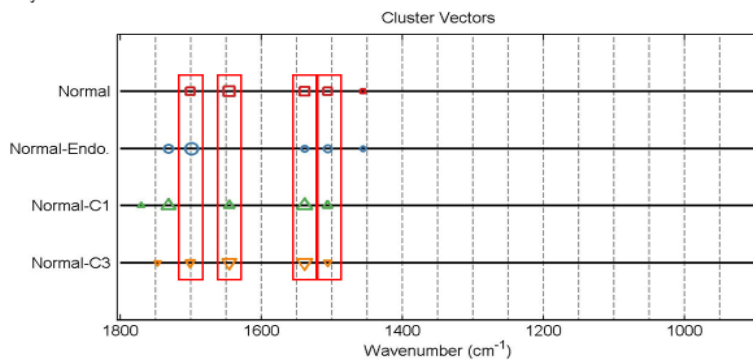
H

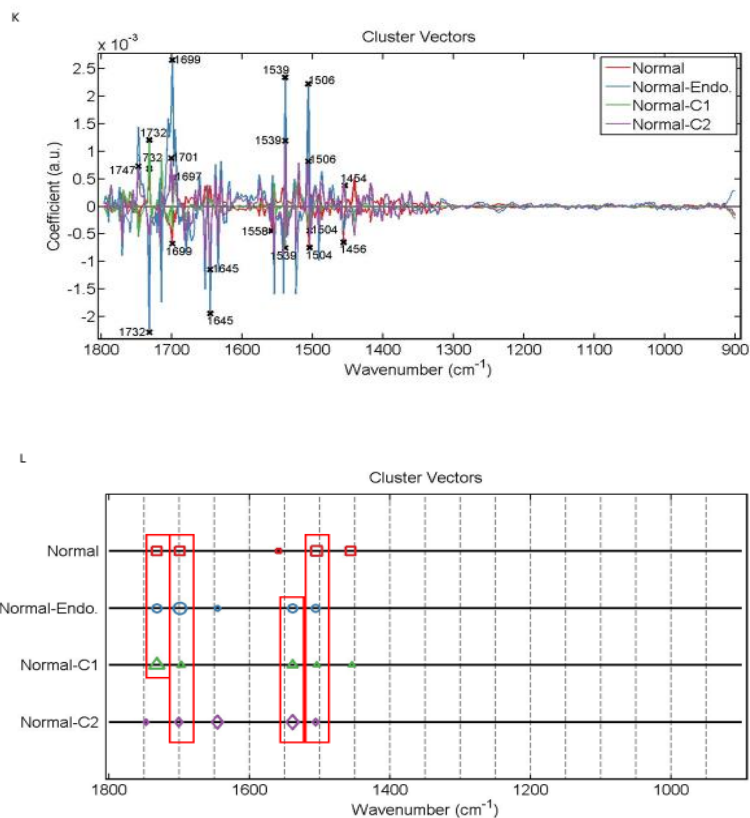


I



J





**Figure 3.60:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells, amongst non-diseased tissue sections. (A) & (B) no sample used as a reference, (C) & (D) Normal (H09-11708-4) reference class, (E) & (F) Normal-Endo. (H09-12292-1) reference class, (G) & (H) Normal-C1 (H09-12890-11) reference class, (I) & (J) Normal-C2 (H09-9102-6) reference class and (K) & (L) Normal-C3 (H09-13558-A8) reference class. The red rectangles in the cluster vector peak location plots highlight the wavenumbers occurring commonly in different tissue samples.

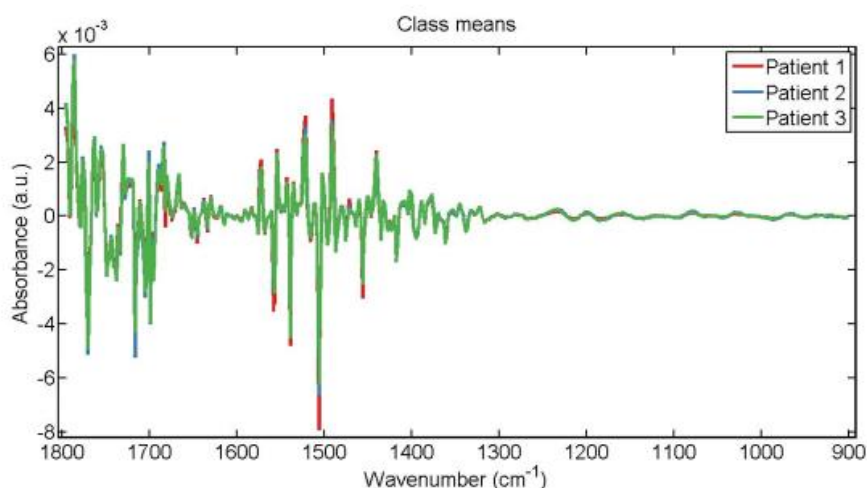
### *Between non-diseased samples from patients with endometrial cancer*

The biochemical composition of basal cells in only non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer was analysed and compared. Classes were labelled as Patient 1, Patient 2 and Patient 3 to represent the samples H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In this section classes of tissue samples will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots.



All classes had a similar shape of mean FTIR spectra with most evident variations in absorbance intensities occurring at  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1558\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$ ,  $\sim 1491\text{ cm}^{-1}$  and  $\sim 1456\text{ cm}^{-1}$  (Fig. 3.61). No peaks were detected in the spectral region  $900\text{--}1300\text{ cm}^{-1}$ .

Statistical analysis indicated that all tissues were significantly different from each other (Fig. 3.62).

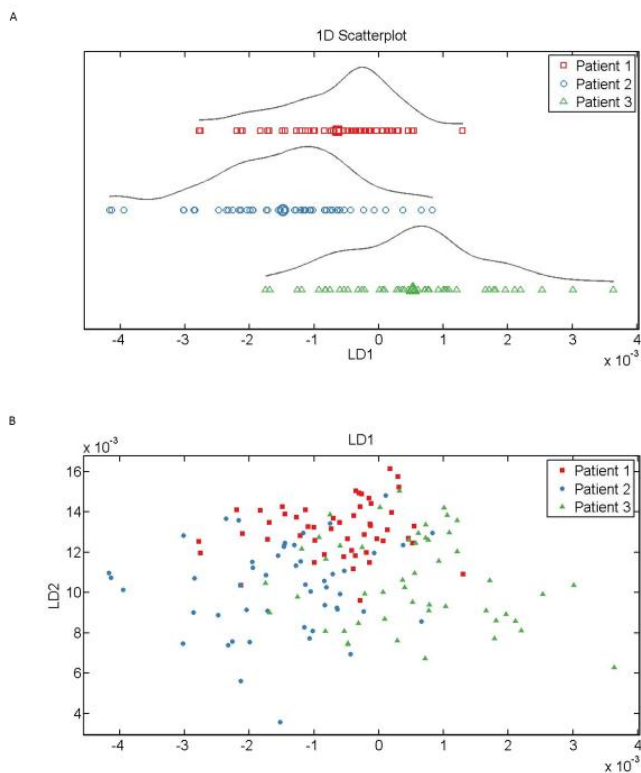


**Figure 3.61:** Mean FTIR spectra for comparison of spectra acquired from basal cells in non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively.

Parameters	P value
LDA1 vs Patient 1	$P > 0.05$
LDA1 vs Patient 2	$P < 0.001$
LDA1 vs Patient 3	$P < 0.001$
Patient 1 vs Patient 2	$P < 0.01$
Patient 1 vs Patient 3	$P < 0.001$
Patient 2 vs Patient 3	$P < 0.001$

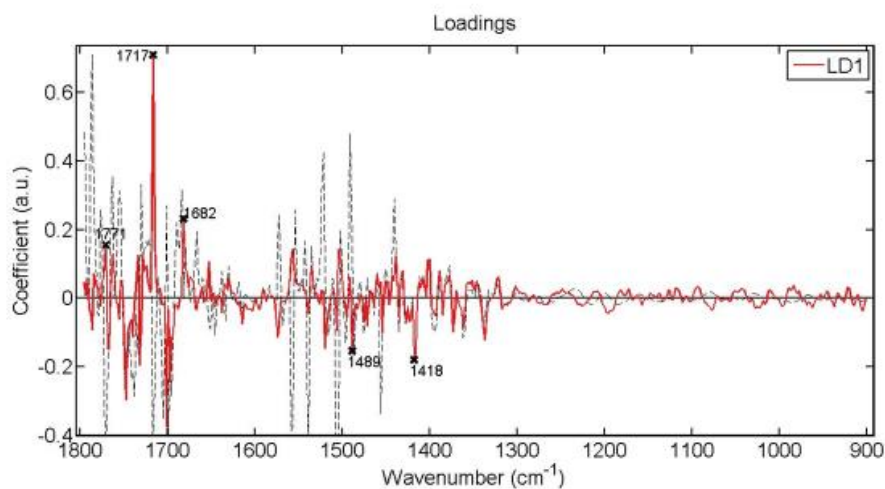
**Figure 3.62:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal cells between non-diseased tissue sections H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D and 2D space (Fig. 3.63). Patient 1 and Patient 2 overlapped with each other throughout most of their spectra. Both classes overlapped with almost half of the spectra of patient 3. Spectral points of Patient 1 appeared to have a more compact arrangement relative to the other class which signified that Patient 1 had the least intra-class variation.



**Figure 3.63:** Scores plots for segregation of spectra in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

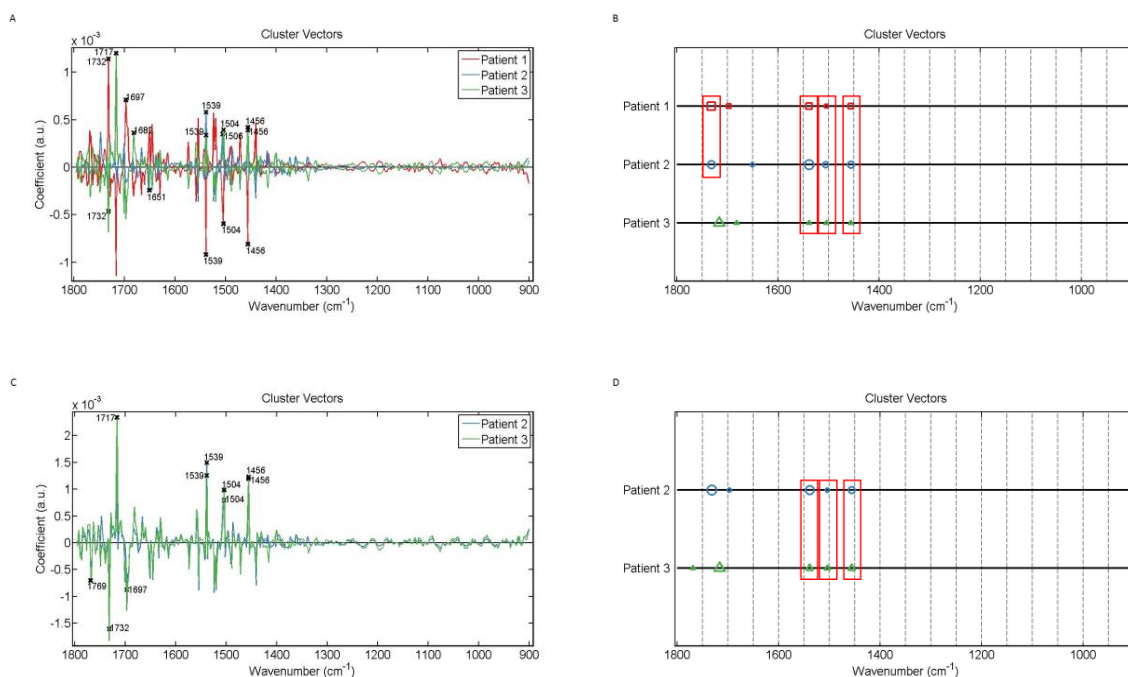
The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were  $1771\text{ cm}^{-1}$ ,  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base),  $1682\text{ cm}^{-1}$ ,  $1489\text{ cm}^{-1}$  (CH bending/deformation C-H) and  $1418\text{ cm}^{-1}$  (deformation C-H) (Fig. 3.64).

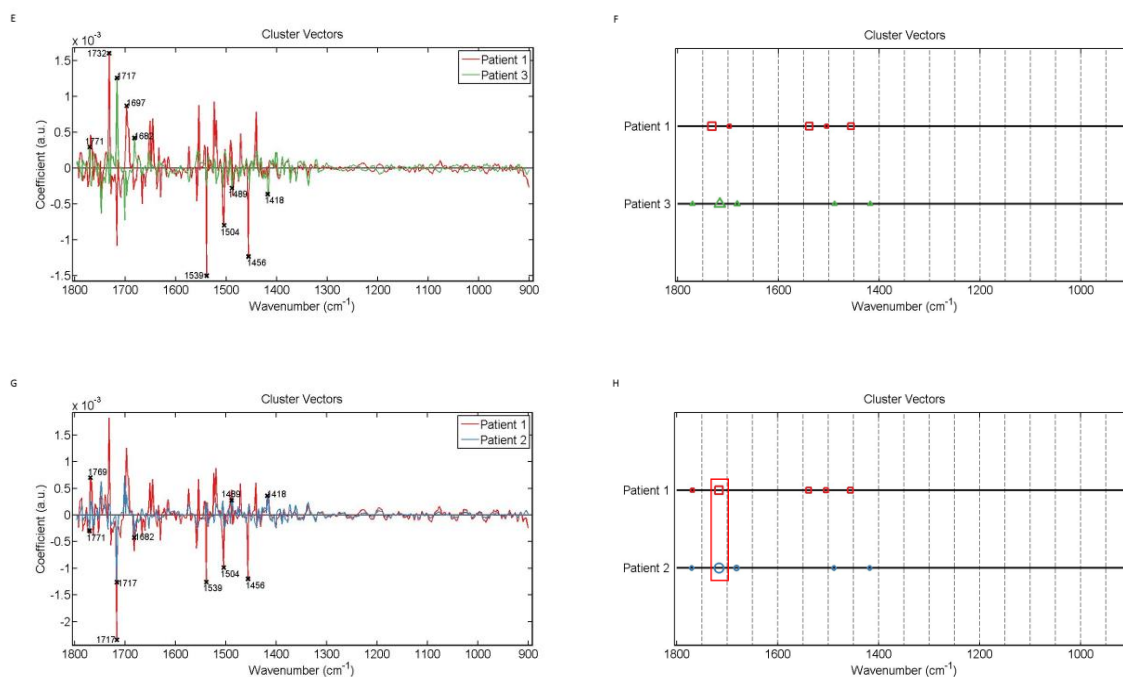


**Figure 3.64:** Loadings plots showing wavenumbers that discriminate basal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference

Cluster vector plots (Fig. 3.65) were presented and analysed in a different way in order to identify potential biomarkers for basal cells by observing the occurrence and expression levels of wavenumbers and/or spectral regions between the tissues by having a class as a reference origin or when no class was used as a reference origin. Wavenumbers would display common occurrence if they were observed in three classes representing the tissues as long as they had a smooth vertical alignment in the plots. When a class was set as the reference origin, a wavenumber would display common occurrence if it was observed in the cluster vectors of the two other classes.

Having no class as a reference origin the wavenumbers displaying common occurrence were  $1732\text{ cm}^{-1}$  (lipids),  $1539\text{ cm}^{-1}$  (amide II),  $1506\text{ cm}^{-1}$  with  $1504\text{ cm}^{-1}$  (phenyl rings), and  $1456\text{ cm}^{-1}$  (lipids and proteins). Using Patient 1 as reference the commonly shared wavenumbers were  $1539\text{ cm}^{-1}$  (amide II),  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1456\text{ cm}^{-1}$  (lipids and proteins). Using Patient 3 as reference common occurrence was observed only for  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base). No wavenumber was observed to be common among classes when Patient 2 was used as a reference class.



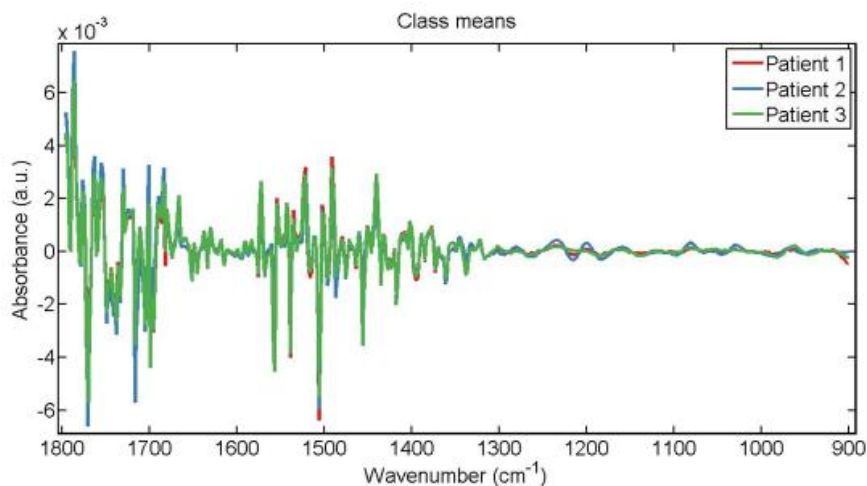


**Figure 3.65:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells among non-diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-11) reference class, (E) & (F) Patient 2 (H09-9102-6) reference class and (G) & (H) Patient 3 (H09-13558-A8) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Between diseased samples from patients with endometrial cancer*

The biochemical composition of basal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) taken from cancerous lesions in the endometrium was analysed and compared. Classes were labelled as Patient 1, Patient 2 and Patient 3 to represent the samples H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In this section classes of tissues will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots. All classes had a very similar shape in mean FTIR spectra with most pronounced variations in peak intensities occurring at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1701\text{ cm}^{-1}$  and  $\sim 1506\text{ cm}^{-1}$  (Fig. 3.66). No peaks were detected in the spectral region  $900\text{--}1300\text{ cm}^{-1}$ .

Statistical analysis indicated that tissues were significantly different from each other (Fig. 3.67).

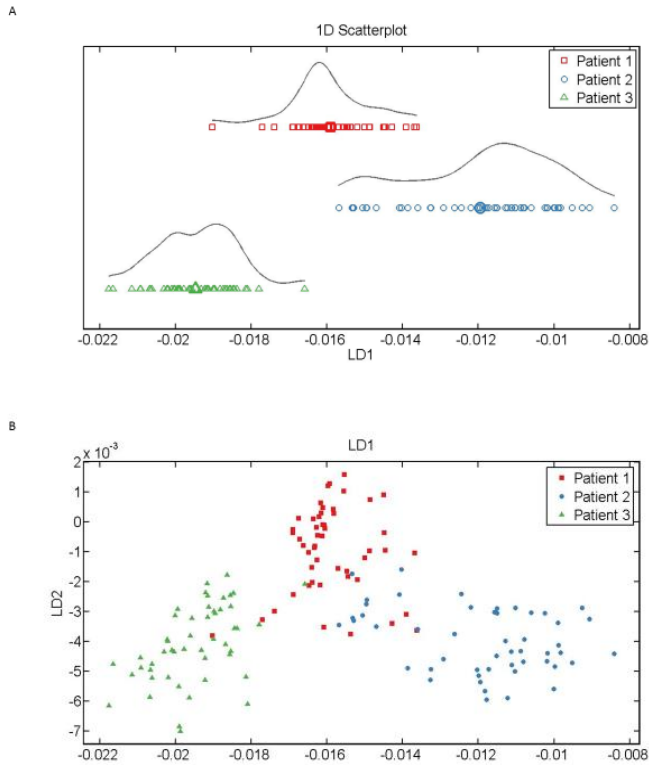


**Figure 3.66:** Mean FTIR spectra for comparison of spectra from basal cells in diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively.

Parameters	P value
LDA1 vs Patient 1	P > 0.05
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P < 0.001
Patient 2 vs Patient 3	P < 0.001

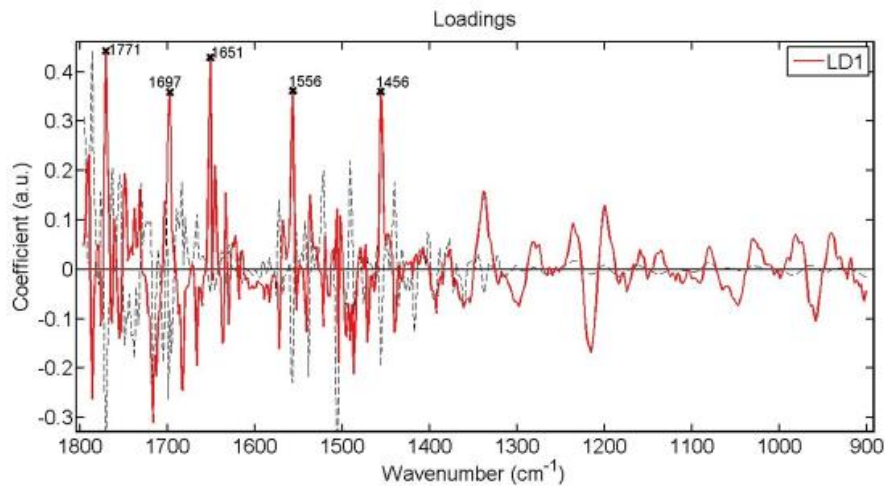
**Figure 3.67:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal cells between diseased tissue sections H09-9102-1, H09-12890-9, H09-13558-A5 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Segregation of classes resulting from application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.68). No overlapping was observed between Patient 2 and Patient 3 but Patient 1 overlapped with both classes. Spectral points of Patient 2 had a more spread arrangement which signified more intra-class variation relative to the other classes.



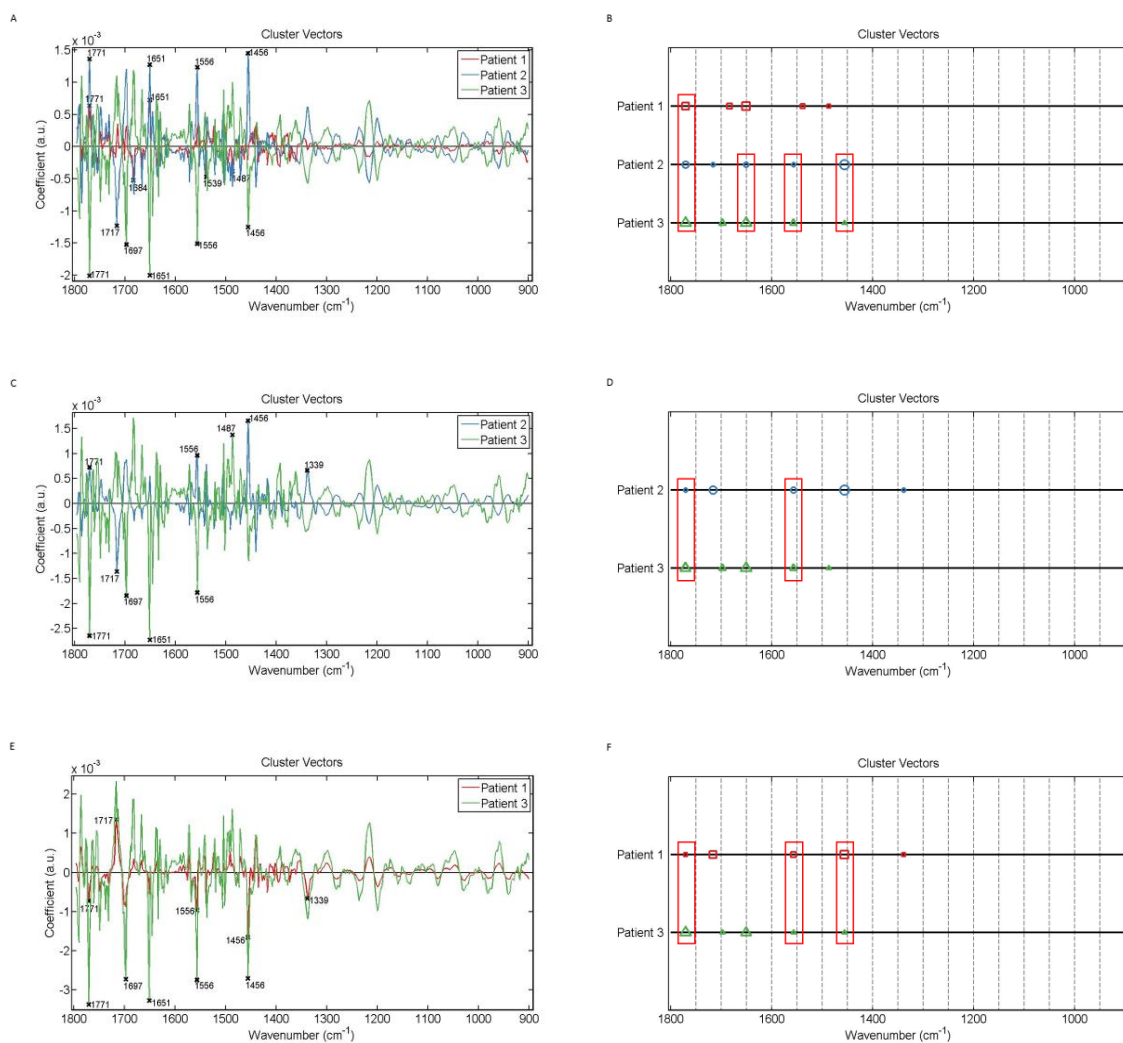
**Figure 3.68:** Scores plots for segregation of spectra acquired from basal cells in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

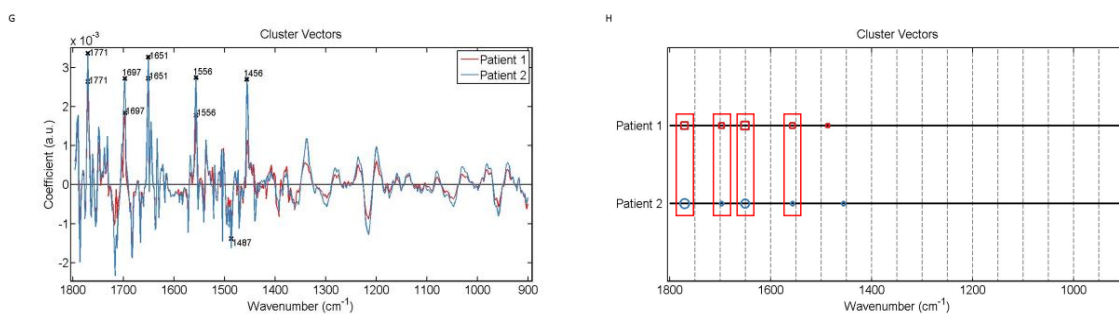
Generation of loading plots revealed five major discriminant wavenumbers accounting for variations between the classes (Fig. 3.69). The identified wavenumbers were  $1771\text{ cm}^{-1}$ ,  $1697\text{ cm}^{-1}$ ,  $1651\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins).



**Figure 3.69:** Loadings plots showing wavenumbers that discriminate basal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

The same approach as before was applied to analyse cluster vector plots for identification of biomarkers (Fig. 3.70). Having no class as a reference origin, all classes were common for 1771  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$ , 1556  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). The wavenumbers observed to have common occurrence when Patient 1 was the reference class were 1771  $\text{cm}^{-1}$  and 1556  $\text{cm}^{-1}$ . These wavenumbers along with 1456  $\text{cm}^{-1}$  displayed common occurrence when Patient 2 was the class reference. Having Patient 3 as reference, common occurrence was displayed by 1771  $\text{cm}^{-1}$ , 1697  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$  and 1556  $\text{cm}^{-1}$ .



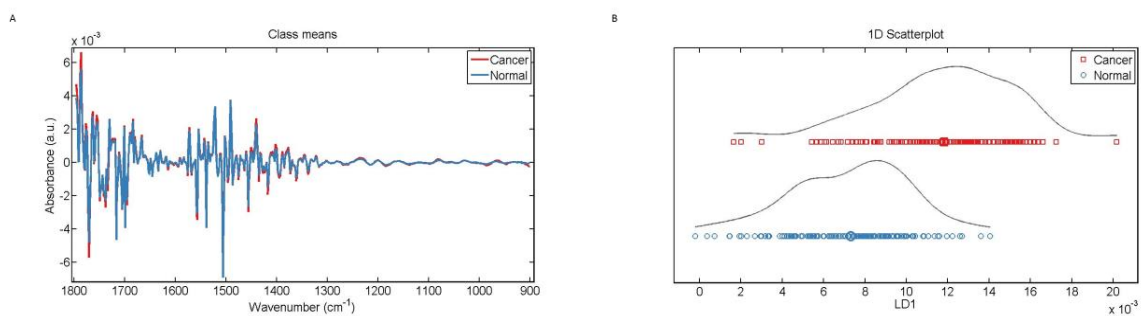


**Figure 3.70:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells among diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-9) reference class, (E) & (F) Patient 2 (H09-9102-1) reference class and (G) & (H) Patient 3 (H09-13558-A5) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Diseased tissues Vs Non-Diseased tissues*

The spectra acquired from the basal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) was compared with the spectra acquired from basal cells in their corresponding non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer. The class representing diseased tissue sections was labelled as ‘Cancer’ whilst the class representing non-diseased tissue sections was labelled as ‘Normal’. In this section classes of tissues will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots. The FTIR mean spectra of both classes was very similar with most evident variations on peak intensities occurring only at  $\sim 1786\text{ cm}^{-1}$  and  $\sim 1771\text{ cm}^{-1}$  whereby Cancer had the highest peaks (Fig. 3.71 A). No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ . Application of PCA-LDA resulted in segregation of classes which was visualised in scores plot of 1D space whereby classes would overlap with each other (Fig. 3.71 B). However, classes were significantly different from each other according to the statistical analysis (Fig. 3.72).



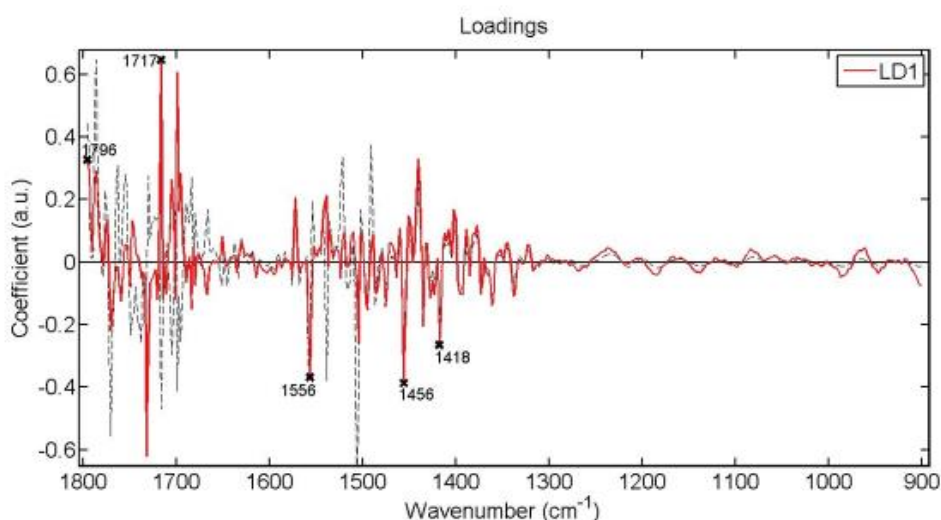


**Figure 3.71:** (A) Mean FTIR spectra for comparison of spectra from basal cells in all diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5) and their corresponding non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). (B) Scores plots for segregation of spectra in 1D space. Diseased and non-diseased tissue samples are represented by the class labels 'Cancer' (red) and 'Normal' (blue) respectively.

Parameters	P Value
Cancer Vs Normal	P<0.0001

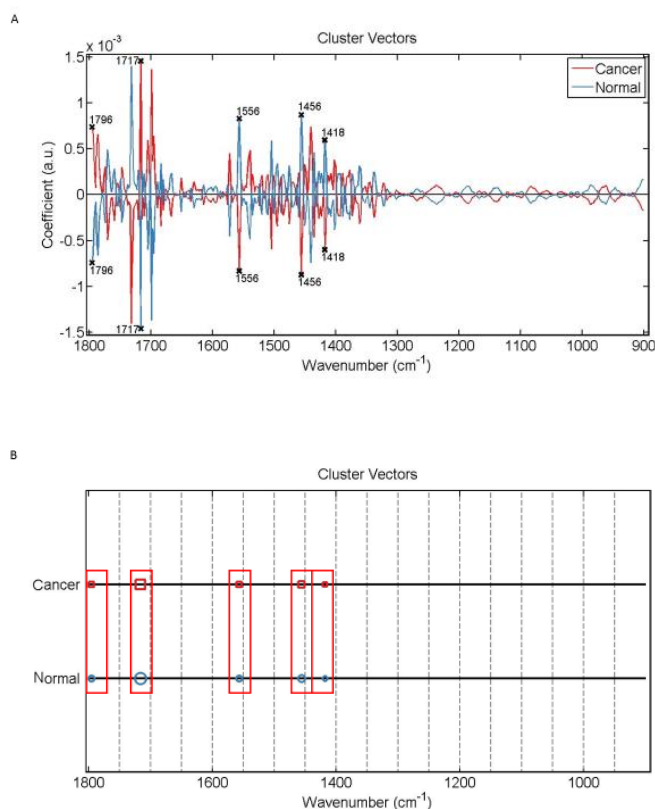
**Figure 3.72:** Obtained p-value by employment of Unpaired t-test to compare spectra acquired from basal cells between non diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and their corresponding diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections represented by the class labels Normal and Cancer respectively.

The discriminant wavenumbers identified in loading plots responsible for variations between the two classes were 1796 cm<sup>-1</sup>, 1717 cm<sup>-1</sup> (amide I, DNA/RNA, purine base), 1556 cm<sup>-1</sup>, 1456 cm<sup>-1</sup> (lipids and proteins) and 1418 cm<sup>-1</sup> (deformation C-H) (Fig. 3.73).



**Figure 3.73:** Loadings plots showing wavenumbers that discriminate spectra from basal cells in all diseased (H09-9102-1, H09-12890-9, H09-13558-A5) and their corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) tissue sections. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

The five major wavenumbers in loading plots were also observed in cluster vector plots occurring in both classes (Fig. 3.74). Also expression levels of these wavenumbers were found to be the same between the classes as indicated by the size of the shapes representing the classes. Highest expression levels were observed for  $1717\text{ cm}^{-1}$ , medium levels were observed for  $1796\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$ . The lowest expression levels were indicated for  $1418\text{ cm}^{-1}$  (deformation C-H). Biomarkers that would specify either 'Normal' or 'Cancer' basal cells could not be identified since the identified discriminant wavenumbers displayed common occurrence between the classes. For a wavenumber to have a potential as a biomarker for either of the classes, it should have been observed in only one class.



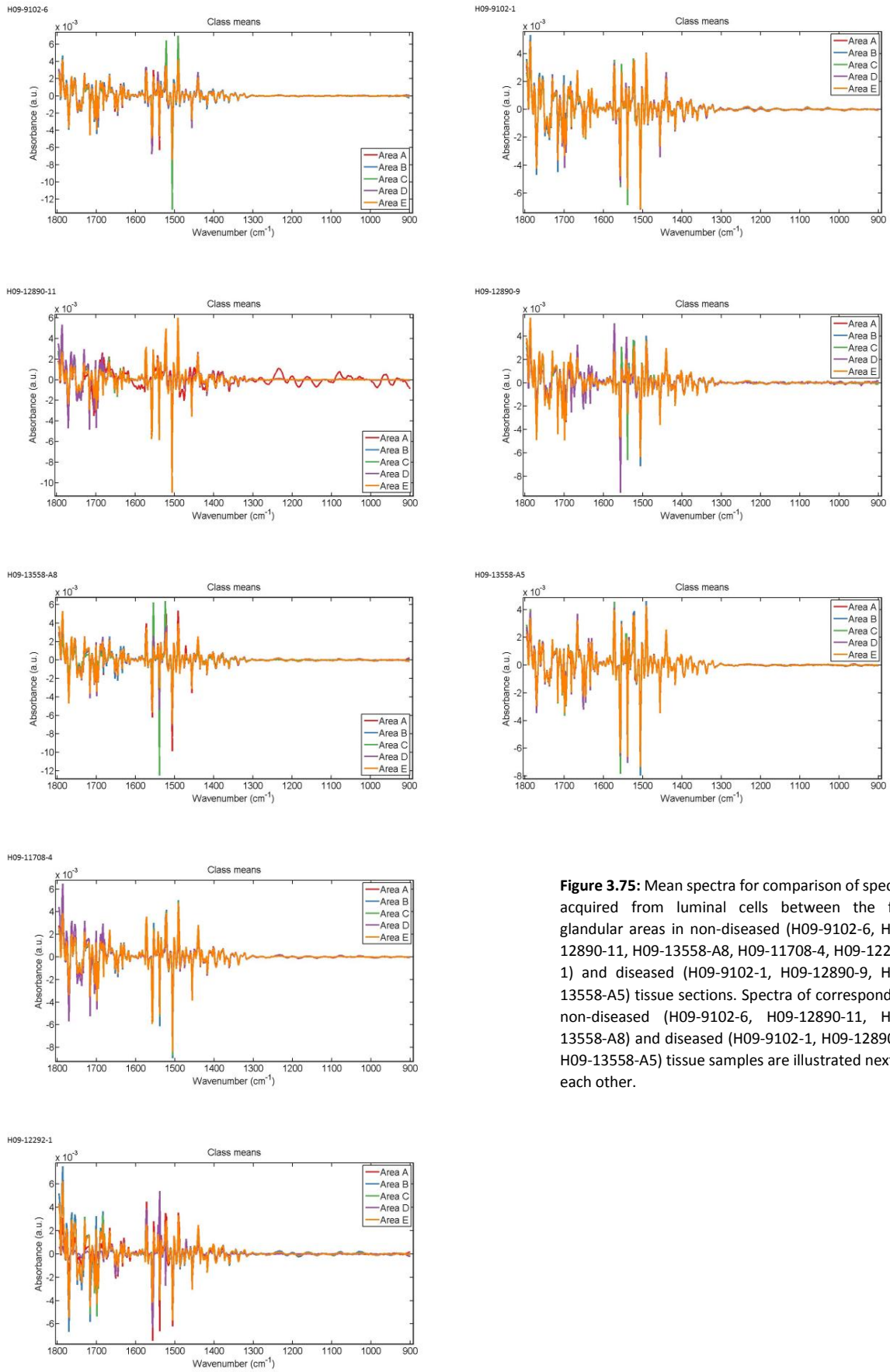
**Figure 3.74:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells among all corresponding diseased (H09-12890-9, H09-9102-1, H09-13558-A5) and non-diseased (H09-12890-11, H09-9102-6, H09-13558-A8) tissue sections. The red rectangles highlight the wavenumbers occurring commonly in

### 3.2.2 Luminal Cells Vs Luminal Cells

#### *In individual tissue samples*

Biochemical composition of luminal cells was also analysed in the hope of identifying potential biomarkers using the same approach as before. Spectra acquired from all five glandular areas in individual tissue samples was analysed. All tissues exhibited relatively the same shape in mean FTIR spectra whereby variations in peak intensities would mainly occur at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1700\text{ cm}^{-1}$ ,  $\sim 1572\text{ cm}^{-1}$ ,  $\sim 1556\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$  and  $\sim 1491\text{ cm}^{-1}$  (Fig. 3.75). No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ .

Statistical analysis (Fig. 3.76) indicated that in the majority of the tissues at least one area was significantly different from the tissue as a whole. Whereas when individual areas within a tissue were compared with each other, they were not significantly different in most of the cases.



**Figure 3.75:** Mean spectra for comparison of spectra acquired from luminal cells between the five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.

## H09-9102-6

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P < 0.001
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.01
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.001
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.01

## H09-12890-11

Parameters	P value
LDA1 vs Area A	P < 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.01
LDA1 vs Area D	P < 0.01
LDA1 vs Area E	P < 0.01
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.05
Area B vs Area D	P < 0.01
Area B vs Area E	P > 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.001

## H09-13558-A8

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P < 0.01
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.01
Area B vs Area C	P < 0.01
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.01
Area C vs Area D	P < 0.001
Area C vs Area E	P < 0.001
Area D vs Area E	P > 0.05

## H09-11708-4

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-12292-1

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-9102-1

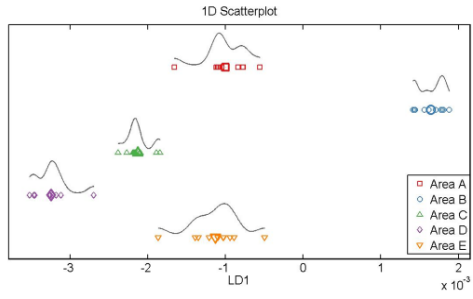
Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P < 0.001
Area D vs Area E	P > 0.05

H09-12890-9		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Area A	P > 0.05	LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05	LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05	LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001	LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05	LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05	Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05	Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001	Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05	Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05	Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.001	Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05	Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001	Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05	Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.001	Area D vs Area E	P > 0.05

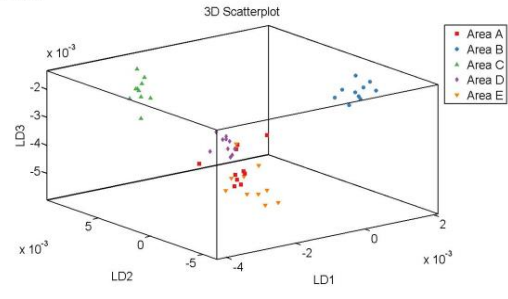
**Figure 3.76:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal in five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Segregation of classes of spectra, resulted from application of PCA-LDA, was visualised in scores plot of 1D and 3D space. Figure 3.77 illustrates scores plots of non-diseased tissue samples and figure 3.78 illustrates scores plots of diseased tissue samples. No consistency amongst samples was observed; spectra of classes, representing the glandular areas, would either display overlapping amongst a few classes or no overlap was observed. In the 3D scores plot spectra would either have a compact or spread arrangement. No observations were made that would correlate corresponding non-diseased and diseased tissue sections.

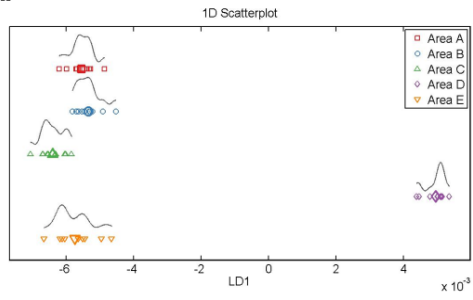
H09-9102-6



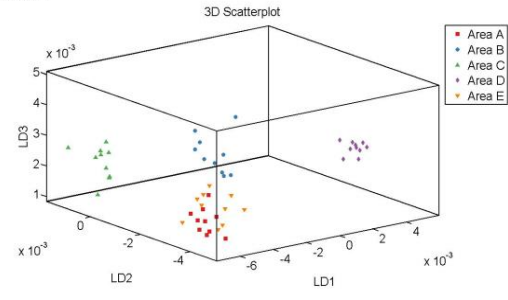
H09-9102-6



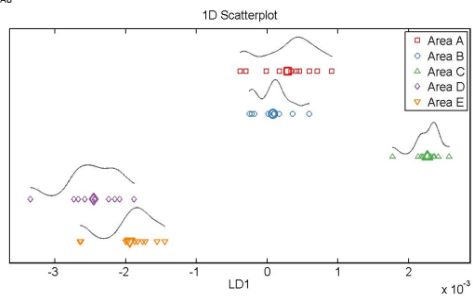
H09-12890-11



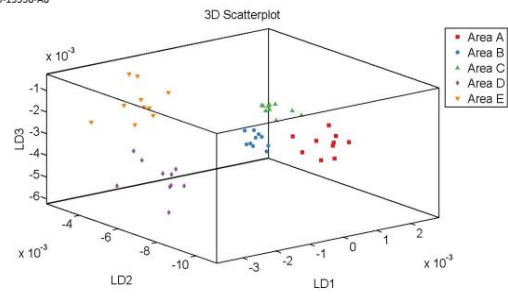
H09-12890-11



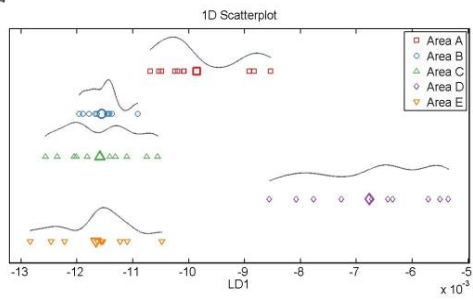
H09-13558-A8



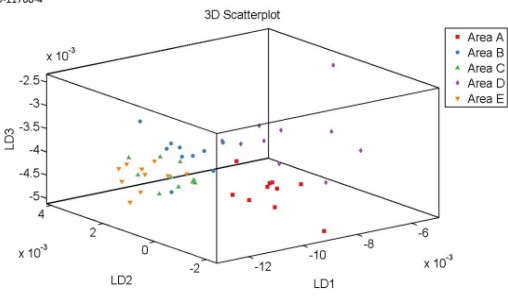
H09-13558-A8

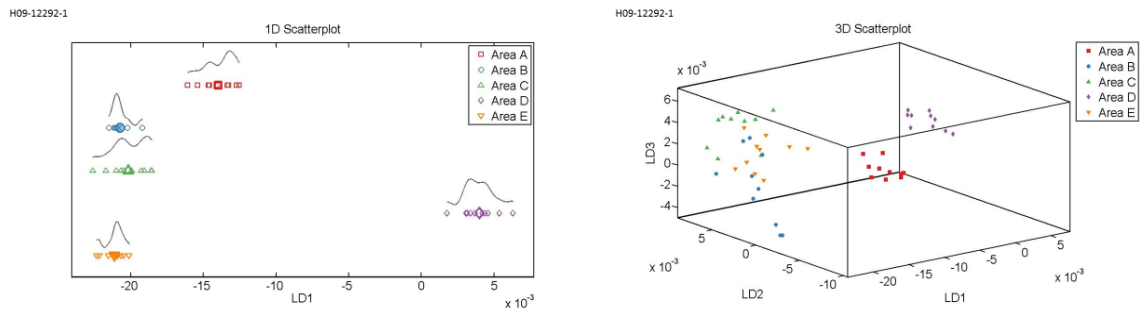


H09-11708-4

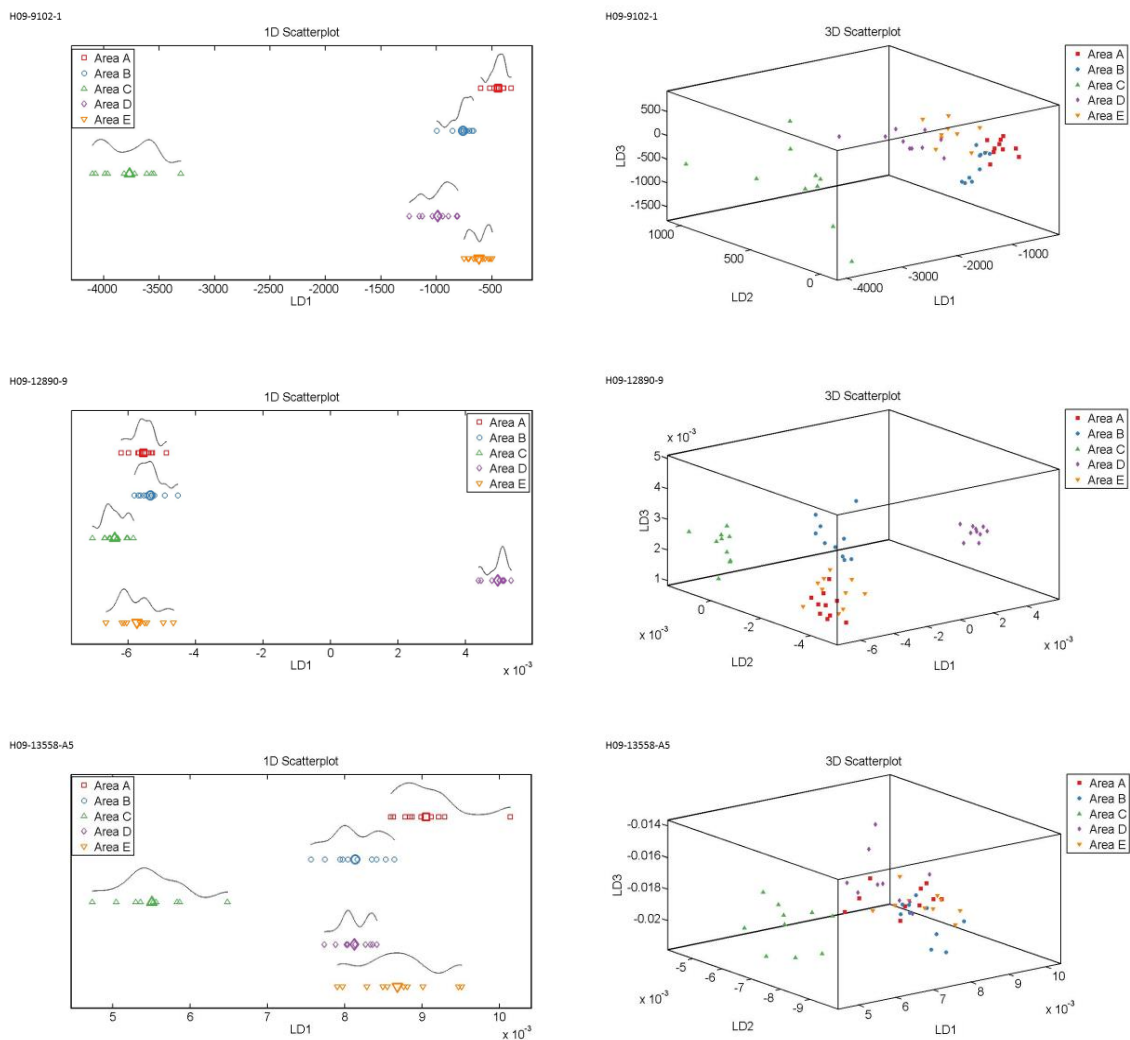


H09-11708-4





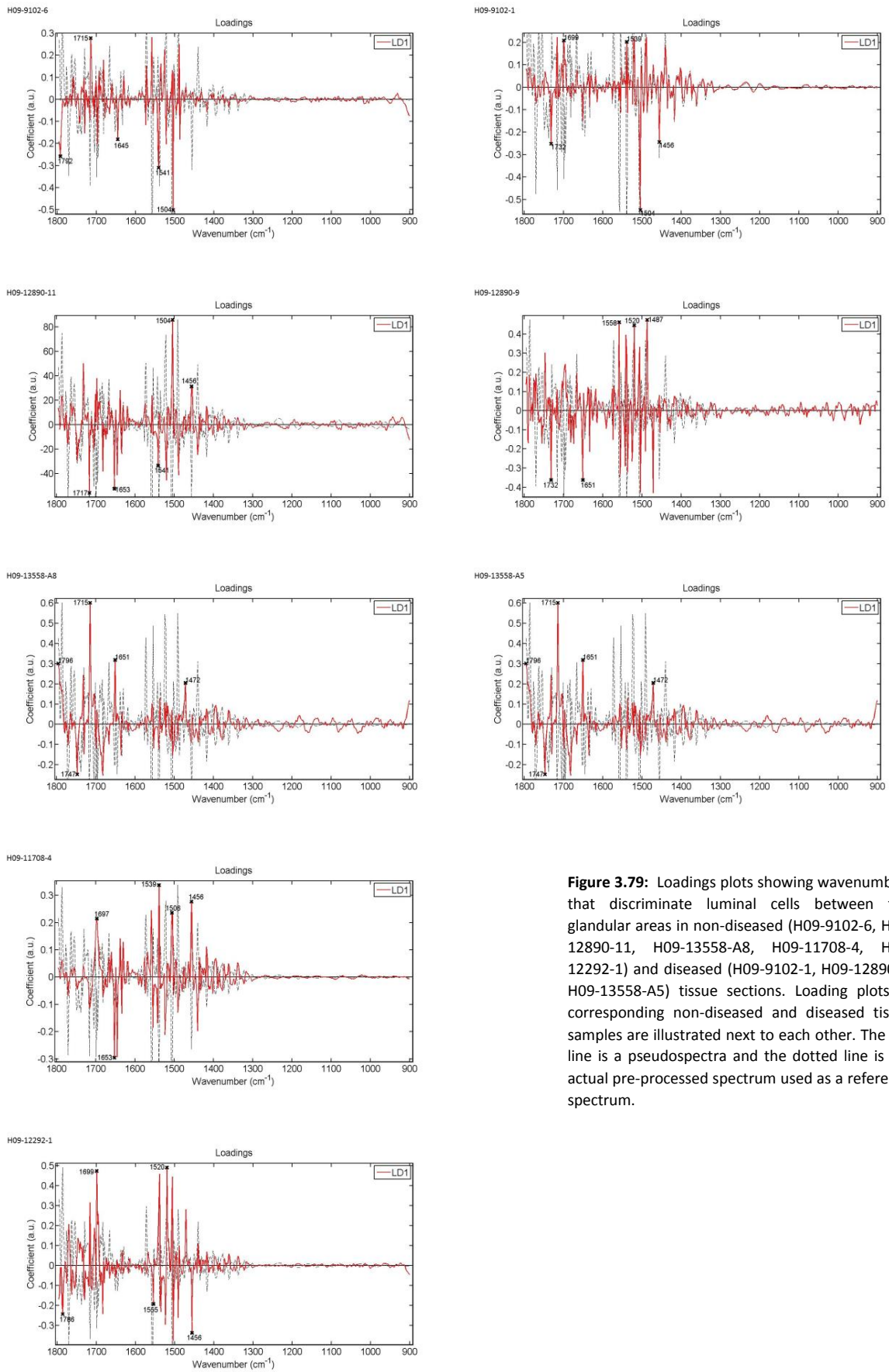
**Figure 3.77:** 1D scores plots and their corresponding 3D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from luminal cells from five different glandular elements.



**Figure 3.78:** 1D scores plots and their corresponding 3D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from luminal cells from five different glandular elements.



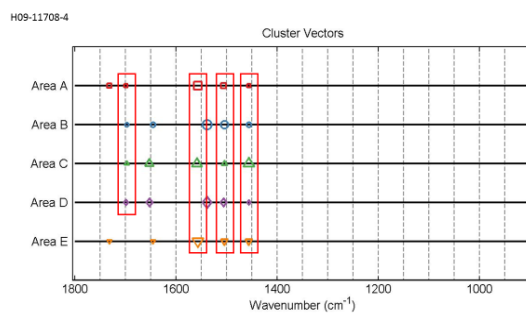
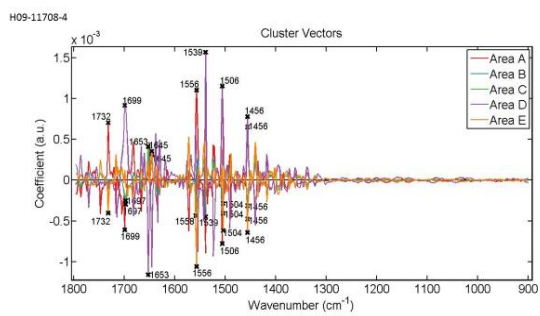
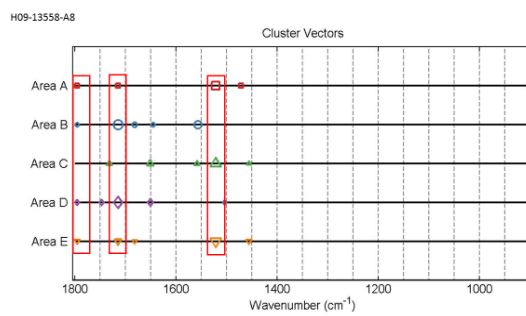
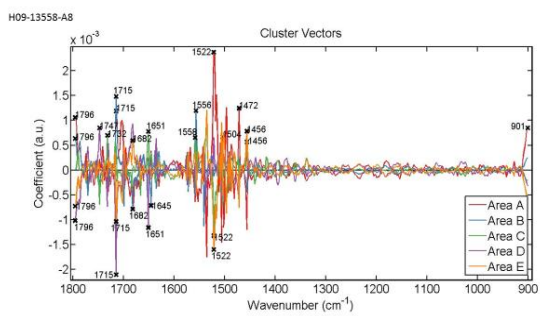
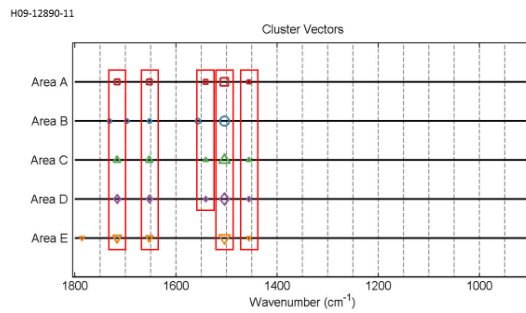
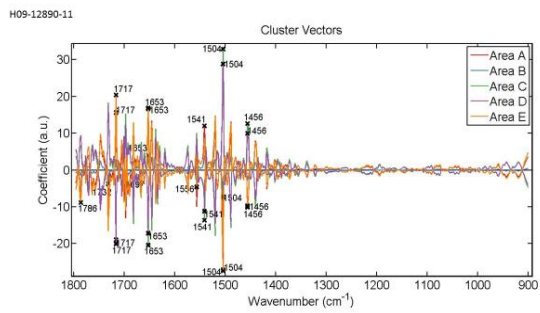
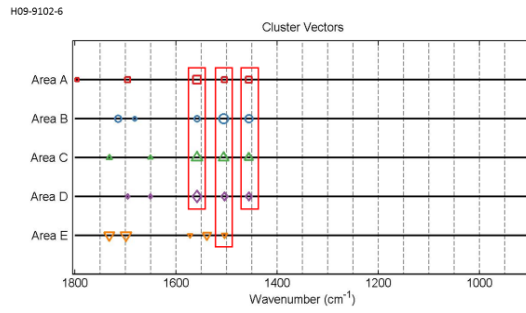
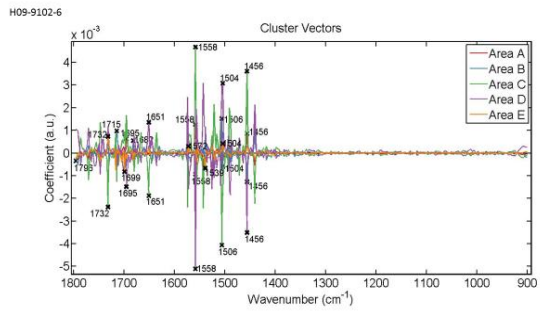
Figure 3.79 illustrates the loading plot for each tissue sample revealing the discriminant wavenumbers responsible for variations between the spectra acquired from luminal cells at glandular areas. In sample H09-9102-6 the discriminant wavenumbers were 1792  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1645  $\text{cm}^{-1}$ , 1541  $\text{cm}^{-1}$  (amide II) and 1504  $\text{cm}^{-1}$  (phenyl rings). In sample H09-12890-11 the wavenumbers were 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1653  $\text{cm}^{-1}$  (amide I), 1541  $\text{cm}^{-1}$  (amide II), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). For sample H09-13558-A8 the identified wavenumbers were 1796  $\text{cm}^{-1}$ , 1747  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$  and 1472  $\text{cm}^{-1}$ . In sample H09-11708-4 the wavenumbers were 1697  $\text{cm}^{-1}$ , 1653  $\text{cm}^{-1}$  (amide I), 1539  $\text{cm}^{-1}$  (amide II), 1506  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). The wavenumbers identified in sample H09-12292-1 were 1786  $\text{cm}^{-1}$ , 1699  $\text{cm}^{-1}$  (guanine/thymine), 1555  $\text{cm}^{-1}$  (ring base/amide II), 1520  $\text{cm}^{-1}$  (amide II) and 1456  $\text{cm}^{-1}$  (lipids and proteins). In the diseased tissue H09-9102-1 the wavenumbers were 1732  $\text{cm}^{-1}$  (lipids), 1699  $\text{cm}^{-1}$  (guanine/thymine), 1539  $\text{cm}^{-1}$  (amide II), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-12890-9 the identified wavenumbers were 1732  $\text{cm}^{-1}$  (lipids), 1651  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$ , 1520  $\text{cm}^{-1}$  (amide II) and 1487  $\text{cm}^{-1}$  (deformation C-H/protein). In sample H09-13558-A5 the wavenumbers were 1796  $\text{cm}^{-1}$ , 1747  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$  and 1472  $\text{cm}^{-1}$ .

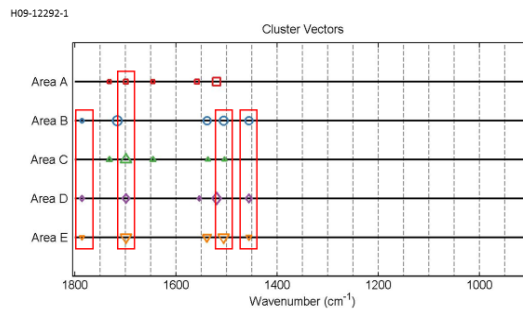
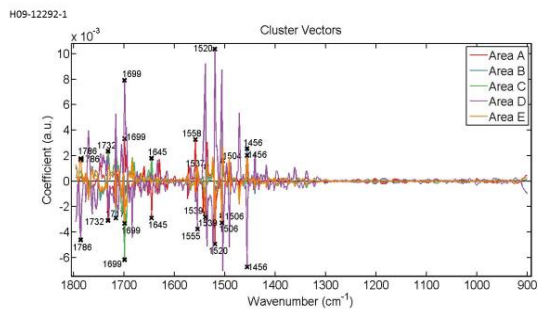


**Figure 3.79:** Loadings plots showing wavenumbers that discriminate luminal cells between five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

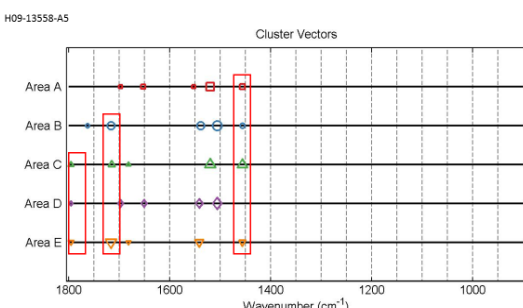
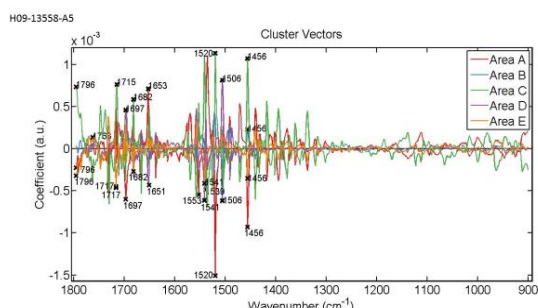
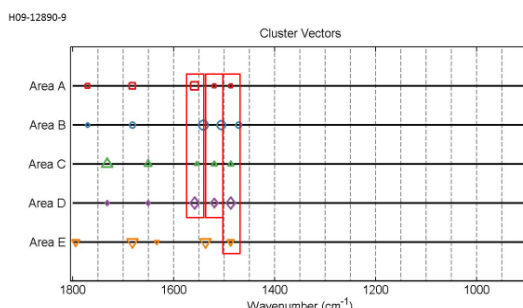
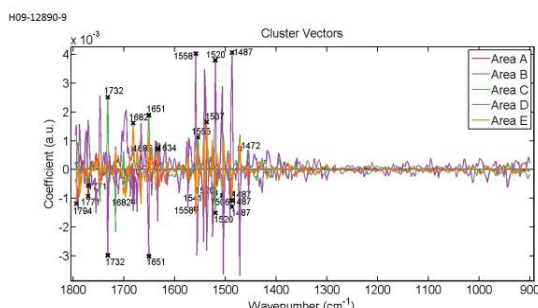
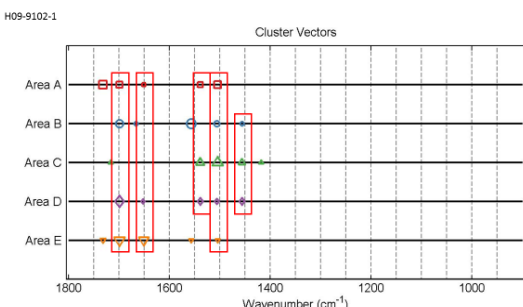
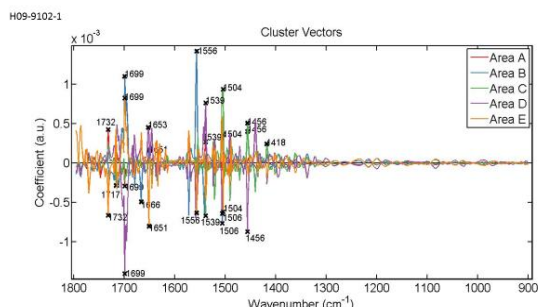
Cluster vector plots of non-diseased (Fig. 3.80) and diseased (Fig. 3.81) tissue samples were presented and analysed in the same approach as before in the hope of identifying possible biomarkers for luminal cells, based on the occurrence of wavenumbers and their expression levels in the glandular areas. In most plots, the wavenumbers displaying common occurrence were the same as the discriminant wavenumbers identified in loading plots.

In sample H09-9102-6 four areas were common for 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins) and all five areas were common for 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings). In sample H09-12890-11 four areas were common for 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1653  $\text{cm}^{-1}$  (amide I), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins) whilst three areas were common for 1541  $\text{cm}^{-1}$  (amide II). In sample H09-13558-A8 four areas were common for 1796  $\text{cm}^{-1}$ , and 1715  $\text{cm}^{-1}$  whilst 1522  $\text{cm}^{-1}$  was commonly shared between three areas. In sample H09-11708-4, 1699  $\text{cm}^{-1}$  (guanine/thymine) with 1697  $\text{cm}^{-1}$  were commonly shared between four areas, five areas were common for 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings), and 1456  $\text{cm}^{-1}$  (lipids and proteins) whilst 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$  were commonly shared among three areas. In sample H09-12292-1, four areas were common only for 1699  $\text{cm}^{-1}$  (guanine/thymine) whilst three areas were common 1786  $\text{cm}^{-1}$ , 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings), and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-9102-1 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings) were commonly shared in all five areas, only 1699  $\text{cm}^{-1}$  (guanine/thymine) was shared between four areas, whilst three areas were common for 1653  $\text{cm}^{-1}$  (amide I) with 1651  $\text{cm}^{-1}$ , 1539  $\text{cm}^{-1}$  (amide II) and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-12890-9 three areas displayed common occurrence for were 1558  $\text{cm}^{-1}$  with 1555  $\text{cm}^{-1}$  (ring base/amide II), and 1520  $\text{cm}^{-1}$  (amide II) and four areas were common for 1487  $\text{cm}^{-1}$  (deformation C-H, protein). Lastly in sample H09-13558-A5 three areas were common for 1796  $\text{cm}^{-1}$  and 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base) with 1715  $\text{cm}^{-1}$ , whilst 1456  $\text{cm}^{-1}$  (lipids and proteins) was commonly shared between four areas.





**Figure 3.80:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells, in five glandular areas in individual non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1). The red rectangles highlight the wavenumbers shared in common among glandular areas.

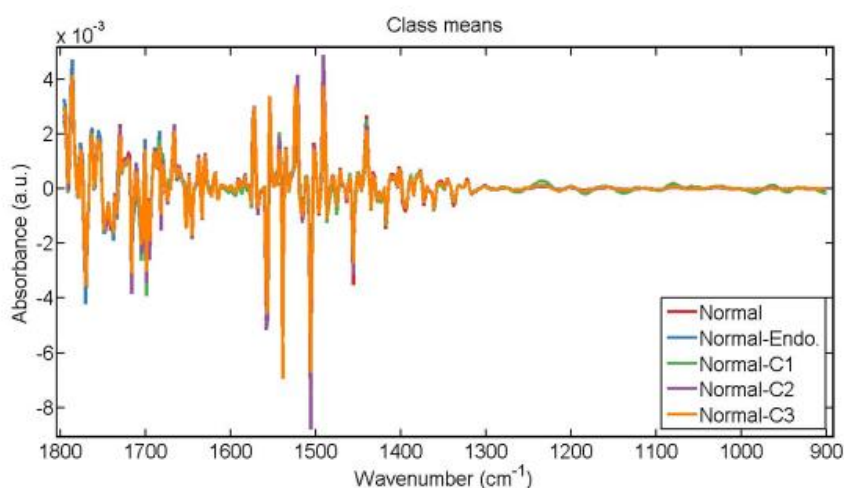


**Figure 3.81:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells, in five glandular areas in individual diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight the wavenumbers shared in common among glandular areas.

*Between all normal tissue samples*

All the non-diseased tissue samples were compared between them for the spectra interrogated from their luminal cells. Classes representing tissues were labelled as before; Normal (H09-11708-4), Normal-Endo. (H09-12292-1), Normal-C1 (H09-12890-11), Normal-C2 (H09-9102-6) and Normal-C3 (H09-13558-A8). In this section tissues will be referred according to their class labels to associate plots with observations. All classes of tissues exhibited relatively the same shape in mean FTIR spectra with most evident variations in peak intensities occurring at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1730\text{ cm}^{-1}$ ,  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1699\text{ cm}^{-1}$ ,  $\sim 1558\text{ cm}^{-1}$ ,  $\sim 1543\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$ ,  $\sim 1491\text{ cm}^{-1}$ ,  $\sim 1456\text{ cm}^{-1}$  and  $\sim 1441\text{ cm}^{-1}$  (Fig. 3.82). No specific class was observed as having the highest or lowest peak intensities throughout the spectra. No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ .

Statistical analysis (Fig. 3.83) indicated that classes were significantly different from each other for the spectra acquired from luminal cells but Normal was not significantly different from Normal-C1 neither from Normal-C2.

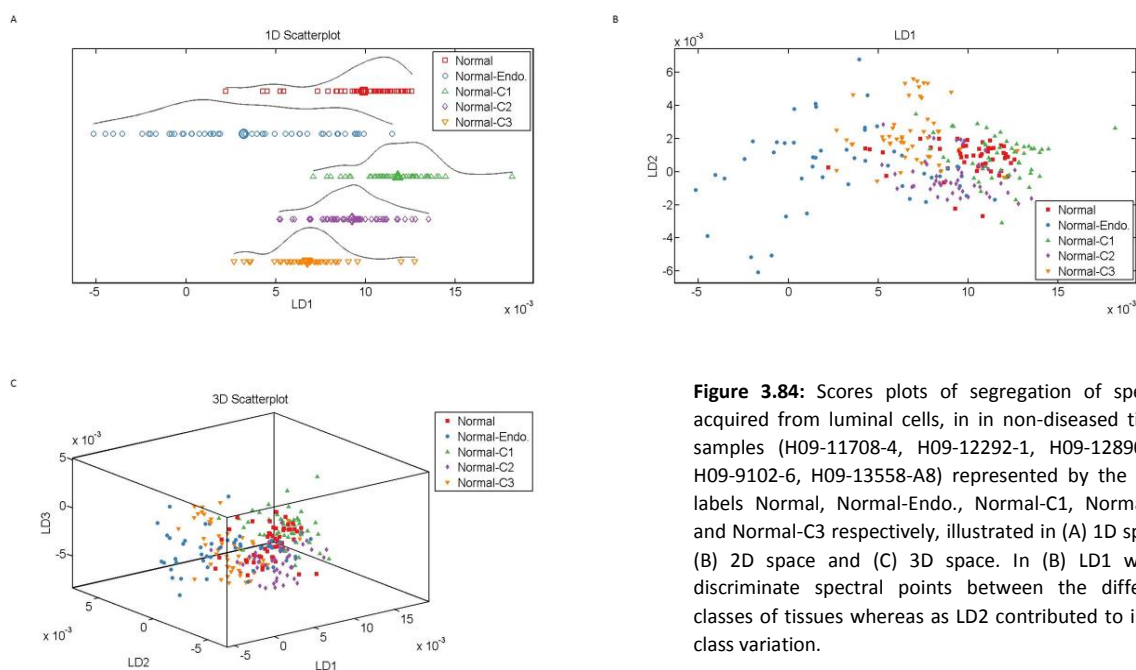


**Figure 3.82:** Mean FTIR spectra for comparison of spectra acquired from luminal cells in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Parameters	P value
LDA1 vs Normal	P < 0.05
LDA1 vs Normal-Endo.	P < 0.001
LDA1 vs Normal-C1	P < 0.001
LDA1 vs Normal-C2	P > 0.05
LDA1 vs Normal-C3	P > 0.05
Normal vs Normal-Endo.	P < 0.001
Normal vs Normal-C1	P > 0.05
Normal vs Normal-C2	P > 0.05
Normal vs Normal-C3	P < 0.001
Normal-Endo. vs Normal-C1	P < 0.001
Normal-Endo. vs Normal-C2	P < 0.001
Normal-Endo. vs Normal-C3	P < 0.001
Normal-C1 vs Normal-C2	P < 0.01
Normal-C1 vs Normal-C3	P < 0.001
Normal-C2 vs Normal-C3	P < 0.01

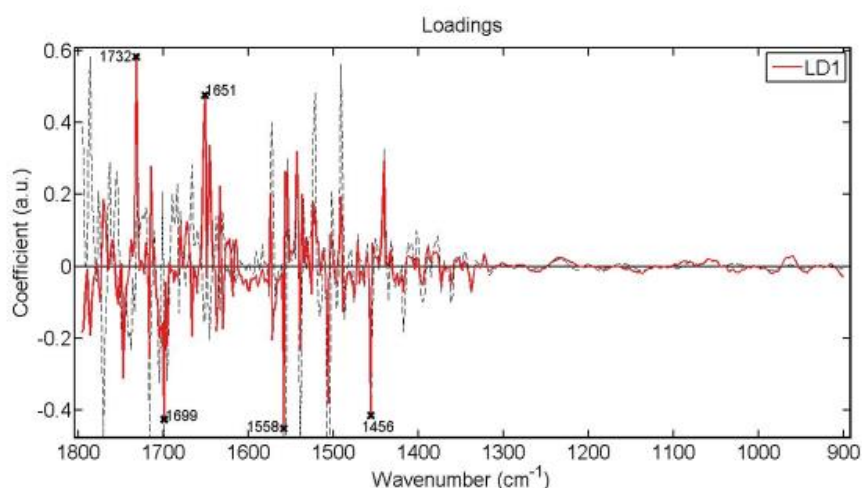
**Figure 3.83:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal cells in all five glandular elements between non-diseased tissue sections H09-11708-4, H09-9102-1, H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D, 2D and 3D space (Fig. 3.84). All classes would overlap with each other mainly throughout their whole spectra. Spectral points of class Normal-Endo. exhibited a relatively more spread arrangement which signified the most intra-class variation. Spectral points of classes Normal, Normal-C1 and Normal-C2 had a more compact arrangement and thus less intra-class variation.



**Figure 3.84:** Scores plots of segregation of spectra acquired from luminal cells, in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively, illustrated in (A) 1D space, (B) 2D space and (C) 3D space. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

The discriminant wavenumbers identified in loading plots accounting for variations between the samples were 1732  $\text{cm}^{-1}$  (lipids), 1699  $\text{cm}^{-1}$  (guanine/thymine), 1651  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins) (Fig. 3.85).



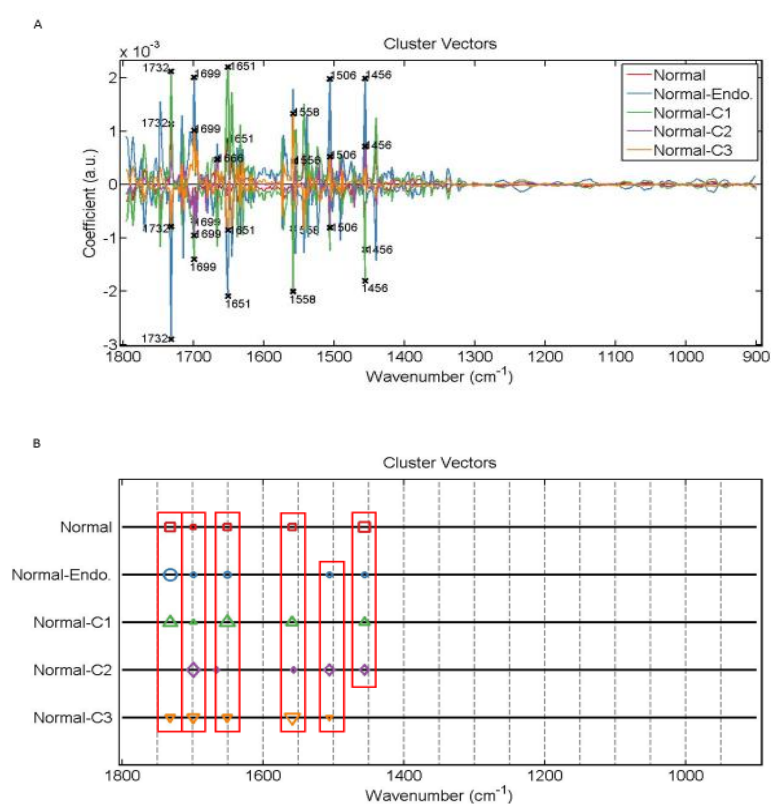
**Figure 3.85:** Loadings plots showing wavenumbers that discriminate luminal cells in non-diseased tissue sections (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

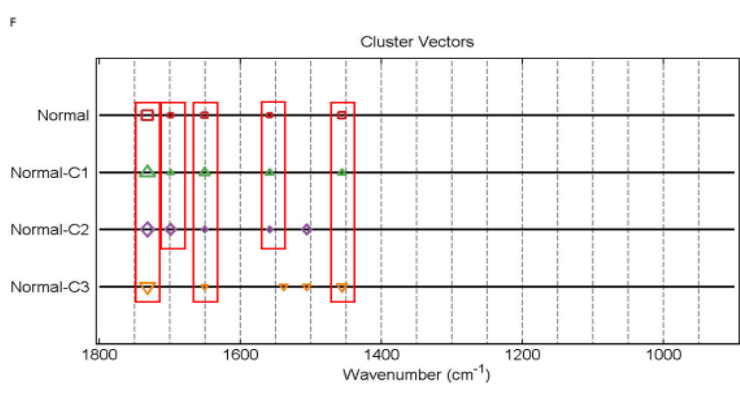
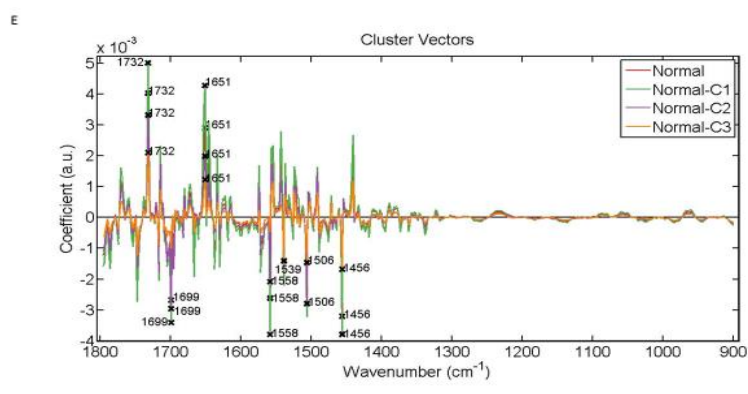
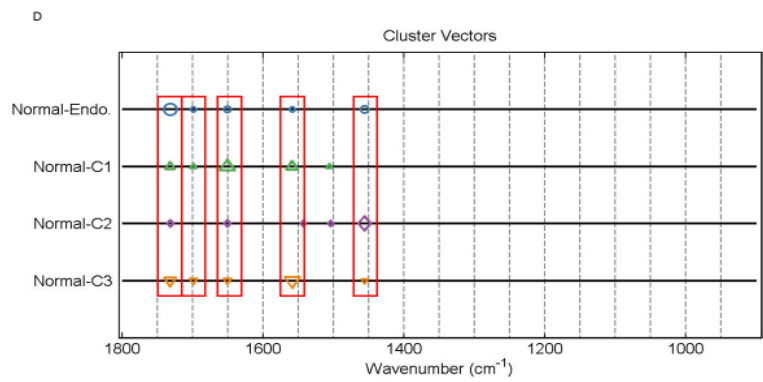
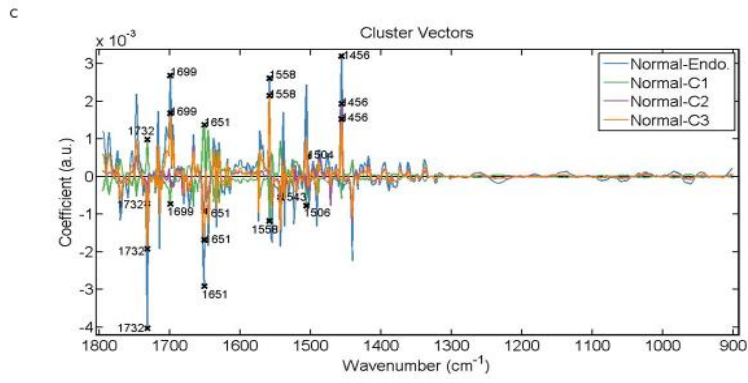
Cluster vector plots (Fig. 3.86) were presented and analysed in a different way in order to identify potential biomarkers for luminal cells by observing the occurrence and expression levels of wavenumbers and/or spectral regions between the tissues by having a class as a reference origin or when no class was used as a reference origin. Wavenumbers would display common occurrence if they were observed in at least three classes representing the tissues as long as they had a smooth vertical alignment in the plots.

Having no class as a reference, all five classes of tissues were common for 1699  $\text{cm}^{-1}$  and four tissues were common for 1732  $\text{cm}^{-1}$  (lipids), 1651  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). Using class Normal as a reference, four tissues were common for 1732  $\text{cm}^{-1}$  (lipids) and 1651  $\text{cm}^{-1}$  and three tissues were common for 1699  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). Having class Normal-Endo. as a reference origin four tissues were common for 1732  $\text{cm}^{-1}$  (lipids) and 1651  $\text{cm}^{-1}$  and three tissues were common for 1699  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). Having class

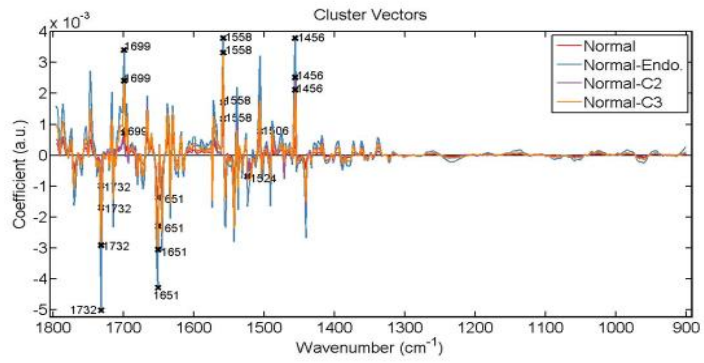


Normal-C1 as reference the wavenumbers 1732  $\text{cm}^{-1}$  (lipids), 1651  $\text{cm}^{-1}$  and 1558  $\text{cm}^{-1}$  were commonly shared by four tissues whilst three tissues were common for 1699  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). Having class Normal-C2 as reference four tissues were common for 1732  $\text{cm}^{-1}$  (lipids) and 1651  $\text{cm}^{-1}$  whilst three tissues were common for 1558  $\text{cm}^{-1}$  and 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings). Having class Normal-C3 as a reference, 1732  $\text{cm}^{-1}$  (lipids) and 1651  $\text{cm}^{-1}$  were commonly shared by four tissues whilst 1699  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins) were commonly shared by three tissues.

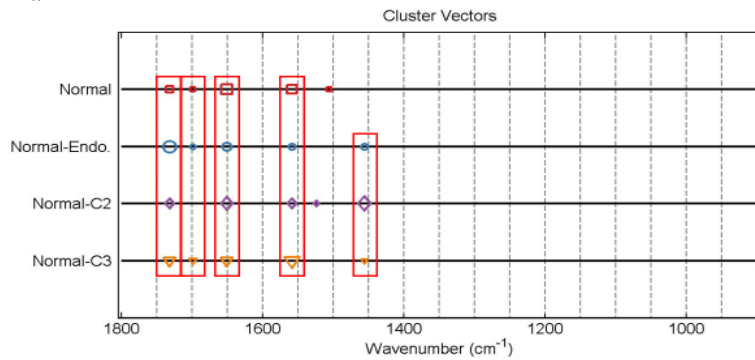




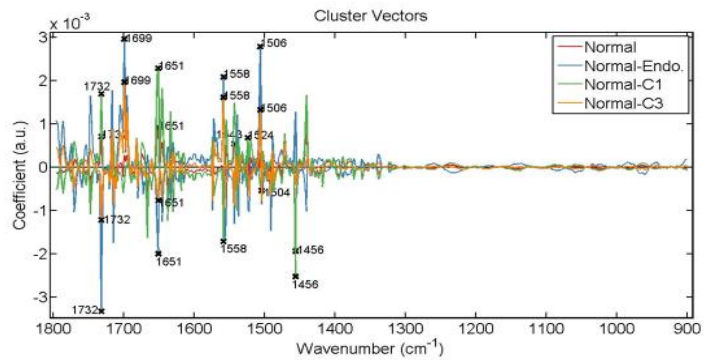
G



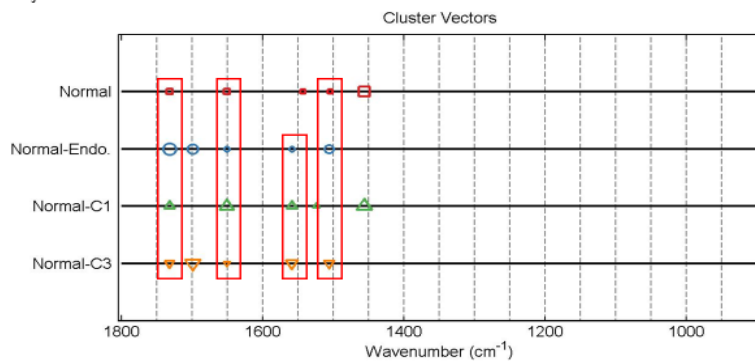
H

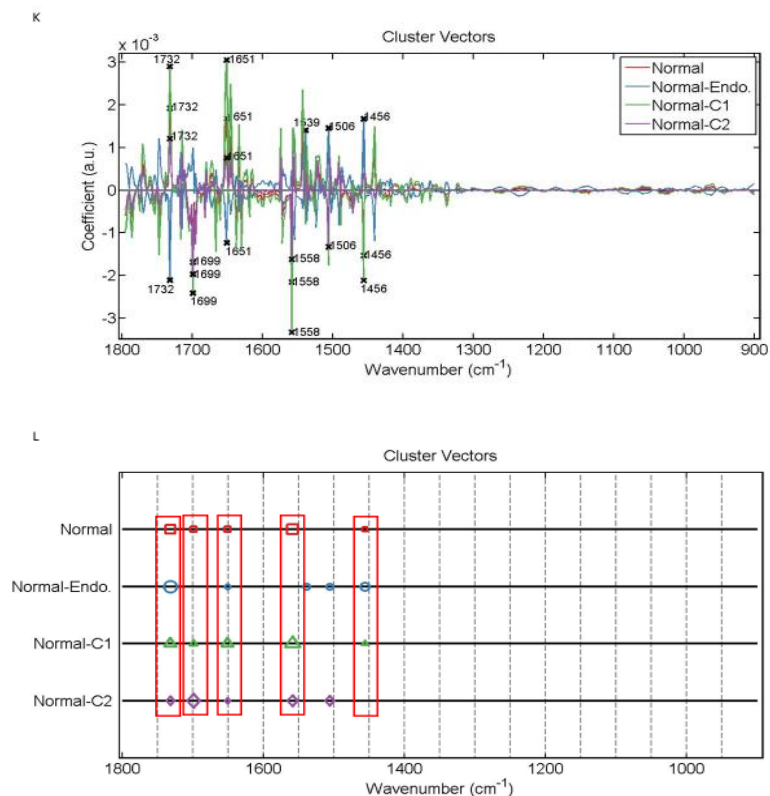


I



J





**Figure 3.86:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells, amongst non-diseased tissue sections. (A) & (B) no sample used as a reference, (C) & (D) Normal (H09-11708-4) reference class, (E) & (F) Normal-Endo. (H09-12292-1) reference class, (G) & (H) Normal-C1 (H09-12890-11) reference class, (I) & (J) Normal-C2 (H09-9102-6) reference class and (K) & (L) Normal-C3 (H09-13558-A8) reference class. The red rectangles in the cluster vector peak location plots highlight the wavenumbers occurring commonly in different tissue samples.

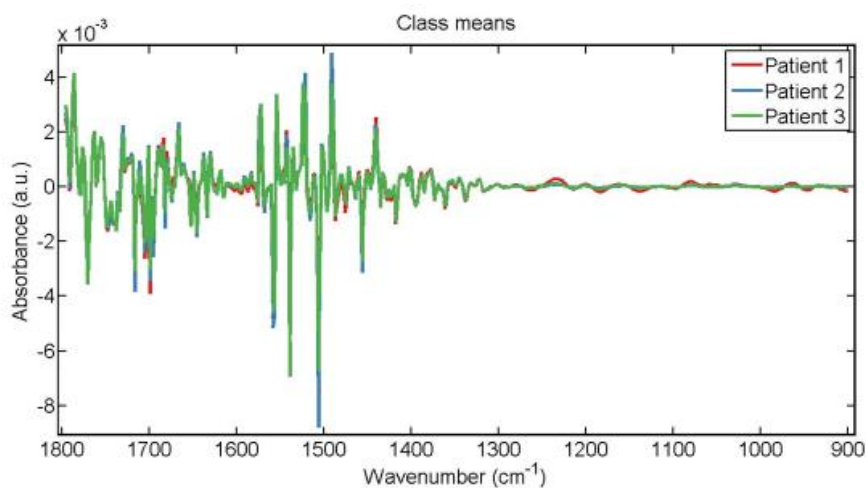
*Between non-diseased samples from patients with endometrial cancer*

Spectra from luminal cells in only non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer were compared. Classes representing the samples were labelled as before so they will be referred accordingly to associate data analysis with what is shown on the plots.

All classes had relatively the same shape for their FTIR mean spectra with most evident variations in absorbance intensities occurring at  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1699\text{ cm}^{-1}$ ,  $\sim 1558\text{ cm}^{-1}$ ,  $\sim 1543\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506$

$\text{cm}^{-1}$ ,  $\sim 1491 \text{ cm}^{-1}$ ,  $\sim 1456 \text{ cm}^{-1}$  and  $\sim 1441 \text{ cm}^{-1}$  (Fig. 3.87). Highest peaks were mainly observed in Patient 2. No peaks were detected in the spectral region  $900\text{-}1300 \text{ cm}^{-1}$ .

Statistical analysis indicated that all tissues were significantly different from each other (Fig. 3.88).

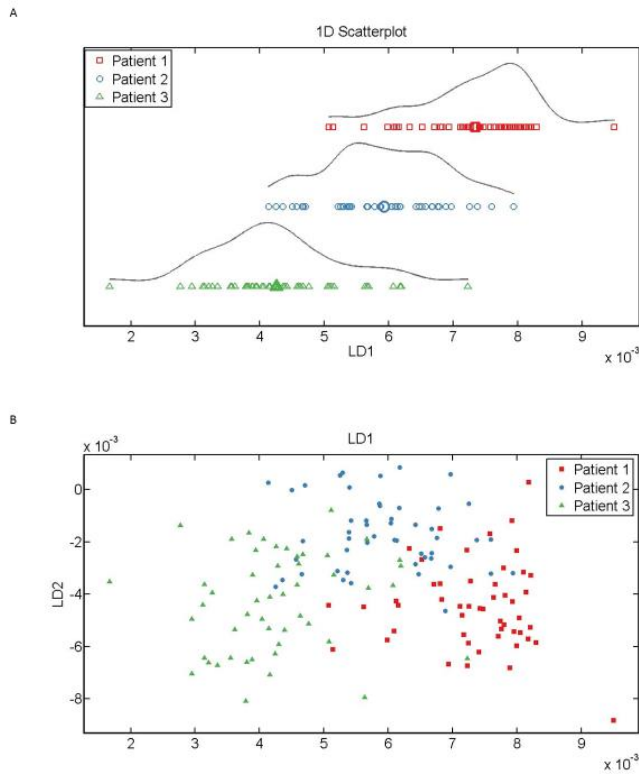


**Figure 3.87:** Mean FTIR spectra for comparison of spectra acquired from luminal cells in non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively.

Parameters	P value
LDA1 vs Patient 1	$P < 0.001$
LDA1 vs Patient 2	$P > 0.05$
LDA1 vs Patient 3	$P < 0.001$
Patient 1 vs Patient 2	$P < 0.001$
Patient 1 vs Patient 3	$P < 0.001$
Patient 2 vs Patient 3	$P < 0.001$

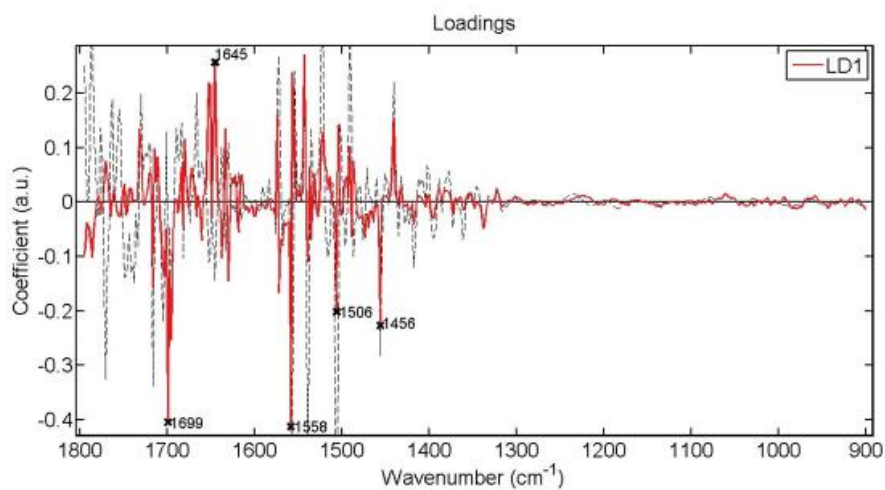
**Figure 3.88:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal cells between non-diseased tissue sections H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D and 2D space (Fig. 3.89). All classes would overlap with each other throughout most of their spectra. Spectral points of Patient 3 had a relatively more spread arrangement in the dimensional space which signified a more intra-class variation.



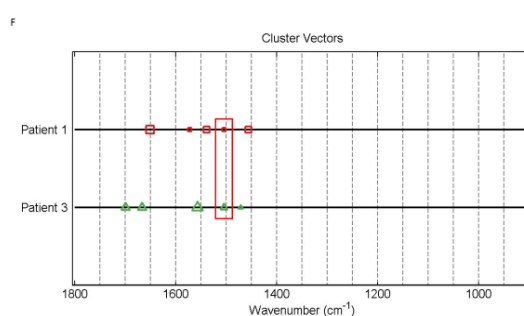
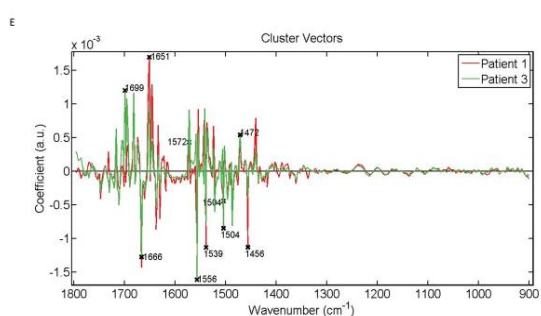
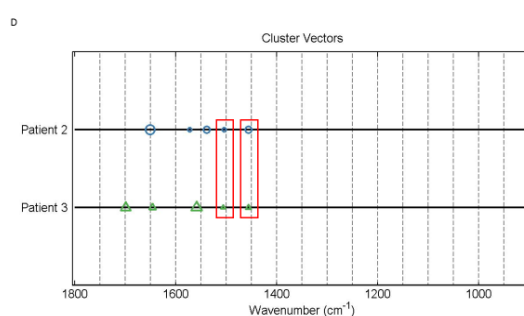
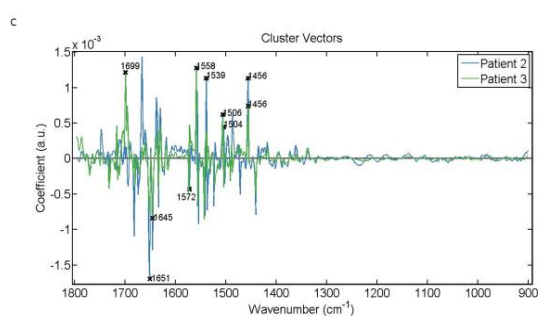
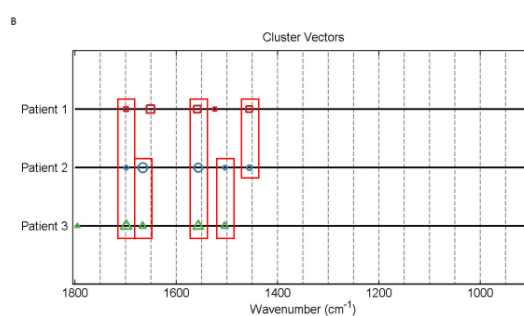
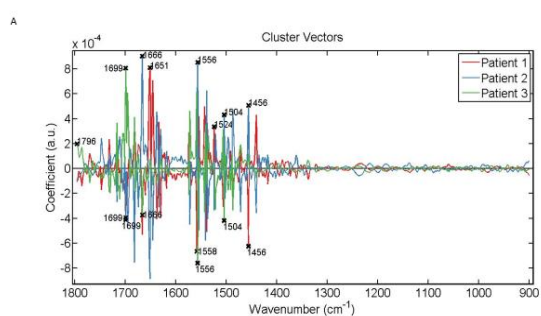
**Figure 3.89:** Scores plots for segregation of spectra in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

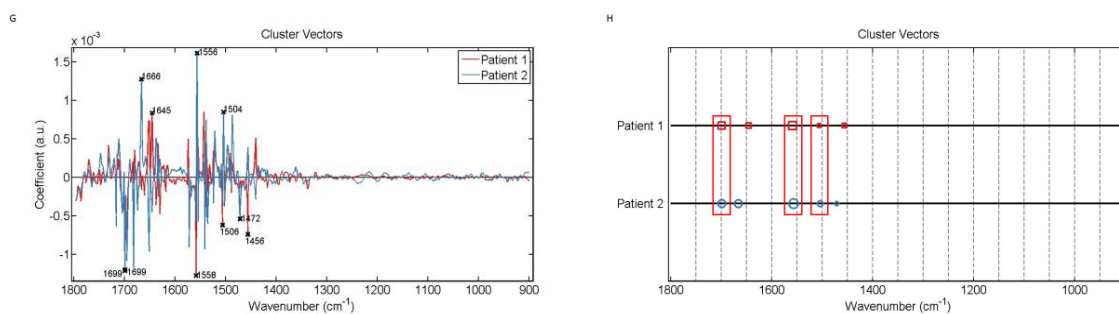
The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were 1699 cm<sup>-1</sup> (guanine/thymine), 1645 cm<sup>-1</sup>, 1558 cm<sup>-1</sup>, 1506 cm<sup>-1</sup> and 1456 cm<sup>-1</sup> (lipids and proteins) (Fig. 3.90).



**Figure 3.90:** Loadings plots showing wavenumbers that discriminate luminal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

These wavenumbers were also observed in cluster vectors plots (Fig. 3.91). These plots were presented and analysed in the same approach as before in an attempt to identify potential biomarkers for luminal cells. Having no class as a reference origin the wavenumbers displaying common occurrence amongst tissue samples were 1699  $\text{cm}^{-1}$  (guanine/thymine), 1666  $\text{cm}^{-1}$  (amide I), 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$ , 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). Using Patient 1 as reference, 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings), and 1456  $\text{cm}^{-1}$  (lipids and proteins) were commonly shared between classes. Using Patient 2 as reference the only commonly shared wavenumber was 1504  $\text{cm}^{-1}$  (phenyl rings). Using Patient 3 as reference common occurrence was observed for 1699  $\text{cm}^{-1}$  (guanine/thymine), 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$ , and 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings).





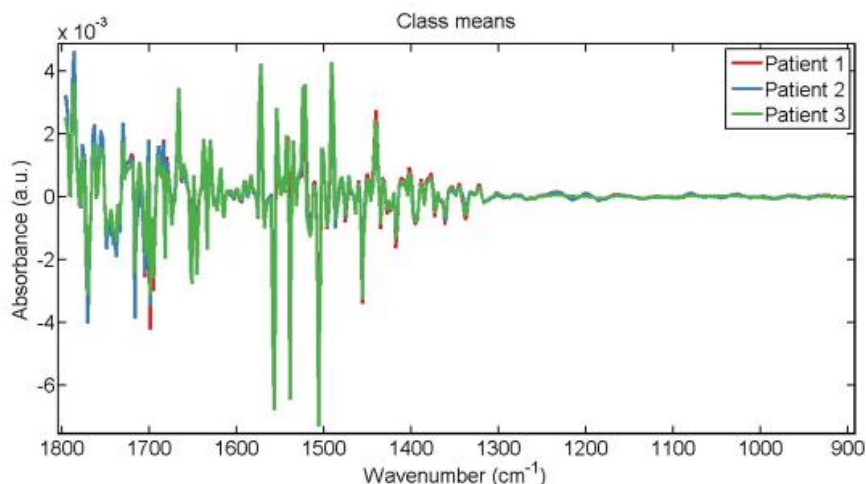
**Figure 3.91:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells among non-diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-11) reference class, (E) & (F) Patient 2 (H09-9102-6) reference class and (G) & (H) Patient 3 (H09-13558-A8) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Between diseased samples from patients with endometrial cancer*

Spectra acquired from luminal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) taken from cancerous lesions in the endometrium were analysed and compared. Classes representing the samples were labelled as before so they will be referred accordingly to associate data analysis with what is shown on the plots. All classes had relatively the same shape of mean spectra and the most evident variations in peak intensities could be identified at  $\sim 1786 \text{ cm}^{-1}$ ,  $\sim 1771 \text{ cm}^{-1}$ ,  $\sim 1763 \text{ cm}^{-1}$ ,  $\sim 1730 \text{ cm}^{-1}$ ,  $\sim 1717 \text{ cm}^{-1}$ ,  $\sim 1699 \text{ cm}^{-1}$  and  $\sim 1441 \text{ cm}^{-1}$  (Fig. 3.92). No peaks were detected in the spectral region  $900\text{-}1300 \text{ cm}^{-1}$ .

Statistical analysis indicated that tissues were significantly different from each other (Fig. 3.93).



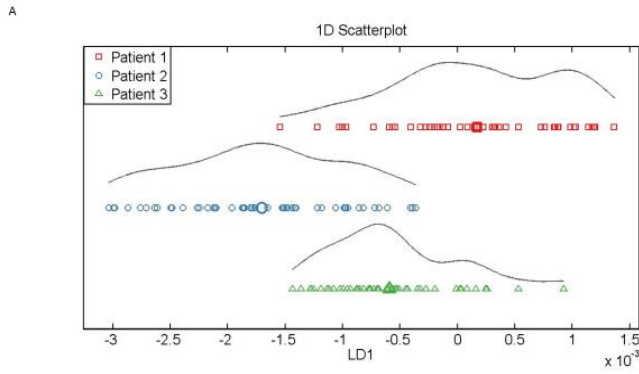


**Figure 3.92:** Mean FTIR spectra for comparison of spectra from luminal cells in diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively.

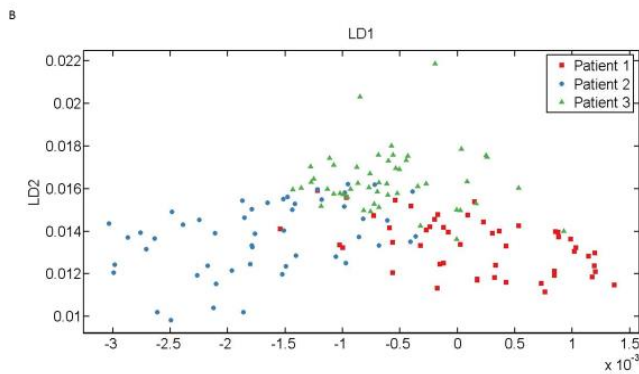
Parameters	P value
LDA1 vs Patient 1	P < 0.001
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P > 0.05
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P < 0.001
Patient 2 vs Patient 3	P < 0.001

**Figure 3.93:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal cells between diseased tissue sections H09-9102-1, H09-12890-9, H09-13558-A5 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

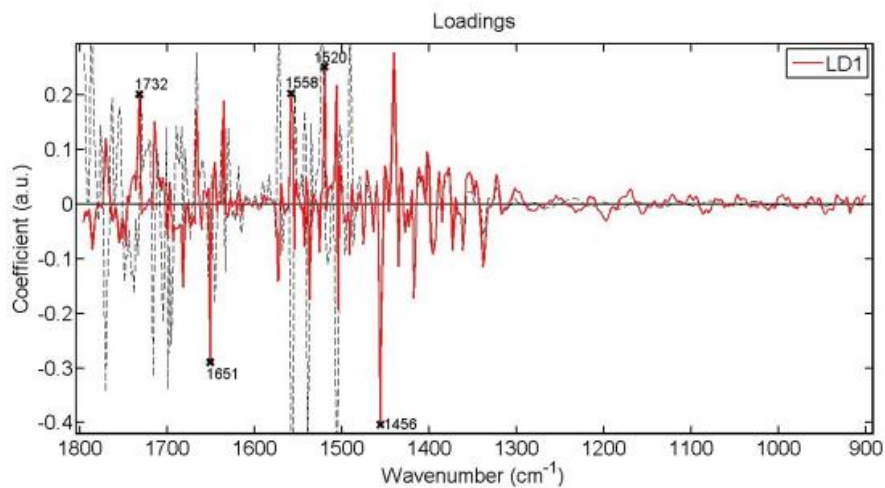
Segregation of classes resulting from application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.94). All classes would overlap with each other. Mainly most of the spectra of Patient 3 overlapped with spectra of the other two classes. Spectral points of Patient 2 had a relative more spread arrangement in the dimensional shape whilst spectral points of Patient 3 exhibited a more compact arrangement. This signified that the first class had more intra-class variation whereas the latter class had less intra-class variation.



**Figure 3.94:** Scores plots for segregation of spectra acquired from luminal cells in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

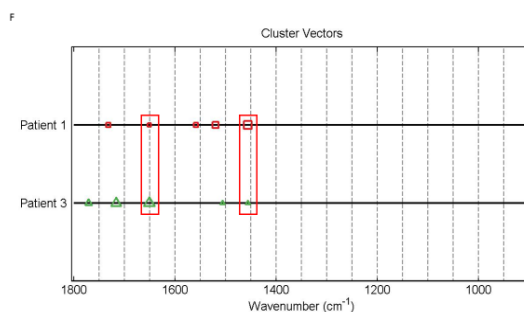
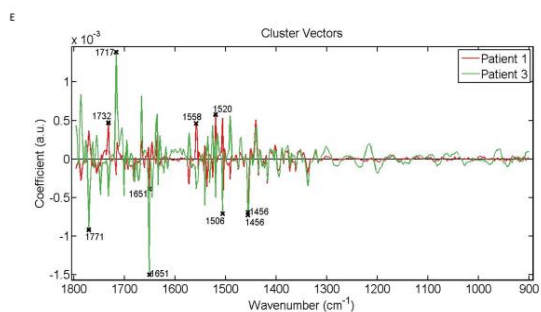
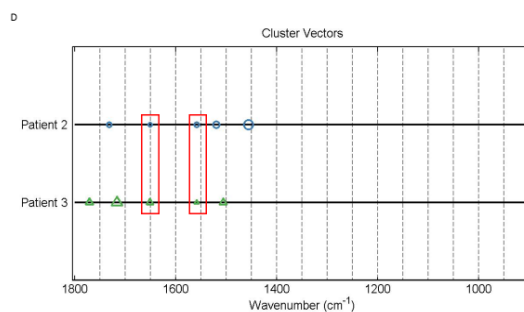
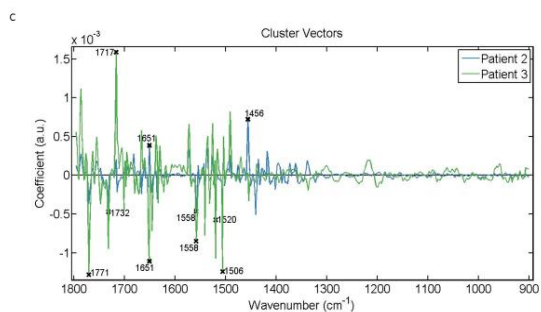
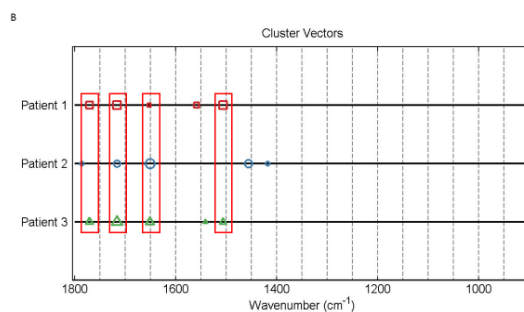
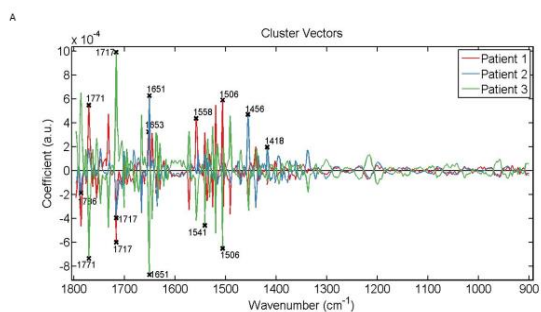


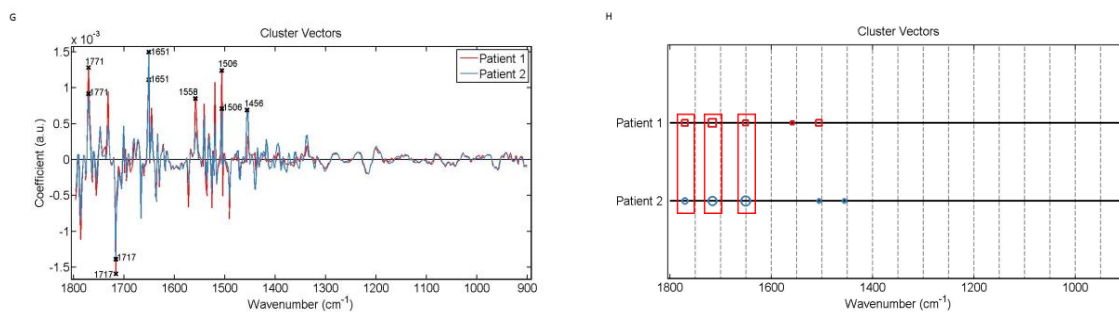
Generation of loading plots revealed five major wavenumbers responsible for variations between the interrogated tissue samples (Fig. 3.95). The identified wavenumbers were  $1732\text{ cm}^{-1}$  (lipids),  $1651\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ ,  $1520\text{ cm}^{-1}$  (amide II) and  $1456\text{ cm}^{-1}$  (lipids and proteins).



**Figure 3.95:** Loadings plots showing wavenumbers that discriminate luminal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference

Cluster vectors plots (Fig. 3.96) were presented and analysed as before to observe occurrence of wavenumbers amongst tissue samples and identify potential biomarkers. Having no class as a reference origin the wavenumbers displaying common occurrence were 1771  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1653  $\text{cm}^{-1}$  (amide I) with 1651  $\text{cm}^{-1}$ , and 1506  $\text{cm}^{-1}$ . The wavenumbers observed to have common occurrence when Patient 1 was the reference class were only 1651  $\text{cm}^{-1}$  and 1558  $\text{cm}^{-1}$ . When Patient 2 was the class reference the wavenumbers displaying common occurrence were 1651  $\text{cm}^{-1}$  and 1558  $\text{cm}^{-1}$ . Having Patient 3 as reference, common occurrence was displayed by 1771  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base) and 1651  $\text{cm}^{-1}$ .

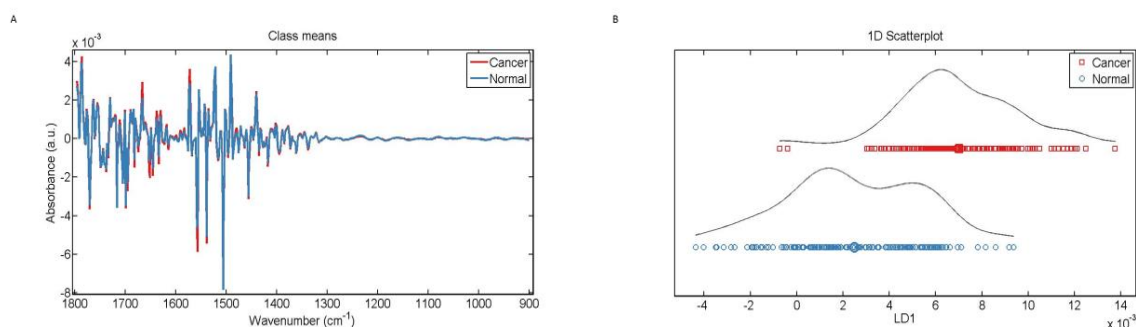




**Figure 3.96:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells among diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-9) reference class, (E) & (F) Patient 2 (H09-9102-1) reference class and (G) & (H) Patient 3 (H09-13558-A5) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Diseased tissues Vs Non-Diseased tissues*

The spectra of the luminal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) were compared with the spectra of luminal cells in their corresponding non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer and the classes were labelled as 'Cancer' and 'Normal' according to the tissue samples they represented. Classes of tissues will be referred according to their labels to associate data analysis with what is shown on the plots. Both classes had the same shape in mean FTIR spectra whereby most evident variations in peak intensities identified at  $\sim 1785\text{ cm}^{-1}$ ,  $\sim 1695\text{ cm}^{-1}$ ,  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1651\text{ cm}^{-1}$ ,  $\sim 1572\text{ cm}^{-1}$  and  $\sim 1556\text{ cm}^{-1}$  (Fig. 3.97 A). No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ . Application of PCA-LDA resulted in segregation of classes which was visualised in scores plot of 1D space whereby classes would overlap with each other (Fig. 3.97 B). However, classes were significantly different from each other according to the statistical analysis (Fig. 3.98).

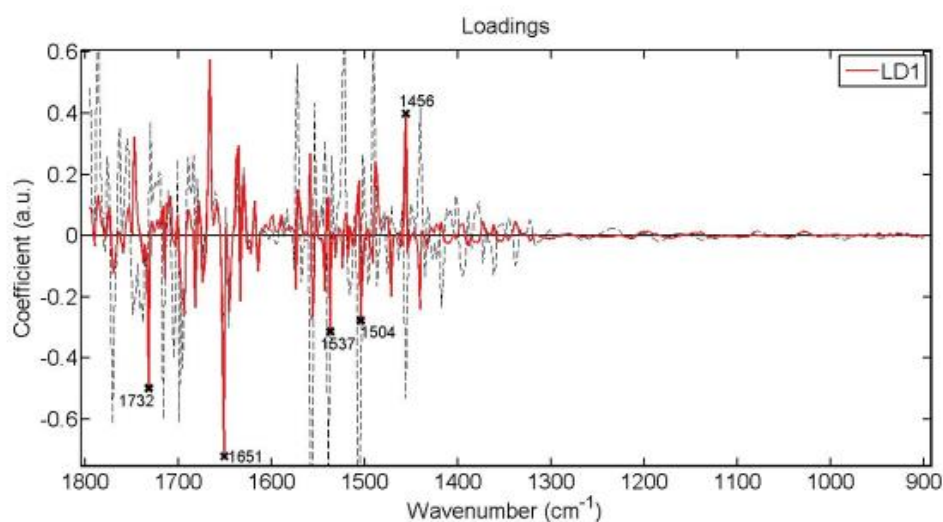


**Figure 3.97:** (A) Mean FTIR spectra for comparison of spectra from luminal cells in all diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5) and their corresponding non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). (B) Scores plots for segregation of spectra in 1D space. Diseased and non-diseased tissue samples are represented by the class labels 'Cancer' (red) and 'Normal' (blue) respectively.

Parameters	P Value
Cancer Vs Normal	P<0.0001

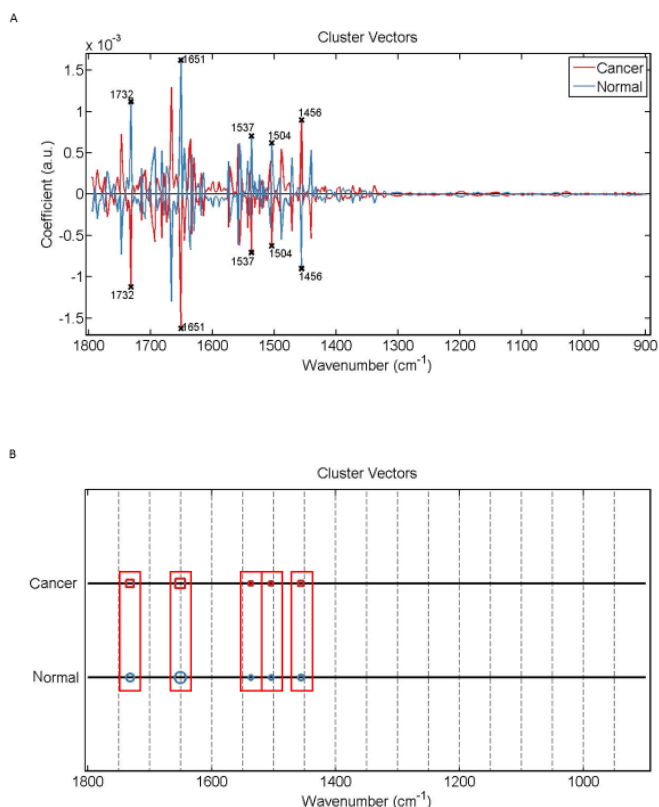
**Figure 3.98:** Obtained p-value by employment of Unpaired t-test to compare spectra acquired from luminal cells between non diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and their corresponding diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections represented by the class labels Normal and Cancer respectively.

The discriminant wavenumbers identified in loading plots responsible for variations between the two classes were 1732 cm<sup>-1</sup> (lipids), 1651 cm<sup>-1</sup>, 1537 cm<sup>-1</sup>, 1504 cm<sup>-1</sup> (phenyl rings) and 1456 cm<sup>-1</sup> (lipids and proteins) (Fig. 3.99).



**Figure 3.99:** Loadings plots showing wavenumbers that discriminate spectra from luminal cells in all diseased (H09-9102-1, H09-12890-9, H09-13558-A5) and their corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) tissue sections. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

These wavenumbers were also observed in cluster vector plots (Fig. 3.100). The wavenumbers displayed common occurrence in both classes so none of the wavenumbers was identified to have a potential as a biomarker to specify either 'Normal' or 'Cancer' luminal cells. Also the wavenumbers exhibited the same concentration levels in both classes with 1651  $\text{cm}^{-1}$  exhibiting the highest concentration levels.



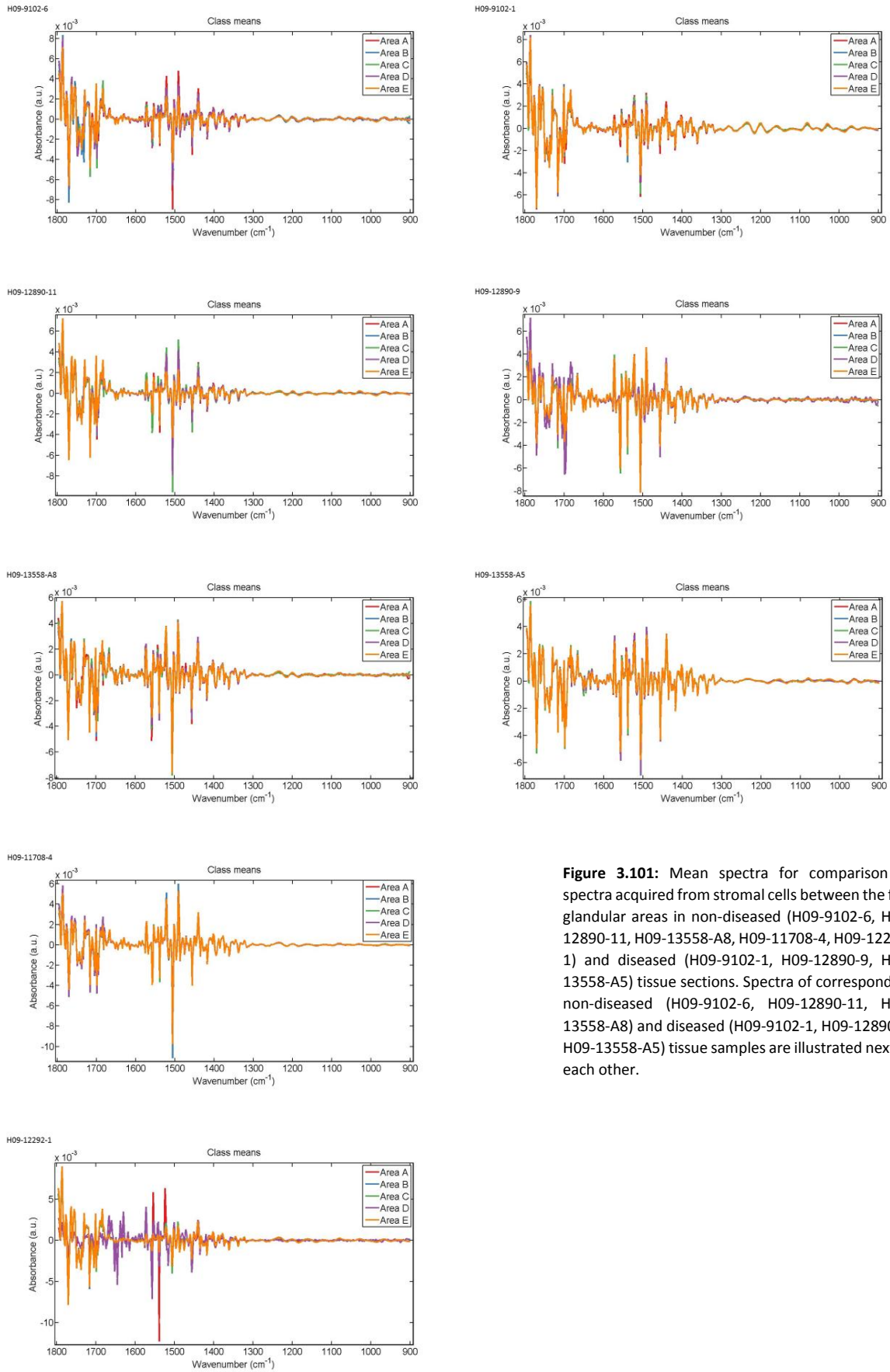
**Figure 3.100:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells among all corresponding diseased (H09-12890-9, H09-9102-1, H09-13558-A5) and non-diseased (H09-12890-11, H09-9102-6, H09-13558-A8) tissue sections. The red rectangles highlight the wavenumbers occurring commonly in classes.

### 3.2.3 Stromal Cells Vs Stromal Cells

#### *In individual tissue samples*

Spectra acquired from stromal cells from the tissue surrounding an individual glandular element was compared with the spectra acquired from the other glandular elements within individual tissue samples. The shape of the mean FTIR spectra of stromal cells was relatively the same amongst glandular areas and amongst tissue samples as well. Most pronounced in absorbance intensities could be identified at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1699\text{ cm}^{-1}$ ,  $\sim 1684\text{ cm}^{-1}$ ,  $\sim 1572\text{ cm}^{-1}$ ,  $\sim 1556\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$  and  $\sim 1524\text{ cm}^{-1}$  (Fig. 3.101). No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ .

Statistical analysis (Fig. 3.102) indicated that in the non-diseased tissues the majority of the glandular areas were significantly different from the tissue as a whole whereas the majority of glandular areas in diseased tissues were not significantly different from the tissue as a whole. In most of the tissues the majority of the areas were not significantly different from each other.



**Figure 3.101:** Mean spectra for comparison of spectra acquired from stromal cells between the five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.



## H09-9102-6

Parameters	P value
LDA1 vs Area A	P < 0.05
LDA1 vs Area B	P < 0.001
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P > 0.05
Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.001
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

## H09-12890-11

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P < 0.001
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.05
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.05
Area B vs Area E	P < 0.001
Area C vs Area D	P < 0.05
Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.05

## H09-13558-A8

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

## H09-11708-4

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P < 0.01
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-12292-1

Parameters	P value
LDA1 vs Area A	P < 0.01
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P < 0.01
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-9102-1

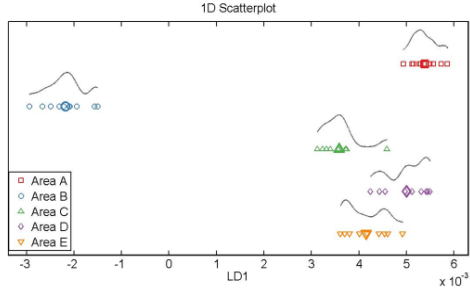
Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P < 0.001
Area D vs Area E	P > 0.05

H09-12890-9		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Area A	P > 0.05	LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05	LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05	LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001	LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05	LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05	Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05	Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001	Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05	Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05	Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.001	Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05	Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001	Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05	Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.001	Area D vs Area E	P > 0.05

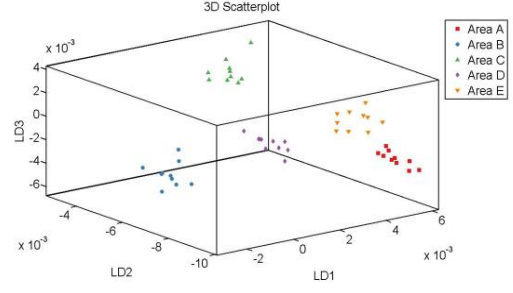
**Figure 3.102:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal in five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Segregation of classes of spectra, resulted from application of PCA-LDA, was visualised in scores plot of 1D and 3D space. Figure 3.103 illustrates scores plots of non-diseased tissue samples and figure 3.104 illustrates scores plots of diseased tissue samples. In the majority of the tissues, spectral points of at least three glandular areas would overlap with each other. In some tissues, spectral points of individual areas were spread in the dimensional space (H09-12890-11, H09-11708-4, H09-9102-1, H09-13558-A5) whereas in other tissues spectral points of individual areas displayed a compact arrangement (H09-9102-6, H09-12890-9, H09-13558-A8, H09-12292-1). No observations were made that would correlate corresponding non-diseased and diseased tissue sections.

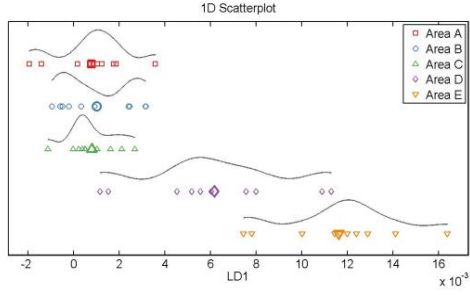
H09-9102-6



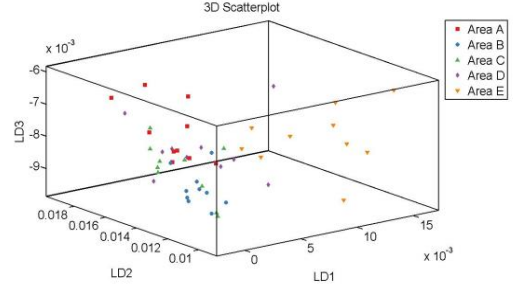
H09-9102-6



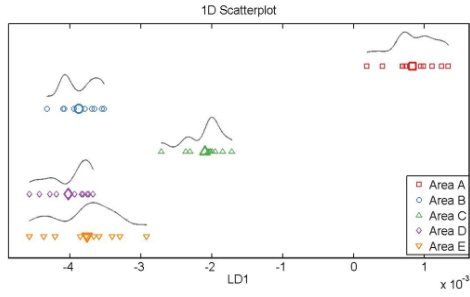
H09-12890-11



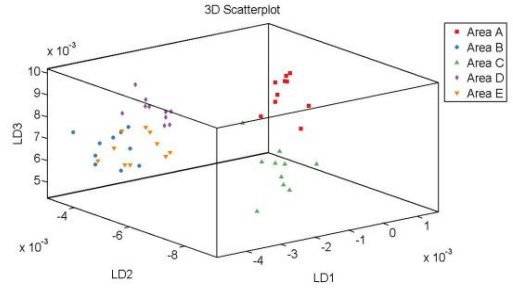
H09-12890-11



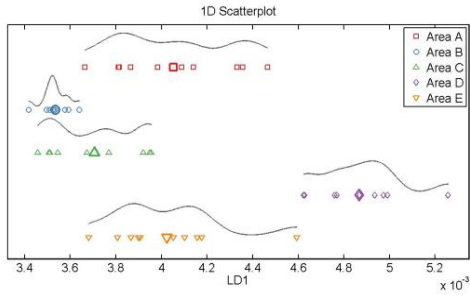
H09-13558-A8



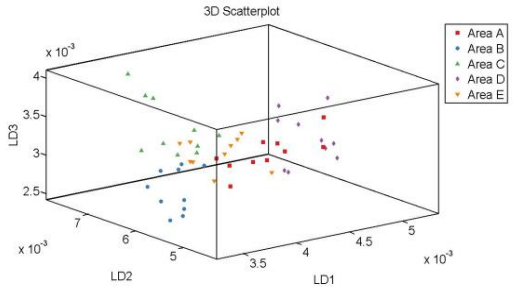
H09-13558-A8

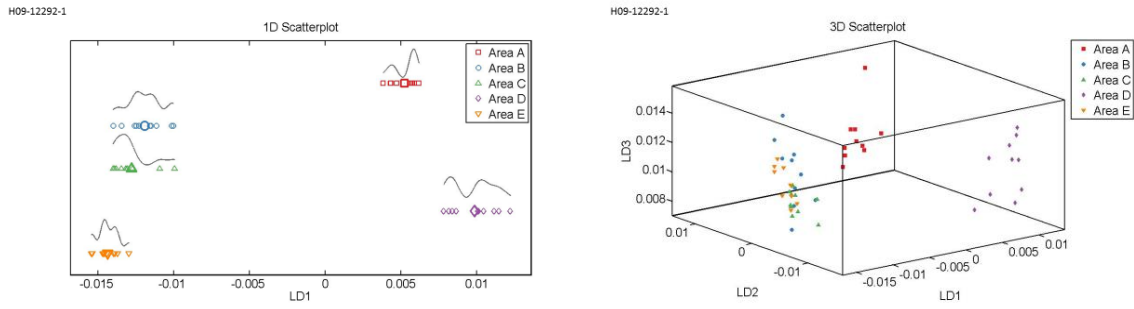


H09-11708-4

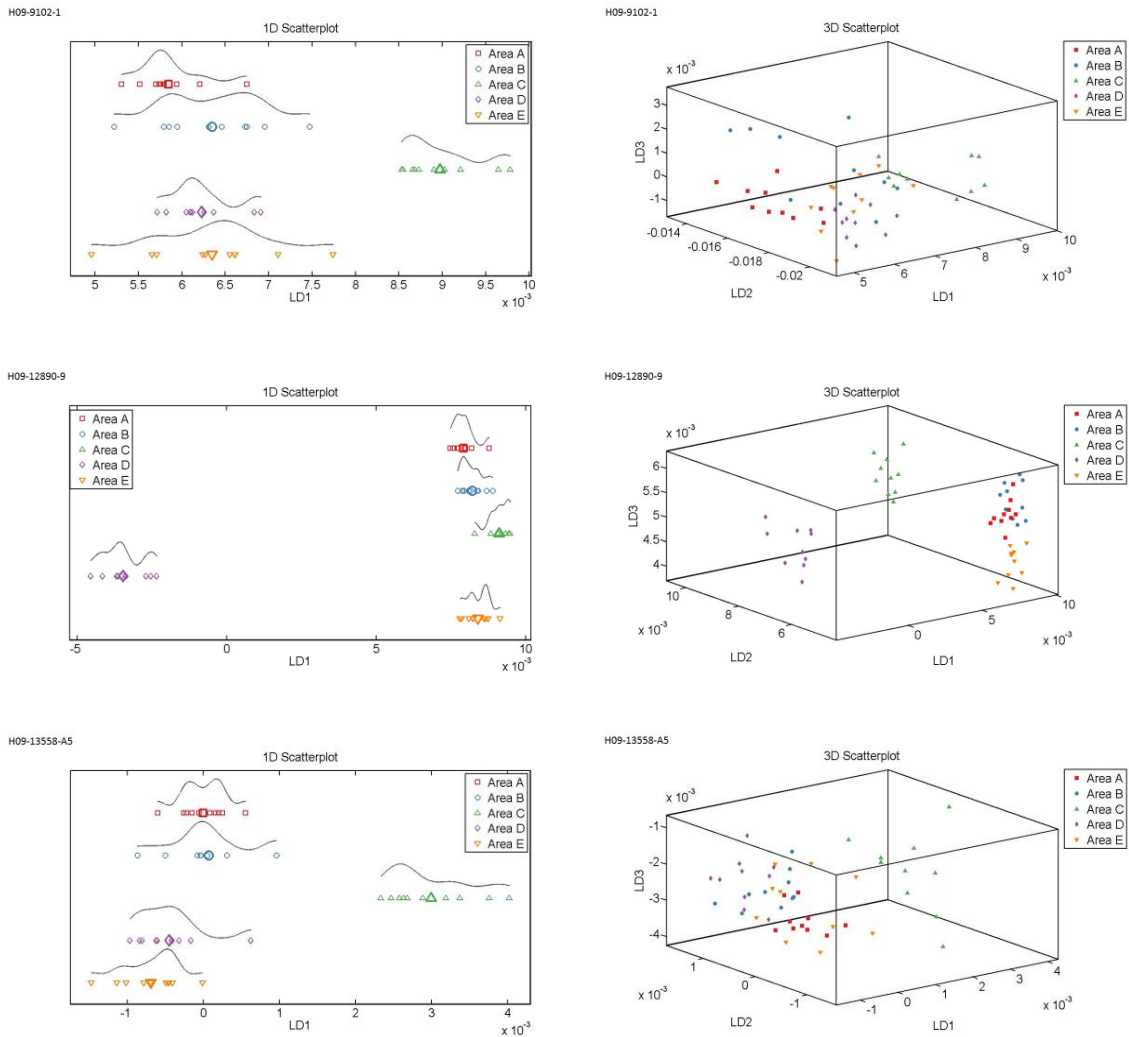


H09-11708-4



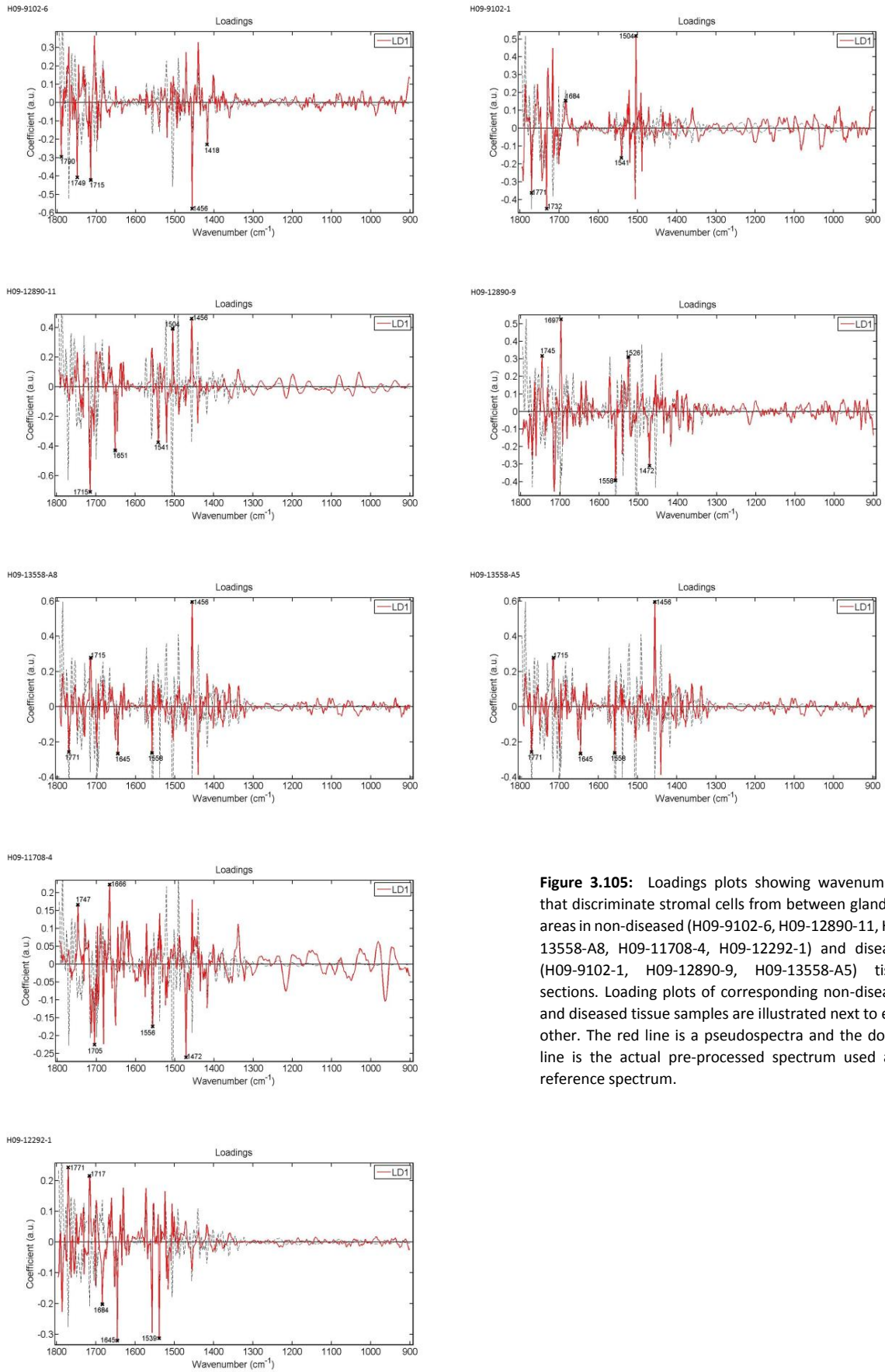


**Figure 3.103:** 1D scores plots and their corresponding 3D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from stromal cells from five different glandular elements.



**Figure 3.104:** 1D scores plots and their corresponding 3D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from stromal cells from five different glandular elements.

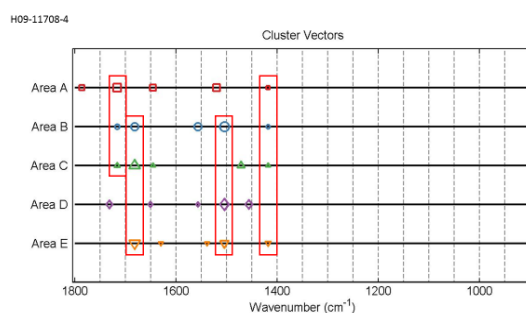
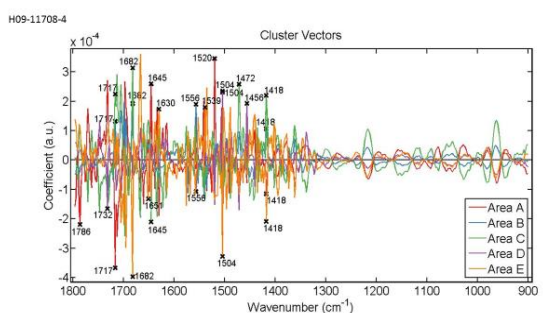
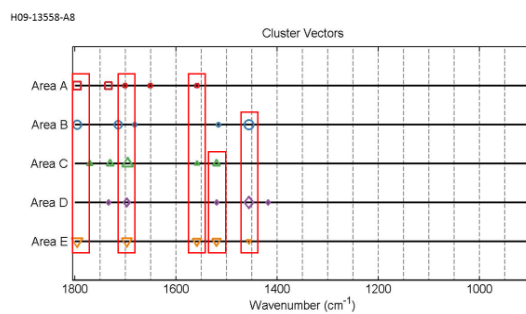
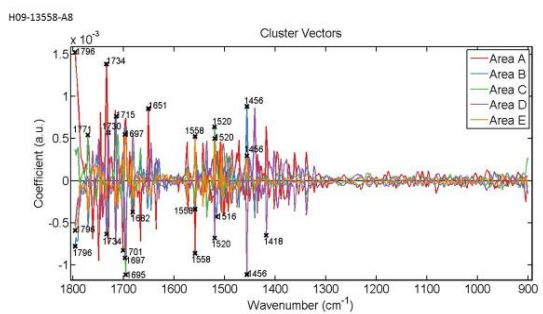
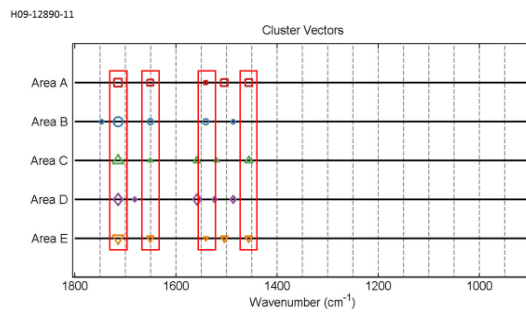
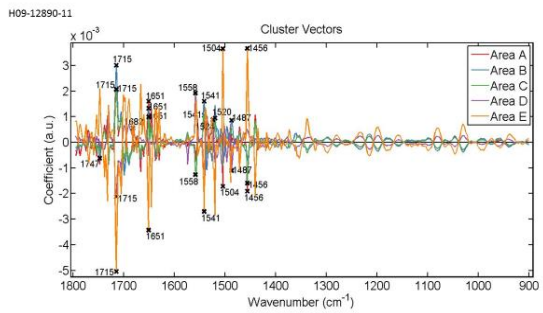
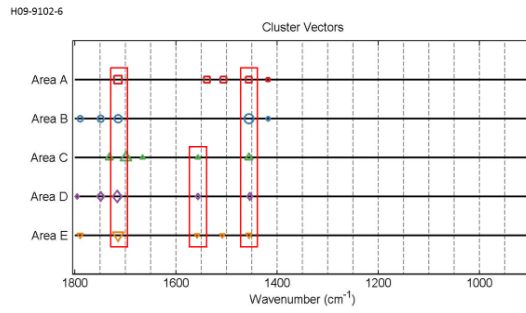
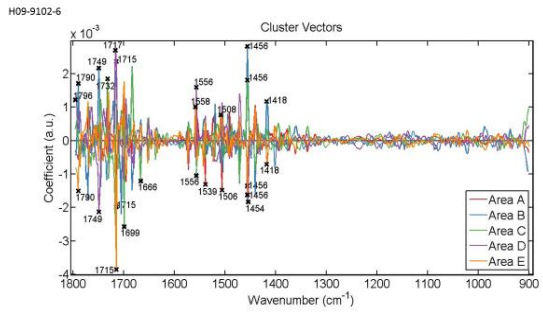
Figure 3.105 illustrates the loading plots for each tissue sample revealing the discriminant wavenumbers responsible for variations between the spectra acquired from stromal cells at glandular areas. In sample H09-9102-6 the discriminant wavenumbers were 1790  $\text{cm}^{-1}$ , 1749  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1456  $\text{cm}^{-1}$  (lipids and proteins) and 1418  $\text{cm}^{-1}$  (deformation C-H). In sample H09-12890-11 the wavenumbers were 1715  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$ , 1541  $\text{cm}^{-1}$  (amide II), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). For sample H09-13558-A8 the identified wavenumbers were 1771  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1645  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-11708-4 the wavenumbers were 1747  $\text{cm}^{-1}$ , 1705  $\text{cm}^{-1}$  (lipids), 1666  $\text{cm}^{-1}$  (amide I), 1556  $\text{cm}^{-1}$  and 1472  $\text{cm}^{-1}$ . The wavenumbers identified in sample H09-12292-1 were 1771  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1684  $\text{cm}^{-1}$  (guanine deformation), 1645  $\text{cm}^{-1}$  and 1539  $\text{cm}^{-1}$  (amide II). In the diseased tissue H09-9102-1 the wavenumbers were 1771  $\text{cm}^{-1}$ , 1732  $\text{cm}^{-1}$  (lipids), 1684  $\text{cm}^{-1}$  (guanine deformation), 1541  $\text{cm}^{-1}$  (amide II) and 1504  $\text{cm}^{-1}$  (phenyl rings). In sample H09-12890-9 the observed wavenumbers were 1745  $\text{cm}^{-1}$  (triglycerides/polysaccharides, pectin), 1697  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$ , 1526  $\text{cm}^{-1}$  (guanine) and 1472  $\text{cm}^{-1}$ . Lastly in sample H09-13558-A5 the wavenumbers were 1771  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1645  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins).



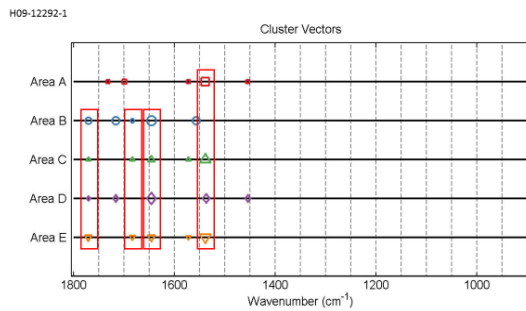
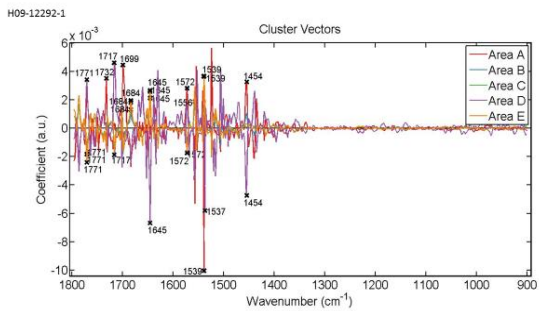
**Figure 3.105:** Loadings plots showing wavenumbers that discriminate stromal cells from between glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots of non-diseased (Fig. 3.106) and diseased (Fig. 3.107) tissue samples were presented and analysed in the same approach as before in the hope of identifying possible biomarkers for stromal cells, based on the occurrence of wavenumbers and their expression levels in the glandular areas. Wavenumbers identified in loading plots were also observed in cluster vector plots of their corresponding tissue.

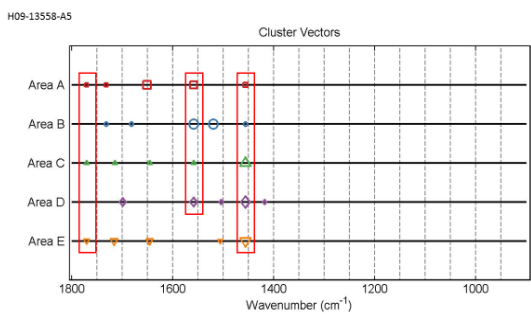
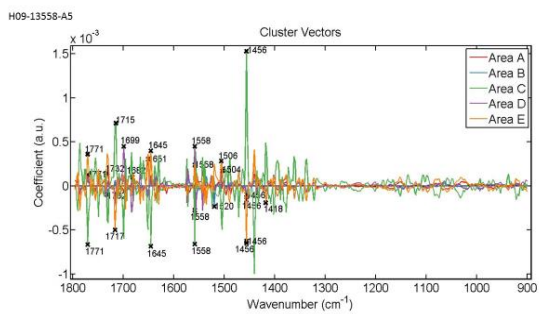
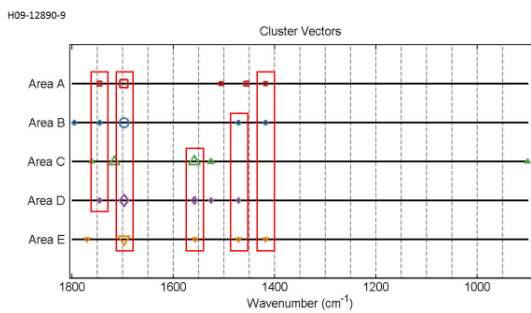
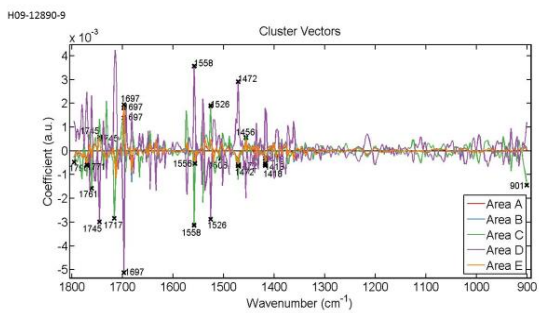
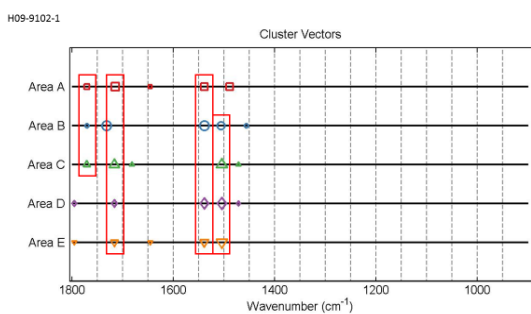
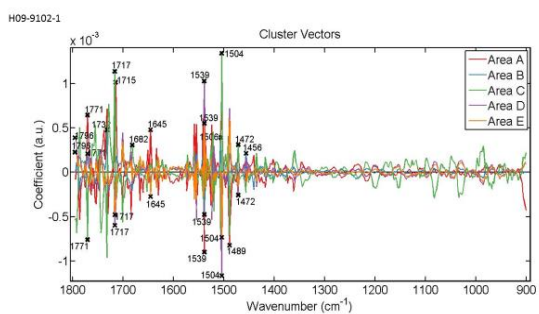
In sample H09-9102-6 three areas were common 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base) with 1715  $\text{cm}^{-1}$  and 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$  whilst all five areas were common for 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-12890-11, 1715  $\text{cm}^{-1}$  was commonly shared by all five areas, four areas were common for 1651  $\text{cm}^{-1}$  and three areas were common for 1541  $\text{cm}^{-1}$  (amide II) and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-13558-A8 wavenumbers observed to occur between glandular areas were 1796  $\text{cm}^{-1}$ , 1697  $\text{cm}^{-1}$  with 1695  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$ , 1520  $\text{cm}^{-1}$  (amide II) and 1456  $\text{cm}^{-1}$  (lipids and proteins). In sample H09-11708-4 three areas were common for 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1682  $\text{cm}^{-1}$  and 1504  $\text{cm}^{-1}$  (phenyl rings) whilst four areas were common for 1418  $\text{cm}^{-1}$  (deformation C-H). In sample H09-12292-1 four areas were common for 1771  $\text{cm}^{-1}$ , 1645  $\text{cm}^{-1}$  (phenyl rings) and 1539  $\text{cm}^{-1}$  (amide II) with 1537  $\text{cm}^{-1}$  whilst three areas were common for 1684  $\text{cm}^{-1}$  (guanine/thymine). In sample H09-9102-1, 1771  $\text{cm}^{-1}$  and 1504  $\text{cm}^{-1}$  (phenyl rings) were commonly shared by three areas whilst 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base) with 1715  $\text{cm}^{-1}$ , and 1539  $\text{cm}^{-1}$  (amide II) were commonly shared by four areas. In sample H09-12890-9 common occurrence for 1745  $\text{cm}^{-1}$  (ring base/amide II), 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$ , 1472  $\text{cm}^{-1}$  and 1418  $\text{cm}^{-1}$  (deformation C-H) was displayed by three areas whilst 1697  $\text{cm}^{-1}$  was commonly shared by four areas. In sample H09-13558-A5, three, four and five areas were common for 1771  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins) respectively.







**Figure 3.106:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells, in five glandular areas in individual non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1). The red rectangles highlight the wavenumbers shared in common among glandular areas.

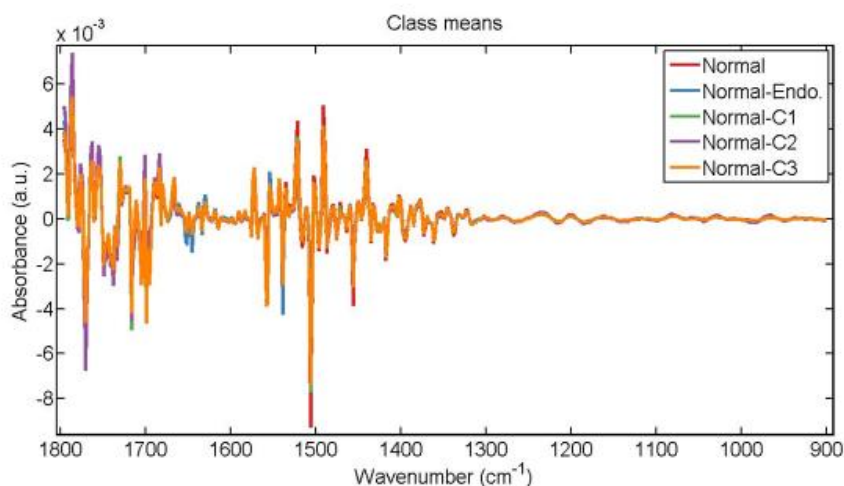


**Figure 3.107:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells, in five glandular areas in individual diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight the wavenumbers shared in common among glandular areas.

*Between all normal tissue samples*

All the non-diseased tissue samples were compared between them for the spectra interrogated from their stromal cells. Classes representing tissues were labelled in the same way as before and will be referred according to their class labels to associate plots with observations. All classes of tissues exhibited relatively the same mean FTIR spectra with most evident variations in peak intensities occurring at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1701\text{ cm}^{-1}$ ,  $\sim 1684\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$ ,  $\sim 1491\text{ cm}^{-1}$ ,  $\sim 1456\text{ cm}^{-1}$  and  $\sim 1441\text{ cm}^{-1}$  (Fig. 3.108). Classes Normal and Normal-C2 had mainly the highest peak intensities throughout the spectra. No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ .

Statistical analysis (Fig. 3.109) indicated that classes were significantly different from each other for the spectra acquired from stromal cells but Normal was not significantly different from Normal-C3.

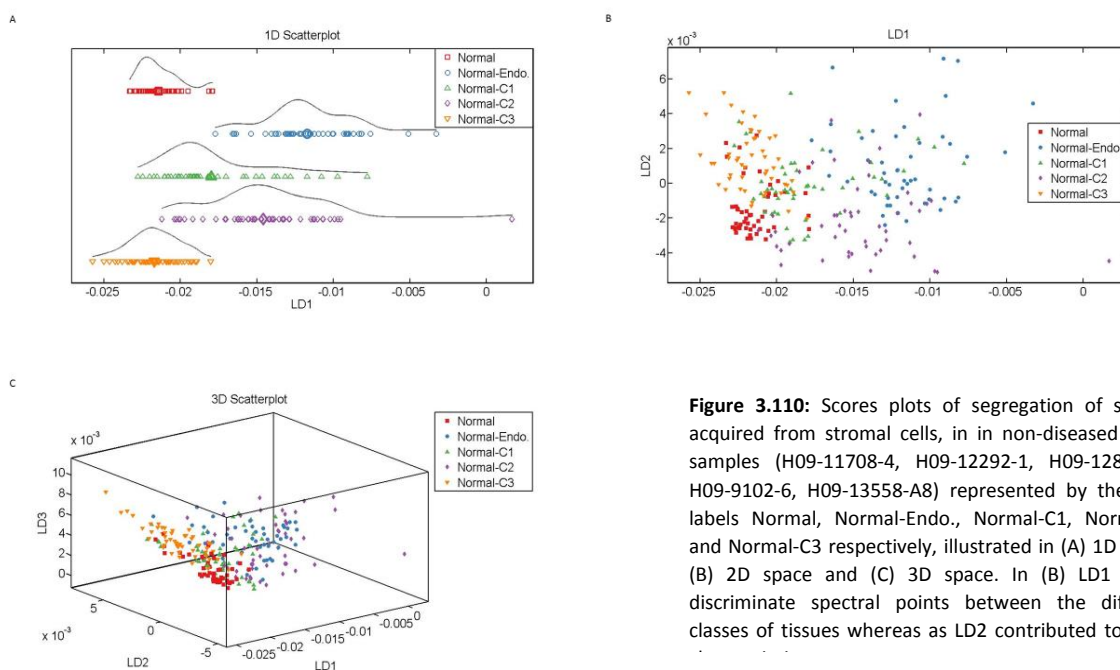


**Figure 3.108:** Mean FTIR spectra for comparison of spectra acquired from stromal cells in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Parameters	P value
LDA1 vs Normal	P < 0.001
LDA1 vs Normal-Endo.	P < 0.001
LDA1 vs Normal-C1	P > 0.05
LDA1 vs Normal-C2	P < 0.001
LDA1 vs Normal-C3	P < 0.001
Normal vs Normal-Endo.	P < 0.001
Normal vs Normal-C1	P < 0.001
Normal vs Normal-C2	P < 0.001
Normal vs Normal-C3	P > 0.05
Normal-Endo. vs Normal-C1	P < 0.001
Normal-Endo. vs Normal-C2	P < 0.01
Normal-Endo. vs Normal-C3	P < 0.001
Normal-C1 vs Normal-C2	P < 0.001
Normal-C1 vs Normal-C3	P < 0.001
Normal-C2 vs Normal-C3	P < 0.001

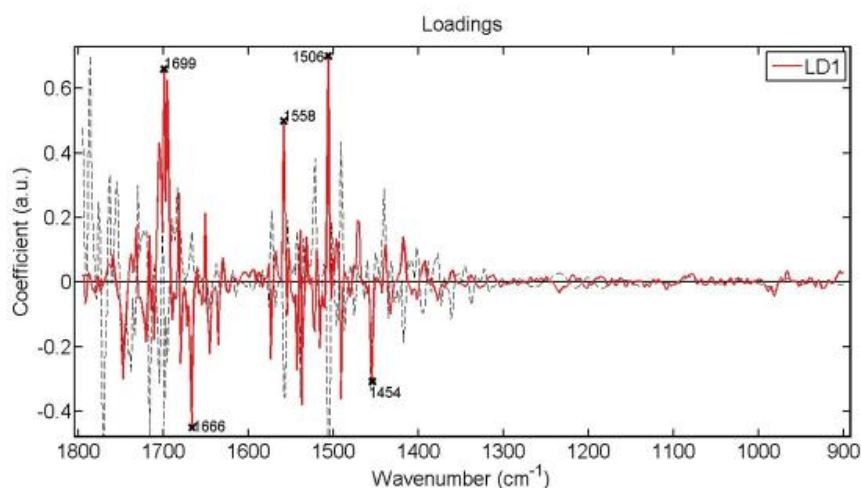
**Figure 3.109:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal cells in all five glandular elements between non-diseased tissue sections H09-11708-4, H09-9102-1, H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D, 2D and 3D space (Fig. 3.110). All classes would overlap with each other. Spectral points of Normal had the most compact arrangement which signified the least intra-class variations whereas spectral points of Normal-Endo., Normal-C1 and Normal-C2 exhibited a spread arrangement in the dimensional space which signified more intra-class variation.



**Figure 3.110:** Scores plots of segregation of spectra acquired from stromal cells, in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively, illustrated in (A) 1D space, (B) 2D space and (C) 3D space. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-

The discriminant wavenumbers identified in loading plots accounting for variations between the samples were 1699  $\text{cm}^{-1}$  (guanine/thymine), 1666  $\text{cm}^{-1}$  (amide I), 1558  $\text{cm}^{-1}$ , 1506  $\text{cm}^{-1}$  and 1454  $\text{cm}^{-1}$  (asymmetric methyl deformation) (Fig. 3.111).



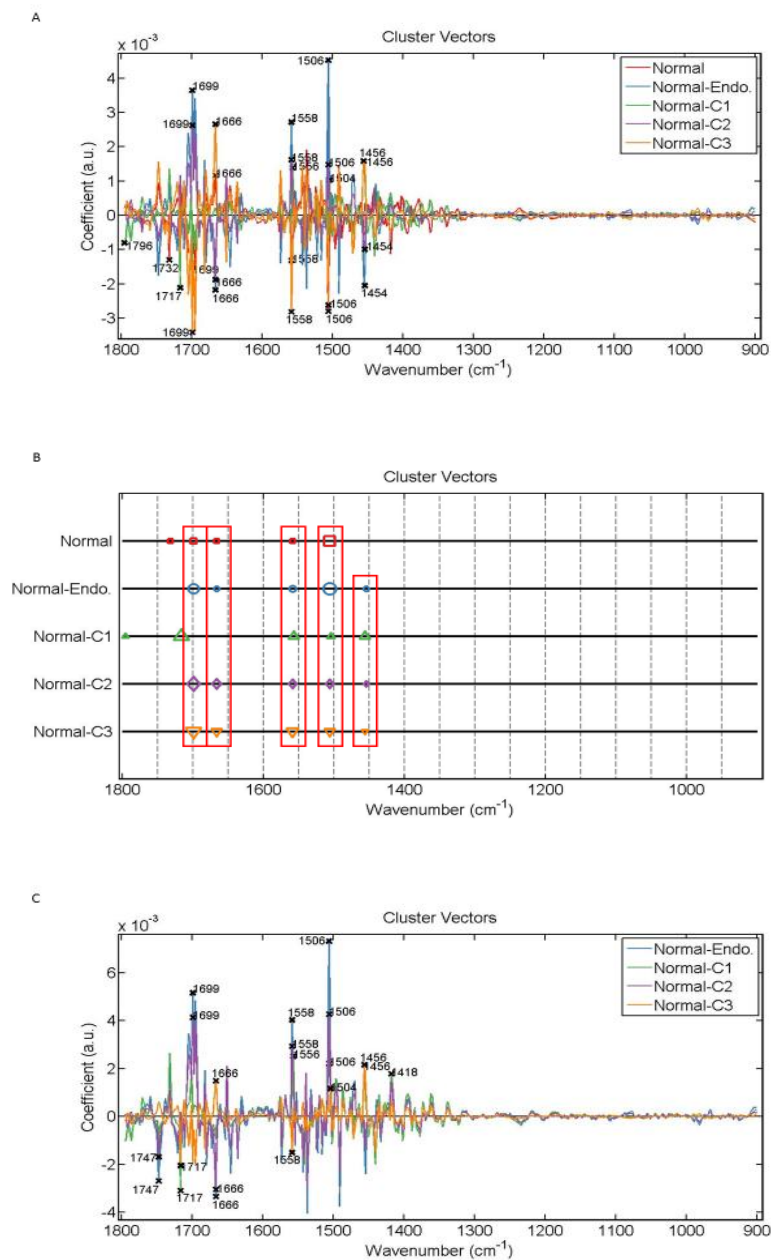
**Figure 3.111:** Loadings plots showing wavenumbers that discriminate stromal cells in non-diseased tissue sections (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

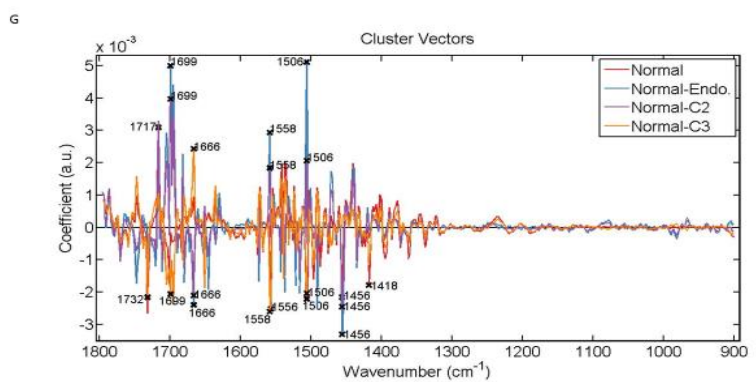
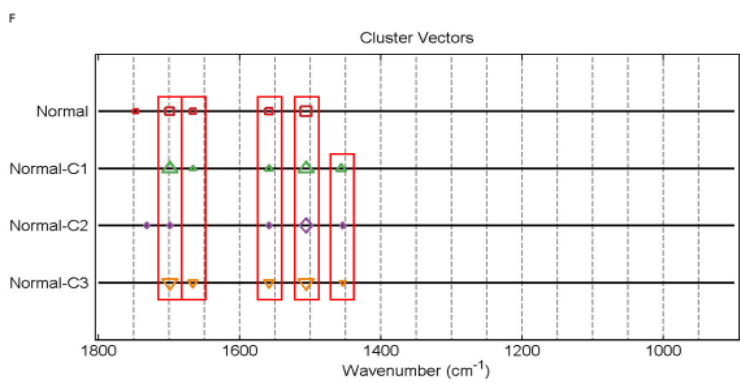
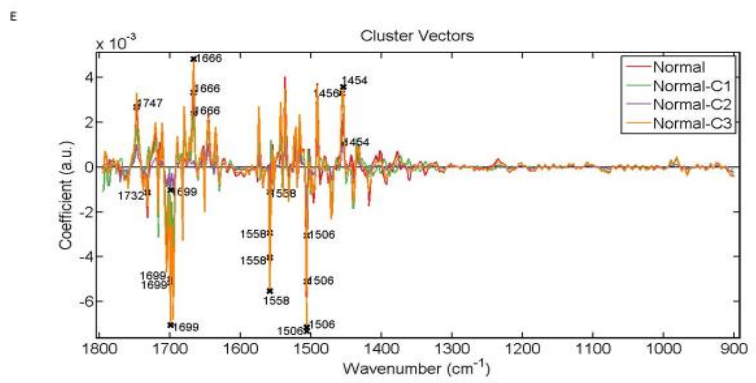
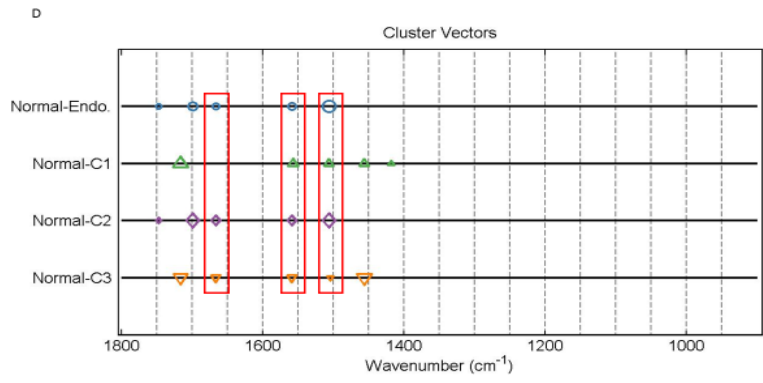
These wavenumbers, as well as new wavenumbers, were also identified in cluster vectors (Fig. 3.112).

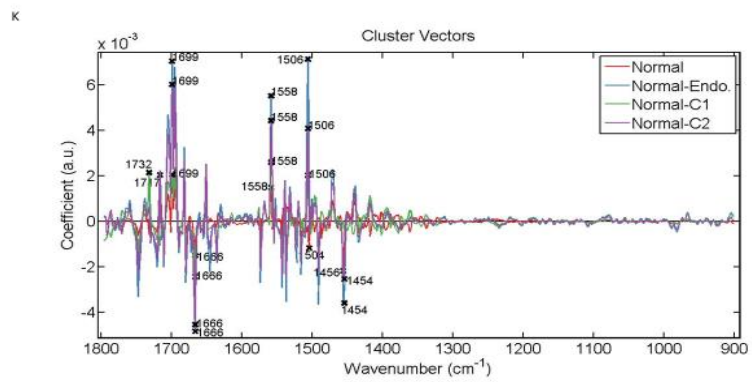
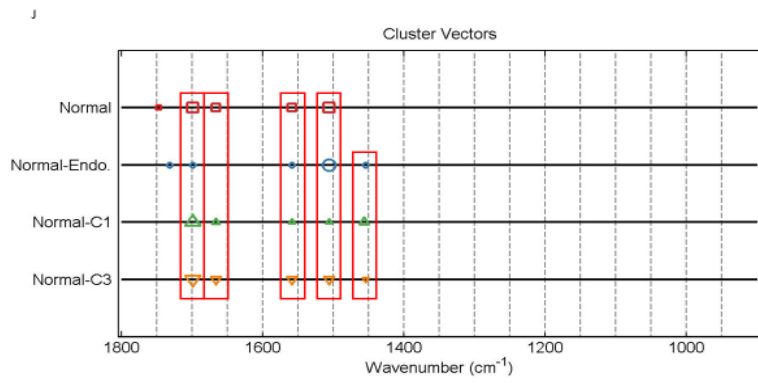
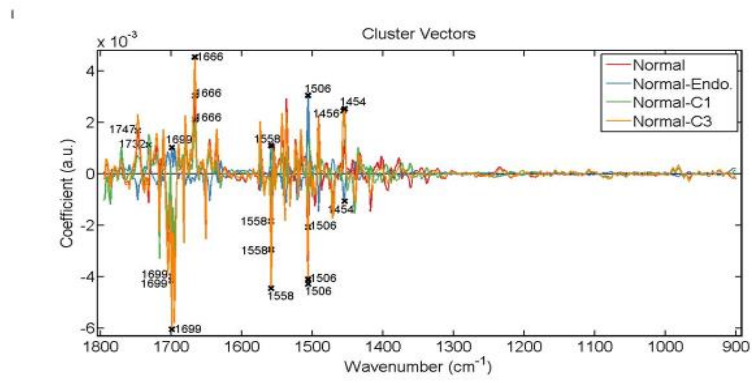
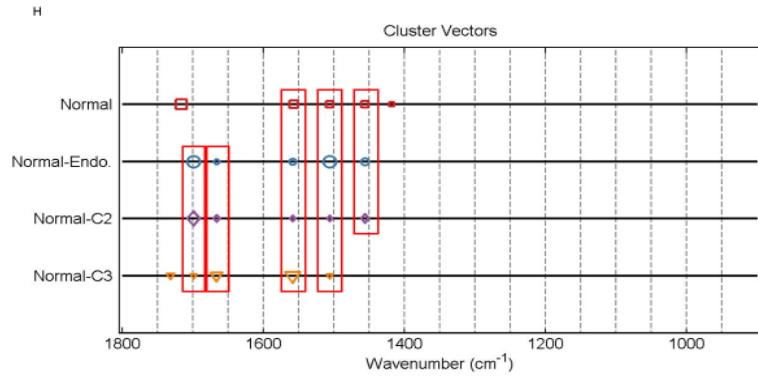
Common occurrence of wavenumbers and their expression levels amongst classes was assessed in the same approach as before in order to identify wavenumbers displaying biomarker potential for stromal cells in normal endometrial tissues.

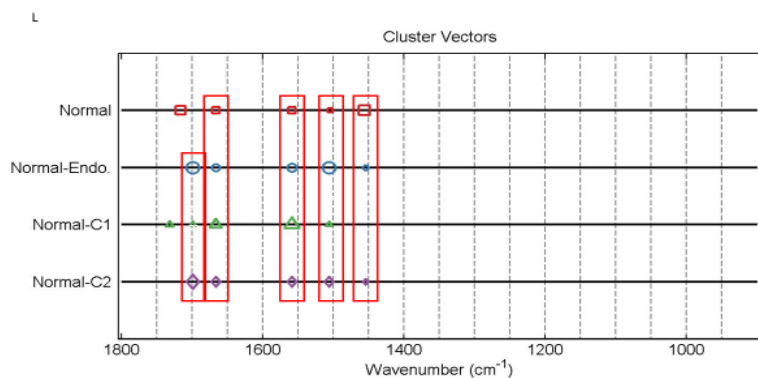
The wavenumbers observed to have common occurrence amongst classes of tissues in all the plots were 1666  $\text{cm}^{-1}$  (amide I), 1558  $\text{cm}^{-1}$  and 1506  $\text{cm}^{-1}$ . Common occurrence of 1699  $\text{cm}^{-1}$  (guanine/thymine) and 1456  $\text{cm}^{-1}$  (lipid and proteins) was not observed only when class Normal was used as a reference origin. The wavenumber 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$  were commonly shared when no class was used as a reference and when classes Normal and Normal-C1 were the reference origin. Also common occurrence of 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings) was observed three times; when no class was used as a reference and when classes Normal and Normal-C3 were set as the reference

origin. The wavenumber 1456 cm<sup>-1</sup> (lipids and proteins) displayed common occurrence with 1454 cm<sup>-1</sup> (asymmetric methyl deformation) in all plots where 1456 cm<sup>-1</sup> was identified, except when class Normal-C1 was the reference origin whereby 1454 cm<sup>-1</sup> (asymmetric methyl deformation) was not identified.









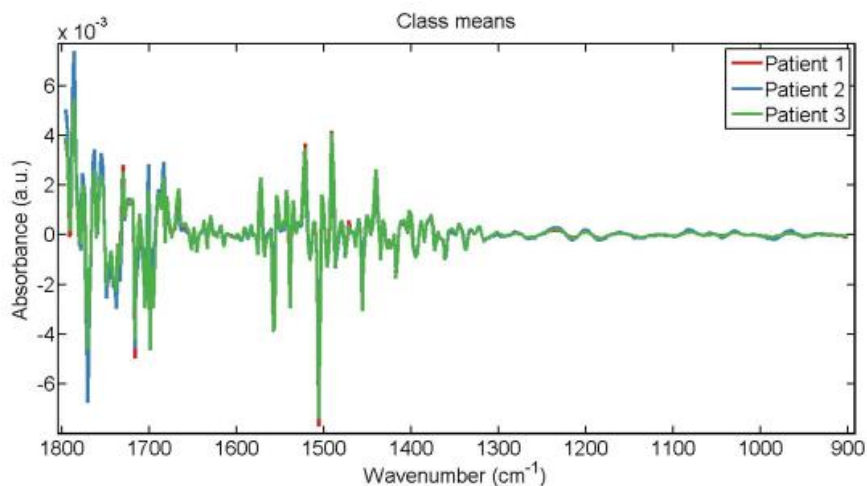
**Figure 3.112:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells, amongst non-diseased tissue sections. (A) & (B) no sample used as a reference, (C) & (D) Normal (H09-11708-4) reference class, (E) & (F) Normal-Endo. (H09-12292-1) reference class, (G) & (H) Normal-C1 (H09-12890-11) reference class, (I) & (J) Normal-C2 (H09-9102-6) reference class and (K) & (L) Normal-C3 (H09-13558-A8) reference class. The red rectangles in the cluster vector peak location plots highlight the wavenumbers occurring commonly in different tissue samples.

*Between non-diseased samples from patients with endometrial cancer*

Spectra from stromal cells in only non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer were compared. Classes representing the samples were labelled as before and will be referred accordingly to associate data analysis with what is shown on the plots. All classes had relatively the same shape for their FTIR mean spectra with most evident variations in absorbance intensities occurring at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1760\text{ cm}^{-1}$ ,  $\sim 1717\text{ cm}^{-1}$ ,  $\sim 1701\text{ cm}^{-1}$  and  $\sim 1684\text{ cm}^{-1}$ . Highest peaks were mainly exhibited by Patient 2. No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$  (Fig. 3.113).

Statistical analysis indicated that all tissues were significantly different from each other (Fig. 3.114).



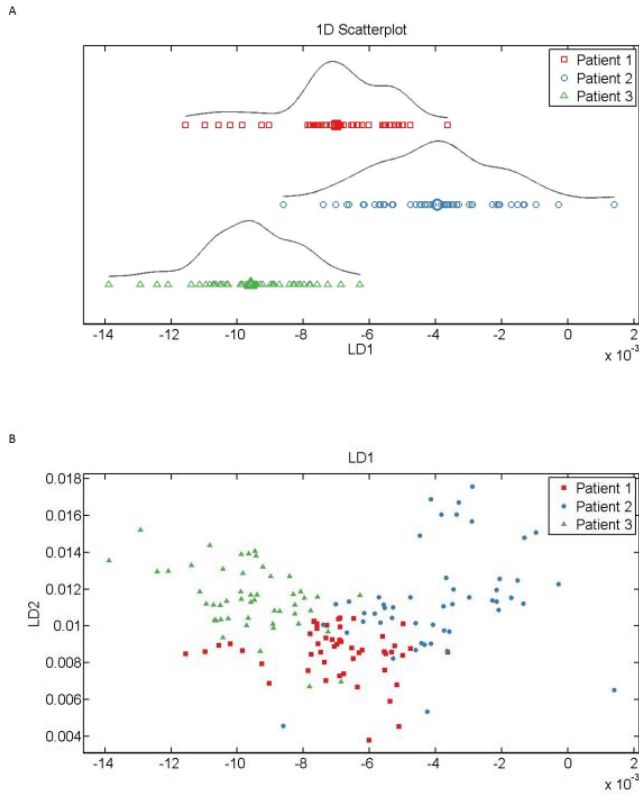


**Figure 3.113:** Mean FTIR spectra for comparison of spectra acquired from stromal cells in non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively.

Parameters	P value
LDA1 vs Patient 1	P > 0.05
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P < 0.001
Patient 2 vs Patient 3	P < 0.001

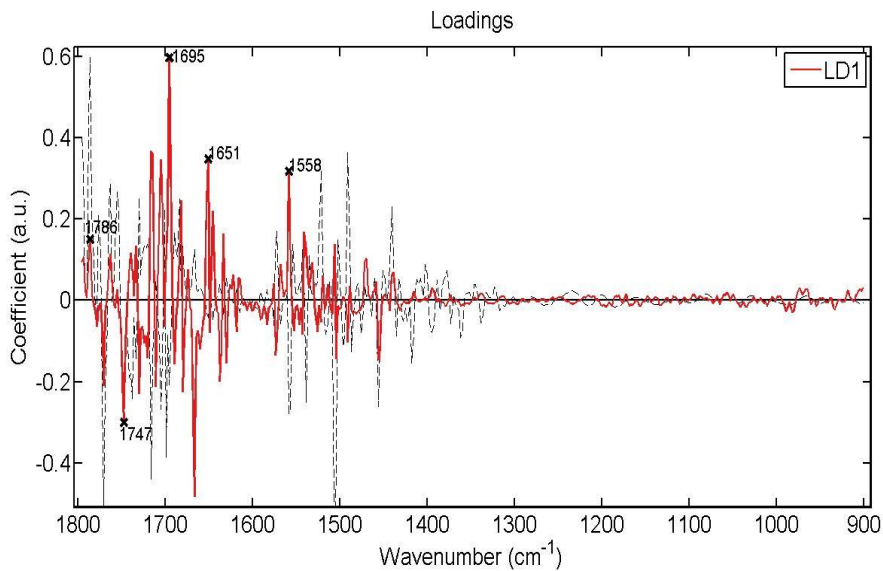
**Figure 3.114:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal cells between non-diseased tissue sections H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D and 2D space (Fig. 3.115). Class labelled as Patient 1 overlapped with the other two classes throughout all of its spectra. Almost half of the spectra of Patient 2 and Patient 3 overlapped with each other and with Patient 1. Spectral points of Patient 1 had a relatively more compact arrangement and thus less intra-class variation.



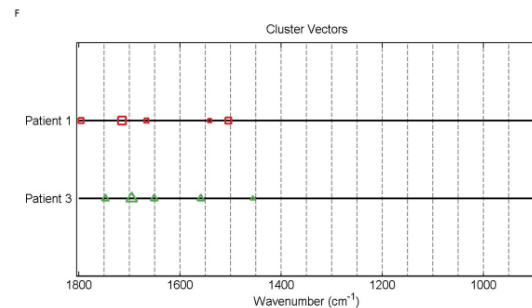
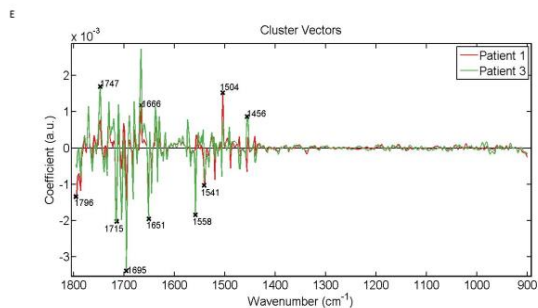
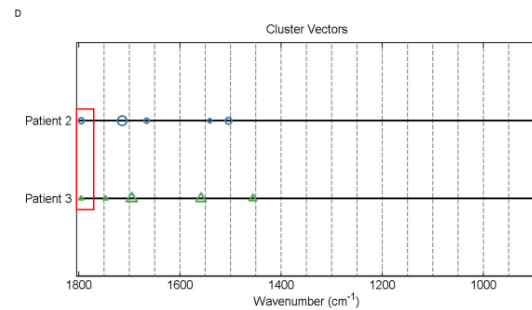
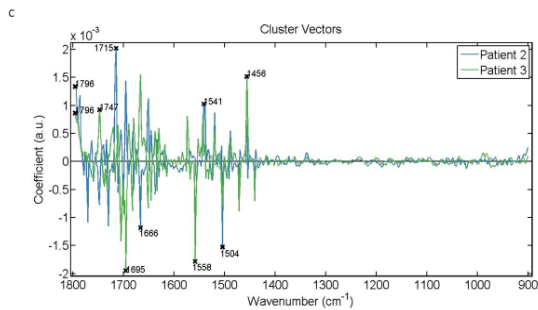
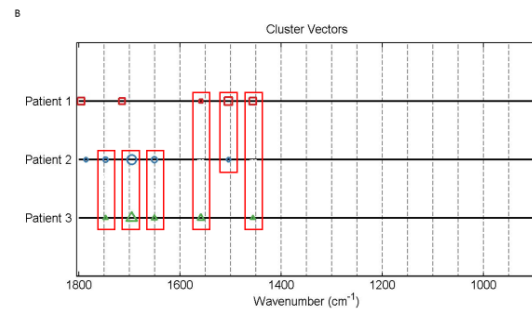
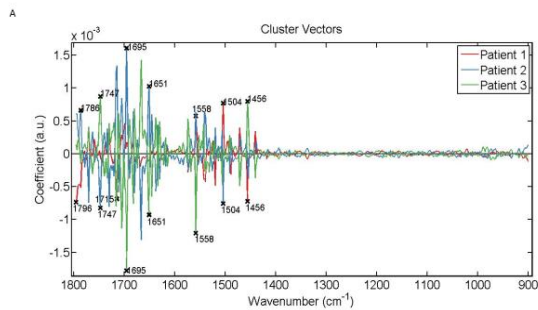
**Figure 3.115:** Scores plots for segregation of spectra in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

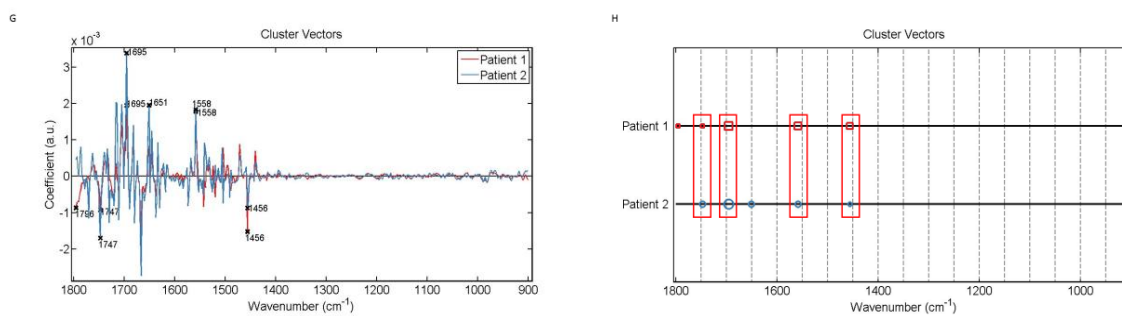
The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were  $1786\text{ cm}^{-1}$ ,  $1747\text{ cm}^{-1}$ ,  $1695\text{ cm}^{-1}$ ,  $1651\text{ cm}^{-1}$  and  $1558\text{ cm}^{-1}$  (Fig. 3.116).



**Figure 3.116:** Loadings plots showing wavenumbers that discriminate stromal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

These wavenumbers along with other wavenumbers were observed in cluster vectors plots (Fig. 3.117). Plots were presented and analysed like previously to identify potential biomarkers for stromal cells. Having no class as a reference origin the wavenumbers displaying common occurrence amongst tissue samples were 1747  $\text{cm}^{-1}$ , 1695  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$ , 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). Using Patient 1 as reference the only wavenumber observed to have common occurrence in classes was 1796  $\text{cm}^{-1}$ . Classes were not found to be common for any wavenumber when Patient 2 was set as the reference origin. Using Patient 3 as reference common occurrence was observed for 1747  $\text{cm}^{-1}$ , 1695  $\text{cm}^{-1}$ , 1558  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins).



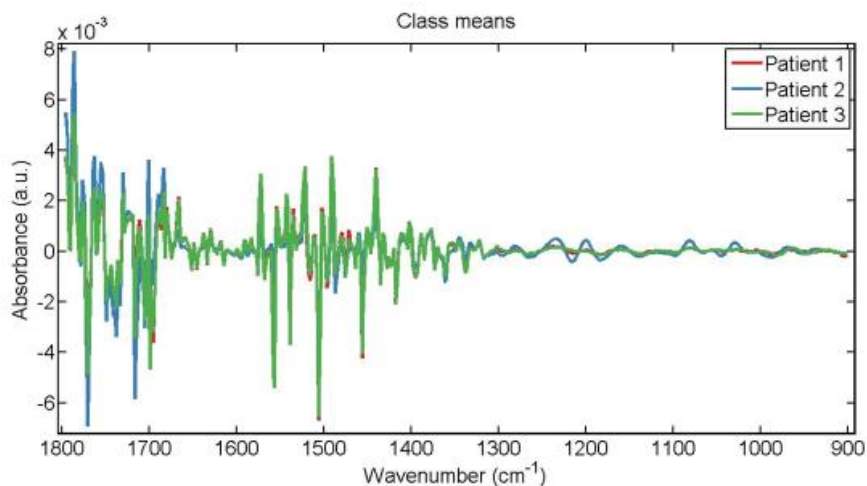


**Figure 3.117:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells among non-diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-11) reference class, (E) & (F) Patient 2 (H09-9102-6) reference class and (G) & (H) Patient 3 (H09-13558-A8) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Between diseased samples from patients with endometrial cancer*

Spectra acquired from stromal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) taken from cancerous lesions in the endometrium were analysed and compared. Classes representing the samples were labelled as before so they will be referred accordingly to associate data analysis with what is shown on the plots. Relatively the same shape in mean FTIR spectra was exhibited by all classes and most pronounced variations in absorbance intensities occurred at  $\sim 1786 \text{ cm}^{-1}$ ,  $\sim 1771 \text{ cm}^{-1}$ ,  $\sim 1755 \text{ cm}^{-1}$ ,  $\sim 1735 \text{ cm}^{-1}$ ,  $\sim 1717 \text{ cm}^{-1}$ ,  $\sim 1701 \text{ cm}^{-1}$ ,  $\sim 1684 \text{ cm}^{-1}$  and  $\sim 1487 \text{ cm}^{-1}$  (Fig. 3.118). Highest peak intensities were mainly observed in Patient 3. No peaks were detected in the spectral region  $900\text{-}1300 \text{ cm}^{-1}$ .

Statistical analysis indicated that Patient 2 was significantly different from Patient 3 and Patient 1 but the two latter were not found to be significantly different (Fig. 3.119).

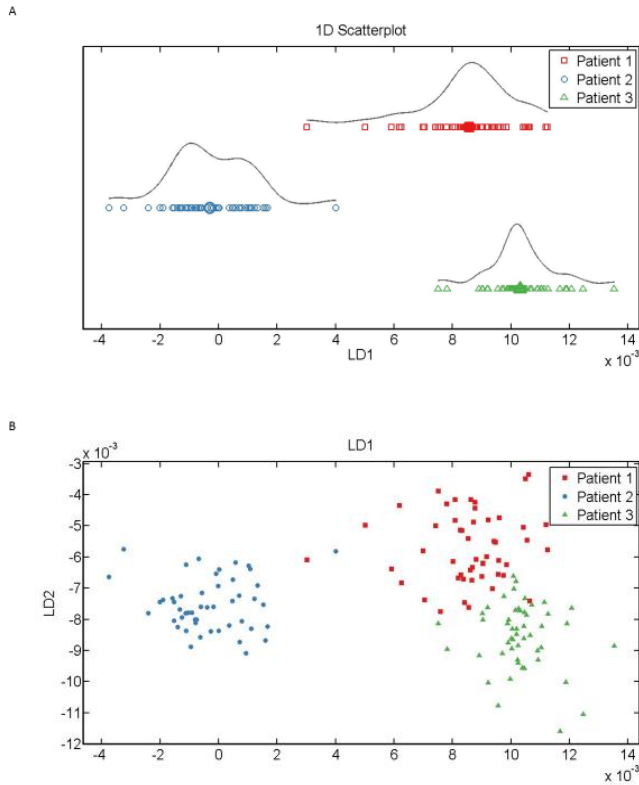


**Figure 3.118:** Mean FTIR spectra for comparison of spectra from stromal cells in diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.001
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P > 0.05
Patient 2 vs Patient 3	P < 0.001

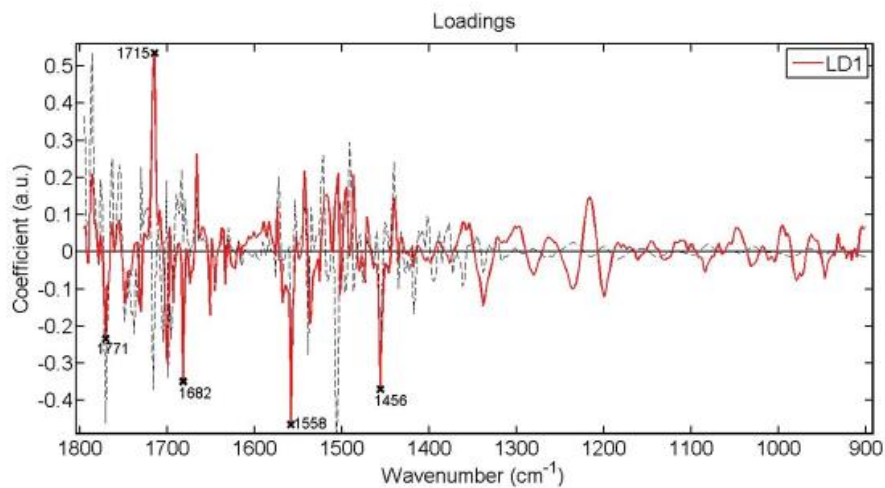
**Figure 3.119:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal cells between diseased tissue sections H09-9102-1, H09-12890-9, H09-13558-A5 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Segregation of classes resulting from application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.120). Spectra of Patient 1 overlapped with both Patient 2 and Patient 3, whereas the latter two classes did not overlap with each other. Spectral points of Patient 2 had a relatively more compact arrangement in the dimensional space which signified the least intra-class variation.



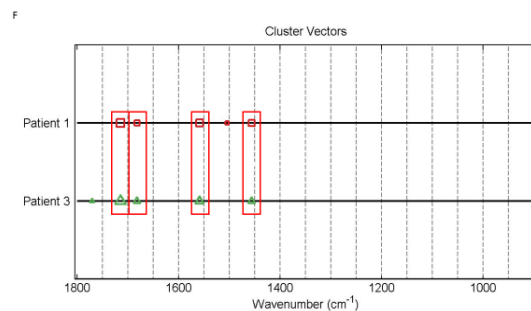
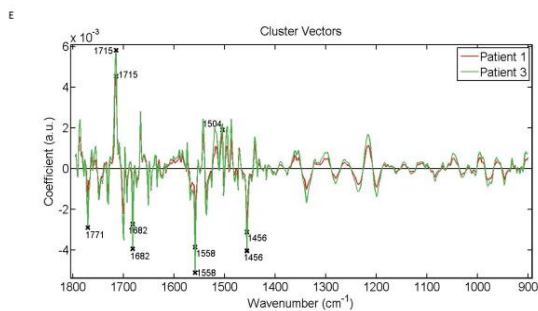
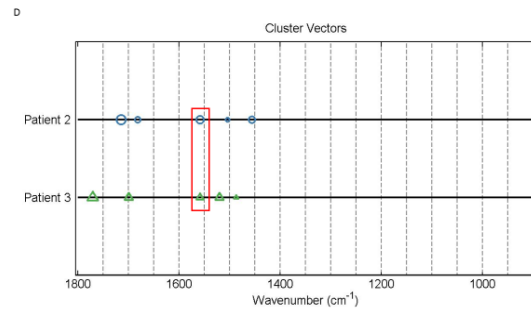
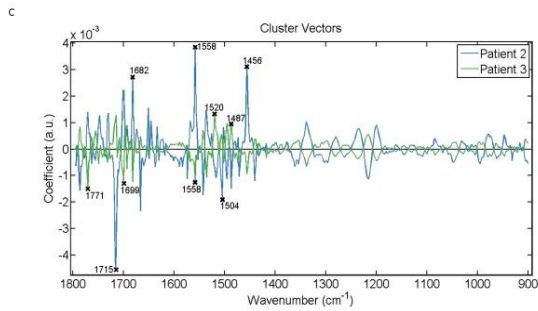
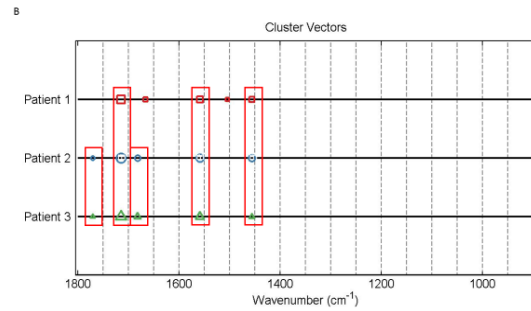
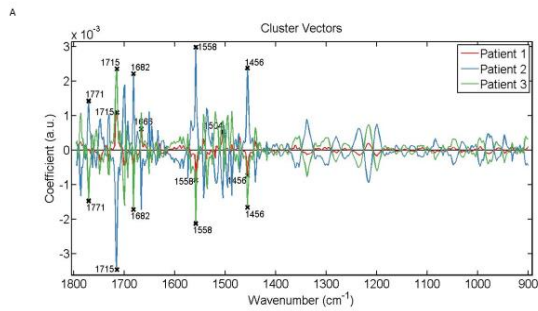
**Figure 3.120:** Scores plots for segregation of spectra acquired from stromal cells in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

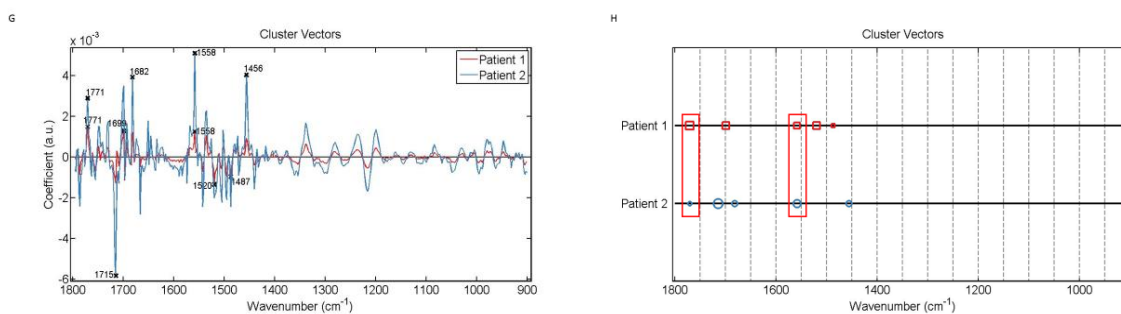
Generation of loading plots identified five major wavenumbers which would discriminate classes of tissues (Fig. 3.121). The identified wavenumbers were 1771 cm<sup>-1</sup>, 1715 cm<sup>-1</sup>, 1682 cm<sup>-1</sup>, 1558 cm<sup>-1</sup> and 1456 cm<sup>-1</sup> (lipids and proteins).



**Figure 3.121:** Loadings plots showing wavenumbers that discriminate stromal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Classes of tissue samples displayed common occurrence for the above mentioned wavenumbers as illustrated in cluster vector plots (Fig. 3.122) Plots were presented and analysed as before to observe occurrence and expression levels of wavenumbers amongst tissue samples and identify potential biomarkers. All the wavenumbers displayed common occurrence amongst the classes when no class was set as the reference origin. The only wavenumbers to display common occurrence when Patient 1 was the reference class was 1558  $\text{cm}^{-1}$ . When Patient 2 was the class reference all the wavenumbers exhibited common occurrence except 1771  $\text{cm}^{-1}$ . Having Patient 3 as reference, common occurrence was displayed only by 1771  $\text{cm}^{-1}$  and 1558  $\text{cm}^{-1}$ .



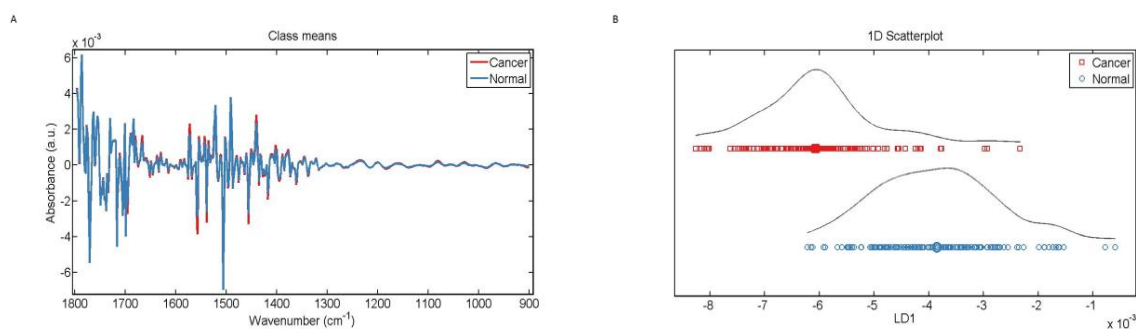


**Figure 3.122:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells among diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-9) reference class, (E) & (F) Patient 2 (H09-9102-1) reference class and (G) & (H) Patient 3 (H09-13558-A5) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Diseased tissues Vs Non-Diseased tissues*

The spectra of the stromal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) were compared with the spectra of stromal cells in their corresponding non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer and tissue samples were represented by the classes 'Cancer' and 'Normal' respectively. Classes of tissues will be referred according to their labels to associate data analysis with what is shown on the plots. Relatively the same shape in mean FTIR spectra was exhibited by both classes and most evident variations in peak intensities were identified at  $\sim 1572 \text{ cm}^{-1}$ ,  $\sim 1556 \text{ cm}^{-1}$ ,  $\sim 1539 \text{ cm}^{-1}$ ,  $\sim 1456 \text{ cm}^{-1}$ ,  $\sim 1441 \text{ cm}^{-1}$  and  $\sim 1418 \text{ cm}^{-1}$  (Fig. 3.123 A). Highest peak intensities were mainly observed in Normal. No peaks were detected in the spectral region  $900\text{-}1300 \text{ cm}^{-1}$ . Application of PCA-LDA resulted in segregation of classes which was visualised in scores plot of 1D space whereby almost the whole spectra of Normal would overlap with almost half of the spectra of Cancer (Fig. 3.123 B). Statistical analysis indicated that classes were not significantly different from each other (Fig. 3.124).



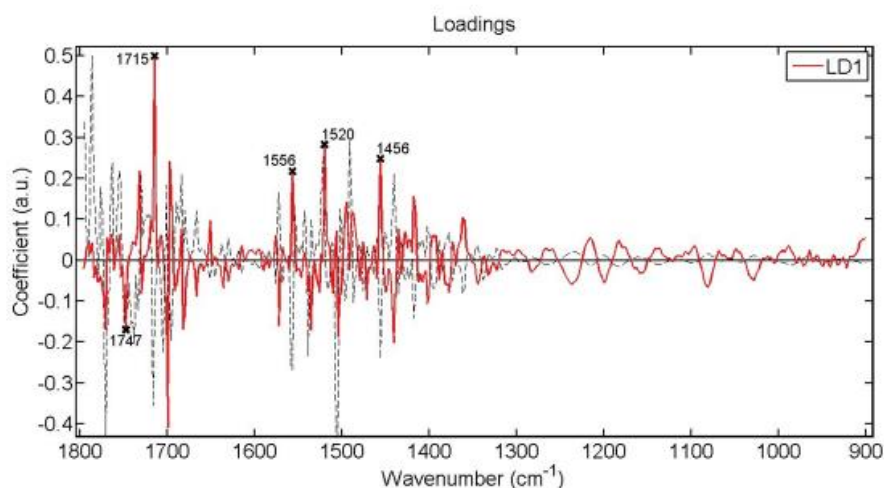


**Figure 3.123:** (A) Mean FTIR spectra for comparison of spectra from stromal cells in all diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5) and their corresponding non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). (B) Scores plots for segregation of spectra in 1D space. Diseased and non-diseased tissue samples are represented by the class labels 'Cancer' (red) and 'Normal' (blue) respectively.

Parameters	P Value
Cancer Vs Normal	$P < 0.0001$

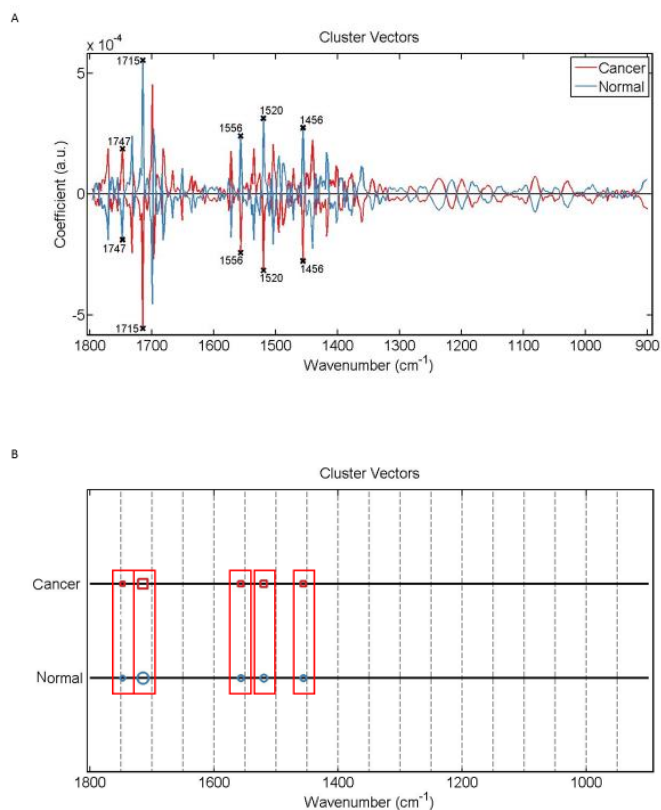
**Figure 3.124:** Obtained p-value by employment of Unpaired t-test to compare spectra acquired from stromal cells between non diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and their corresponding diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections represented by the class labels Normal and Cancer respectively.

The five major wavenumbers identified in loading responsible for discriminating the classes of tissue samples were  $1747 \text{ cm}^{-1}$ ,  $1715 \text{ cm}^{-1}$ ,  $1556 \text{ cm}^{-1}$ ,  $1520 \text{ cm}^{-1}$  (amide II) and  $1456 \text{ cm}^{-1}$  (lipids and proteins) (Fig. 3.125).



**Figure 3.125:** Loadings plots showing wavenumbers that discriminate spectra from stromal cells in all diseased (H09-9102-1, H09-12890-9, H09-13558-A5) and their corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) tissue sections. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

These wavenumbers were also observed in cluster vector plots which were presented and analysed in the same approach as previously (Fig. 3.126). The wavenumbers displayed common occurrence in both classes and exhibited same expression levels amongst classes so none of the wavenumbers was identified to have a potential as a biomarker to specify either 'Normal' or 'Cancer' stromal cells though highest expression levels were exhibited by 1715  $\text{cm}^{-1}$  whereas 1747  $\text{cm}^{-1}$  exhibited the lowest levels.



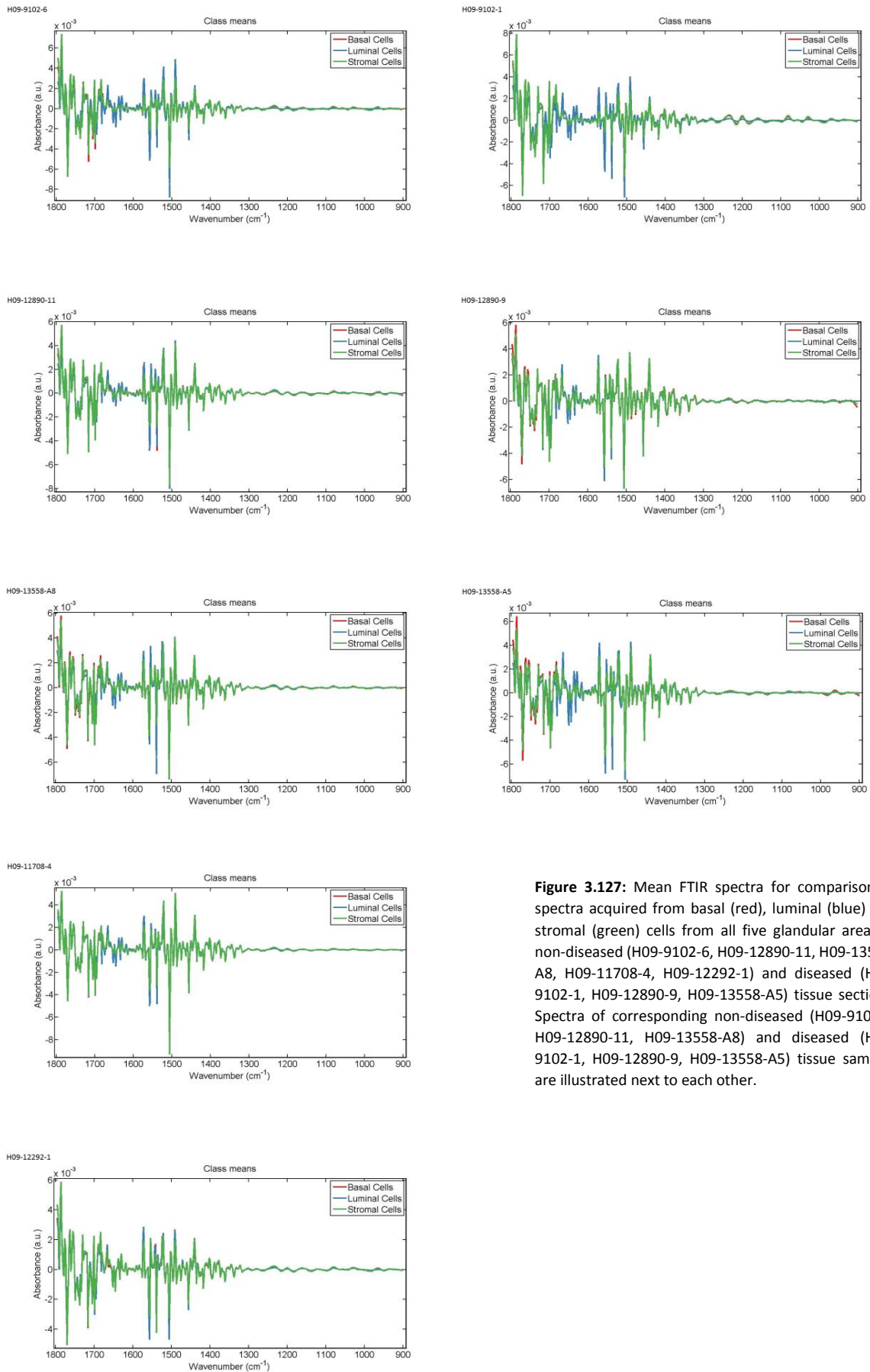
**Figure 3.126:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells among all corresponding diseased (H09-12890-9, H09-9102-1, H09-13558-A5) and non-diseased (H09-12890-11, H09-9102-6, H09-13558-A8) tissue sections. The red rectangles highlight the wavenumbers occurring commonly in

### 3.2.4 Basal Cells Vs Luminal Cells Vs Stromal Cells

#### *In individual tissue samples*

The biochemical composition of basal, luminal and stromal cells from all five glandular elements in individual tissue sections were analysed and compared. Figure 3.127 shows the mean FTIR spectra of all the classes of cells acquired from all five glandular elements in each tissue section. All three classes of cells in all tissue samples exhibited relatively the same shape in mean FTIR spectra whereby most evident variations in absorbance intensities occurred at  $\sim 1682\text{ cm}^{-1}$ ,  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1640\text{ cm}^{-1}$ ,  $\sim 1572\text{ cm}^{-1}$ ,  $\sim 1556\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$ ,  $\sim 1491\text{ cm}^{-1}$  and  $\sim 1456\text{ cm}^{-1}$ . No specific observations were made that could correlate in any way non-diseased tissue samples with their corresponding diseased tissue. No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ .

Classes of cells were significantly different from each other except in tissue sample H09-9102-1 whereby the obtained P value suggested no significant difference between basal and stromal cells (Fig. 3.128).



**Figure 3.127:** Mean FTIR spectra for comparison of spectra acquired from basal (red), luminal (blue) and stromal (green) cells from all five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.

H09-9102-6		H09-12292-1	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P > 0.05
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P < 0.01
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

H09-12890-11		H09-12890-9	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P > 0.05
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.01
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P < 0.001
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

H09-13558-A8		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P < 0.05
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P < 0.05
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

H09-11708-4		H09-12292-1	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P > 0.05
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P < 0.01
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

**Figure 3.128:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal, luminal and stromal cells from five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Figure 3.129 and figure 3.130 illustrate scores plots of non-diseased and diseased tissue sections respectively. In sample H09-9102-6, basal cells overlapped with both luminal and stromal cells. Luminal and stromal cells would overlap slightly with each other. None of the classes displayed to have neither more nor less intra-class variation based on arrangement of spectral points within the dimensional space.

In sample H09-12890-11 basal cells overlapped with both luminal and stromal cells whereas the two latter classes displayed less overlap between them. None of the classes displayed to have neither more nor less intra-class variation based on arrangement of spectral points within the dimensional space.

In sample H09-13558-A8, basal cells overlapped with luminal and stromal cells whilst little overlap was observed between luminal and stromal cells. Spectral points of stromal cells had a relatively more compact arrangement signifying less intra-class variation.

In sample H09-11708-4 basal cells overlapped with both luminal and stromal cells. Less overlapping was displayed between luminal and stromal cells. Spectral points of basal cells had a relatively more spread arrangement which signified more intra-class variation.

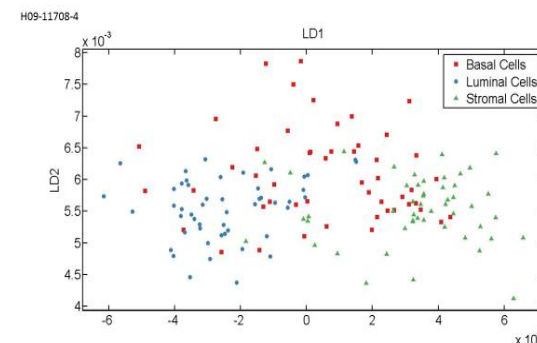
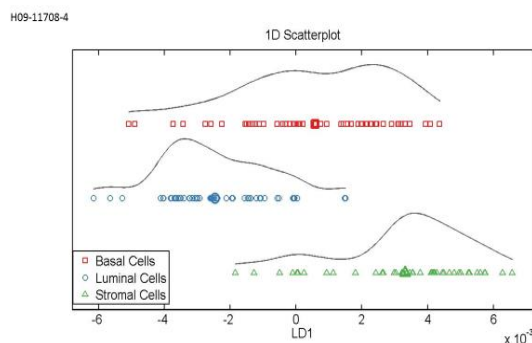
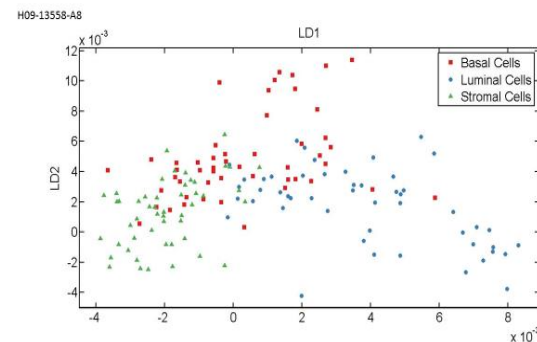
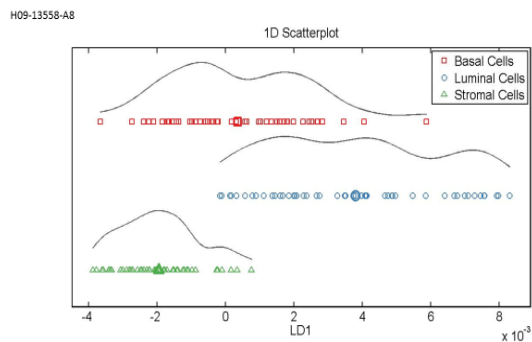
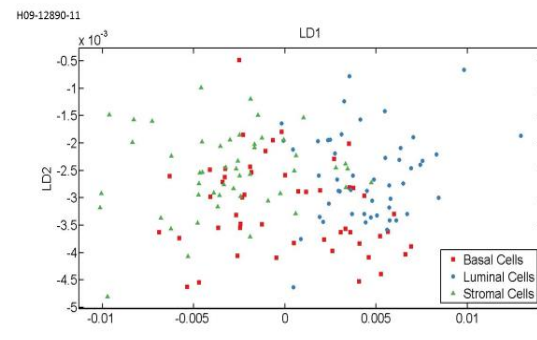
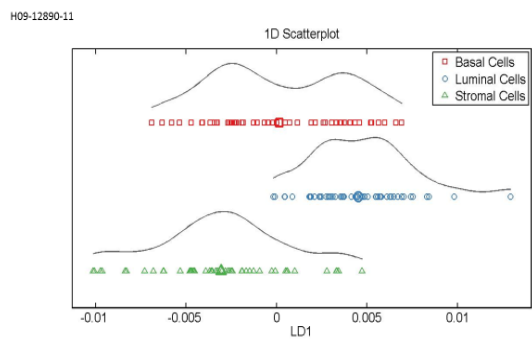
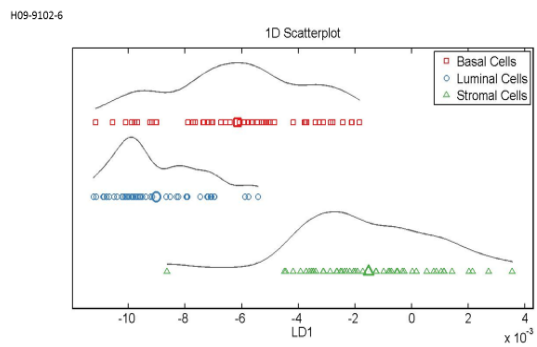
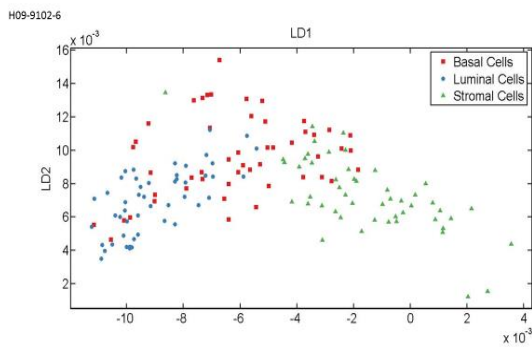
In sample H09-12292-1, degree of overlapping was relatively the same between all classes. Spectral point of basal cells displayed a more compact arrangement which signified less intra-class variation.

In sample H09-9102-1, most overlapping was observed between basal and stromal cells. Both classes displayed little overlap with luminal cells. Spectral point of luminal cells had a more spread arrangement within the dimensional space which signified the most intra-class variation.

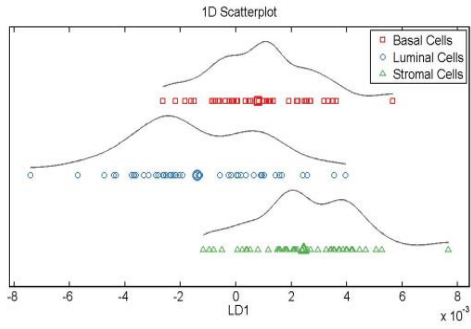
In sample H09-12890-9 all classes overlapped with each other to a certain degree. None of the classes displayed to have neither more nor less intra-class variation based on arrangement of spectral points within the dimensional space.

In sample H09-13558-A5 overlapping was mainly displayed between basal and stromal cells, both of which would slightly overlap with luminal cells. Spectral point of stromal cells had a relatively more

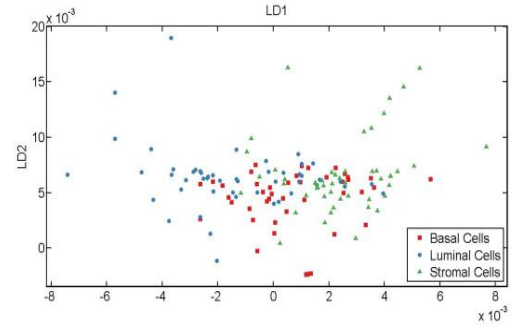
compact arrangement which signified less intra-class variation. No specific observations were made that would associate the corresponding non-diseased and diseased tissue sections.



H09-12292-1

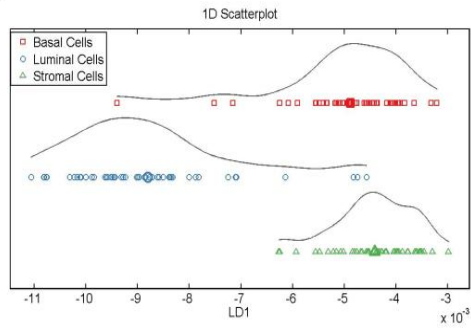


H09-12292-1

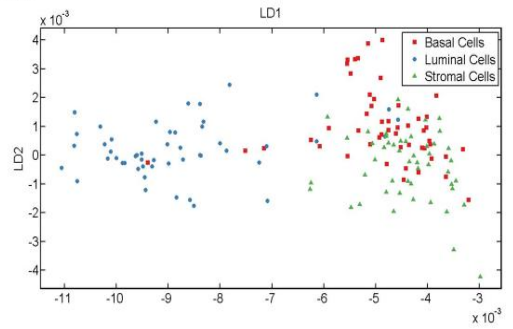


**Figure 3.129:** 1D and 2D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from basal (red), luminal (blue) and stromal (green) cells from all glandular elements in individual samples. In 2D plots LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

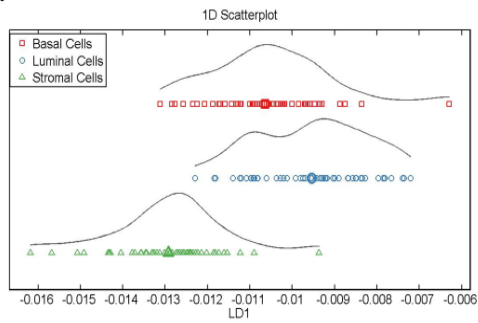
H09-9102-1



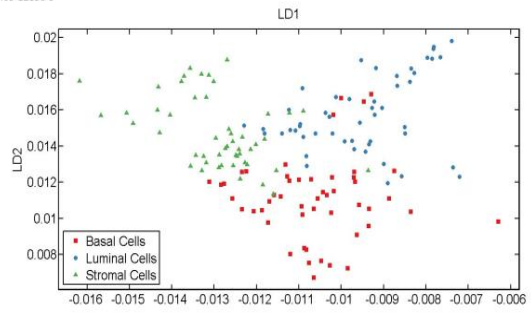
H09-9102-1



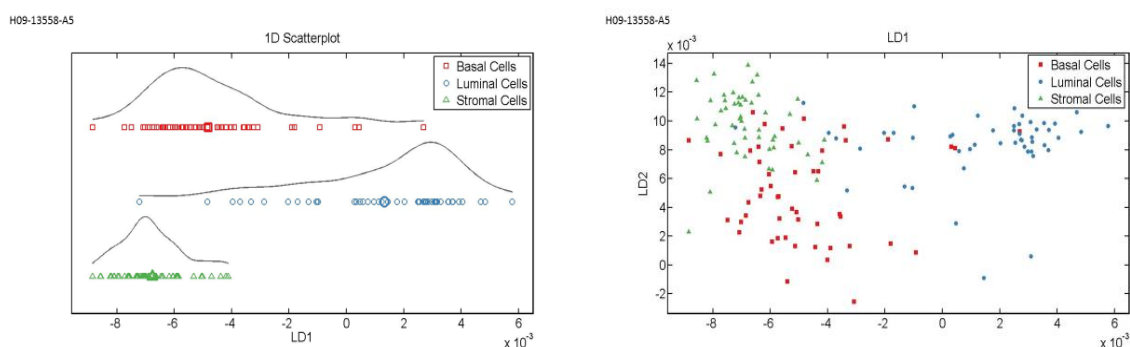
H09-12890-9



H09-12890-9

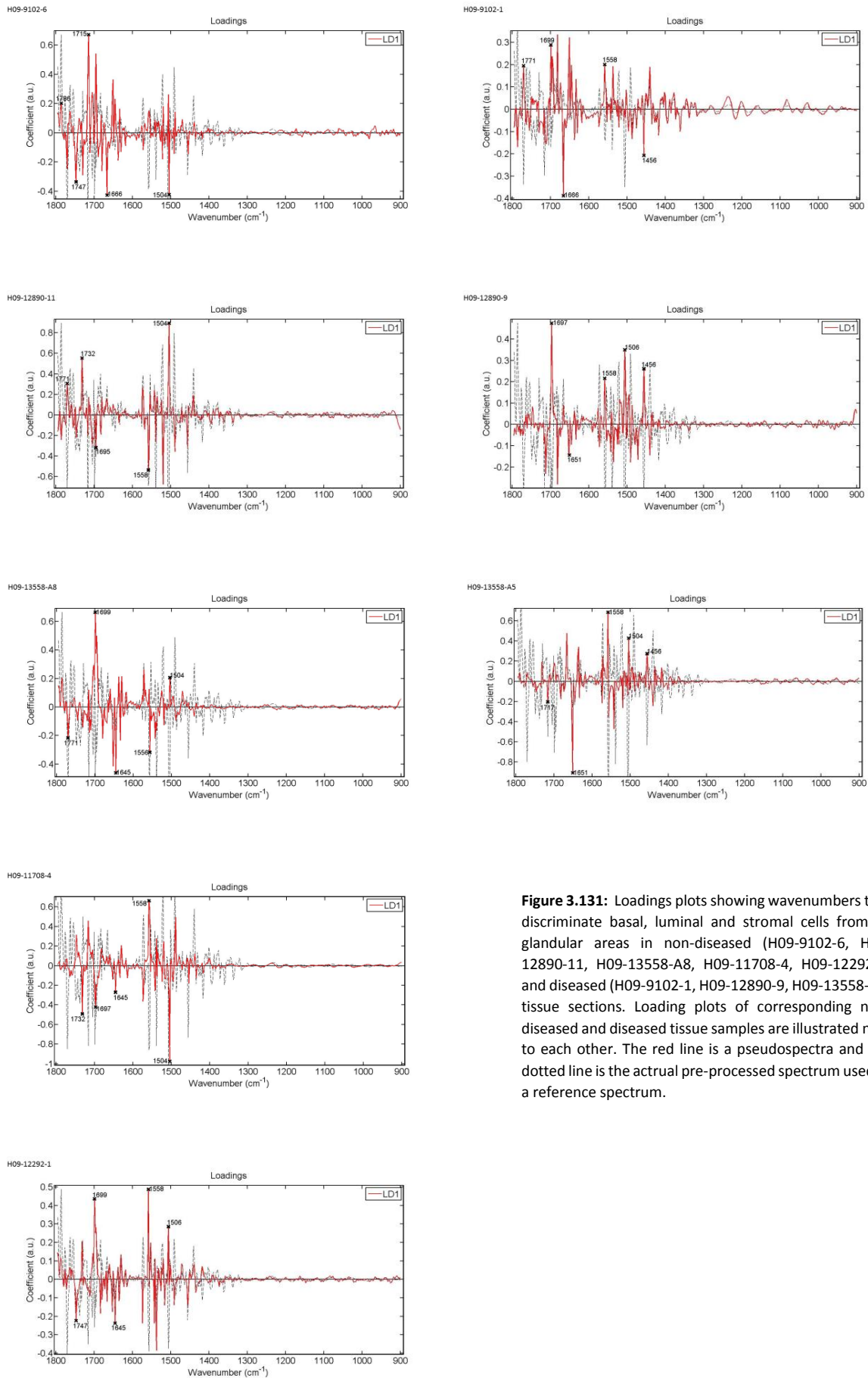






**Figure 3.130:** 1D and 2D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from basal (red), luminal (blue) and stromal (green) cells from all glandular elements in individual samples. In 2D plots LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

Generation of loading plots identified five major discriminant wavenumbers responsible for variations between classes of basal, luminal and stromal cells from all the interrogated glandular elements in individual samples (Fig. 3.131). The identified wavenumbers in sample H09-9102-6 were  $1786\text{ cm}^{-1}$ ,  $1747\text{ cm}^{-1}$ ,  $1715\text{ cm}^{-1}$ ,  $1666\text{ cm}^{-1}$  (amide I) and  $1504\text{ cm}^{-1}$  (phenyl rings). In H09-12890-11 the wavenumbers were  $1771\text{ cm}^{-1}$ ,  $1732\text{ cm}^{-1}$  (lipids),  $1695\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). In sample H09-13558-A8 the discriminant wavenumbers were  $1771\text{ cm}^{-1}$ ,  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1645\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). In H09-11708-4 the major wavenumbers were  $1732\text{ cm}^{-1}$  (lipids),  $1697\text{ cm}^{-1}$ ,  $1645\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). In H09-12292-1 the wavenumbers were  $1747\text{ cm}^{-1}$ ,  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1645\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). In the diseased sample H09-9102-1 the wavenumbers were  $1771\text{ cm}^{-1}$ ,  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1666\text{ cm}^{-1}$  (amide I),  $1558\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins). In sample H09-12890-9 the identified wavenumbers were  $1697\text{ cm}^{-1}$ ,  $1651\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ ,  $1506\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (proteins and lipids). In sample H09-13558-A5 the discriminant wavenumbers were  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA),  $1651\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ ,  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1456\text{ cm}^{-1}$  (proteins and lipids). The wavenumbers  $1558\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings) were observed to occur commonly in loading plots among tissue samples.



**Figure 3.131:** Loadings plots showing wavenumbers that discriminate basal, luminal and stromal cells from all glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots were analysed in an alternative way in an attempt to identify potential biomarkers for basal, luminal and stromal cells (Fig. 3.132). Observations were based on expression levels of wavenumbers in the different classes of cells and which wavenumbers and/or spectral regions were commonly shared among classes and which wavenumbers occurred only in one class. The expression levels of shared wavenumbers were the same among classes, unless stated differently.

In sample H09-9102-6 the wavenumbers shared in common between all classes of cells were  $1747\text{ cm}^{-1}$ ,  $1715\text{ cm}^{-1}$ ,  $1666\text{ cm}^{-1}$  (amide I) and  $1504\text{ cm}^{-1}$  (phenyl rings). Wavenumbers occurring only in basal, luminal and stromal cells were  $1541\text{ cm}^{-1}$  (amide II),  $1556\text{ cm}^{-1}$  and  $1786\text{ cm}^{-1}$  respectively.

In sample H09-12890-11 all three classes were common for  $1695\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). Luminal and stromal cells were common for  $1732\text{ cm}^{-1}$  (lipids) and the first were also common with basal cells for  $1456\text{ cm}^{-1}$  (lipids and proteins). Wavenumbers for which basal and stromal cells were unique were  $901\text{ cm}^{-1}$  and  $1771\text{ cm}^{-1}$  respectively.

In sample H09-13558-A8 all classes of cells were common for  $1558\text{ cm}^{-1}$  only. Luminal and stromal cells were common for  $1699\text{ cm}^{-1}$  (guanine/thymine) and  $1645\text{ cm}^{-1}$  whilst basal and stromal cells were common for  $1456\text{ cm}^{-1}$  (lipids and proteins) whose expression levels were higher in basal cells. Wavenumbers found only in basal cells were  $1747\text{ cm}^{-1}$ ,  $1715\text{ cm}^{-1}$  and  $1651\text{ cm}^{-1}$  whilst luminal cells were unique for  $1771\text{ cm}^{-1}$  and  $1472\text{ cm}^{-1}$ .

In sample H09-11708-4 all three classes of cells were common for  $1558\text{ cm}^{-1}$  and  $1556\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). Wavenumbers for which luminal and stromal cells were common were  $1732\text{ cm}^{-1}$  (lipids),  $1697\text{ cm}^{-1}$  and  $1645\text{ cm}^{-1}$ . Basal cells were found to be unique for  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA),  $1682\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins).

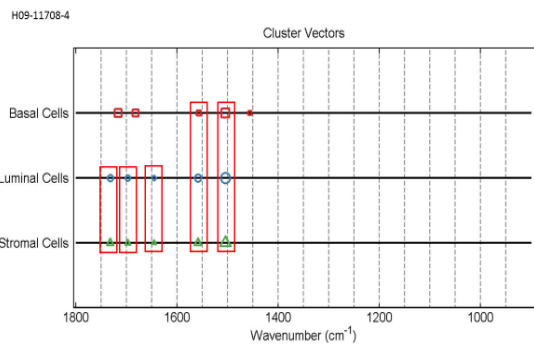
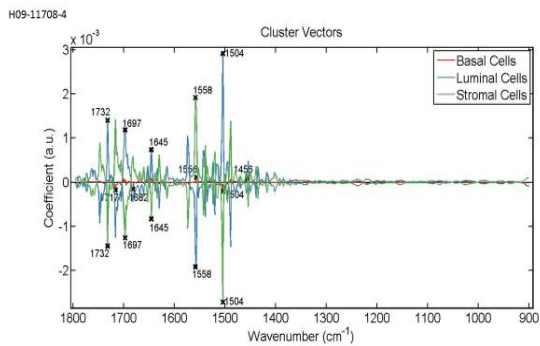
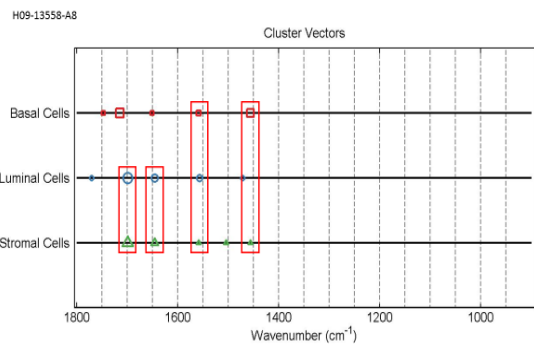
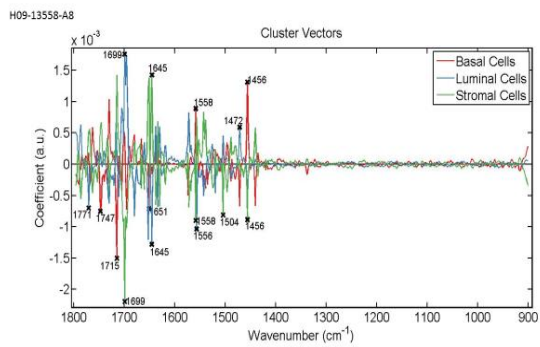
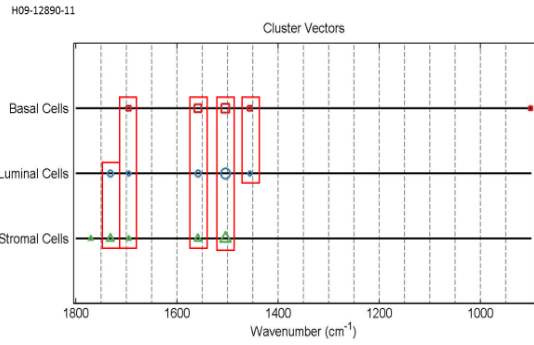
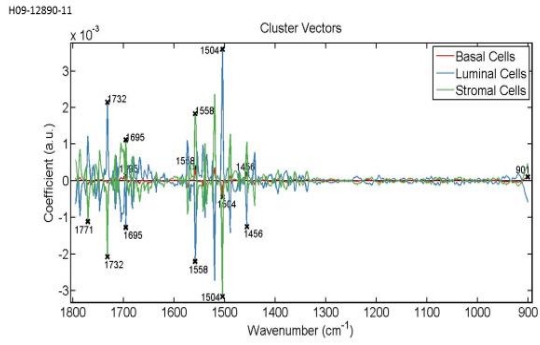
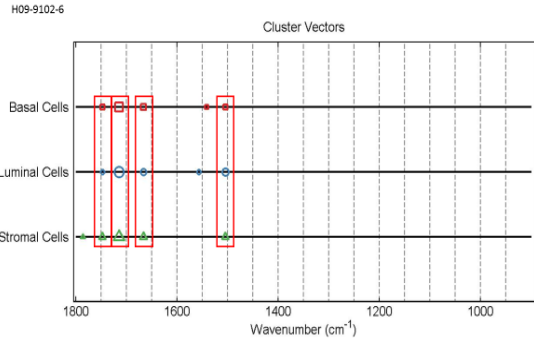
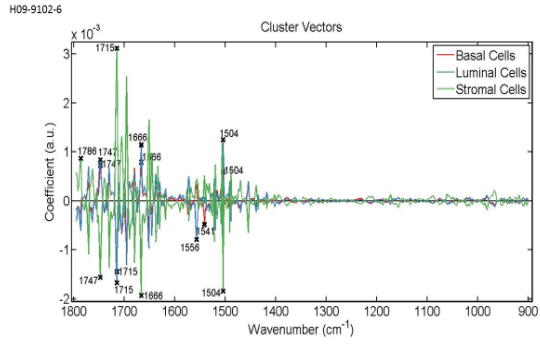
In sample H09-12292-1 all classes of cells were common only for  $1699\text{ cm}^{-1}$  (guanine/thymine). Basal and luminal cells were common for  $1558\text{ cm}^{-1}$ , luminal and stromal cells were common for  $1645\text{ cm}^{-1}$  whilst basal and stromal cells were common for  $1506\text{ cm}^{-1}$ . Basal cells were unique for  $1796\text{ cm}^{-1}$  and

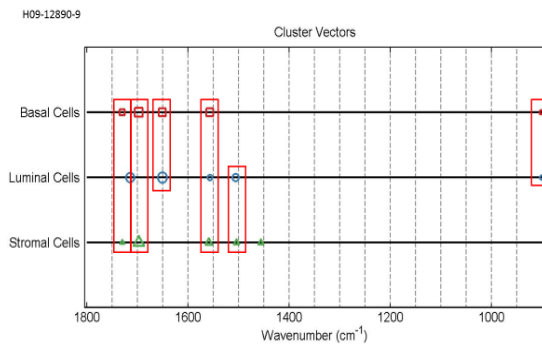
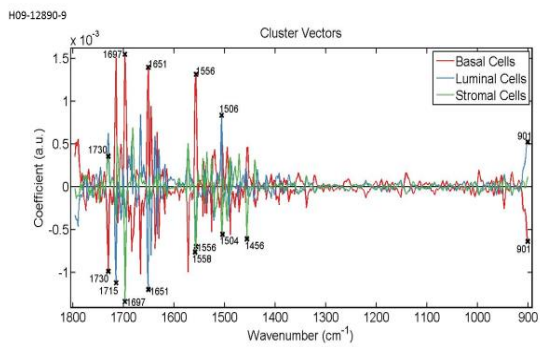
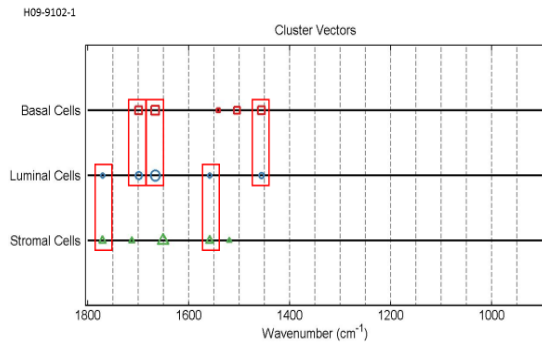
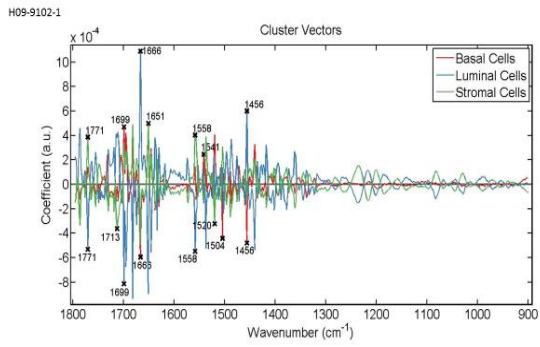
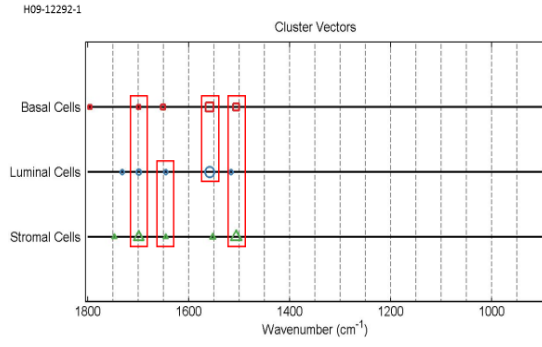
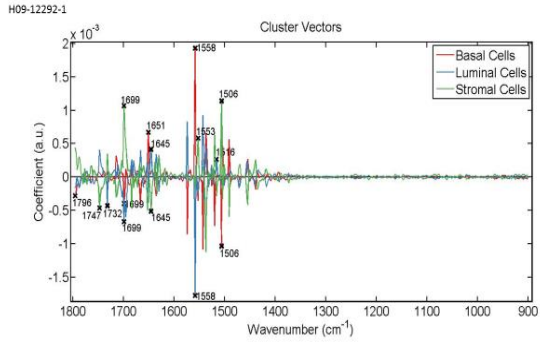
1651  $\text{cm}^{-1}$ . Luminal cells were unique for 1732  $\text{cm}^{-1}$  (lipids) and 1516  $\text{cm}^{-1}$  (amide II) whilst stromal cells were unique only for 1747  $\text{cm}^{-1}$ .

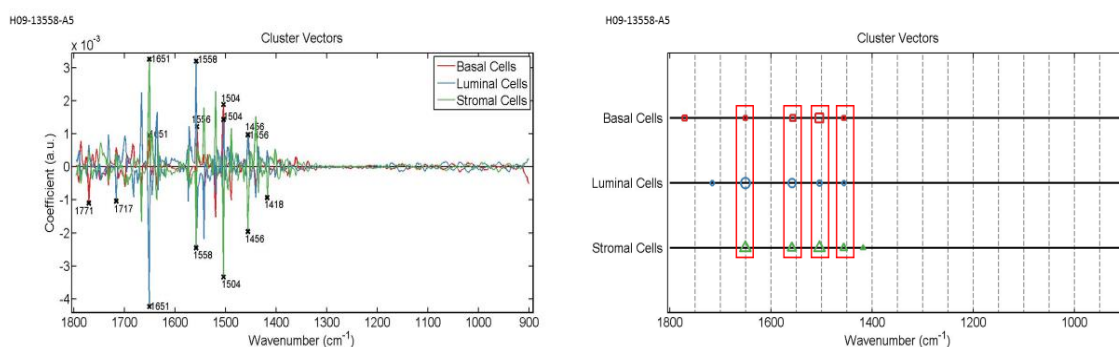
In sample H09-9102-1 no wavenumber was found to be common among all three classes of cells. Basal and luminal cells were common for 1699  $\text{cm}^{-1}$  (guanine/thymine), 1666  $\text{cm}^{-1}$  (amide I) and 1456  $\text{cm}^{-1}$  (lipids and proteins). Expression levels of the latter were higher in basal cells. Luminal and stromal cells were common for 1771  $\text{cm}^{-1}$  and 1558  $\text{cm}^{-1}$  whose expression levels were higher in stromal cells. Basal cells were unique for 1541  $\text{cm}^{-1}$  (amide II) and 1504  $\text{cm}^{-1}$  (phenyl rings) whilst stromal cells were unique for 1713  $\text{cm}^{-1}$  (C=O thymine), 1651  $\text{cm}^{-1}$  and 1520  $\text{cm}^{-1}$  (amide II). No wavenumbers were observed to occur only in luminal cells.

In sample H09-12890-9 the wavenumbers 1558  $\text{cm}^{-1}$  and 1556  $\text{cm}^{-1}$  were commonly shared among all classes of cells. Basal and stromal cells were common for 1730  $\text{cm}^{-1}$  (fatty acid ester) and 1697  $\text{cm}^{-1}$ . Basal and luminal cells were common for 1651  $\text{cm}^{-1}$  and 901  $\text{cm}^{-1}$  whilst common occurrence of 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  was observed between stromal and luminal cells. Also luminal and stromal cells were unique for 1715  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins) respectively.

In sample H09-13558-A5 all classes of cells were common for 1651  $\text{cm}^{-1}$ , 1556  $\text{cm}^{-1}$  with 1558  $\text{cm}^{-1}$ , 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins). Expression levels of 1651  $\text{cm}^{-1}$  were higher in basal cells and levels of 1504  $\text{cm}^{-1}$  (phenyl rings) were lower in luminal cells. Wavenumbers found only in basal, luminal and stromal cells were 1771  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA) and 1418  $\text{cm}^{-1}$  (deformation C-H) respectively.



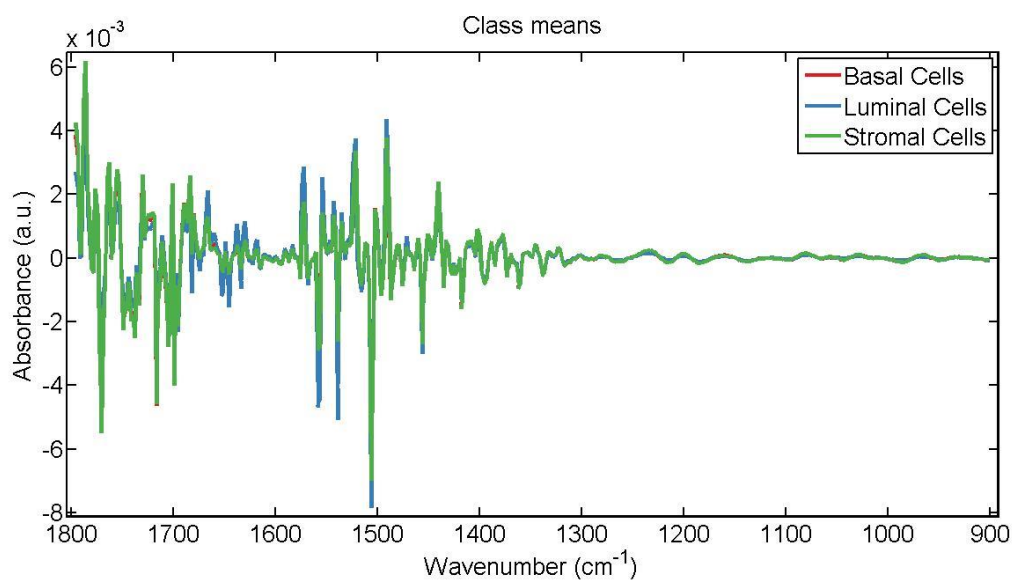




**Figure 3.132:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers among basal (red), luminal (blue) and stromal (green) cells from all glandular areas in individual non-diseased (H09-9102-1, H09-12890-9, H09-13558-A5, H09-11708-4, H09-12292-1) and diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight wavenumbers shared in common between classes of cells.

### *In non-diseased tissues*

The biochemical composition of all basal, luminal and stromal cells in non-diseased tissue samples (H09-9102-6, H09-12890-11, H09-13558-A8) taken from patients with endometrial cancer was analysed. All classes of cells exhibited relatively the same shape in mean FTIR spectra whereby evident variations in absorbance intensities were observed at  $\sim 1682\text{ cm}^{-1}$ ,  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1635\text{ cm}^{-1}$ ,  $\sim 1572\text{ cm}^{-1}$ ,  $\sim 1556\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$ ,  $\sim 1522\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$ ,  $\sim 1491\text{ cm}^{-1}$  and  $\sim 1456\text{ cm}^{-1}$  (Fig. 3.133). Luminal cells had the highest peak intensities at these spectral regions. No peaks were detected in the spectral region  $900\text{-}1300\text{ cm}^{-1}$ . Classes of cells were significantly different from each other as indicated by the *P* values obtained from statistical analysis (Fig. 3.134).



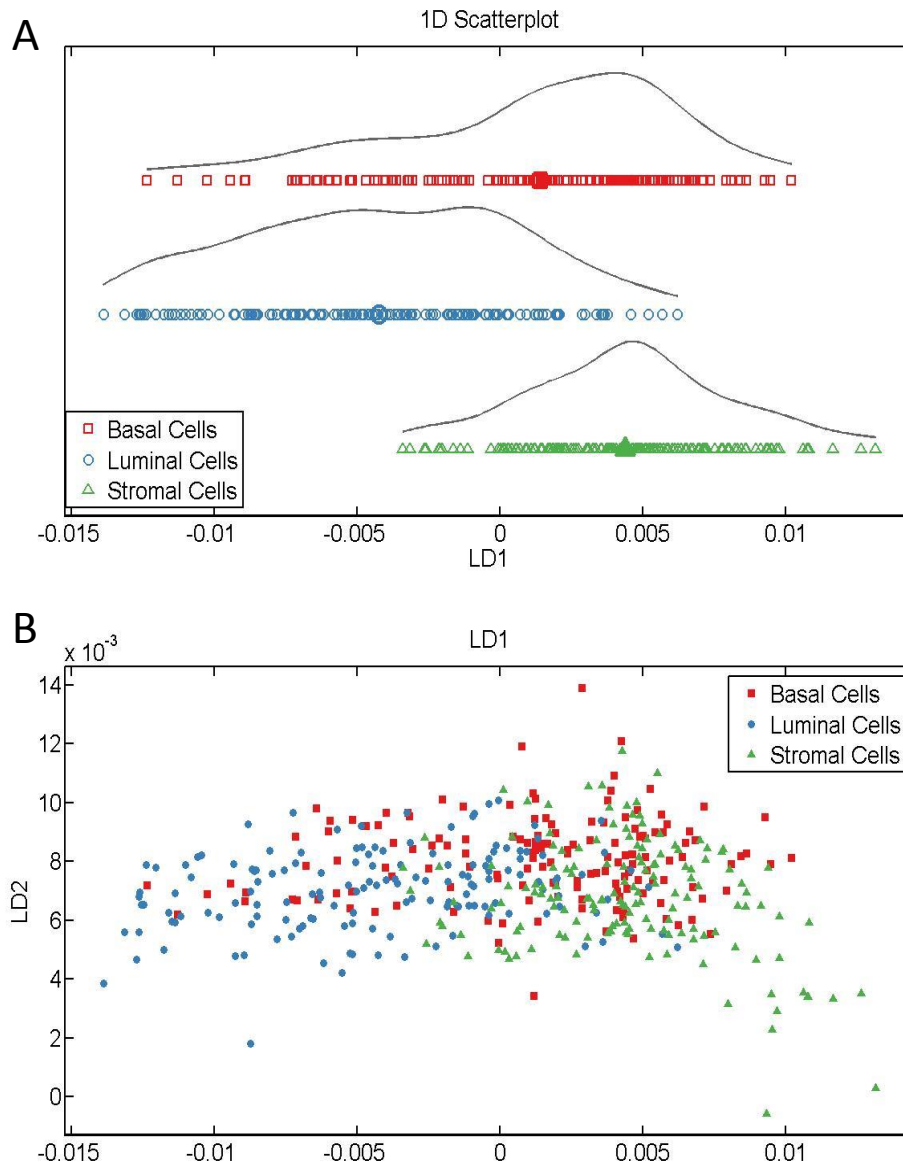
**Figure 3.133:** Mean FTIR spectra for comparison of spectra acquired from all basal (red), luminal (blue) and stromal (green) cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8) to analyse their biochemical composition.

Parameters	P value
LDA1 vs Basal Cells	P > 0.05
LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001
Luminal Cells vs Stromal Cells	P < 0.001

**Figure 3.134:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from all basal, luminal and stromal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8).

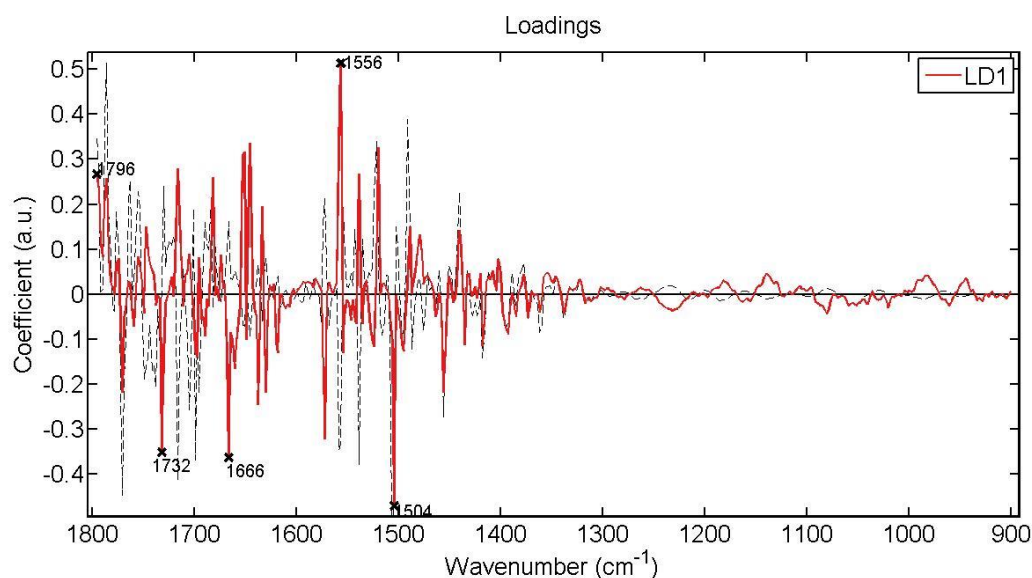


Segregation of classes after application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.135). Basal and luminal cells overlapped with each other throughout most of their spectra. Stromal cells overlapped more with basal cells than with luminal cells. Spectral point of all classes were spread within the dimensional space but relatively more compact arrangement was displayed by spectra of stromal cells and thus signifying less intra-class variation.



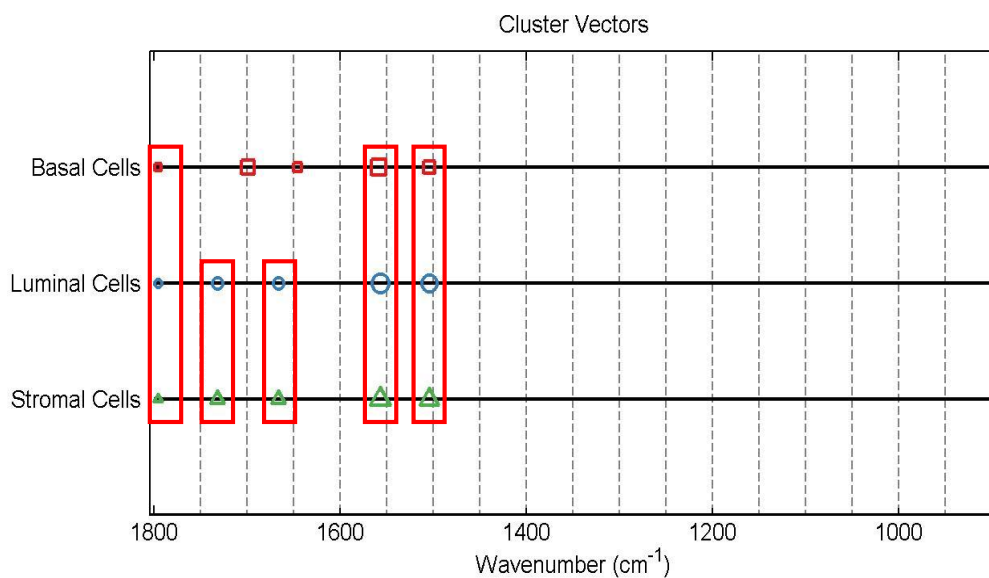
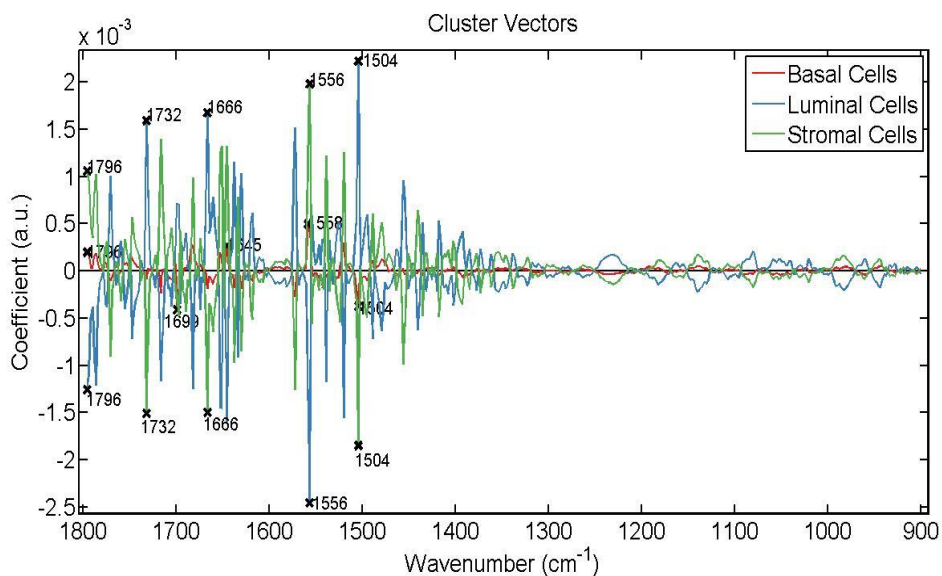
**Figure 3.135:** Scores plots of (A) 1D and (B) 2D space illustrating segregation of spectra acquired from basal (red), luminal (blue) and stromal (green) cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8) produced after application of PCA-LDA. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class

The five major discriminant wavenumbers identified in loading plots (Fig. 3.136) accounting for variations between classes of cells were  $1796\text{ cm}^{-1}$ ,  $1732\text{ cm}^{-1}$  (lipids),  $1666\text{ cm}^{-1}$  (amide I),  $1556\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings).



**Figure 3.136:** Loadings plots showing five major wavenumbers that discriminate basal, luminal and stromal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

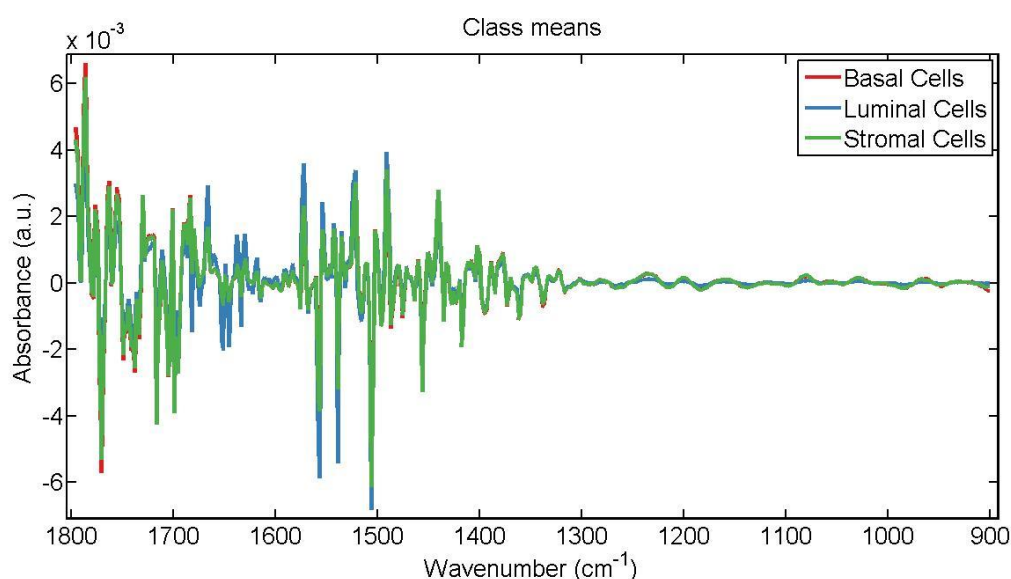
Cluster vector plots were analysed in an alternative way in an attempt to identify potential biomarkers for basal, luminal and stromal cells in non-diseased tissue samples (Fig. 3.137). Observations were based on expression levels of wavenumbers in the different classes of cells and which wavenumbers and/or spectral regions were commonly shared among classes and which wavenumbers occurred only in one class. The expression levels of shared wavenumbers were the same among classes. All classes of cells were common for  $1796\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings). Luminal and stromal cells were common for  $1732\text{ cm}^{-1}$  (lipids) and  $1666\text{ cm}^{-1}$  (amide I). Wavenumbers occurring only in basal cells were  $1699\text{ cm}^{-1}$  (guanine/thymine) and  $1645\text{ cm}^{-1}$ .



**Figure 3.137:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers among basal (red), luminal (blue) and stromal (green) cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red rectangles highlight wavenumbers shared in common between classes of cells.

### *In diseased tissues*

The biochemical composition of all basal, luminal and stromal cells in diseased tissue samples (H09-9102-1, H09-12890-9, H09-13558-A5) taken from cancerous lesions of the endometrium was analysed. All classes of cells exhibited relatively the same shape in mean FTIR spectra (Fig. 3.138). Evident variations in absorbance intensities were observed at  $\sim 1786\text{ cm}^{-1}$ ,  $\sim 1771\text{ cm}^{-1}$ ,  $\sim 1682\text{ cm}^{-1}$ ,  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1647\text{ cm}^{-1}$ ,  $\sim 1634\text{ cm}^{-1}$ ,  $\sim 1572\text{ cm}^{-1}$ ,  $\sim 1556\text{ cm}^{-1}$ ,  $\sim 1539\text{ cm}^{-1}$ ,  $\sim 1506\text{ cm}^{-1}$  and  $\sim 1491\text{ cm}^{-1}$ . Within the spectral region  $\sim 1796\text{--}1738\text{ cm}^{-1}$  slightly higher peaks were exhibited by basal cells whereas within the region  $\sim 1700\text{--}1500\text{ cm}^{-1}$  highest peaks were exhibited by luminal cells. No peaks were detected in the spectral region  $900\text{--}1300\text{ cm}^{-1}$ . Statistical analysis suggested that basal and stromal cells were not significantly different from each other but both were significantly different from luminal cells (Fig. 3.139).

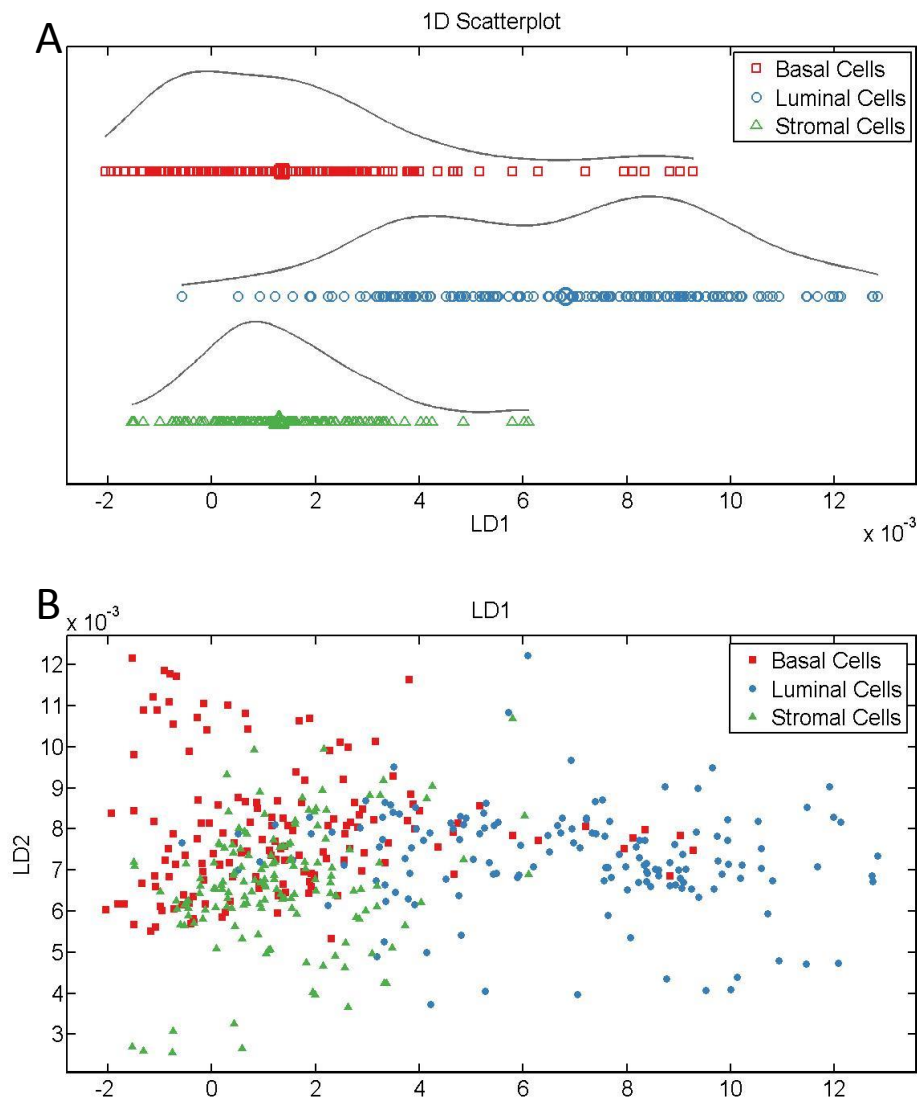


**Figure 3.138:** Mean FTIR spectra for comparison of spectra acquired from all basal (red), luminal (blue) and stromal (green) cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) to analyse their biochemical composition.

Parameters	P value
LDA1 vs Basal Cells	P < 0.001
LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P > 0.05
Luminal Cells vs Stromal Cells	P < 0.001

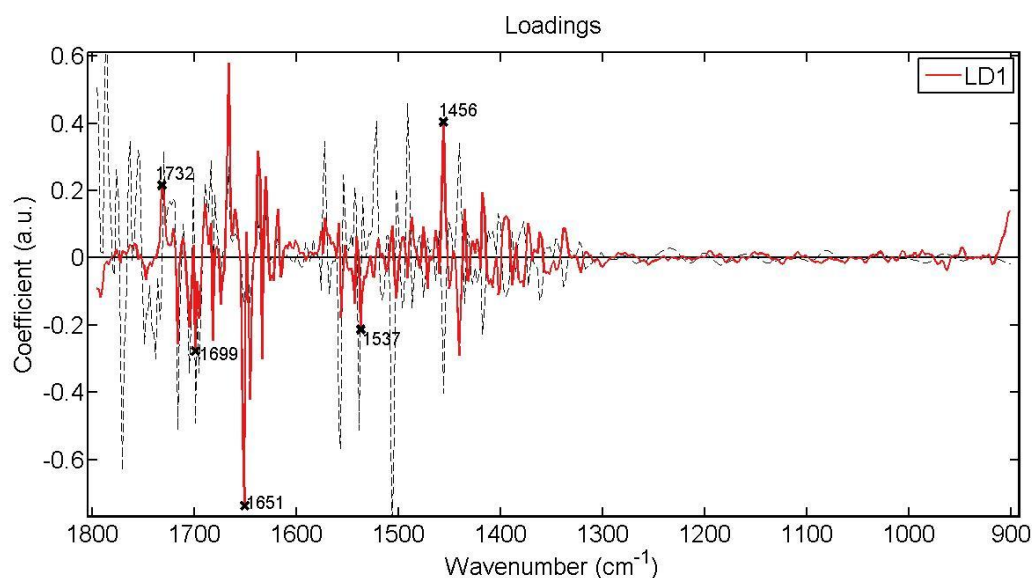
**Figure 3.139:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from all basal, luminal and stromal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5).

Segregation of classes after application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.140). Basal and stromal cells overlapped more with each other relative to the overlap exhibited by either class with luminal cells. Spectral points of luminal cells had a relative more spread arrangement which signified more intra-class variation.



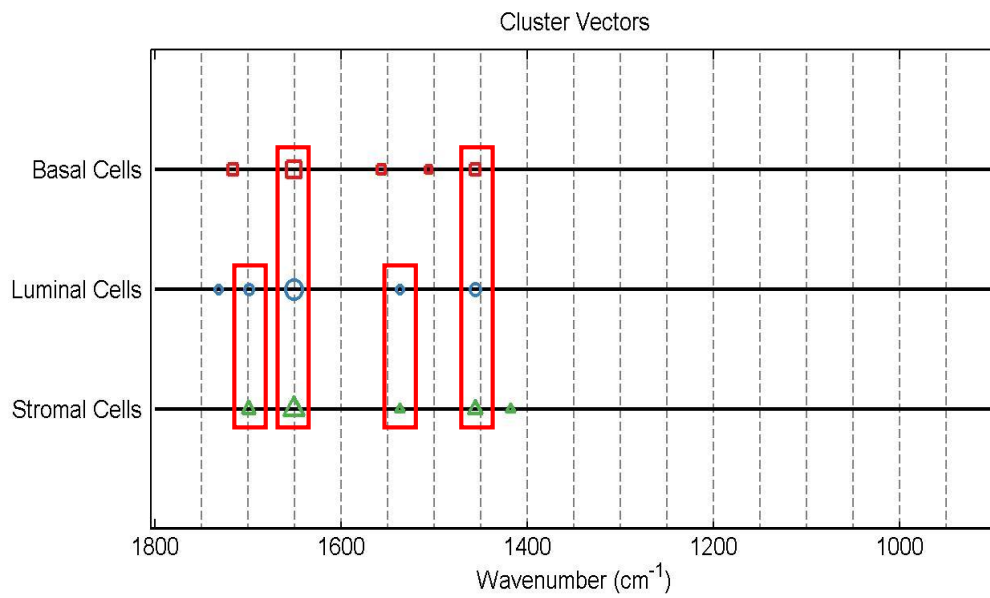
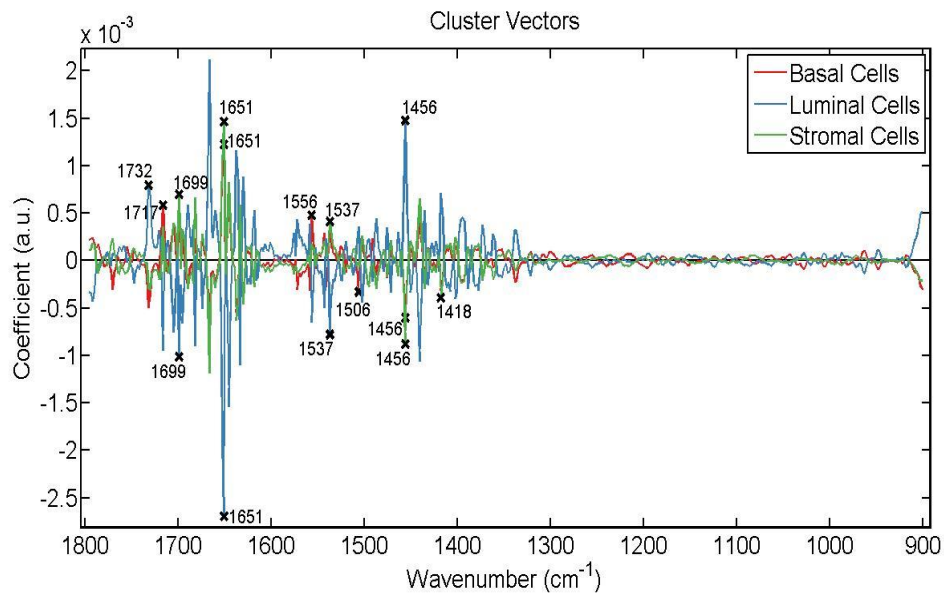
**Figure 3.140:** Scores plots of (A) 1D and (B) 2D space illustrating segregation of spectra acquired from basal (red), luminal (blue) and stromal (green) cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

The five major discriminant wavenumbers identified in loading plots (Fig. 3.141) accounting for variations between classes of cells were 1732  $\text{cm}^{-1}$  (lipids), 1699  $\text{cm}^{-1}$  (guanine/thymine), 1651  $\text{cm}^{-1}$ , 1537  $\text{cm}^{-1}$  (stretching C=N) and 1456  $\text{cm}^{-1}$  (lipids and proteins).



**Figure 3.141:** Loadings plots showing five major wavenumbers that discriminate basal, luminal and stromal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

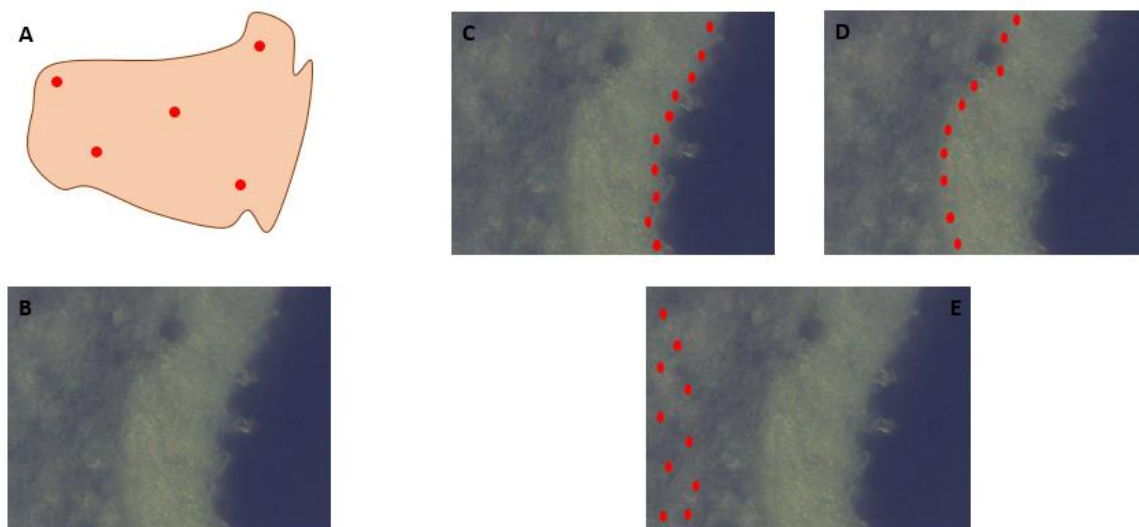
Cluster vector plots were analysed in an alternative way in an attempt to identify potential biomarkers for basal, luminal and stromal cells in diseased tissue samples (Fig. 3.142). Observations were based on expression levels of wavenumbers in the different classes of cells and which wavenumbers and/or spectral regions were commonly shared among classes and which wavenumbers occurred only in one class. The expression levels of shared wavenumbers were the same among classes. The wavenumbers commonly shared between all three classes of cells were 1651  $\text{cm}^{-1}$  and 1456  $\text{cm}^{-1}$  (lipids and proteins). Luminal and stromal cells were common for 1699  $\text{cm}^{-1}$  (guanine/thymine) and 1537  $\text{cm}^{-1}$  (stretching C=N). Wavenumbers occurring only in basal cells were 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA), 1556  $\text{cm}^{-1}$  and 1506  $\text{cm}^{-1}$ . Luminal and stromal cells were unique for 1732  $\text{cm}^{-1}$  (lipids) and 1418  $\text{cm}^{-1}$  (deformation C-H) respectively.



**Figure 3.142:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers among basal (red), luminal (blue) and stromal (green) cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight wavenumbers shared in common between classes of cells.

### 3.3 Raman spectroscopy

The same tissue sections interrogated using FTIR spectroscopy were also interrogated using Raman spectroscopy. A x50 objective lens was used to visualise the five randomly selected glandular elements in a tissue section and locations for spectral acquisition. For each tissue sample a point map was used as a methodological approach to select points for spectral acquisition (Fig. 3.143). Spectra was collected from 10 points corresponding to the location of basal epithelial cells (hypothesised location of epithelial stem/progenitor cells) in a gland (Fig.3.143 D), 10 points from locations of luminal epithelial cells (Fig.3.143 C), which is adjacent to the lumen of the gland, and 10 points from stromal cells from the area surrounding the glandular element (Fig.3.143 E). A total of 30 spectral points were selected in each glandular element and a total of 150 spectral points per tissue sample.



**Figure 3.143: Experimental design – an example of a tissue sample.** (A): shape of interrogated tissue section indicating the locations of the 5 different glandular areas, (B): glandular element as seen through the white light camera, (C): 10 selected points for spectral acquisition from basal cells, (D): 10 selected points for spectral acquisition from luminal cells and (E): 10 selected points for spectral acquisition from surrounding stromal cells.

The raw spectral fingerprint acquired from the cells was pre-processed by application of rubber band baseline correction followed by vector normalisation and wavelet-denoising to remove any issues that arose during spectral collection (Fig. 3.144). Pre-processing was followed by multivariate analysis via

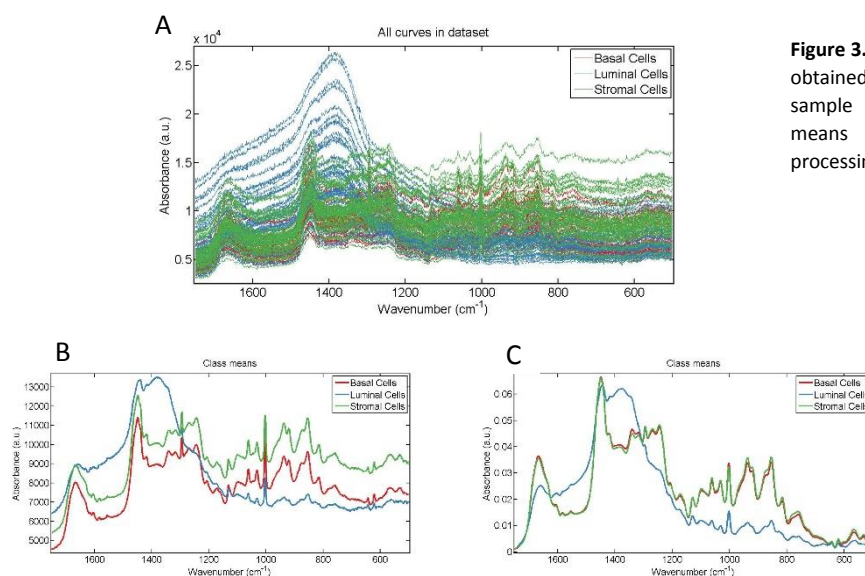


employment of PCA-LDA to allow segregation of classes which was visualised on scores plots of one-dimensional (1D), two-dimensional (2D) or three-dimensional (3D) space, depending on the number of classes. In each case generation of loading plots revealed five major discriminant wavenumbers responsible for variations and thus segregation of classes. Cluster vector plots were presented in a different way illustrating the expression levels of wavenumbers in the classes. Classes were represented by different shapes whose sizes correlated to expression levels i.e. the bigger the size of the shape the higher the expression.

The aim of data analysis was to identify potential biomarkers for basal, luminal and stromal cells that would enable their characterisation based on their location within the glandular elements and eventually assess their 'stemness' bases on the hypothesised location of endometrial stem/progenitor cells. In this concept basal cells were expected to have more endometrial-like nature. For wavenumbers and/or spectral regions to present a potential as biomarkers, they should be shared in common among classes being compared.

Statistical analysis was also employed to assess significance of inter-class variations. One-way ANOVA or unpaired t-test were performed, depending on number of classes being analysed at each section.

On a first approach, the biochemical composition of basal, luminal and stromal cells in tissue samples was analysed individually. Then their spectra was compared and analysed.



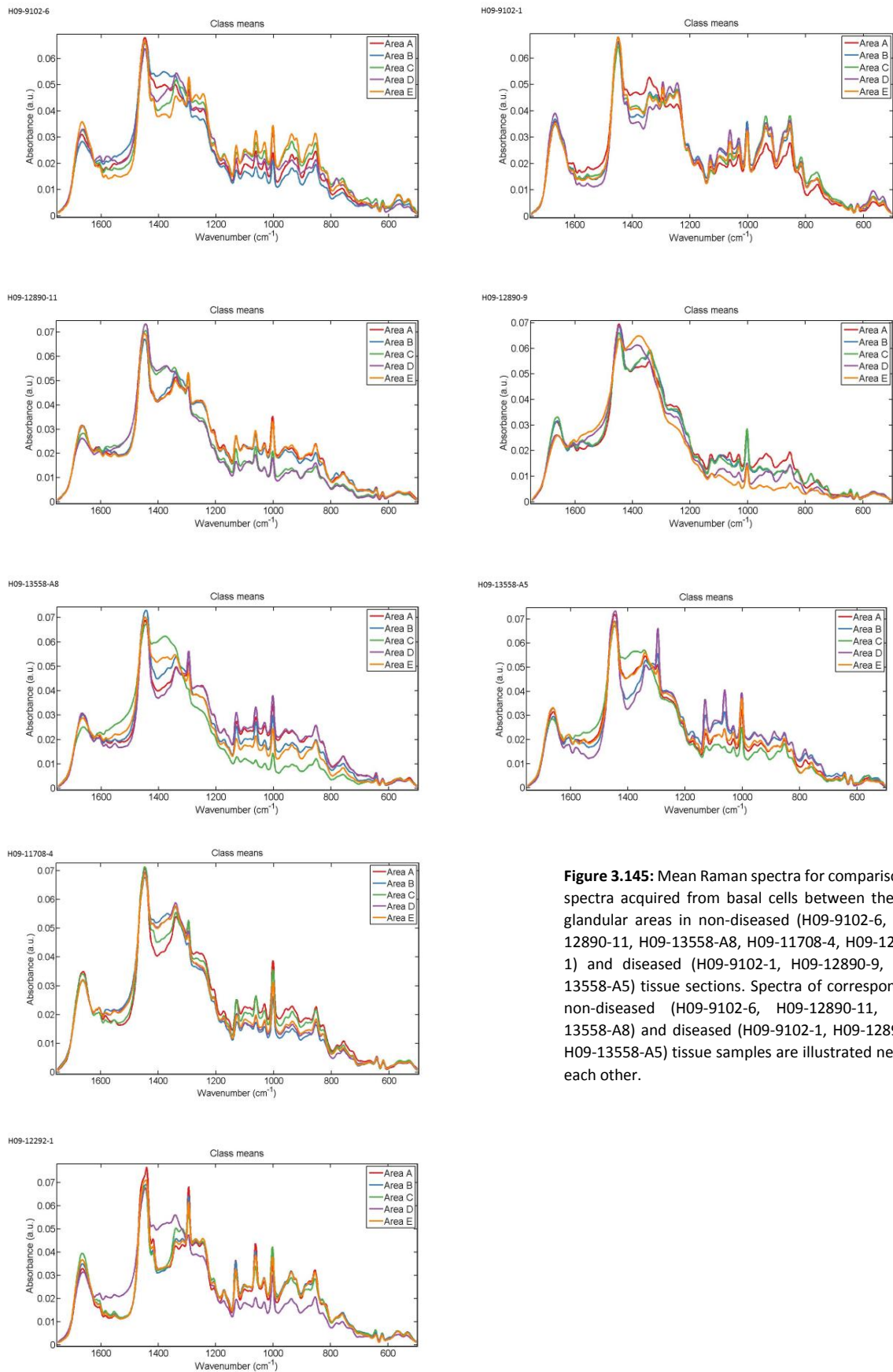
**Figure 3.144:** An example of (A) raw spectra obtained during interrogation of a tissue sample by FTIR spectroscopy and class means (B) before and (C) after pre-processing.

### 3.3.1 Basal Cells Vs Basal Cells

#### *In individual tissue samples*

By looking at the mean Raman spectra of basal cells from the five different glandular elements in each tissue sample (Fig. 3.145) it was observed that areas within individual tissues exhibited relatively the same shape in mean Raman spectra. The shape of the spectra was relatively consistent between samples. It was also observed that the mean spectra of basal cells in non-diseased tissue samples exhibited a very similar shape with their corresponding diseased tissue samples. In all tissues most evident variations in absorbance intensities occurred at  $\sim 1660\text{ cm}^{-1}$ ,  $\sim 1445\text{ cm}^{-1}$ ,  $\sim 1343\text{ cm}^{-1}$ ,  $\sim 1293\text{ cm}^{-1}$ ,  $\sim 1250\text{ cm}^{-1}$ ,  $\sim 1129\text{ cm}^{-1}$ ,  $\sim 1063\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 943\text{ cm}^{-1}$ ,  $\sim 852\text{ cm}^{-1}$  and  $757\text{ cm}^{-1}$ .

Statistical analysis (Fig. 3.146) indicated that in most tissues at least three areas were not significantly different from the tissue as a whole. Whereas in the majority of the tissues, most areas were significantly different from each other.



**Figure 3.145:** Mean Raman spectra for comparison of spectra acquired from basal cells between the five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.

## H09-9102-6

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-12890-11

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P < 0.01
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.01
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-13558-A8

Parameters	P value
LDA1 vs Area A	P < 0.01
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.01
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-11708-4

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001
Area B vs Area C	P < 0.01
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

## H09-12292-1

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P < 0.05
LDA1 vs Area C	P < 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

## H09-9102-1

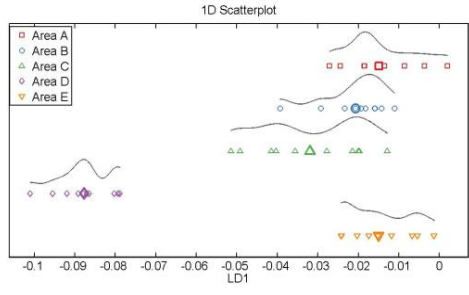
Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

H09-12890-9		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Area A	P < 0.001	LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05	LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001	LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.01	LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05	LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001	Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001	Area A vs Area C	P > 0.05
Area A vs Area D	P > 0.05	Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001	Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05	Area B vs Area C	P < 0.05
Area B vs Area D	P < 0.001	Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05	Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001	Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05	Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001	Area D vs Area E	P < 0.001

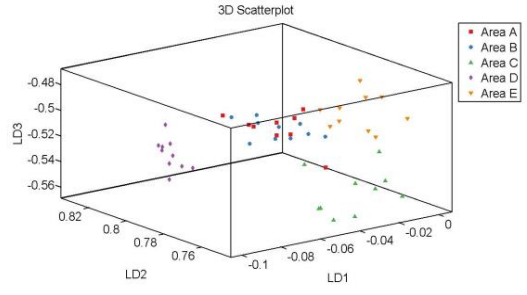
**Figure 3.146:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal in five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Segregation of spectra acquired from the glandular elements was visualised in scores plots of 1D and 3D space. Figure 3.147 illustrates scores plots of non-diseased tissue samples and figure 3.148 illustrates scores plots of diseased tissue samples. Not all areas overlapped with each other in individual tissues and spectral points of all areas had relatively the same arrangement in the dimensional space which was rather spread. No specific observations were made that would correlate corresponding non-diseased and diseased tissue sections.

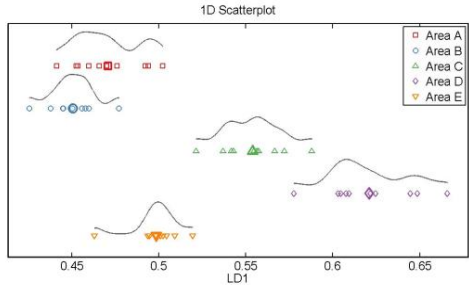
H09-9102-6



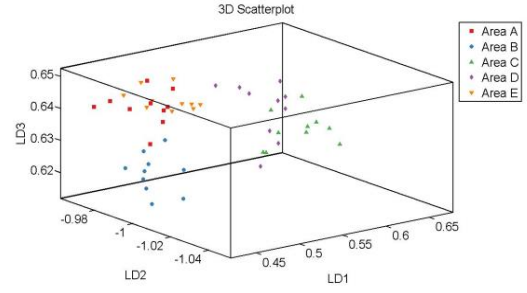
H09-9102-6



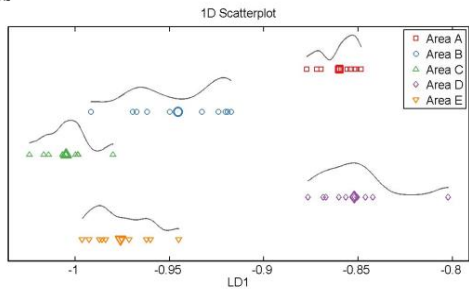
H09-12890-11



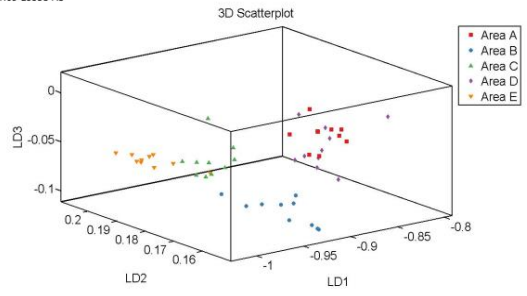
H09-12890-11



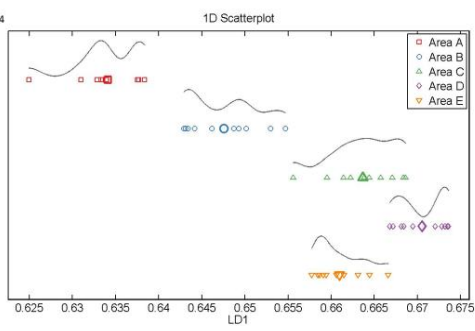
H09-13558-A8



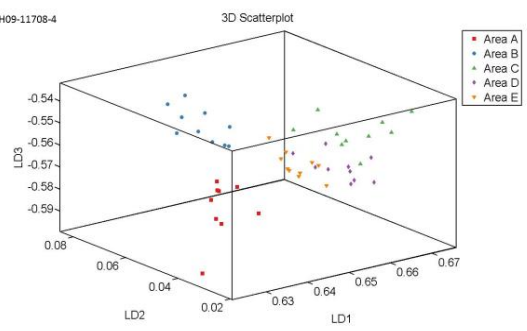
H09-13558-A8

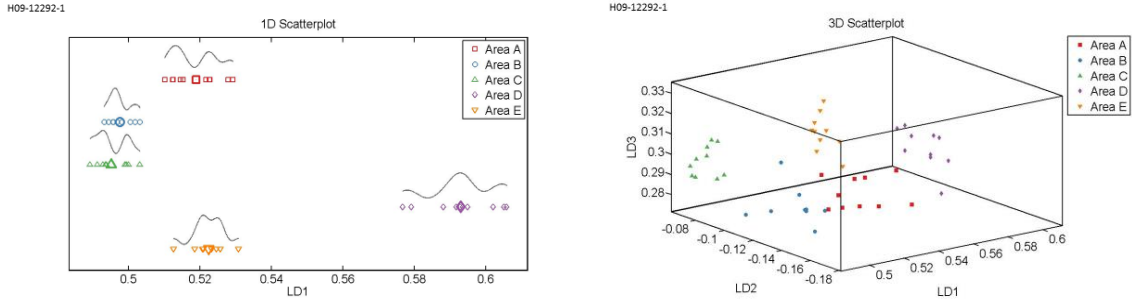


H09-11708-4

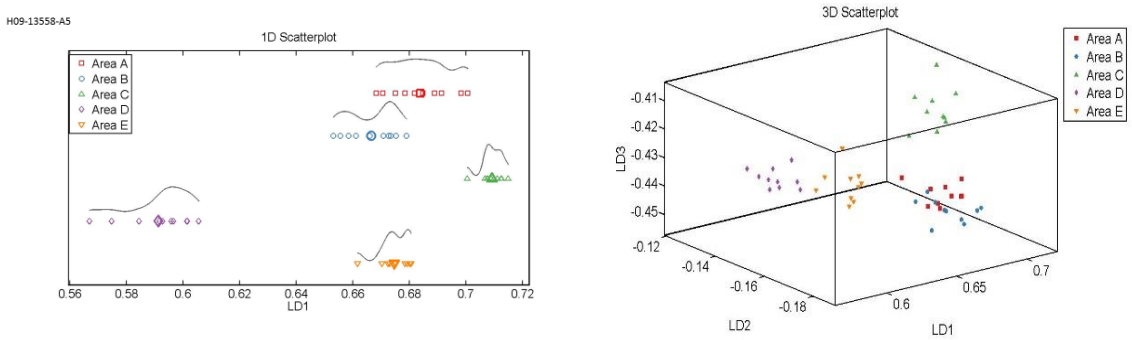
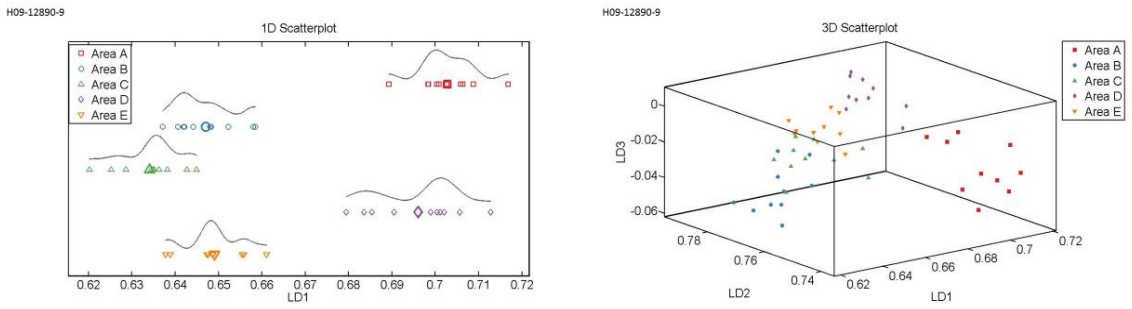
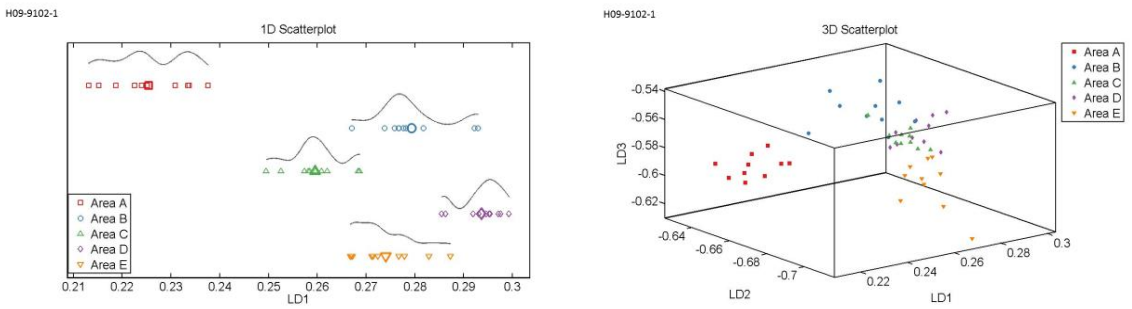


H09-11708-4





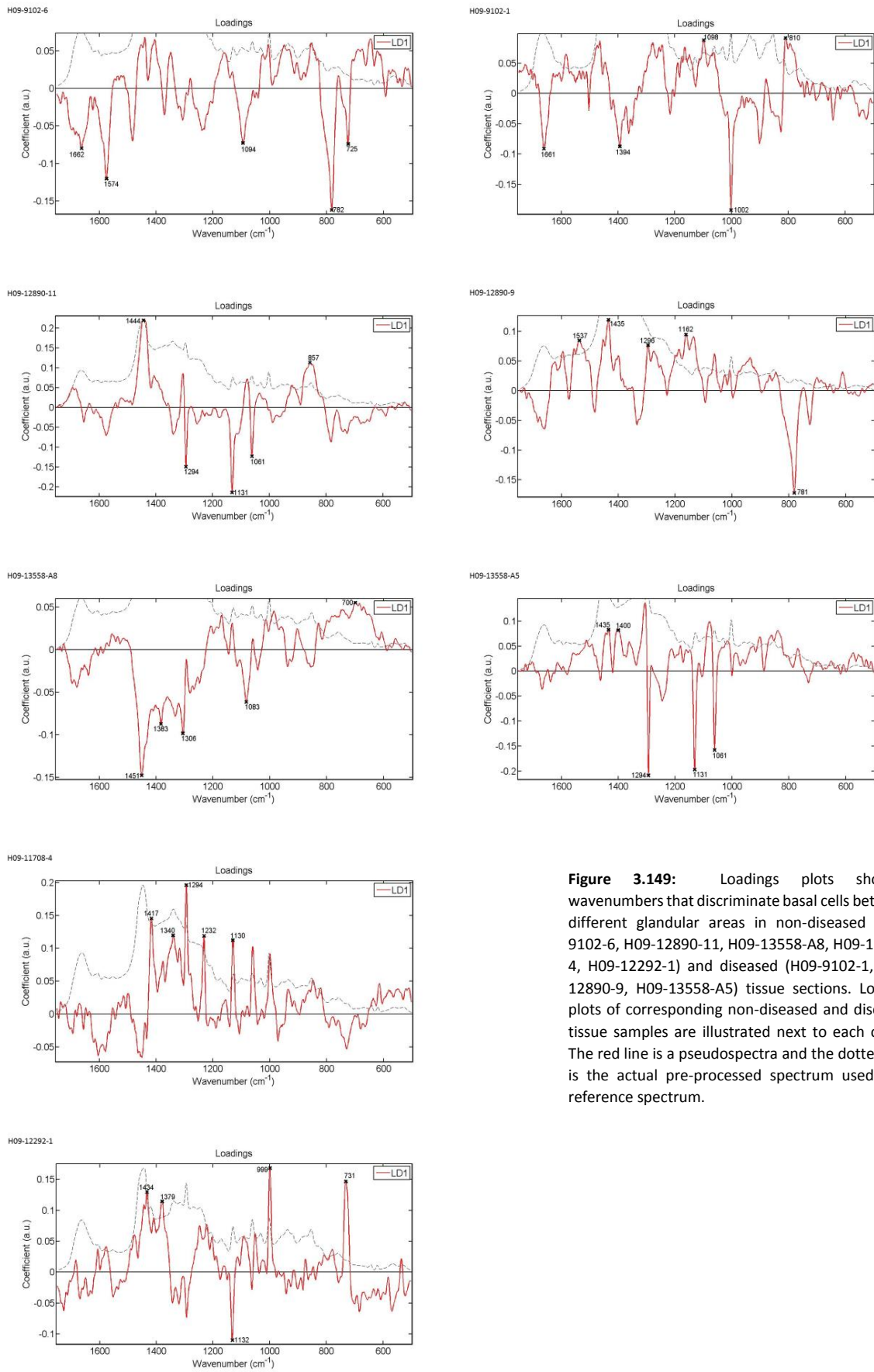
**Figure 3.147:** 1D scores plots and their corresponding 3D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from basal cells from five different glandular elements.



**Figure 3.148:** 1D scores plots and their corresponding 3D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from basal cells from five different glandular elements.

Figure 3.149 illustrates the loading plot for each tissue sample revealing the discriminant wavenumbers responsible for variations between the spectra acquired from basal cells at glandular areas. In tissue sample H09-9102-6 the discriminant wavenumbers and their assignments were as follows; 1662  $\text{cm}^{-1}$  (amide I), 1574  $\text{cm}^{-1}$  (purine rings, DNA/RNA), 1094  $\text{cm}^{-1}$  (DNA), 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA) and 725  $\text{cm}^{-1}$  (DNA/RNA bases). In sample H09-12890-11 the discriminant wavenumbers were 1444  $\text{cm}^{-1}$  (lipids), 1294  $\text{cm}^{-1}$  (methylene twisting), 1131  $\text{cm}^{-1}$  (lipids), 1061  $\text{cm}^{-1}$  (paraffin) and 857  $\text{cm}^{-1}$ . For tissue H09-13558-A8 the discriminant wavenumbers were 1451  $\text{cm}^{-1}$  (proteins and lipids), 1383  $\text{cm}^{-1}$ , 1306  $\text{cm}^{-1}$ , 1083  $\text{cm}^{-1}$  (proteins and lipids) and 700  $\text{cm}^{-1}$ . For tissue H09-11708-4 the identified wavenumbers were 1417  $\text{cm}^{-1}$  (quinoid ring), 1340  $\text{cm}^{-1}$  (nucleic acids/collagen), 1294  $\text{cm}^{-1}$  (methylene twisting), 1232  $\text{cm}^{-1}$  and 1130  $\text{cm}^{-1}$  (lipids). For tissue H09-12292-1 the wavenumbers were 1434  $\text{cm}^{-1}$ , 1379  $\text{cm}^{-1}$  (lipids), 1132  $\text{cm}^{-1}$  (proteins and lipids), 999  $\text{cm}^{-1}$  (phospholipid, glucose-1-phosphate) and 731  $\text{cm}^{-1}$ . In sample H09-9102-1 the identified wavenumbers were 1661  $\text{cm}^{-1}$ , 1394  $\text{cm}^{-1}$ , 1098  $\text{cm}^{-1}$ , 1002  $\text{cm}^{-1}$  (phenylalanine) and 810  $\text{cm}^{-1}$  (phosphodiester/phosphate backbone). In sample H09-12890-9, the wavenumbers were 1537  $\text{cm}^{-1}$ , 1435  $\text{cm}^{-1}$ , 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation), 1162  $\text{cm}^{-1}$  and 781  $\text{cm}^{-1}$  (cytosine/uracil). In sample H09-13558-A5 the wavenumbers were 1435  $\text{cm}^{-1}$ , 1400  $\text{cm}^{-1}$  (proteins), 1294  $\text{cm}^{-1}$  (methylene twisting), 1131  $\text{cm}^{-1}$  (lipids) and 1061  $\text{cm}^{-1}$  (paraffin).

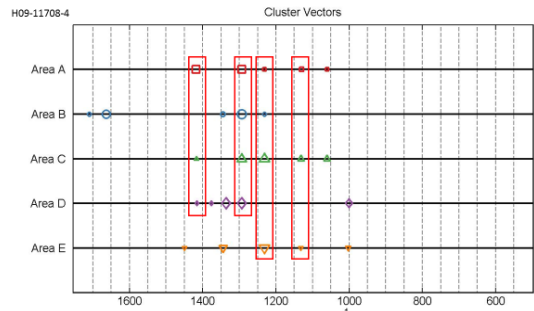
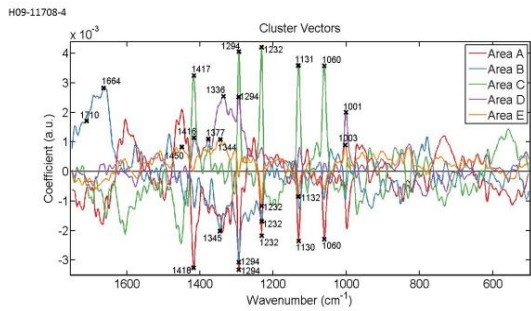
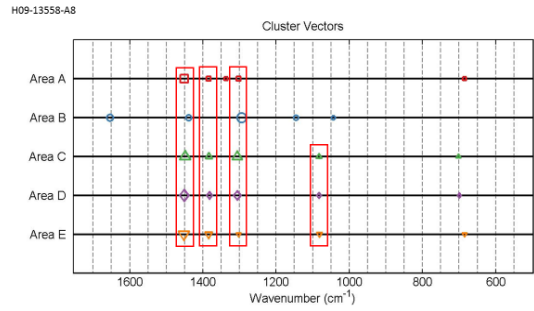
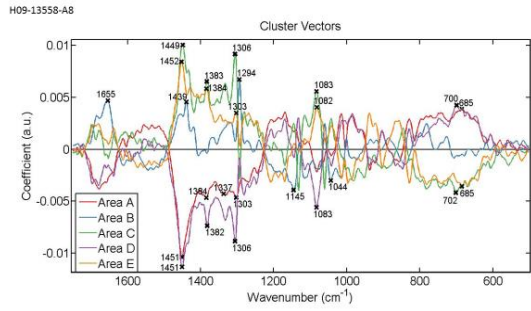
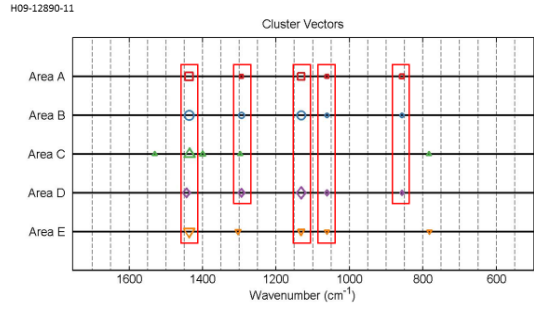
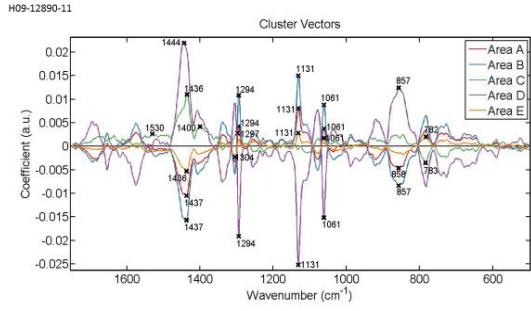
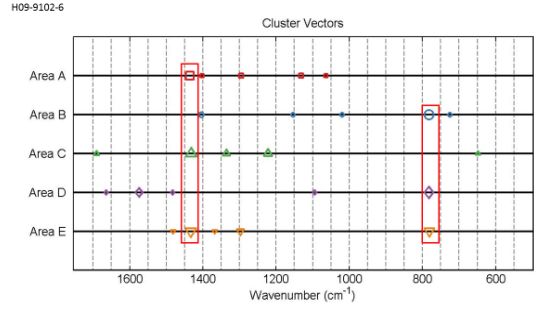
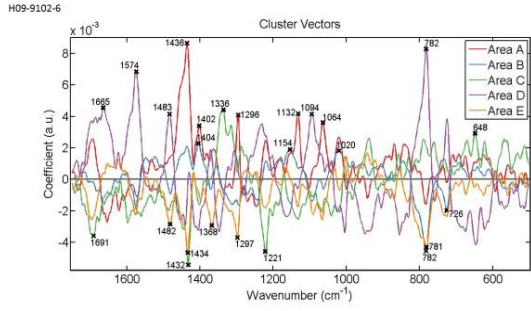


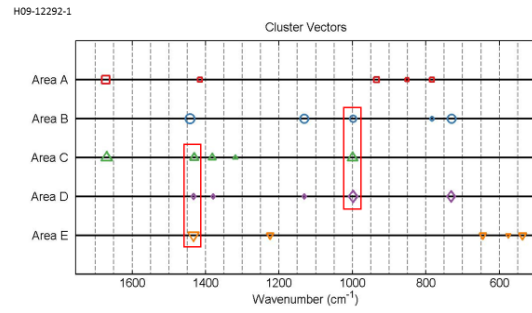
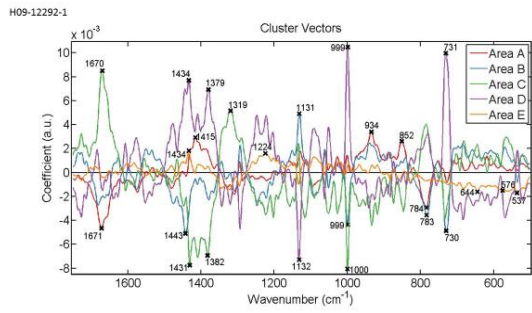


**Figure 3.149:** Loadings plots showing wavenumbers that discriminate basal cells between different glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

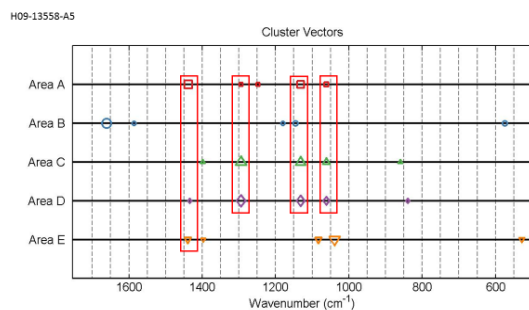
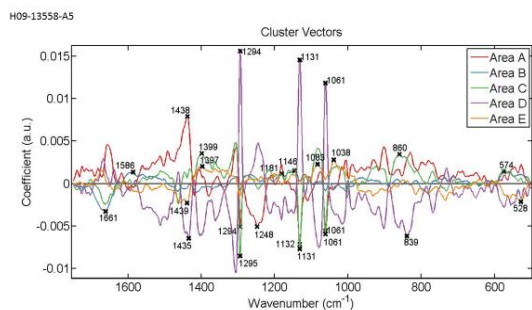
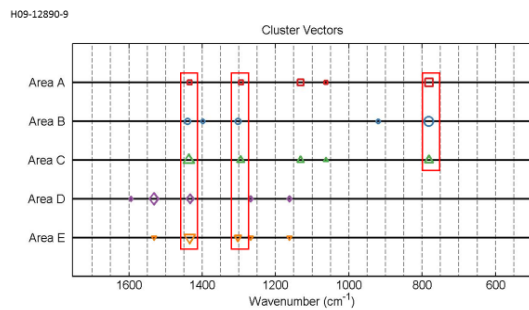
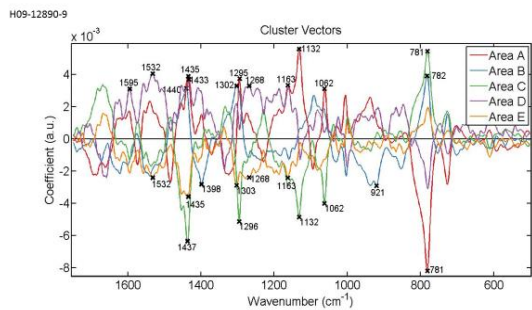
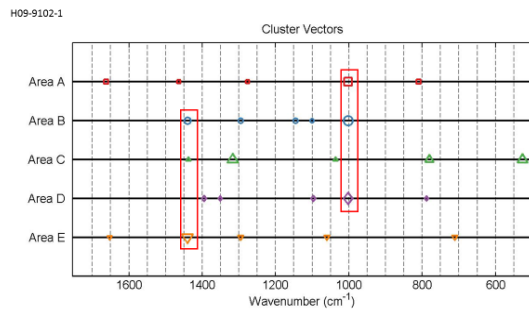
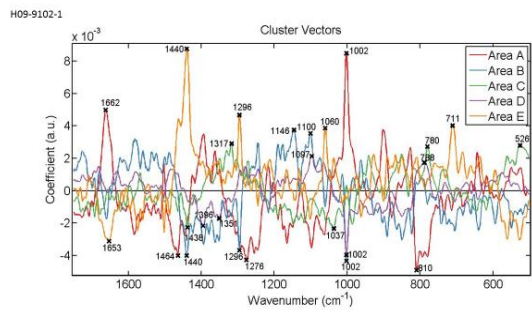
Cluster vector plots of non-diseased (Fig. 3.150) and diseased (Fig. 3.151) tissue samples were presented and analysed in the same approach as before based on the common occurrence and expression levels of wavenumbers in an attempt to identify potential biomarkers for basal cells.

In sample H09-9102-6, three areas were common for 1436  $\text{cm}^{-1}$  (lipids) with 1434  $\text{cm}^{-1}$  and 1432  $\text{cm}^{-1}$ , and 782  $\text{cm}^{-1}$  (cytosine/uracil) with 781  $\text{cm}^{-1}$  (cytosine/uracil). In sample H09-12890-11 three areas were common for 858  $\text{cm}^{-1}$  with 857  $\text{cm}^{-1}$  and four areas were common for 1436  $\text{cm}^{-1}$  (lipids) with 1437  $\text{cm}^{-1}$  (proteins and lipids), 1294  $\text{cm}^{-1}$  (methylene twisting) with 1297  $\text{cm}^{-1}$ , 1131  $\text{cm}^{-1}$  (lipids) and 1061  $\text{cm}^{-1}$  (paraffin). In sample H09-13558-A8 common occurrence of 1449  $\text{cm}^{-1}$  (proteins and lipids) with 1451  $\text{cm}^{-1}$  and 1452  $\text{cm}^{-1}$ , 1382  $\text{cm}^{-1}$  with 1383  $\text{cm}^{-1}$  and 1384  $\text{cm}^{-1}$ , and 1303  $\text{cm}^{-1}$  (collagen) with 1306  $\text{cm}^{-1}$  was displayed by four areas whilst three areas were common for 1082  $\text{cm}^{-1}$  (proteins) and 1083  $\text{cm}^{-1}$  (proteins and lipids). In sample H09-11708-4 three areas were common for 1418  $\text{cm}^{-1}$  with 1417  $\text{cm}^{-1}$  and 1416  $\text{cm}^{-1}$ , and 1132  $\text{cm}^{-1}$  with 1131  $\text{cm}^{-1}$  (lipids) and 1130  $\text{cm}^{-1}$  (phospholipids) whilst four areas were common for the 1294  $\text{cm}^{-1}$  (methylene twisting) and 1232  $\text{cm}^{-1}$  (amide III). In sample H09-12292-1 three areas showed similarity for the wavenumbers 1431  $\text{cm}^{-1}$  with 1434  $\text{cm}^{-1}$ , and 999  $\text{cm}^{-1}$  (phospholipids) with 1000  $\text{cm}^{-1}$  (phenylalanine). In H09-9102-1 three areas were common for 1438  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) with 1440  $\text{cm}^{-1}$  (lipids), and 1002  $\text{cm}^{-1}$  (phenylalanine). Five areas in sample H09-12890-9 were common for 1440  $\text{cm}^{-1}$  (lipids) with 1437  $\text{cm}^{-1}$  (proteins and lipids), 1435  $\text{cm}^{-1}$  and 1433  $\text{cm}^{-1}$ , four areas were common for 1303  $\text{cm}^{-1}$  with 1302  $\text{cm}^{-1}$  (protein, collagen/lipids), 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) and 1295  $\text{cm}^{-1}$  whilst 781  $\text{cm}^{-1}$  (cytosine/uracil) displayed common occurrence with 782  $\text{cm}^{-1}$  (cytosine/uracil) in three areas. In sample H09-13558-A5, three areas displayed common for 1439  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) with 1438  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) and 1435  $\text{cm}^{-1}$ , 1294  $\text{cm}^{-1}$  (methylene twisting) with 1295  $\text{cm}^{-1}$ , 1131  $\text{cm}^{-1}$  (lipids) with 1132  $\text{cm}^{-1}$ , and 1061  $\text{cm}^{-1}$  (paraffin).





**Figure 3.150:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells, in five glandular areas in individual non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1). The red rectangles highlight the wavenumbers shared in common among glandular areas.

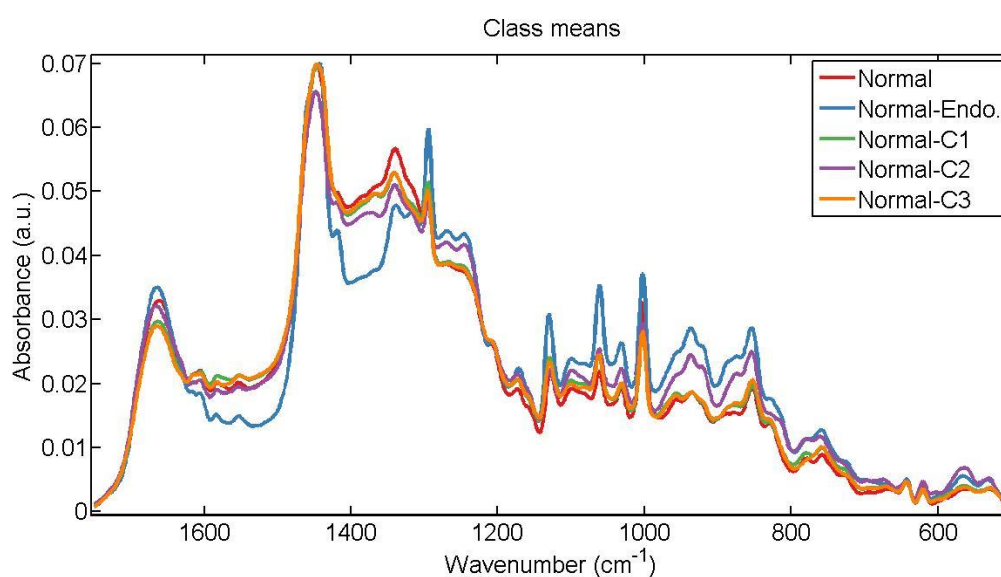


**Figure 3.151:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells, in five glandular areas in individual diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight the wavenumbers shared in common among glandular areas.

*Between all normal tissue samples*

All the non-diseased tissue samples were compared between them for the spectra interrogated from their basal cells. Classes representing tissues were labelled as before. In this section tissues will be referred according to their class labels to associate plots with observations. It was observed that all tissues had similar shape in mean Raman spectra with most prominent variations in absorbance intensities occurring at  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1130\text{ cm}^{-1}$ ,  $\sim 1061\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 936\text{ cm}^{-1}$  and  $\sim 853\text{ cm}^{-1}$  whereby highest peaks were mainly exhibited by Normal-Endo. (Fig. 3.152).

Statistical analysis indicated that Normal and Normal-Endo. were not significantly different from each other neither were Normal-C1, Normal-C2 and Normal C3 (Fig. 3.153).

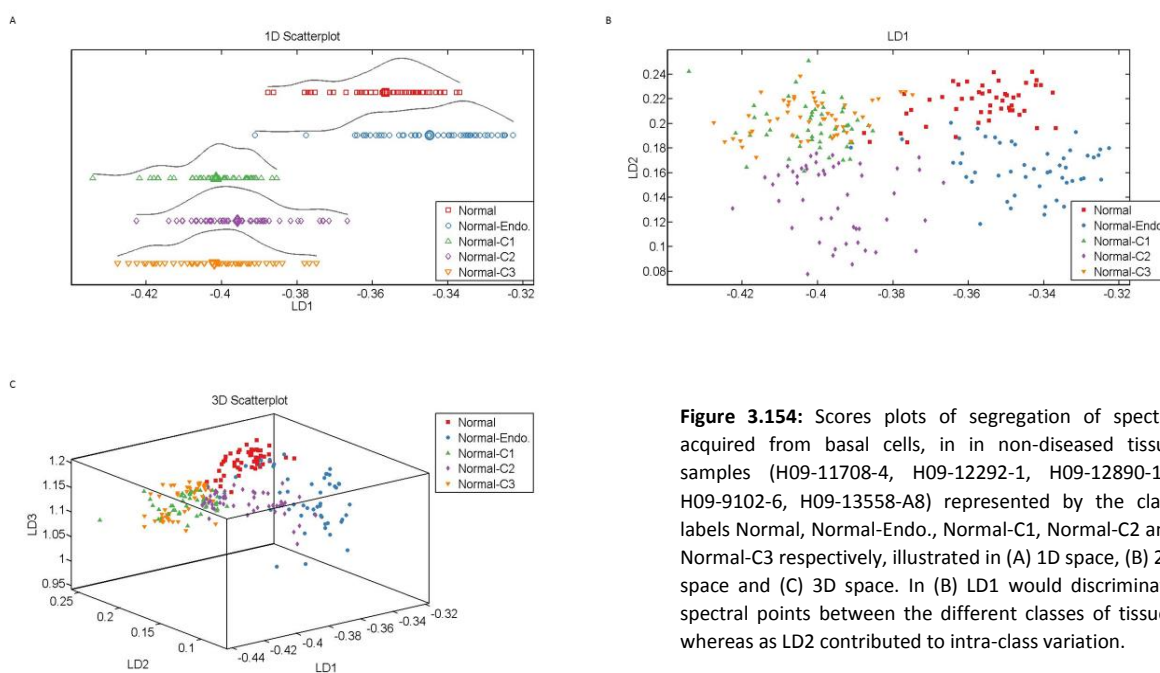


**Figure 3.152:** Mean RAMAN spectra for comparison of spectra acquired from basal cells in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Parameters	P value
LDA1 vs Normal	P < 0.001
LDA1 vs Normal-Endo.	P < 0.001
LDA1 vs Normal-C1	P < 0.001
LDA1 vs Normal-C2	P < 0.001
LDA1 vs Normal-C3	P < 0.001
Normal vs Normal-Endo.	P > 0.05
Normal vs Normal-C1	P < 0.001
Normal vs Normal-C2	P < 0.001
Normal vs Normal-C3	P < 0.001
Normal-Endo. vs Normal-C1	P < 0.001
Normal-Endo. vs Normal-C2	P < 0.001
Normal-Endo. vs Normal-C3	P < 0.001
Normal-C1 vs Normal-C2	P > 0.05
Normal-C1 vs Normal-C3	P > 0.05
Normal-C2 vs Normal-C3	P > 0.05

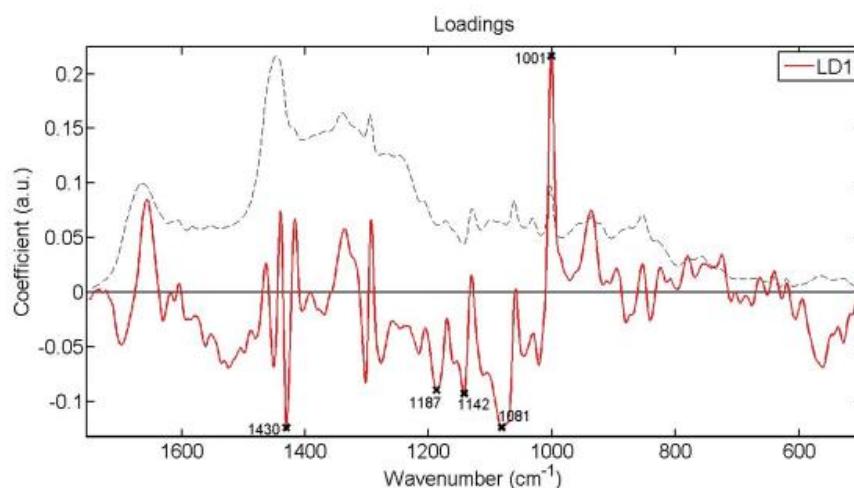
**Figure 3.153:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal cells in all five glandular elements between non-diseased tissue sections H09-11708-4, H09-9102-1, H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D, 2D and 3D space (Fig. 3.154). Normal-C1, Normal-C2 and Normal-C3 overlapped mostly between them and displayed little overlap with the two classes. Similarly Normal and Normal-Endo. overlapped more with each other. Mainly spectral points of Normal-Endo. and Normal-C2 exhibited the most spread arrangement relative to the other classes which signified the most intra-class variation.



**Figure 3.154:** Scores plots of segregation of spectra acquired from basal cells, in in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively, illustrated in (A) 1D space, (B) 2D space and (C) 3D space. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

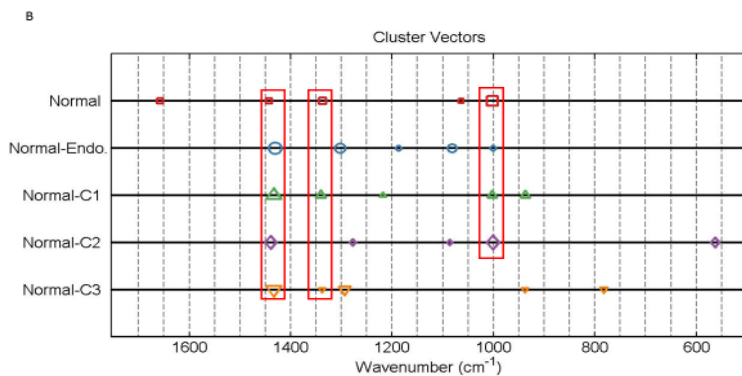
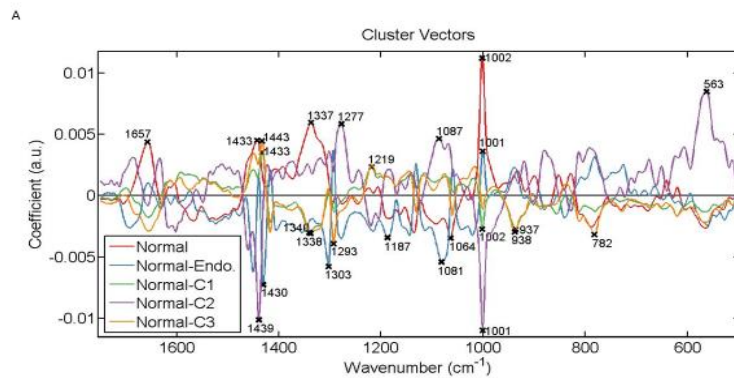
The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were 1430  $\text{cm}^{-1}$ , 1187  $\text{cm}^{-1}$ , 1142  $\text{cm}^{-1}$ , 1081  $\text{cm}^{-1}$  (lipids) and 1001  $\text{cm}^{-1}$  (phenylalanine) (Fig. 3.155).



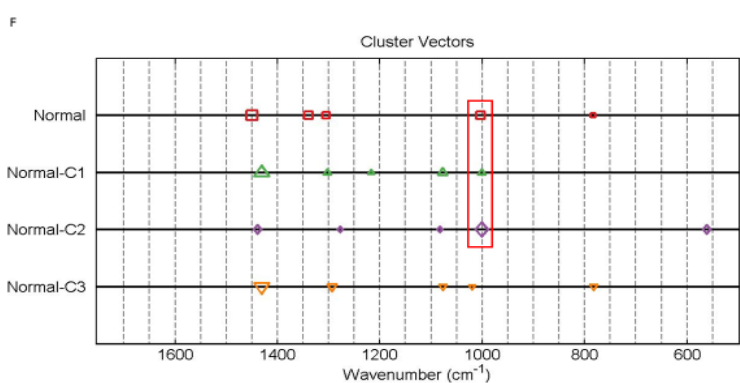
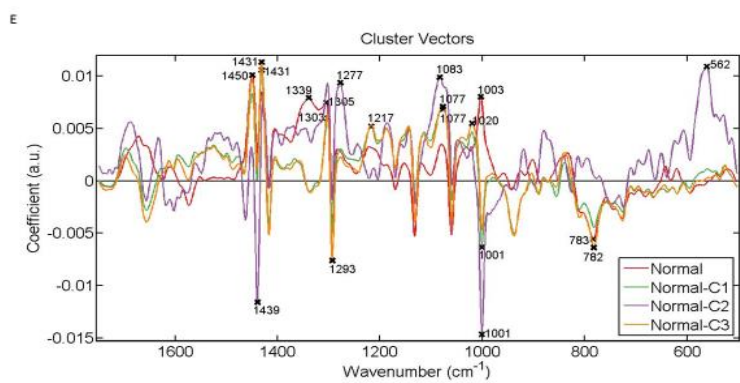
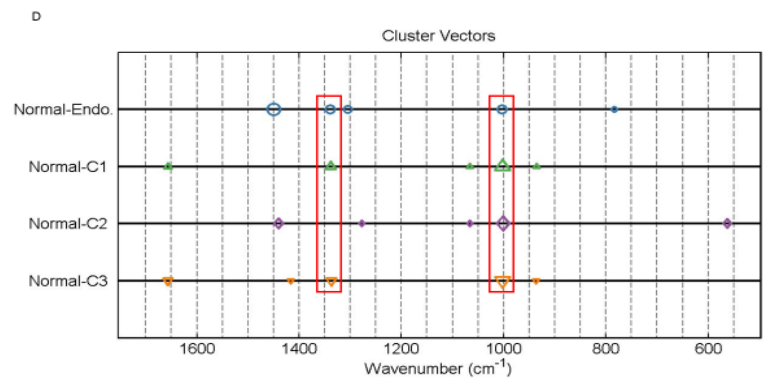
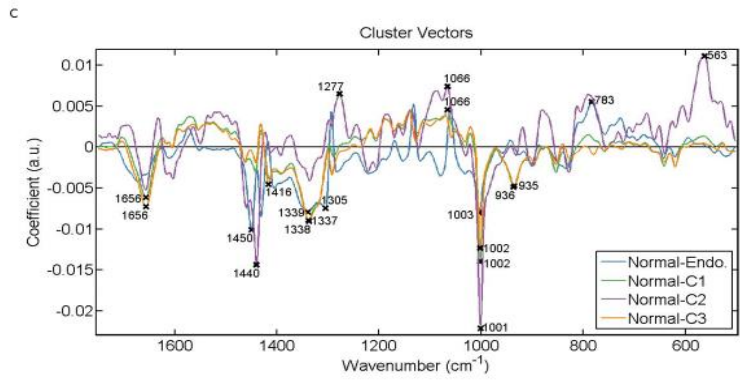
**Figure 3.155:** Loadings plots showing wavenumbers that discriminate basal cells in non-diseased tissue sections (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Some of the aforementioned wavenumbers were also observed in cluster vector plots which were presented and analysed in the same approach as before in order to identify potential biomarkers for basal cells (Fig. 3.156). Without having any class as a reference, 1433  $\text{cm}^{-1}$  (lipids) with 1430  $\text{cm}^{-1}$  were commonly shared by three tissues, 1340  $\text{cm}^{-1}$  (nucleic acids/collagen) with 1338  $\text{cm}^{-1}$  (nucleic acids/collagen) and 1337  $\text{cm}^{-1}$  (nucleic acids/collagen) were commonly shared by three tissues whilst 1002  $\text{cm}^{-1}$  (phenylalnine) with 1001  $\text{cm}^{-1}$  (phenylalanine) were commonly shared by four tissues. Using Normal as a reference, three tissues were common for 1339  $\text{cm}^{-1}$  (tryptophan/collagen/nucleic acids) with 1338  $\text{cm}^{-1}$  (tryptophan/collagen/nucleic acids) and 1337  $\text{cm}^{-1}$  (tryptophan/collagen/nucleic acids) whilst four tissues were common for 1003  $\text{cm}^{-1}$  (phenylalnine), 1002  $\text{cm}^{-1}$  (phenylalnine) and 1001  $\text{cm}^{-1}$  (phenylalanine). Using Normal-Endo. as reference, 1003  $\text{cm}^{-1}$  was commonly shared with 1001  $\text{cm}^{-1}$  by three tissues, both of which are assigned to phenylalanine. Similarly using Normal-C1 as reference, common occurrence of 1003  $\text{cm}^{-1}$  (phenylalnine) with 1002  $\text{cm}^{-1}$  (phenylalnine) and

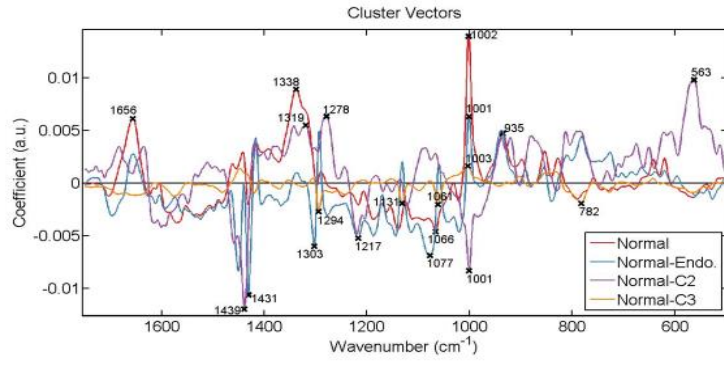
1001  $\text{cm}^{-1}$  (phenylalanine). Having Normal-C2 as reference, four tissues were common for 1440  $\text{cm}^{-1}$  (lipids) with 1439  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation), 1278  $\text{cm}^{-1}$  (proteins) with 1277  $\text{cm}^{-1}$ , and 563  $\text{cm}^{-1}$  with 562  $\text{cm}^{-1}$  whilst 1001  $\text{cm}^{-1}$  (phenylalanine) was commonly shared by three tissues. Having Normal-C3 as reference, three tissues were common for 1003  $\text{cm}^{-1}$  (phenylalanine) with 1002  $\text{cm}^{-1}$  (phenylalanine) and 1001  $\text{cm}^{-1}$  (phenylalanine), and 784  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA) with 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA).



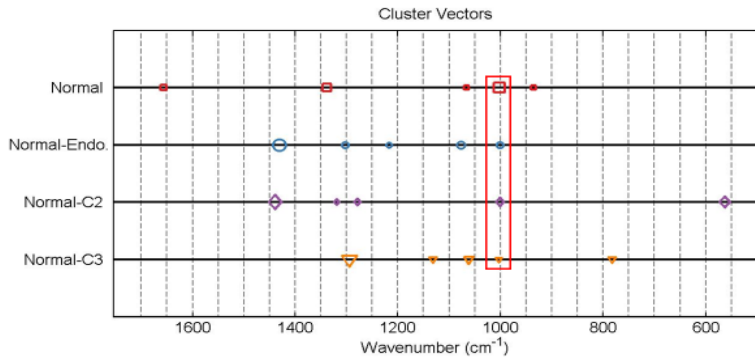




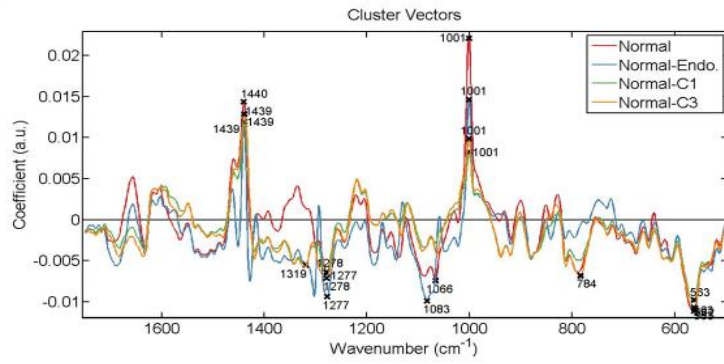
G



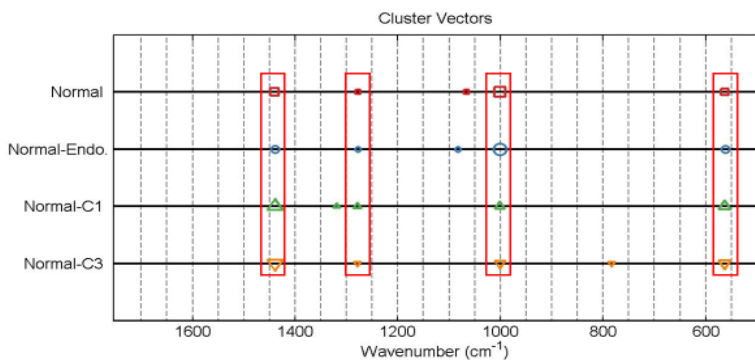
H

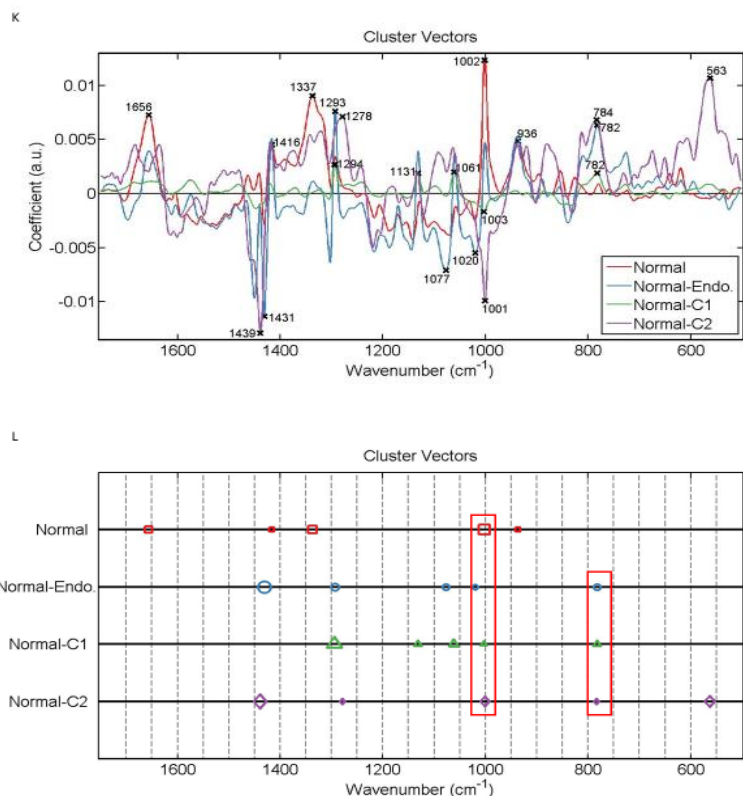


I



J



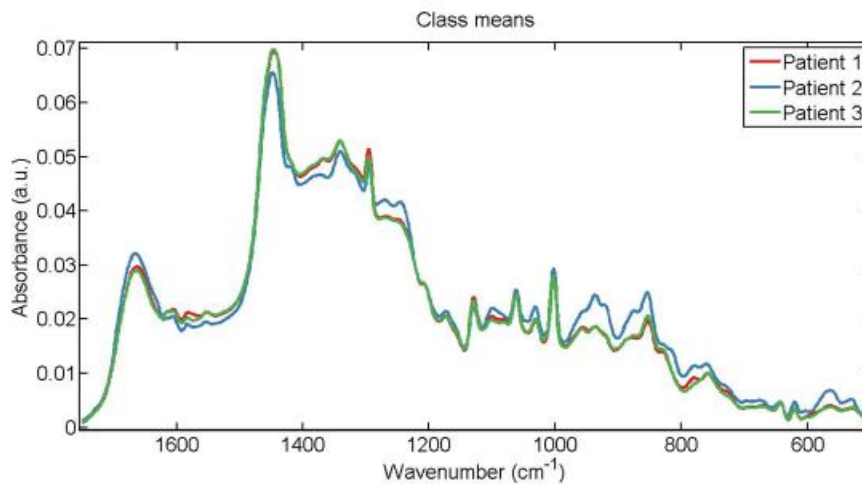


**Figure 3.156:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells, amongst non-diseased tissue sections. (A) & (B) no sample used as a reference, (C) & (D) Normal (H09-11708-4) reference class, (E) & (F) Normal-Endo. (H09-12292-1) reference class, (G) & (H) Normal-C1 (H09-12890-11) reference class, (I) & (J) Normal-C2 (H09-9102-6) reference class and (K) & (L) Normal-C3 (H09-13558-A8) reference class. The red rectangles in the cluster vector peak location plots highlight the wavenumbers occurring commonly in different

*Between non-diseased samples from patients with endometrial cancer*

Spectra from basal cells in only non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer were compared. Classes were labelled as Patient 1, Patient 2 and Patient 3 to represent the samples H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In this section classes of tissue samples will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots. All classes exhibited a very similar mean Raman spectra with only small variations in peak intensities occurring at  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1447\text{ cm}^{-1}$ ,  $\sim 1258\text{ cm}^{-1}$ ,  $\sim 937\text{ cm}^{-1}$  and  $\sim 853\text{ cm}^{-1}$  (Fig. 3.157).

Statistical analysis indicated that Patient 2 was significantly different from Patient 1 and Patient 3 whereas the two latter classes were not significantly different from each other (Fig. 3.158).

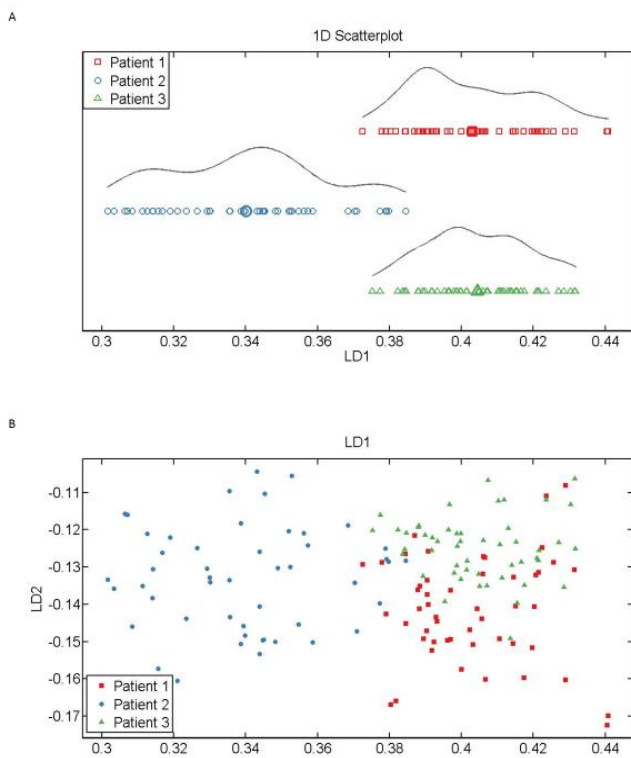


**Figure 3.157:** Mean Raman spectra for comparison of spectra acquired from basal cells in non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.001
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P > 0.05
Patient 2 vs Patient 3	P < 0.001

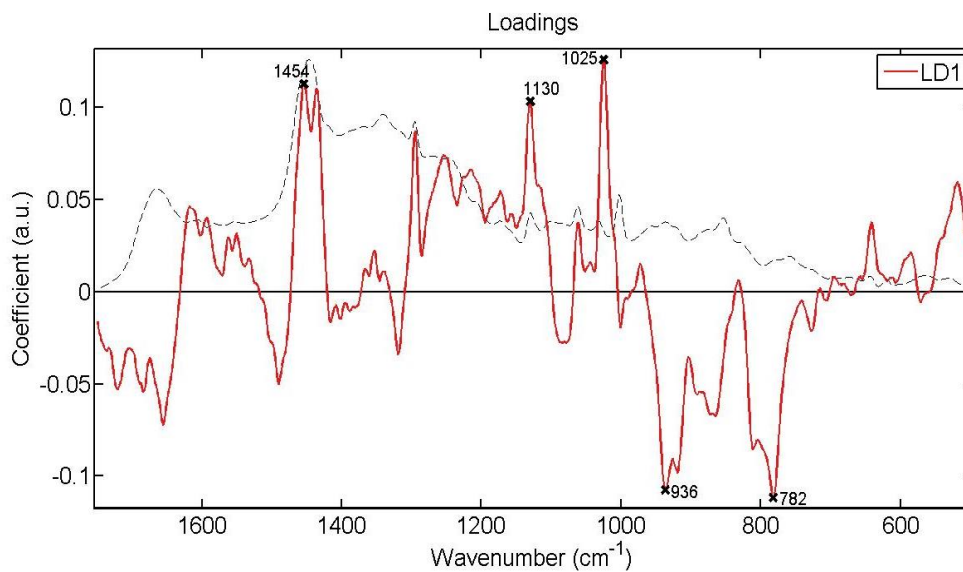
**Figure 3.158:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal cells between non-diseased tissue sections H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D and 2D space (Fig. 3.159). most overlapping occurred between Patient 1 and Patient 3 whilst spectral point of Patient 2 exhibited the most spread arrangement within the dimensional space which signified the most intra-class variation.



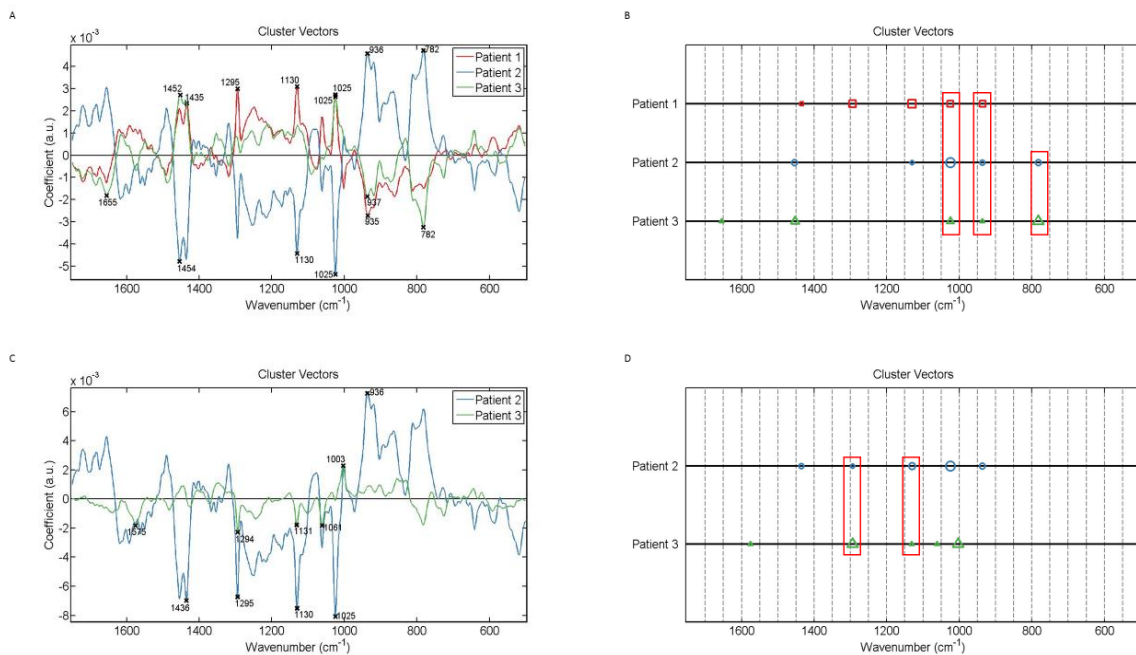
**Figure 3.159:** Scores plots for segregation of spectra in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

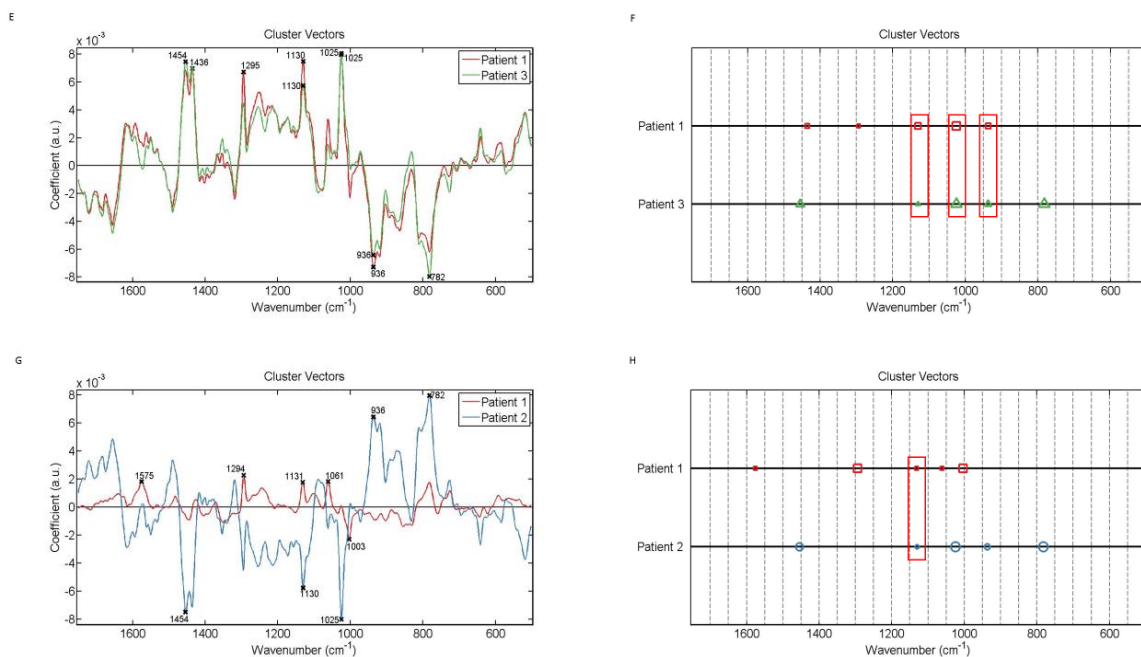
The five major discriminant wavenumbers identified in loading plots (Fig. 3.160) responsible for variations between tissue samples were  $1454\text{ cm}^{-1}$  (collagen and phospholipids),  $1130\text{ cm}^{-1}$  (lipids),  $1025\text{ cm}^{-1}$  (glycogen),  $936\text{ cm}^{-1}$  (valine/proline and protein) and  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA).



**Figure 3.160:** Loadings plots showing wavenumbers that discriminate basal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots were presented and analysed like previously as to observe occurrence and expression levels of wavenumbers and/or spectral regions in the hope of identifying biomarkers for basal cells (Fig. 3.161). Having no class as a reference origin, common occurrence was displayed by 1454  $\text{cm}^{-1}$  with 1452  $\text{cm}^{-1}$ , 1130  $\text{cm}^{-1}$  (lipids), 1025  $\text{cm}^{-1}$  (glycogen), and 937  $\text{cm}^{-1}$  (proline, protein backbone/glycogen) with 936  $\text{cm}^{-1}$  (valine/proline and protein) and 935  $\text{cm}^{-1}$  (proline, valine, protein backbone/glycogen). Using Patient 1 as reference the commonly shared wavenumbers were 1295  $\text{cm}^{-1}$  with 1294  $\text{cm}^{-1}$  (methylene twisting), and 1131  $\text{cm}^{-1}$  (lipids) with 1130  $\text{cm}^{-1}$  (lipids). Using Patient 2 as reference 1130  $\text{cm}^{-1}$  (lipids), 1025  $\text{cm}^{-1}$  (glycogen) and 936  $\text{cm}^{-1}$  (valine/proline and protein) were commonly shared. Using Patient 3 as reference 1131  $\text{cm}^{-1}$  (lipids) was commonly shared with 1130  $\text{cm}^{-1}$  (lipids) between classes of tissues.



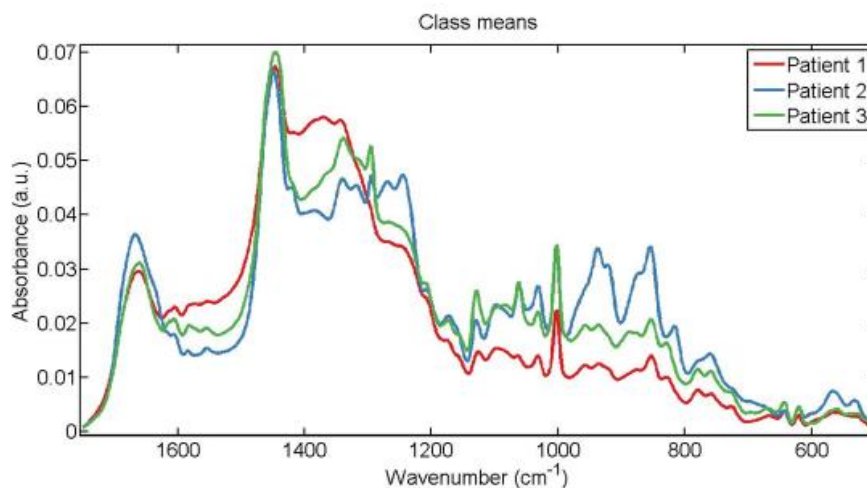


**Figure 3.161:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells among non-diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-11) reference class, (E) & (F) Patient 2 (H09-9102-6) reference class and (G) & (H) Patient 3 (H09-13558-A8) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Between diseased samples from patients with endometrial cancer*

Spectra from basal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) taken from cancerous lesions in the endometrium were compared. Classes were labelled as Patient 1, Patient 2 and Patient 3 to represent the samples H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In this section classes of tissues will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots. All classes exhibited a relatively similar mean spectra with pronounced variations in peak intensities occurring at  $1668\text{ cm}^{-1}$ ,  $1128\text{ cm}^{-1}$  (proteins/carbohydrates),  $1061\text{ cm}^{-1}$  (paraffin),  $1002\text{ cm}^{-1}$  (phenylalanine),  $939\text{ cm}^{-1}$ ,  $853\text{ cm}^{-1}$  (tyrosine/proline) and at the spectral region  $\sim 1300\text{ cm}^{-1}$  (Fig. 3.162). Lowest absorbance intensities were mainly exhibited by Patient 1.

Statistical analysis indicated that Patient 2 was significantly different from Patient 1 and Patient 3 whereas the two latter classes were not significantly different from each other (Fig. 3.163).



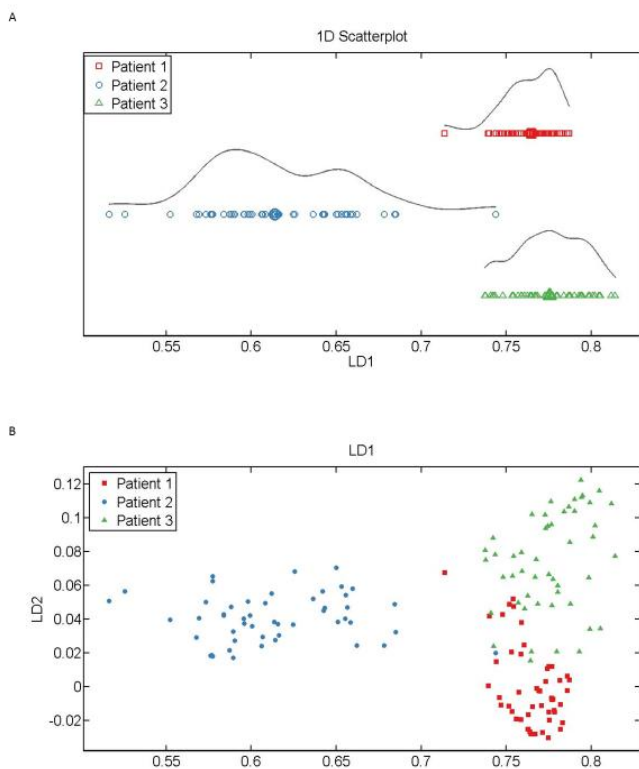
**Figure 3.162:** Mean Raman spectra for comparison of spectra from basal cells in diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.001
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P > 0.05
Patient 2 vs Patient 3	P < 0.001

**Figure 3.163:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal cells between diseased tissue sections H09-9102-1, H09-12890-9, H09-13558-A5 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

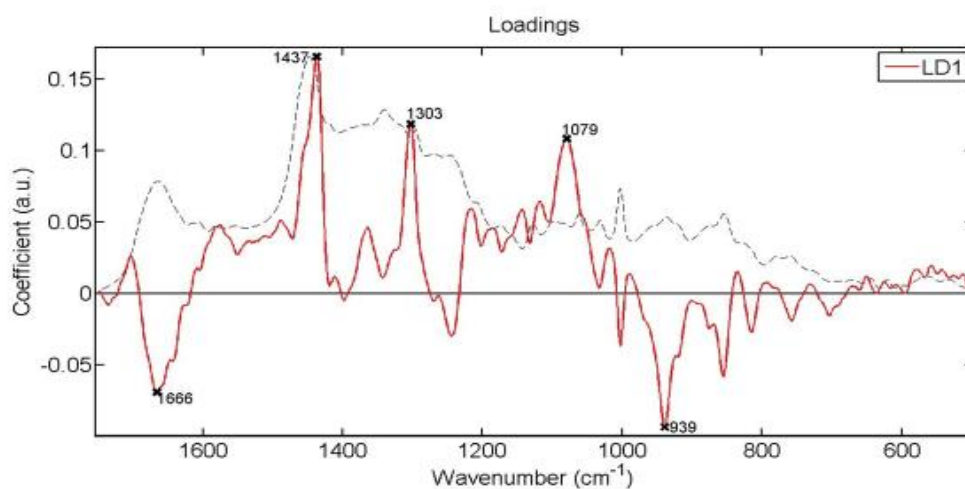
Segregation of classes resulting from application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.164). Minimal overlap was displayed by Patient 2 whilst more overlapping was displayed between Patient 1 and Patient 3. Relative more compact arrangement was exhibited by spectral points of Patient 1 which signified less intra-class variation.





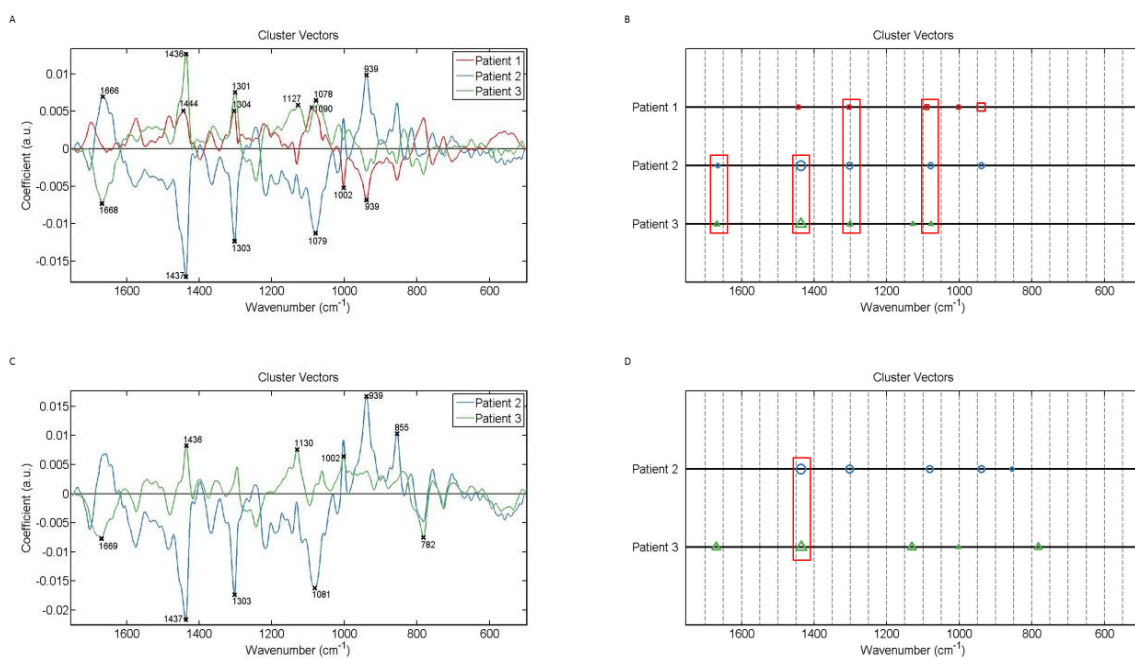
**Figure 3.164:** Scores plots for segregation of spectra acquired from basal cells in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

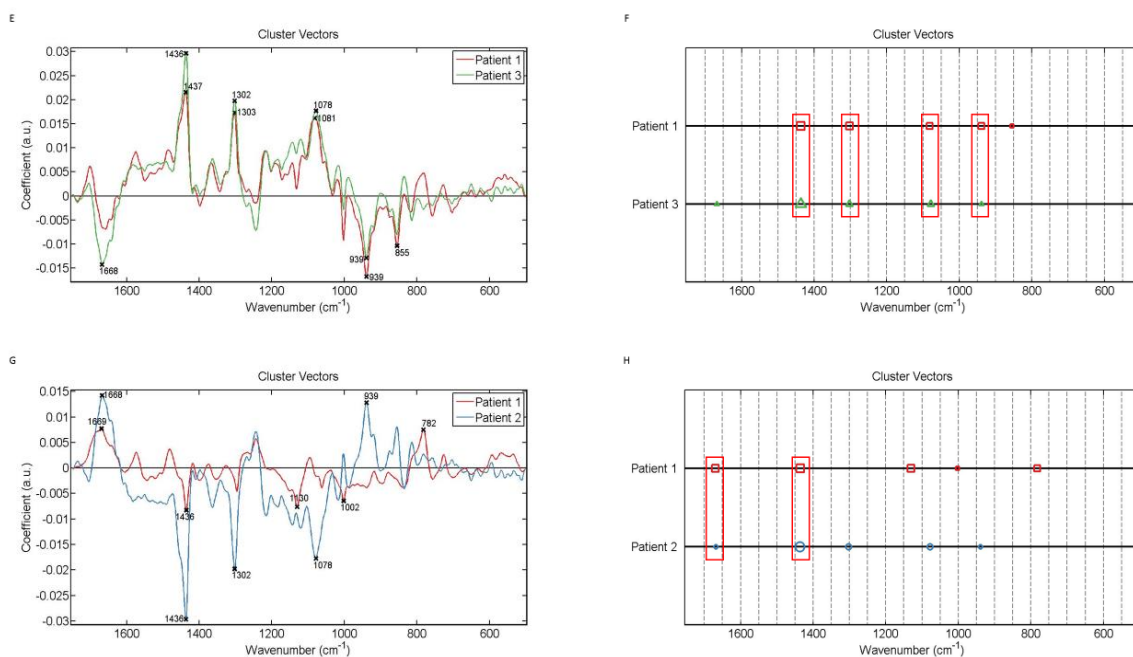
The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were  $1666\text{ cm}^{-1}$  (collagen),  $1437\text{ cm}^{-1}$  (proteins and lipids),  $1303\text{ cm}^{-1}$  (collagen),  $1079\text{ cm}^{-1}$  and  $939\text{ cm}^{-1}$  (Fig. 3.165).



**Figure 3.165:** Loadings plots showing wavenumbers that discriminate basal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

The same approach as before was applied to present and analyse cluster vector plots for identification of biomarkers (Fig. 3.166). Having no class as a reference, the wavenumbers observed to occur commonly between classes were 1668  $\text{cm}^{-1}$  with 1666  $\text{cm}^{-1}$  (collagen), 1437  $\text{cm}^{-1}$  (proteins and lipids) with 1436  $\text{cm}^{-1}$  (lipids), 1304  $\text{cm}^{-1}$  (lipids, adenine, cytosine) with 1303  $\text{cm}^{-1}$  (collagen) and 1301  $\text{cm}^{-1}$  (lipids) and 1079  $\text{cm}^{-1}$  with 1078  $\text{cm}^{-1}$  (phospholipids, nucleic acids). Using Patient 1 as reference, only 1437  $\text{cm}^{-1}$  (proteins and lipids) was commonly shared with 1436  $\text{cm}^{-1}$  (lipids). Using Patient 2 as reference, common occurrence was observed for 1437  $\text{cm}^{-1}$  (proteins and lipids) with 1436  $\text{cm}^{-1}$  (lipids), 1303  $\text{cm}^{-1}$  (collagen) with 1302  $\text{cm}^{-1}$  (protein/collagen/lipids), 1081  $\text{cm}^{-1}$  (lipids) with 1078  $\text{cm}^{-1}$  (phospholipids/nucleic acids), and 939  $\text{cm}^{-1}$ . Using Patient 3 as a reference, 1669  $\text{cm}^{-1}$  (lipids) with 1668  $\text{cm}^{-1}$ , and 1463  $\text{cm}^{-1}$  (lipids) were found to be common between the classes.





**Figure 3.166:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells among diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-9) reference class, (E) & (F) Patient 2 (H09-9102-1) reference class and (G) & (H) Patient 3 (H09-13558-A5) reference class. The red rectangles highlight the wavenumbers occurring

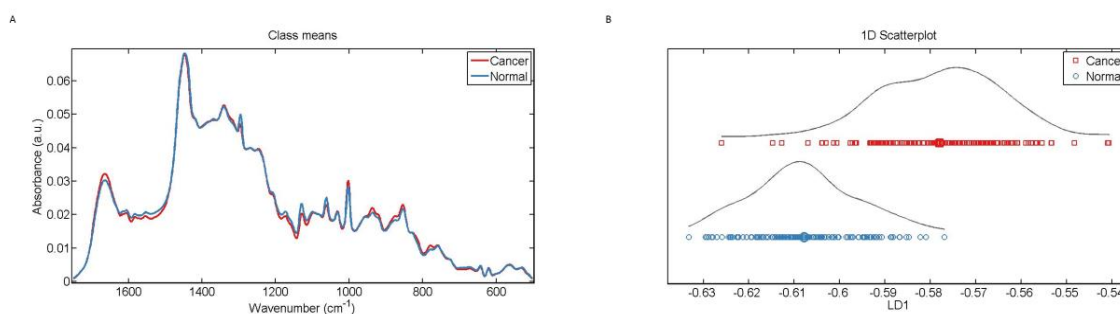
### *Diseased tissues Vs Non-Diseased tissues*

The spectra of the basal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) was compared with the spectra of basal cells in their corresponding non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer. The class representing the spectra collected from basal cells in diseased tissue sections was labelled as ‘Cancer’ whilst the class representing non-diseased tissue sections was labelled as ‘Normal’. In this section classes of spectra will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots.

Both classes exhibited relatively the same mean Raman spectra with the Cancer class having slightly higher peak intensities at  $1664\text{ cm}^{-1}$  (amide I),  $1003\text{ cm}^{-1}$  (phenylalanine),  $936\text{ cm}^{-1}$  (valine/proline and protein) and  $854\text{ cm}^{-1}$  (tyrosine/proline) whilst the Normal class had slighter higher intensities at  $1130$

$\text{cm}^{-1}$  (lipids),  $1296 \text{ cm}^{-1}$  ( $\text{CH}_2$  deformation) and  $1062 \text{ cm}^{-1}$  (paraffin) (Fig. 3.167 A). Application of PCA-LDA resulted in segregation of classes which was visualised in scores plot of 1D space whereby about half of the spectra of Cancer overlapped with spectra of Normal (Fig. 3.167 B).

Statistical analysis indicated that classes were significantly different from each other (Fig. 3.168).

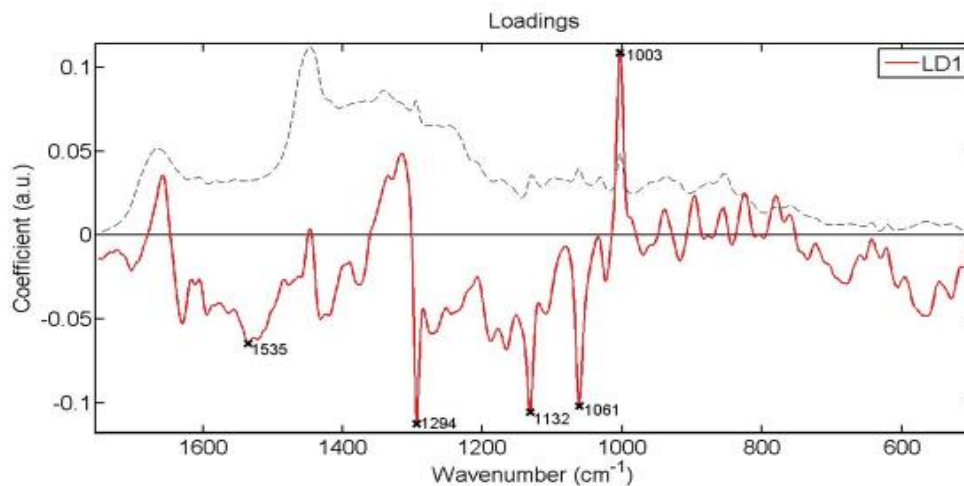


**Figure 3.167:** (A) Mean Raman spectra for comparison of spectra from basal cells in all diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5) and their corresponding non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). (B) Scores plots for segregation of spectra in 1D space. Diseased and non-diseased tissue samples are represented by the class labels 'Cancer' (red) and 'Normal' (blue) respectively.

Parameters	P Value
Cancer Vs Normal	$P < 0.0001$

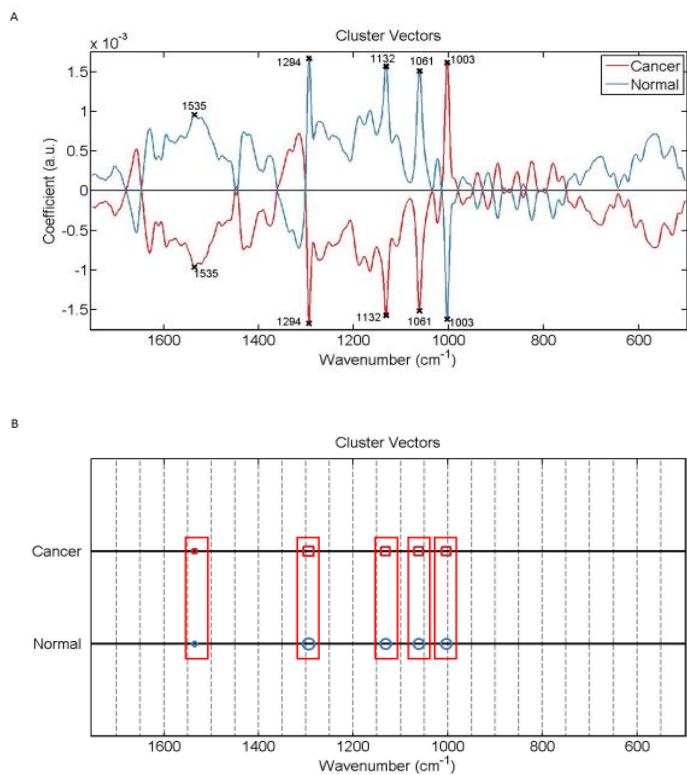
**Figure 3.168:** Obtained p-value by employment of Unpaired t-test to compare spectra acquired from basal cells between non diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and their corresponding diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections represented by the class labels Normal and Cancer respectively.

Generation of loading plots revealed five major wavenumbers responsible for variations between the classes;  $1535 \text{ cm}^{-1}$ ,  $1294 \text{ cm}^{-1}$  (methylene twisting),  $1132 \text{ cm}^{-1}$  (proteins and lipids),  $1061 \text{ cm}^{-1}$  (residual paraffin) and  $1003 \text{ cm}^{-1}$  (phenylalanine) (Fig.3.169).



**Figure 3.169:** Loadings plots showing wavenumbers that discriminate spectra from basal cells in all diseased (H09-9102-1, H09-12890-9, H09-13558-A5) and their corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) tissue sections. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

The five major wavenumbers in loading plots were also observed in cluster vector plots occurring in both classes (Fig. 3.170). Also concentrations of biomolecules assigned to these wavenumbers were found to be the same between the classes. Lowest expression levels were exhibited by 1535  $\text{cm}^{-1}$ . Biomarkers that would specify either 'Normal' or 'Cancer' basal cells could not be identified since the identified discriminant wavenumbers displayed common occurrence between the classes. For a wavenumber to have a potential as a biomarker for either of the classes, should be observed in only one class.



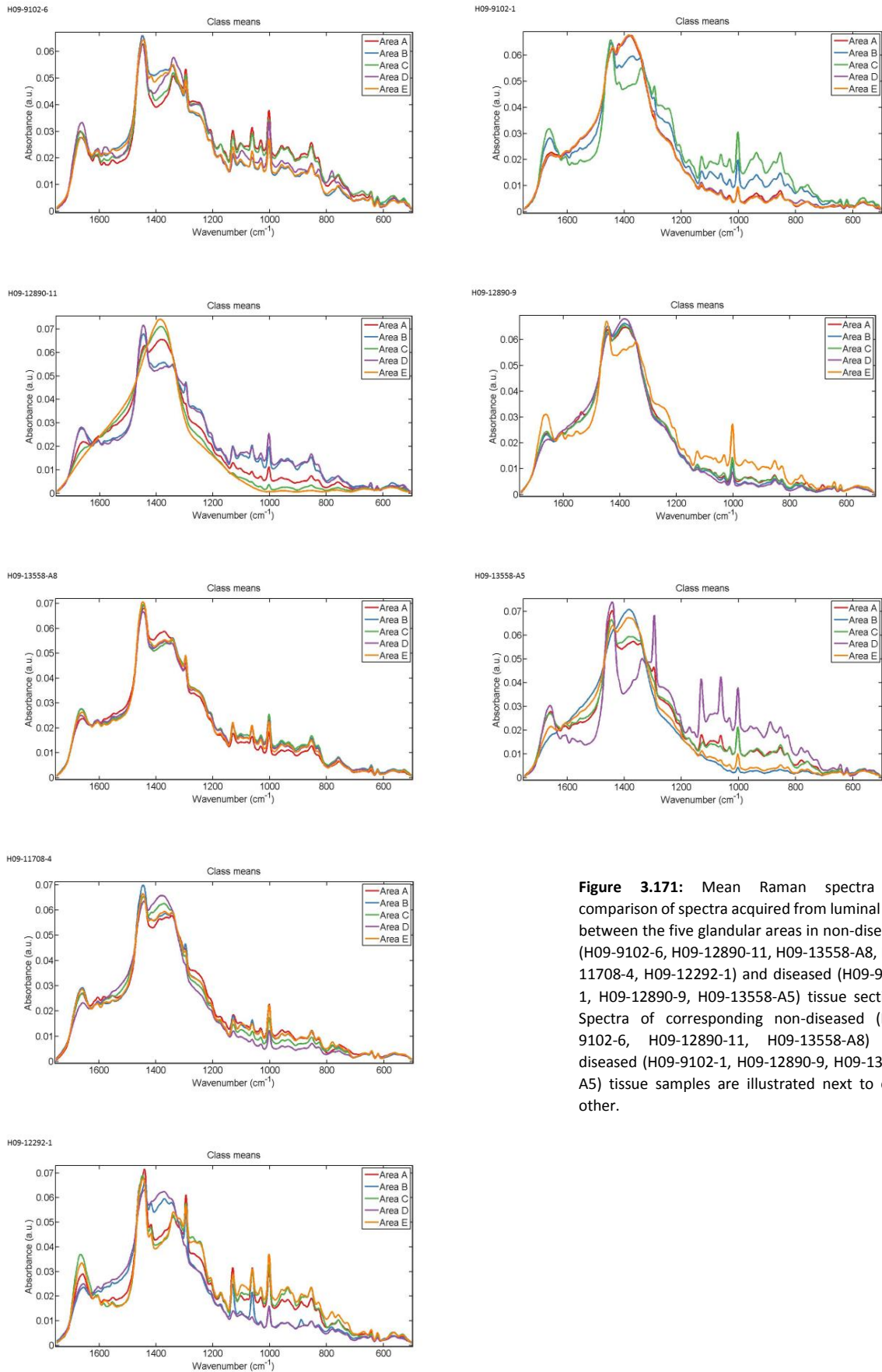
**Figure 3.170:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from basal cells among all corresponding diseased (H09-12890-9, H09-9102-1, H09-13558-A5) and non-diseased (H09-12890-11, H09-9102-6, H09-13558-A8) tissue sections. The red rectangles highlight the wavenumbers occurring commonly in classes.

### 3.3.2 Luminal Cells Vs Luminal Cells

#### *In individual tissue samples*

Spectra of luminal cells were also analysed in the hope of identifying potential biomarkers. Data was analysed in the same approach as applied for basal cells. Tissue samples exhibited a relatively similar mean Raman spectra with most evident variations in absorbance intensities occurring at  $\sim 1660\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1131\text{ cm}^{-1}$ ,  $\sim 1060\text{ cm}^{-1}$ ,  $1003\text{ cm}^{-1}$ ,  $936\text{ cm}^{-1}$  and  $\sim 854\text{ cm}^{-1}$  (Fig. 3.171). These observations were made from both types of tissue samples, either diseased or non-diseased. No significant observations were made that would correlate corresponding diseased and non-diseased tissue samples.

Statistical analysis (Fig. 3.172) indicated that in the majority of the tissues at least one area was significantly different from the tissue as a whole. Also in the majority of the tissues most areas were significantly different from each other.



**Figure 3.171:** Mean Raman spectra for comparison of spectra acquired from luminal cells between the five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.



H09-9102-6

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

H09-12890-11

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P < 0.001
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.001
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

H09-13558-A8

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P < 0.001
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.05
Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.01
Area B vs Area E	P < 0.001
Area C vs Area D	P > 0.05
Area C vs Area E	P < 0.05
Area D vs Area E	P > 0.05

H09-11708-4

Parameters	P value
LDA1 vs Area A	P < 0.01
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001
Area B vs Area C	P < 0.01
Area B vs Area D	P < 0.01
Area B vs Area E	P < 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

H09-12292-1

Parameters	P value
LDA1 vs Area A	P < 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P < 0.001
Area D vs Area E	P > 0.05

H09-9102-1

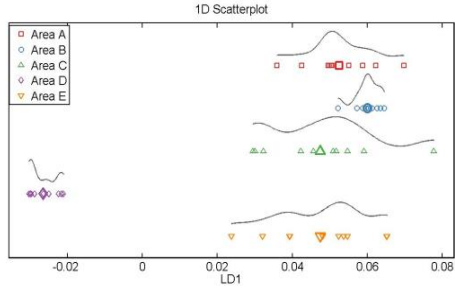
Parameters	P value
LDA1 vs Area A	P < 0.01
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.01
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.05
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P < 0.001
Area D vs Area E	P > 0.05

H09-12890-9		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Area A	P > 0.05	LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05	LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05	LDA1 vs Area C	P < 0.01
LDA1 vs Area D	P < 0.001	LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P < 0.001	LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05	Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05	Area A vs Area C	P > 0.05
Area A vs Area D	P > 0.05	Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001	Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05	Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.01	Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.05	Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001	Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05	Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001	Area D vs Area E	P < 0.001

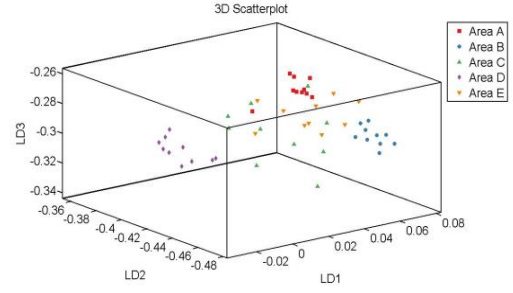
**Figure 3.172:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal in five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Segregation of classes of spectra, resulted from application of PCA-LDA, was visualised in scores plot of 1D and 3D space. Figure 3.173 illustrates scores plots of non-diseased tissue samples and figure 3.174 illustrates scores plots of diseased tissue samples. In all tissue samples a single glandular area would overlap with at least three or four other areas, Only in samples H09-9102-6, H09-12292-1 and H09-13558-A5 a single area would not overlap with the other areas. In the majority of samples, spectral points within individual glandular areas displayed a relatively spread arrangement rather than compact. No specific observations were made that could correlate the corresponding non-diseased and diseased tissue samples.

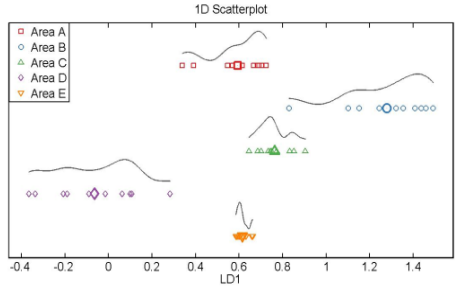
H09-9102-6



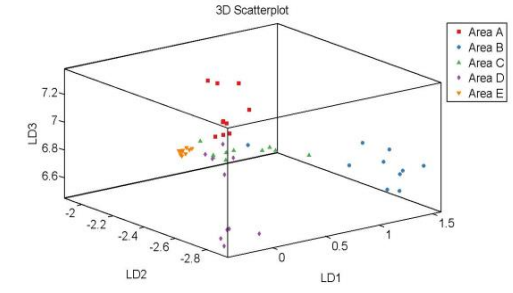
H09-9102-6



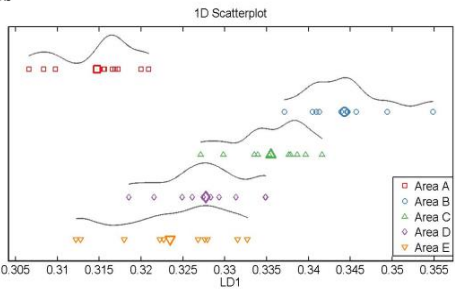
H09-12890-11



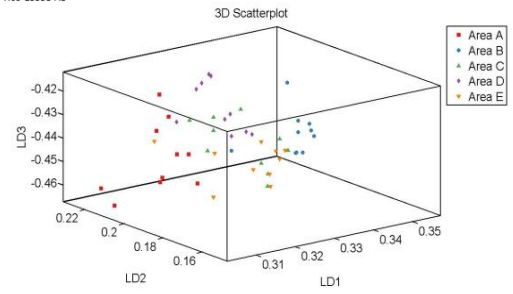
H09-12890-11



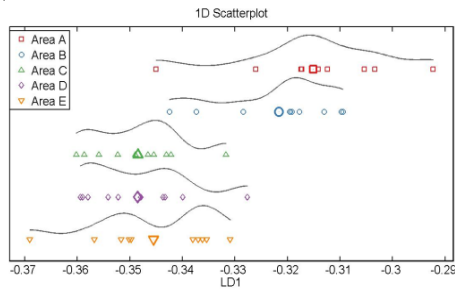
H09-13558-A8



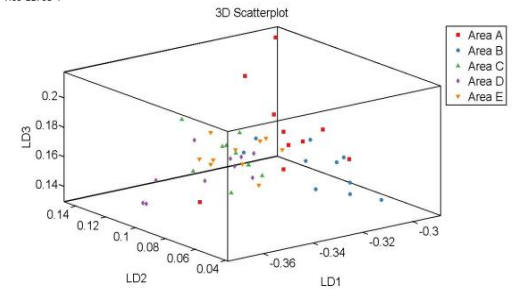
H09-13558-A8

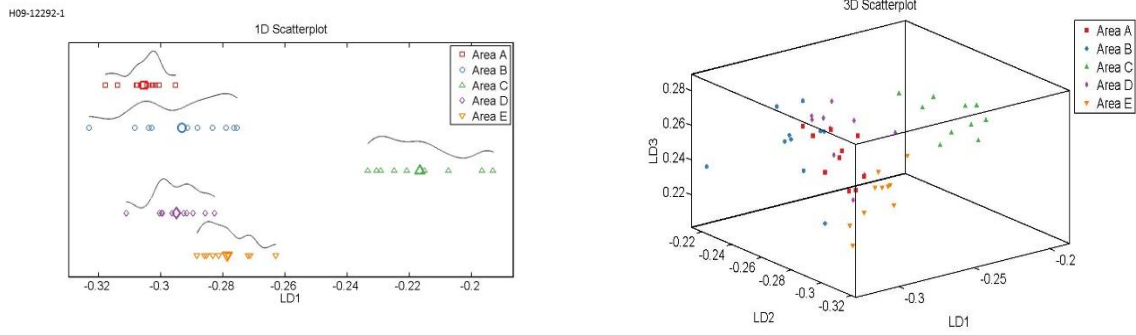


H09-11708-4

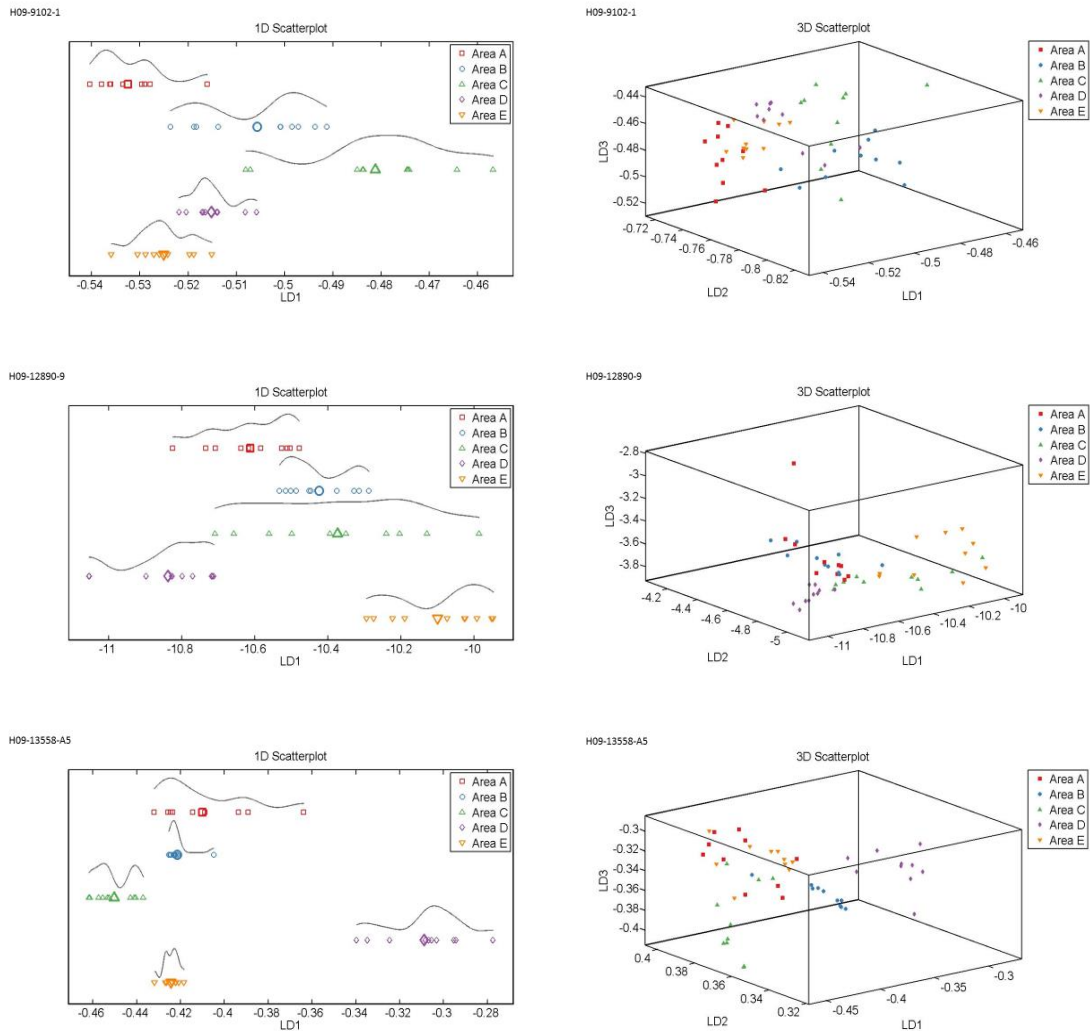


H09-11708-4



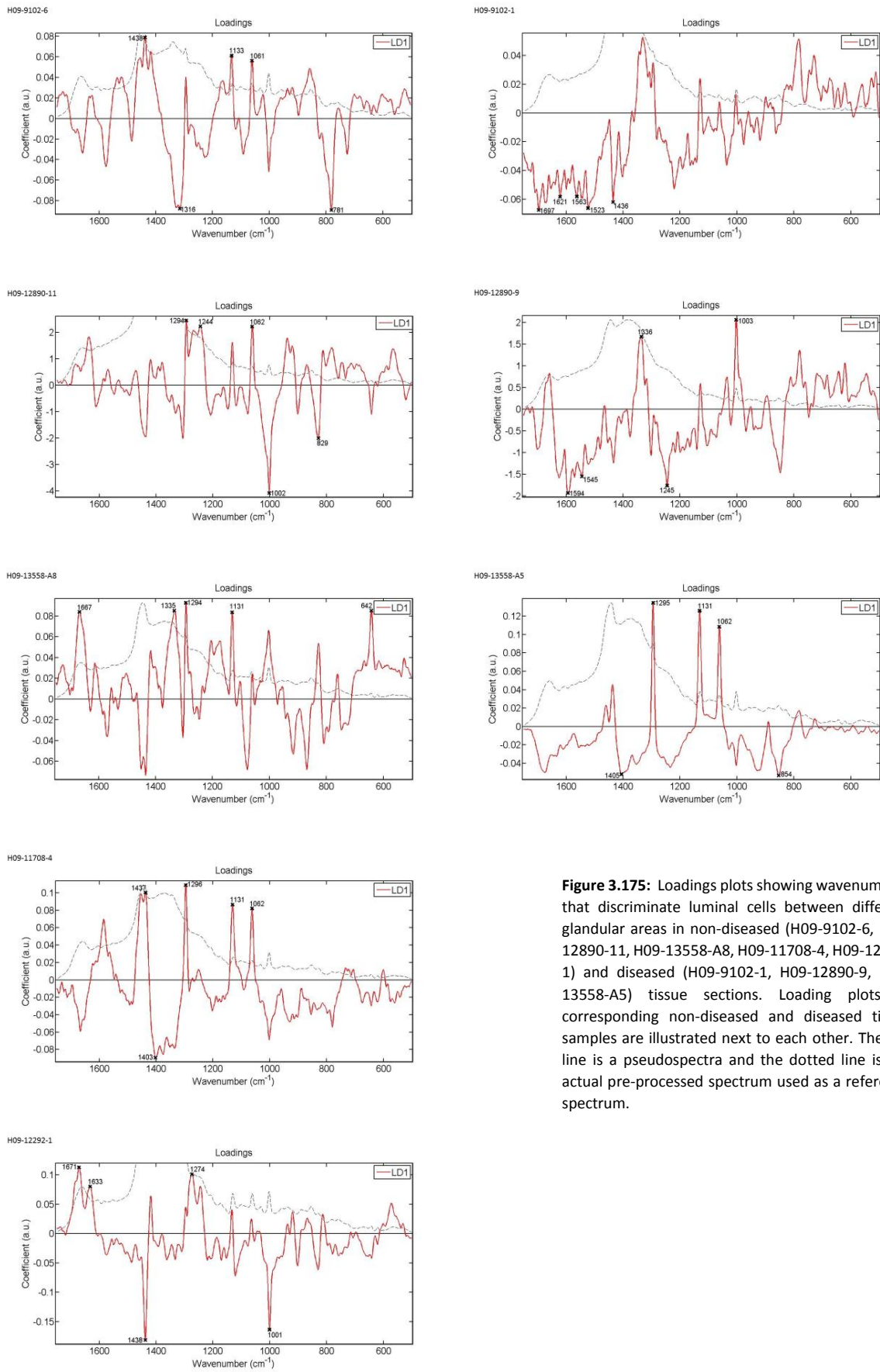


**Figure 3.173:** 1D scores plots and their corresponding 3D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from luminal cells from five different glandular elements.



**Figure 3.174:** 1D scores plots and their corresponding 3D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from luminal cells from five different glandular elements.

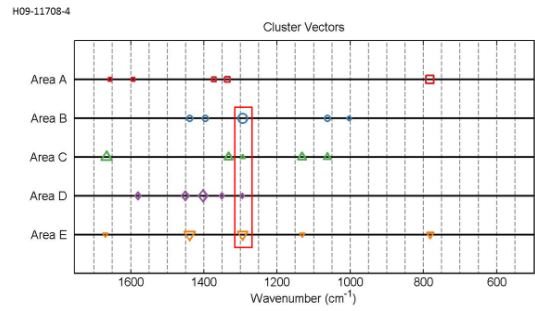
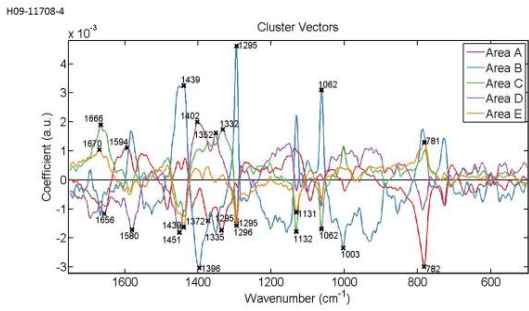
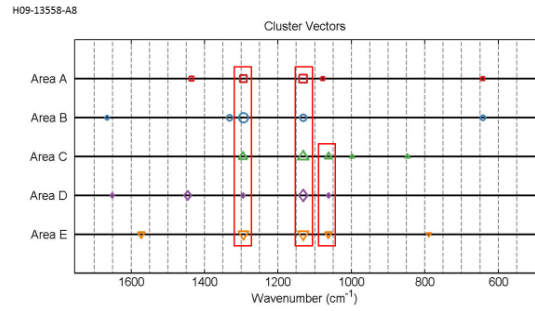
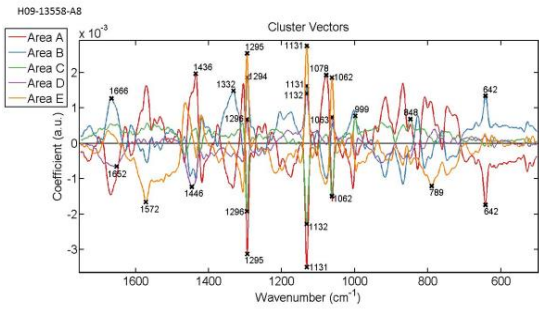
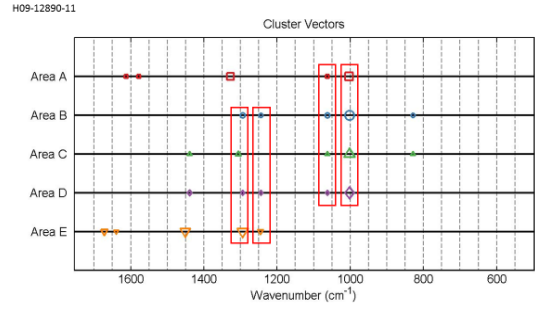
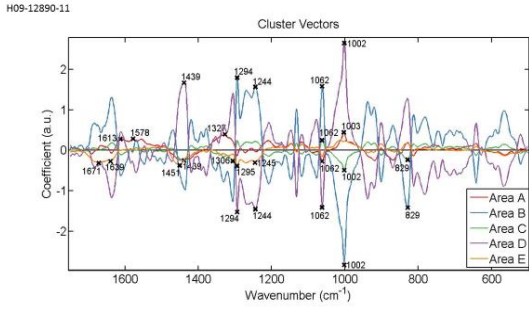
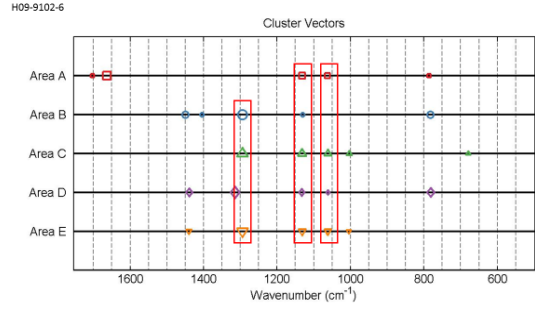
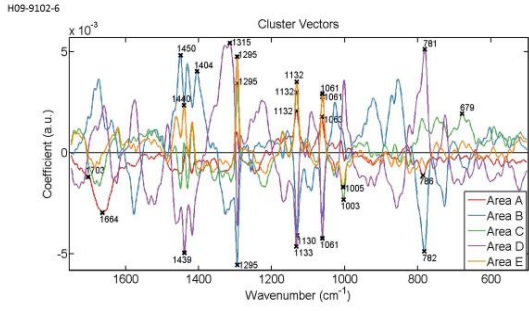
Figure 3.175 illustrates the loading plot for each tissue sample revealing the discriminant wavenumbers responsible for variations between the spectra acquired from luminal cells at glandular areas. In tissue sample H09-9102-6 the discriminant wavenumbers and their assignments were as follows; 1438  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation), 1316  $\text{cm}^{-1}$ , 1133  $\text{cm}^{-1}$ , 1061  $\text{cm}^{-1}$  (paraffin) and 781  $\text{cm}^{-1}$  (cytosine/uracil). In sample H09-12890-11 the discriminant wavenumbers were 1294  $\text{cm}^{-1}$  (methylene twisting), 1244  $\text{cm}^{-1}$ , 1062  $\text{cm}^{-1}$  (paraffin), 1002  $\text{cm}^{-1}$  (phenylalanine) and 829  $\text{cm}^{-1}$ . For tissue H09-13558-A8 the discriminant wavenumbers were 1667  $\text{cm}^{-1}$  (protein, amide I;  $\alpha$ -helix), 1335  $\text{cm}^{-1}$  (collagen/nucleic acids) 1294  $\text{cm}^{-1}$  (methylene twisting), 1131  $\text{cm}^{-1}$  (lipids) and 642  $\text{cm}^{-1}$  (tyrosine). For tissue H09-11708-4 the wavenumbers were 1437  $\text{cm}^{-1}$  (proteins and lipids), 1403  $\text{cm}^{-1}$ , 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation), 1131  $\text{cm}^{-1}$  (lipids) and 1062  $\text{cm}^{-1}$  (paraffin). For tissue H09-12292-1 the discriminant wavenumbers were 1671  $\text{cm}^{-1}$  (amide I), 1663  $\text{cm}^{-1}$  (DNA), 1438  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation), 1274  $\text{cm}^{-1}$  and 1001  $\text{cm}^{-1}$  (phenylalanine). In sample H09-9102-1 the discriminant wavenumbers were 1697  $\text{cm}^{-1}$  (amide I), 1621  $\text{cm}^{-1}$ , 1563  $\text{cm}^{-1}$ , 1523  $\text{cm}^{-1}$  and 1436  $\text{cm}^{-1}$  (lipids). In sample H09-12890-9 the identified wavenumbers were 1594  $\text{cm}^{-1}$ , 1545  $\text{cm}^{-1}$  ( $\text{C}_6\text{-H}$  deformation), 1336  $\text{cm}^{-1}$  (collagen/DNA), 1245  $\text{cm}^{-1}$  (amide III) and 1003  $\text{cm}^{-1}$  (phenylalanine). In sample H09-13558-A5 the discriminant wavenumbers were 1405  $\text{cm}^{-1}$ , 1295  $\text{cm}^{-1}$ , 1131  $\text{cm}^{-1}$  (lipids), 1062  $\text{cm}^{-1}$  (residual paraffin) and 854  $\text{cm}^{-1}$  (tyrosine/proline).



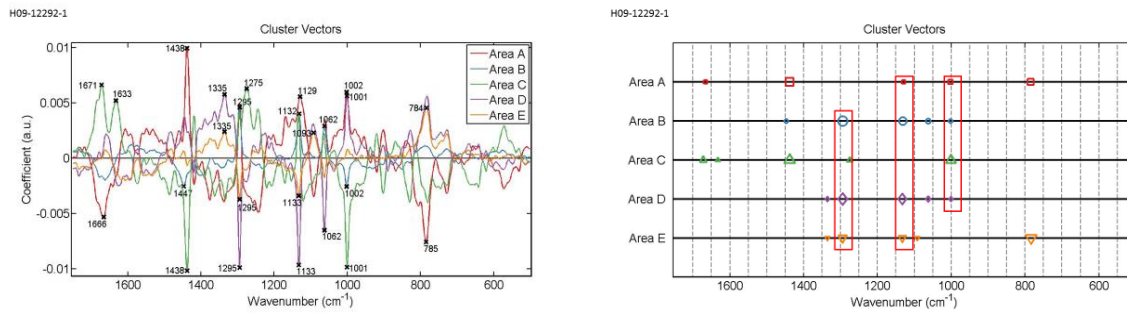
**Figure 3.175:** Loadings plots showing wavenumbers that discriminate luminal cells between different glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots of non-diseased (Fig. 3.176) and diseased (Fig. 3.177) tissue samples were presented and analysed in the same approach as before in the hope of identifying possible biomarkers for luminal cells, based on the occurrence of wavenumbers and their expression levels in the glandular areas. In most plots, the wavenumbers displaying common occurrence were the same as the discriminant wavenumbers identified in loading plots.

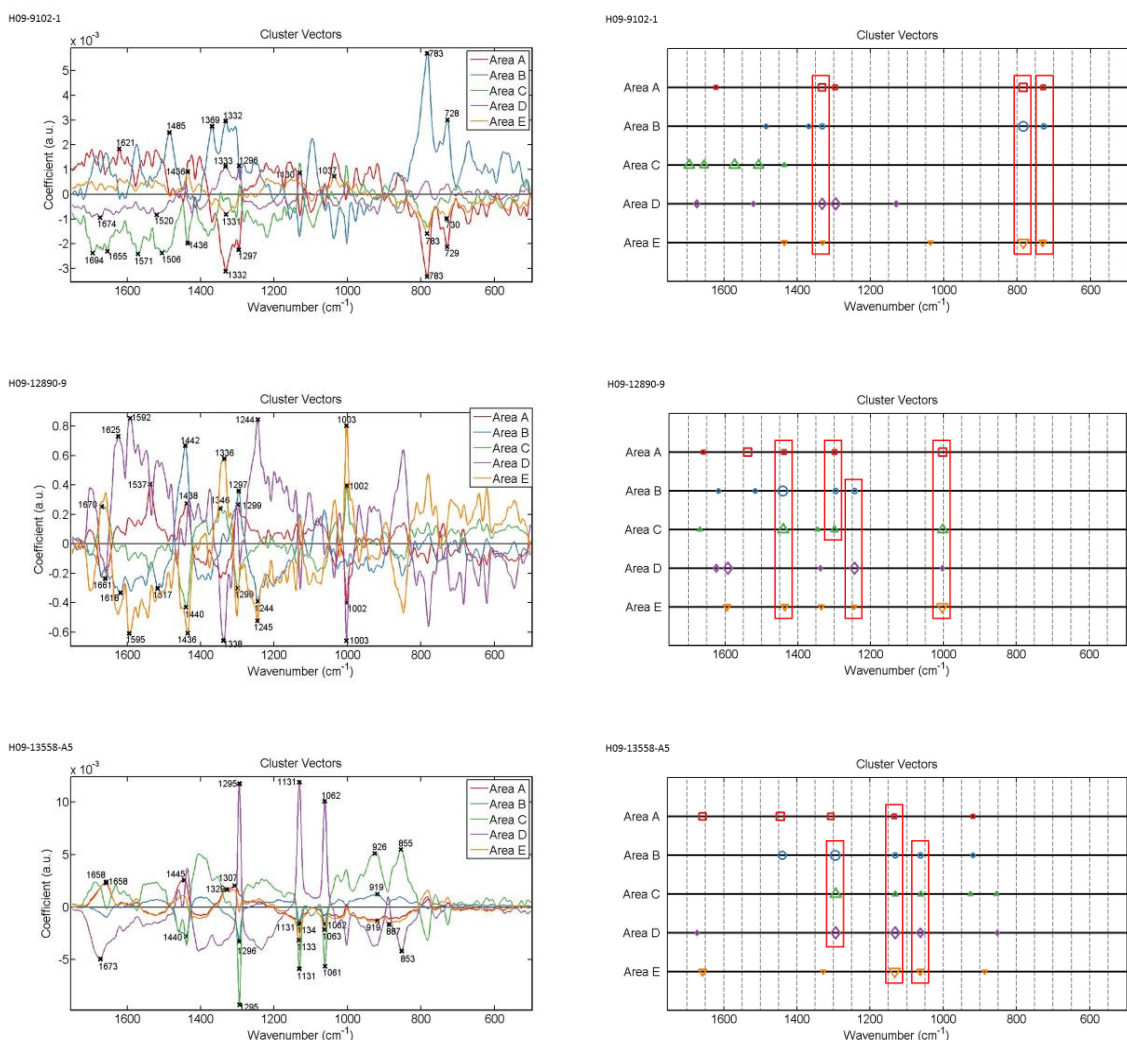
In sample H09-9102-6 three areas were common for 1295  $\text{cm}^{-1}$ , 1063  $\text{cm}^{-1}$  (C-C skeletal stretch) with 1061 whilst five areas were common for 1133  $\text{cm}^{-1}$  with 1132  $\text{cm}^{-1}$  (proteins and lipids) and 1130  $\text{cm}^{-1}$  (lipids). In sample H09-12890-11 three areas were common 1295  $\text{cm}^{-1}$  with 1294  $\text{cm}^{-1}$  (methylene twisting), and 1245  $\text{cm}^{-1}$  (amide III) with 1244  $\text{cm}^{-1}$  whilst four areas were common for 1062  $\text{cm}^{-1}$  (paraffin) and 1003  $\text{cm}^{-1}$  (phenylalanine) with 1002  $\text{cm}^{-1}$  (phenylalanine). In sample H09-13558-A8 common occurrence of 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) with 1295  $\text{cm}^{-1}$  and 1294  $\text{cm}^{-1}$  (methylene twisting) and 1132  $\text{cm}^{-1}$  (proteins and lipids) with 1131  $\text{cm}^{-1}$  (lipids) was displayed by five areas. In sample H09-11708-4 only 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) with 1295  $\text{cm}^{-1}$  were commonly shared between four areas. In sample H09-12292-1 three areas were common for 1295  $\text{cm}^{-1}$  whilst four areas were common for 1133  $\text{cm}^{-1}$  with 1132  $\text{cm}^{-1}$  (proteins and lipids) and 1129  $\text{cm}^{-1}$  (lipids), and 1002  $\text{cm}^{-1}$  (phenylalanine) with 1001  $\text{cm}^{-1}$  (phenylalanine). Four areas in sample H09-9102-1 were common for 1333  $\text{cm}^{-1}$  (guanine) with 1332  $\text{cm}^{-1}$  (phenyl) and 1331  $\text{cm}^{-1}$  whilst three areas were common for 783  $\text{cm}^{-1}$  and 730  $\text{cm}^{-1}$  with 729  $\text{cm}^{-1}$  (DNA/RNA bases) and 728  $\text{cm}^{-1}$  (proline-collagen). In sample 12890-9 four areas were common for 1442  $\text{cm}^{-1}$  with 1440  $\text{cm}^{-1}$ , 1438  $\text{cm}^{-1}$  and 1436  $\text{cm}^{-1}$ , and 1003  $\text{cm}^{-1}$  (phenylalanine) with 1002  $\text{cm}^{-1}$  (phenylalanine) whilst three areas were common for 1299  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation; lipids) with 1297  $\text{cm}^{-1}$ , and 1245  $\text{cm}^{-1}$  (amide III) with 1244  $\text{cm}^{-1}$ . In sample H09-13558-A5, 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) was commonly shared with 1295  $\text{cm}^{-1}$  in three areas whilst four areas were common for 1134  $\text{cm}^{-1}$  with 1133  $\text{cm}^{-1}$  and 1131  $\text{cm}^{-1}$  (lipids).







**Figure 3.176:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells, in five glandular areas in individual non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1). The red rectangles highlight the wavenumbers shared in common among glandular areas.

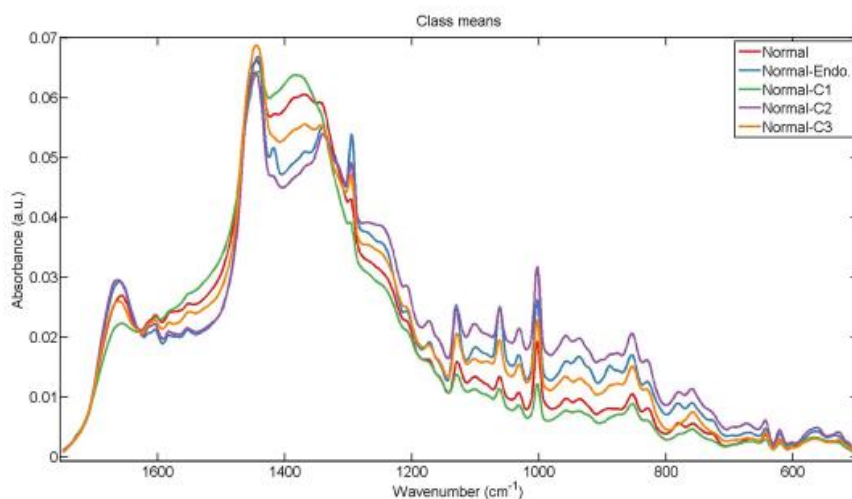


**Figure 3.177:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells, in five glandular areas in individual diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight the wavenumbers shared in common among glandular areas.

*Between all normal tissue samples*

All the non-diseased tissue samples were compared between them for the spectra interrogated from their luminal cells. Classes representing tissues were labelled as before; Normal (H09-11708-4), Normal-Endo. (H09-12292-1), Normal-C1 (H09-12890-11), Normal-C2 (H09-9102-6) and Normal-C3 (H09-13558-A8). From the mean Raman spectra it was observed that all classes displayed a similar shape in spectra with most pronounced variations in absorbance intensities occurring at  $\sim 1660\text{ cm}^{-1}$ ,  $\sim 1446\text{ cm}^{-1}$ ,  $\sim 1380\text{ cm}^{-1}$ ,  $\sim 1297\text{ cm}^{-1}$ ,  $\sim 1130\text{ cm}^{-1}$ ,  $1061\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 940\text{ cm}^{-1}$ ,  $853\text{ cm}^{-1}$  and  $757\text{ cm}^{-1}$  (Fig. 3.178). No class was distinguished in terms of maintaining higher or lower intensities throughout the interrogated spectra.

Statistical analysis (Fig. 3.179) indicated that classes were significantly different from each other for the spectra acquired from luminal cells but Normal-C1 was not significantly different from Normal neither from Normal-C3.

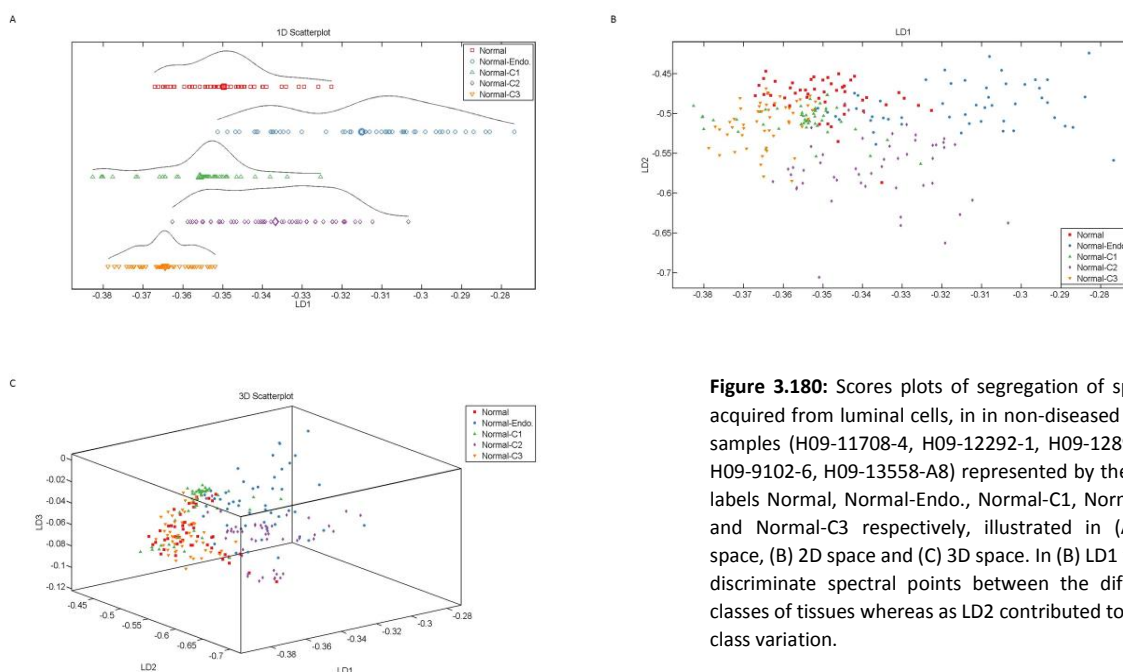


**Figure 3.178:** Mean Raman spectra for comparison of spectra acquired from luminal cells in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Parameters	P value
LDA1 vs Normal	P > 0.05
LDA1 vs Normal-Endo.	P < 0.001
LDA1 vs Normal-C1	P < 0.01
LDA1 vs Normal-C2	P > 0.05
LDA1 vs Normal-C3	P < 0.001
Normal vs Normal-Endo.	P < 0.001
Normal vs Normal-C1	P > 0.05
Normal vs Normal-C2	P < 0.01
Normal vs Normal-C3	P < 0.01
Normal-Endo. vs Normal-C1	P < 0.001
Normal-Endo. vs Normal-C2	P < 0.001
Normal-Endo. vs Normal-C3	P < 0.001
Normal-C1 vs Normal-C2	P < 0.001
Normal-C1 vs Normal-C3	P > 0.05
Normal-C2 vs Normal-C3	P < 0.001

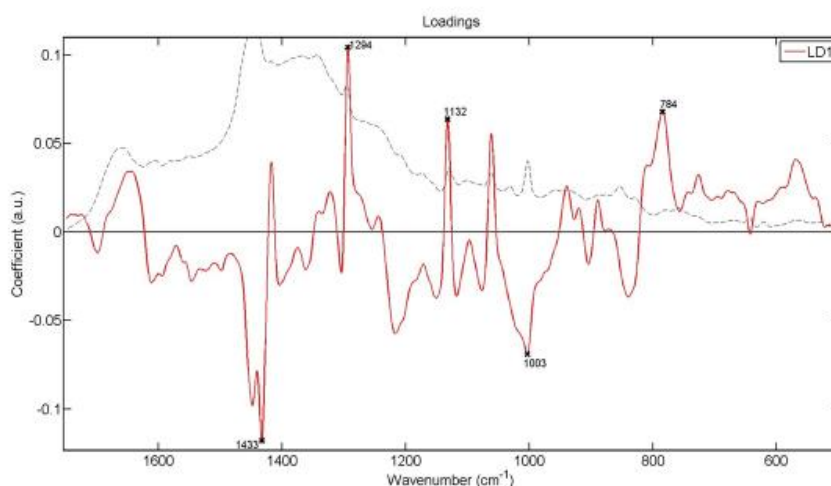
**Figure 3.179:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal cells in all five glandular elements between non-diseased tissue sections H09-11708-4, H09-9102-1, H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D, 2D and 3D space (Fig. 3.180). An individual class would overlap at some point throughout its spectra if not with all four other classes then at least with three other classes. Normal class displayed the least within class variation since its spectral points had a more compact arrangement relative to the other classes. Most evident overlap occurred between classes Normal, Normal-C1 and Normal-C3.



**Figure 3.180:** Scores plots of segregation of spectra acquired from luminal cells, in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively, illustrated in (A) 1D space, (B) 2D space and (C) 3D space. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

The discriminant wavenumbers identified in loading plots accounting for variations between the samples were 1294  $\text{cm}^{-1}$  (methylene twisting), 1433  $\text{cm}^{-1}$  (lipids), 1132  $\text{cm}^{-1}$  (proteins and lipids), 1003  $\text{cm}^{-1}$  (phenylalanine) and 784  $\text{cm}^{-1}$  (cytosine/uracil) (Fig. 3.181).

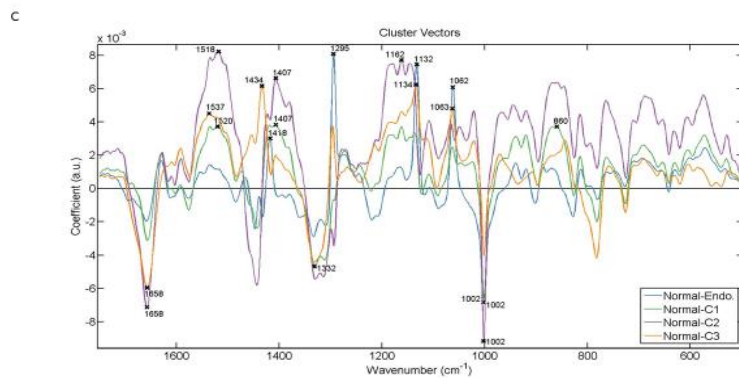
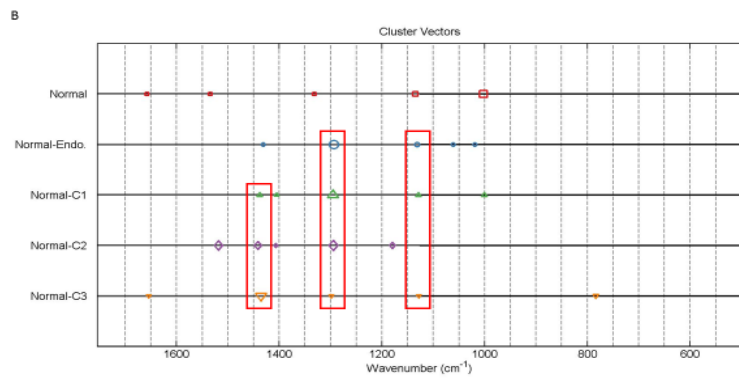
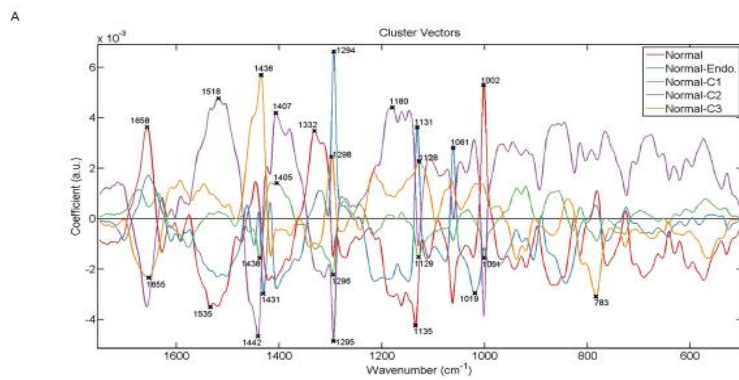


**Figure 3.181:** Loadings plots showing wavenumbers that discriminate luminal cells in non-diseased tissue sections (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

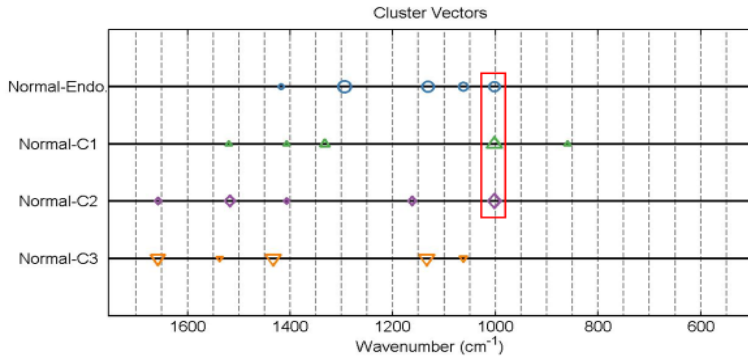
Cluster vector plots were presented and analysed in a different way in order to identify potential biomarkers for luminal cells by observing the occurrence and expression levels of wavenumbers and/or spectral regions between the tissues by having a class as a reference origin or when no class was used as a reference origin (Fig. 3.182). Wavenumbers would display common occurrence if they were observed in at least three classes representing the tissues as long as they had a smooth vertical alignment in the plots.

Without having any class as a reference origin, 1298  $\text{cm}^{-1}$  (fatty acids) with 1296  $\text{cm}^{-1}$  (residual paraffin), 1295  $\text{cm}^{-1}$  and 1294  $\text{cm}^{-1}$  (methylene twisting) were commonly shared by four tissues whilst 1442  $\text{cm}^{-1}$  (fatty acids,  $\text{CH}_2$  bending mode) with 1438  $\text{cm}^{-1}$  and 1436  $\text{cm}^{-1}$  (lipids), and 1131  $\text{cm}^{-1}$  (lipids) with 1129  $\text{cm}^{-1}$  (lipids) and 1128  $\text{cm}^{-1}$  (proteins/carbohydrates) were commonly shared by three tissues. Having class Normal as a reference three tissues were only common for 1002  $\text{cm}^{-1}$  (phenylalanine). Using class Normal-Endo. as reference, 1295  $\text{cm}^{-1}$  was commonly shared with 1294

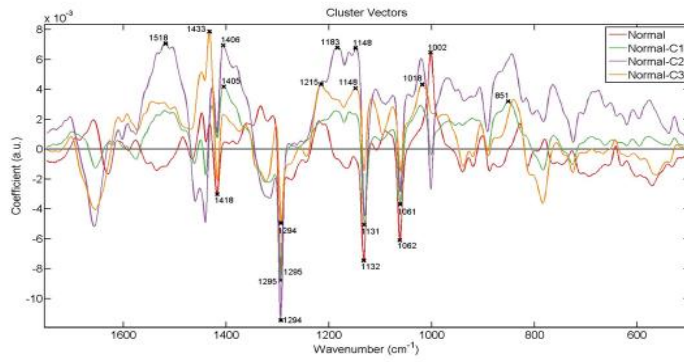
$\text{cm}^{-1}$  (methylene twisting) in four tissues. Having Normal-C1 as refernece, no wavenumbers showed a potetial as biomarkers since they were commonly shared by only two tissues. Using Normal-C2 as reference three tissues were common for  $1518 \text{ cm}^{-1}$  (porphyrin/carotenoid). Having Normal-C3 as reference,  $1436 \text{ cm}^{-1}$  (lipids) was commonly shared with  $1434 \text{ cm}^{-1}$  and  $1433 \text{ cm}^{-1}$  (lipids) by three tissues and  $1297 \text{ cm}^{-1}$  was commonly shared with  $1296 \text{ cm}^{-1}$  ( $\text{CH}_2$  deformation) and  $1294 \text{ cm}^{-1}$  (methylene twisting) by three tissues. In the majority of the plots some wavenumbers were shared by only two classes of tissues and thus displaying minimal potential as biomarkers for luminal cells.



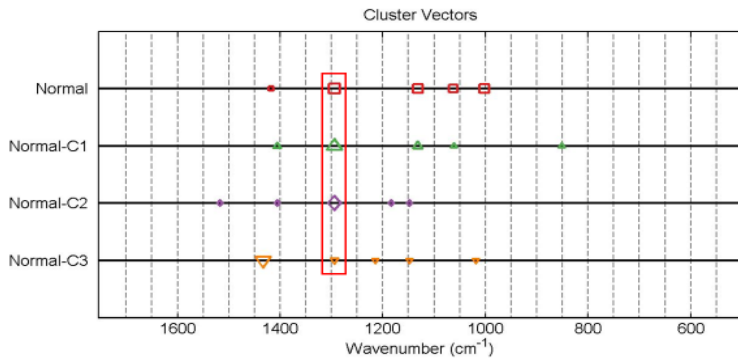
D



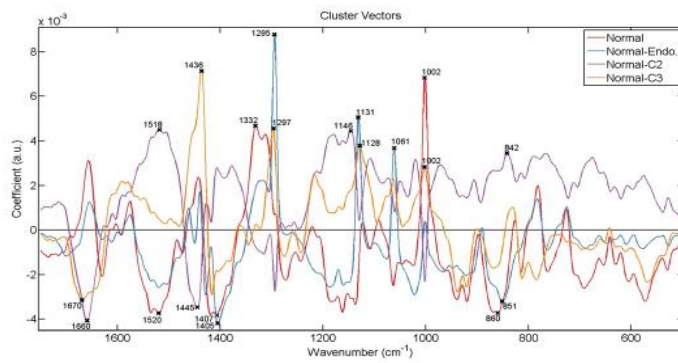
E



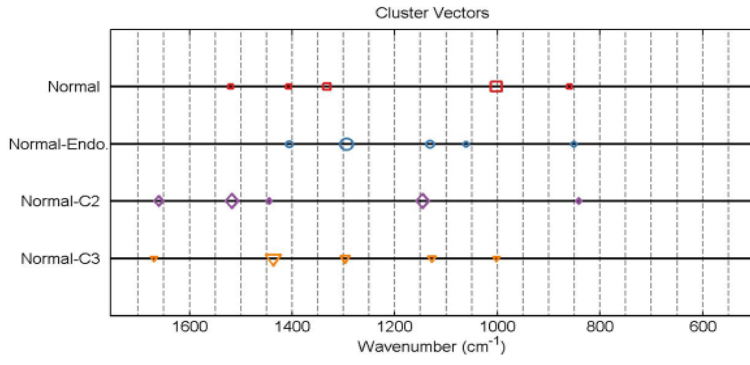
F



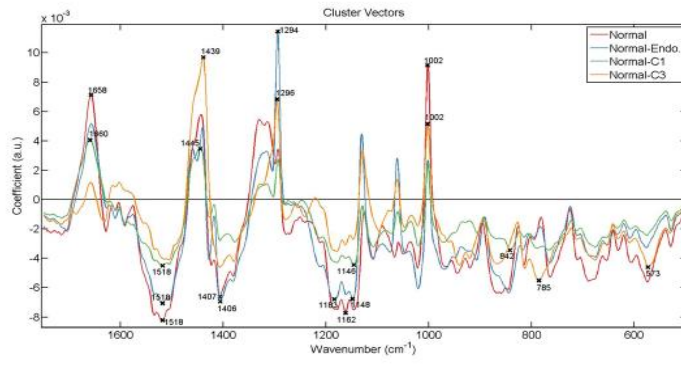
G



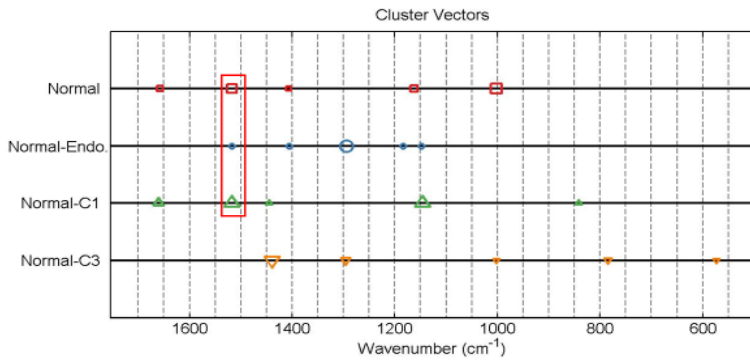
H

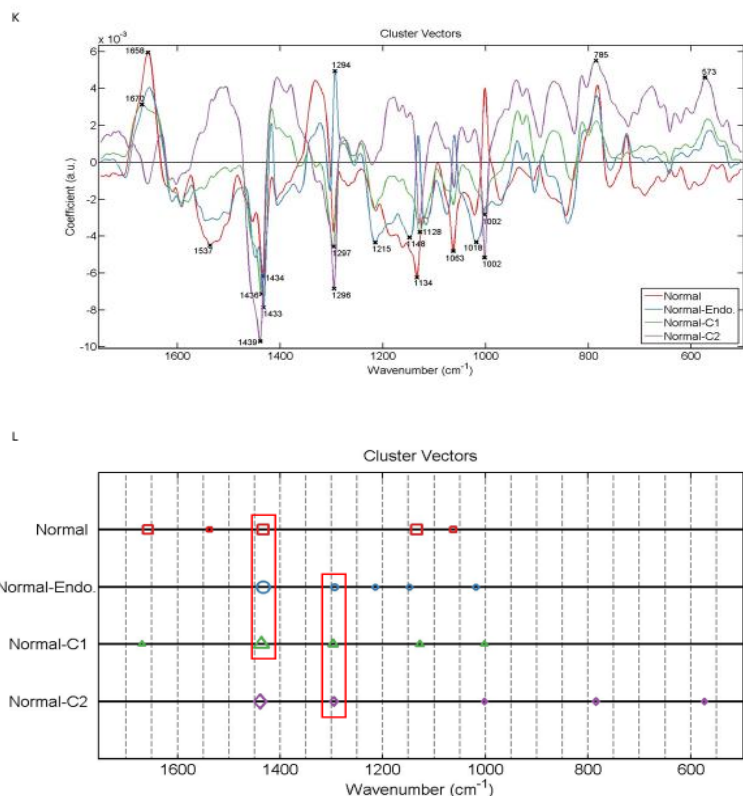


I



J





**Figure 3.182:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells, amongst non-diseased tissue sections. (A) & (B) no sample used as a reference, (C) & (D) Normal (H09-11708-4) reference class, (E) & (F) Normal-Endo. (H09-12292-1) reference class, (G) & (H) Normal-C1 (H09-12890-11) reference class, (I) & (J) Normal-C2 (H09-9102-6) reference class and (K) & (L) Normal-C3 (H09-13558-A8) reference class. The red rectangles in the cluster vector peak location plots highlight the wavenumbers occurring commonly in different tissue samples.

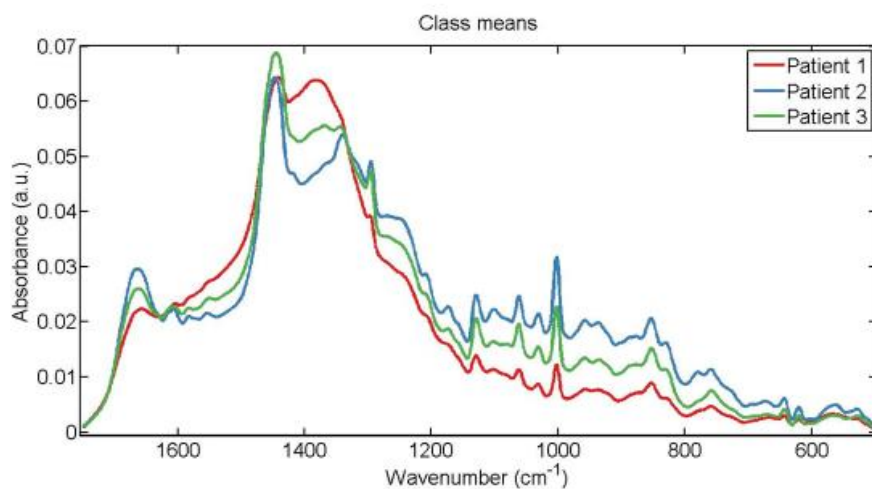
### *Between non-diseased samples from patients with endometrial cancer*

Spectra from luminal cells in only non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer were compared. Classes representing the samples were labelled as before so they will be referred accordingly to associate data analysis with what is shown on the plots.

All classes exhibited a relatively similar Raman mean spectra whereby most pronounced variations in absorbance intensities occurred at  $\sim 1663 \text{ cm}^{-1}$ ,  $\sim 1445 \text{ cm}^{-1}$ ,  $\sim 1380 \text{ cm}^{-1}$ ,  $\sim 1338 \text{ cm}^{-1}$ ,  $\sim 1295 \text{ cm}^{-1}$ ,  $\sim 1128 \text{ cm}^{-1}$ ,  $\sim 1062 \text{ cm}^{-1}$ ,  $\sim 1002 \text{ cm}^{-1}$ ,  $\sim 949 \text{ cm}^{-1}$ ,  $\sim 854 \text{ cm}^{-1}$  and  $\sim 757 \text{ cm}^{-1}$  (Fig. 3.183). Patient 2 was found to have the highest peak intensities except at  $\sim 1445 \text{ cm}^{-1}$  where Patient 3 had the highest peak. Patient 1 exhibited the lowest peak intensities except at  $\sim 1380 \text{ cm}^{-1}$  where the absorbance was the highest.



Statistical analysis indicated that Patient 2 was significantly different from Patient 1 and Patient 3 whereas the two latter classes were not significantly different from each other (Fig. 3.184).

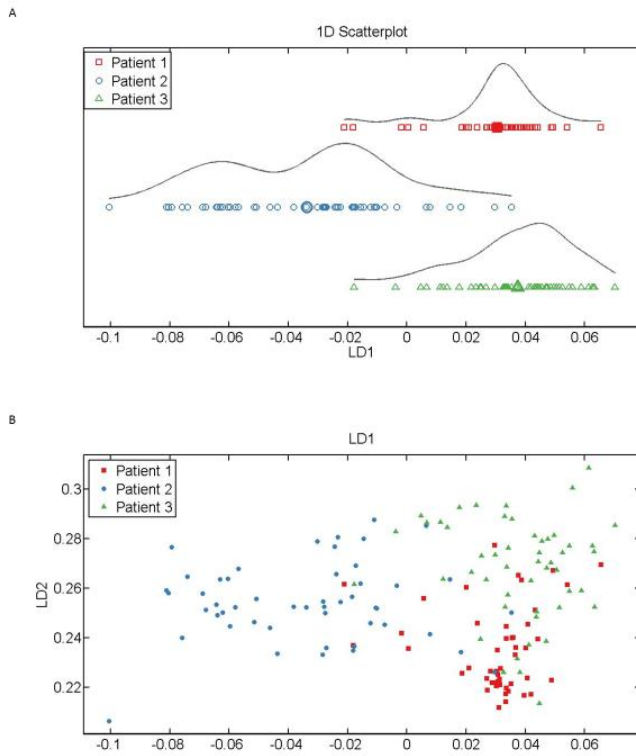


**Figure 3.183:** Mean Raman spectra for comparison of spectra acquired from luminal cells in non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.01
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P > 0.05
Patient 2 vs Patient 3	P < 0.001

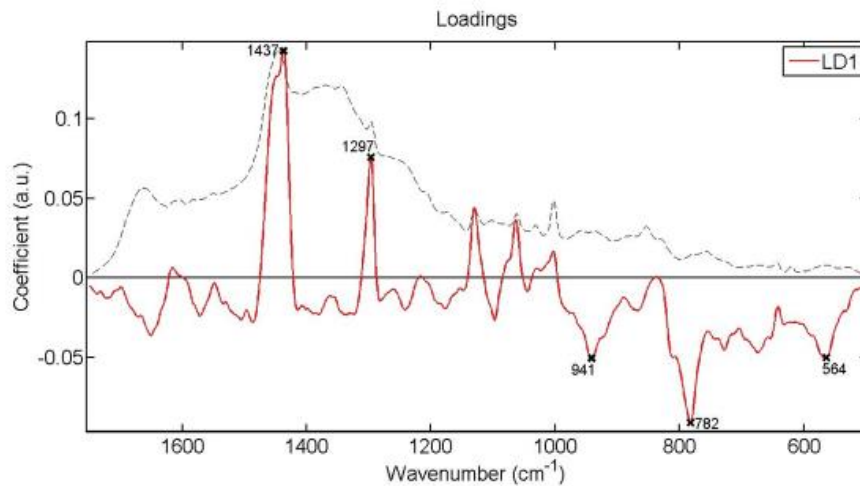
**Figure 3.184:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal cells between non-diseased tissue sections H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D and 2D space (Fig. 3.185). Patient 2 displayed the least overlap whilst Patient 1 and 3 would overlap almost throughout their entire spectra. Spectral points of patient 1 had a more compact arrangement, relative to the other classes, which signified the least intra-class variation.



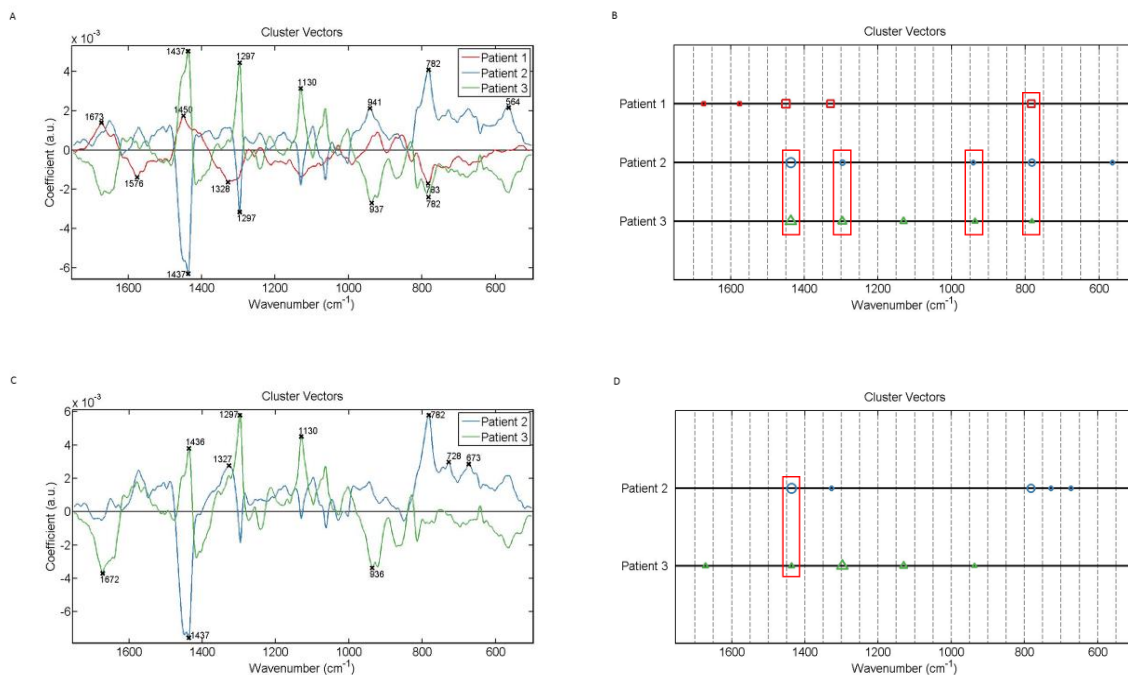
**Figure 3.185:** Scores plots for segregation of spectra in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

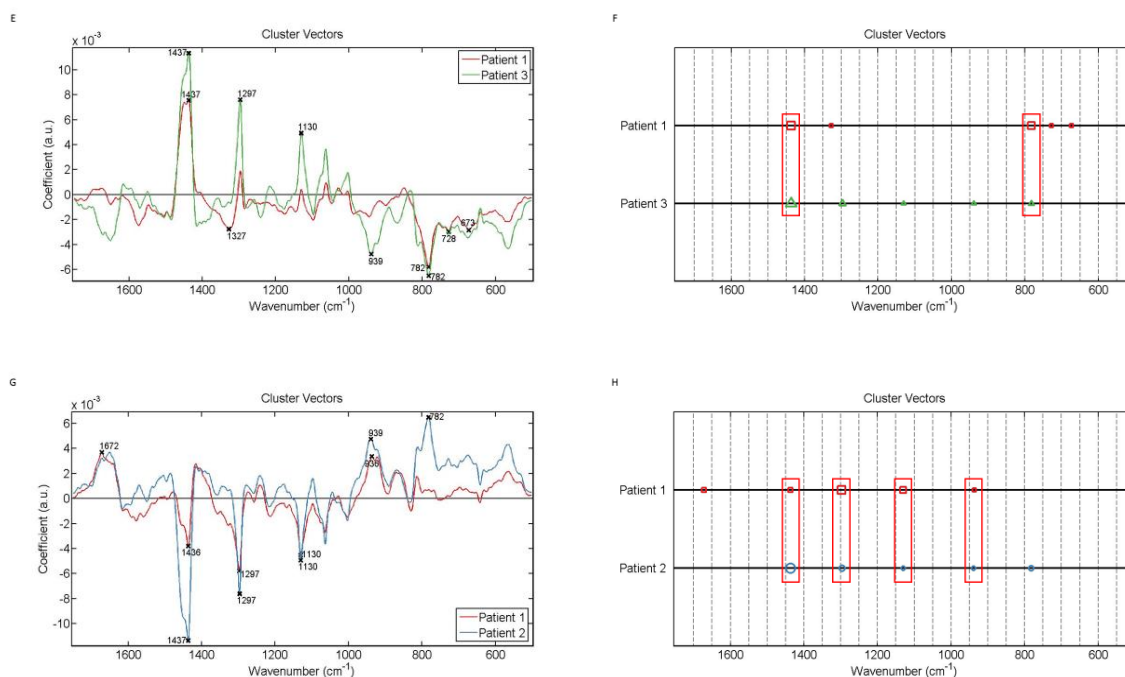
The five major discriminant wavenumbers identified in loading plots accounting for variations between tissue samples were  $1437\text{ cm}^{-1}$  (proteins and lipids),  $1297\text{ cm}^{-1}$ ,  $941\text{ cm}^{-1}$  (polysaccharides, amylose),  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA) and  $564\text{ cm}^{-1}$  (Fig. 3.186).



**Figure 3.186:** Loadings plots showing wavenumbers that discriminate luminal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vectors plots were presented and analysed in the same approach as before in an attempt to identify potential biomarkers for luminal cells (Fig. 3.187). Having no class as a reference, two tissues were common for 1437  $\text{cm}^{-1}$  (proteins and lipids), 1297  $\text{cm}^{-1}$  and 941  $\text{cm}^{-1}$  (polysaccharides, amylose) with 937  $\text{cm}^{-1}$  (proline, hydroxyproline, collagen backbone, glycogen) whilst 783  $\text{cm}^{-1}$  was commonly shared with 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA) by all three tissues. Using Patient 1 as reference, 1437  $\text{cm}^{-1}$  (proteins and lipids) was commonly shared with 1436  $\text{cm}^{-1}$  (lipids) by two tissues. Using Patient 2 as reference, two tissues were common for 1437  $\text{cm}^{-1}$  (proteins and lipids) and 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA). Having Patient 3 as class reference, the observed wavenumbers which showed common occurrence were 1437  $\text{cm}^{-1}$  (proteins and lipids) with 1436  $\text{cm}^{-1}$  (lipids), 1297  $\text{cm}^{-1}$ , 1130  $\text{cm}^{-1}$  (lipids) and 939  $\text{cm}^{-1}$  with 936  $\text{cm}^{-1}$  (valine/proline and protein).



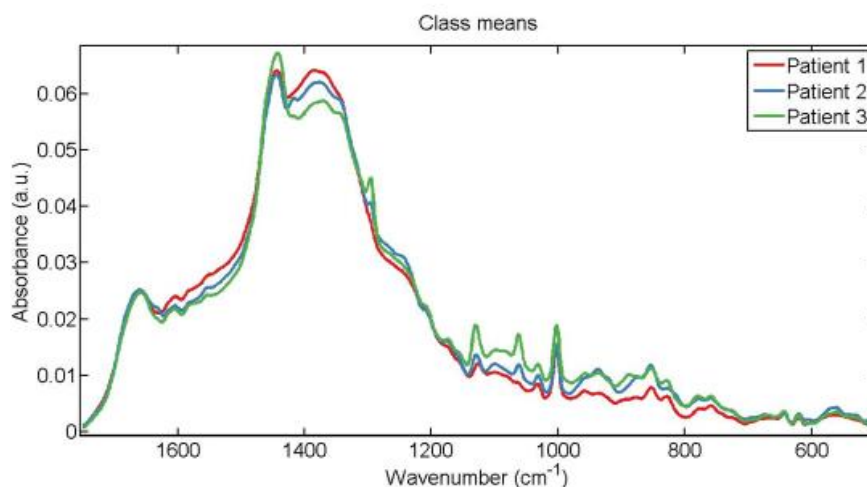


**Figure 3.187:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells among non-diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-11) reference class, (E) & (F) Patient 2 (H09-9102-6) reference class and (G) & (H) Patient 3 (H09-13558-A8) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Between diseased samples from patients with endometrial cancer*

Spectra acquired from luminal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) taken from cancerous lesions in the endometrium were analysed and compared. Classes were labelled as Patient 1, Patient 2 and Patient 3 to represent the samples H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In this section classes of tissues will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots. All classes exhibited a very similar mean Raman spectra with most pronounced variations in absorbance intensities occurring at  $\sim 1443 \text{ cm}^{-1}$ ,  $\sim 1380 \text{ cm}^{-1}$ ,  $\sim 1295 \text{ cm}^{-1}$ ,  $\sim 1130 \text{ cm}^{-1}$ ,  $\sim 1061 \text{ cm}^{-1}$ ,  $\sim 1001 \text{ cm}^{-1}$ ,  $\sim 935 \text{ cm}^{-1}$  and  $853 \text{ cm}^{-1}$  (Fig. 3.188). Mainly Patient 3 had the highest peak intensities except at  $\sim 1380 \text{ cm}^{-1}$  whereby Patient 1 had the highest peak.

Statistical analysis indicated that all tissues were significantly different from each other (Fig. 3.189).

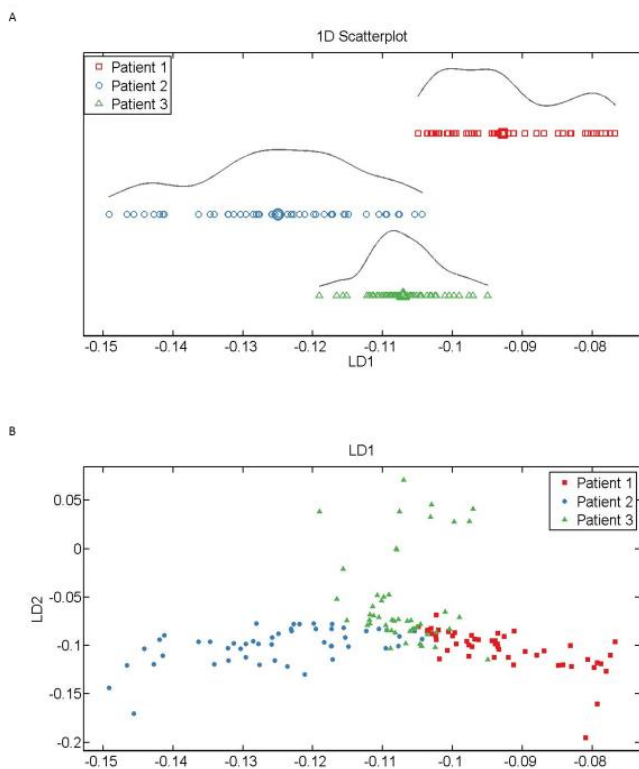


**Figure 3.188:** Mean Raman spectra for comparison of spectra from luminal cells in diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.001
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P > 0.05
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P < 0.001
Patient 2 vs Patient 3	P < 0.001

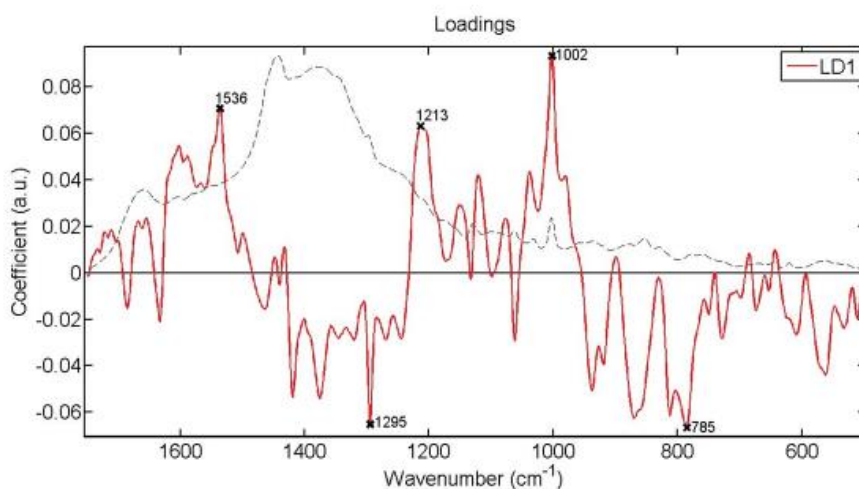
**Figure 3.189:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from luminal cells between diseased tissue sections H09-9102-1, H09-12890-9, H09-13558-A5 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Segregation of classes resulting from application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.190). Patient 2 displayed slight overlap with Patient 1. Almost half of the spectra of Patient 3 would overlap with the spectra of Patient 2 whilst the other half would overlap with spectra of Patient 1. Spectral points of Patient 1 and 3 had a more compact arrangement relative to the spectral points of Patient 2 which signified less intra-class variation for Patient 1 and Patient 3.



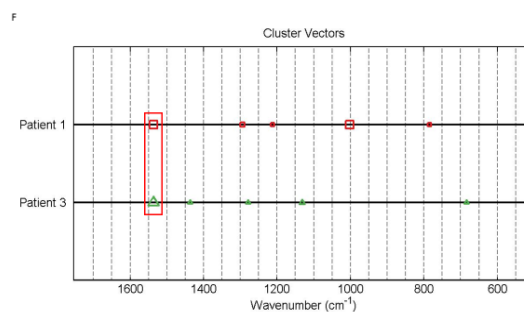
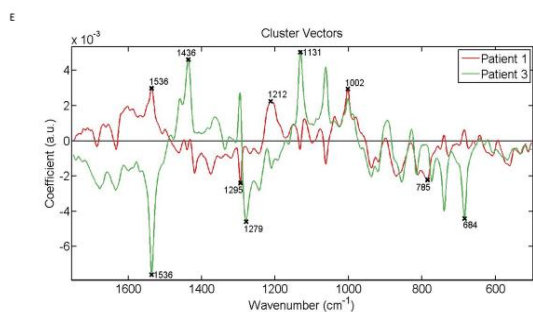
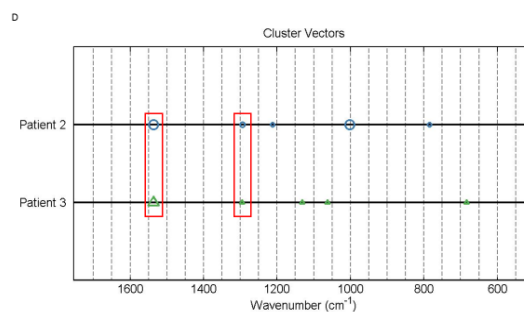
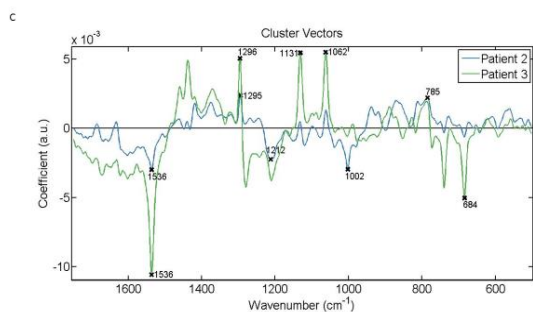
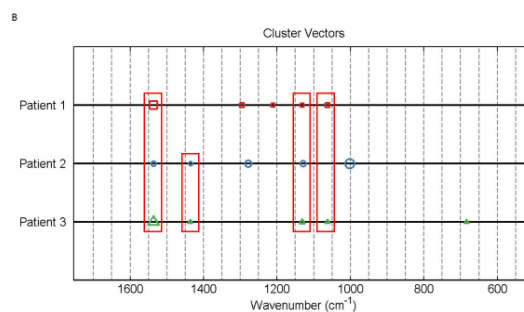
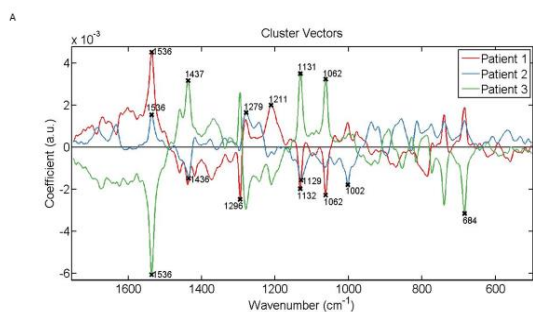
**Figure 3.190:** Scores plots for segregation of spectra acquired from luminal cells in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

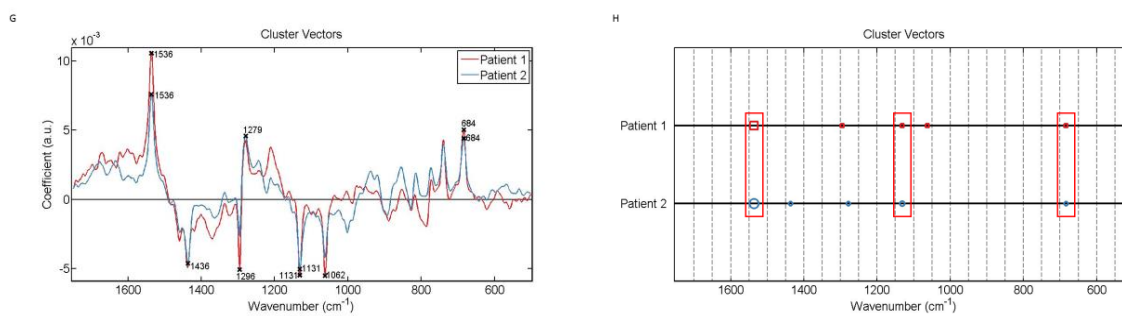
Generation of loading plots identified five major wavenumbers responsible for variations between the interrogated tissue samples (Fig. 3.191). The identified wavenumbers were 1536  $\text{cm}^{-1}$ , 1295  $\text{cm}^{-1}$ , 1213  $\text{cm}^{-1}$ , 1002  $\text{cm}^{-1}$  (phenylalanine) and 785  $\text{cm}^{-1}$ .



**Figure 3.191:** Loadings plots showing wavenumbers that discriminate luminal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots were presented and analysed in the same approach as before in order to identify biomarkers for luminal cells (Fig. 3.192). Having no class as a reference origin the wavenumbers displaying common occurrence were 1536  $\text{cm}^{-1}$ , 1437  $\text{cm}^{-1}$  (proteins and lipids) with 1436  $\text{cm}^{-1}$  (lipids), 1132  $\text{cm}^{-1}$  (proteins and lipids) with 1131  $\text{cm}^{-1}$  (lipids) and 1129  $\text{cm}^{-1}$  (lipids) and 1062  $\text{cm}^{-1}$  (paraffin). Having Patient 1 as reference, common occurrence was displayed by 1536  $\text{cm}^{-1}$  and 1296  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) with 1295  $\text{cm}^{-1}$ . Having Patient 2 as reference, only 1536  $\text{cm}^{-1}$  was commonly shared. Having Patient 3 as reference the commonly shared wavenumbers were 1536  $\text{cm}^{-1}$ , 1131  $\text{cm}^{-1}$  (lipids) and 684  $\text{cm}^{-1}$ .





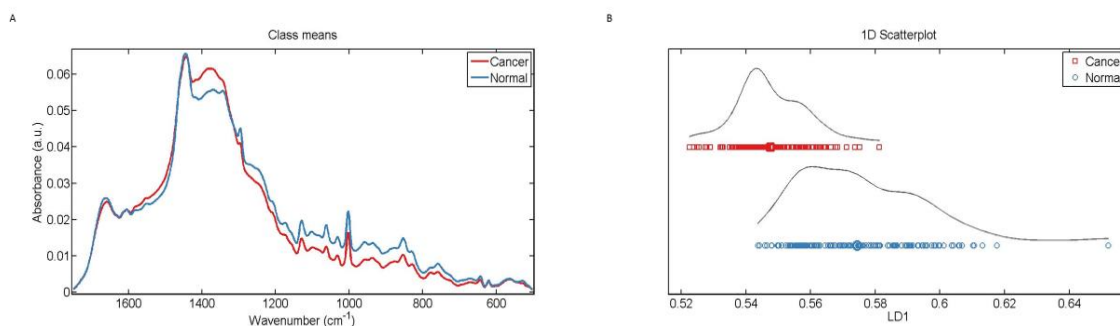
**Figure 3.192:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells among diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-9) reference class, (E) & (F) Patient 2 (H09-9102-1) reference class and (G) & (H) Patient 3 (H09-13558-A5) reference class. The red rectangles highlight the wavenumbers occurring

### *Diseased tissues Vs Non-Diseased tissues*

The spectra of the luminal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) were compared with the spectra of luminal cells in their corresponding non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer. The class representing the spectra collected from luminal cells in diseased tissue sections was labelled as ‘Cancer’ whilst the class representing non-diseased tissue sections was labelled as ‘Normal’. In this section classes of spectra will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots.

Both classes exhibited a very similar mean Raman spectra with most pronounced variations in absorbance intensities occurring at  $\sim 1377\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1129\text{ cm}^{-1}$ ,  $\sim 1061\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 949\text{ cm}^{-1}$ ,  $\sim 853\text{ cm}^{-1}$  and  $\sim 760\text{ cm}^{-1}$  (Fig.3.193 A). The Normal class exhibited the highest peak intensities throughout the spectra except at  $\sim 1377\text{ cm}^{-1}$  whereby the Cancer class had the highest peak. Application of PCA-LDA resulted in segregation of classes which was visualised in scores plot of 1D space whereby more than half of the spectra of Cancer would overlap with less than half of the spectra of Normal (Fig. 3.193 B). Classes were significantly different from each other according to the statistical analysis (Fig. 3.194).



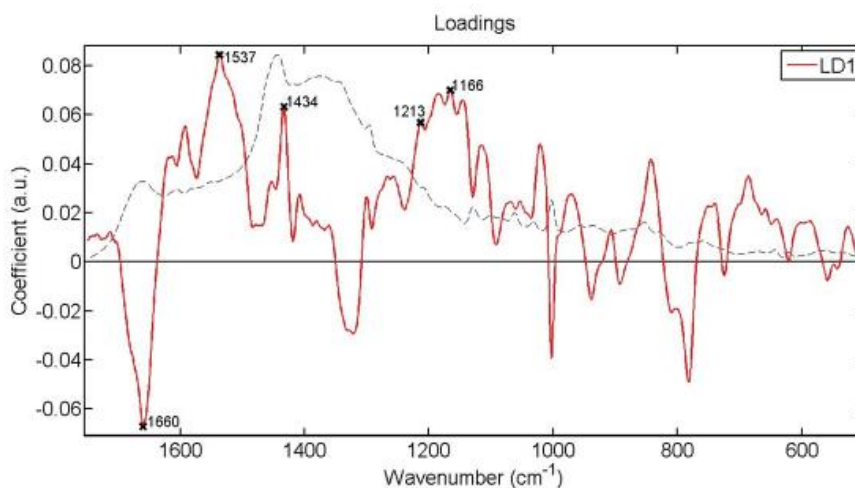


**Figure 3.193:** (A) Mean Raman spectra for comparison of spectra from luminal cells in all diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5) and their corresponding non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). (B) Scores plots for segregation of spectra in 1D space. Diseased and non-diseased tissue samples are represented by the class labels 'Cancer' (red) and 'Normal' (blue) respectively.

Parameters	P Value
Cancer Vs Normal	P<0.0001

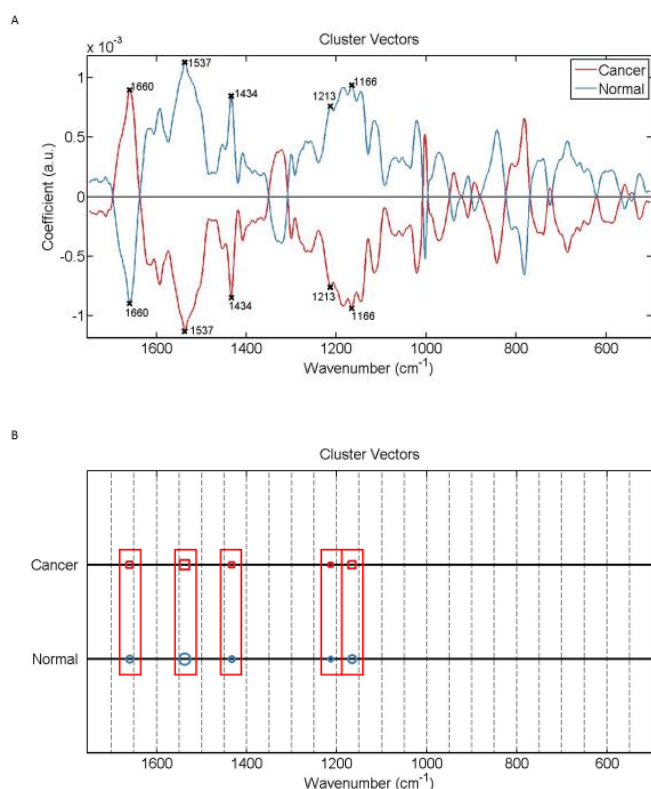
**Figure 3.194:** Obtained p-value by employment of Unpaired t-test to compare spectra acquired from luminal cells between non diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and their corresponding diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections represented by the class labels Normal and Cancer respectively.

The discriminant wavenumbers identified in loading plots responsible for variations between the two classes were 1660 cm<sup>-1</sup> (amide I), 1537 cm<sup>-1</sup>, 1434 cm<sup>-1</sup>, 1213 cm<sup>-1</sup> and 1166 cm<sup>-1</sup> (Fig. 3.195).



**Figure 3.195:** Loadings plots showing wavenumbers that discriminate spectra from luminal cells in all diseased (H09-9102-1, H09-12890-9, H09-13558-A5) and their corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) tissue sections. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

The five major wavenumbers in loading plots were also observed in cluster vector plots occurring in both classes (Fig. 3.196). Expression levels of the wavenumbers were found to be the same between the classes as indicated by the size of the shapes representing the classes. Highest concentration levels were observed for 1537  $\text{cm}^{-1}$ , 1660  $\text{cm}^{-1}$  (amide I) and 1166  $\text{cm}^{-1}$  showed medium levels whilst 1434  $\text{cm}^{-1}$  and 1213  $\text{cm}^{-1}$  were displayed at low levels. Biomarkers that would specify either 'Normal' or 'Cancer' luminal cells could not be identified since the identified wavenumbers were commonly shared by both classes. For a wavenumber to have a potential as a biomarker for either of the classes, should have been observed in only one class.



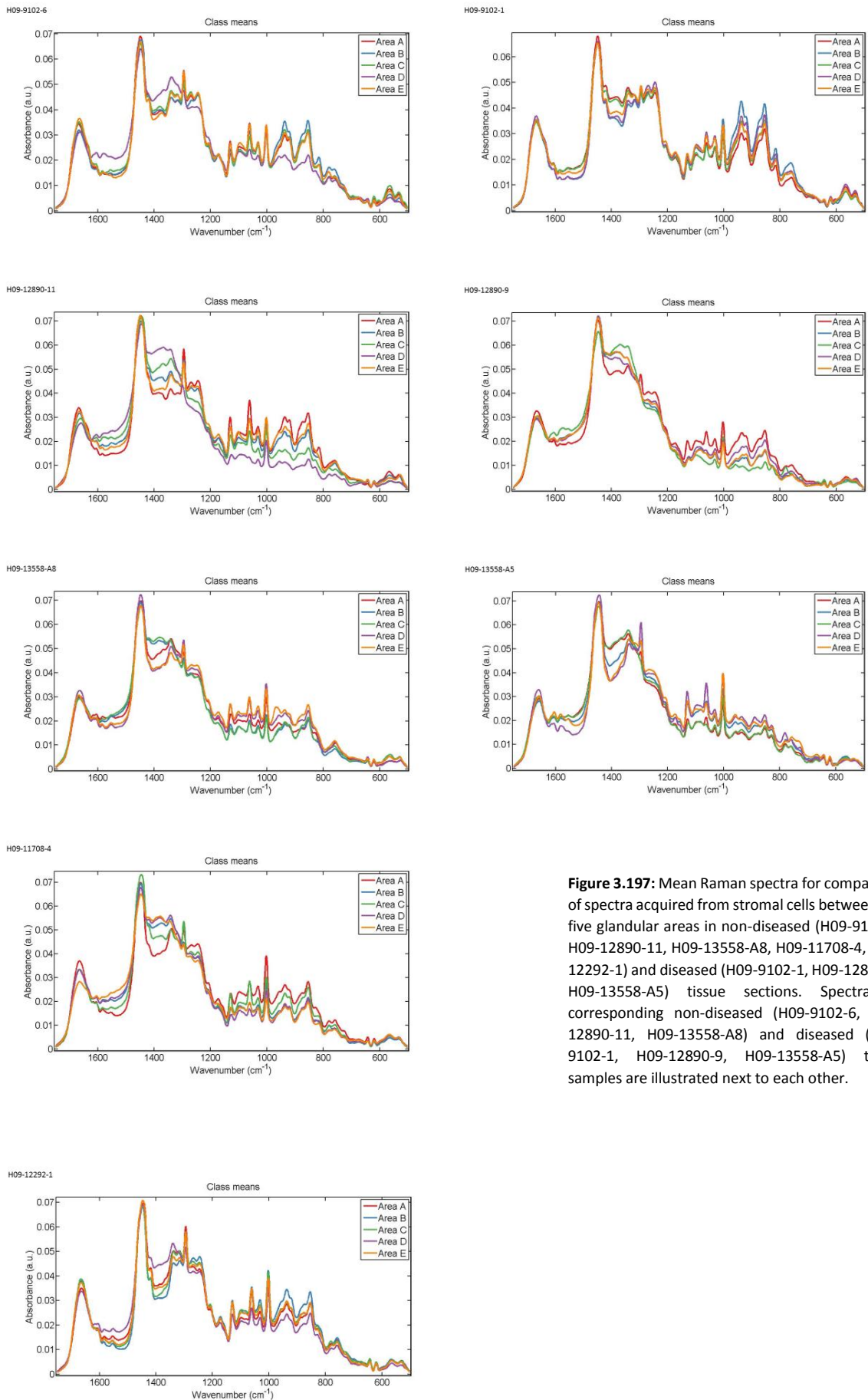
**Figure 3.196:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from luminal cells among all corresponding diseased (H09-12890-9, H09-9102-1, H09-13558-A5) and non-diseased (H09-12890-11, H09-9102-6, H09-13558-A8) tissue sections. The red rectangles highlight the wavenumbers occurring commonly in classes.

### 3.3.3 Stromal Cells Vs Stromal Cells

#### *In individual tissue samples*

Looking at the mean Raman spectra of stromal cells from the five different glandular elements in each tissue sample (Fig. 3.197) it was observed that areas within individual samples exhibited a very similar shape of the collected spectra whilst in general the shape of the spectra was relatively similar amongst tissue samples with main differences in shape occurring between the spectral region 1244-1380  $\text{cm}^{-1}$ . Most evident variations in absorbance intensities were observed at  $\sim 1666 \text{ cm}^{-1}$ ,  $\sim 1449 \text{ cm}^{-1}$ ,  $1294 \text{ cm}^{-1}$ ,  $\sim 1130 \text{ cm}^{-1}$ ,  $\sim 1061 \text{ cm}^{-1}$ ,  $\sim 1002 \text{ cm}^{-1}$ ,  $935 \text{ cm}^{-1}$  and  $\sim 855$ . These variations in intensities were observed in the mean spectra of both types of tissue samples, either diseased or non-diseased. Not a significant similarity in shape of spectra was observed between corresponding diseased and non-diseased tissue samples.

Statistical analysis (Fig. 3.198) indicated that in the majority of the tissues, individual glandular areas were not significantly different from the spectra collected from the tissue as a whole whereas most of the glandular areas within individual tissues were significantly different from each other.



**Figure 3.197:** Mean Raman spectra for comparison of spectra acquired from stromal cells between the five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.

H09-9102-6

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.01
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.001
Area B vs Area D	P < 0.01
Area B vs Area E	P > 0.05
Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.001

H09-12890-11

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P > 0.05
Area B vs Area C	P < 0.01
Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P < 0.05
Area D vs Area E	P < 0.001

H09-13558-A8

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P < 0.001
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.05
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.001

H09-11708-4

Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P < 0.01
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001
Area B vs Area C	P > 0.05
Area B vs Area D	P > 0.05
Area B vs Area E	P > 0.05
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P < 0.01

H09-12292-1

Parameters	P value
LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P < 0.05
LDA1 vs Area C	P < 0.05
LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05
Area A vs Area B	P > 0.05
Area A vs Area C	P > 0.05
Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.05
Area B vs Area C	P > 0.05
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.001
Area C vs Area D	P < 0.001
Area C vs Area E	P < 0.001
Area D vs Area E	P < 0.05

H09-9102-1

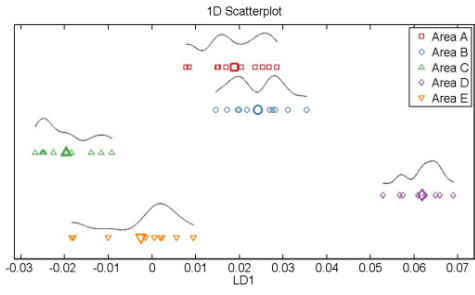
Parameters	P value
LDA1 vs Area A	P < 0.001
LDA1 vs Area B	P < 0.001
LDA1 vs Area C	P > 0.05
LDA1 vs Area D	P > 0.05
LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001
Area A vs Area C	P < 0.001
Area A vs Area D	P > 0.05
Area A vs Area E	P < 0.05
Area B vs Area C	P < 0.01
Area B vs Area D	P < 0.001
Area B vs Area E	P < 0.001
Area C vs Area D	P > 0.05
Area C vs Area E	P > 0.05
Area D vs Area E	P > 0.05

H09-12890-9		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Area A	P < 0.001	LDA1 vs Area A	P > 0.05
LDA1 vs Area B	P > 0.05	LDA1 vs Area B	P > 0.05
LDA1 vs Area C	P < 0.01	LDA1 vs Area C	P < 0.001
LDA1 vs Area D	P > 0.05	LDA1 vs Area D	P < 0.001
LDA1 vs Area E	P > 0.05	LDA1 vs Area E	P > 0.05
Area A vs Area B	P < 0.001	Area A vs Area B	P > 0.05
Area A vs Area C	P < 0.001	Area A vs Area C	P < 0.05
Area A vs Area D	P < 0.001	Area A vs Area D	P < 0.001
Area A vs Area E	P < 0.001	Area A vs Area E	P > 0.05
Area B vs Area C	P > 0.05	Area B vs Area C	P < 0.05
Area B vs Area D	P > 0.05	Area B vs Area D	P < 0.001
Area B vs Area E	P > 0.05	Area B vs Area E	P > 0.05
Area C vs Area D	P > 0.05	Area C vs Area D	P < 0.001
Area C vs Area E	P > 0.05	Area C vs Area E	P < 0.001
Area D vs Area E	P > 0.05	Area D vs Area E	P > 0.05

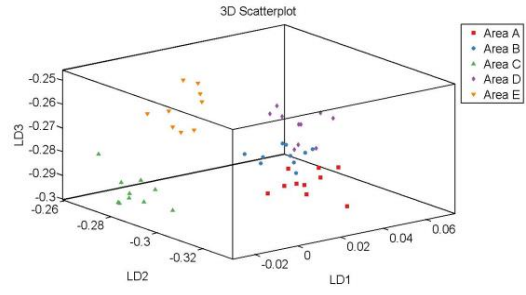
**Figure 3.198:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal in five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Segregation of classes of spectra, resulted from application of PCA-LDA, was visualised in scores plot of 1D and 3D space. Figure 3.199 illustrates scores plots of non-diseased tissue samples and figure 3.200 illustrates scores plots of diseased tissue samples. Scores plots illustrated a well separation of classes representing the areas with mainly at least two or three glandular areas displaying overlap in individual tissues. In non-diseased tissue samples, spectral points of individual glandular areas exhibited a rather compact arrangement relative to the arrangement displayed by spectral point of glandular areas in diseased tissue samples.

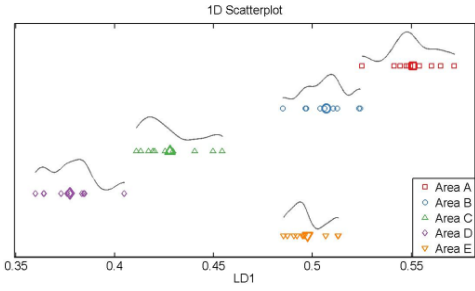
H09-9102-6



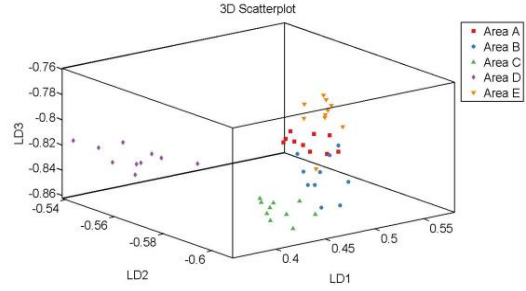
H09-9102-6



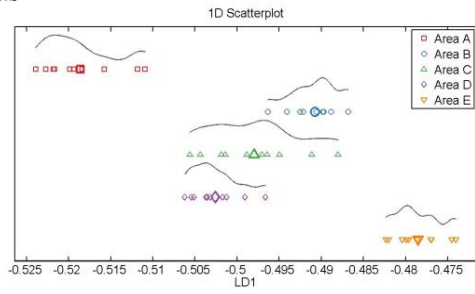
H09-12890-11



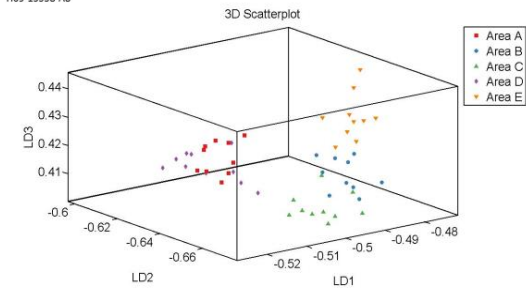
H09-12890-11



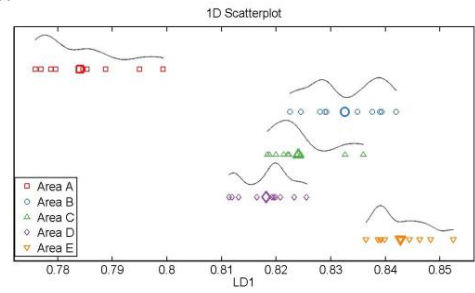
H09-13558-A8



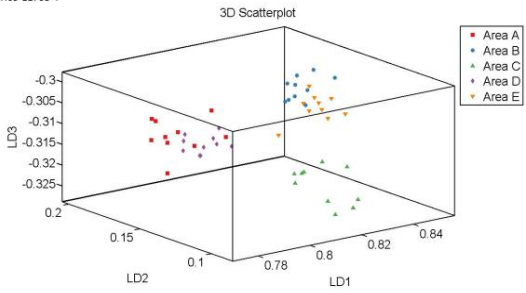
H09-13558-A8

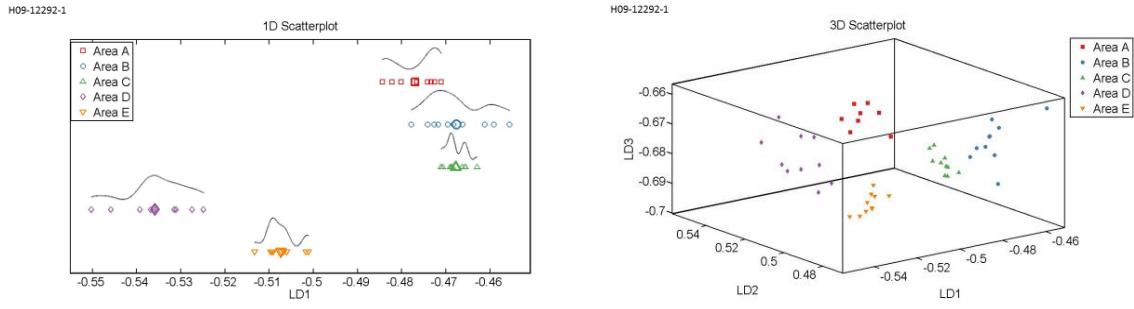


H09-11708-4

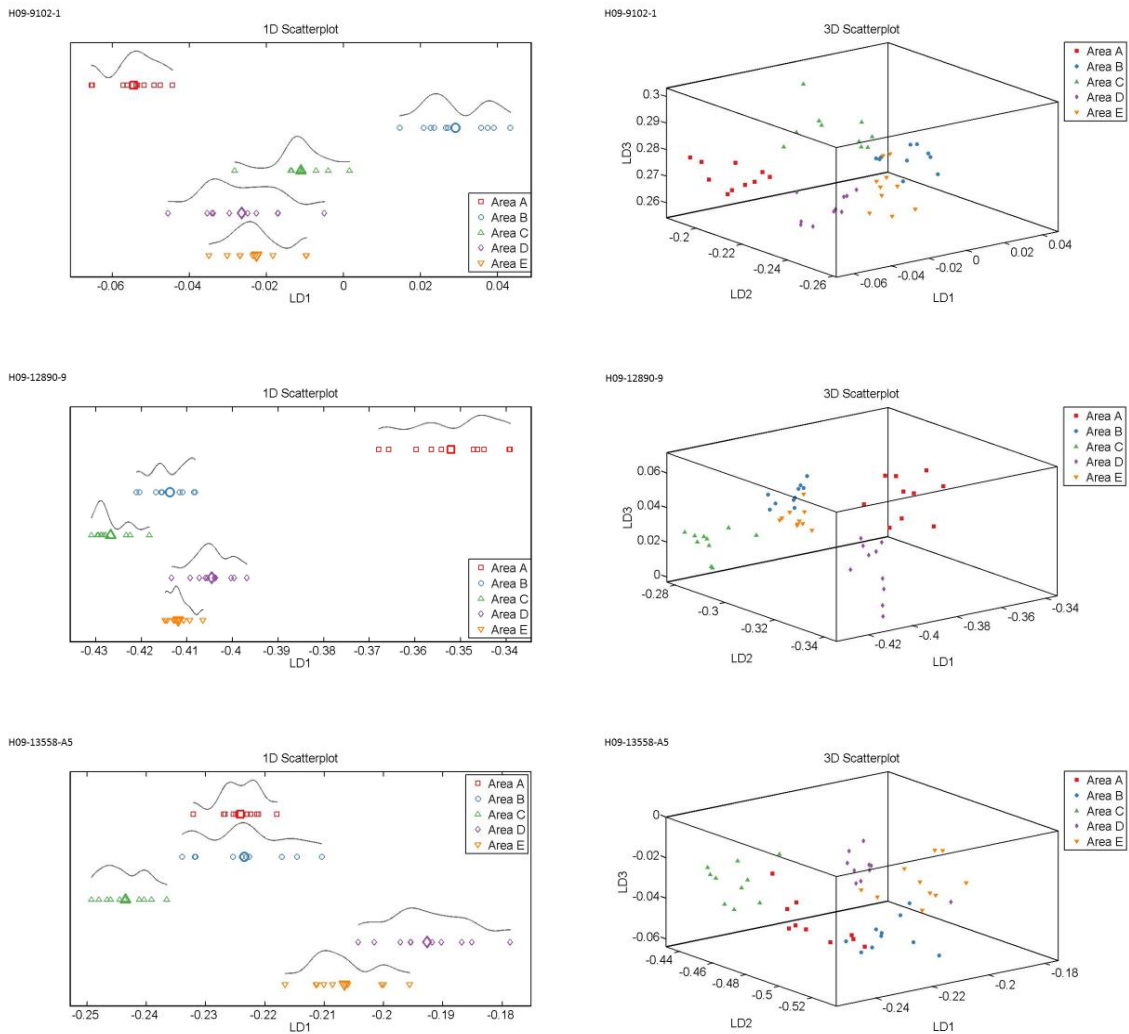


H09-11708-4





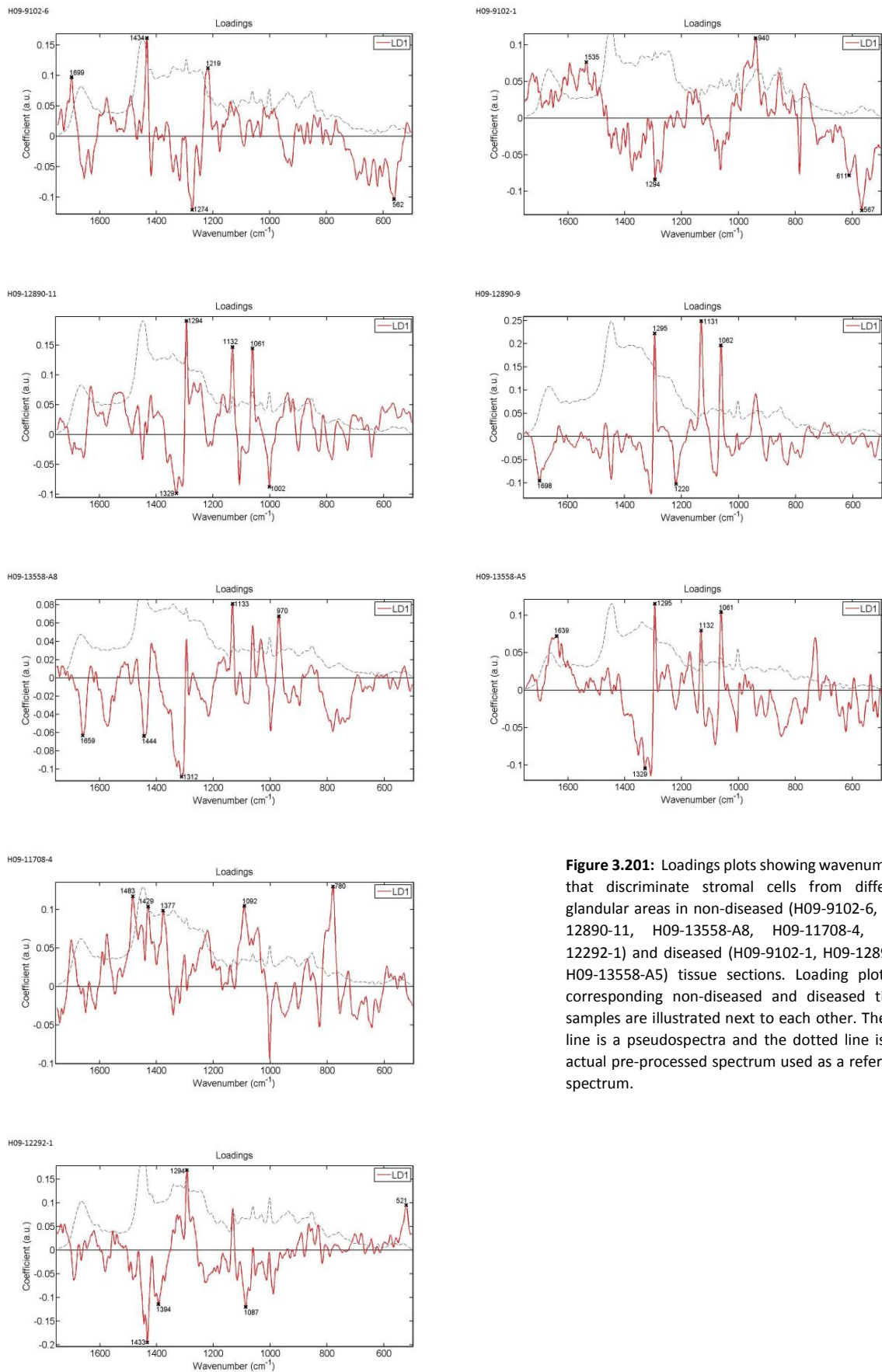
**Figure 3.199:** 1D scores plots and their corresponding 3D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from stromal cells from five different glandular elements.



**Figure 3.200:** 1D scores plots and their corresponding 3D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from stromal cells from five different glandular elements.



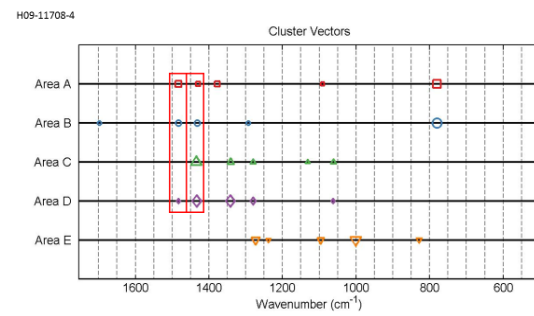
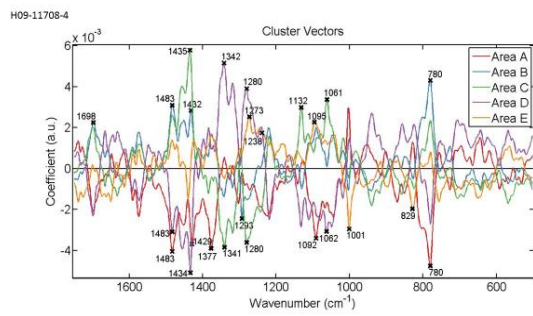
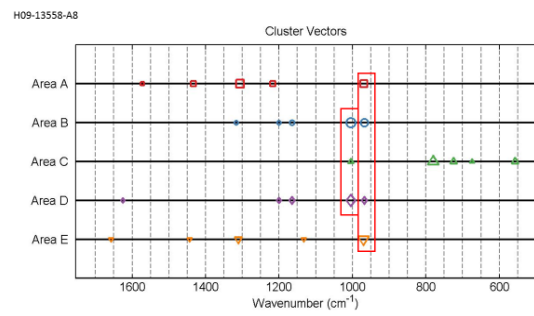
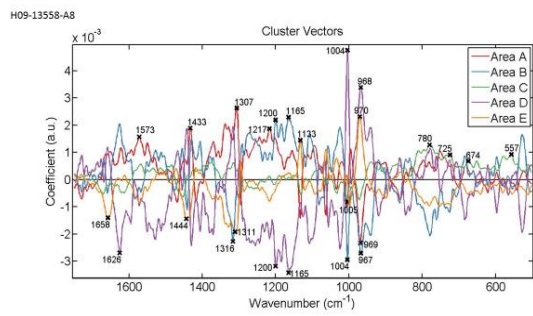
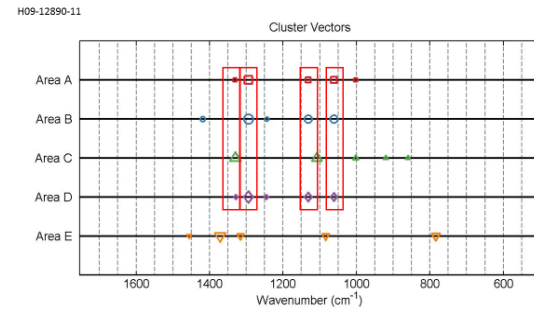
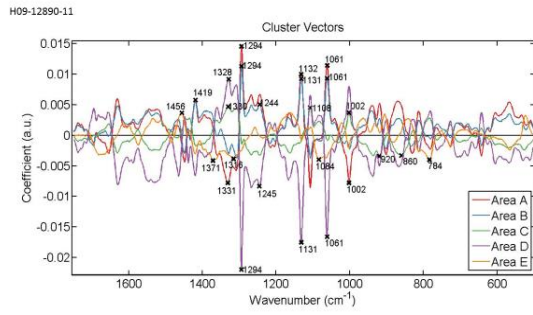
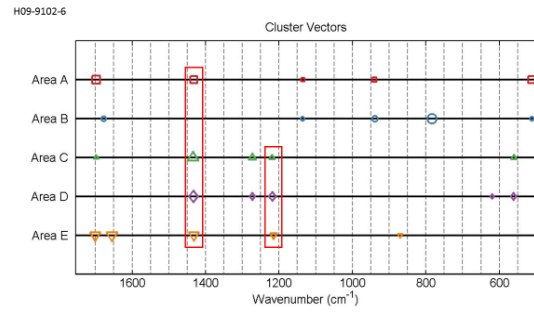
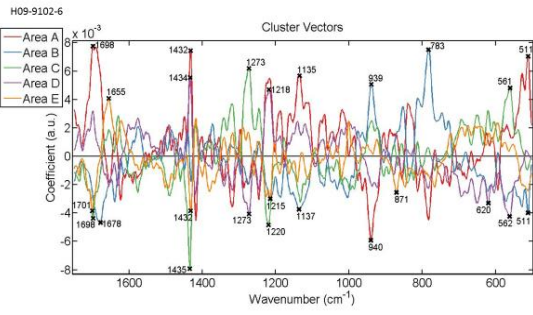
Figure 3.201 illustrates the loading plots for each tissue sample revealing the discriminant wavenumbers responsible for variations between the spectra acquired from stromal cells at glandular areas. In tissue sample H09-9102-6 the discriminant wavenumbers were 1699  $\text{cm}^{-1}$ , 1434  $\text{cm}^{-1}$ , 1274  $\text{cm}^{-1}$ , 1219  $\text{cm}^{-1}$  and 562  $\text{cm}^{-1}$ . In sample H09-12890-11 the discriminant wavenumbers were 1329  $\text{cm}^{-1}$  (purine bases in nucleic acids), 1294  $\text{cm}^{-1}$  (methylene twisting), 1132  $\text{cm}^{-1}$  (proteins and lipids), 1061  $\text{cm}^{-1}$  (residual paraffin) and 1002  $\text{cm}^{-1}$  (phenylalanine). For tissue H09-13558-A8 the discriminant wavenumbers were 1659  $\text{cm}^{-1}$  (amide I, collagen like proteins), 1444  $\text{cm}^{-1}$  (lipids), 1312  $\text{cm}^{-1}$  (collagen/lipids), 1133  $\text{cm}^{-1}$  and 970  $\text{cm}^{-1}$  (phosphorylated proteins and cellular nucleic acids). For tissue H09-11708-4 the wavenumbers were 1483  $\text{cm}^{-1}$ , 1429  $\text{cm}^{-1}$ , 1377  $\text{cm}^{-1}$ , 1092  $\text{cm}^{-1}$  (phosphodioxo) and 780  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA). For tissue H09-12292-1 the discriminant wavenumbers were 1433  $\text{cm}^{-1}$  (lipids), 1394  $\text{cm}^{-1}$ , 1294  $\text{cm}^{-1}$  (methylene twisting), 1087  $\text{cm}^{-1}$  (lipids) and 521  $\text{cm}^{-1}$ . In sample H09-9102-1 the identified wavenumbers were 1535  $\text{cm}^{-1}$ , 1294  $\text{cm}^{-1}$  (methylene twisting), 940  $\text{cm}^{-1}$ , 611  $\text{cm}^{-1}$  and 567  $\text{cm}^{-1}$ . In sample H09-12890-9, the wavenumbers were 1698  $\text{cm}^{-1}$ , 1295  $\text{cm}^{-1}$ , 1220  $\text{cm}^{-1}$  (amide III;  $\beta$  sheet), 1131  $\text{cm}^{-1}$  (lipids) and 1062  $\text{cm}^{-1}$  (paraffin). In sample H09-13558-A5 the identified wavenumbers were 1639  $\text{cm}^{-1}$ , 1329  $\text{cm}^{-1}$  (purine bases in nucleic acids), 1295  $\text{cm}^{-1}$ , 1132  $\text{cm}^{-1}$  (proteins and lipids) and 1061  $\text{cm}^{-1}$  (residual paraffin).

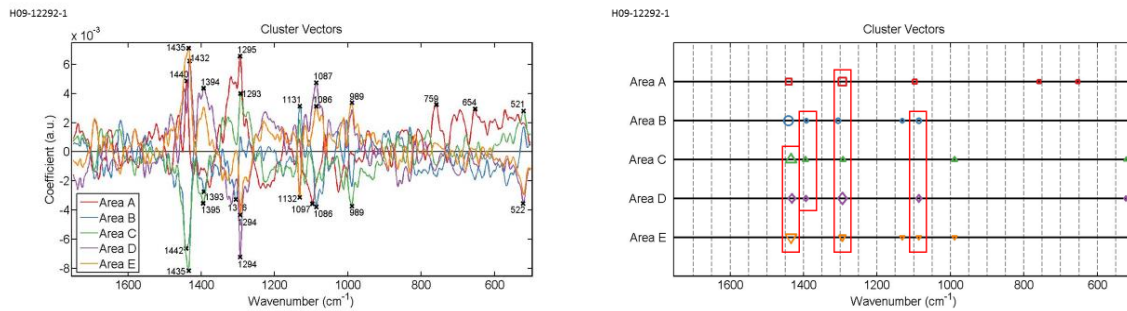


**Figure 3.201:** Loadings plots showing wavenumbers that discriminate stromal cells from different glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

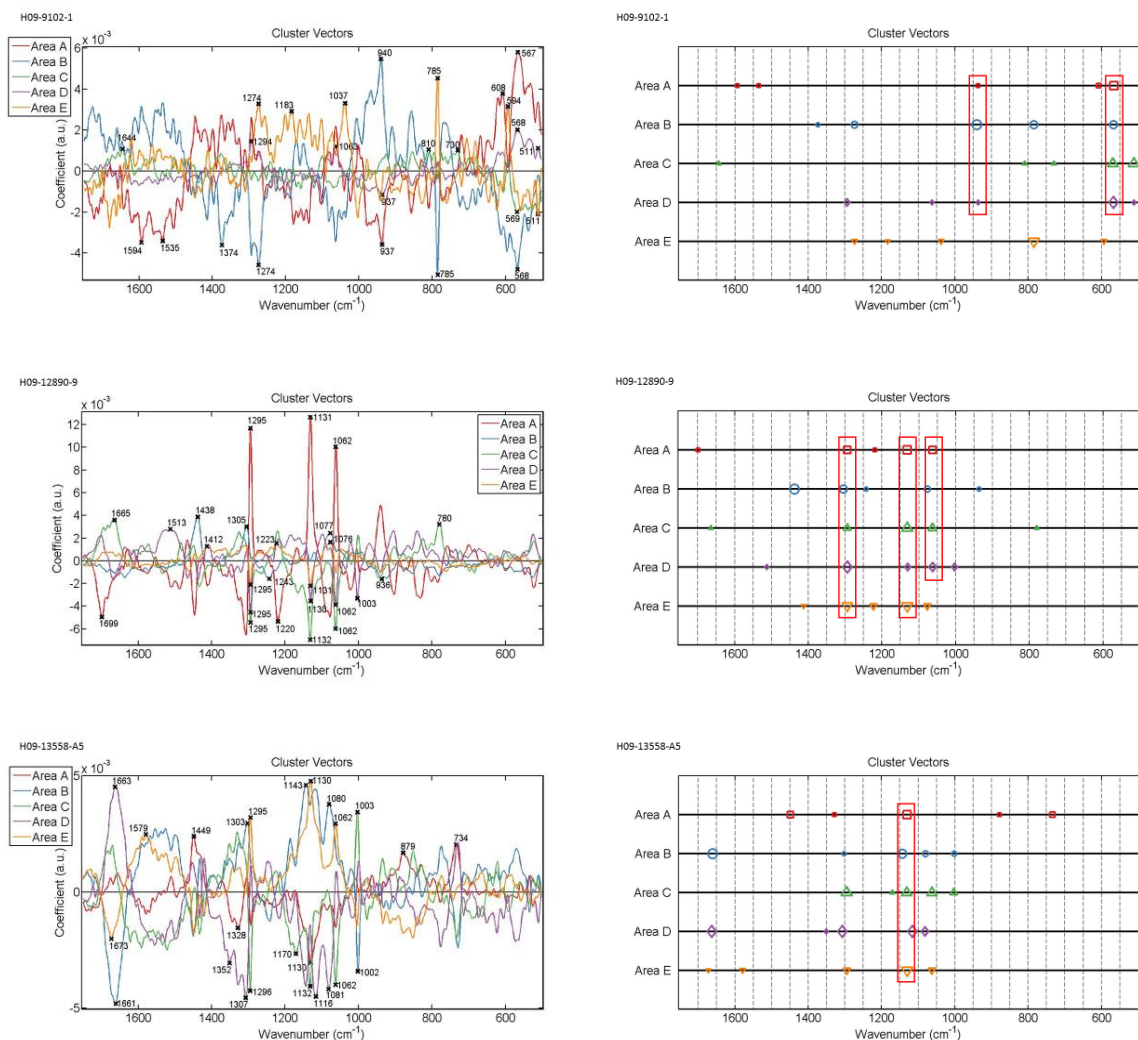
Cluster vector plots of non-diseased (Fig. 3.202) and diseased (Fig. 3.203) tissue samples were presented and analysed in the same approach as before in the hope of identifying possible biomarkers for stromal cells, based on the occurrence of wavenumbers and their expression levels in the glandular areas.

In sample H09-9102-6, 1435  $\text{cm}^{-1}$  was commonly shared with 1434  $\text{cm}^{-1}$  and 1432  $\text{cm}^{-1}$  in four areas whilst 1220  $\text{cm}^{-1}$  (amide III,  $\beta$  sheet) was commonly shared with 1218  $\text{cm}^{-1}$  and 1215  $\text{cm}^{-1}$  by three areas. In sample H09-12890-11, 1331  $\text{cm}^{-1}$  was commonly shared with 1330  $\text{cm}^{-1}$  (DNA, phospholipids, collagen, phosphates) and 1328  $\text{cm}^{-1}$  by three areas, similarly three areas were common for 1294  $\text{cm}^{-1}$  (methylene twisting), 1132  $\text{cm}^{-1}$  (proteins and lipids) with 1131  $\text{cm}^{-1}$  (lipids) and 1061  $\text{cm}^{-1}$  (paraffin). In sample H09-13558-A8 three areas were common for 1005  $\text{cm}^{-1}$  (phenylalanine) with 1004  $\text{cm}^{-1}$  (phenylalanine) whilst 970  $\text{cm}^{-1}$  (phosphorylated proteins and cellular nucleic acids) was commonly shared with 969  $\text{cm}^{-1}$ , 968  $\text{cm}^{-1}$  (lipids) and 967  $\text{cm}^{-1}$  in four areas. In sample H09-11708-4 three areas were common for 1483  $\text{cm}^{-1}$ , 1435  $\text{cm}^{-1}$  with 1434  $\text{cm}^{-1}$ , 1432  $\text{cm}^{-1}$  and 1429  $\text{cm}^{-1}$ . In sample H09-12292-1 common occurrence of 1442  $\text{cm}^{-1}$  (fatty acids/collagen) with 1440  $\text{cm}^{-1}$  (lipids), 1435  $\text{cm}^{-1}$  and 1432  $\text{cm}^{-1}$  was displayed by all five areas, four areas were common for 1295  $\text{cm}^{-1}$  with 1294  $\text{cm}^{-1}$  (methylene twisting) and 1293  $\text{cm}^{-1}$  whilst three areas were common for 1395  $\text{cm}^{-1}$  with 1394  $\text{cm}^{-1}$  and 1393  $\text{cm}^{-1}$  (CH rocking), and 1087  $\text{cm}^{-1}$  (acyl backbone in lipids) with 1086  $\text{cm}^{-1}$  (phosphate backbone). In sample H09-9102-1, 940  $\text{cm}^{-1}$  was commonly shared with 937  $\text{cm}^{-1}$  (proline, protein backbone/glycogen) by three areas whilst four areas were common for 569  $\text{cm}^{-1}$  with 568  $\text{cm}^{-1}$  and 567  $\text{cm}^{-1}$ . In sample H09-12890-9, four areas were common for 1295  $\text{cm}^{-1}$ , 1132  $\text{cm}^{-1}$  (proteins and lipids) with 1131  $\text{cm}^{-1}$  (lipids) and 1130  $\text{cm}^{-1}$  (lipids) whilst three areas were common for 1062  $\text{cm}^{-1}$  (paraffin). In sample H09-13558-A5 three areas were only common for 1132  $\text{cm}^{-1}$  (proteins and lipids) with 1130  $\text{cm}^{-1}$  (lipids).





**Figure 3.202:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells, in five glandular areas in individual non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1). The red rectangles highlight the wavenumbers shared in common among glandular areas.

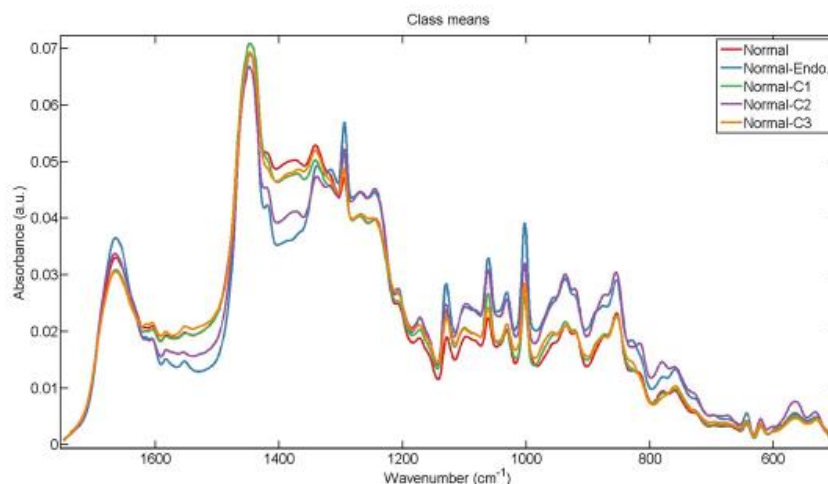


**Figure 3.203:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells, in five glandular areas in individual diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight the wavenumbers shared in common among glandular areas.

*Between all normal tissue samples*

All the non-diseased tissue samples were compared between them for the spectra interrogated from their stromal cells. Classes representing tissues were labelled in the same way as before and will be referred according to their class labels to associate plots with observations. All classes of tissues exhibited a very similar shape of mean Raman spectra with most evident variations in peak intensities occurring at  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1447\text{ cm}^{-1}$ ,  $\sim 1341\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1259\text{ cm}^{-1}$ ,  $\sim 1131\text{ cm}^{-1}$ ,  $\sim 1061\text{ cm}^{-1}$ ,  $\sim 1003\text{ cm}^{-1}$ ,  $\sim 937\text{ cm}^{-1}$ ,  $\sim 854\text{ cm}^{-1}$  and  $\sim 757\text{ cm}^{-1}$  (Fig. 3.204). Mainly the class Normal-Endo. had the highest peaks throughout the spectra.

Statistical analysis (Fig. 3.205) indicated that classes were significantly different from each other for the spectra acquired from stromal cells but Normal was not significantly different from Normal-C2 neither was Normal-C1 from Normal-C3.

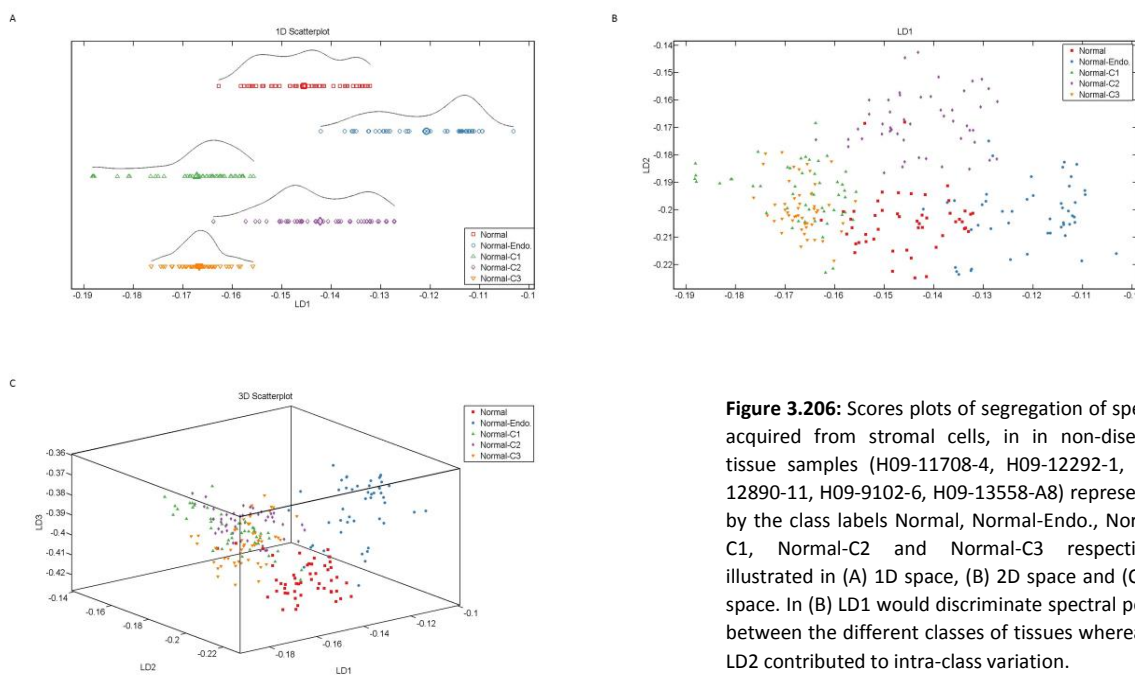


**Figure 3.204:** Mean Raman spectra for comparison of spectra acquired from stromal cells in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Parameters	P value
LDA1 vs Normal	P > 0.05
LDA1 vs Normal-Endo.	P < 0.001
LDA1 vs Normal-C1	P < 0.001
LDA1 vs Normal-C2	P > 0.05
LDA1 vs Normal-C3	P < 0.001
Normal vs Normal-Endo.	P < 0.001
Normal vs Normal-C1	P < 0.001
Normal vs Normal-C2	P > 0.05
Normal vs Normal-C3	P < 0.001
Normal-Endo. vs Normal-C1	P < 0.001
Normal-Endo. vs Normal-C2	P < 0.001
Normal-Endo. vs Normal-C3	P < 0.001
Normal-C1 vs Normal-C2	P < 0.001
Normal-C1 vs Normal-C3	P > 0.05
Normal-C2 vs Normal-C3	P < 0.001

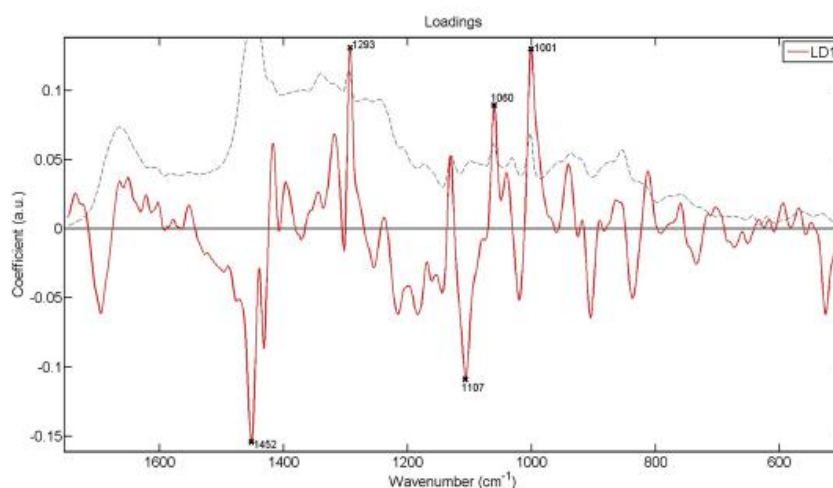
**Figure 3.205:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal cells in all five glandular elements between non-diseased tissue sections H09-11708-4, H09-9102-1, H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D, 2D and 3D space (Fig. 3.206). It was observed that an individual class would overlap at some point throughout its spectra with at least three other classes but class Normal-Endo. overlapped only with classes Normal and Normal-C2. Spectral points of Normal-C3 had the most compact arrangement relative to the other classes which signified the least intra-class variation.



**Figure 3.206:** Scores plots of segregation of spectra acquired from stromal cells, in non-diseased tissue samples (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8) represented by the class labels Normal, Normal-Endo., Normal-C1, Normal-C2 and Normal-C3 respectively, illustrated in (A) 1D space, (B) 2D space and (C) 3D space. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

The discriminant wavenumbers identified in loading plots accounting for variations between the samples were  $1452\text{ cm}^{-1}$ ,  $1293\text{ cm}^{-1}$ ,  $1107\text{ cm}^{-1}$ ,  $1060\text{ cm}^{-1}$  (DNA/RNA, lipids, carbohydrates) and  $1001\text{ cm}^{-1}$  (phenylalanine) (Fig. 3.207).



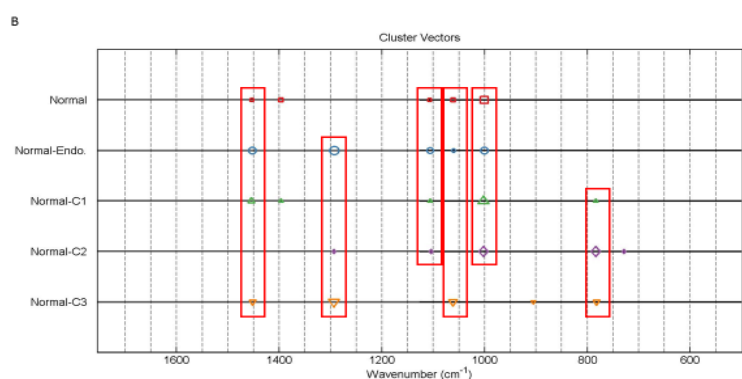
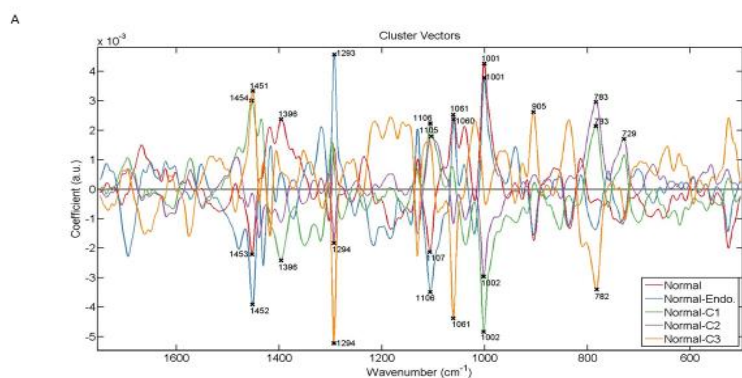
**Figure 3.207:** Loadings plots showing wavenumbers that discriminate stromal cells in non-diseased tissue sections (H09-11708-4, H09-12292-1, H09-12890-11, H09-9102-6, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vectors plots were presented and analysed in the same approach as before to identify potential biomarkers for stem cells based on the occurrence and expression level of wavenumbers and/or spectral region amongst tissue samples (Fig. 3.208).

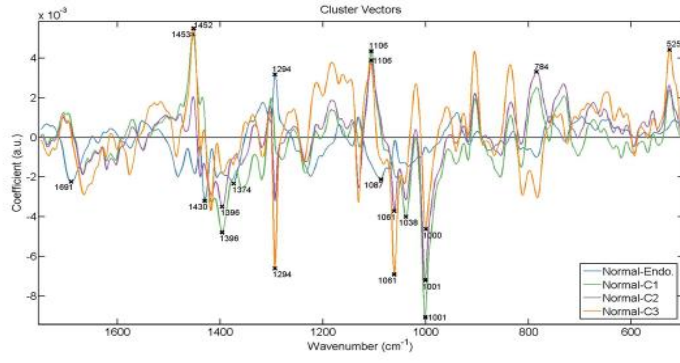
Without having any class as a reference origin,  $1454\text{ cm}^{-1}$  (elastin, collagen, phospholipids) was commonly shared with  $1453\text{ cm}^{-1}$  (structural proteins),  $1452\text{ cm}^{-1}$  and  $1451\text{ cm}^{-1}$  ( $\text{CH}_2\text{CH}_3$  deformation in collagen) in four tissues,  $1294\text{ cm}^{-1}$  (methylene twisting) with  $1293\text{ cm}^{-1}$  were commonly shared in three tissues,  $1107\text{ cm}^{-1}$  with  $1106\text{ cm}^{-1}$  (lipids and proteins),  $1105\text{ cm}^{-1}$  (carbohydrates) and  $1104\text{ cm}^{-1}$  (phenylalanine in proteins) were commonly shared in four tissues,  $1061\text{ cm}^{-1}$  (paraffin) with  $1060\text{ cm}^{-1}$  (DNA/RNA, lipids, carbohydrates) were commonly shared in three tissues,  $1002\text{ cm}^{-1}$  (phenylalanine) was commonly shared with  $1001\text{ cm}^{-1}$  (phenylalanine) in four tissues whilst  $783\text{ cm}^{-1}$  was commonly shared with  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA) in three tissues. Having class Normal (H09-11708-4) as reference, the only wavenumbers occurring commonly in three classes were  $1001\text{ cm}^{-1}$



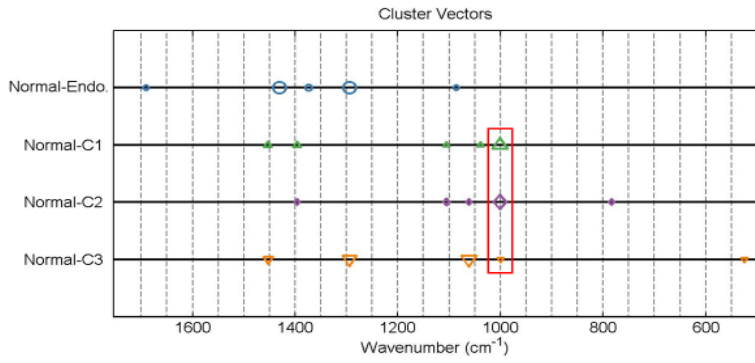
(phenylalanine) with  $1000\text{ cm}^{-1}$  (phenylalanine). Using Normal-Endo as reference, three tissues were common for  $1453\text{ cm}^{-1}$  (structural proteins) with  $1452\text{ cm}^{-1}$  and  $1451\text{ cm}^{-1}$  ( $\text{CH}_2\text{CH}_3$  deformation in collagen),  $1294\text{ cm}^{-1}$  (methylene twisting) and  $1107\text{ cm}^{-1}$  with  $1106\text{ cm}^{-1}$  (lipids and proteins). Having Normal-C1 as reference,  $1003\text{ cm}^{-1}$  (phenylalanine) was commonly shared with  $1002\text{ cm}^{-1}$  (phenylalanine) and  $1001\text{ cm}^{-1}$  (phenylalanine) in four tissues. Using Normal-C2 as reference,  $1295\text{ cm}^{-1}$  was commonly shared with  $1294\text{ cm}^{-1}$  (methylene twisting) in three tissues,  $1062\text{ cm}^{-1}$  (paraffin) with  $1061\text{ cm}^{-1}$  (paraffin) displayed common occurrence in three tissues and  $1004\text{ cm}^{-1}$  (phenylalanine) was commonly shared with  $1002\text{ cm}^{-1}$  (phenylalanine) and  $1001\text{ cm}^{-1}$  (phenylalanine) four tissues whilst  $784\text{ cm}^{-1}$  (cytosine/uracil) was commonly shared with  $783\text{ cm}^{-1}$  in three tissues. Having class Normal-C3 as reference,  $1452\text{ cm}^{-1}$  was commonly shared with  $1451\text{ cm}^{-1}$  ( $\text{CH}_2\text{CH}_3$  deformation in collagen) in three tissues,  $1295\text{ cm}^{-1}$  with  $1294\text{ cm}^{-1}$  (methylene twisting) had common occurrence in four tissues,  $1061\text{ cm}^{-1}$  (residual paraffin) with  $1060\text{ cm}^{-1}$  (DNA/RNA, lipids, carbohydrates) had common occurrence in four tissues and  $1004\text{ cm}^{-1}$  (phenylalanine) was commonly shared with  $1003\text{ cm}^{-1}$  (phenylalanine) and  $1000\text{ cm}^{-1}$  (phenylalanine) in three tissues.



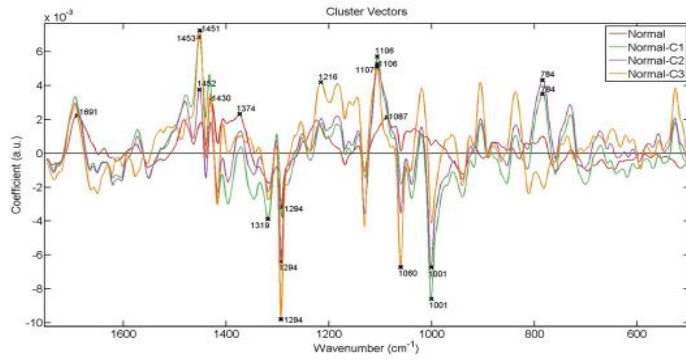
C



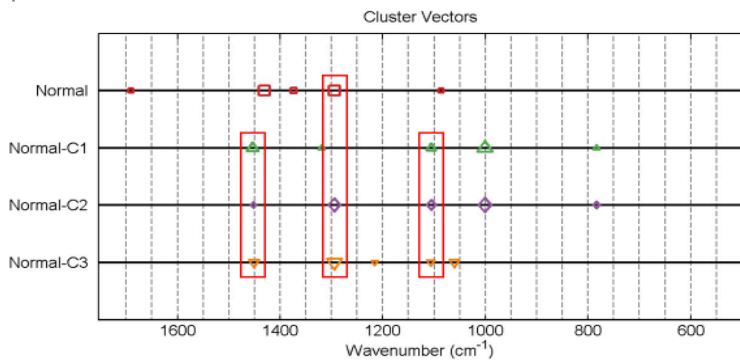
D



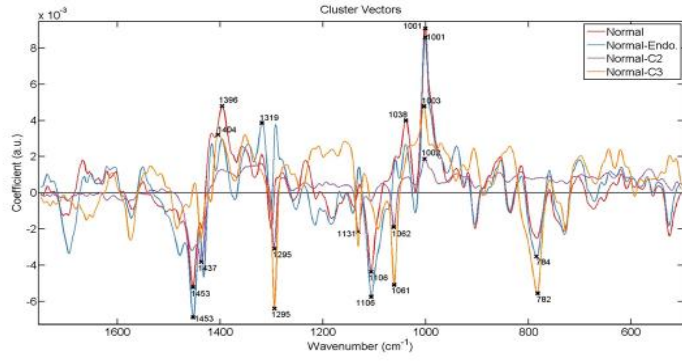
E



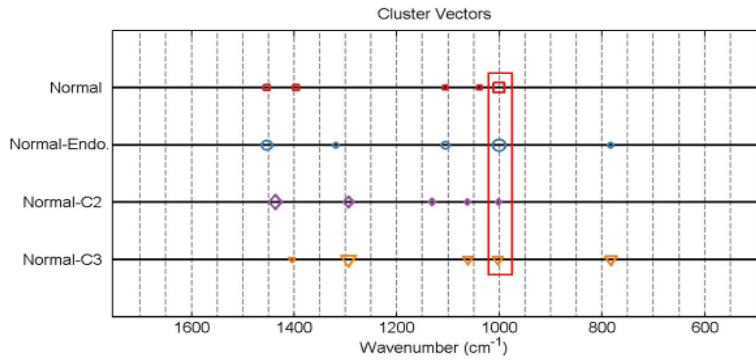
F



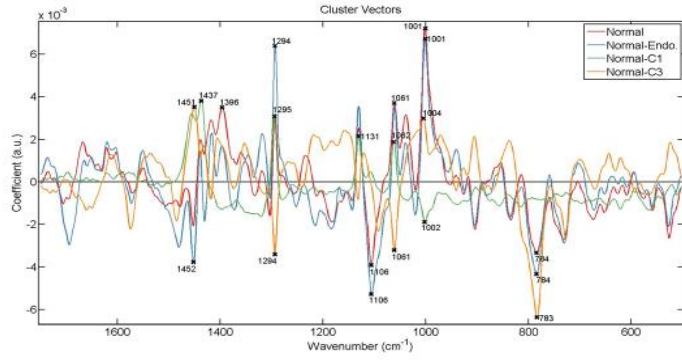
G



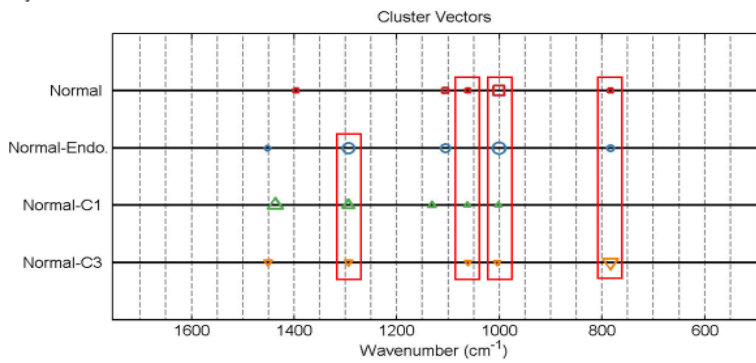
H

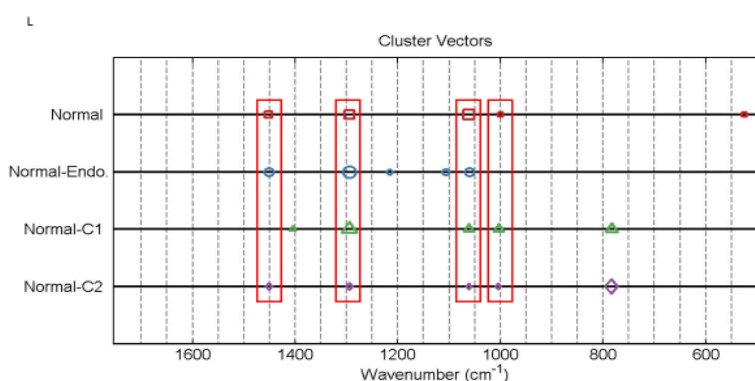
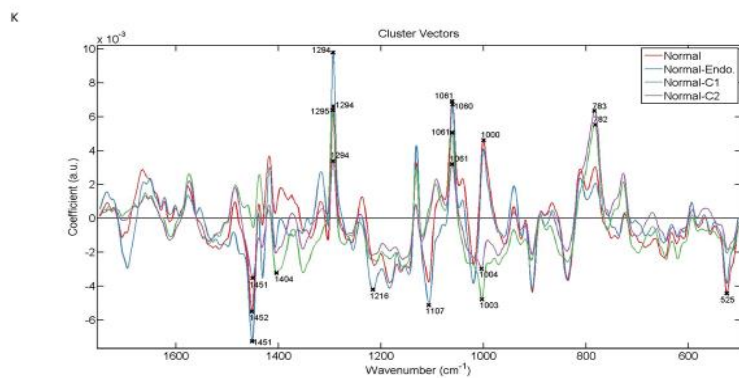


I



J





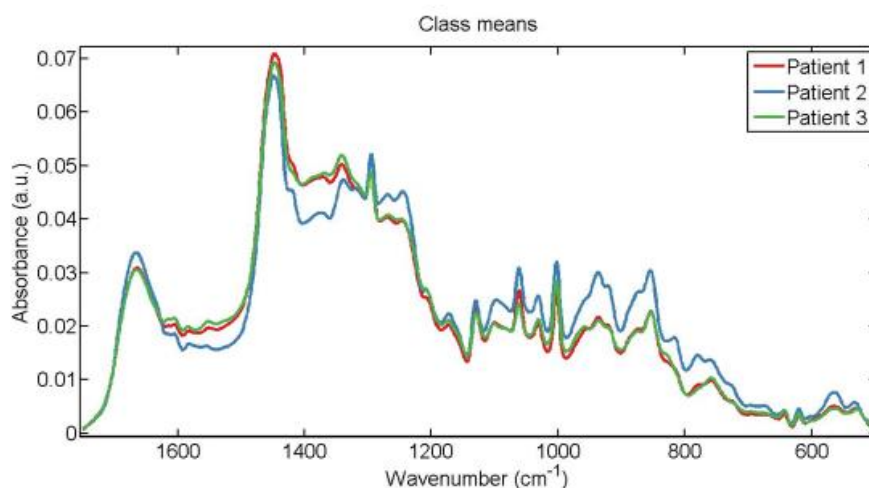
**Figure 3.208:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells, amongst non-diseased tissue sections. (A) & (B) no sample used as a reference, (C) & (D) Normal (H09-11708-4) reference class, (E) & (F) Normal-Endo. (H09-12292-1) reference class, (G) & (H) Normal-C1 (H09-12890-11) reference class, (I) & (J) Normal-C2 (H09-9102-6) reference class and (K) & (L) Normal-C3 (H09-13558-A8) reference class. The red rectangles in the cluster vector peak location plots highlight the wavenumbers occurring commonly in different tissue samples.

*Between non-diseased samples from patients with endometrial cancer*

Spectra from stromal cells in only non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer were compared. Classes representing the samples were labelled as before and will be referred accordingly to associate data analysis with what is shown on the plots. All classes exhibited a very similar Raman mean spectra whereby most pronounced variations in absorbance intensities would occur at  $\sim 1666\text{ cm}^{-1}$ ,  $\sim 1447\text{ cm}^{-1}$ ,  $\sim 1340\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1258\text{ cm}^{-1}$ ,  $\sim 1130\text{ cm}^{-1}$ ,  $\sim 1061\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 937\text{ cm}^{-1}$ ,  $\sim 857\text{ cm}^{-1}$  and  $\sim 758\text{ cm}^{-1}$  (Fig. 3.209). Patient 2 was found to have the highest peak intensities except at  $\sim 1447\text{ cm}^{-1}$  where it

exhibited the lowest peak while Patient 1 had the highest peak. It was observed that Patient 1 and Patient 3 exhibited very similar peak intensities.

Statistical analysis indicated that Patient 2 was significantly different from Patient 1 and Patient 3 whereas the two latter were no significantly different from each other (Fig. 3.210).

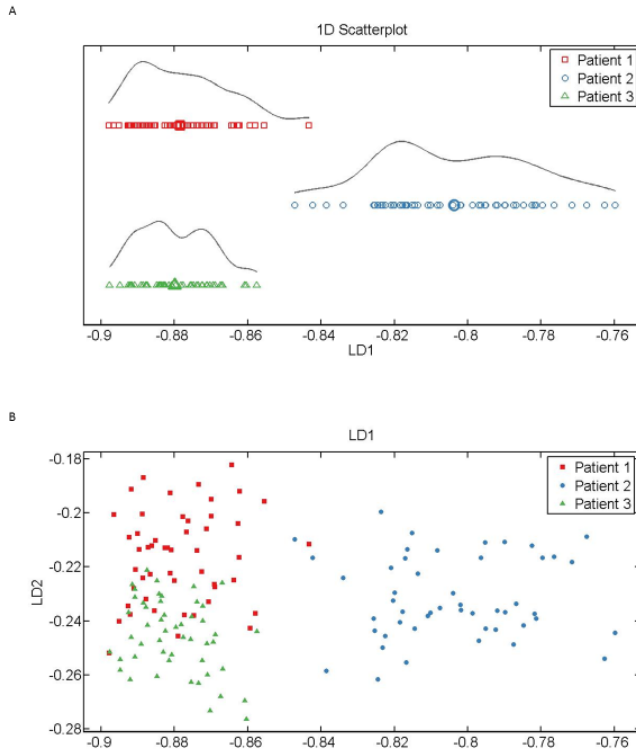


**Figure 3.209:** Mean Raman spectra for comparison of spectra acquired from stromal cells in non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.001
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P > 0.05
Patient 2 vs Patient 3	P < 0.001

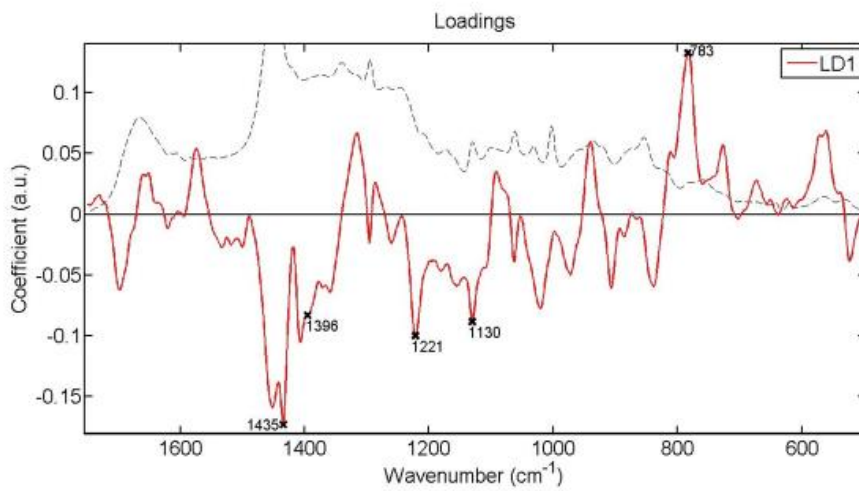
**Figure 3.210:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal cells between non-diseased tissue sections H09-9102-6, H09-12890-11, H09-13558-A8 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Application of PCA-LDA resulted in segregation of classes which was visualised in scores plots of 1D and 2D space (Fig. 3.211). Patient 2 would just overlap with Patient 1 only. The whole spectra of Patient 3 overlapped with most of the spectra of patient 1. Spectral points in Patient 1 and Patient 3 appeared to have a more compact arrangement and thus less intra-class variation relative to Patient 2. Despite the overlapping between classes, the dimensional space occupied by a class as an individual could be easily distinguished.



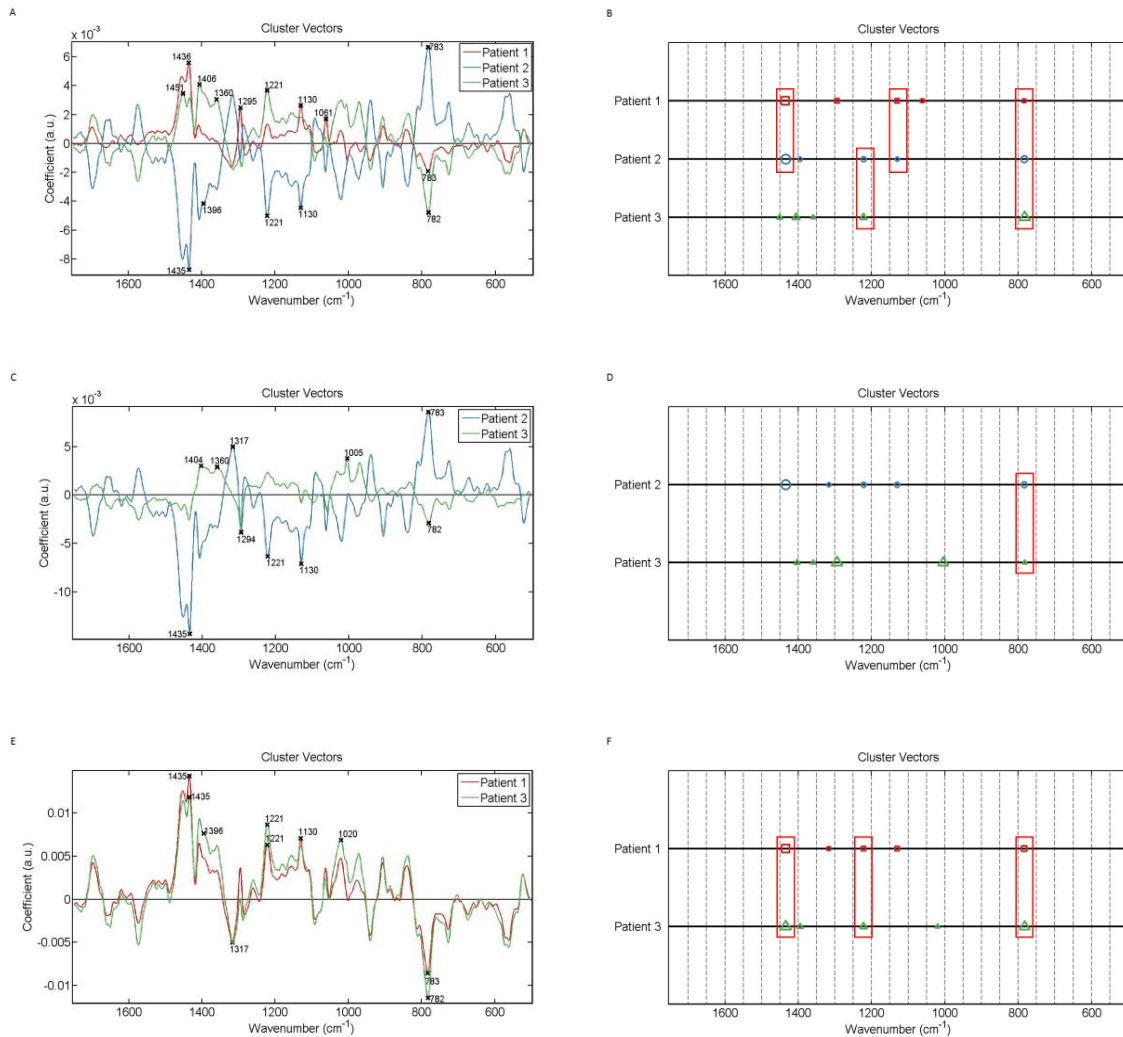
**Figure 3.211:** Scores plots for segregation of spectra in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-11, H09-9102-6 and H09-13558-A8 respectively. In (B) LD1 would discriminate spectral points between the different classes of tissues whereas as LD2 contributed to intra-class variation.

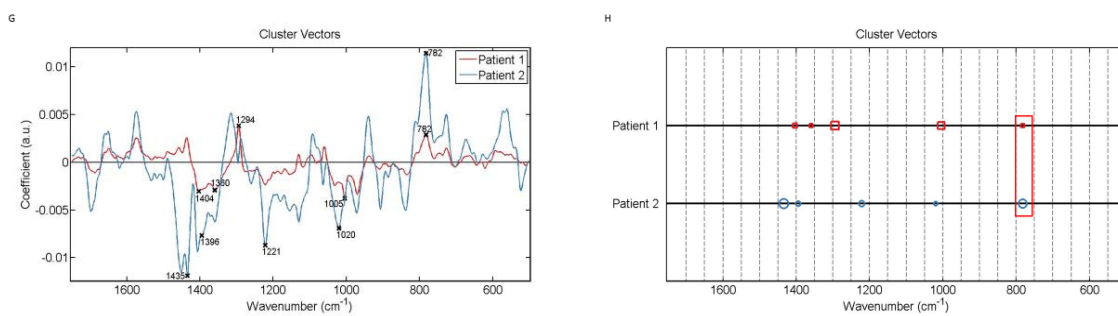
The five major discriminant wavenumbers identified in loading plots (Fig. 3.212) accounting for variations between classes and thus tissue samples were  $1435\text{ cm}^{-1}$ ,  $1396\text{ cm}^{-1}$ ,  $1221\text{ cm}^{-1}$  (amide III;  $\beta$  sheet),  $1130\text{ cm}^{-1}$  (lipids) and  $783\text{ cm}^{-1}$ .



**Figure 3.212:** Loadings plots showing wavenumbers that discriminate stromal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots were presented and analysed like previously to identify potential biomarkers for stromal cells based on occurrence and expression levels of wavenumbers and/or spectral regions between classes representing the tissues (Fig. 3.213). Having no class as a reference,  $1436\text{ cm}^{-1}$  (lipids) was commonly shared with  $1435\text{ cm}^{-1}$  in two tissues,  $1221\text{ cm}^{-1}$  (amide III;  $\beta$  sheet) and  $1130\text{ cm}^{-1}$  were common in two tissues as well, whilst all three classes were common for  $783\text{ cm}^{-1}$  with  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA). Using Patient 1 as reference,  $783\text{ cm}^{-1}$  was commonly shared with  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA). Using Patient 2 as reference,  $1435\text{ cm}^{-1}$ ,  $1221\text{ cm}^{-1}$  (amide III;  $\beta$  sheet), and  $783\text{ cm}^{-1}$  with  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA) occurred commonly in classes. Using Patient 3 as reference only  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA) showed common occurrence.



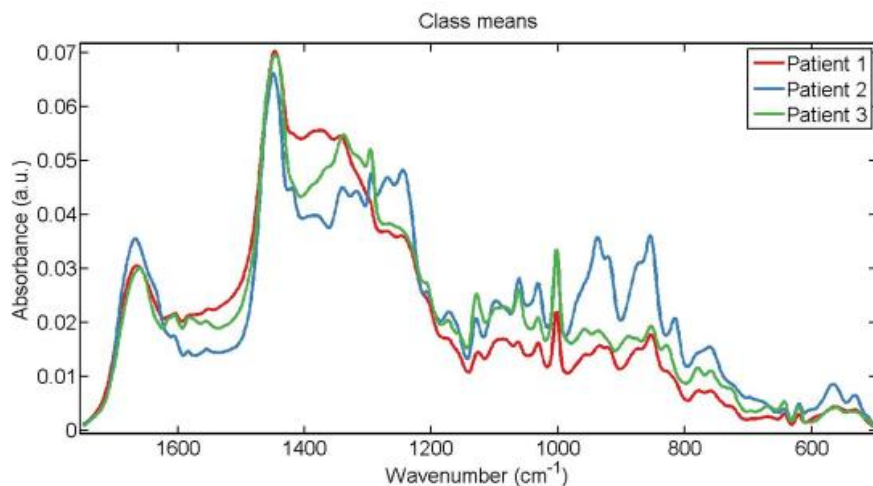


**Figure 3.213:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells among non-diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-11) reference class, (E) & (F) Patient 2 (H09-9102-6) reference class and (G) & (H) Patient 3 (H09-13558-A8) reference class. The red rectangles highlight the wavenumbers occurring commonly in classes.

### *Between diseased samples from patients with endometrial cancer*

Spectra acquired from stromal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) taken from cancerous lesions in the endometrium were analysed and compared. Classes were labelled as Patient 1, Patient 2 and Patient 3 to represent the samples H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In this section classes of tissues will be referred according to their labels for the purpose of associating data analysis with what is shown on the plots. All classes exhibited a relatively similar shape in mean Raman spectra with most pronounced variations in absorbance intensities occurring at  $\sim 1667\text{ cm}^{-1}$ ,  $\sim 1447\text{ cm}^{-1}$ ,  $\sim 1341\text{ cm}^{-1}$ ,  $\sim 1294\text{ cm}^{-1}$ ,  $\sim 1128\text{ cm}^{-1}$ ,  $\sim 1062\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 938\text{ cm}^{-1}$ ,  $855\text{ cm}^{-1}$  and  $762\text{ cm}^{-1}$  (Fig. 3.214). Mainly Patient 2 had the highest peak intensities. Statistical analysis indicated that all tissues were significantly different from each other (Fig. 3.215).



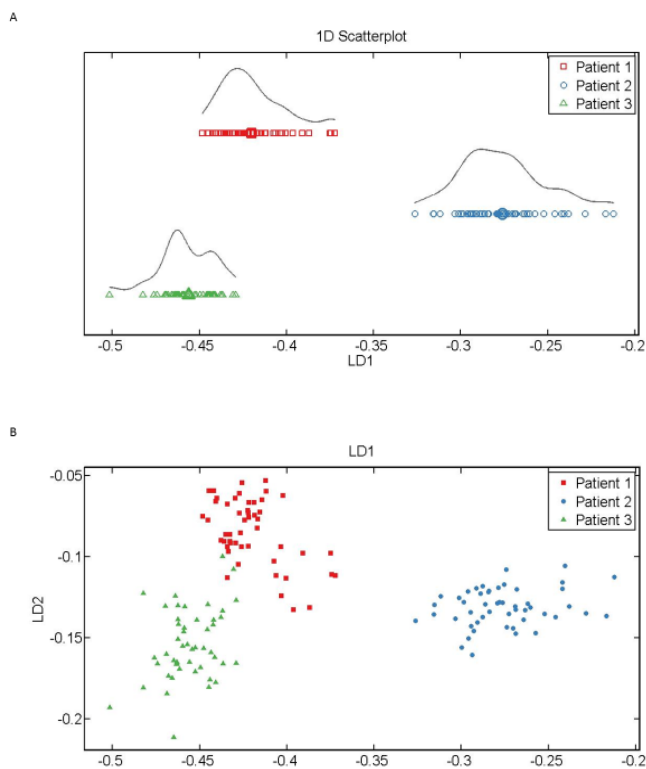


**Figure 3.214:** Mean Raman spectra for comparison of spectra from stromal cells in diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5). Classes of tissues were labelled as Patient 1, Patient 2 and Patient 3 representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively.

Parameters	P value
LDA1 vs Patient 1	P < 0.01
LDA1 vs Patient 2	P < 0.001
LDA1 vs Patient 3	P < 0.001
Patient 1 vs Patient 2	P < 0.001
Patient 1 vs Patient 3	P < 0.05
Patient 2 vs Patient 3	P < 0.001

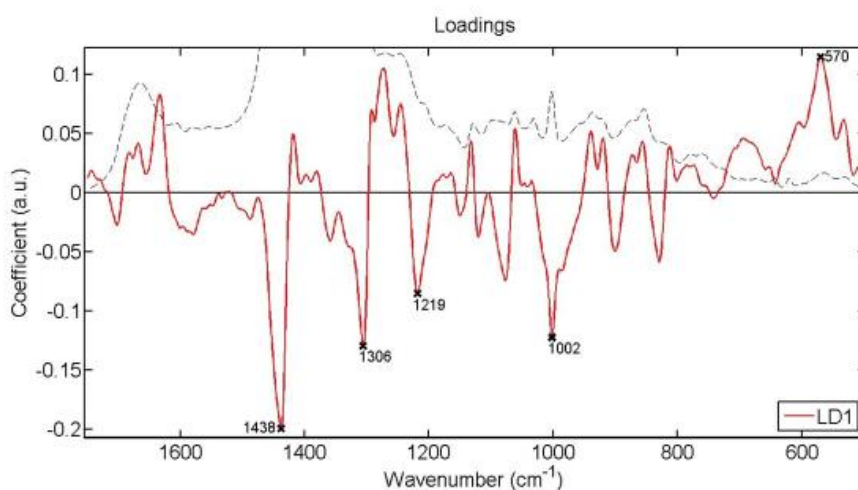
**Figure 3.215:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from stromal cells between diseased tissue sections H09-9102-1, H09-12890-9, H09-13558-A5 represented by the class labels Patient 1, Patient 2 and Patient 3 respectively.

Segregation of classes resulting from application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.216). Patient 2 did not overlap with any of the classes whilst little overlap was observed between Patient 1 and Patient 3. Spectral points in all classes had a compact arrangement rather than spread and the dimensional space occupied by each class was easily distinguished.



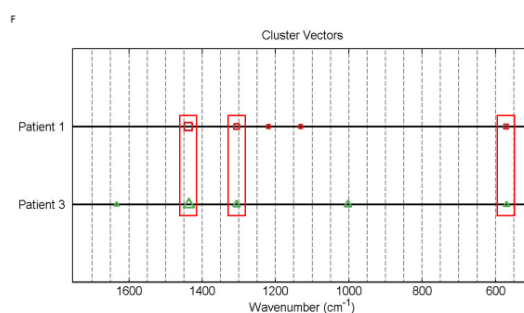
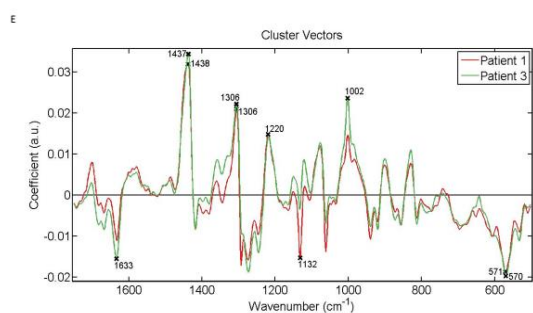
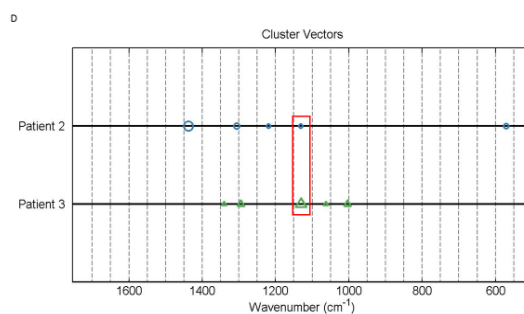
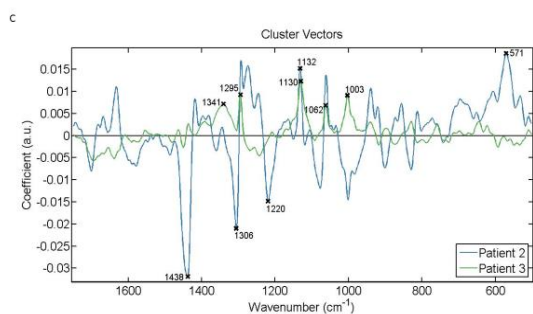
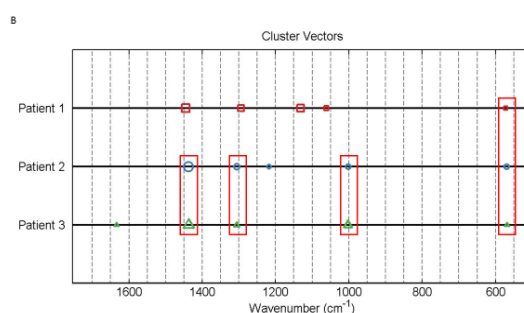
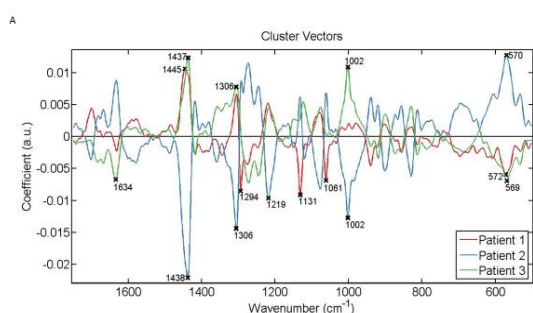
**Figure 3.216:** Scores plots for segregation of spectra acquired from stromal cells in (A) 1D and (B) 2D space. Classes of tissues were labelled as Patient 1 (red), Patient 2 (blue) and Patient 3 (green) representing the tissues H09-12890-9, H09-9102-1 and H09-13558-A5 respectively. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

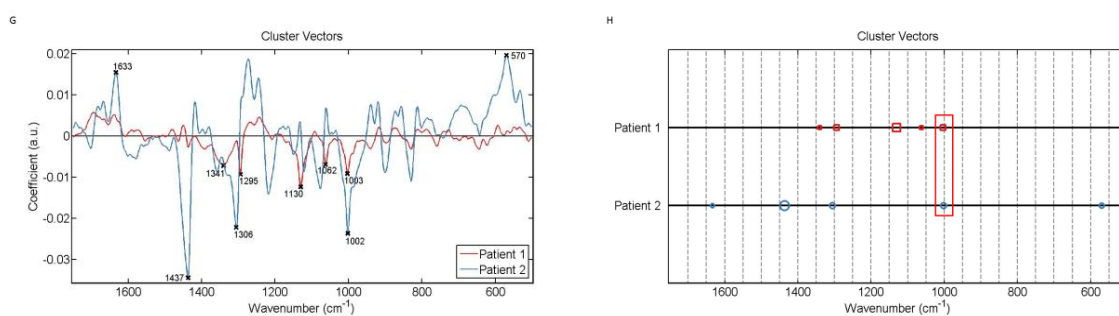
Generation of loading plots identified five major wavenumbers which would discriminate classes of tissues (Fig. 3.217). The identified wavenumbers were  $1438\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation),  $1306\text{ cm}^{-1}$ ,  $1219\text{ cm}^{-1}$ ,  $1002\text{ cm}^{-1}$  (phenylalanine) and  $570\text{ cm}^{-1}$ .



**Figure 3.217:** Loadings plots showing wavenumbers that discriminate stromal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots were presented analysed in the same approach as before in order to identify biomarkers for stromal cells (Fig. 3.218). Having no class as a reference origin, two tissues were common for  $1438\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation) with  $1437\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation in lipids),  $1306\text{ cm}^{-1}$  and  $1002\text{ cm}^{-1}$  (phenylalanine) whilst  $572\text{ cm}^{-1}$  was commonly shared with  $570\text{ cm}^{-1}$  and  $569\text{ cm}^{-1}$  in all three tissues. Having Patient 1 as reference,  $1132\text{ cm}^{-1}$  (proteins and lipids) was commonly shared with  $1130\text{ cm}^{-1}$  (lipids). Having Patient 2 as reference, common occurrence was observed for  $1438\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation) with  $1437\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation in lipids),  $1306\text{ cm}^{-1}$  and  $571\text{ cm}^{-1}$  with  $570\text{ cm}^{-1}$ . Having Patient 3 as reference, only  $1003\text{ cm}^{-1}$  (phenylalanine) was commonly shared with  $1002\text{ cm}^{-1}$  (phenylalanine).

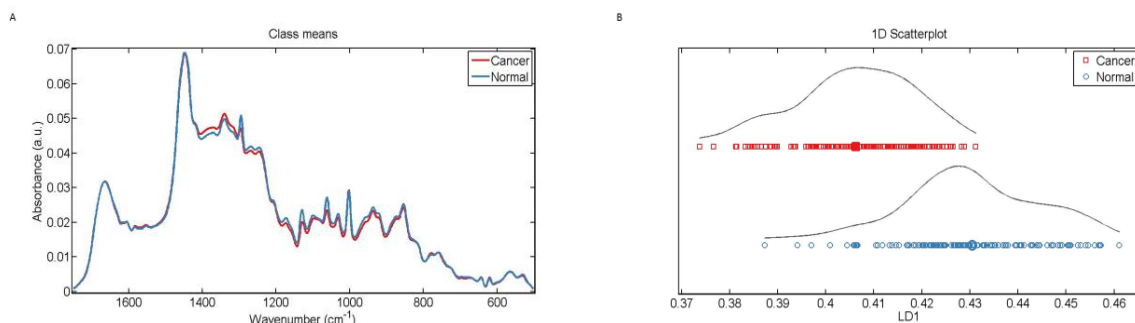




**Figure 3.218:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells among diseased tissue sections from patients with endometrial cancer. (A) & (B) no class used as a reference, (C) & (D) Patient 1 (H09-12890-9) reference class, (E) & (F) Patient 2 (H09-9102-1) reference class and (G) & (H) Patient 3 (H09-13558-A5) reference class. The red rectangles highlight the wavenumbers occurring

### *Diseased tissues Vs Non-Diseased tissues*

The spectra of the stromal cells in diseased tissue samples (H09-12890-9, H09-9102-1, H09-13558-A5) were compared with the spectra of stromal cells in their corresponding non-diseased tissue samples (H09-12890-11, H09-9102-6, H09-13558-A8) taken from patients with endometrial cancer and tissue samples were represented by the classes ‘Cancer’ and ‘Normal’ respectively. Classes of tissues will be referred according to their labels to associate data analysis with what is shown on the plots. Both classes exhibited a very similar mean Raman spectra with slight variations in absorbance intensities. Cancer had a higher intensity at  $\sim 1341\text{ cm}^{-1}$  only whilst Normal had just higher intensities at  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1129\text{ cm}^{-1}$ ,  $\sim 1062\text{ cm}^{-1}$  and  $\sim 937\text{ cm}^{-1}$  (Fig. 3.219 A). Application of PCA-LDA resulted in segregation of classes which was visualised in scores plot of 1D space whereby most of the spectra of Cancer overlapped with almost half of the spectra of Normal (Fig. 3.219 B). Statistical analysis indicated that classes were not significantly different from each other (Fig. 3.220).

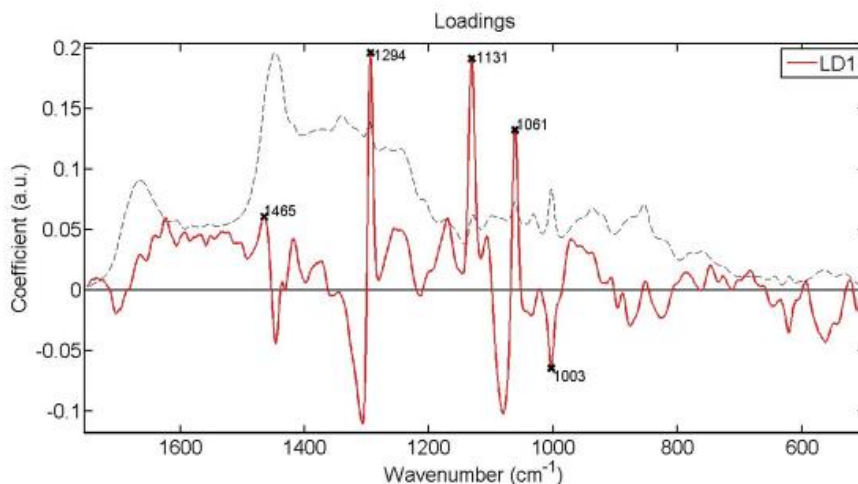


**Figure 3.219:** (A) Mean Raman spectra for comparison of spectra from stromal cells in all diseased tissue sections (H09-12890-9, H09-9102-1, H09-13558-A5) and their corresponding non-diseased tissue sections (H09-12890-11, H09-9102-6, H09-13558-A8). (B) Scores plots for segregation of spectra in 1D space. Diseased and non-diseased tissue samples are represented by the class labels 'Cancer' (red) and 'Normal' (blue) respectively.

Parameters	P Value
Cancer Vs Normal	P<0.0001

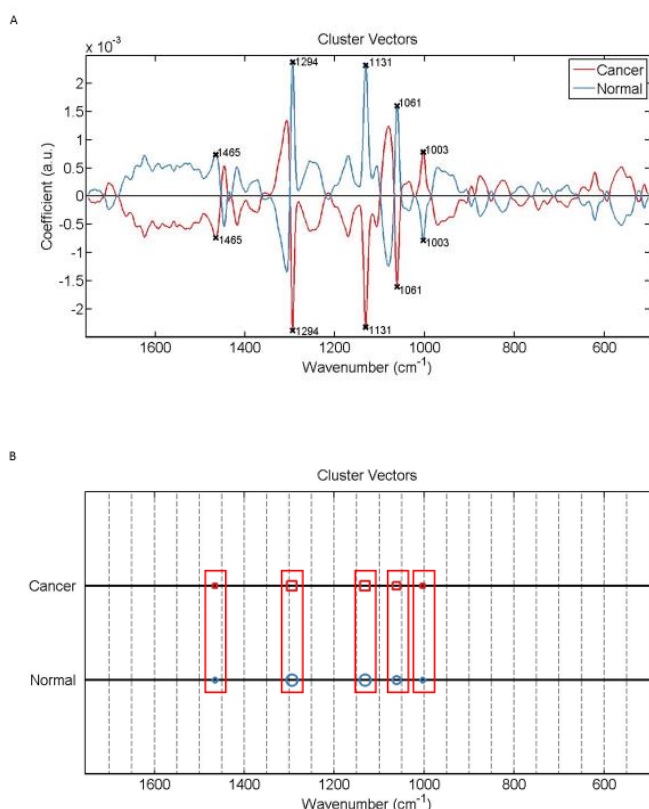
**Figure 3.220:** Obtained p-value by employment of Unpaired t-test to compare spectra acquired from stromal cells between non diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and their corresponding diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections represented by the class labels Normal and Cancer respectively.

The five major wavenumbers identified in loading responsible for discriminating the classes of tissue samples were 1465 cm<sup>-1</sup> (lipids), 1294 cm<sup>-1</sup> (methylene twisting), 1131 cm<sup>-1</sup> (lipids), 1061 cm<sup>-1</sup> (residual paraffin) and 1003 cm<sup>-1</sup> (phenylalanine) (Fig. 3.221).



**Figure 3.221:** Loadings plots showing wavenumbers that discriminate spectra from stromal cells in all diseased (H09-9102-1, H09-12890-9, H09-13558-A5) and their corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) tissue sections. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

The five major wavenumbers in loading plots were also observed in cluster vector plots occurring in both classes (Fig. 3.222). The wavenumbers displayed common occurrence in both classes and exhibited same expression levels amongst classes so none of the wavenumbers was identified to have a potential as a biomarker to specify either 'Normal' or 'Cancer' stromal cells though highest expression levels were exhibited by 1294  $\text{cm}^{-1}$  (methylene twisting) and 1131  $\text{cm}^{-1}$  (lipids) whereas 1003  $\text{cm}^{-1}$  (phenylalanine) exhibited the lowest levels. Biomarkers that would specify either 'Normal' or 'Cancer' stromal cells could not be identified since the identified discriminant wavenumbers displayed common occurrence between the classes. For a wavenumber to have a potential as a biomarker for either of the classes, should have been observed in only one class.



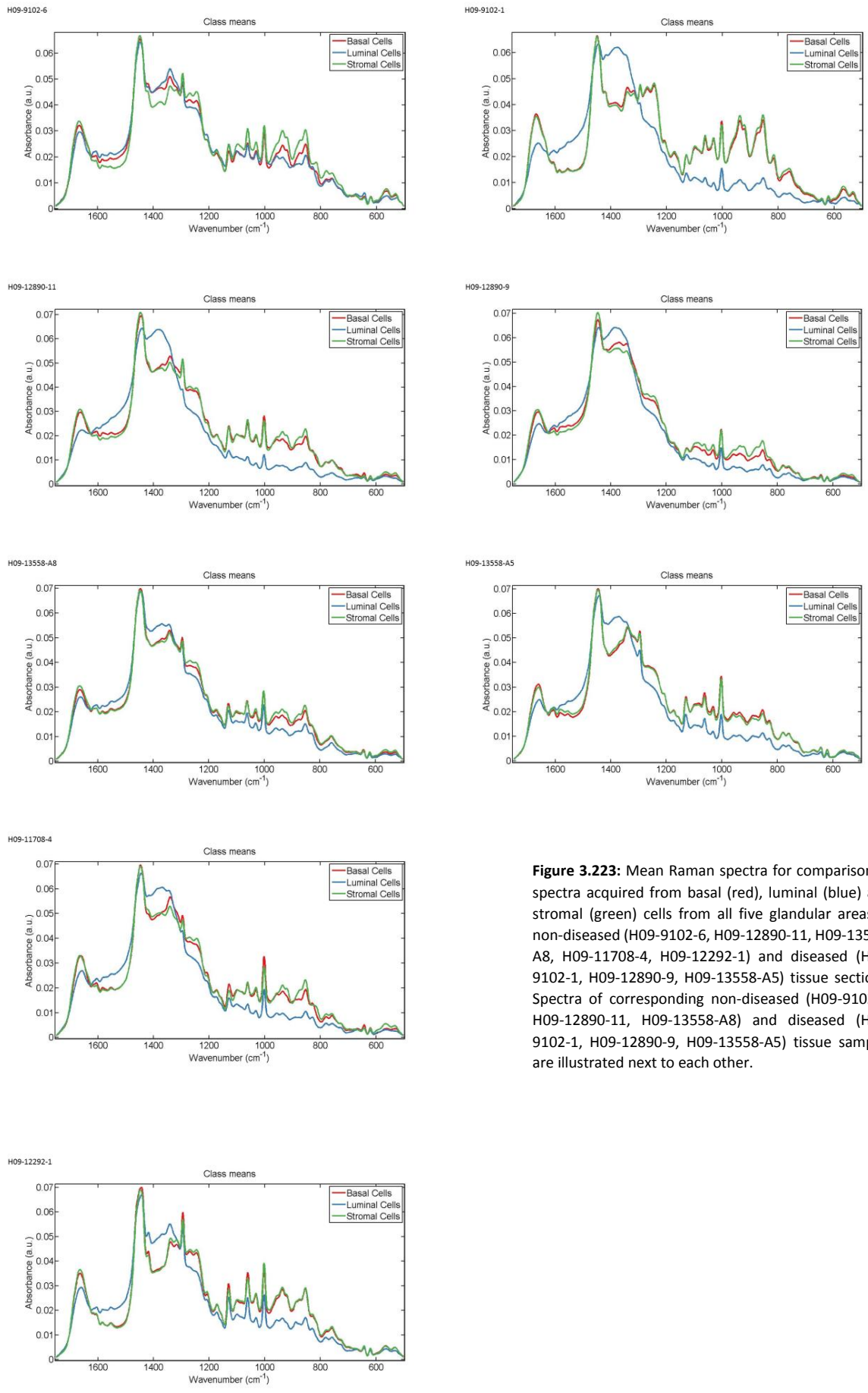
**Figure 3.222:** Alternative presentation of cluster vector plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers, from spectra acquired from stromal cells among all corresponding diseased (H09-12890-9, H09-9102-1, H09-13558-A5) and non-diseased (H09-12890-11, H09-9102-6, H09-13558-A8) tissue sections. The red rectangles highlight the wavenumbers occurring commonly in classes.

### 3.3.4 Basal Cells Vs Luminal Cells Vs Stromal Cells

#### *In individual tissue samples*

The biochemical composition of basal, luminal and stromal cells from all five glandular elements in individual tissue sections was analysed and compared. Figure 3.223 shows the mean Raman spectra of all the classes of cells acquired from all five glandular elements in each tissue section. It was observed that in each tissue the shape of the mean spectra of all three classes was relatively similar. Any variations in shape of the mean spectra would be mainly exhibited by luminal cells especially in the spectral region  $\sim 1300-1400\text{ cm}^{-1}$ . Basal and stromal cells had almost the same peak intensities which were higher than the peak intensities of luminal cells except at the spectral regions at  $\sim 1600\text{ cm}^{-1}$  and  $\sim 1340\text{ cm}^{-1}$  whereby the peaks of luminal cells were higher and at  $\sim 620\text{ cm}^{-1}$  whereby the peak intensities for all classes was relatively the same. Most pronounced variations in absorbance intensities were observed at  $\sim 1661\text{ cm}^{-1}$ ,  $\sim 1445\text{ cm}^{-1}$ ,  $\sim 1297\text{ cm}^{-1}$ ,  $\sim 1129\text{ cm}^{-1}$ ,  $\sim 1062\text{ cm}^{-1}$ ,  $\sim 1003\text{ cm}^{-1}$ ,  $\sim 942\text{ cm}^{-1}$ ,  $\sim 853\text{ cm}^{-1}$  and  $\sim 770\text{ cm}^{-1}$ . Either non-diseased or diseased tissue samples, the same observations were made. No specific observations were made that could correlate in any way non-diseased tissue samples with their corresponding diseased tissue.

Classes of cells were significantly different from each other except in tissue samples H09-12890-11, H09-12292-1, H09-9102-1 and H09-13558-A5 whereby the obtained P value suggested no significant difference between basal and stromal cells (Fig. 3.224).



**Figure 3.223:** Mean Raman spectra for comparison of spectra acquired from basal (red), luminal (blue) and stromal (green) cells from all five glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Spectra of corresponding non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue samples are illustrated next to each other.



H09-9102-6		H09-9102-1	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P < 0.001
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P > 0.05
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

H09-12890-11		H09-12890-9	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P > 0.05
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P > 0.05	Basal Cells vs Stromal Cells	P < 0.001
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

H09-13558-A8		H09-13558-A5	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P > 0.05	LDA1 vs Basal Cells	P < 0.01
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.01
Basal Cells vs Luminal Cells	P < 0.001	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P > 0.05
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

H09-11708-4		H09-12292-1	
Parameters	P value	Parameters	P value
LDA1 vs Basal Cells	P < 0.05	LDA1 vs Basal Cells	P < 0.05
LDA1 vs Luminal Cells	P < 0.001	LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001	LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.01	Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001	Basal Cells vs Stromal Cells	P > 0.05
Luminal Cells vs Stromal Cells	P < 0.001	Luminal Cells vs Stromal Cells	P < 0.001

**Figure 3.224:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from basal, luminal and stromal cells from five glandular elements in individual non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-9102-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections.

Figure 3.225 and figure 3.226 illustrate scores plots of non-diseased and diseased tissue sections respectively. In sample H09-9102-6, basal cells overlapped with both luminal and stromal cells whilst the luminal and stromal cells showed no overlapping. None of the classes displayed to have neither more nor less intra-class variation based on arrangement of spectral points within the dimensional space.

In sample H09-12890-11 all classes would overlap with each other. In the 2D scores plot, spectral points of basal cells had a more compact arrangement and thus signifying less within class variation relative to the other classes of cells.

In sample H09-13558-A8, basal cells overlapped with luminal and stromal cells whilst slight overlap was observed between luminal and stromal cells. Spectra of luminal cells had a more compact arrangement signifying less within class variation relative to the other classes.

In sample H09-11708-4 basal cells overlapped more with luminal cells than with stromal cells but no overlapping was observed between luminal and stromal cells. Spectra of stromal cells had a more spread arrangement relative to the other classes of cells signifying more intra-class variation.

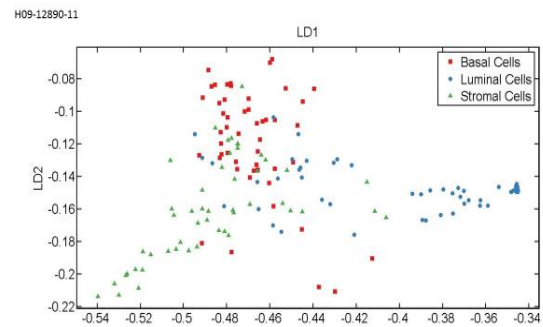
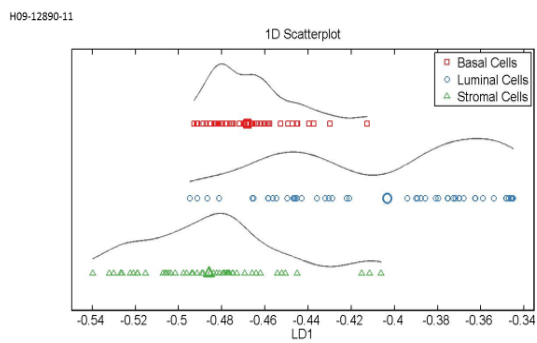
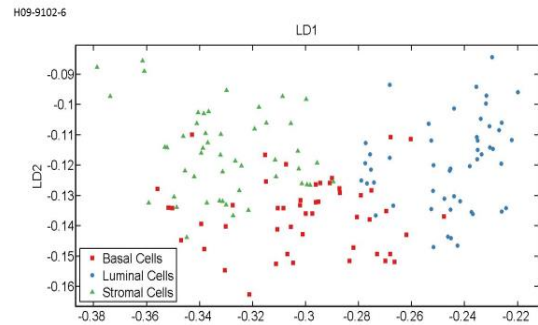
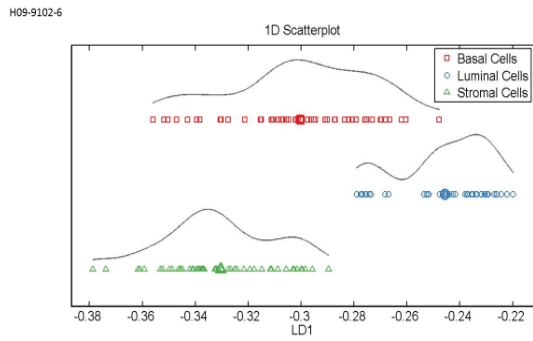
In sample H09-12292-1, all classes would overlap with each other and no significant differences were observed concerning dimensional arrangement of spectra neither intra-class variation.

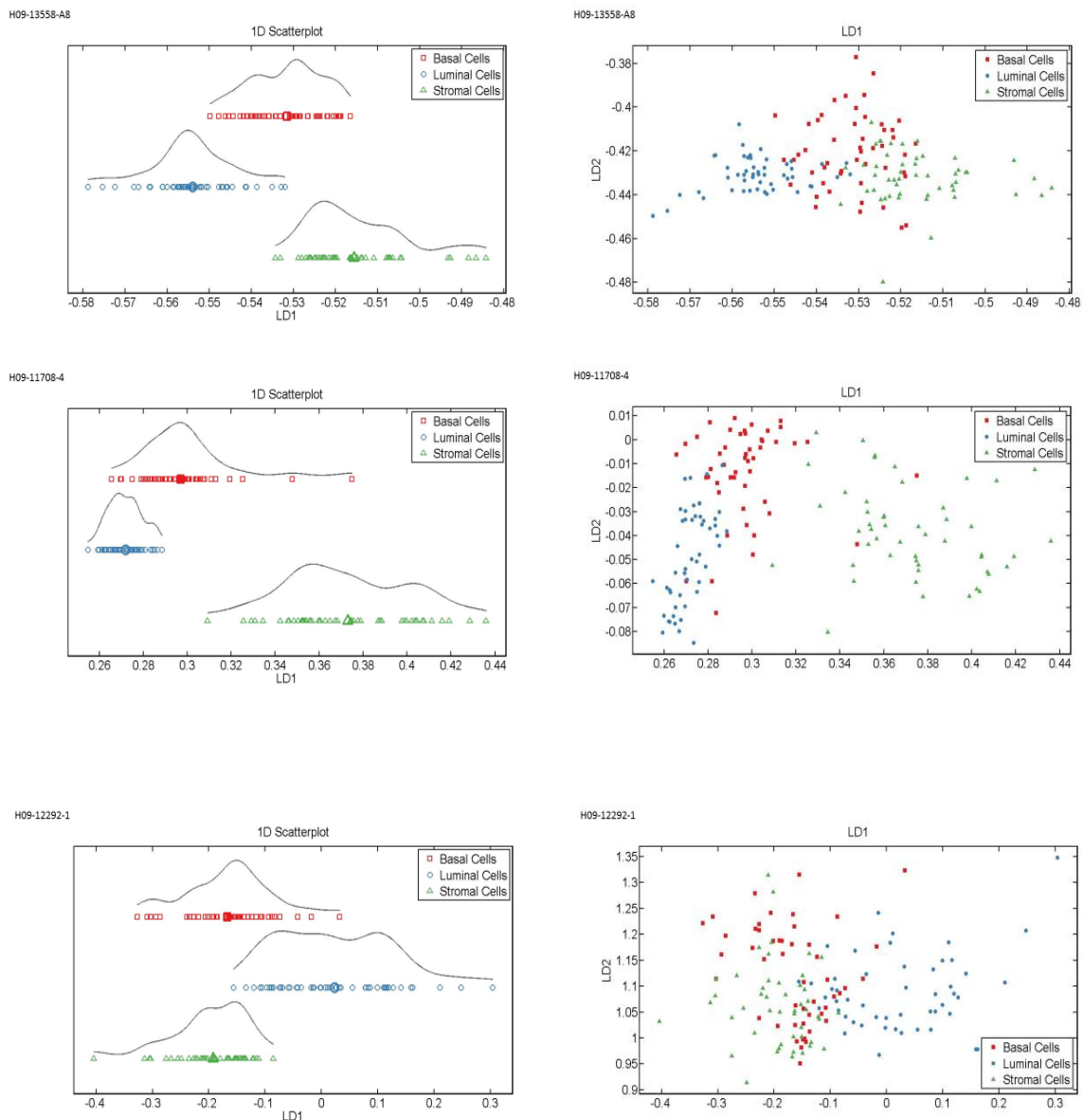
In sample H09-9102-1, all classes of cells displayed overlap with basal and stromal cells just overlap with luminal cells whilst they displayed more overlap with each other. Spectra of luminal cells was more spread relative to the other classes and thus displaying the most intra-class variation.

In sample H09-12890-9 basal cells overlapped with both luminal and stromal cells whereas no overlapping was observed between the two latter classes. Spectral points of luminal cells had a more compact arrangement which signified less intra-class variation.

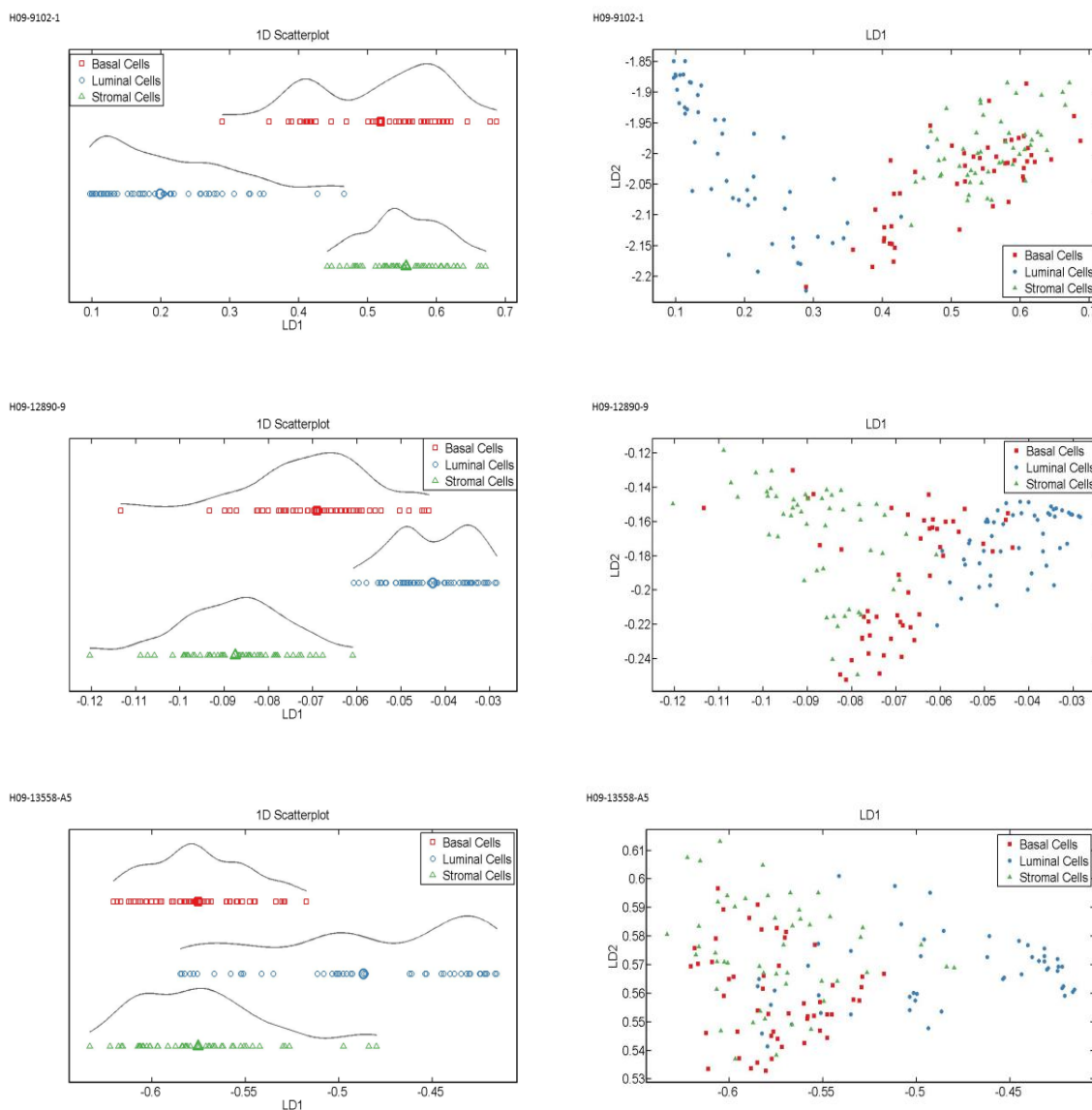
In sample H09-13558-A5 all classes overlapped with each other and spectra of basal cells had a more compact arrangement. Luminal cells showed the most segregation from basal and stromal cells so overlapping between the two latter cells was more obvious.

No specific observations were made that would associate the corresponding non-diseased and diseased tissue sections.





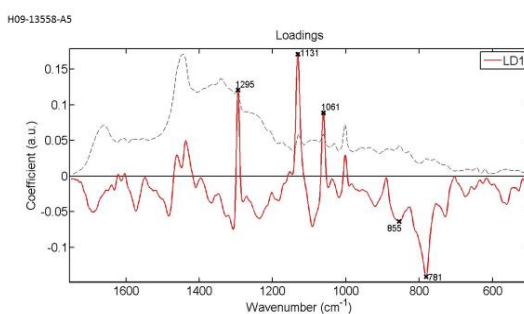
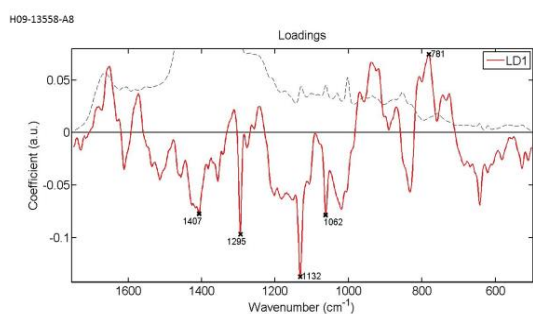
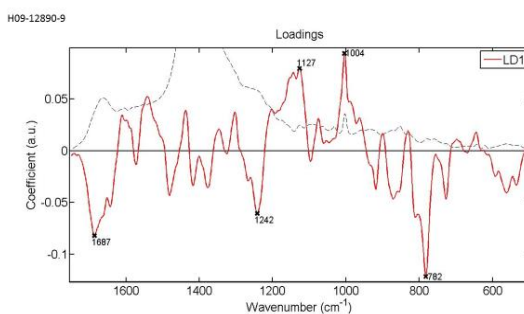
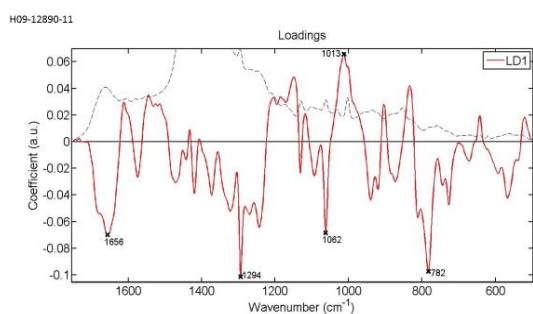
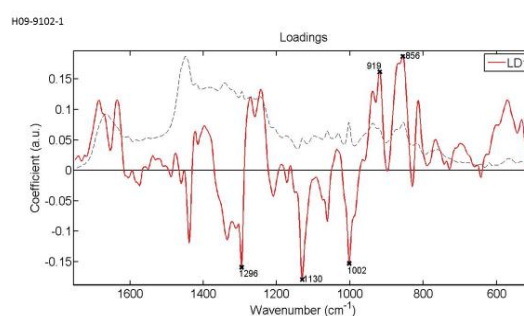
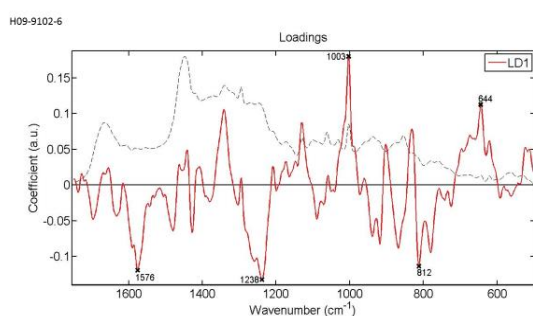
**Figure 3.225:** 1D and 2D scores plots in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) produced after application of PCA-LDA on the spectra acquired from basal (red), luminal (blue) and stromal (green) cells from all glandular elements in individual samples. In 2D plots LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

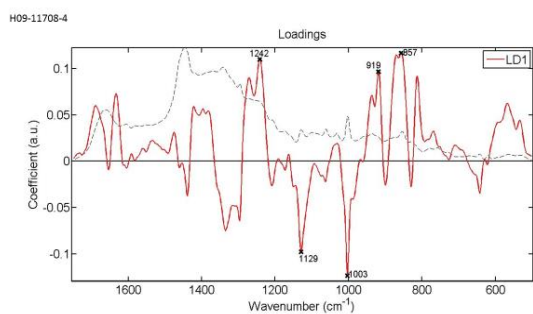


**Figure 3.226:** 1D and 2D scores plots in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA on the spectra acquired from basal (red), luminal (blue) and stromal (green) cells from all glandular elements in individual samples. In 2D plots LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

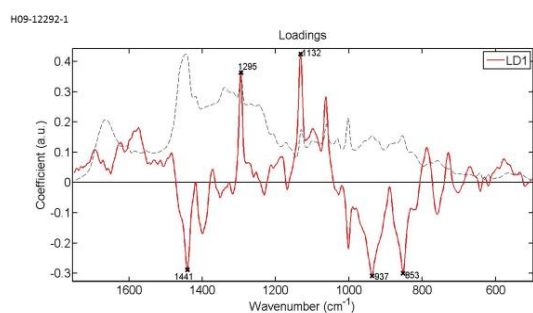
Generation of loading plots identified five major discriminant wavenumbers responsible for variations between classes of basal, luminal and stromal cells from all the interrogated glandular elements in individual samples (Fig. 3.227). The identified wavenumbers in sample H09-9102-6 were  $1576\text{ cm}^{-1}$  (nucleic acids/guanine),  $1238\text{ cm}^{-1}$ ,  $1003\text{ cm}^{-1}$  (phenylalanine),  $812\text{ cm}^{-1}$  (phosphodiester) and  $644\text{ cm}^{-1}$ . In H09-12890-11 the wavenumbers were  $1656\text{ cm}^{-1}$  (amide I/lipids),  $1294\text{ cm}^{-1}$  (methylene twisting),  $1062\text{ cm}^{-1}$  (residual paraffin),  $1013\text{ cm}^{-1}$  and  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA). In sample H09-

13558-A8 the discriminant wavenumbers were 1407  $\text{cm}^{-1}$ , 1295  $\text{cm}^{-1}$ , 1132  $\text{cm}^{-1}$  (proteins and lipids), 1062  $\text{cm}^{-1}$  (residual paraffin) and 781  $\text{cm}^{-1}$  (cytosine/uracil). In H09-11708-4 the major wavenumbers were 1242  $\text{cm}^{-1}$  (amide III;  $\beta$  sheet), 1129  $\text{cm}^{-1}$  (lipids), 1003  $\text{cm}^{-1}$  (phenylalanine), 919  $\text{cm}^{-1}$  and 857  $\text{cm}^{-1}$ . In H09-12292-1 the wavenumbers were 1441  $\text{cm}^{-1}$  (lipids), 1295  $\text{cm}^{-1}$ , 1132  $\text{cm}^{-1}$  (proteins and lipids), 937  $\text{cm}^{-1}$  (proline, protein backbone/glycogen) and 853  $\text{cm}^{-1}$  (tyrosine/proline). In the diseased sample H09-9102-1 the wavenumbers were 1296  $\text{cm}^{-1}$  (residual paraffin), 1130  $\text{cm}^{-1}$  (lipids), 1002  $\text{cm}^{-1}$  (phenylalanine), 919  $\text{cm}^{-1}$  and 856  $\text{cm}^{-1}$  (proline, hydroxyproline, tyrosine, collagen backbone). In sample H09-12890-9 the identified wavenumbers were 1687  $\text{cm}^{-1}$ , 1242  $\text{cm}^{-1}$  (amide III;  $\beta$  sheet), 1127  $\text{cm}^{-1}$ , 1004  $\text{cm}^{-1}$  (phenylalanine) and 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA). In sample H09-13558-A5 the discriminant wavenumbers were 1295  $\text{cm}^{-1}$ , 1131  $\text{cm}^{-1}$  (lipids), 1061  $\text{cm}^{-1}$  (residual paraffin), 856  $\text{cm}^{-1}$  (proline, hydroxyproline, tyrosine, collagen backbone) and 781  $\text{cm}^{-1}$  (cytosine/uracil).





**Figure 3.227:** Loadings plots showing wavenumbers that discriminate basal, luminal and stromal cells from all glandular areas in non-diseased (H09-9102-6, H09-12890-11, H09-13558-A8, H09-11708-4, H09-12292-1) and diseased (H09-9102-1, H09-12890-9, H09-13558-A5) tissue sections. Loading plots of corresponding non-diseased and diseased tissue samples are illustrated next to each other. The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.



Cluster vector plots were analysed in an alternative way in an attempt to identify potential biomarkers for basal, luminal and stromal cells (Fig. 3.228). Observations were based on expression levels of wavenumbers in the different classes of cells and which wavenumbers and/or spectral regions were commonly shared among classes and which wavenumbers occurred only in one class. The expression levels of shared wavenumbers were the same among classes, unless stated differently.

In sample H09-9102-6, the class of basal cells did not share any wavenumbers with the other classes. Wavenumbers unique for basal cells were  $1685\text{ cm}^{-1}$  (amide I),  $1440\text{ cm}^{-1}$  (lipids),  $1296\text{ cm}^{-1}$  (residual paraffin),  $1131\text{ cm}^{-1}$  (lipids) and  $919\text{ cm}^{-1}$ . The shared wavenumbers between luminal and stromal cells were  $1576\text{ cm}^{-1}$  (nucleic acids),  $1238\text{ cm}^{-1}$ ,  $1003\text{ cm}^{-1}$  (phenylalanine),  $812\text{ cm}^{-1}$  (phosphodiester) and  $644\text{ cm}^{-1}$ . Luminal cells neither stromal cells were not unique for any wavenumbers. The wavenumber shared between luminal and stromal cells was only  $1013\text{ cm}^{-1}$ .

In sample H09-12890-11 the shared wavenumbers between basal and stromal cells were  $1296\text{ cm}^{-1}$  (residual paraffin),  $1294\text{ cm}^{-1}$  (methylene twisting),  $1062\text{ cm}^{-1}$  (residual paraffin),  $783\text{ cm}^{-1}$  and  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA). Wavenumbers observed only in basal cells were  $1130\text{ cm}^{-1}$  (lipids) and

727  $\text{cm}^{-1}$  (proline), in luminal cells was only 16556  $\text{cm}^{-1}$  (amide I/lipids) and in stromal cells were 1663  $\text{cm}^{-1}$  (DNA), 1420  $\text{cm}^{-1}$  (proteins and lipids) and 938  $\text{cm}^{-1}$  (protein backbone). Basal cells and stromal cells were not found to share any wavenumbers.

In sample H09-11708-4 the wavenumbers commonly shared between basal and stromal cells were 971  $\text{cm}^{-1}$  (C-C wagging), whose expression levels were higher in basal cells, and 781  $\text{cm}^{-1}$  (cytosine/uracil), between luminal and stromal cells the shared wavenumbers were 1296  $\text{cm}^{-1}$  (paraffin), 1295  $\text{cm}^{-1}$ , 1132  $\text{cm}^{-1}$  (proteins and lipids), 1131  $\text{cm}^{-1}$  (lipids), 1062  $\text{cm}^{-1}$  (paraffin) and 1061  $\text{cm}^{-1}$  (paraffin). Wavenumbers occurring only in basal cells were 1608  $\text{cm}^{-1}$  (cytosine), 1257  $\text{cm}^{-1}$  (amide III) and 725  $\text{cm}^{-1}$  (DNA/RNA bases), in luminal and stromal cells were 1407  $\text{cm}^{-1}$  (proline and valine) and 642  $\text{cm}^{-1}$  (tyrosine) respectively.

In sample H09-117084-4 the wavenumbers shared commonly between all three classes of cells were 1243  $\text{cm}^{-1}$  (amide III) and 1242  $\text{cm}^{-1}$  (amide III;  $\beta$  sheet). Basal and stromal cells were common for 1003  $\text{cm}^{-1}$  (phenylalanine). Luminal and stromal cells were common for 1130  $\text{cm}^{-1}$  (lipids), 1129  $\text{cm}^{-1}$  (lipids), 919  $\text{cm}^{-1}$ , 857  $\text{cm}^{-1}$  and 856  $\text{cm}^{-1}$  (proline, hydroxyproline, tyrosine). Expression levels of 1130  $\text{cm}^{-1}$  (lipids) and 856  $\text{cm}^{-1}$  (proline, hydroxyproline, tyrosine) were higher in luminal cells. Wavenumbers occurring only in basal cells were 1339  $\text{cm}^{-1}$  (tryptophan), 869  $\text{cm}^{-1}$  (proline) and 812  $\text{cm}^{-1}$  (phosphodiester) and in luminal cells it was only 1296  $\text{cm}^{-1}$  (paraffin). Stromal cells did not present any unique wavenumbers.

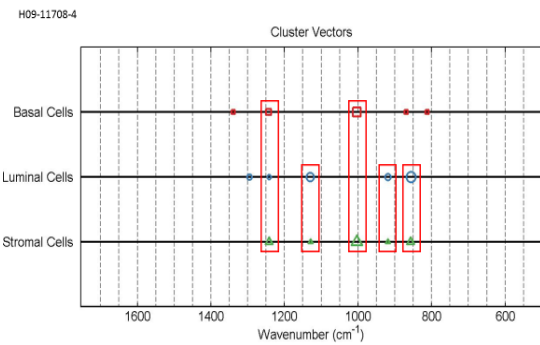
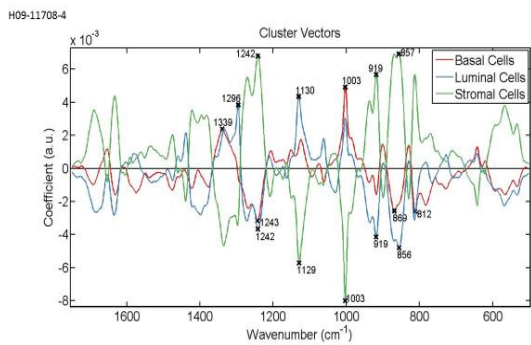
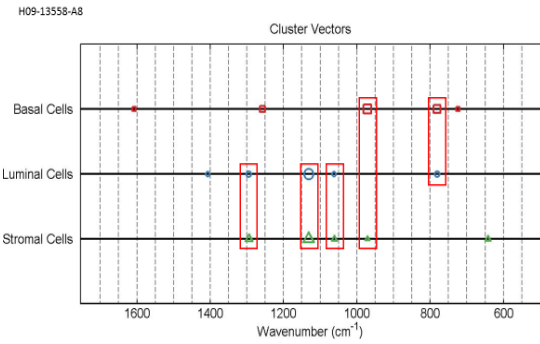
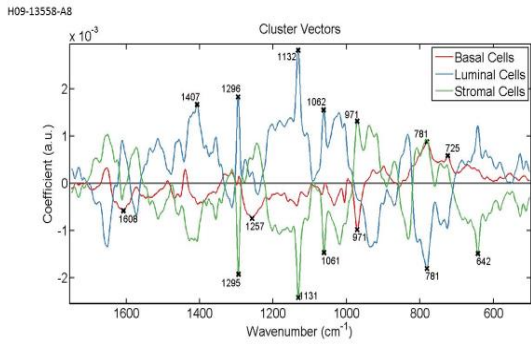
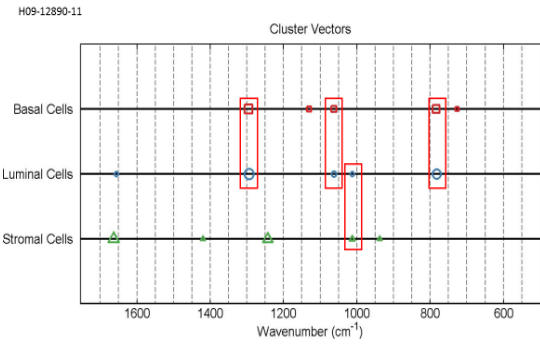
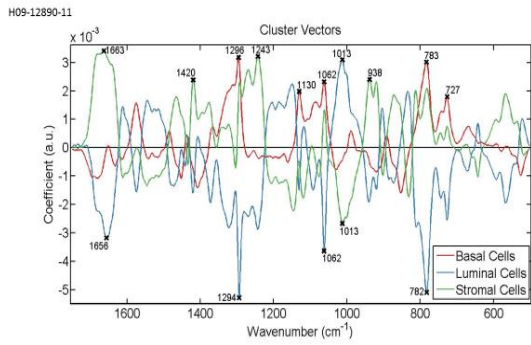
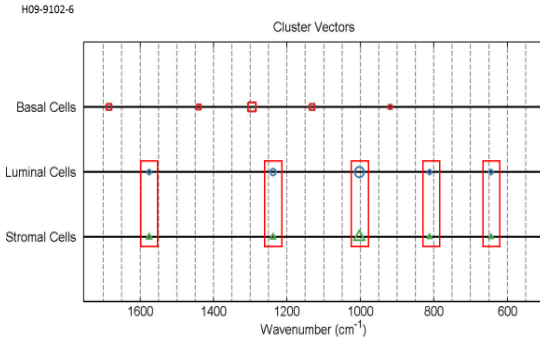
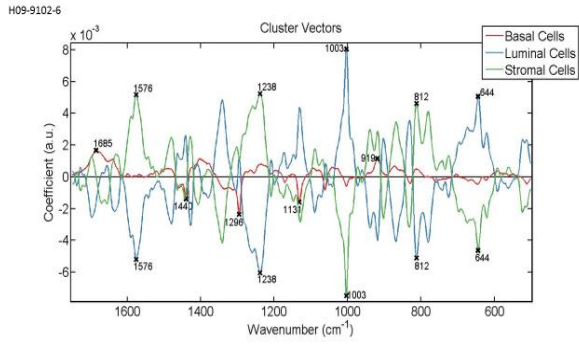
In sample H09-12292-1 all three classes of cells were common for 1296  $\text{cm}^{-1}$  (paraffin), 1295  $\text{cm}^{-1}$ , 938  $\text{cm}^{-1}$  (protein backbone), 937  $\text{cm}^{-1}$  (proline, protein backbone/glycogen), 935  $\text{cm}^{-1}$  (proline, valine, protein backbone/glycogen) and 853  $\text{cm}^{-1}$  (tyrosine/proline). Basal and luminal cells were common for 1440  $\text{cm}^{-1}$  and 1436  $\text{cm}^{-1}$ , both assigned to lipids, but expression levels were higher in basal cells. Basal and stromal cells were unique for 1397  $\text{cm}^{-1}$  and 1063  $\text{cm}^{-1}$  (C-C stretch) respectively. Luminal cells did not present any unique wavenumbers.

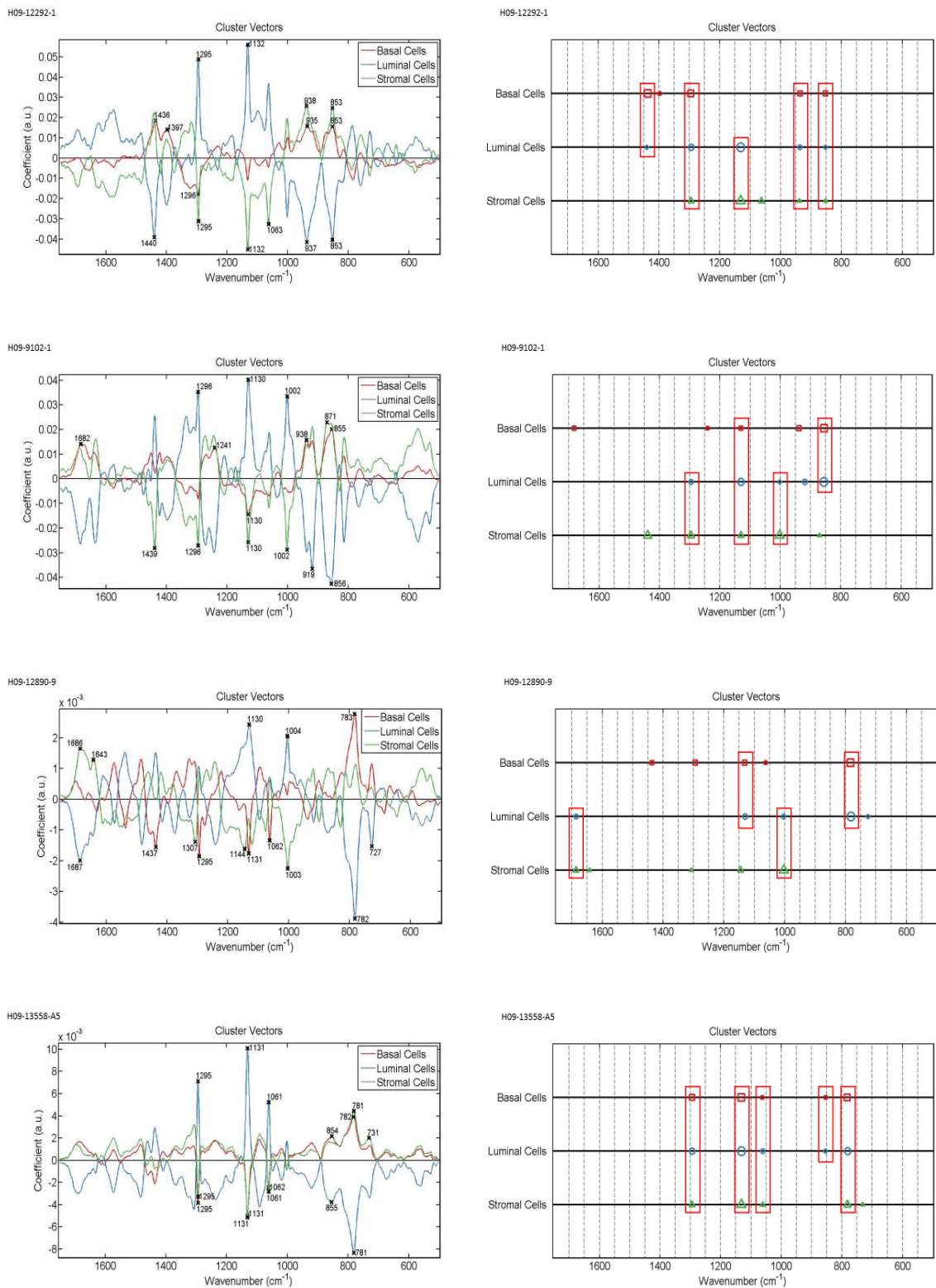


In sample H09-9102-1 all classes of cells were common for 1130  $\text{cm}^{-1}$  (lipids) only. Basal and luminal cells were common for 856  $\text{cm}^{-1}$  and 855  $\text{cm}^{-1}$ , both assigned to proline, hydroxyproline, tyrosine and collagen backbone. Luminal and stromal cells were common for 1296  $\text{cm}^{-1}$  (paraffin) and 1002  $\text{cm}^{-1}$  (phenylalanine). Expression levels of the latter were higher in stromal cell. Wavenumbers occurring only in basal cells were 1682  $\text{cm}^{-1}$  (cortisone), 1241  $\text{cm}^{-1}$  (nucleic acids) and 938  $\text{cm}^{-1}$  (protein backbone). Wavenumbers occurring only in stromal cells were 1493  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation) and 871  $\text{cm}^{-1}$ . Luminal cells were unique only for 919  $\text{cm}^{-1}$ .

In sample H09-12890-9 luminal and stromal cells were common for 1687  $\text{cm}^{-1}$ , 1686  $\text{cm}^{-1}$ , 1004  $\text{cm}^{-1}$  and 1003  $\text{cm}^{-1}$ , both assigned to phenylalanine and expression levels were higher in stromal cells. Basal and luminal cells were common for 1131  $\text{cm}^{-1}$  (lipids), 1130  $\text{cm}^{-1}$  (lipids), 783  $\text{cm}^{-1}$  and 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA). No wavenumber was found to be shared in common between basal and stromal cells. Wavenumbers occurring only in basal cells were 1437  $\text{cm}^{-1}$  (lipids and proteins), 1295  $\text{cm}^{-1}$  and 1062  $\text{cm}^{-1}$  (paraffin). Luminal cells were unique only for 727  $\text{cm}^{-1}$  (proline). Stromal cells were found unique for 1643  $\text{cm}^{-1}$ , 1307  $\text{cm}^{-1}$  (lipids/collagen) and 1144  $\text{cm}^{-1}$ .

In sample H09-13558-A5 the wavenumbers shared in common between all classes of cells were 1295  $\text{cm}^{-1}$ , 1131  $\text{cm}^{-1}$  (lipids), 1062  $\text{cm}^{-1}$  (paraffin), 1061  $\text{cm}^{-1}$  (paraffin), 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA) and 781  $\text{cm}^{-1}$  (cytosine/uracil). Basal and stromal cells were common for 855  $\text{cm}^{-1}$  and 854  $\text{cm}^{-1}$  both assigned to tyrosine/proline. Only stromal cells were found unique for 731  $\text{cm}^{-1}$ .

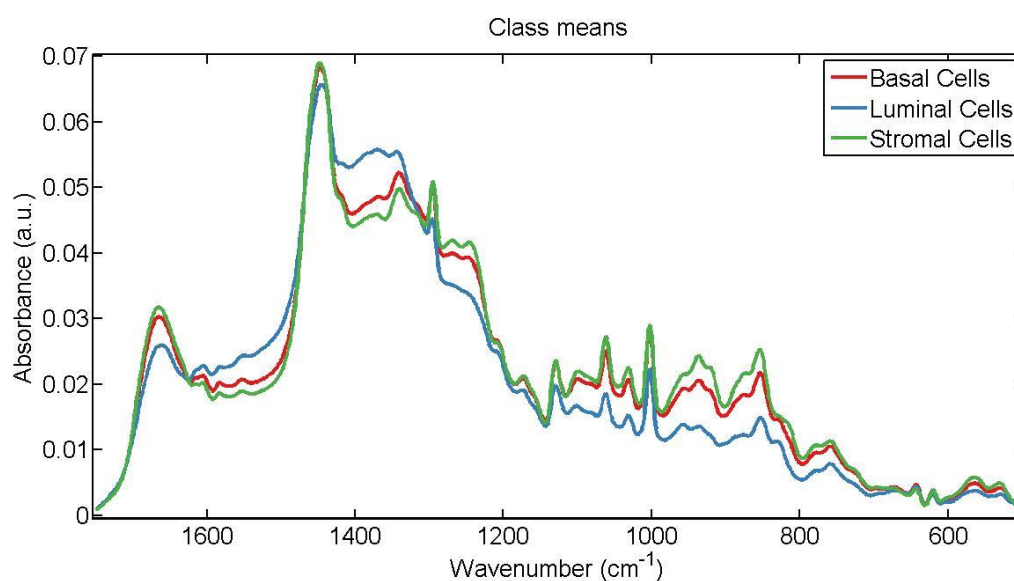




**Figure 3.228:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers among basal (red), luminal (blue) and stromal (green) cells from all glandular areas in individual non-diseased (H09-9102-1, H09-12890-9, H09-13558-A5, H09-11708-4, H09-12292-1) and diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight wavenumbers shared in common between classes of cells.

*In non-diseased tissues*

The biochemical composition of all basal, luminal and stromal cells in non-diseased tissue samples (H09-9102-6, H09-12890-11, H09-13558-A8) taken from patients with endometrial cancer was analysed. All classes of cells exhibited relatively the same shape in mean Raman spectra whereby evident variations in absorbance intensities were observed at  $\sim 1655\text{ cm}^{-1}$ ,  $\sim 1448\text{ cm}^{-1}$ ,  $\sim 1342\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1128\text{ cm}^{-1}$ ,  $\sim 1061\text{ cm}^{-1}$ ,  $\sim 1031\text{ cm}^{-1}$ ,  $\sim 936\text{ cm}^{-1}$ ,  $\sim 854\text{ cm}^{-1}$  and  $\sim 758\text{ cm}^{-1}$  (Fig. 3.229). Stromal cells had the highest peak intensities which were slightly higher than intensities of basal cells. Luminal cells had the lowest peak intensities except at  $\sim 1342\text{ cm}^{-1}$  whereby the peak was the highest. Classes of cells were significantly different from each other as indicated by the *P* values obtained from statistical analysis (Fig. 3.230).

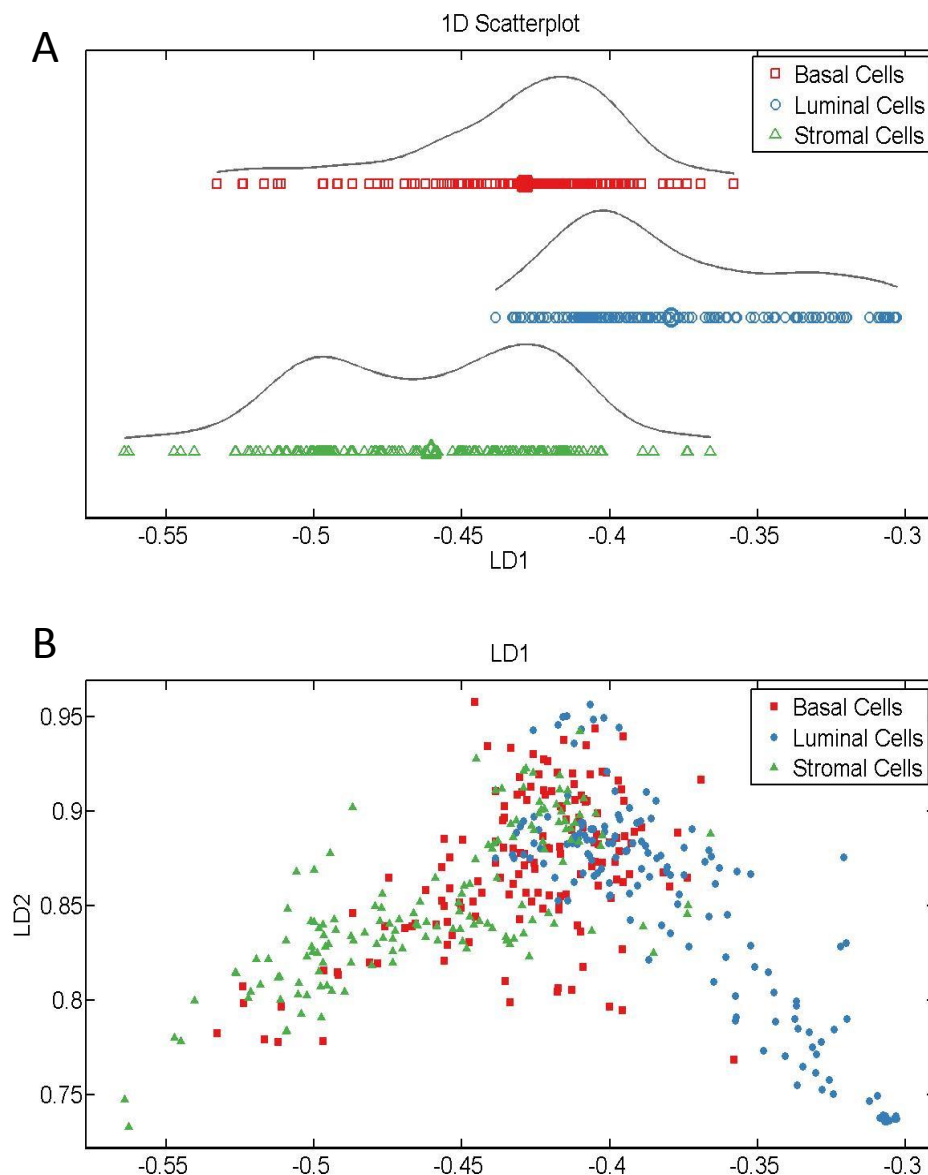


**Figure 3.229:** Mean Raman spectra for comparison of spectra acquired from all basal (red), luminal (blue) and stromal (green) cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8) to analyse their biochemical composition.

Parameters	P value
LDA1 vs Basal Cells	P > 0.05
LDA1 vs Luminal Cells	P < 0.001
LDA1 vs Stromal Cells	P < 0.001
Basal Cells vs Luminal Cells	P < 0.001
Basal Cells vs Stromal Cells	P < 0.001
Luminal Cells vs Stromal Cells	P < 0.001

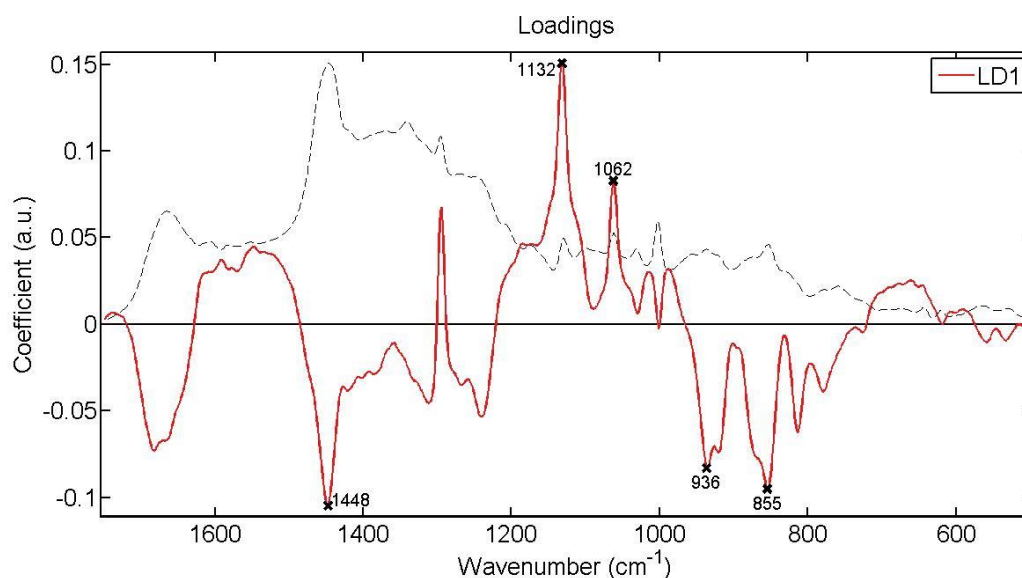
**Figure 3.230:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from all basal, luminal and stromal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8).

Segregation of classes after application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.231). Basal and stromal cells overlapped with each other throughout most of their spectra. Almost more than half of spectra collected from luminal cells overlapped with both other classes. Spectral points of all classes had a compact arrangement in the dimensional space where overlapping between all three classes occurred. However, no class was identified as having more or less intra-class variation since spectral points in individual classes exhibited relatively the same dimensionality.



**Figure 3.231:** Scores plots of (A) 1D and (B) 2D space illustrating segregation of spectra acquired from basal (red), luminal (blue) and stromal (green) cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8) produced after application of PCA-LDA. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

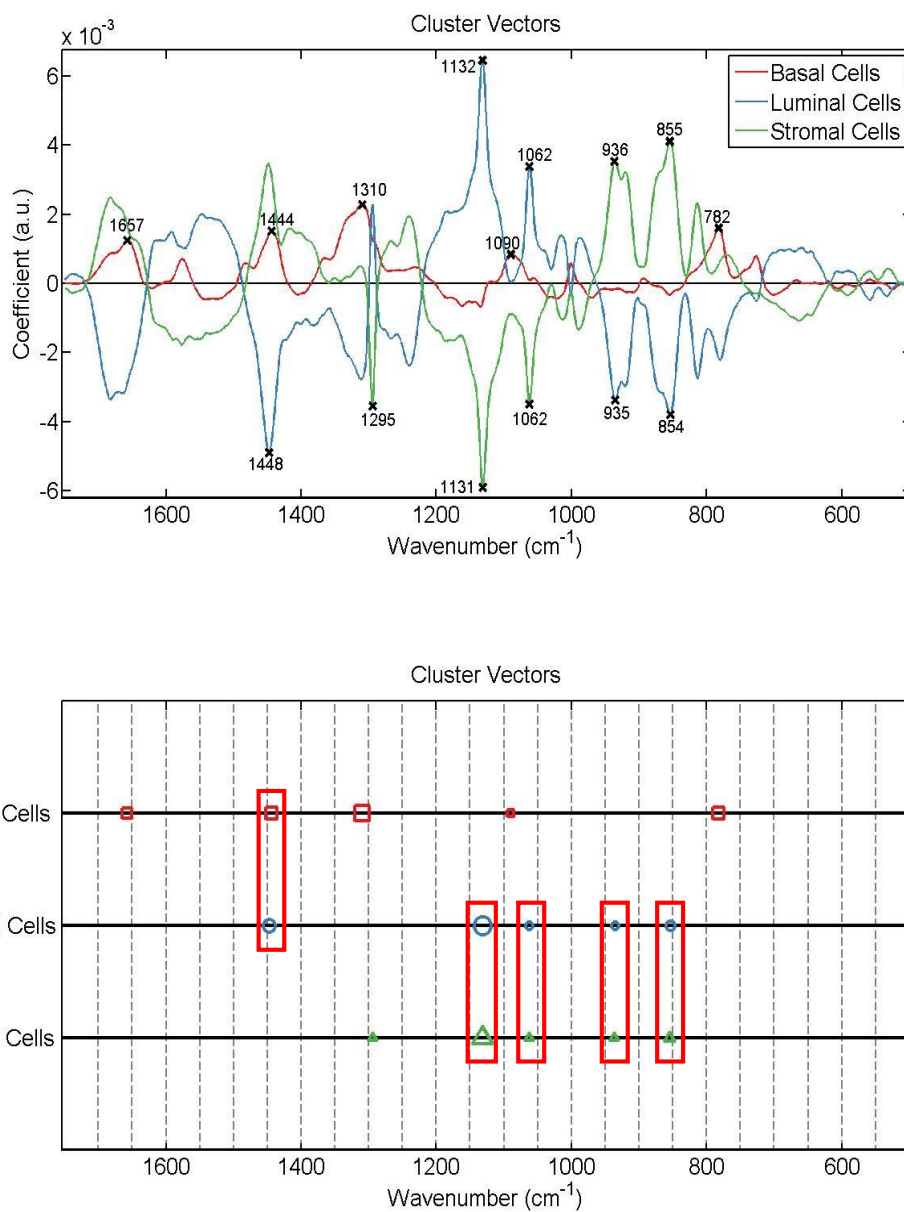
The five major discriminant wavenumbers identified in loading plots (Fig. 3.232) accounting for variations between classes of cells were  $1448\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation, collagen),  $1132\text{ cm}^{-1}$  (proteins and lipids),  $1062\text{ cm}^{-1}$  (paraffin),  $936\text{ cm}^{-1}$  (valine/proline and protein) and  $855\text{ cm}^{-1}$  (proline, hydroxyproline, tyrosine/collagen backbone).



**Figure 3.232:** Loadings plots showing five major wavenumbers that discriminate basal, luminal and stromal cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

Cluster vector plots were analysed in an alternative way in an attempt to identify potential biomarkers for basal, luminal and stromal cells in non-diseased tissue samples (Fig. 3.233). Observations were based on expression levels of wavenumbers in the different classes of cells and which wavenumbers and/or spectral regions were commonly shared among classes and which wavenumbers occurred only in one class. The expression levels of shared wavenumbers were the same among classes. Basal cells and luminal cells were found to be common for  $1444\text{ cm}^{-1}$  (cholesterol) and  $1448\text{ cm}^{-1}$  ( $\text{CH}_2$  deformation). Luminal and stromal cells were common for  $1132\text{ cm}^{-1}$  (proteins and lipids),  $1131\text{ cm}^{-1}$  (lipids),  $1062\text{ cm}^{-1}$  (paraffin),  $935\text{ cm}^{-1}$  (proline, valine, protein backbone/glycogen),  $936\text{ cm}^{-1}$  (valine, proline, protein),  $854\text{ cm}^{-1}$  and  $855\text{ cm}^{-1}$ , both assigned to tyrosine/proline. Wavenumbers occurring only in basal cells were  $1657\text{ cm}^{-1}$  (amide I),  $1310\text{ cm}^{-1}$ ,  $1090\text{ cm}^{-1}$  (symmetric phosphate stretching

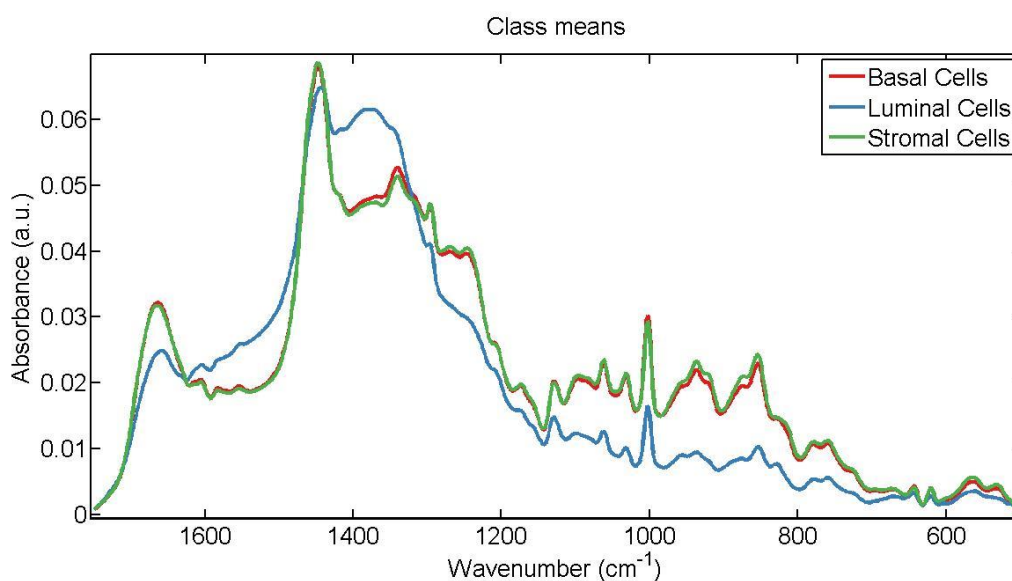
vibrations) and  $782\text{ cm}^{-1}$  (cytosine/uracil, DNA/RNA). Only  $1295\text{ cm}^{-1}$  was found to occur only in stromal cells whereas no wavenumber was found to be unique for luminal cells.



**Figure 3.233:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers among basal (red), luminal (blue) and stromal (green) cells in non-diseased tissue sections (H09-9102-6, H09-12890-11, H09-13558-A8). The red rectangles highlight wavenumbers shared in common between classes of cells.

*In diseased tissues*

The biochemical composition of all basal, luminal and stromal cells in diseased tissue samples (H09-9102-1, H09-12890-9, H09-13558-A5) taken from cancerous lesions of the endometrium was analysed. All classes of cells exhibited relatively the same shape in mean Raman spectra whereby evident variations in absorbance intensities were observed at  $\sim 1665\text{ cm}^{-1}$ ,  $\sim 1448\text{ cm}^{-1}$ ,  $\sim 1342\text{ cm}^{-1}$ ,  $\sim 1295\text{ cm}^{-1}$ ,  $\sim 1128\text{ cm}^{-1}$ ,  $\sim 1061\text{ cm}^{-1}$ ,  $\sim 1002\text{ cm}^{-1}$ ,  $\sim 936\text{ cm}^{-1}$ ,  $\sim 853\text{ cm}^{-1}$  and  $\sim 767\text{ cm}^{-1}$  (Fig. 3.234). Stromal and basal cells had similar peak intensities which were higher than intensities of luminal cells. Luminal cells had the lowest peak intensities except at  $\sim 1376\text{ cm}^{-1}$  whereby the peak was the highest. Classes of cells were significantly different from each other as indicated by the *P* values obtained from statistical analysis (Fig. 3.235).



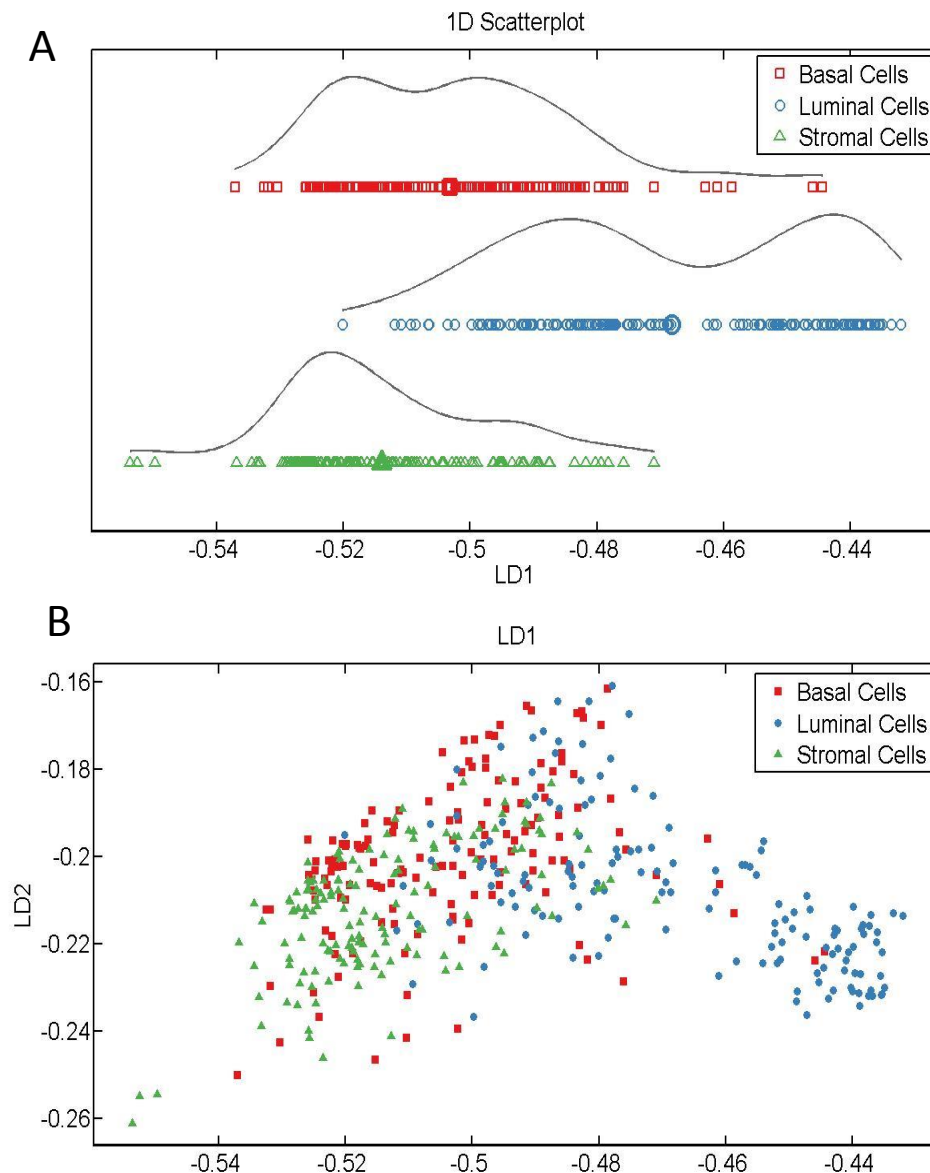
**Figure 3.234:** Mean Raman spectra for comparison of spectra acquired from all basal (red), luminal (blue) and stromal (green) cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) to analyse their biochemical composition.

Parameters	P value
LDA1 vs Basal Cells	$P < 0.01$
LDA1 vs Luminal Cells	$P < 0.001$
LDA1 vs Stromal Cells	$P < 0.001$
Basal Cells vs Luminal Cells	$P < 0.001$
Basal Cells vs Stromal Cells	$P < 0.001$
Luminal Cells vs Stromal Cells	$P < 0.001$

**Figure 3.235:** Obtained p-values by employment of One-way ANOVA test coupled with Tukey's multiple comparison test to compare spectra acquired from all basal, luminal and stromal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5).

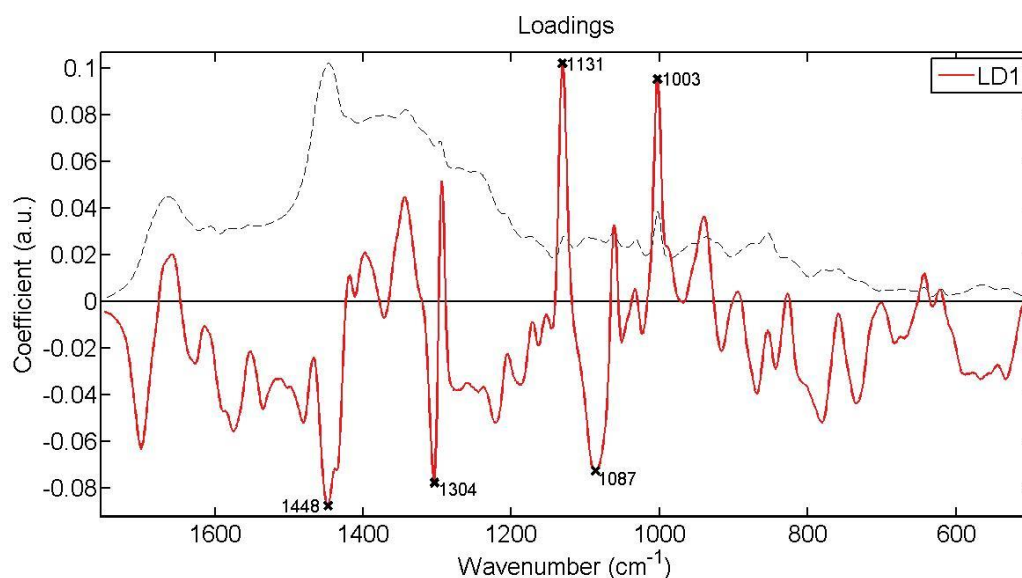


Segregation of classes after application of PCA-LDA was visualised in scores plots of 1D and 2D space (Fig. 3.236). All classes would overlap with each other mainly throughout their whole spectra. Almost all spectral points of basal and stromal cells overlapped with each other. Some spectral points of luminal cells displayed no overlap. Also spectral points of luminal cells had a relatively more spread arrangement signifying more intra-class variation.



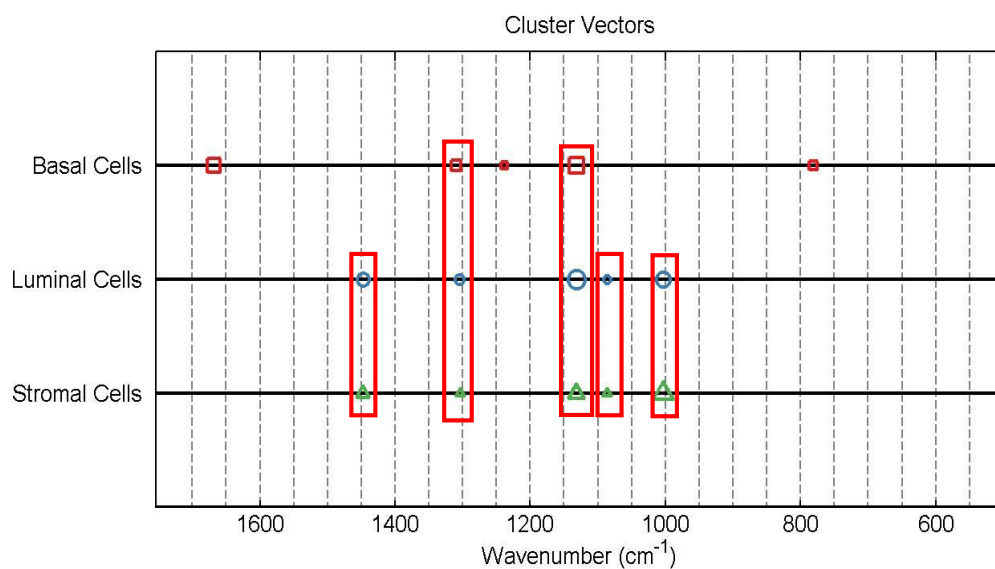
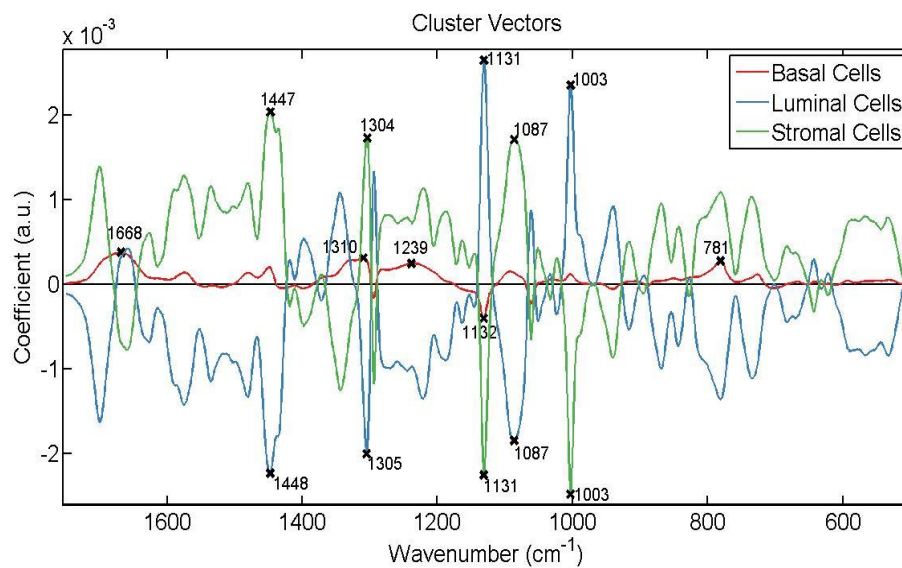
**Figure 3.236:** Scores plots of (A) 1D and (B) 2D space illustrating segregation of spectra acquired from basal (red), luminal (blue) and stromal (green) cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5) produced after application of PCA-LDA. In (B) LD1 would discriminate spectral points between the different classes of cells whereas as LD2 contributed to intra-class variation.

The five major discriminant wavenumbers identified in loading plots (Fig. 3.237) accounting for variations between classes of cells were 1448  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation, collagen), 1304  $\text{cm}^{-1}$  (lipids, adenine, cytosine), 1131  $\text{cm}^{-1}$  (lipids), 1087  $\text{cm}^{-1}$  (lipids) and 1003  $\text{cm}^{-1}$  (phenylalanine).



**Figure 3.237:** Loadings plots showing five major wavenumbers that discriminate basal, luminal and stromal cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red line is a pseudospectra and the dotted line is the actual pre-processed spectrum used as a reference spectrum.

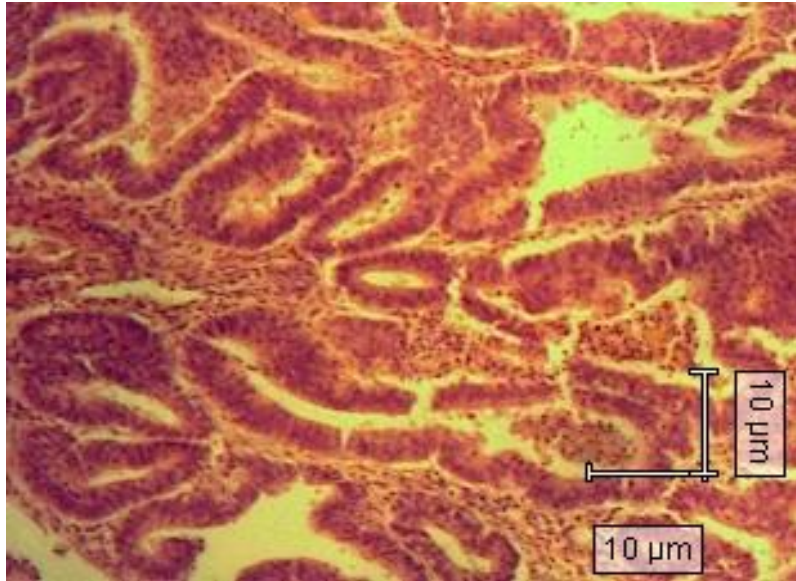
Cluster vector plots were analysed in an alternative way in an attempt to identify potential biomarkers for basal, luminal and stromal cells in diseased tissue samples (Fig. 3.238). Observations were based on expression levels of wavenumbers in the different classes of cells and which wavenumbers and/or spectral regions were commonly shared among classes and which wavenumbers occurred only in one class. The expression levels of shared wavenumbers were the same among classes. The wavenumbers commonly shared between all three classes of cells were 1310  $\text{cm}^{-1}$ , 1305  $\text{cm}^{-1}$  (adenine, cytosine), 1304  $\text{cm}^{-1}$  (lipids, adenine, cytosine), 1132  $\text{cm}^{-1}$  (lipids and proteins) and 1131  $\text{cm}^{-1}$  (lipids). Luminal and stromal cells were common for 1448  $\text{cm}^{-1}$  ( $\text{CH}_2$  deformation), 1447  $\text{cm}^{-1}$  (proteins and lipids), 1087  $\text{cm}^{-1}$  (lipids) and 1003  $\text{cm}^{-1}$  (phenylalanine). Wavenumbers occurring only in basal cells were 1668  $\text{cm}^{-1}$ , 1239  $\text{cm}^{-1}$  (amide III) and 781  $\text{cm}^{-1}$  (cytosine/uracil). No wavenumber was found to be unique neither for luminal cells nor for stromal cells.



**Figure 3.238:** Alternative presentation of cluster vectors plots, produced after application of PCA-LDA, showing occurrence and expression levels of wavenumbers among basal (red), luminal (blue) and stromal (green) cells in diseased tissue sections (H09-9102-1, H09-12890-9, H09-13558-A5). The red rectangles highlight wavenumbers shared in common between classes of cells.

### **3.4 Light Microscopy**

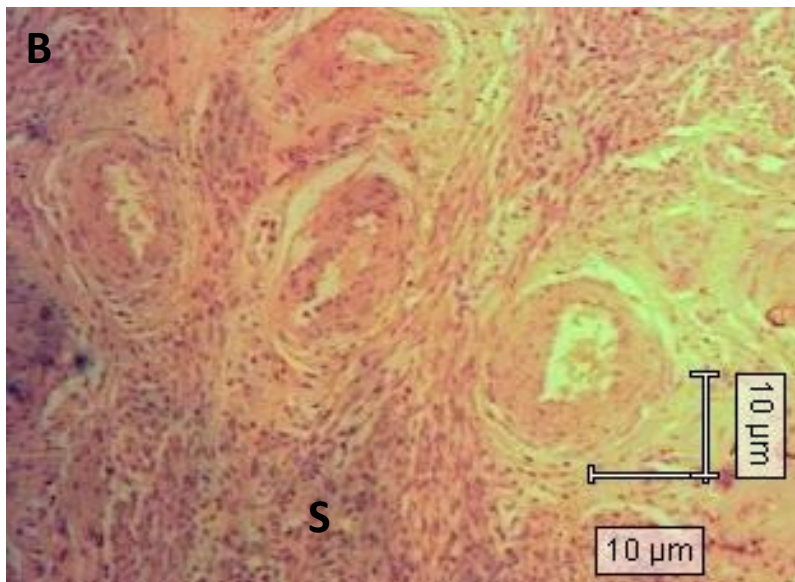
Parallel endometrial tissue sections of 4 $\mu$ m in thickness were cut from the paraffin embedded blocks of the samples used in biospectroscopy and were stained with hematoxylin and eosin (H&E). Glandular areas of interest were visualised through the x5 objective lens and images were taken using the Live video tool on Renishaw Wire 3.1 software of the Raman spectrometer. Representative images of the tissue samples are illustrated below (Fig. 3.239-3.246)

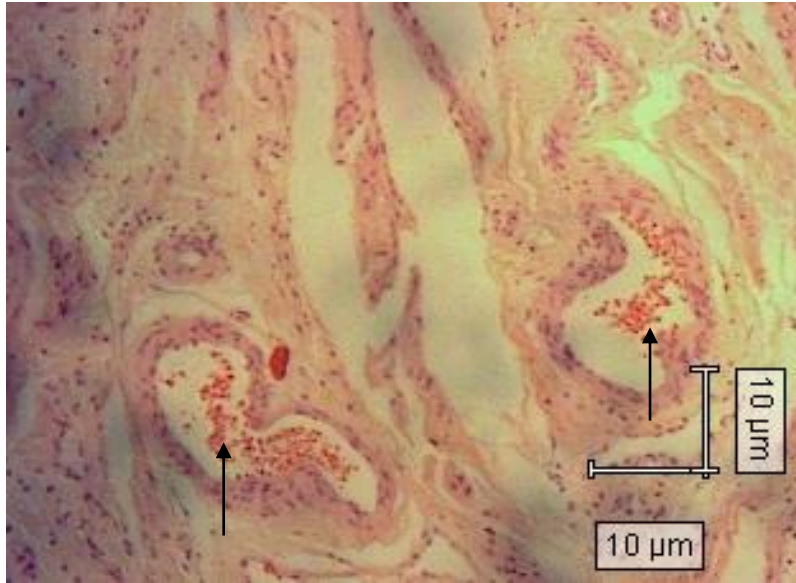


**Figure 3.239 (sample H09-12890-9):** Photomicrograph taken from cancerous lesion of post-menopausal endometrium. Tall epithelial cells make up the endometrial glands which are hyperplastic and irregular in shape. Endometrial hyperplasia resulted in little stroma presentation between glandular elements. Cellular atypia is not observed. No further information was given about the health status of the patient so from the observations the tissue exhibits characteristics of endometrial tumour or adenocarcinoma. Scale bars are illustrated on the micrograph.

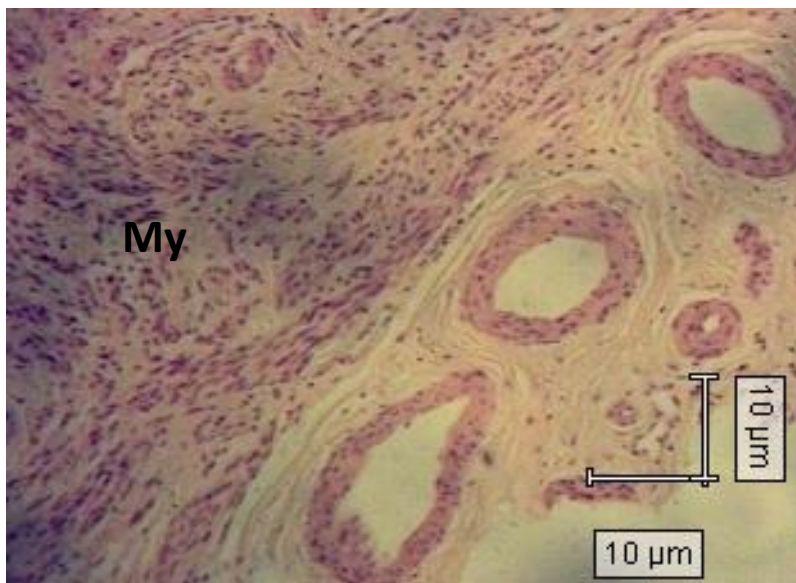


**Figure 3.240 (sample H09-12890-11):** Photomicrographs of tissue taken from post-menopausal patient with endometrial cancer but at site unaffected by tumour. **A:** Endometrial cancer causes bleeding resulting in appearance of red blood cells (black arrows) in endometrial glands. **B:** Cystic glands surrounded by dense stroma (**S**); typical feature of atrophic post-menopausal endometrium. Scale bars are illustrated on the micrograph.

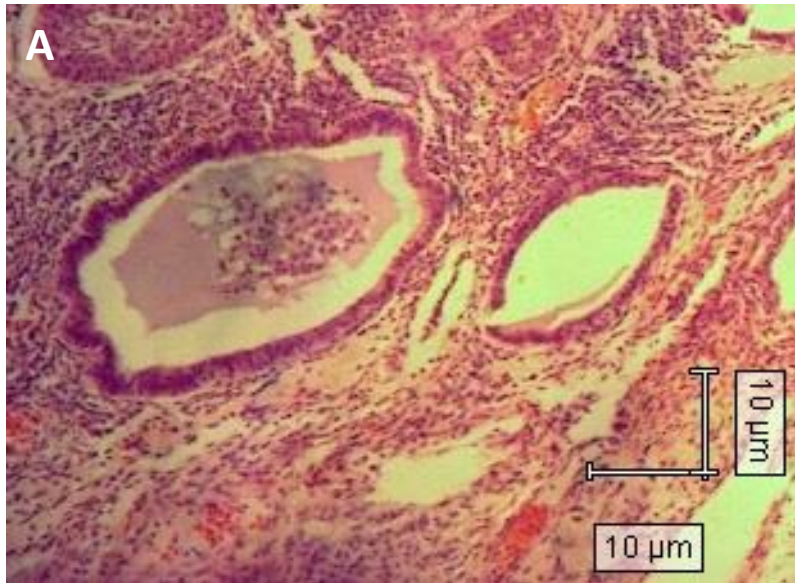




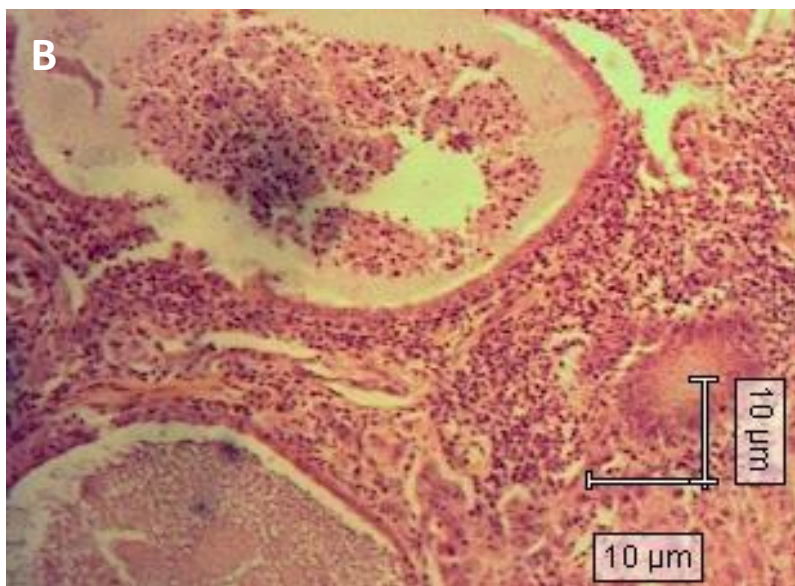
**Figure 3.241 (sample H09-9102-1):** Photomicrograph of tissue section at the site of tumour taken from post-menopausal patient with grade 1 endometrial cancer. Endometrial glands are large and irregular in shape. Papillary projections into the lumen of the glands can be noticed. Glandular lumens (black arrows) are filled with either red blood cells are found within glands because of bleeding or masses of cells due to loss of contact inhibition. Scale bars are illustrated on the micrograph.



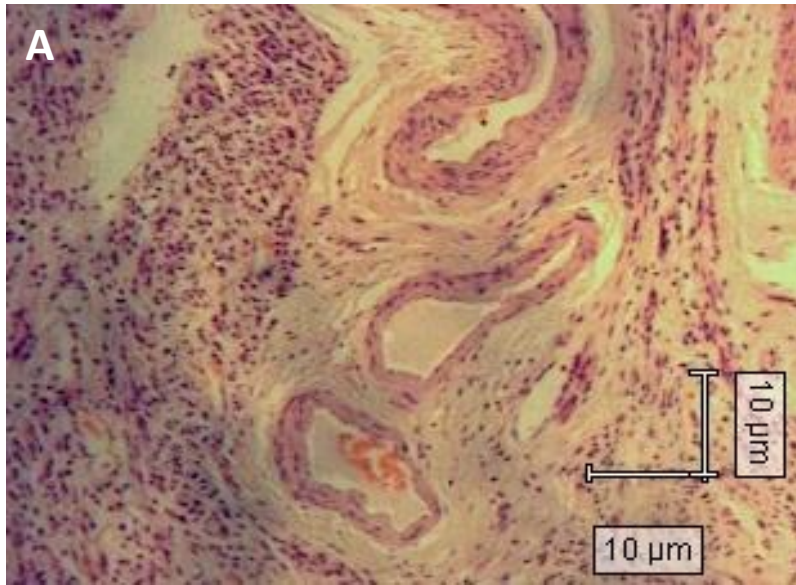
**Figure 3.242 (sample H09-9102-6):** Photomicrograph of tissue section from post-menopausal patient with grade 1 endometrial cancer but from an endometrial site away from tumour progression. The endometrium looks normal with no invasion of the myometrium (**My**). The glands have also a normal shape. Scale bars are illustrated on the micrograph.



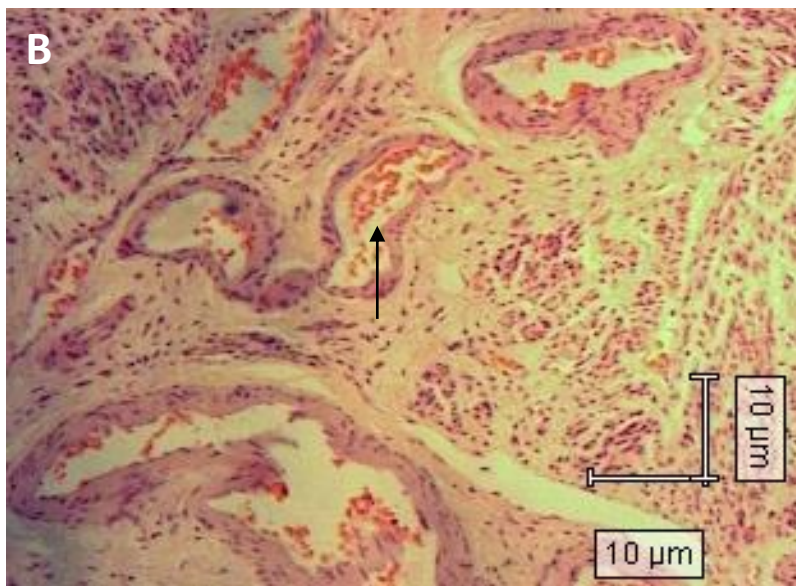
**Figure 3.243 (sample H09-13558-A5):** (A) and (B) are photomicrographs of post-menopausal endometrial tissue section of clear cell carcinoma. Masses of malignant cells fill the stroma surrounding the endometrial glands. Scale bars are illustrated on the micrographs.

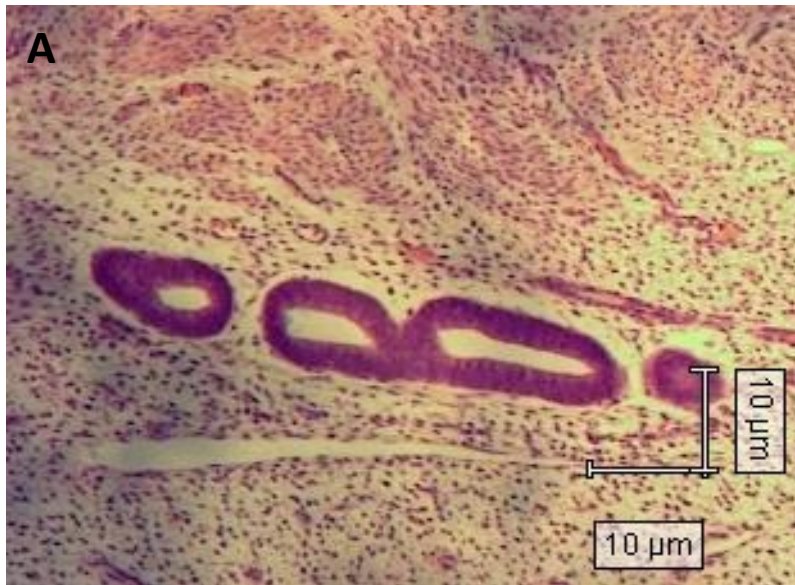




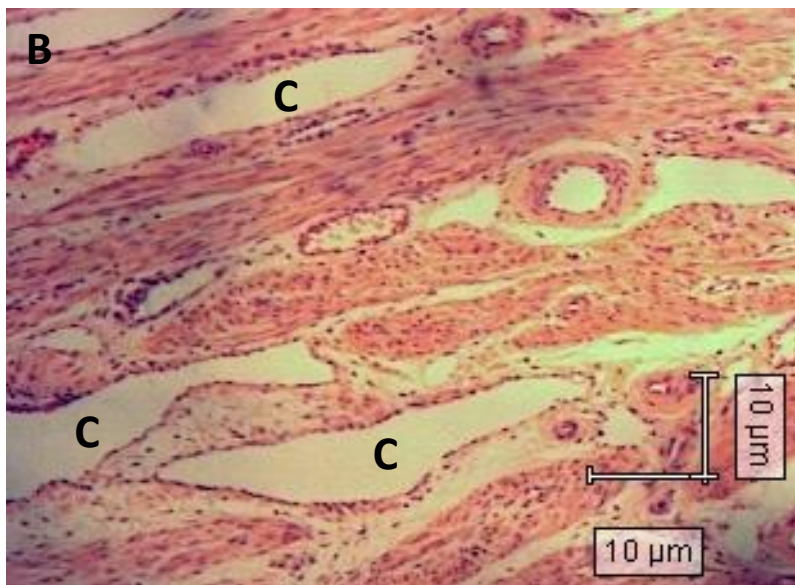


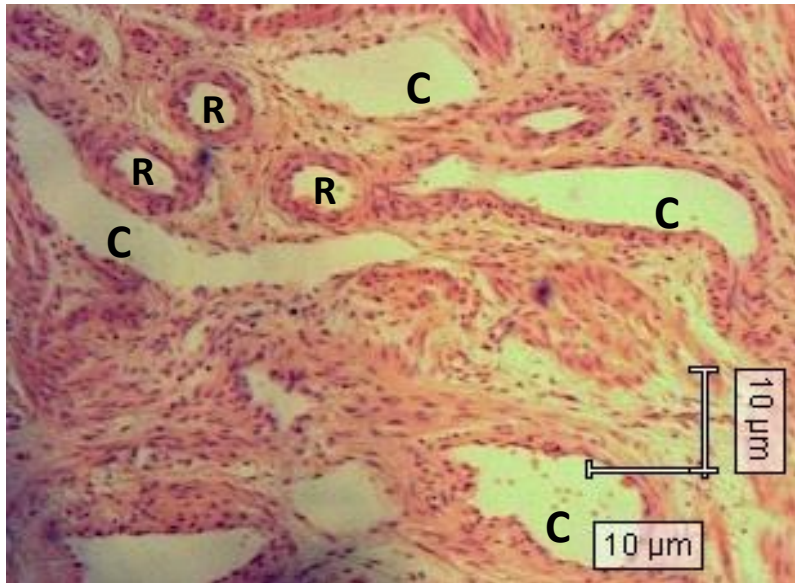
**Figure 3.244 (sample H09-13558-A8):** (A) and (B) are photomicrographs of tissue sections of non-diseased endometrium from patient with clear cell endometrial carcinoma. Red blood cells (arrow) were observed in glandular lumen because of bleeding. Scale bars are illustrated on the micrographs.





**Figure 3.245 (sample H09-11708-4):** Photomicrographs of normal, non-proliferating, post-menopausal endometrium. **A:** glands present cuboidal or low columnar epithelial cells. **B:** dilated glands due to formation of cystic spaces (C). Scale bars are illustrated on the micrographs.





**Figure 3.246 (sample H09-12292-1):** Photomicrograph of tissue section of normal post-menopausal endometrium taken from patient diagnosed with endometriosis. Some endometrial glands are dilated forming cystic spaces (C) whilst other glands have a regular shape (R). Scale bars are illustrated on the micrograph.

# CHAPTER 4

## Discussion

A variety of techniques were employed in this work including Scanning electron microscopy, stereology, FTIR spectroscopy and Raman spectroscopy to investigate the ultrastructure and molecular composition of post-menopausal endometrium with the aim of providing new insights on the endometrium with special attention to endometrial stem cells.

#### **4.1 Scanning Electron Microscopy**

##### 4.1.1 Scanning Electron Photomicrographs

Even though the examined endometrial biopsies were taken from non-diseased endometrial sites of postmenopausal patients, their endometrial morphology varied between samples. However the observations were similar to results obtained from other studies. Adams and Murphy (2001) assessed the epithelial architecture of biopsies taken from the anterior fundus of a 47 year old woman during three consecutive cycles under the influence of increasing doses of exogenous oestrogen and progesterone in an attempt for embryo transfer which eventually resulted in the birth of healthy twins. Evident changes in the plasma membrane of epithelial cells were observed throughout the cycles. Biopsies taken from the first two cycles showed characteristics of atrophic epithelial cells. The endometrial surface was inhomogeneous whilst certain areas were not covered by an epithelial lining. Epithelial cells exhibited various shapes from distended to flat and dome shaped. Borders of cells could not be identified whereas in other areas cell separations were pronounced. Ciliated cells were rare whilst the surface membrane of non-ciliated cells was covered by microvilli. However some cells had apical defects or wrinkled appearance (pinopods). Secretory droplets were also observed. Biopsies taken from the last cycle had a mature secretory epithelium whereby cells were small, uniform in shape, not separated by clefts, covered with microvilli, had clearly defined borders and covered the

surface area of the examined tissue. Fully developed pinopods were observed as well as an abundance in secretory droplets.

The majority of the images obtained in this project showed a hypotrophic epithelium which was similar to what was observed in the first two cycles. Administration of higher doses of oestrogen and progesterone resulted in a secretory epithelial response suggesting that even though postmenopausal endometrium is inactive it retains its capacity to differentiate in the presence of exogenous steroid hormones. It cannot be said with certainty if our observations were or were not influenced by administration of any therapy since no information was provided if patients were receiving a medical treatment.

In another study changes in endometrial surface during different phases of menstrual shedding and repair were examined in female patients with regular menstrual cycles (Garry et al., 2009). It was reported that tall columnar epithelial cells covered with microvilli composed the lining of endometrial glands in the late secretory phase of the menstrual cycle. Also ciliated cells were observed within the epithelial surface. Endometrial shedding resulted in exposure of glandular and blood vessel openings at the level of the basal layer and the tissue that was lost is eventually covered by a fibrinous matrix. During re-epithelisation, the new epithelial cells displayed irregularities in shape and size but they were smaller than previously observed and had a smooth surface lacking microvilli. Ciliated cells were rarely observed during the healing phase. In our results the presence of a fibrin mesh was very evident. Even though the cells presented variations in size and shape, they were covered by microvilli so they cannot be characterised as newly synthesised epithelial cells. Moreover samples were taken from postmenopausal patients whose endometrium is inactive and thus does not undergo cyclical changes.

Another study, whereby SEM was also employed, focused on structural changes in endometrial glands at the basal layer throughout menstrual shedding and regeneration, more specifically during, just before and after active bleeding phase (Garry et al., 2010). In this study endometrial biopsies were taken from females diagnosed with benign gynaecological conditions but had regular cycles.

Immunohistochemical labelling was also employed to assess expression of Ki-67 and CD68, markers of proliferation and macrophage activity respectively. It was reported that in the late secretory phase the glandular epithelium consisted of tall columnar epithelial cells which were covered by microvilli and interspersed with ciliated cells. Whereas during the early proliferative phase the newly synthesised epithelial cells were cuboidal, flat and without microvilli. In our findings though, features of atrophic endometrium were identified. During menstruation almost all of the functional layer is shed off and remnants of endometrial glands, with deficiencies in morphology, were observed at the surface of the basal layer. Glandular stumps were also identified in our results since the post-menopausal endometrium is mainly composed of the basal layer which retains its glandular elements even though they are inactive. Macrophages were also reported between epithelial cells of the glands or even within the lumen during the late secretory and early menstrual phase. It has been suggested that their numbers increase from the proliferative phase until the menstrual phase (Bonatz et al., 1992) and their function is to remove cells undergoing programmed cell death (Teo and Hughes, 2003) during the very early phase of menstruation (day1-2). Our findings have also reported the presence of macrophages suggesting that they are responsible in clearing away apoptotic cells of atrophic post-menopausal endometrium. In menstruating endometrium it was observed that apoptosis was followed by cellular renewal however the cessation in production of oestrogen and progesterone in post-menopausal endometrium, both of which induce differentiation and proliferation of endometrial cells, results in retention of the fibrinous stroma and continuation of cell loss. As mentioned earlier, a fibrin mesh was found in the majority of the examined endometrial surfaces.

Scanning electron microscopy was also employed to examine the effects of high doses of levonorgestrel on the uterine lining surface based on comparisons with control specimens (Ugocsai et al., 2002). Levonorgestrel is used as an emergency postcoital contraception which, like other contraceptive methods, prevents pregnancy by inhibiting fertilisation and providing unfavourable conditions for implantation (Nikas et al., 1995; Sphycher and Bigler, 2001). Its mechanism of action at molecular level is unclear but it has been previously reported to induce changes in the endometrium

(Van Look PFA, 1990; Raymond et al., 2000). However it does not cause any changes in production of oestrogen and progesterone since their serum levels in study participants and controls were not different for the menstrual phase at which they were compared. Changes in endometrial surface were compared during proliferative, luteal and secretory phase of the menstrual cycle. Epithelial cells in proliferative phase endometrium of study participants had a cobblestone-like appearance, were covered with numerous microvilli and were protruding from the lumen of endometrial glands. Pseudosecretory structures and only a few ciliated cells were observed. Whereas, in the control endometrial specimens ciliated cells were numerous around glandular openings and epithelial cells were large and richly covered by microvilli. Luteal phase endometrium of study participants was reported to have an even surface appearance, a few secretory droplets, epithelial cells of various sizes, less and shorter microvilli. Whereas control endometrium at the same phase had a cobblestone-appearance but surface was relatively even, epithelial cells had the same size and maintained their rich microvilli but cilia were shorter than in proliferative phase. Pinopodes were also observed.

Post-menopause results as a consequence of termination of ovulation and eventually steroid hormones are no longer produced so the endometrium can no longer support pregnancy. Our results were very similar to the observations made from the study participants as the supplementation of contraceptive pill induced changes to prevent pregnancy, even though it caused no alterations in circulating levels of progesterone and oestrogen.

Pinopode-like structures were observed in the endometrial specimens of our study, however no suggestions have been made so far about their function in post-menopausal endometrium. As mentioned earlier, no information was given whether patients were receiving any medication so it cannot be stated with certainty if the observed formation of pinopodes in examined biopsies was caused by a medical treatment.

Scanning electron microscopy in 1972 (Johanisson and Nilsson, 1972) reported for the first time the appearance of pinopodes from the apical surface of the luminal epithelium in human endometrium.



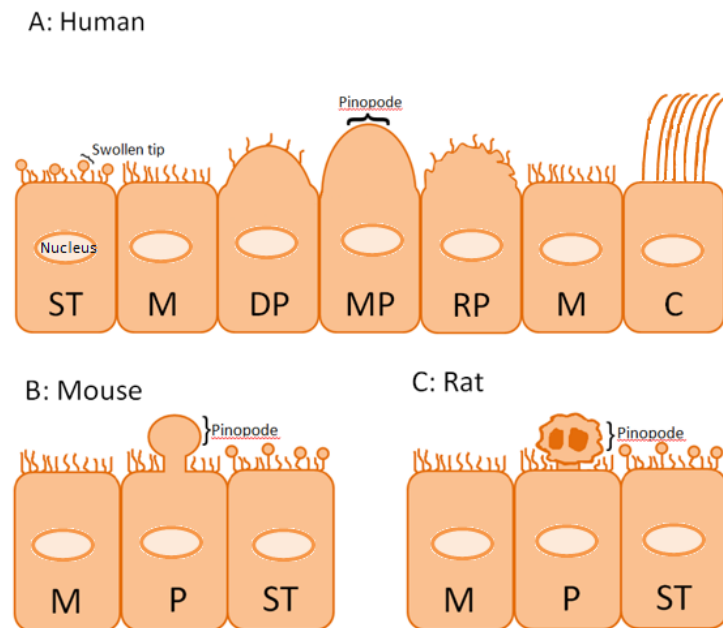
Earlier though, pinopodes had been found in rodents (Nilsson, 1958; Warren and Enders, 1964) arising mainly from the cell surface of epithelial stalks with a dependence on progesterone but inhibition by oestrogen (Martel et al., 1991) whilst in humans they seem to arise from non-ciliated epithelial cells (Murphy, 2000; Nikas et al., 1995) depending only on progesterone (Murphy, 1995). Usually one pinopode arises from the surface of the cells and covers most of it (Psychoyos and Mandon, 1971; Nikas et al., 2000; Quinn et al., 2007). The term uterodomes has been used to characterise pinopode-like structures in humans and rodents (Murphy, 2000).

To achieve a successful pregnancy during *in vitro* fertilisation, the embryo needs to be transferred after endometrial receptivity is established. Several studies in humans and rodents discovered the appearance of pinopodes on the endometrial surface during the “nidation window”, that is when the endometrium prepares to receive a blastocyst for implantation (Navot and Bergh, 1991). However their use as consistent markers for endometrial receptivity in mouse and human remains questionable whereas in the rat there is substantial evidence for their role. For pinopodes to be used as markers to determine endometrial receptivity, they need to persist for <48 hours (Quinn et al., 2007).

Pinopode morphology (size, shape, content) differs between rodents and humans. Pinopodes in rats have large vacuoles (Lopata et al., 2002), lack organelles (Nilsson, 1966; 1972) and extend on an actin stalk (Parr and Parr, 1974) above the microvilli (Warren and Enders, 1964; Enders and Nelson, 1973). In humans though, uterodomes extend from the whole cell surface (Murphy, 2000), have organelles but lack vacuoles (Friedrich, 1967; Bentin-Ley et al., 2000). Mitochondria and glycogen have been found in pinopodes of rodents, humans and other species (Guillmont et al., 1982; Parr and Parr, 1982; Dockery et al., 1997) suggesting that their formation requires a certain amount of cellular energy (Novonty et al., 1999) and/or they provide energy for growth to the implanted blastocyst.

In mice pinopodes are balloon-shaped (Quinn et al., 2007) whereas in rats they are wrinkled (Enders and Nelson, 1973). In humans their shape changes during the luteal phase of the menstrual cycle; uterodomes start to develop, reach full maturity and then start to regress (Fig. 4.1). Each developing

stage lasts about 24 hours (Nikas et al., 1995; 2000). Several studies debate on the lifespan of pinopodes in humans. However they all seem to agree that appearance of pinopodes lasts for about five days during the luteal phase; their numbers peak on the first two days and then start to disappear (Stavreus-Evers et al., 2001; Stavreus-Evers et al., 2002; Aghajanova et al., 2003).



**Figure 4.1:** Schematic drawing of pinopode formation in A: human B: mouse and C: rat endometrium in luminal epithelium. Pinopodes vary in morphology (size, shape). ST: cell whose microvilli have swollen tips M: cell with microvilli DP: developing pinopode MP: mature pinopode RP: regressing pinopode C: ciliated cell P: pinopode extends from the stalk at the apical surface of the cell (reproduced from Quinn and Casper, 2009).

The pinocytotic function of pinopodes has only been demonstrated in mice and rats by the uptake of ferritin, an electron dense tracer, from the uterine lumen into the vacuoles (Enders and Nelson, 1973; Parr and Parr, 1982). The possible role of uterodomes in humans is to enhance the adherence of the embryo on the endometrial surface by preventing cilia from moving the embryo (Stavreus-Evers, 2005).

Only a few markers though can correlate appearance of uterodomes and thus their function during the period of implantation window. Expression of integrins on the blastocyst and on epithelial cells of endometrial glands and lumen enhances the contact between these two surfaces for implantation to take place (Lessey et al., 1994). In normally menstruating women it was also observed that expression of uterodomes during the period of implantation occurred simultaneously with expression of alpha v beta 3 and alpha beta 1 integrins (Nardo et al., 2003).

Kabir-Salmani et al., 2005 demonstrated a secretory function of uterodomes in humans using endometrial biopsies throughout the whole period of luteal phase. Transmission electron microscopy (TEM), SEM, immunofluorescence and immunogold TEM revealed high expression of leukemia inhibitory factor (LIF) in uterodomes which was colocalised with biochemical markers of exocytosis. On the other hand, women with fertility problems have lower LIF levels in uterine secretion (Laird et al., 1997). In mouse endometrium, presence of LIF is necessary for embryo implantation (Stewart et al., 1992).

Heparin-binding epidermal growth factor (HB-EGF) expression was observed in glandular and luminal epithelium of human endometrium during the time of nidation (Leach et al., 1999; Yoo et al., 1997) as well as on the surface of uterodomes (Stavreus-Evres et al., 2002). These findings suggested a possible role of HB-EGF for facilitating the attachment of the blastocyst on luminal epithelium cells. Uterodomes may also provide protection against oxidative stress to the endometrium and pre-implanted embryo via expression of glutaredoxin (Stavreus-Evers et al., 2002).

Recently a novel function for uterodomes and how they influence endometrial receptivity was proposed. Subcellular localisation of MECA-79, an L-selectin ligand, was demonstrated in uterodomes during the midluteal phase of fertile females whereby the endometrium is the most receptive. It was concluded that this L-selectin ligand adhesion system facilitates the first attachment of an embryo on the luminal epithelium of the endometrium (Nejatbakhsh et al., 2012) in a similar way as it facilitates the rolling and adhesion of leukocytes on the endothelium (Barreiro et al., 2002; Barreiro et al., 2004; Ley and Kansas, 2004).

It has been long suggested and is the most accepted theory that proliferation of epithelial cells from the remnants of glandular stumps at the basal layer results in the synthesis of new epithelial lining to replace the surface that was lost during menstruation (Ferenczy, 1976; Ludwig et al., 1988; Salamonsen et al., 2007). Also it has been suggested that the regenerative potential of endometrium is due to the presence of populations of stem/progenitor cells at the base of endometrial glands and

near the vasculature which give rise to epithelial and stromal cell respectively (Gargett et al., 2007). However the work done by Garry et al., (2009) did not support this hypothesis, instead it was suggested that new epithelial cells develop on the fibrinous matrix which is formed on the endometrial surface during the healing process. These newly synthesised epithelial cells fuse with each other and eventually reach the glands to restore the glandular epithelium rather than arising from differentiation of already existing glandular epithelial cells.

Similarly the current theory on endometrial regeneration was again opposed in a further study by Garry et al. (2010). During the proliferative phase, stromal and glandular epithelial cells expressed the proliferation marker Ki-67 during cycles of mitotic cell division. However in the secretory phase, the marker was expressed mainly entirely by the stromal cells and minimal mitotic activity was observed in glandular epithelial cells. Also during the first days of menstruation, whereby most of the regeneration takes place, little proliferation was reported in the endometrium and none at the glandular areas thus suggesting that newly synthesised epithelial and stromal cells do not arise from remnants of endometrial glands.

Our findings are more in agreement with the long accepted theory on endometrial regeneration without though rejecting the opposing theories. From the results, especially at areas of endometrial crypts, it was observed that glandular openings were surrounded by fibrinous stroma whereas the margins and in some cases the internal of glands was lined by epithelial cells. As deep as it could be seen, it has also been observed that epithelial cells reached the depths of the crypts without any evidence though on their migration to the endometrial surface. It was rather unlikely to observe signs of migration since post-menopausal endometrium does not undergo cyclical changes like menstrual endometrium.

#### 4.1.2 Image Analysis

Scanning electron photomicrographs of endometrial crypts were used to analyse the morphology of cells surrounding the crypts and cells away from the crypts. The obtained results showed that the two groups of cells were significantly different for their perimeter, diameter max, diameter mean, sphericity and shape factor but they were not significantly different for their area, diameter min and shape factor. The mean of diameter max and diameter mean was higher for the cells around from the crypts thus it could be suggested that they are rather more elongated than the cells away from the crypts. This could also explain the higher mean perimeter for cells around from the crypts. Even though the two groups of cells were not significantly different for their areas, the higher mean in diameter max for cells around from the crypts comes in agreement with the higher mean in area for cells around from the crypts, taking into account that  $area = \left(\frac{\pi}{4}\right) \times diameter^2$ . Sphericity describes the 'roundness' of the particle being analysed. To this end, the two groups of cells were significantly different whilst the mean for cells away from the crypts was higher and thus suggesting that cells away from the crypts are rather more spherical which comes in agreement to the more elongated shape of cells around from the crypts. This can also be supported by the analysis of shape factor which provides information about the 'roundness' of the examined particle. Even though the two groups of cells were not significantly different in the concept of shape factor, the mean for cells away from the crypts was higher which supports the suggested more rounded morphology. The higher mean in aspect ratio for cells around from the crypts and obtained significant difference also supports the rather more elongated shape for cells around from the crypts and the more spherical shape for cells away from the crypts. The more elongated shape of cells around from the crypts can also be supported by observations from the photomicrographs whereby epithelial cells making the internal and external lining of endometrial glandular stumps were elongated and sometimes adjacent cells fused with each other.

## **4.2 FTIR spectroscopy**

Employment of biospectroscopy methods has become very popular over the recent years for examination of biological samples to investigate progression of cancer in several tissues as well as to study the differentiation process of stem cells. Interestingly biospectroscopy was shown to be an important tool in discriminating populations of stem cells and differentiated cells in regenerative tissues. To this end, biospectroscopy techniques were used in this project to segregate and characterise basal, luminal and stromal cells, based on their location at glandular areas, in an attempt to identify potential biomarkers for each cell type that would eventually support or provide more evidence about the hypothesised location of endometrial stem/progenitor cells. It has been suggested and accepted that epithelial stem cells are located at the base of glandular elements whilst stromal stem cells are found at vascular areas of the endometrium. In this concept, it was expected that luminal cells would display characteristics of a more differentiated state relative to basal and stromal cells whilst the two latter would not display significant differences between them. On a first approach, the biochemical composition of each cell type was analysed and was later compared with the other cells.

#### 4.2.1 Basal cells

During analysis of cluster vectors plots derived from application of PCA-LDA on spectra of basal cells in individual non-diseased tissue samples, the wavenumbers displaying a potential as biomarkers were 1456  $\text{cm}^{-1}$  (lipids and proteins) whose common occurrence among glandular areas was observed in all five samples, 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$  and 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings) whose common occurrence was observed in three out of the five samples. None of these wavenumbers though maintained constant expression levels among tissues. When all the five non-diseased samples were analysed for the spectra acquired from their basal cells the wavenumbers observed to have a potential as biomarkers were 1699  $\text{cm}^{-1}$  (guanine/thymine), 1539  $\text{cm}^{-1}$  (amide II) and 1506  $\text{cm}^{-1}$  since their common occurrence between tissues was observed in all six plots, either when having a class as reference origin or not. Common occurrence of 1504  $\text{cm}^{-1}$  (phenyl rings) with 1506  $\text{cm}^{-1}$  was also observed in this section. Common occurrence of 1732  $\text{cm}^{-1}$  (lipids) was observed in five plots so it could also be a potential biomarker. The wavenumber 1701  $\text{cm}^{-1}$  displayed similar occurrence with 1699  $\text{cm}^{-1}$  (guanine/thymine) in four plots so it could equally be considered to have a potential as biomarker. Only 1506  $\text{cm}^{-1}$  and 1701  $\text{cm}^{-1}$  maintained constant intermediate expression levels among tissues. When comparing the non-diseased tissues from patients with endometrial cancer only, the wavenumbers observed to have a potential as biomarkers for basal cells were 1539  $\text{cm}^{-1}$  (amide II), 1504  $\text{cm}^{-1}$  (phenyl rings) and 1456  $\text{cm}^{-1}$  (lipids and proteins) with the two latter wavenumbers maintaining constant expression of low and medium levels respectively.

Concerning the diseased-tissues, when glandular areas in individual samples were compared, the wavenumbers displaying a potential as biomarkers were 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA, purine base), 1456  $\text{cm}^{-1}$  (lipids and proteins), 1506  $\text{cm}^{-1}$  with 1504  $\text{cm}^{-1}$  (phenyl rings) and 1558  $\text{cm}^{-1}$  with 1556  $\text{cm}^{-1}$  since their common occurrence in glandular areas was observed in two out of the three tissues but

were found at different expression levels among tissues. When the three diseased tissues were compared for the spectra acquired from their basal cells, the wavenumbers observed to have the most potential as biomarkers were  $1771\text{ cm}^{-1}$  and  $1556\text{ cm}^{-1}$  since they were common among classes of tissues either when having a class as reference or not, but their expression levels among tissues varied. Since common occurrence of  $1558\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$ ,  $1506\text{ cm}^{-1}$ ,  $1504\text{ cm}^{-1}$  (phenyl rings),  $1456\text{ cm}^{-1}$  (lipids and proteins) was observed in both, non-diseased and diseased tissue sections, they cannot be considered as valid potential biomarkers for basal cells. So up to this point the wavenumbers displaying a potential as biomarkers for basal cells in non-diseased tissues were  $1732\text{ cm}^{-1}$  (lipids),  $1701\text{ cm}^{-1}$ ,  $1699\text{ cm}^{-1}$  (guanine/thymine) and  $1539\text{ cm}^{-1}$  (amide II) and for basal cells in diseased tissues were  $1771\text{ cm}^{-1}$  and  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base).

#### 4.2.2 Luminal cells

Analysing the spectra of luminal cells between the glandular areas in individual tissues, the wavenumbers displaying a potential as biomarkers in non-diseased tissues were  $1504\text{ cm}^{-1}$  (phenyl rings),  $1506\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins) since they were commonly shared between areas of four out of the five tissues with variations though in expression levels. Analysing the spectra of luminal cells between all the five non-diseased tissues, the wavenumbers observed to have a potential as biomarkers were  $1732\text{ cm}^{-1}$  (lipids),  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1651\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins) since they were commonly shared between tissues when a tissue was used as a class reference or not. Similarly none of the wavenumbers maintained constant expression levels between tissues. When analysing only the three non-diseased tissues taken from patients with endometrial cancer, the wavenumbers displaying a potential as biomarkers were  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1558\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$ ,  $1506\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings) but only  $1558\text{ cm}^{-1}$  and  $1556\text{ cm}^{-1}$  maintained high expression levels among tissues.



For the diseased tissues, wavenumbers observed to have a potential as biomarkers were  $1456\text{ cm}^{-1}$  (lipids and proteins) since it was commonly shared between glandular areas of two out of the three tissues. Also when only diseased tissues were compared, they were found to be common for  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base) and  $1651$  and thus displaying a potential as biomarkers.

Common occurrence of  $1651\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ ,  $1506\text{ cm}^{-1}$ ,  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1456\text{ cm}^{-1}$  (lipids and proteins) was observed in both, non-diseased and diseased tissue sections, so they cannot be considered as valid potential biomarkers for luminal cells. The wavenumbers found to be unique for luminal cells in non-diseased tissues were  $1732\text{ cm}^{-1}$  (lipids) and  $1699\text{ cm}^{-1}$  (guanine/thymine) whilst for luminal cells in diseased tissues was only  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base).

#### 4.2.3 Stromal cells

Analysing the spectra of stromal cells between the glandular areas in individual tissues, the wavenumbers displaying the most potential as biomarkers in non-diseased tissues were  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base) with  $1715\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins) since they were commonly shared between glandular areas of three out of the five tissues but neither maintained constant expression levels. Analysing the spectra of stromal cells between all the five non-diseased tissues, the wavenumbers observed to have a potential as biomarkers were  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1666\text{ cm}^{-1}$  (amide I),  $1558\text{ cm}^{-1}$  with  $1556\text{ cm}^{-1}$ ,  $1506\text{ cm}^{-1}$  with  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1456\text{ cm}^{-1}$  (lipids and proteins) with  $1454\text{ cm}^{-1}$  (asymmetric methyl deformation) since they were commonly shared between tissues when a tissue was used as a class reference or not. None of the wavenumbers though maintained constant expression levels between tissues. When analysing only the three non-diseased tissues taken from patients with endometrial cancer, the wavenumbers displaying a potential as biomarkers were  $1747\text{ cm}^{-1}$ ,  $1695\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and

proteins). Expression levels of  $1695\text{ cm}^{-1}$  were high but for the other wavenumbers expression levels were either intermediate or low.

For the diseased tissues, wavenumbers observed to have a potential as biomarkers were  $1771\text{ cm}^{-1}$  and  $1558\text{ cm}^{-1}$  since they were commonly shared between glandular areas of two out of the three tissues but at different expression levels. Also when only diseased tissues were compared, they were found to be common for  $1771\text{ cm}^{-1}$ ,  $1715\text{ cm}^{-1}$ ,  $1682\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ , and  $1456\text{ cm}^{-1}$  (lipids and proteins) and thus displaying a potential as biomarkers.

Common occurrence of  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base),  $1715\text{ cm}^{-1}$ ,  $1558\text{ cm}^{-1}$ ,  $1556\text{ cm}^{-1}$ ,  $1456\text{ cm}^{-1}$  (lipids and proteins),  $1454\text{ cm}^{-1}$  (asymmetric methyl deformation) was observed in both, non-diseased and diseased tissue sections, so they cannot be considered as valid potential biomarkers for stromal cells. The wavenumbers found to be unique for stromal cells in non-diseased tissues were  $1747\text{ cm}^{-1}$ ,  $1699\text{ cm}^{-1}$  (guanine/thymine),  $1695\text{ cm}^{-1}$ ,  $1666\text{ cm}^{-1}$  (amide I),  $1506\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings) whilst wavenumbers found to be unique for stromal cells in diseased tissues were  $1771\text{ cm}^{-1}$  and  $1682\text{ cm}^{-1}$ .

From the previous sections,  $1699\text{ cm}^{-1}$  (guanine/thymine) showed a potential as biomarker for non-diseased basal, luminal and stromal cells whilst  $1732\text{ cm}^{-1}$  (lipids) was identified in non-diseased basal and luminal cells. Even though  $1506\text{ cm}^{-1}$  and  $1504\text{ cm}^{-1}$  (phenyl rings) were suggested to be unique for normal stromal cells, our observations cannot support this with certainty since both of these wavenumbers were found in spectra of basal and luminal cells in non-diseased and diseased tissue sections. At this point the wavenumbers found to be unique for normal basal cells were  $1539\text{ cm}^{-1}$  (amide II) and  $1701\text{ cm}^{-1}$ , for normal stromal cells  $1747\text{ cm}^{-1}$  and  $1666\text{ cm}^{-1}$  (amide I) whilst no wavenumbers was found to occur only in normal luminal cells. But it could be suggested that  $1699\text{ cm}^{-1}$  (guanine/thymine) is unique for non-diseased endometrial tissues.

Similarly,  $1771\text{ cm}^{-1}$  displayed a potential as biomarker for basal and stromal cells in the diseased tissues whilst  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA, purine base) was identified in diseased basal and stromal

cells. It could be suggested that these wavenumbers are unique for diseased endometrial tissues whilst  $1682\text{ cm}^{-1}$  remains the only unique wavenumber diseased stromal cells.

#### 4.2.4 Basal cells Vs Luminal cells Vs Stromal cells

According to the hypothesised location of endometrial stem cells, cells at the base of the glands differentiate into more functional epithelial cells making up the lumen of the glands. Similarly, differentiation of stromal stem cells located at perivascular areas produces more functional stromal cells. In the case of post-menopausal endometrium, like the examined samples, the endometrium is inactive so proliferation of cells does not take place but cell populations are not lost since the basal layer which is comprised of inactive endometrial glands is retained. Thus, even if the selected locations for spectral acquisition at the base of the glands and the surrounding tissue of the examined samples did not contain stem cells, it could still be implied that the basal cells are the transient amplifying cells whilst the luminal cells are the terminally differentiated cells.

From the scores plots derived after application of PCA-LDA, even though perfect segregation between classes of cells was not achieved, it was observed that basal cells overlapped with both luminal and stromal cells at relatively the same degree. Luminal and stromal cells overlapped with each other but they appeared to be more dissimilar with each other than they were with basal cells. Overlapping between basal and luminal cells could be explained by the fact that differentiated cells had not yet migrated from the base to the lumen of the gland whilst overlapping between basal and stromal cells could be explained by the fact that both populations contained rather undifferentiated cells. Population of stromal cells may contain more undifferentiated cells than differentiated cells or apoptotic cells which were inactive like the luminal cells which could explain the overlapping between basal and luminal cells with stromal cells.

Unfortunately the majority of identified discriminant wavenumbers in loading plots accounting for variations between cell populations in individual tissues had undefined assignments so to this extent not enough suggestions can be made about the different biochemical composition of the cells. However a consistency of spectral regions was observed throughout the samples. The discriminant wavenumbers identified in loading plots when comparing cell populations in non-diseased tissues and in diseased tissues were mainly assigned to lipids and proteins suggesting that cells were different in their function, plasma membrane or presence of any secretory droplets at the lumen of the glands could account for these variations and that cells were rather in a differentiated state.

In the cluster vectors plots when comparing the three different cell populations in individual non-diseased tissues the wavenumbers found to be unique for basal cells from all the tissues were 1541  $\text{cm}^{-1}$  (amide II), 901  $\text{cm}^{-1}$ , 1747  $\text{cm}^{-1}$ , 1715  $\text{cm}^{-1}$ , 1651  $\text{cm}^{-1}$ , 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA), 1682  $\text{cm}^{-1}$ , 1456  $\text{cm}^{-1}$  (lipids and proteins), and 1796  $\text{cm}^{-1}$  but were not consistent between samples. The wavenumber 1456  $\text{cm}^{-1}$  (lipids and proteins) was also identified in spectra of luminal and stromal cells so its potential as biomarker is not clear which is in agreement to our previous observations. Similarly the potential of 1717  $\text{cm}^{-1}$  (amide I, DNA/RNA) and 1715  $\text{cm}^{-1}$  as biomarkers can be argued since they were previously observed in spectra of basal and luminal cells in diseased tissues. Also the potential of 1682  $\text{cm}^{-1}$  is questionable since it was observed in cluster vectors plots of stromal cells in diseased tissues and 1651  $\text{cm}^{-1}$  was observed in luminal cells in both diseased and non-diseased tissues.

The wavenumbers found to be unique for luminal cells in cluster vectors plots of individual non-diseased tissues were 1771  $\text{cm}^{-1}$ , 1472  $\text{cm}^{-1}$ , 1732  $\text{cm}^{-1}$  (lipids) and 1516  $\text{cm}^{-1}$  (amide II). Based on previous observations though the potential of 1771  $\text{cm}^{-1}$  and 1732  $\text{cm}^{-1}$  (lipids) as biomarkers is questionable and unclear.

The wavenumbers found to be unique for stromal cells in individual non-diseased tissues were 1786  $\text{cm}^{-1}$ , 1771  $\text{cm}^{-1}$ , 1747  $\text{cm}^{-1}$  and 1553  $\text{cm}^{-1}$  (amide II). The potential though of 1771  $\text{cm}^{-1}$  as biomarker

is unclear for the same reasons mentioned earlier as well as the potential of  $1747\text{ cm}^{-1}$  is questionable since it was found to be unique for basal cells in another tissue.

When comparing the three populations of cells for their spectra acquired from all the three non-diseased tissue sections from patients with endometrial cancer, only basal cells were found to be unique for  $1699\text{ cm}^{-1}$  (guanine/thymine) and  $1645\text{ cm}^{-1}$  but based on observations so far the potential of  $1699\text{ cm}^{-1}$  as biomarker is not clear.

In the cluster vectors plots when comparing the three different cell populations in individual diseased tissues the wavenumbers found to be unique for basal cells from all the tissues were  $1541\text{ cm}^{-1}$  (amide II),  $1504\text{ cm}^{-1}$  (phenyl rings) and  $1771\text{ cm}^{-1}$  but based to our previous observations the potential of the two latter wavenumbers as potential biomarkers is of questionable. The wavenumber  $1541\text{ cm}^{-1}$  (amide II) was also found to be unique for basal cells in non-diseased tissues so it could be suggested no be a potential biomarker for basal cells irrespective of the type of tissue. Luminal cells in diseased tissues were found to be unique for  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA) and  $1715\text{ cm}^{-1}$  but their potential as biomarkers based on our observations so far is not clear. Stromal cells in diseased tissues were found to be unique for  $1713\text{ cm}^{-1}$  (C=O thymine),  $1520\text{ cm}^{-1}$  (amide II),  $1418\text{ cm}^{-1}$  (deformation C-H),  $1651\text{ cm}^{-1}$  and  $1456\text{ cm}^{-1}$  (lipids and proteins) but the potential of the two latter as biomarkers is unclear.

When comparing the three populations of cells for their spectra acquired from all the three diseased tissue sections from patients with endometrial cancer, basal cells were found to be unique for  $1717\text{ cm}^{-1}$  (amide I, DNA/RNA),  $1556\text{ cm}^{-1}$  and  $1506\text{ cm}^{-1}$  but their potential as biomarkers is unclear. Luminal cells were found to be unique only for  $1732\text{ cm}^{-1}$  (lipids) but its potential as biomarker is unclear whilst stromal cells were unique only for  $1418\text{ cm}^{-1}$  (deformation C-H).

Table 4.1 summarises the observations as to which wavenumbers were identified to have the most potential as biomarkers for the three types of populations in non-diseased and diseased tissues. Unfortunately, the assignments for most of these spectral regions are undefined so further

interpretations about the biochemical composition of the cells is not feasible. Also as mentioned earlier the shape of the obtained spectra was different from the usual shape of an FTIR spectra of a biological samples so our observations and results may not be valid.

Tissue status	Cell population	Wavenumber/ $\text{cm}^{-1}$	Assignment
Non-diseased	Basal	901	
		1539	amide II
		1541	amide II
		1645	
		1796	
	Luminal	1516	amide II
		1472	
	Stromal	1786	
		1666	amide I
		1553	amide II
Diseased	Basal	1541	amide II
	Stromal	1713	C=O thymine
		1520	amide II
		1418	deformation C-H

**Table 4.1:** A summary of observations based on which wavenumbers displayed the most potential as biomarkers for basal, luminal and stromal cells in non-diseased and diseased- tissue sections.

#### 4.2.5 Applications of Infrared spectroscopy

FTIR biospectroscopy has been employed in other studies as a non-invasive method to monitor differentiation of stem cells. Differentiation of murine embryonic stem cells was studied whereby the absorption intensities in spectral region assigned to amide I and nucleic acids increased and decreased respectively with the progress of differentiation which suggested that mRNA translation was taking place and the  $\alpha$ -helix content of proteins increased (Ami et al., 2008). From our observations, spectral regions assigned to amide I and amide II were found in all the cell populations so it can be suggested that the cells were rather differentiated. High-intensity synchrotron radiation was used to distinguish stem cell, transit amplifying cells and terminal differentiated cells in bovine cornea whereby clear segregation was observed between the different cell types however biomarkers for corneal stem cells were not identified. In our study, perfect segregation between the interrogated cell populations was not observed but a degree of segregation was achieved and cell populations were found to be significantly different from each other. Synchrotron FTIR was also applied to characterise cells according to their position along the intestinal crypts of human tissues (Walsh et al., 2008). This study suggested that modifications in symmetric  $\text{PO}_2^-$  stretch observed at  $1080\text{ cm}^{-1}$  can be used as marker to identify the location of putative stem cells. Similar results were obtained in another study whereby employment of synchrotron radiation-based FTIR microscopy identified  $\text{PO}_2^-$  vibrational modes as potential biomarkers for stem cells in the inter-follicular skin epidermis (Patel et al., 2012). FTIR spectroscopy was used to study the differentiation of human mesenchymal stem cells into osteoblasts in cell cultures whereby a significant decrease in the mineral-to-matrix ratio in the extracellular matrix was observed with the progress of differentiation (Salasznyk et al., 2007). Stem cells and transit amplifying cells in human corneal tissues were interrogated by FTIR spectroscopy and it resulted in segregation of the cell populations whilst nucleic acids accounted for variations between the cell types (Bentley et al., 2007).

## 4.3 Raman spectroscopy

### 4.3.1 Basal cells

From the analysis of cluster vectors plots derived from interrogation of basal cells in glandular areas in individual non-diseased tissues the wavenumbers in the spectral region 1431-1437  $\text{cm}^{-1}$  displayed the most potential as biomarkers for basal cells since their common occurrence between glandular areas was observed in three out of the five tissues. However they did not display constant expression levels between tissues. When comparing all the five non-diseased tissues for the spectra acquired from their basal cells, the wavenumbers in the spectral region 1001-1003  $\text{cm}^{-1}$  displayed the most potential as biomarkers since they were commonly shared between tissues as illustrated in the cluster vectors plots either when a tissue was used as a reference class or not. Their expression levels were rather high to intermediate. When comparing spectra of basal cells in only the three non-diseased tissues from patients with endometrial cancer, the wavenumbers that displayed the most potential as biomarkers were 1131  $\text{cm}^{-1}$  (lipids), 1130  $\text{cm}^{-1}$  (lipids) and 1025  $\text{cm}^{-1}$  (glycogen) and wavenumbers in the spectral region 935-937  $\text{cm}^{-1}$  since they were commonly shared by the tissues according to the cluster vectors plots. Their expression levels though between tissue samples were not constant.

Analysing the spectra acquired from glandular areas in individual diseased tissue samples, the wavenumbers in the spectral region 1433-1440  $\text{cm}^{-1}$  and 1294-1303  $\text{cm}^{-1}$  displayed to have the most potential as biomarkers since their common occurrence between glandular areas was observed in all three and two tissues respectively. When diseased tissues were compared with each other for their



spectra, the wavenumbers in the spectral regions 1436-1444  $\text{cm}^{-1}$  and 1301-1304  $\text{cm}^{-1}$  displayed to have the most potential as biomarkers as illustrated from the cluster vectors plots and maintained rather high and intermediate expression levels respectively among tissues. However wavenumbers in the spectral regions 1433-1444  $\text{cm}^{-1}$  observed to have a potential as biomarkers in the non-diseased tissues so they can not be used to distinguish non-diseased basal cells from diseased cells but they may be used to discriminate basal cells from populations of luminal and stromal cells.

#### 4.3.2 Luminal cells

From the analysis of cluster vectors plots derived from interrogation of luminal cells in glandular areas in individual non-diseased tissues the wavenumbers in the spectral regions 1294-1296  $\text{cm}^{-1}$  and 1129-1133  $\text{cm}^{-1}$  displayed the most potential as biomarkers for luminal cells since their common occurrence between glandular areas was observed in all five and three out of the five tissues respectively. Their expression levels though were not constant. Also wavenumbers in both of these regions displayed a potential as biomarkers for basal cells so to this point their potential can be argued. When comparing all the five non-diseased tissues for the spectra acquired from their luminal cells, the wavenumbers in the spectral regions 1294-1298  $\text{cm}^{-1}$ , 1128-1131  $\text{cm}^{-1}$ , 1405-1407  $\text{cm}^{-1}$  and 1002  $\text{cm}^{-1}$  (phenylalanine) displayed the most potential as biomarkers since they were commonly shared between tissues as illustrated in the cluster vectors plots either when a tissue was used as a reference class or not. Only the wavenumbers in the region 1405-1407  $\text{cm}^{-1}$  maintained constant low expression levels. But the potential of 1002  $\text{cm}^{-1}$  (phenylalanine) as well as the potential of 1294-1298  $\text{cm}^{-1}$  and 1128-1131  $\text{cm}^{-1}$  can be argued since they were identified in analysis of basal cells. When comparing spectra of luminal cells in only the three non-diseased tissues from patients with endometrial cancer, the wavenumbers that displayed the most potential as biomarkers were 1436  $\text{cm}^{-1}$  (lipids) and 1437  $\text{cm}^{-1}$  (proteins and lipids) and displayed high and intermediate expression levels but these wavenumbers were also observed in analysis of basal cells so their potential as biomarkers is questionable.

Analysing the spectra acquired from glandular areas in individual diseased tissue samples, the wavenumbers in the spectral region 1295-1299  $\text{cm}^{-1}$  displayed to have the most potential as biomarkers since their common occurrence between glandular areas was observed in two out of the three tissues at high and intermediate expression levels however their potential can be argued since they were observed in analysis of basal cells. When diseased tissues were compared with each other for their spectra, the wavenumber 1536  $\text{cm}^{-1}$  displayed to have the most potential as biomarker as illustrated from the cluster vectors plots and maintained rather high and intermediate expression levels among tissues.

#### 4.3.3 Stromal cells

From the analysis of cluster vectors plots derived from interrogation of stromal cells in glandular areas in individual non-diseased tissues the wavenumbers in the spectral regions 1131-1137  $\text{cm}^{-1}$  and 1432-1435  $\text{cm}^{-1}$  displayed the most potential as biomarkers for stromal cells since their common occurrence between glandular areas was observed in three out of the five tissues respectively. Their expression levels though were not constant. Also wavenumbers in both of these regions were identified in analysis of both basal and luminal cells so to this point their potential can be argued. When comparing all the five non-diseased tissues for the spectra acquired from their stromal cells, the wavenumbers in the spectral regions 1451-1453  $\text{cm}^{-1}$ , 1294-1295  $\text{cm}^{-1}$ , 1104-1107  $\text{cm}^{-1}$  and 1001-1004  $\text{cm}^{-1}$  displayed the most potential as biomarkers since they were commonly shared between tissues as illustrated in the cluster vectors plots. Wavenumbers in the regions 1451-1453  $\text{cm}^{-1}$  and 1104-1107  $\text{cm}^{-1}$  maintained intermediate and low expression levels. The potential though of 1294-1295  $\text{cm}^{-1}$  and 1104-1107  $\text{cm}^{-1}$  can be argued since they were identified in analysis of basal and luminal cells. When comparing spectra of luminal cells in only the three non-diseased tissues from patients with endometrial cancer, the wavenumbers that displayed the most potential as biomarkers were 783  $\text{cm}^{-1}$  and 782  $\text{cm}^{-1}$  (cytosine/uracil, DNA/RNA) but their expression levels varied among tissues.

Analysing the spectra acquired from glandular areas in individual diseased tissue samples, the wavenumbers in the spectral region  $1295\text{-}1296\text{ cm}^{-1}$ ,  $1130\text{-}1132\text{ cm}^{-1}$  and  $1076\text{-}1081\text{ cm}^{-1}$  displayed to have the most potential as biomarkers since their common occurrence between glandular areas was observed in two out of the three tissues but at various expression levels. However the potential of the two first regions can be argued since they were observed in analysis of basal and luminal cells. When diseased tissues were compared with each other for their spectra, the wavenumbers in the region  $569\text{-}572\text{ cm}^{-1}$  displayed to have the most potential as biomarkers as illustrated from the cluster vectors plots and maintained rather intermediate and low expression levels among tissues.

#### 4.3.4 Basal cells Vs Luminal cells Vs Stromal cells

From the mean Raman spectra in all sections of data analysis it was observed that basal and stromal cells had similar peak intensities throughout most of the spectral region whereas absorbance intensities for luminal cells were lower thus suggesting that biochemical composition of basal and stromal cells is very similar. Also from the scores plots, even perfect segregation between the different populations of cells was not achieved it was observed that luminal and stromal cells were very dissimilar whilst basal cells were positioned rather in the middle. A possible explanation for this is the same as the one mentioned in the FTIR section.

According to the loading plots, generated when cell populations in individual tissues were compared, as well as when they were compared in all non-diseased and diseased tissues, the identified discriminant wavenumbers were not consistent however most of them were assigned to lipids, proteins and nucleic acids, similar to what was observed analysis of results derived from FTIR spectroscopy, which could be mainly attributed to epithelial cells thus suggesting that cell populations were different for their function and that cells were rather in a differentiated state. Wavenumbers

assigned to collagen were also identified and this could be explained by the composition of the fibrinous stroma.

From all the analysis in cluster vectors plots the wavenumbers found to be unique for basal cells in non-diseased tissues, excluding those whose potential as biomarkers is questionable were  $1685\text{ cm}^{-1}$  (amide I),  $919\text{ cm}^{-1}$ ,  $1608\text{ cm}^{-1}$  (cytosine),  $1257\text{ cm}^{-1}$  (amide III),  $725\text{ cm}^{-1}$  (DNA/RNA bases),  $1339\text{ cm}^{-1}$  (tryptophan/collagen/nucleic acids),  $869\text{ cm}^{-1}$  (proline),  $812\text{ cm}^{-1}$  (phosphodiester) and  $1397\text{ cm}^{-1}$ . For basal cells in diseased tissues only  $1682\text{ cm}^{-1}$  (cortisone) was found to be unique. As it can be observed these wavenumbers are not consistent for their assignments so the exact biochemical composition of basal cells cannot be interpreted.

Luminal cells in diseased tissues did not have any unique wavenumbers but the only wavenumber observed to be unique and thus have a potential as a biomarker was  $1407\text{ cm}^{-1}$  (proline, valine). This observation is in agreement to what was observed in the analysis of spectra of luminal cells in all the five non-diseased tissues whereby wavenumbers in the spectral region  $1405\text{-}1407\text{ cm}^{-1}$  displayed a potential as biomarkers. However not enough information is available to derive the biochemical composition of luminal cells.

Wavenumbers found to be unique for stromal cells in non-diseased tissues were  $1663\text{ cm}^{-1}$  (DNA),  $1420\text{ cm}^{-1}$  (proteins and lipids),  $642\text{ cm}^{-1}$  (tyrosine) and  $1063\text{ cm}^{-1}$  (C-C skeletal stretch). In diseased tissues, the wavenumbers found to be unique were  $871\text{ cm}^{-1}$ ,  $1643\text{ cm}^{-1}$ ,  $1307\text{ cm}^{-1}$  (lipids/collagen),  $1144\text{ cm}^{-1}$  and  $731\text{ cm}^{-1}$ . As it can be observed these wavenumbers are not consistent for their assignments so the exact biochemical composition of stromal cells cannot be interpreted.

Table 4.2 summarises the observations as to which wavenumbers were identified to have the most potential as biomarkers for the three types of populations in non-diseased and diseased tissues.

Tissue type	Cell population	Wavenumber/ cm <sup>-1</sup>	Assignment
Non-diseased	Basal	1685	amide I
		1608	cytosine
		1397	
		1339	tryptophan/collagen/nucleic acids
		1257	amide III
		919	
		869	proline
		812	phosphodiester
		725	DNA/RNA bases
	Luminal	1407	proline, valine
	Stromal	1663	DNA
		1420	proteins and lipids
		1063	C-C skeletal stretch
		642	tyrosine
Diseased	Basal	1682	cortisone
	Stromal	1643	
1307		lipids/collagen	
1144			
731			

#### 4.3.5 Applications of Raman spectroscopy

Raman spectroscopy was used to study the differentiation process of live murine embryonic stem cells whereby a decrease in both RNA and DNA peaks were observed with the progress of differentiation (Notingher et al., 2004). Similarly, a decrease in peak intensities of RNA and DNA was observed during

**Table 4.2:** A summary of observations based on which wavenumbers displayed the most potential as biomarkers for basal, luminal and stromal cells in non-diseased and diseased- tissue sections.

differentiation of live human embryonic stem cells into cardiomyocytes (Chan et al., 2009). Mineralisation process during differentiation of mesenchymal cells into osteoblasts was monitored by employment of Raman spectroscopy (Crane, 2009). In a similar study Raman spectroscopy was used to monitor production of minerals from mesenchymal cells under the influence of osteogenic agents (Azrad et al., 2006). Even though our results require further work and are unclear, Raman spectroscopy has been a promising and important tool in biomedical research since it has been able to give results of good quality in several studies. Raman spectroscopy was coupled with FTIR spectroscopy to differentiate pluripotent human embryonic stem cells from multipotent adult mesenchymal stem cells at different O<sub>2</sub> concentrations and results showed that population of cells were mainly different for their lipid composition (Pijanka et al., 2010).

Raman spectroscopy has been a useful tool in cancer research. It was able to identify biochemical markers that distinguish the transition zone, peripheral zone and central zone in normal human prostate tissues which could explain the susceptibility of these zones to development of malignancies and other pathological conditions (Patel and Martin, 2010). Near-infrared Raman spectroscopy was used as a diagnostic tool to study the biochemical changes in cervical tissues associated with dysplastic malignancies (Duraipandian et al., 2011). Raman spectroscopy and FTIR spectroscopy were used in the same study to examine the susceptibility of different populations (India Vs UK) in to development of prostate cancer (Patel et al., 2011).

#### 4.4 Conclusion

The dynamic regenerative capacity of the endometrium during the reproductive years of a female and its ability to respond to exogenous stimulus during post-menopause period is attributed to the existence of populations of stem/progenitor cells. Over the years biomedical research has focused on the identification of endometrial stem cells in order to provide potential for development of therapies to treat gynaecological diseases and infertility. Achievements so far provided evidence for the existence of stem cells throughout functional studies however markers of endometrial stem/progenitor cells are yet to be identified. Based on the structure and function of the human endometrium as well as on observations from studies in mouse endometrium a hypothetical location of endometrial stem cells has been proposed and accepted. It is believed that epithelial stem cells are found at the base of the endometrial glands in the basal layer from where they proliferate and migrate to reach the functional layer in order to replace the tissue which is sloughed off during menstruation, whilst stromal cells are located in perivascular regions of the endometrium and are responsible for regeneration of the lost stroma.

Post-menopausal endometrium has not been investigated a lot by means of Scanning electron microscopy. Rather of more interest were the changes occurring in endometrial morphology during the menstrual cycle so not much information was available from previous literature to support or correlate our findings. Similarly, biospectroscopic techniques were applied to characterise biochemical changes occurring in endometrial pathologies and to characterise populations of stem cells other tissues such as the skin, cornea and intestines whereby their location is known. Employment of biospectroscopic methods in our project was aiming to identify potential biomarkers for cells at the base and lumen of the glands and the surrounding stroma based on the hypothesised location of endometrial stem cells. To this end our attempt has rather failed however even though a perfect segregation between the different populations of cells was not achieved, a certain degree of separation was achieved. Luminal cells were found to be more dissimilar from basal and stromal cells,

and always significantly different, relative to the degree of difference observed between basal and stromal cells which signified that basal and stromal cells had a more similar biochemical composition even though they were mainly found to be significantly different for their spectra.

There is space for improvement in our approaches as well as potential for further work. A significant limitation in the section of FTIR spectroscopy was the fact that the shape of the obtained spectra was not that of a usual biological sample which could be attributed to experimental settings such as the mode for spectra acquisition. Prior to spectra collection, the paraffin embedded sections were de-waxed but it is quite possible that tissue sections still had residual paraffin so it would have been better if spectral regions accounting to residual paraffin were removed before data analysis so that the spectra would be more representative of the biochemical composition of the cells. In FTIR spectroscopy the band region assigned to residual paraffin in 1420-1480  $\text{cm}^{-1}$  and in Raman spectroscopy is at 1061  $\text{cm}^{-1}$ , which was indeed observed during data analysis. Also during data analysis we could have compared the spectra of two populations of cells individually i.e. basal cells Vs stromal cells, basal cells Vs luminal cells and luminal cells Vs stromal cells, instead of just comparing all three populations at the same time. This would have provided more information about differences and/or similarities in biochemical composition. Also the fact that samples were taken from post-menopausal endometrium which is inactive, atrophic and does not regenerate could have affected our observations.

Further work that can be done, is the application of other methods of computational analysis to assess the degree of segregation. Also it would be a good idea to assess if segregation was due to a particular spectral region, such as proteins or nucleic acids or lipids, in both FTIR and Raman spectroscopy. Other biospectroscopy methods can be employed in a similar concept like ATR-FTIR spectroscopy, or use spectral image mapping for both FTIR and Raman spectroscopy which can monitor distribution of chemical entities within the endometrial glands whereby absorbance intensities are illustrated by thermal colour changes from blue to green to yellow to red with blue and red signifying lowest and



highest intensity respectively. Transmission electron microscopy can also be used to examine endometrial tissues.

# References

**Aghajanova L, Stavreus-Evers A, Nikas Y, Hovatta O, Landgren BM.** (2003) Co-expression of pinopodes and leukemia inhibitory factor, as well as its receptor, in human endometrium. *Fertil Steril*, **1**:808-14.

**Ami D, Neri T, Natalello A, Mereghetti P, Doglia SM, Zanoni M, Zuccotti M, Garagna S, Redi CA.** (2008) Embryonic stem cell differentiation studied by FT-IR spectroscopy. *Biochim Biophys Acta*, **1783**:98-106.

**Azrad E, Zahor D, Vago R, Nevo Z, Doron R, Robinson R, Gheber LA, Rosenwaks S, Bar I.** (2006) Probing the effect of an extract of elk velvet antler powder on mesenchymal stem cells using Raman microspectroscopy: enhanced differentiation toward osteogenic fate. *J Raman Spectrosc*, **37**:480-6.

**Bajada S, Harrison PE, Ashton BA, Cassar Pullicino VN, Ashammakhi N, Richardson JB.** (2007) Successful treatment of refractory tibial non-union using calcium sulphate and bone marrow stromal cell implantation. *J Bone Joint Surg Br*, **89**:1382-86.

**Barreiro O, Vicente-Manzanares M, Urzainqui A, Yanez-Mo M, Sanchez-Madrid F.** (2004) Interactive protrusive structures during leukocyte adhesion and transendothelial migration. *Front Biosci*, **9**:1849-63.

**Barreiro O, Yanez-Mo M, Serrador JM, Montoya MC, Vicente-Manzanares M, Tejedor R, Furthmayr H, Sanchez-Madrid F.** (2002) Dynamic interaction of VCAM-1 and ICAM-1 with moesin and ezrin in a novel endothelial docking structure for adherent leukocytes. *J Cell Biol*, **157**:1233-45.

**Bassan P, Byrne HJ, Bonnier F, Lee J, Dumas P, Gardner P.** (2009) Resonant Mie scattering in infrared spectroscopy of biological materials- understanding the 'dispersion artefact'. *Analyst*, **134**:1586-93.

**Bassan P, Kohler A, Martens H, Lee J, Byrne HJ, Dumas P, Gazi E, Brown M, Clarke N, Gardner P.** (2010) Resonant Mie scattering (RMieS) correction of infrared spectra from highly scattering biological samples. *Analyst*, **135**:268-77.

**Bentin-Ley U, Horn T, Sjogren A, Sorensen S, Falck Larsen J, Hamberger L.** (2000) Ultrastructure of human blastocyst-endometrial interactions *in vitro*. *J Reprod Fertil*, **120**:337-50.

**Bentley AJ, Nakamura T, Hammiche A, Pollock HM, Martin FL, Kinoshita S, Fullwood NJ.** (2007) Characterization of human corneal stem cells by synchrotron infrared micro-spectroscopy. *Mol Vis*, **13**:237-42.

**Bentz EK, Kenning M, Schneeberger C, Kolbus A, Huber JC, Hefler LA, Tempfer CB.** (2010) OCT-4 expression in follicular and luteal phase endometrium: a pilot study. *Reprod Biol Endocrinol*, **8**:38.

**Borlogan CV, Kaneko Y, Maki M, Yu S-J, Ali M, Allickson JG, Sanberg CD, Kuzmin-Nichols N, Sanberg PR.** (2010) Menstrual blood cells display stem cell-like phenotypic markers and exert neuroprotection following transplantation in experimental stroke. *Stem Cell Dev*, **19**:439-52.

**Brenner RM, Slayden OD, Rodgers WH, Critchley HOD, Carroll R, Nie XJ, Mah K.** (2003) Immunocytochemical assessment of mitotic activity with an antibody to phosphorylated histone H3 in the macaque and human endometrium. *Hum Reprod*, **18**:1185-93.

**Caplan AI.** (2007) Adult mesenchymal stem cells for tissue engineering versus regenerative medicine. *J Cell Physiol*, **213**:341-7.

- Caplan AI.** (2009) Why are MSCs therapeutic? New data: new insight. *J Pathol*, **217**:318-24.
- Cervello I, Gil-Sanchis C, Mas A, Gado-Rosas F, Martinez-Conejero JA, Galan A, Martinez-Romero A, Martinez S, Navarro I, Ferro J, Horcajadas JA, Esteban FJ, O' Connor JE, Pellicer A, Simon C.** (2010) Human endometrial side population cells exhibit genotypic, phenotypic and functional features of somatic stem cells. *PLoS One*, **5**:e10964.
- Challen GA and Little MH.** (2006) A side order of stem cells: the SP phenotype. *Stem Cells*, **24**:3-12.
- Chan JW, Lieu DK, Huser T, Li RA.** (2009) Label-free separation of human embryonic stem cells and their cardiac derivatives using Raman spectroscopy. *Anal Chem*, **81**:1324-31.
- Chan RW, Schwab KE, Gargett CE.** (2004) Clonogenicity of human endometrial epithelial and stromal cells. *Biol Reprod*, **70**:1738–50.
- Chen YJ, Li HY, Chang YL, Yuan CC, Tai LK, Lu KH, Chang CM, Chiou SH.** (2010) Suppression of migratory/invasive ability and induction of apoptosis in adenomyosis-derived mesenchymal stem cells by cyclooxygenase-2 inhibitors. *Fertil Steril*, **94**:1972-9.
- Chiriboga L, Xie P, Yee H, Vigorita V, Zarou D, Zakim D, Diem M.** (1998) Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelial cells in the human cervix. *Biospectroscopy*, **4**:47-53.
- Chiriboga L, Xie P, Yee H, Vigorita V, Zarou D, Zakim D, Diem M.** (1998) Infrared spectroscopy of human tissue. II. A comparative study of spectra of biopsies of cervical squamous epithelium and of exfoliated cervical cells. *Biospectroscopy*, **4**:55-9.
- Choo-Smith LP, Edwards HM, Endtz HP, Kros JM, Heule F, Barr H, Robinson JS, Jr Bruining HA, Pupells GJ.** (2002) Medical applications of Raman spectroscopy: From proof of principle to clinical implementation. *Biopolymers (Biospectroscopy)*, **67**:1-9.
- Cooke PS, Buchanan DL, Young P, Setiawan T, Brody J, KorachKS, Taylor J, Lubahn DB, Cunha GR.** (1997) Stromal estrogen receptors mediate mitogenic effects of estradiol on uterine epithelium. *Proc Natl Acad Sci USA*, **94**:6535–40.
- Cong F and Varmus H.** (2004) Nuclear-cytoplasmic shuttling of Axin regulates subcellular localization of  $\beta$ -catenin. *Proc Natl Acad Sci USA*, **101**:2882–7.
- Crane NJ.** (2009) A non-destructive method for monitoring in vitro stem cell osteogenic differentiation with Raman spectroscopic mapping. *Spectroscopy*, **24**:9.
- Cui CH, Uyama T, Miyado K, Terai M, Kyo S, Kiyono T, Umezawa A.** (2007) Menstrual blood-derived cells confer human dystrophin expression in the murine model of Duchenne muscular dystrophy via cell fusion and myogenic transdifferentiation. *Mol Biol Cell*, **18**:1586–94.
- Dekker E and Fockens P.** (2005) Advances in colonic imaging: new endoscopic imaging methods. *Eur J Gastroenterol Hepatol*, **17**:803-8.
- Demos SG, Vogel AJ, Gandjbakhche AH.** (2006) Advances in optical spectroscopy and imaging of breast lesions. *J Mammary Gland Biol Neoplasia*, **11**:165-81.
- Di Cristofano A and Ellenson LH.** (2007) Endometrial carcinoma. *Ann Rev Pathol Mech Dis*, **2**:57–85.

- Dimitrov R, Timeva T, Kyurkchiev D, Stamenova M, Shterev A, Kostova P, Zlatkov V, Kehayov I, Kyurkchiev S.** (2008) Characterisation of clonogenic stromal cells isolated from human endometrium. *Reprod*, **135**:551–8.
- Dockery P, Ismail RMJ, Li TC, Warren MA, Coole ID.** (1997) The effect of single dose of mifepristone (RU486) on the fine structure of the human endometrium during the early luteal phase. *Hum Reprod*, **12**:1778-84.
- Dominici M, Le BK, Mueller I, Slaper-Cortenbach I, Marini F, Krause D, Deans R, Keating A, Prockop D, Horwitz E.** (2006) Minimal criteria for defining multipotent mesenchymal stromal cells. The international society for cellular therapy position statement. *Cytotherapy*, **8**:315–7.
- Downes A, Mouras R, Elfick A.** (2010) Optical spectroscopy for noninvasive monitoring of stem cell differentiation. *J Biomed Biotechnol*, **2010**:1-10.
- Du HL and Taylor HS.** (2009) Stem cells and reproduction. *Curr Opin Obstet Gynecol*, **22**:235–41.
- Duda RO, Hart PE, Stork DG.** (2001) Pattern Classification 2<sup>nd</sup> ed. (John Wiley & Sons, New York, USA).
- Dukor RK.** (2002) Vibrational spectroscopy in the detection of cancer. *Biomed Appl*, **5**:3335-59.
- Duraipandian S, Zheng W, Ng J, Low JJH, Ilancheran A, Huang Z.** (2011) In vivo diagnosis of cervical precancer using Raman spectroscopy and genetic algorithm techniques. *Analyst*, **136**:4328-36.
- Eckfeldt CE, Mendenhall EM, Verfaillie CM.** (2005) The molecular repertoire of the ‘almighty’ stem cell. *Nat Rev Mol Cell Biol*, **6**:726-37.
- Edward SL, Werkmeister JA, Rosamilia A, Ramshaw JA, White JF, Gargett CE.** (2013) Characterisation of clinical and newly fabricated meshes for pelvic organ prolapse repair. *J Mech Behav Biomed Mater*, **23**:53-61.
- EMBRYOLOGY (nd) *Module 6: Implantation* [online] Available at <http://www.embryology.ch/anglais/gnidation/role01.html> [Accessed on 3rd November 2014].
- Enders AC and Nelson DM.** (1973) Pinocytotic activity of the uterus of the rat. *Am J Anat*, **138**:277–99.
- Fearn T.** (2002) Discriminant analysis. In Handbook of Vibrational Spectroscopy Vol. 3 (eds. Chalmers, J. & Griffiths, P.R), 2086-93 (John Wiley & Sons, New York, USA).
- Figueira PGM, Abrao MS, Krikun G, Taylor H.** (2011) Stem cells in endometrium and their role in the pathogenesis of endometriosis. *Ann N Y Acad Sci*, **1221**:10-7.
- Friedrich ER.** (1967) Effects of contraceptive hormone preparations on the fine structure of the endometrium. *Obstet Gynecol*, **30**:201–19.
- Friel AM, Sergent PA, Patnaude C, Szotek PP, Oliva E, Scadden DT, Seiden MV, Foster R, Rueda BR.** (2008) Functional analyses of the cancer stem cell-like properties of human endometrial tumor initiating cells. *Cell Cycle*, **7**:242-9.
- Fuchs E, Tumber T, Guash G.** (2004) Socializing with the neighbors: stem cells and their niche. *Cell*, **116**:769-78.

- Fung MFK, Senterman MK, Mikhael NZ, Lacelle S, Wong PTT.** (1996) Pressure-tuning fourier transform infrared spectroscopic study of carcinogenesis in human endometrium. *Biospectroscopy*, **2**:155-65.
- Gage FH.** (2000) Mammalian neural stem cells. *Science*, **287**:1433-8.
- Gaide Chevronnay HP, Galant C, Lemoine P, Courtoy PJ, Marbaix E, Henriët P.** (2009) Spatiotemporal coupling of focal extracellular matrix degradation and reconstruction in the menstrual human endometrium. *Endocrinology*, **150**:5094-105.
- Gargett CE.** (2004) Stem cells in gynaecology. *Aust N Z J Obstet Gynaecol*, **44**:380–386.
- Gargett CE.** (2006) Identification and characterisation of human endometrial stem/progenitor cells. *Aust N Z J Obstet Gynaecol*, **46**:250-53.
- Gargett CE.** (2007) Uterine stem cells: What is the evidence? *Hum Reprod Update*, **13**:87–101.
- Gargett CE, Chan RW, Schwab KE.** (2007) Endometrial stem cells. *Curr Opin Obstet Gynecol*, **19**:377-83.
- Gargett CE and Healy DL.** (2011) Generating receptive endometrium in Asherman's syndrome. *J Hum Reprod Sci*, **4**:49-52.
- Gargett CE and Masuda H.** (2010) Adult stem cells in the endometrium. *Mol Hum Reprod*, **16**:818–34.
- Gargett CE, Nguyen Hong PT, Ye L.** (2012) Endometrial regeneration and endometrial stem/progenitor cells. *Rev Endocr Metab Disord*, **13**:235-51.
- Gargett CE and Rogers PAW.** (2001) Human endometrial angiogenesis. *Reproduction*, **121**:181–6.
- Gargett CE, Schwab KE, Zillwood RM, Nguyen HPT, Wu D.** (2009) Isolation and culture of epithelial progenitors and mesenchymal stem cells from human endometrium. *Biol Reprod*, **80**:1136-45.
- Gargett CE and Ye L.** (2012) Endometrial reconstruction from stem cells. *Fertil Steril*, **98**:11-20.
- German MJ, Pollock HM, Zhao B, Tobin MJ, Hammiche A, Bentley A, Cooper LJ, Martin FL, Fullwood NJ.** (2006) Characterisation of putative stem cell populations in the cornea using synchrotron infrared microspectroscopy. *Invest Ophthalmol Vis Science*, **47**:2417-21.
- Giudice LC and Kao LC.** (2004) Endometriosis. *Lancet*, **364**:1789–99.
- Gniadecka M, Wulf HC, Mortensen NN, Nielsen OF, Christensen DH.** (1997) Diagnosis of basal cell carcinoma by Raman spectroscopy. *J Raman Spectrosc*, **28**:125-9.
- Goodell MA, Rosenzweig M, Kim H, Marks DF, DeMaria MA, Paradis G, Grupp SA, Sieff CA, Mulligan RC, Johnson RP.** (1997) Dye efflux studies suggest that hematopoietic stem cells expressing low or undetectable levels of CD34 antigen exist in multiple species. *Nat Med*, **12**:1337-45.
- Gotte M, Wolf M, Steabler A, Buchweitz O, Kelsch R, Schuring AN, Kiessel L.** (2008) Increased expression of the adult stem cell marker Musashi-1 in endometriosis and endometrial carcinoma. *J Pathol*, **215**:317-29.
- Gruenwald P.** (1942) Origin of endometriosis from the mesenchyme of the coelomic walls. *Am J Obstet Gynecol*, **44**:474.

- Guillmont M and Guay P.** (1982) Ultrastructural features of the cell surfaces of uterine and trophoblast epithelia during embryo attachment of the cow. *Anat Rec*, **204**:315-22.
- Han X, Meng X, Yin Z, Rogers A, Zhing J, Rillema P, Jackson JA, Ichim TE, Minev B, Carrier E.** (2009) Inhibition of intracranial glioma growth by endometrial regenerative cells. *Cell Cycle*, **8**:606-10.
- Henriet P, Gaide Chevronnay HP, Marbaix E.** (2011) The endocrine and paracrine control of menstruation. *Mol Cell Endocrinol*, **358**:197-207.
- Hida N, Nishiyama N, Miyoshi S, Kira S, Segawa K, Uyama T, Mori T, Miyado K, Ikegami Y, Cui CH, Kiyono S, Shimizu T, Okano T, Sakamoto M, Ogawa S, Umezawa A.** (2008) Novel cardiac precursor-like cells from human menstrual blood-derived mesenchymal cells. *Stem Cells*, **26**:1695-1704.
- Horwitz EM, Prockop DJ, Fitzpatrick LA, Koo WW, Gordon PL, Neel M, Sussman M, Orchard P, Marx JC, Pyeritz RE, Brenner MK.** (1999) Transplantability and therapeutic effects of bone marrow-derived mesenchymal cells in children with osteogenesis imperfecta. *Nat Med*, **5**:309-313.
- Horwitz EM, Prockop DJ, Gordon PL, Koo WW, Fitzpatrick LA, Neel MD, McCarville ME, Orchard PJ, Pyeritz RE, Brenner MK.** (2001) Clinical responses to bone marrow transplantation in children with severe osteogenesis imperfecta. *Blood*, **97**:1227-31.
- Huang Z, McWilliams A, Lui M, McLean DI, Lam S, Zeng H.** (2003) Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *Int J Cancer*, **107**:1047-52.
- Hubbard SA, Friel AM, Kumar B, Zhang L, Rueda BR, Gargett CE.**(2009) Evidence for cancer stem cells in human endometrial carcinoma. *Cancer Res*, **69**:8241–8.
- Ichim Te, Alexandrescu DT, Solano F, Lara F, Campion RDN, Paris E, Woods EJ, Murphy MP, Dasanu CA, Patel AN.** (2010) Mesenchymal stem cells as anti-inflammatories: implications for treatments of Duchenne muscular dystrophy. *Cell Immunol*, **260**:75-82.
- Ichim TE, Solano F, Lara F, Rodriguez JP, Cristea O, Minev B, Ramos F, Woods EJ, Murphy MP, Alexandrescu DT, Patel AN, Riordan NH.** (2010) Combination stem cell therapy for heart failure. *Int Arch Med*, **3**:5.
- Ito M, Liu Y, Yang Z, Nguyen J, Liang F, Morris RJ, Cotsarelis G.** (2005) Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nat Med*, **11**:1351-54.
- Jabbour HN, Kelly RW, Fraser HM, Critchley HOD.** (2006) Endocrine regulation of menstruation. *Endocr Rev*, **27**:17–46.
- Jensen JR and Coddington CC 3rd.** (2010) Evolving spectrum: the pathogenesis of endometriosis. *Clin Obstet Gynecol*, **53**:379–88.
- Janssens S, Dubois C, Bogaert J, Theunissen K, Deroose C, Desmet W, Kalantzi M, Herbots L, Sinnaeve P, Dens J, Maertens J, Rademakers F, Dymarkowski S, Gheysens O, Van Cleemput J, Bormans G, Nuyts J, Belmans A, Mortelmans L, Boogaert M, Van de Werf F.** (2006) Autologous bone marrow-derived stem-cell transfer in patients with ST-segment elevation myocardial infarction: double-blind, randomised controlled trial. *Lancet*, **367**:113-121.
- Jordan CT.** (2006) Searching for leukemia stem cells - Not yet the end of the road? *Cancer Cell*, **10**:253–4.

- Kachi S, Hirano K, Takeuse Y, Miura M.** (2000) Unusual corneal deposit after the topical use of cyclosporine as eyedrops. *Am J Ophthalmol*, **130**:667-9.
- Kao AP, Wang KH, Chang CC, Lee JN, Long CY, Chen HS, Tsai CF, Hsieh TH, Tsai EM.** (2011) Comparative study of human eutopic and ectopic endometrial mesenchymal stem cells and the development of an in vivo endometriotic invasion model. *Fertil Steril*, **95**:1308-15.e1301.
- Kato K, Takao T, Kuboyama A, Tanaka Y, Ohgami T, Yamaguchi S, Adachi S, Yoneda T, Ueoka Y, Kato K, Hayashi S, Asanoma K, Wake N.** (2010) Endometrial cancer side-population cells show prominent migration and have a potential to differentiate into the mesenchymal cell lineage. *Am J Pathol*, **176**:381–92.
- Kato K, Yoshimoto M, Adachi S, Yamayoshi A, Arima T, Asanoma K, Kyo S, Nakahata T, Wake N.** (2007) Characterization of side-population cells in human normal endometrium. *Hum Reprod*, **22**:1214–23.
- Kelly JG, Maneesh NS, Stringfellow HF, Walsh MJ, Nicholson JM, Bahrami F, Ashton KM, Pitt MA, Martin-Hirsch PL, Martin FL.** (2009) Derivation of a subtype-specific biochemical signature of endometrial carcinoma using synchrotron-based Fourier-transform infrared microspectroscopy. *Cancer Lett*, **274**:208-17.
- Kelly JG, Martin-Hirsch PL, Martin FL.** (2009) Discrimination of base differences in oligonucleotides using mid-infrared spectroscopy and multivariate analysis. *Anal Chem*, **81**:5314-9.
- Kelly JG, Trevisan J, Scott AD, Carmichael PL, Pollock HM, Martin-Hirsch PL, Martin FL.** (2011) Biospectroscopy to metabolically profile biomolecular structure: a multistage approach linking computational analysis with biomarkers. *J Proteome Res*, **10**:1437-48.
- Kim JY, Tavare S, Shibata D.** (2005) Counting human somatic cell replications: methylation mirrors endometrial stem cell divisions. *Proc Natl Acad Sci USA*, **102**:17739-44.
- Kneipp K, Haka AS, Kneipp H, Badizadegan K, Yoshizawa N, Boone C, Shafer-Peltier KE, Motz JT, Dasari RR, Feld MS.** (2002) Surface-Enhanced Raman spectroscopy in single living cells using gold nanoparticles. *Appl Spectrosc*, **56**:150-4.
- Kurita T, Medina R, Schabel AB, Young P, Gama P, Parekh TV, Brody J, Cunha GR, Osteen KG, Bruner-Tran KL, Gold LI.** (2005) The activation function-1 domain of estrogen receptor alpha in uterine stromal cells is required for mouse but not human uterine epithelial response to estrogen. *Differentiation*, **73**:313–22.
- Laird SM, Tuckerman EM, Dalton CF, Dunphy BC, Li TC, Zhang X.** (1997) The production of leukaemia inhibitory factor by human endometrium: presence in uterine flushings and production by cells in culture. *Hum Reprod*, **12**:569-74.
- Leach RE, Khalifa R, Ramirez ND, Das SK, Wang J, Dey SK, Romero R, Armant DR.** (1999) Multiple roles for heparin-binding epidermal growth factor-like growth factor are suggested by its cell-specific expression during the human endometrial cycle and early placentation. *J Clin Endocrinol Metab*, **84**:3355-63.
- Lessey BA, Castelbaum AJ, Buck CA.** (1994) Further characterization of endometrial integrins during menstrual cycle and pregnancy. *Fertil Steril*, **62**:497-506.



- Ley K and Kansas GS.** (2004) Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. *Nat Rev Immunol*, **4**:325–35.
- Leyendecker G, Wildt L, Mall G.** (2009) The pathophysiology of endometriosis and adenomyosis: tissue injury and repair. *Arch Gynecol Obstet*, **280**:529-38.
- Lin-Vien D, Colthup NB, Fateley WG, Graselli JG.** (1991) The handbook of infrared and Raman characteristic frequencies of organic molecules, chapter 1. *Academic, New York*, pp 1-7.
- Llabjani V, Trevisan J, Jones KC, Shore RF, Martin FL.** (2011) Derivation by infrared spectroscopy with multivariate analysis of biomodal contaminant-induced dose effects in MCF-7 cells. *Environ Sci Technol*, **45**:6129-35.
- Lopata A, Bentin-Ley U, Enders A.** (2002) 'Pinopodes' and implantation. *Rev Endocr Metab Disord*, **3**:77–86.
- Ludwig H, Metzger H, Frauli M.** (1990) Endometrium: tissue remodelling and regeneration. In: D' Arcangues C, Fraser IS, Newton JR, Odland V, editors. Contraception and mechanisms of endometrial bleeding Cambridge University Press; pp 441-46.
- Ludwig H and Spornitz M.** (1991) Microarchitecture of the human endometrium by scanning electron microscopy: menstrual desquamation and remodelling. *Ann N Y Acad Sci*, **622**:28-46.
- Martel D, Monier MN, Roche D, Psychos A.** (1991) Hormonal control of pinopodeformation at the uterine luminal surface. *Hum Reprod*, **6**:597-603.
- Martin FL, Kelly JG, Llabjani V, Martin-Hirsch PL, Patel II, Trevisan J, Fullwood NJ, Wals MJ.** (2010) Distinguishing cell types or populations based on the computational analysis of their infrared spectra. *Nat Protoc*, **5**:1748-60.
- Maruyama T, Masuda H, Ono M, Kajitani T, Yoshimura Y.** (2010) Human uterine stem/progenitor cells: their possible role in uterine physiology and pathology. *Reproduction*, **140**:11–22.
- Massasa EE and Taylor HS.** (2012) Use of endometrial stem cells in regenerative medicine. *Regen Med*, **7**:133-5.
- Masuda H, Anwar SS, Buhning HJ, Rao JR, Gargett CE.** (2012) A novel marker of human endometrial mesenchymal stem-like cells. *Cell Transplant*, **21**:2201-14.
- Masuda H, Matsuzaki Y, Hiratsu E, Ono M, Nagashima T, Kajitani T, Arase T, Oda H, Uchida H, Asada H, Ito M, Yoshimura Y, Maruyama T, Okano H.** (2010) Stem cell-like properties of the endometrial side population: implication in endometrial regeneration. *PLoS One* **5**:e10387.
- Mather JP.** (2012) Cancer stem cells: *in vitro* models. *Stem Cells* **30**:95–9.
- Matthai C, Horvat R, Noe M, Nagele F, Radjabi A, van Trotsenburg M, Huber J, Kolbus A.** (2006) Oct-4 expression in human endometrium. *Mol Hum Reprod*, **12**:7–10.
- Maybin J and Critchley H.** (2009) Repair and regeneration of the human endometrium. *Expert Rev Obstet Gynecol*, **4**:283-98.
- Meng X, Ichim TE, Zhong J, Rogers A, Yin Z, Jackson J, Wang H, Ge W, Bogin V, Chan KW, Thebaud B, Riordan NH.** (2007) Endometrial regenerative cells: a novel stem cell population. *J Transl Med*, **5**:57.

- Moore Ka and Lemischka IR.** (2006) Stem cells and their niches. *Science*, **311**:1180-85.
- Mordechai S, Sahu RK, Hammody Z, Mark S, Kantarovich K, Guterman H, Podshyvalov J, Goldstein J, Argov S.** (2004) Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma. *J Microsc*, **215**:86-91.
- Morris RJ, Liu YP, Marles L, Yang ZX, Trempus C, Li SL, Lin JS, Sawicki JA, Cotsarelis G.** (2004) Capturing and profiling adult hair follicle stem cells. *Nat Biotechnol*, **22**:411-17.
- Morrison SJ, Shah NM, Anderson DL.** (1997) Regulatory mechanisms in stem cell biology. *Cell*, **88**:287-98.
- Movasaghi Z, Rehman S, Rehman IU.** (2007) Raman spectroscopy of biological tissues. *Appl Spectrosc Rev*, **42**:493-541.
- Movasaghi Z, Rehman S, Rehman IU.** (2008) Fourier Transform Infrared (FTIR) spectroscopy of biological tissues. *Appl Spectrosc Rev*, **43**:134-79.
- Murphy CR.** (1995) The cytoskeleton of uterine epithelial cells: a new player in uterine receptivity and the plasma membrane transformation. *Hum Reprod Update*, **1**:567-80.
- Murphy CR.** (2000) Understanding the apical surface markers of uterine receptivity; pinopodes or uterodomes? *Hum Reprod*, **15**:2451-4.
- Murphy MP, Wang H, Patel AN, Kambhampati S, Angle N, Chan K, Marleau AM, Pyszniak A, Carrier E, Ichim TE, Riordan NH.** (2008) Allogenic endometrial regenerative cells: an "Off the shelf solution" for critical limb ischemia? *J Transl Med*, **6**:45.
- Murry CE and Keller G.** (2008) Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell*, **132**:661-80.
- Musina RA, Belyavski AV, Tarusova OV, Solovyova EV, Sukhikh GT.** (2008) Endometrial mesenchymal stem cells isolated from the menstrual blood. *Bull Exp Biol Med*, **145**:539-43.
- Mutter GL, Ince TA, Baak JPA, Kust GA, Zhou XP, Eng C.** (2001) Molecular identification of latent precancers in histologically normal endometrium. *Cancer Res*, **61**:4311-14.
- Mutter GL, Lin MC, Fitzgerald JT, Kum JB, Eng C.** (2000) Changes in endometrial PTEN expression throughout the human menstrual cycle. *J Clin Endocrinol Metab*, **85**:2334-38.
- Nardo LG, Nikas G, Makrigiannakis A, Sinatra F, Nardo F.** (2003) Synchronous expression of pinopodes and alpha v beta 3 and alpha 4 beta 1 integrins in the endometrial surface epithelium of normally menstruating women during the implantation window. *J Reprod Med*, **48**:355-61.
- Navot D and Bergh P.** (1991) Preparation of the human endometrium for implantation. *Ann N Y Acad Sci*, **622**:212-9.
- Nejatbakhsh R, Kabir-Salmani M, Dimitriadis E, Hosseini A, Taheripanah R, Sadeghi Y, Akimoto Y, Iwashita M.** (2012) Subcellular localization of L-selectin ligand in endometrium implies a novel function for pinopodes in endometrial receptivity. *Reprod Biol Endocrinol*, **10**:46-54.
- Nguyen HPT, Sprung CN, Gargett CE.** (2012) Differential expression of Wnt signaling molecules between pre- and postmenopausal endometrial epithelial cells suggests a population of putative epithelial stem/progenitor cells reside in the basal layer. *Endocrinology*, **153**:2870-83.

- Nguyen LV, Vanner R, Dirks P, Eaves CJ.** (2012) Cancer stem cells: an evolving concept. *Nat Rev Cancer*, **12**:133–43.
- Nikas G, Drakakis P, Loutradis D, Mara-Skoufari C, Koumantakis E, Michalas S, Psychoyos A.** (1995) Uterine pinopodes as markers of the 'nidation window' in cycling women receiving exogenous oestradiol and progesterone. *Hum Reprod*, **10**:1208-13.
- Nikas G, Makrigiannakis A, Hovatta O, Jones HW Jr.** (2000) Surface morphology of the human endometrium. Basic and clinical aspects. *Ann N Y Acad Sci*, **900**:316–324.
- Nilsson O.** (1958) Ultrastructure of mouse uterine surface epithelium under different estrogenic influences. 1. Spayed animals and oestrous animals. *J Ultrastruct Res*, **1**:375-96.
- Nilsson O.** (1966) Structural differentiation of luminal membrane in rat uterus during normal and experimental implantations. *Z Anat Entwicklungsgesch*, **125**:152–9.
- Nilsson O.** (1972) Ultrastructure of the process of secretion in the rat uterine epithelium at preimplantation. *J Ultrastruct Res*, **40**:572–80.
- Notingher I, Bisson I, Polak JM, Hench LL.** (2004) In situ spectroscopic study of nucleic acids in differentiating embryonic stem cells. *Vib Spectrosc*, **35**:199-203.
- Novotny R, Malinsky J, Oborna I, Dostal J.** (1999) Ultrastructure of endometrial surface relief in normal menstrual cycle and after hormonal stimulation. *Acta Univ Palacki Olomuc Fac Med*, **142**:47-55.
- Okulicz WC and Scarrell R.** (1998) Estrogen receptor  $\alpha$  and progesterone receptor in the rhesus endometrium during the late secretory phase and menses. *Proc Soc Exp Biol Med*, **218**:316-21.
- Panayiotides I, Weyers S, Bosteels J, Van Herendael B.** (2009) Intrauterine adhesions (IUA): has there been progress in understanding and treatment over the last 20 years? *Gynecol Surg*, **6**:197-211.
- Padykula HA, Coles LG, Okulicz WC, Rapaport SI, McCracken JA, King Jr NW, Longcope C, Kaiserman-Abramof IR.** (1989) The basalis of the primate endometrium: a bifunctional germinal compartment. *Biol Reprod*, **40**:681-90.
- Park HC, Shim YS, Ha Y, Yoon SH, Park SR, Choi BH, Park HS.** (2005) Treatment of complete spinal cord injury patients by autologous bone marrow cell transplantation and administration of granulocyte-macrophage colony stimulating factor. *Tissue Eng*, **11**:913-922.
- Parker MF.** (2005) Emerging technology in cervical cancer screening: Spectroscopy. *Clin Obset Gynecol*, **48**:209-17.
- Parr MB and Parr EL.** (1974) Uterine luminal epithelium: protrusions mediate endocytosis, not apocrine secretion, in the rat. *Biol Reprod*, **11**:220–33.
- Parr MB and Parr EL.** (1982) Relationship of apical domes in the rabbit uterine epithelium during the peri-implantation period to endocytosis, apocrine secretion and fixation. *J Reprod Fertil*, **66**:739-44.
- Patel AN, Park E, Kuzman M, Benetti F, Silva FJ, Allickson JG.** (2008) Multipotent menstrual blood stem cells: isolation, characterisation, and differentiation. *Cell Transplant*, **17**:303-11.

**Patel II, Harrison WJ, Kerns JG, Filik J, Wehbe K, Carmicheal PL, Scott AD, Philpott MP, Frogley MD, cinque G, Maritn FL.** (2012) Isolating stem cells in the inter-follicular epidermis employing synchrotron radiation-based Fourier-transform infrared micorspectroscopy and focal plane array imaging. *Anal Bioanal Chem*, **404**:1745-58.

**Patel II and Martin FL.** (2010) Discrimination of zone specific spectral signatures in normal human prostate using Raman spectroscopy. *Analyst*, **135**:3060-9.

**Patel II, Trevisal J, Evans G, Llabjani V, Martin-Hirsch PL, Stringfellow HF, Martin FL.** (2011) High contrast images of uterine tissue derived using Raman microspectroscopy with the empty modelling approach of multivariate curve resolution-alternating least squares. *Analyst*, **136**:4950-9.

**Patel II, Trevisan J, Singh PB, Nicholson CM, Gopala Krishnan RK, Matanhelia SS, Martin FL.** (2011) Segregation of human prostate tissues classified high-risk (UK) versus low-risk (India) for adenocarcinoma using Fourier-transform infrared or Raman microspectroscopy coupled with discriminant analysis. *Anal Bioanal Chem*, **401**:969-82.

**Pellegrini G, Traverso CE, Franzi AT, Zingirian M, Cancedda R, De Luca M.** (1997) Long-term restoration of damaged corneal surfaces with autologous cultivated corneal epithelium. *Lancet*, **349**:990-3.

**Peron J, Jazedje T, Branadao W, Perin P, Maluf M, Evangelitsa L, Halpern S, Nisenbaum M, Czeresnia C, Zatz M.** (2012) Human endometrial-derived mesenchymal stem cells suppress inflammation in the central nervous system of EAE mice. *Stem Cell Rev*, **8**:940-52.

**Pijanka JK, Kumar D, Dale T, Yousef I, Parkes G, Untereiner V, Yang Y, Dumas P, Collins D, Manfait M, Sockalingum GD.** (2010) Vibrational spectroscopy differentiates between multipotent and pluripotent stem cells. *Analyst*, **135**:3126-32.

**Pfeffer FI, Dombkowski D, Sykes M, Scadden D, Yang YG.** (2002) Lineage-negative-side population (SP) cells with restricted hematopoietic capacity circulate in normal human adult blood: immunophenotypic and functional characterization. *Stem Cells*, **20**:417-27.

**Psychoyos A and Mandon P.** (1971) Study of the surface of the uterine epithelium by scanning electron microscope observations in the rat at the 4<sup>th</sup> and 5<sup>th</sup> day of pregnancy. *C R Acad Sci Hebd Seances Acad Sci D*, **272**:2723-5.

**Purandare NC, Patel II, Trevisan J, Bolger N, Kelehan R, von Bunau G, Martin-Hirsch PL, Prendiville WJ.** (2013) Biospectroscopy insights into the multi-stage process of cervical cancer development: probing for spectral biomarkers in cytology to distinguish grades. *Analyst*, **138**:3909-16.

**Quinn CE, Detmar J, Casper RF.** (2007) Pinopodes are present in Lif null and Hoca 10 null mice. *Fertil Steril*, **88**:1021-8.

**Ro S and Rannal B.** (2001) Methylation patterns and mathematical models reveal dynamics of stem cell turnover in the human colon. *Proc Natl Acad Sci USA*, **98**:10519-21.

**Ross MH, Romrell LJ, Kaye GI.** (1995) Histology, A text and atlas, Third edition, Baltimore, Williams & Wilkins.

**Roupa Z., Polikandrioti M., Sotiropoulou P., Faros E., Koulouri A., Wozniak G., Gourni M.** (2009) Causes of infertility in women at reproductive age. *HSJ*, **3**:80-7.

- Rutella S, Bonanno G, Procoli A, Mariotti A, Corallo M, Prisco MG, Eramo A, Napoletano C, Gallo D, Perillo A, Nuti M, Pierelli L, Testa U, Scambia G, Ferrandina G.** (2009) Cells with characteristics of cancer stem/progenitor cells express the CD133 antigen in human endometrial tumors. *Clin Cancer Res*, **15**:4299–311.
- Salamonsen LA.** (2003) Tissue injury and repair in the female human reproductive tract. *Reproduction*, **125**:301–11.
- Salasznyk RM, Klees RF, Williams WA, Boskey A, Plopper GE.** (2007) Focal adhesion kinase signalling pathways regulate the osteogenic differentiation of human mesenchymal stem cells. *Exp Cell Res*, **313**:22-37.
- Sampson JA.** (1927a) Metastatic or embolic endometriosis, due to the menstrual dissemination of endometrial tissue into the venous circulation. *Am J Pathol*, **3**:93–110.
- Sampson JA.** (1927b) Peritoneal endometriosis due to the menstrual dissemination of endometrial tissue into the peritoneal cavity. *Am J Obstet Gynecol*, **14**:469.
- Sancho E, Battle E, Clevers H.** (2004) Signalling pathways in intestinal development and cancer. *Annu Rev Cell Dev Biol*, **20**:695-723.
- Santamaria X, Massasa EE, Feng Y, Wolff E, Taylor HS.** (2011) Derivation of insulin producing cells from human endometrial stem cells and use in the treatment of murine diabetes. *Mol Ther*, **19**:2065-71.
- Sasson IE and Taylor HS.** (2008) Stem cells and the pathogenesis of endometriosis. *Ann N Y Acad Sci*, **1127**:106–115.
- Schenker JG and Margalioth EJ.** (1982) Intrauterine adhesions: an updated appraisal. *Fertil Steril*, **37**:593-610.
- Schuring AN, Schulte N, Kelsch R, Ropke A, Kiesel L, Gotte M.** (2011) Characterization of endometrial mesenchymal stem-like cells obtained by endometrial biopsy during routine diagnostics. *Fertil Steril*, **95**:423–6.
- Schwab KE, Chan RW, Gargett CE.** (2005) Putative Stem cell activity of human endometrial epithelial and stromal cells during the menstrual cycle. *Fertil Steril*, **84**:1124-30.
- Schwab KE and Gargett CE.** (2007) Co-expression of two perivascular cell markers isolates mesenchymal stem-like cells from human endometrium. *Hum Reprod*, **22**:2903–11.
- Sebag J, Nie S, Reise K, Yu NT.** (1993) Raman spectroscopy characterization of diabetes effects on human vitreous in diabetic retinopathy. *Proceedings of the SPIE*, **1877**:284-8.
- Shoae-Hassani A, Sharif S, Seifalian AM, Mortazavi-Tabatabaei SA, Rezaie S, Verdi J.** (2013) Endometrial stem cell differentiation into smooth muscle cell: a novel approach for bladder tissue engineering in women. *BJU Int*, **112**:854-63.
- Short MA, Lui H, McLean D, Zeng H, Alajlan A, Chen XK.** (2006) Changes in nuclei and peritumoral collagen within nodular basal cell carcinomas via confocal micro-Raman spectroscopy. *J Biomed Opt*, **11**:34004-13.
- Shostak S.** (2006) (Re)defining stem cells. *Bioessays*, **28**:301-8.

- Sigurdsson S, Philipson PA, Hansen LK, Laesen L, Gniadecka M, Wulf HC.** (2004) Detection of skin cancer by classification of Raman spectra. *IEEE Trans Biomed Eng*, **51**:1784-93.
- Singh R and Riess F.** (1998) Raman and the story of Nobel Prize. *Current Science*, **75**:965-71.
- Slayden OD and Brenner RM.** (2004) Hormonal regulation and localization of estrogen, progesterin and androgen receptors in the endometrium of nonhuman primates: effects of progesterone receptor antagonists. *Arch Histol Cytol*, **67**:393-409.
- Sobiesiak M, Sivasubramaniyan K, Hermann C, Tan C, Orgel M, Tremel S, Cerabona F, de Zwart P, Ochs U, Muller CA, Gargett CE, Kalbacher H, Buhning HJ.** (2010) The mesenchymal stem cell antigen MSCA-1 is identical to tissue non-specific alkaline phosphatase. *Stem Cells Dev*, **19**:669-77.
- Spencer TE, Hayashi K, Hu J, Carpenter KD.** (2005) Comparative developmental biology of the mammalian uterus. *Curr Top Dev Biol*, **68**:85-122.
- Spitzer TL, Rojas A, Zelenko Z, Aghajanova L, Erikson DW, Barragan F, Meyer M, Tamareisis JS, Hamilton AE, Irwin JC, Giudice LC.** (2012) Perivascular human endometrial mesenchymal stem cells express pathways relevant to self-renewal, lineage specification, and functional phenotype. *Biol Reprod*, **86**:58.
- Stavreus-Evers A.** (2005) Characteristics and possible function of pinopodes seen on the surface of the receptive human endometrium. *MEFSJ*, **10**:22-8. Opinion
- Stavreus-Evers A, Aghajanova L, Brismar H, Eriksson H, Landgren BM, Hovatta O.** (2002) Co-existence of heparin-binding epidermal growth factor-like growth factor and pinopodes in human endometrium at the time of implantation. *Mol Hum Reprod*, **8**:765-9.
- Stavreus-Evers A, Masironi B, Landgren BM, Holmgren A, Eriksson H, Sahlin L.** (2002) Immunohistochemical localization of glutaredoxin and thioredoxin in human endometrium: a possible association with pinopodes. *Mol Hum Reprod*, **8**:546-51.
- Stavreus-Evers A, Nikas G, Sahlin L, Eriksson H, Landgren BM.** (2001) Formation of pinopodes in human endometrium is associated with the concentrations of progesterone and progesterone receptors. *Fertil Steril*, **76**:782-91.
- Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Kontgen F, Abbondanzo SJ.** (1992) Blastocyst implantation depends on maternal expression of leukemia inhibitory factor. *Nature*, **359**:76-9.
- Stone N, Kendell C, Shepherd N, Crow P, Barr H.** (2002) Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *J Raman Spectrosc*, **33**:564-73.
- Stone N, Kendell C, Smith J, Crow P, Barr H.** (2004) Raman spectroscopy for the identification of epithelial cancers. *Faraday Discuss*, **126**:141-57.
- Stuart B.** (2004) Infrared spectroscopy: fundamentals and applications, chapter 7. *Wiley, Chichester*, pp 137-65.
- Tai MH, Chang CC, Kiupel M, Webster JD, Olson LK, Trosko JE.** (2005) Oct4 expression in adult human stem cells: evidence in support of the stem cell theory of carcinogenesis. *Carcinogenesis*, **26**:495-502.
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S.** (2007) Induction of pluripotent stem cells by defined factors. *Cell*, **131**:861-72.

- Tanaka M, Kyo S, Kanaya T, Yatabe N, Nakamura M, Maida Y, Okabe M, Inoue M.** (2003) Evidence of the monoclonal composition of human endometrial epithelial glands and mosaic pattern of clonal distribution. *Am J Pathol*, **163**:295-301.
- Taylor HS.** (2004) Endometrial cells derived from donor stem cells in bone marrow transplant recipients. *JAMA*, **292**:81-5.
- Tosh D and Slack JMW** (2002) How cells change their phenotype. *Nat Rev Mol Cell Biol*, **3**:187-94.
- Trounson A.** (2006) The production and directed differentiation of human embryonic stem cells. *Endocr Rev*, **27**:208-19.
- Tsuji S, Yoshimoto M, Takahashi K, Noda Y, Nakahata T, Heike T.** (2008) Side population cells contribute to the genesis of human endometrium. *Fertil Steril*, **90**:1528–37.
- Urich D, Edward SL, White JF, Supit T, Ramshaw JA, Lo C, Rosamilia A, Werkmeister JA, Gargett CE.** (2012) A preclinical evaluation of alternative synthetic biomaterials for fascial defect repair using a rat abdominal hernia model. *PLoS One*, **7**:e50044.
- Utzinger URS, Heintzelman DL, Mahadevan-Jansen A, Malpica A, Follen M, Richards-Kortum R.** (2001) Near-infrared Raman spectroscopy for *in vivo* detection of cervical precancers. *Appl Spectrosc*, **55**:955-9.
- Vacanti CA, Bonassar LJ, Vacanti MP, Shufflebarger J.** (2001) Replacement of an avulsed phalanx with tissue-engineered bone. *N Engl J Med*, **344**:1511-4.
- van der Horst PH, Wang Y, van der Zee M, Burger CW, Blok LJ.** (2011) Interaction between sex hormones and WNT/beta-catenin signal transduction in endometrial physiology and disease. *Mol Cell Endocrinol*, **358**:176-84.
- Verdi J, Tan A, Shoaie-Hassani A, Seifalian AM.** (2014) Endometrial stem cells in regenerative medicine. *Biol Eng*, **8**:20-9.
- Visvader JE and Lindeman GJ.** (2008) Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer*, **8**:755–68.
- von Recklinghausen F.** (1986) Adenomyomas and cystadenomas of the wall of the uterus and tube: their origin as remnants of the wolffian body. *Wien Klin Wochenschr*, **8**:530.
- Wagers AJ and Weissman IL.** (2004) Plasticity of adult stem cells. *Cell*, **116**:639-48.
- Wakitani S, Imoto K, Yamamoto T, Saito M, Murata N, Yoneda M.** (2002) Human autologous culture expanded bone marrow mesenchymal cell transplantation for repair of cartilage defects in osteoarthritic knees. *Osteoarthr Cartilage*, **10**:199-206.
- Walsh MJ, Fellous TG, Hammiche A, Lin WR, Fullwood NJ, Grude O, Bahrami F, Nicholson JM, Cotte M, Susini J, Pollock HM, Brittan M, Martin-Hirsch PL, Alison MR, Martin FL.** (2008) Fourier Transform Infrared Microspectroscopy identifies symmetric PO<sub>2</sub><sup>-</sup> modifications as marker of the putative stem cell region of human intestinal crypts. *Stem Cells*, **26**:108-18.
- Walsh MJ, German MJ, Singh M, Pollock HM, Hammiche A, Kyrgiou M, Stringfellow HF, Paraskevaidis E, Martin-Hirsch PL, Martin FL.** (2007) IR microspectroscopy: potential applications in cervical cancer screening. *Cancer Lett*, **246**:1-11.

- Warren RH and Enders AC.** (1964) An electron microscope study of the rat endometrium during delayed implantation. *Anat Rec*, **148**:177–95.
- Weissman IL.** (2002) Stem cells-scientific, medical, and political issues. *N Engl J Med*, **346**:1576-79.
- Wolf EF, Gao XB, Yao KV, Andrews ZB, Du H, Elsworth JD, Taylor HS.** (2010) Endometrial stem cell transplantation restores dopamine production in a Parkinson’s disease model. *J Cell Mol Med*, **15**:747-55.
- Wood BR, Hermelink A, Lasch P, Bambery KR, Webster GT, Khiavi MA, Cooke BM, Deed S, Naumann D, McNaughton D.** (2009) Resonance Raman microscopy in combination with partial dark-field microscopy lights up a new path in malaria diagnostics. *Analyst*, **134**:1119-25.
- Wood BR, Quinn MA, Burden FR, McNaughton D.** (1996) An investigation into FTIR spectroscopy as a bio-diagnostic tool for cervical cancer. *Biospectroscopy*, **2**:143-53.
- Ye L, Evans J, Gargett CE.** (2012) Lim1/LIM1 is expressed in developing and adult mouse and human endometrium. *Histochem Cell Biol*, **137**:527-36.
- Ye L, Mayberry R, Lo CY, Britt KL, Stanley EG, Elefanty AG, Gargett CE.** (2011) Generation of human female reproductive tract epithelium from human embryonic stem cells. *PLoS One*, **6**:e21136.
- Yoo HJ, Barlow DH, Mardon HJ.** (1997) Temporal and spatial regulation of expression of heparin-binding epidermal growth factor-like growth factor in the human endometrium: a possible role in blastocyst implantation. *Dev Genet*, **21**:102-8.
- Young B, Lowe JS, Stevens A, Health JW.** (2006) Wheater’s Functional Histology, A Text and Colour Atlas, Fifth edition, Philadelphia, Churchill Livingstone Elsevier.
- Yu D, Wong Y-M, Cheong Y, Xia E, Li T-C.** (2008) Asherman’s syndrome – one century later. *Fertil Steril*, **89**:759-79.
- Zeng H, McWilliams A, Lam S.** (2004) Optical spectroscopy and imaging for early lung cancer detection: A review. *Photodiagnosis Photodin Ther*, **1**:111-122.
- Zhong Z, Patel AN, Ichim TE, Riordan NH, Wang H, Min W-P, Woods EJ, Reid M, Mansilla E, Marin GH.** (2009) Feasibility investigation of allogeneic endometrial regenerative cells. *J Transl Med*, **7**:29-37.
- Zhou S, Schuetz JD, Bunting KD, Colapietro AM, Sampath J, Morris JJ, Lagutina I, Grosveld GC, Osawa M, Nakauchi H, Sorrentino BP.** (2001) The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat Med*, **7**:1028–34.