

Central Lancashire Online Knowledge (CLOK)

Title	Synchrotron- and focal plane array-based Fourier-transform infrared spectroscopy differentiates the Basalis and Functional epithelial endometrial regions and identifies putative stem cell regions of human endometrial glands
Type	Article
URL	https://clock.uclan.ac.uk/id/eprint/22528/
DOI	https://doi.org/10.1007/s00216-018-1111-x
Date	2018
Citation	Theophilou, Georgios, Medeiros-De-morais, Camilo De Ielis ORCID: 0000-0003-2573-787X, Halliwell, Diane E., Lima, Kássio M. G., Drury, Josephine, Martin-Hirsch, Pierre L., Stringfellow, Helen F., Hapangama, Dharani K. and Martin, Francis L (2018) Synchrotron- and focal plane array-based Fourier-transform infrared spectroscopy differentiates the Basalis and Functional epithelial endometrial regions and identifies putative stem cell regions of human endometrial glands. <i>Analytical and Bioanalytical Chemistry</i> , 410 (18). pp. 4541-4554. ISSN 1618-2642
Creators	Theophilou, Georgios, Medeiros-De-morais, Camilo De Ielis, Halliwell, Diane E., Lima, Kássio M. G., Drury, Josephine, Martin-Hirsch, Pierre L., Stringfellow, Helen F., Hapangama, Dharani K. and Martin, Francis L

It is advisable to refer to the publisher's version if you intend to cite from the work.
<https://doi.org/10.1007/s00216-018-1111-x>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLOK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>



Synchrotron- and focal plane array-based Fourier-transform infrared spectroscopy differentiates the basalis and functionalis epithelial endometrial regions and identifies putative stem cell regions of human endometrial glands

Georgios Theophilou¹ · Camilo L. M. Morais² · Diane E. Halliwell² · Kássio M. G. Lima³ · Josephine Drury⁴ · Pierre L. Martin-Hirsch⁵ · Helen F. Stringfellow⁵ · Dharani K. Hapangama⁴ · Francis L. Martin²

Received: 6 February 2018 / Revised: 18 April 2018 / Accepted: 25 April 2018
© The Author(s) 2018

Abstract

The cyclical process of regeneration of the endometrium suggests that it may contain a cell population that can provide daughter cells with high proliferative potential. These cell lineages are clinically significant as they may represent clonogenic cells that may also be involved in tumourigenesis as well as endometriotic lesion development. To determine whether the putative stem cell location within human uterine tissue can be derived using vibrational spectroscopy techniques, normal endometrial tissue was interrogated by two spectroscopic techniques. Paraffin-embedded uterine tissues containing endometrial glands were sectioned to 10- μ m-thick parallel tissue sections and were floated onto BaF₂ slides for synchrotron radiation-based Fourier-transform infrared (SR-FTIR) microspectroscopy and global focal plane array-based FTIR spectroscopy. Different spectral characteristics were identified depending on the location of the glands examined. The resulting infrared spectra were subjected to multivariate analysis to determine associated biophysical differences along the length of longitudinal and crosscut gland sections. Comparison of the epithelial cellular layer of transverse gland sections revealed alterations indicating the presence of putative transient-amplifying-like cells in the basalis and mitotic cells in the functionalis. SR-FTIR microspectroscopy of the base of the endometrial glands identified the location where putative stem cells may reside at the same time pointing towards $\nu_s\text{PO}_2^-$ in DNA and RNA, nucleic acids and amide I and II vibrations as major discriminating factors. This study supports the view that vibration spectroscopy technologies are a powerful adjunct to our understanding of the stem cell biology of endometrial tissue.

Keywords Biospectroscopy · Cell lineage · Chemometrics · Endometrium · FTIR spectroscopy · Stem cells

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00216-018-1111-x>) contains supplementary material, which is available to authorized users.

✉ Francis L. Martin
flmartin@uclan.ac.uk

¹ Department of Gynaecology, Leeds Teaching Hospitals NHS Foundation Trust, Leeds LS1 3EX, UK

² School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK

³ Biological Chemistry and Chemometrics, Institute of Chemistry, Federal University of Rio Grande do Norte, Natal 59072-970, Brazil

⁴ Department of Obstetrics and Gynaecology, Liverpool Women's NHS Foundation Trust, Liverpool L8 7SS, UK

⁵ Department of Obstetrics and Gynaecology, Royal Preston Hospital, Lancashire Teaching Hospitals NHS Foundation Trust, Fulwood, Preston PR2 9HT, UK

Abbreviations

DF	Discriminant function
DNA	Deoxyribonucleic acid
FN	False negative
FP	False positive
FPA	Focal plane array
F P A -	Focal plane array-based Fourier-transform infrared
FTIR	
FTIR	Fourier-transform infrared
GA	Genetic algorithm
G A -	Genetic algorithm-linear discriminant analysis
LDA	
H&E	Haematoxylin and eosin
IR	Infrared
KS	Kennard-Stone
LDA	Linear discriminant analysis

MCT	Mercury cadmium telluride
NA	Numerical aperture
PCA	Principal component analysis
P C A -	Principal component analysis–linear discriminant analysis
LDA	analysis
PCs	Principal components
RNA	Ribonucleic acid
SNR	Signal-to-noise ratio
SPA	Successive projections algorithm
S P A -	Successive projections algorithm–linear discriminant analysis
LDA	
S R -	Synchrotron radiation-based Fourier-transform infrared
FTIR	
TN	True negative
TP	True positive

Introduction

The endometrium is a highly regenerative tissue that undergoes monthly cycles of shedding and regeneration under the influence of oestrogen and progesterone [1, 2]. It comprises two regions, the functionalis and the basalis. The functionalis is the region that is shed with menstruation and is composed of glands lined by pseudostratified columnar epithelium during the proliferative phase of the menstrual cycle and columnar epithelium in the secretory phase. It is surrounded by vascularised stroma. The basalis comprises of the bases of the glands, surrounded by vascular endothelium and denser stroma [3]. The basalis layer extends into the myometrium. The endometrial/myometrial junction is the transition zone between the glandular epithelium and the stroma of the endometrium and the inner myometrium, which are not separated by a submucosal layer. Despite this, they are clearly distinguishable by light microscopy. Embryologically, these areas are of similar origin arising from the paramesonephric ducts, while the outer myometrium is thought to be of non-paramesonephric origin.

Adult stem cells are rare undifferentiated cells that function to maintain tissue homeostasis by generating replacement cells for tissues during routine cellular turnover or for the repair of tissues during injury [4]. Their defining properties are proliferation and self-renewal and differentiation into multiple lineages depending on the tissues they aim to regenerate [5]. They reside amongst niche cells and extracellular matrix which transmit signals that regulate their activity, at the same time guarding their genetic stability [4]. In the endometrium, there are three types of stem cells, namely epithelial, endothelial and mesenchymal stem cells [6]. It has been hypothesised that stem cells reside within the epithelial layer lining the glands of the basalis layer as well as the surrounding stroma and are responsible for the cyclical regeneration of the functionalis layer [1]. At the present time, there are no

conclusive markers to isolate epithelial progenitor cells in menstrual blood while stromal progenitor cells have been identified [7]. This suggests that the former may reside exclusively in the luminal epithelium of the basalis layer [8]. Their daughter cells are termed transient-amplifying cells and migrate to the functionalis layer resulting in its proliferation. Several techniques have been exploited to confirm the presence, location and activity of endometrial stem cells. These include cloning studies [9, 10], immunochemistry studies [11, 12] and regeneration studies [13]. Expression of immunochemical markers for such stem cells was demonstrated in human endometrium, but the exact location of most of these cells has not been pinpointed [14]. The SSEA1 expressing epithelial cells from basalis glands have shown some progenitor activity *in vitro* [12], but their abundance suggests that only a subpopulation of SSEA1 expressing epithelial cells from basalis glands has adult stem cell properties [12].

Stem cells may be implicated in uterine carcinogenesis. Cancer stem cells possess similar properties to stem cells in terms of differentiation and self-renewal potential. They differ from benign adult stem cells in that their growth potential is no longer controlled by signals from the surrounding niche cells; therefore, they proliferate uncontrollably and differentiate into unpredictable cellular lineages [15]. Cancer stem cells may be derived from resident adult stem cells through genetic or epigenetic changes [15].

Endometrial stem cells may also be implicated in the pathogenesis of endometriosis. Sampson's retrograde menstruation theory states that endometrial fragments that enter the peritoneal cavity through the fallopian tubes during menstruation implant into peritoneal surfaces and undergo similar cycles of proliferation and shedding as normal endometrium. It is hypothesised that these fragments may contain endometrial stem cells that are abnormally shed during menses and have the ability to implant into ectopic surfaces and proliferate forming endometriotic lesions [16, 17]. Although endometrial stem cells have been identified in menstrual blood, they have not been recognised in peritoneal fluid in higher quantities during menstruation [18]. This may be due to the lack of structural markers for these cells.

Infrared (IR) spectroscopy is a powerful technique to investigate biological tissues, since it can detect many important biochemical signatures including amide I ($\sim 1650\text{ cm}^{-1}$), amide II ($\sim 1550\text{ cm}^{-1}$), protein ($\sim 1425\text{ cm}^{-1}$), amide III ($\sim 1260\text{ cm}^{-1}$), asymmetric phosphate stretching vibrations ($\nu_{\text{as}}\text{PO}_2^-$; $\sim 1225\text{ cm}^{-1}$), carbohydrates ($\sim 1155\text{ cm}^{-1}$), symmetric phosphate stretching vibrations ($\nu_{\text{s}}\text{PO}_2^-$; $\sim 1080\text{ cm}^{-1}$) and protein phosphorylation ($\sim 970\text{ cm}^{-1}$) [19]. Successful applications of IR spectroscopy towards the analysis of tissue samples include cancer identification in breast, lung, colon and prostate tissues [20, 21]. Endometrial tissues have been also investigated using IR spectroscopy, where results for differentiating benign and malignant tissues were mainly

assigned to lipid and amide I/II regions [22]. Putative stem cells have been studied in human intestinal crypts using IR spectroscopy where $\nu_s\text{PO}_2^-$ were found to be the main biomarker for distinguishing different putative cell types [23].

Amongst different types of IR techniques, synchrotron radiation-based Fourier-transform infrared (SR-FTIR) and focal plane array (FPA) FTIR spectroscopy are excellent techniques for investigating tissue samples [21]. In SR-FTIR, a synchrotron source emits a collimated light beam more intense than that of a bench-top spectrometer. This provides an excellent signal-to-noise ratio (SNR) that is 1000 times greater to that of conventional IR sources and allows spatial resolutions as small as 10 μm [24, 25]. The absence of thermal noise and the order of intensity magnitude greater for synchrotron radiation source in infrared microspectroscopy increased the SNR performance in comparison to operation with a built-in global (thermal) source, therefore generating a spectra with higher resolution than regular FTIR [21]. On the other hand, FPA uses an IR focal plane array detector to generate hyperspectral imaging. The multiple detector elements of the FPA detector enable the concurrent acquisition of several spectra at each spatial point of the area of the sample under investigation. This allows the examination of larger areas at reasonable experimental time periods. This method results in the formation of a “hypercube” which contains information in two spatial dimensions: a pseudo-image and one spectral dimension, corresponding to the spectrum for each point (pixel) of that image [21].

Biospectroscopy is a powerful technology that can be used in conjunction to other biomarker-based isolation techniques to attempt the identification of endometrial stem cells amongst their progeny. The characterisation of endometrial stem cells with regard to their molecular, genetic and epigenetic make-up may assist in unravelling their role in many endometrial proliferative disorders such as endometriosis, adenomyosis and carcinogenesis. In this paper, SR-FTIR and FPA-FTIR microspectroscopy were used to investigate endometrial tissue samples. SR-FTIR microspectroscopy was used for identification of putative stem cells amongst the epithelial cells on the basalis portion of endometrial glands, while FPA-FTIR microspectroscopy was used to distinguish epithelial cells lining in the basalis portion from those lining in the functionalis portion of the endometrial glands.

Methods

Tissue collection and preparation

Ethical approval was obtained from Liverpool Adult Research Ethics Committee (LREC 09/H1005/55 and 11/H1005/4), and the study was conducted according to the principles of the Declaration of Helsinki and all other applicable national or

local laws and regulations. Written, informed consent was obtained from all participants prior to inclusion in the study at the Liverpool Women's NHS Foundation Trust.

Three secretory phase pre-menopausal uteruses from women undergoing hysterectomy for non-endometrial benign causes (uterine prolapse) who had not been on hormonal treatments in the preceding 3 months were collected. A wedge of tissue from the lumen to the muscular myometrial layer that included superficial and basal endometrium as well as myometrium was taken from the detached uterus, collected in normal buffered formalin and subsequently embedded in paraffin. Several 10- μm -thick parallel tissue sections were cut from each sample and floated onto 1 cm $\times$ 1 cm BaF_2 slides (Photox Optical Systems). These were de-waxed by serial immersion in three sequential fresh xylene baths for 5 min and washed in an acetone bath for a further 5 min. The resulting samples were allowed to air dry and placed in a desiccator until analysis. Four-micrometre-thick parallel tissue sections were floated to glass slides and stained with haematoxylin and eosin (H&E) for histological comparison.

Synchrotron radiation-based Fourier-transform infrared microspectroscopy

The SR-FTIR microspectroscopy setup is illustrated in Fig. 1. Spectral mapping data from 4000 to 600 cm^{-1} were obtained at the Diamond Light Source Ltd., UK (www.diamond.ac.uk) at beamline 22IR. Optically interfaced to the beamline were a Bruker Vertex 80v FTIR spectrometer, a Bruker Hyperion 3000 microscope equipped with front surface reflecting optics, a liquid nitrogen mercury cadmium telluride (MCT) detector and a $\times 36$ objective lens. Spectra were collected in transmission mode from specimens onto BaF_2 slides. At 10- μm step sizes, 256 spectra were co-added with the aperture set for 10 $\mu\text{m} \times 10 \mu\text{m}$ (numerical aperture [NA] = 0.50, lateral resolution 6.7–12.2 μm). Mapping data were acquired within 6 h (Fig. 1a) with background spectra updated after every 10 spectra. Spectra recorded in transmittance were converted and displayed as absorbance with the internal Bruker Optics OPUS 8 software (Bruker Optics, Ettlingen, Germany).

Thermal source focal plane array Fourier-transform infrared microspectroscopy

A Bruker Vertex 80v FTIR spectrometer and Hyperion 3000 microscope were employed using the built-in global (thermal) infrared source. A 64 $\times$ 64 cooled photodiode FPA collected spectra from 4000 to 600 cm^{-1} in parallel at 4 cm^{-1} resolution (NA = 0.65, lateral resolution 5.2–9.4 μm). Co-addition of 128 sample scans and 256 background scans was used. Matching $\times 15$ Cassegrain objective and condenser mirror lenses were used for transmission measurement of microtomed tissue sections mounted on BaF_2 slides (1 cm $\times$

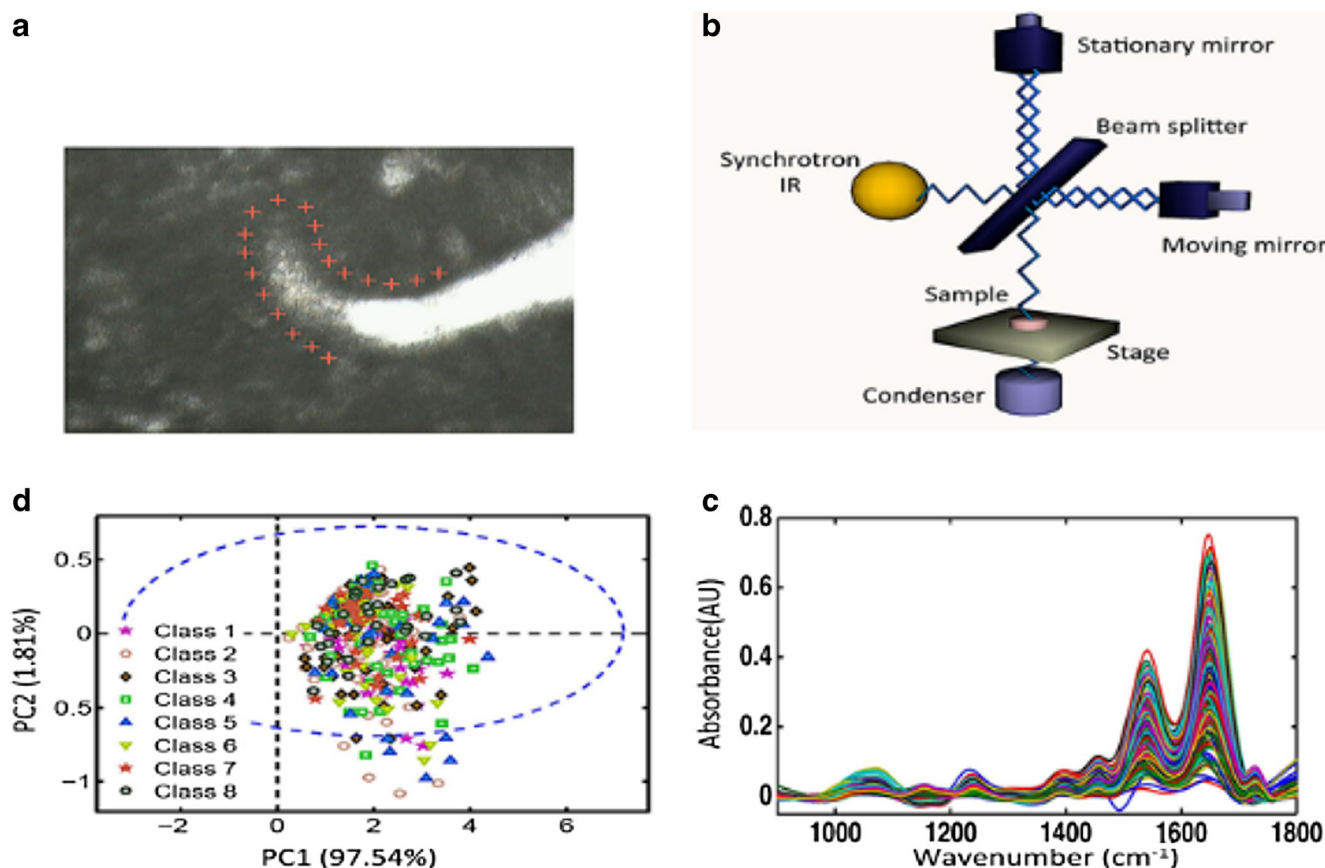


Fig. 1 Synchrotron radiation-based FTIR spectral interrogation and analysis of endometrial epithelial gland bases to identify the location of putative stem cells. **a** Micrograph with superimposed marks representing areas interrogated by synchrotron radiation (magnification $\times 36$). **b**

Diagram representing a Michelson interferometer, which is the basic structure of an IR spectroscope. **c** The interferogram is transformed to an absorbance spectrum by Fourier-transform. **d** PCA identifies “outliers” which may represent putative stem cells

1 cm). FPA imaging maps were acquired from several sections from the three uterine samples involving basalis/functionalis junctions containing epithelial glands. The fidelity of the FPA spectral image region between 1800 and 900 cm^{-1} at 4 cm^{-1} spectral resolution (468 wavenumbers with data spacing of 1.9 cm^{-1}) is the area associated with the biological spectral fingerprint [26] (Fig. 2).

Data pre-processing

Absorbance spectral images were converted to suitable digital files (.txt) for analysis within MATLAB R2014a environment (MathWorks Inc., Natick, USA). They included wavenumber regions scanned between 4000 and 600 cm^{-1} that were truncated to include only the fingerprint region between 1800 and 900 cm^{-1} . The resulting dataset was smoothed by Savitzky-Golay filter, rubber band baseline-corrected and normalised to the amide I peak (i.e. $\approx 1650 \text{ cm}^{-1}$) [21, 27]. The normalisation to amide I peak is commonly performed after baseline correction in order to correct spectra having different absorbance values at amide I region [21]. This prevents the effect of samples having different thicknesses or concentrations that

can mask the actual biochemical difference of interest [27]. The importing and pre-treatment of the spectral data and the construction of chemometric classification models were performed using PLS Toolbox 7.8 (Eigenvector Research Inc., Wenatchee, USA) within MATLAB software.

Computational analysis

Computational analysis consisted of three models: principal component analysis (PCA), successive projections algorithm (SPA) and genetic algorithm (GA). All models were followed by linear discriminant analysis (LDA) [28]. Before applying each analytical model, spectral data were divided into training (70%), validation (15%) and prediction (15%) sets by applying the classic Kennard-Stone (KS) uniform sampling algorithm [29]. The training set was dependent on the number of pixels of the FPA detector for the same IR source, whereas the prediction set was only used for the final classification evaluation.

PCA is a multivariate analysis technique that aims to reduce the number of variables present in the spectral dataset [30]. Principal components (PCs) capture most of the variance

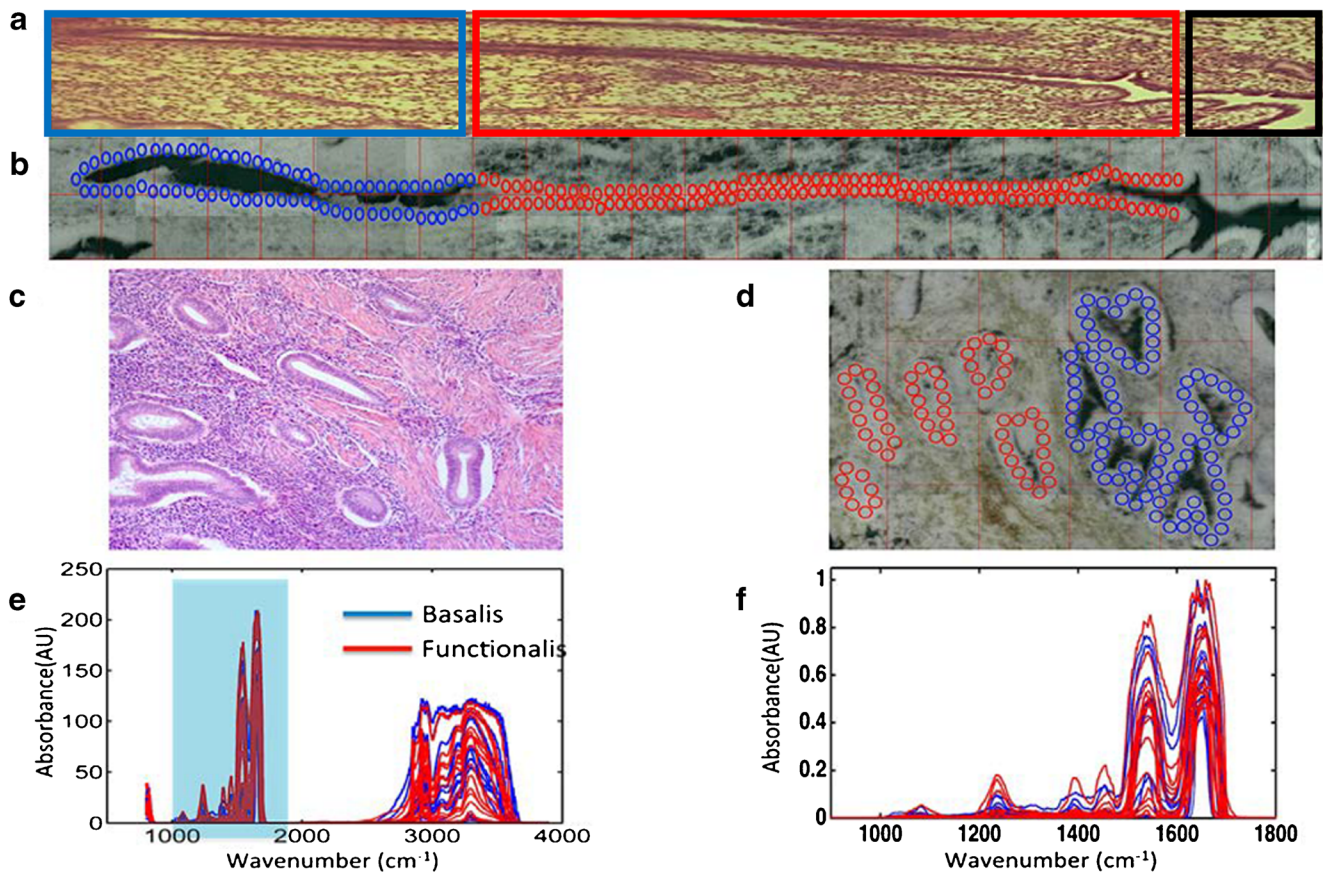


Fig. 2 Representation of region identification for spectral selection by FPA and pre-processing of acquired spectra. **a** H&E-stained parallel section of the examined longitudinally cut crypt, in the secretory phase (circled in blue: cells in the deepest portion of the gland; circled in red: basalis/functionalis cells; circled in black: cells close to the gland opening into the endometrial cavity). **b** Micrograph of the examined longitudinally cut crypt with overlaid circular markers representing epithelial areas which were selected for comparison (magnification $\times 15$) (the actual spectral area selection on the pseudo-image was performed at different magnifications to ensure the spectra selected were good representatives of the

areas to be examined). **c** H&E-stained parallel section of the examined coronally cut crypt, in the secretory phase. **d** Micrograph of the examined coronally cut crypt with overlaid circular markers representing epithelial areas which were selected for comparison (magnification $\times 15$) (the actual spectral area selection on the pseudo-image was performed at different magnifications to ensure the spectra selected were good representatives of the areas to be examined). **e** Averaged absorbance spectra acquired by FPA imaging of all samples. **f** Pre-processed spectra resulting from the raw spectra acquired by FPA imaging of all samples

present in the original dataset. After PCA decomposition, the original data is represented by a combination of scores and loadings. The scores represent the variance on the samples' direction; therefore, they show the similarities and dissimilarities amongst the samples; and the loadings represent the variance on the variables direction; therefore, they show the weight of each wavenumber for the pattern observed on the scores. The scores are used as input for LDA discriminant analysis in PCA-LDA [31]. In addition, after PCA, it is possible to obtain the Q residual value calculated as the sum of the squares of the residual values at each variable for each sample (squared prediction error). This indicates the similarity of the residuals between the samples and its projection into the principal component space, representing the lack of fit in the PCA model. This is therefore a tool for outlier detection [32].

SPA is a forward feature selection method [33], which operates solving co-linearity problems by selecting

wavelengths whose information content is minimally redundant. The model is built through a series of interactions, starting with one wavelength and then incorporating a new one at each iteration until a specified number (N) of wavelengths is reached. Different from PCA, SPA does not modify the original data space, since the projections are used only for selection purposes; thus, the relationship between spectral variables and the original data space is preserved [34].

Genetic algorithms are combinational algorithms inspired by Mendelian genetics. They use a combination of selection, recombination and mutation to evolve a solution to a problem. They treat data as chromosomes allocating reproductive opportunities in such a way that those chromosomes, which represent a better solution to the target problem, are given more chances to "reproduce" than those which represent poorer solutions [34, 35]. The GA routine was carried out using 100 generations containing 200 chromosomes each.

Crossover and mutation probabilities were set to 60 and 10%, respectively. Moreover, the algorithm was repeated three times, starting from different random initial populations. The best solution, in terms of the fitness value, resulting from the three realisations of the GA was employed.

LDA was performed following the application of each of the analytical models. LDA scores, loadings and discriminant function (DF) values were obtained. Usually, the first LDA factor is used to visualise the main biochemical alterations within the sample on a 1-D scores plot. The optimum number of variables for SPA-LDA and GA-LDA was determined by the minimum cost function G calculated for a given validation dataset as:

$$G = \frac{1}{N_V} \sum_{n=1}^{N_V} g_n \quad (1)$$

where N_V is the number of validation samples and g_n is defined as:

$$g_n = \frac{r^2(x_n, m_{I(n)})}{\min_{I(m) \neq I(n)} r^2(x_n, m_{I(m)})} \quad (2)$$

in which $r^2(x_n, m_{I(n)})$ is the squared Mahalanobis distance between object x_n (of class index $I(n)$) and the centre of its true class ($m_{I(n)}$), and $r^2(x_n, m_{I(m)})$ is the squared Mahalanobis distance between object x_n and the centre of the closest wrong class ($m_{I(m)}$) [36].

For each model, sensitivity (the confidence in a positive result for a sample of the label class is obtained) and specificity (the confidence that a negative result for a sample of the non-label class is obtained) were calculated as important quality standards in test evaluation. The quality metrics used in this study for evaluating the classification results can be calculated following the equations [36]:

$$\text{Sensitivity}(\%) = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (3)$$

$$\text{Specificity}(\%) = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \quad (4)$$

where FN is defined as false negative, FP as false positive, TP as true positive and TN as true negative.

Results

Differences between basalis and functionalis for all specimens by FPA imaging

In order to investigate the presence of spectral differences between the functionalis and basalis epithelial layers, all specimens were interrogated using FPA imaging. A raster scan approach was applied to include the epithelial layers

surrounding the glandular base within the specimens. Micrographs of the involved areas with overlaid markers were used for identification of the specific epithelial regions from which spectra were acquired (Fig. 2b, d). Visual correlation with parallel H&E tissue sections was used for selection of different epithelial regions (Fig. 2a, c). Chemometric analyses of the location-derived spectra (Fig. 2e) allowed classification into basalis and functionalis regions.

The chemometric analyses employed for the differentiation of the epithelial cells residing in the two regions included PCA-LDA, SPA-LDA and GA-LDA that were applied to the pre-processed dataset. In total, $n = 65,832$ spectra resided within the functionalis class, and $n = 38,357$ spectra within the basalis class for the 18 areas were interrogated for comparison. Figure 2f shows the pre-processed spectra with the basalis area depicted in blue and the functionalis area in red. Overall, the IR spectra appear to have similarities in the biochemical fingerprint region (1800 to 900 cm^{-1}). Figure 3a shows the explained variance for each PC using PCA-LDA. Eight PCs (explained variance = 89.41%) were used for model construction. Figure 3b shows the 1-D scores plot derived by PCA-LDA. It reveals some segregation of the basalis from the functionalis, although with some overlap of scores along DF1. The majority of the difference between the two regions was attributed to absorptions at 1540 and 1625 – 1675 cm^{-1} according to the PCA loadings [see Electronic supplementary material (ESM) Fig. S1a].

Although GA-LDA classified very successfully the two classes using 14 variables determined from the minimum cost function G (Fig. 3e) (see ESM Table S2), in which the related 1-D scores plot from different classes dissociated along DF1 is shown in Fig. 3f, the chemometric technique that classified the two classes most successfully was SPA-LDA using 12 variables (Fig. 3c) (see ESM Table S2), since better classifying performance was achieved using this technique in the prediction set (Table 1). Also, the related 1-D scores plot illustrates that spectral points from different classes dissociate very well along DF1 (Fig. 3d).

All three techniques identified differences that aided classification within similar spectral regions. These differences were tentatively identified in the spectral regions of 1650 cm^{-1} (amide I band) and 1510 cm^{-1} (C-N stretch contribution to the amide II vibrational band). The overall classification rates by each chemometric technique employed to differentiate the basalis from the functionalis using FPA spectral data are shown in ESM Table S3.

Differences between basalis and functionalis in different individuals by FPA imaging

To investigate whether similar spectral areas are able to identify segregation between the basalis and functionalis regions in different individuals, the same chemometric methods were

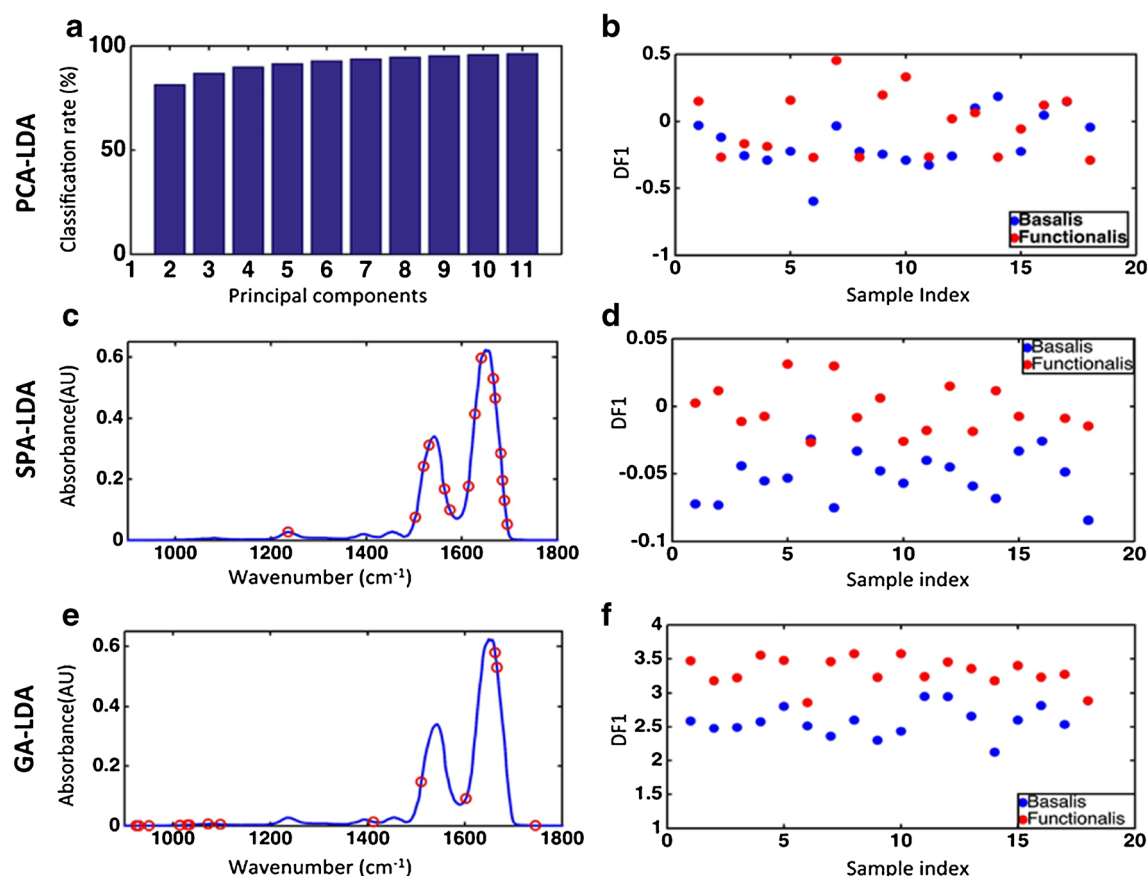


Fig. 3 Classification of basalis and functionalis regions by spectral analysis using PCA-LDA, SPA-LDA and GA-LDA on FPA-FTIR derived data (red = functionalis, blue = basalis): All specimens. **a** Cost/function plot identifying the optimal number of PCs to be used for PCA. **b** Scores plot graphically representing classification by PCA-LDA. The *x*-axis represents the sample index and the *y*-axis the

discriminant factor 1 (DF1). **c** Wavenumber selection for SPA-LDA. **d** Scores plot graphically representing classification by SPA-LDA. The *x*-axis represents the sample index and the *y*-axis DF1. **e** Wavenumber selection for GA-LDA. **f** Scores plot graphically representing classification by GA-LDA. The *x*-axis represents sample index and the *y*-axis DF1

applied on each of the specimens separately. We used the same fingerprint spectral regions for each of the specimens (see ESM Figs. 2a, 3a and 4a). When the average values of the spectra derived from the basalis layer were compared with those from the functionalis layers, subtle differences were visualised (see ESM Figs. S2b, S3b and S4b). The classification rates produced by the different chemometric techniques

Table 1 Predictive performance presented as sensitivity and specificity rates calculated for each chemometric model developed to differentiate basalis and functionalis tissues from the FPA-FTIR spectral data

	PCA-LDA	SPA-LDA	GA-LDA
Basalis			
Sensitivity (%)	66.7	100.0	75.0
Specificity (%)	66.7	100.0	100.0
Functionalis			
Sensitivity (%)	0.0	100.0	60.0
Specificity (%)	0.0	100.0	100.0

(see ESM Table S4) varied from 50.0 to 100.0% in the prediction set. The wavenumbers used for SPA-LDA and GA-LDA in this case are presented in ESM Table S5 and graphically in ESM Figs. S2g and S2e, S3g and S3e, and S4g and S4e, respectively. PCA loadings for these specimens are shown in ESM Fig. S1b, c and d, in which it is observed that most of the differences between the two regions are associated with absorptions on 1235, 1550 and 1628–1685 cm^{-1} . Associated scores plots graphically represent the separation identified between the two layers utilising the three chemometric techniques (see ESM Figs. S2, S3 and S4: d, f and h). Similar spectral areas seem to be causing segregation of the two histologic regions when specimens are analysed separately. These areas mostly include the spectral regions of 1650 cm^{-1} (amide I band) and 1510 cm^{-1} (C-N stretch contribution to the amide II vibrational band). These spectral areas correspond to those responsible for class segregation when all specimens were examined together. This means that spectral variations causing segregation of the two histological areas are

constant independent of other inter-individual variability that may exist within the sample pool.

Inter-individual differences within the basalis layer by FPA imaging

To investigate if the basalis layer is dissimilar in different individuals, PCA-LDA, SPA-LDA and GA-LDA approaches were applied to the pre-processed spectra derived from the basalis regions of all specimens (see ESM Fig. S5a). ESM Fig. S5b shows the average spectrum taken from the basalis layer of each individual. The 2-D scores plots in ESM Fig. S5d, f and h all show clustering of spectra from each individual as well as segregation of the spectra from different specimens. PCA-LDA identified eight PCs (explained variance = 92.61%) as the main responsible for this variation (ESM Fig. S5c). SPA-LDA utilised 12 wavenumbers (ESM Fig. S5e) and GA-LDA utilised 11 wavenumbers (ESM Fig. S5g). The classification rates between the specimens within the basalis layer are shown in ESM Table S6. ESM Table S7 shows the wavenumbers used for the analyses using SPA-LDA and GA-LDA. Wavenumbers between 1600 and 1700 cm^{-1} are consistently responsible for the segregation between the basalis layers of different specimens as well as for the classification into basalis and functionalis layers as shown above.

Investigation of epithelial variability at the base of the glands using synchrotron radiation

The spatial resolution of synchrotron-based FTIR microspectroscopy is in the region of 10 $\mu\text{m} \times 10 \mu\text{m}$. The high resolution allowed the investigation of differences that may exist between epithelial cells that reside close to the base of the glands within the basalis layer. A micrograph of the involved areas with overlaid markers was used for identification of the specific regions of which spectra were acquired. Spectra were acquired along glands, which were cut longitudinally ensuring that the deepest portion of the gland was examined (Fig. 4a). The derived spectra were divided into three classes according to their position (Fig. 4b). Class 1 included spectra taken closer to the deeper margins of the glands with the other classes being sequentially more superficial. Class 1 contained 60 spectra, class 2 contained 310 spectra and class 3 contained 30 spectra. PCA-LDA, SPA-LDA and GA-LDA of the three classes were employed using 2 wavenumbers for SPA and 20 for GA (Fig. 4c, e, g). Initial attempts to classification of these classes were underpowered as revealed by the 2-D scores plots in Fig. 4d, f and h, where segregation of classes is not readily visible, despite noticeable progression of the scores especially in the GA-LDA-derived plot.

To identify whether there is any variation between the epithelial cells resting at the terminal end of the glands, the same

analysis was employed comparing the spectra derived from the classes from above that were furthest apart (classes 1 and 3) (Fig. 5a, b). Interestingly, there was separation between these two classes when compared by all three chemometric techniques as evidenced by the 2-D scores plots in Fig. 5d, f and h, especially using GA-LDA. In this case, eight PCs (explained variance = 93.72%) were used for PCA-LDA and two wavenumbers were utilised for SPA-LDA and eight for GA-LDA (Fig. 5c, e, g and ESM Table S8), being predominantly close to the region of ~ 1234 , ~ 1550 and $\sim 1650 \text{ cm}^{-1}$ which corresponds to the biomarker responses from amide III, amide II and amide I, respectively [26]. The classification rates for each technique in this case are presented in ESM Table S9.

Synchrotron radiation-based IR spectra for the identification of the location of stem cells

To investigate the existence of putative stem cells within the deepest/terminal portion of the glands, we searched for the existence of spectral discriminating factors between all epithelial cells in this area. Each spectrum was treated individually, but spectra were also placed into classes (1 to 8) depending on their distance from the deepest part of the gland (Fig. 6a). Ten adjacent spectra were sequentially placed in each class. The reason for this was to attempt the localisation of putative stem cells in addition to their existence. The original spectra (Fig. 6b) were cut to only include the biological fingerprint region and were pre-processed as before (Fig. 6c). They underwent exploratory PCA using the first 3 PCs (explained variance = 99.61%). Two-dimensional scores plots were extracted representing the first 2 PCs (Fig. 6d). Each point on the resulting scores plot represented a single point on the image maps. Using a 95% confidence interval (CI), spectral points that segregated from the clustered spectra were identified. These “outliers” were particularly obvious on the PC2 axis and included mostly cells from classes 1, 2, 5 and 6. Interestingly, when Q residual analysis was applied, classes 1 and 5 also featured outlier spectra (Fig. 6e). Following analysis by these unsupervised techniques, spectra were classified depending on whether they were identified as outliers or not (Fig. 6f). The averaged spectra resulting from these two classes (“normal” and “outliers”) exposed visible differences (Fig. 6g). A resulting subtraction spectrum identified differences in the absorption bands near 1080 cm^{-1} (symmetric PO_2^- stretching vibrations in RNA and DNA), 1390 cm^{-1} (nucleic acids), 1550 cm^{-1} (amide II) and 1650 cm^{-1} (amide I) (Fig. 6h).

These absorption bands were utilised to produce intensity images based on the FPA spectra to show where in the gland their intensity is maximal to attempt visualisation of “outlier” cellular regions that include putative stem cells (Fig. 7). Although the definition of the FPA images does not allow precise localisation of the highest intensities, it is broadly

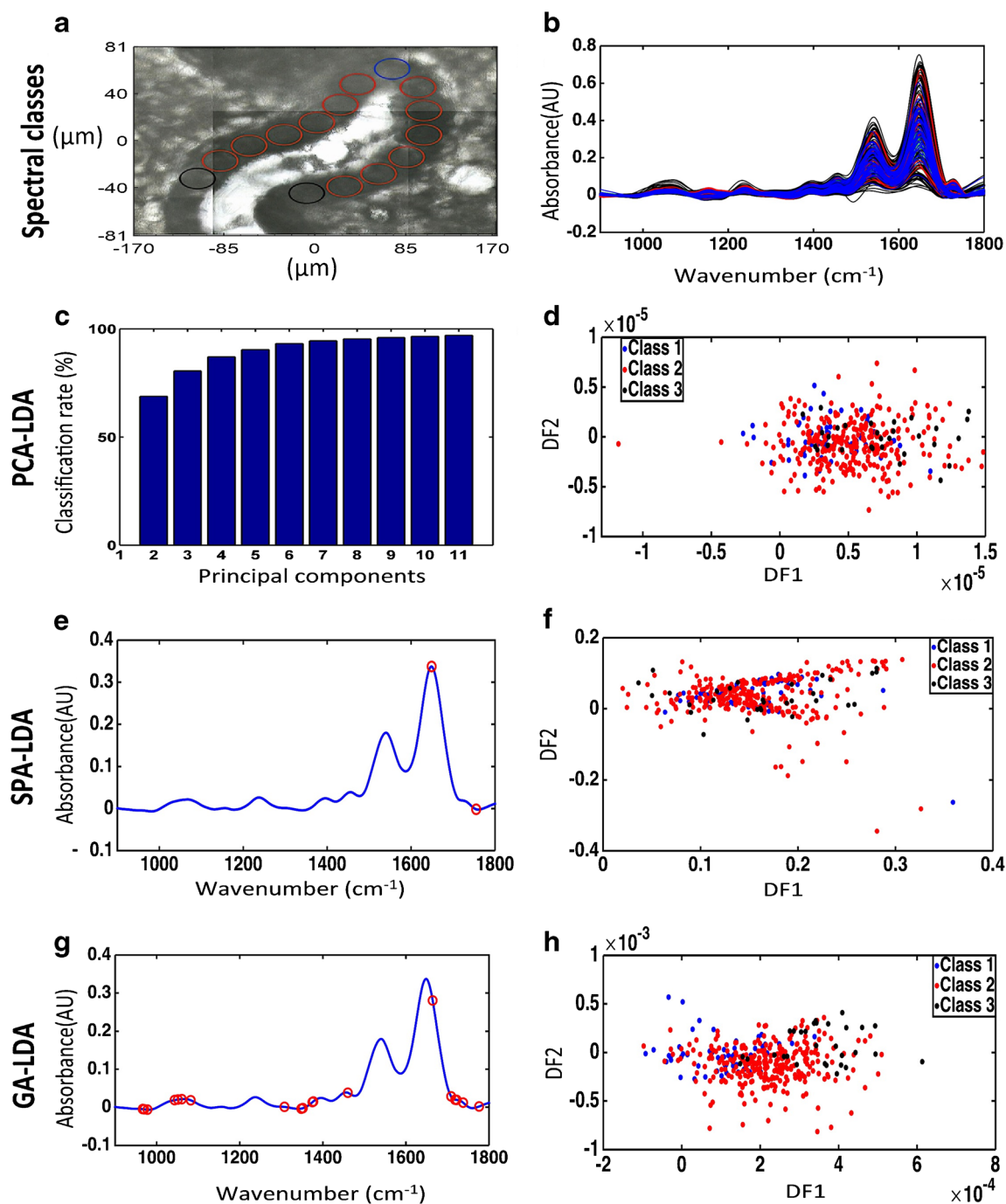


Fig. 4 Classification of epithelial cells according to their position in the endometrial crypt base by synchrotron radiation spectral analysis using PCA-LDA, SPA-LDA and GA-LDA on FPA-FTIR microscopy-derived spectral data. **a** Micrograph with circles representing the regions sampled (this is a representation as the areas analysed were selected at different magnifications to reveal the synchrotron radiation point spectral area). **b** Spectra for the three area classes in the basalis regions of the three specimens. **c** Cost/function plot identifying the optimal number of PCs to be

used for PCA. **d** Scores plot graphically representing classification by PCA-LDA. The *x*-axis represents DF1 and the *y*-axis DF2. **e** Wavenumber selection for SPA-LDA. **f** Scores plot graphically representing classification by SPA-LDA. The *x*-axis represents DF1 and the *y*-axis DF2. **g** Wavenumber selection for GA-LDA. **h** Scores plot graphically representing classification by GA-LDA. The *x*-axis represents DF1 and the *y*-axis DF2

evident that there are in general three areas within the gland where these outlier cells reside: these are along the deepest portion of the gland (Fig. 7a—circle in blue), the basalis/functionalis junction (Fig. 7a—circled in red) and close to

the most superficial part of the gland opening into the endometrial cavity (Fig. 7a—circled in black). The stem cells were found within the population of epithelial cells lining the basalis portion of endometrial glands.

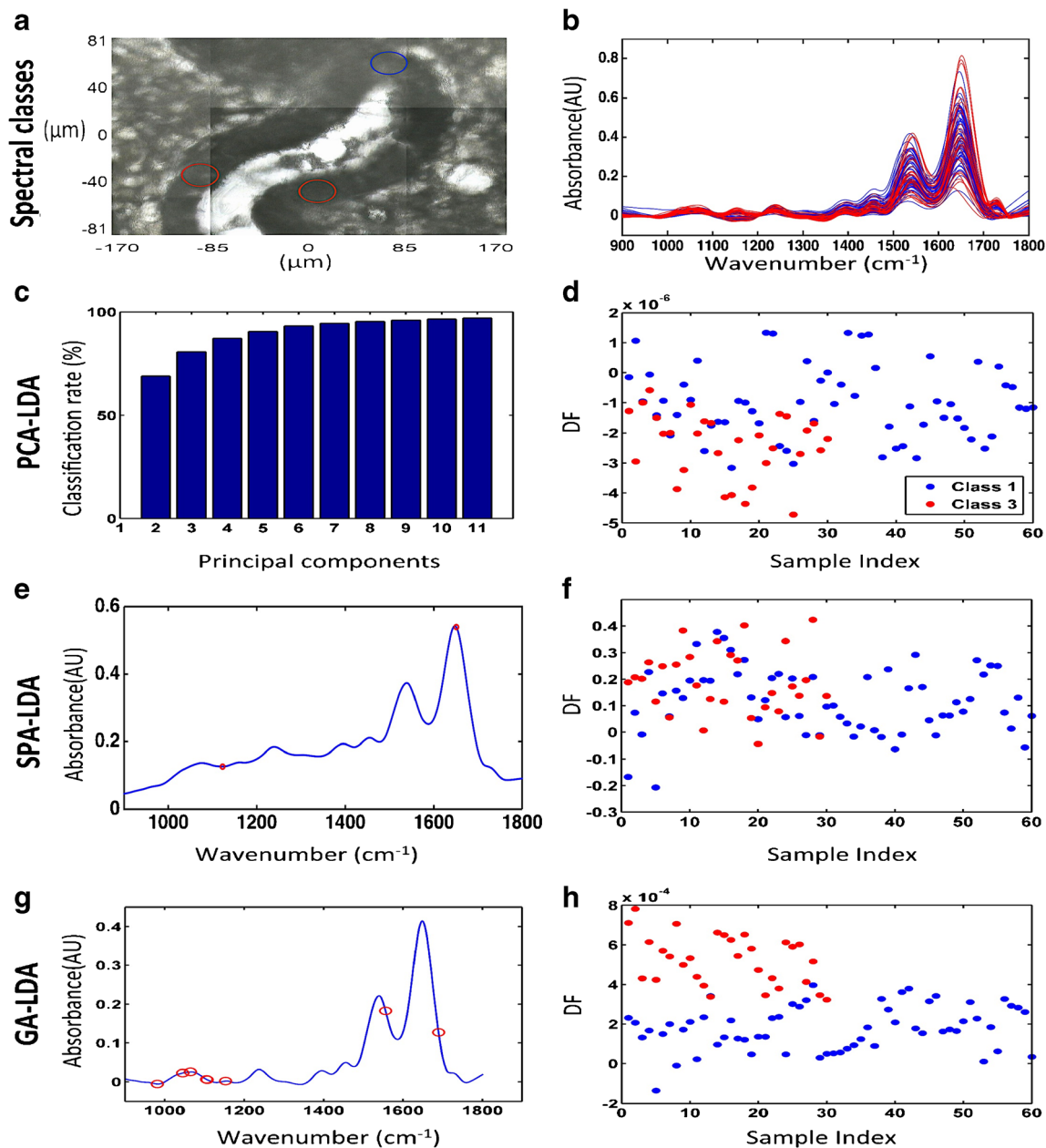


Fig. 5 Classification of epithelial cell classes furthest apart in the endometrial crypt base by synchrotron radiation spectral analysis using PCA-LDA, SPA-LDA and GA-LDA on FPA-FTIR microspectroscopy-derived data. **a** Micrograph with circles representing the regions sampled (this is a representation as the areas analysed were selected at different magnifications to reveal the synchrotron point spectral area). **b** Pre-processed spectra of the classes furthest apart in crypt bases. **c** Cost/function plot identifying the optimal number of PCs to be used for

PCA. **d** Scores plot graphically representing classification by PCA-LDA. The *x*-axis represents the sample index and the *y*-axis DF1. **e** Wavenumber selection for SPA-LDA. **f** Scores plot graphically representing classification by SPA-LDA. The *x*-axis represents the sample index and the *y*-axis DF1. **g** Wavenumber selection for GA-LDA. **h** Scores plot graphically representing classification by GA-LDA. The *x*-axis represents the sample index and the *y*-axis DF1

Discussion

This study employed powerful chemometric techniques in conjunction with synchrotron radiation and global IR technologies to classify epithelial histological areas within the endometrial glands. It also explored the potential of such technologies in isolating putative stem cells and determining their

location within the endometrial epithelium. Since the progress in characterising the human endometrial epithelial stem cells has been particularly poor, our method may attract the attention of endometrial biologists.

Pathophysiological, morphological and molecular alterations of epithelial cells that line the glands in the basalis and functionalis layers may be associated with both benign

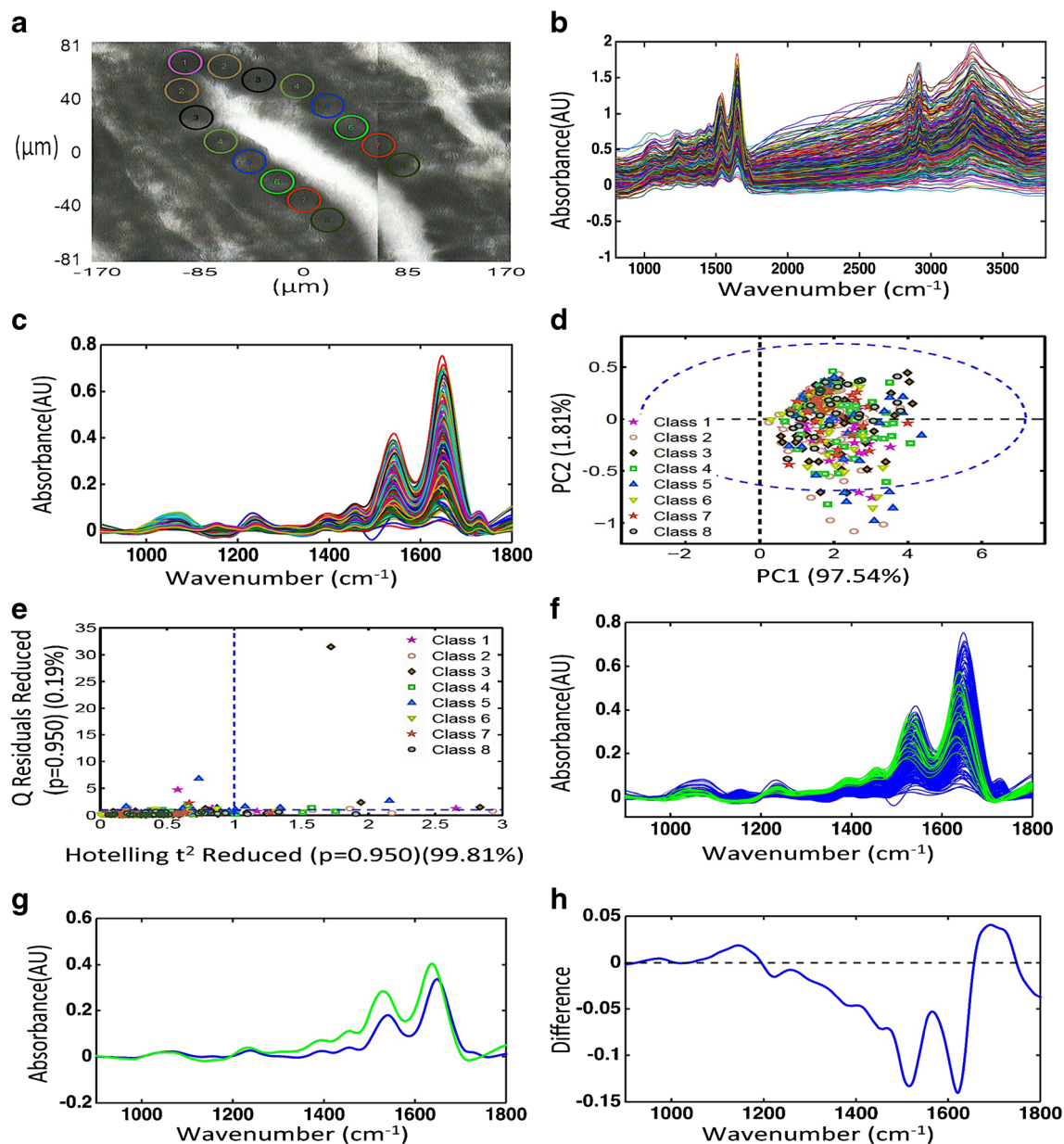


Fig. 6 Exploratory analysis of synchrotron radiation-derived data from crypt bases to isolate putative stem cells and estimate their location. **a** Micrograph with circles representing the regions sampled (this is a representation as the regions analysed were selected at different magnifications to reveal the synchrotron radiation point spectral area). **b** Raw spectra of the classes (regions) sampled in crypt bases. **c** Pre-processed spectra of the classes (regions) sampled in crypt bases. **d** Scores plot graphically representing classification by PCA. The *x*-axis represents PC1 and the *y*-axis PC2, the dotted ellipse the 95 confidence interval (CI). **e** Scores plot

graphically representing classification by Q residuals. The horizontal dotted line represents 2 standard deviations. **f** Pre-processed spectra of the epithelial cells (blue) and “outliers” (green) sampled in crypt bases. **g** Averaged spectra of the epithelial cells (blue) and “outliers” (green) sampled in crypt bases. **h** Loadings curve to identify wavenumbers responsible for segregation of “outliers” from other epithelial cells. The peaks and troughs furthest from the dotted line are the most responsible wavenumber regions

and malignant diseases arising from the endometrium. Therefore, understanding the differences in molecular composition of the cells within the examined histological areas may assist directing research to further pinpoint the exact locations of such pathological changes and to potentially formulate curative strategies.

Computational analysis of the obtained spectra identified significant differences between the epithelial cells lining the glands in the basal and functionalis regions. Some of the discriminating wavenumbers for this variation remained constant when uterine specimens from different individuals were examined separately. Namely, amide I and stretching of C=N bonds correctly classified the two histological regions. This

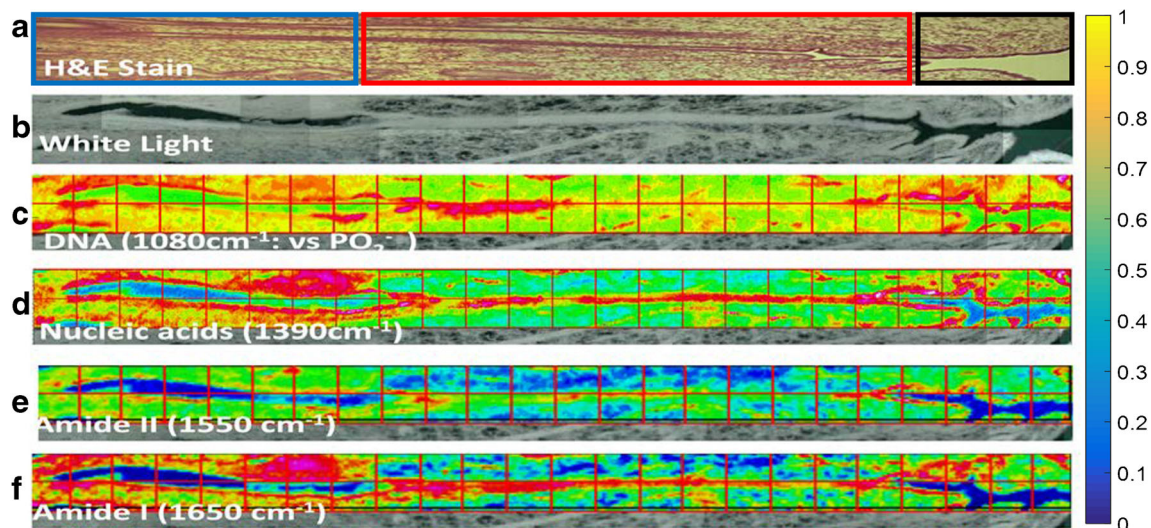


Fig. 7 Intensity images based on FPA imaging (pink represents highest intensity while blue the lowest). **a** H&E-stained longitudinally cut uterine glandular crypt (circled in blue: cells in the deepest portion of the gland; circled in red: basalis/functionalis cells; circled in black: cells close to the gland opening into the endometrial cavity). **b** White light image of the longitudinally cut uterine glandular crypt (magnification $\times 15$). **c** 2-D intensity map derived following FPA spectral analysis at the wavenumber 1080 cm^{-1} overlaid onto a white light micrograph (magnification $\times 15$). **d**

2-D intensity map derived following FPA spectral analysis at the wavenumber 1390 cm^{-1} overlaid onto a white light micrograph (magnification $\times 15$). **e** 2-D intensity map derived following FPA spectral analysis at the wavenumber 1550 cm^{-1} overlaid onto a white light micrograph (magnification $\times 15$). **f** 2-D intensity map derived following FPA spectral analysis at the wavenumber 1650 cm^{-1} overlaid onto a white light micrograph (magnification $\times 15$). Colour bar: normalised absorbance for c–f

molecular variability between the two layers may assist with the further characterisation of the basalis epithelial cells, which are postulated to give rise to many endometrial proliferative conditions such as endometriosis. The functionalis layer is shed during the menstrual shedding leaving the basalis from which the new functionalis is thought to be regenerated. Noe et al. [3] have suggested that the cells from the basalis layer are responsible for the establishment of ectopic endometriotic lesions and that it is further supported by the baboon model of induction of endometriosis [17]. In the baboons, the endometrium biopsied during day 2 of the menses (when the basalis is easily accessible) when instilled in the pelvic cavity gives rise to identical endometriotic deposits to humans. This is likely to be due to the enrichment of stem cells in the basalis layer, that is sampled during menses, and these cells are likely to survive and establish endometriotic lesions. Therefore, identifying such discriminatory spectroscopic biomarkers can also be significant due to their potential for translation into clinical practice. Biomarkers identified using the above techniques may be further explored to interrogate the basalis endometrial layer of women with and without endometriosis.

Several studies have attempted the identification of adult stem cells within the basalis layer [1, 6]. These studies have utilised different stem cell-specific activities for identification [9, 10, 13]. Our study intended to isolate and locate putative epithelial stem cells using morphological, molecular and chemical variability that exists within the cells lining the

glandular glands. There is expanding evidence that FTIR microspectroscopy is capable of identifying stem cells in several tissues including the cornea, epidermis and intestine [23, 37–40]. The chemometric analysis of the IR data we acquired was able to identify specific putative stem cell locations within the population of epithelial cells. Both FPA and synchrotron radiation-based techniques identified putative stem cells at the deepest part of the gland and at a distance of about $50\text{ }\mu\text{m}$ from that location. This indicates that there are at least two sets of stem cells residing in these gland bases. We speculate that the stem cells have different functions: one being dormant and the other actively differentiating to more specialised daughter cells.

Analysis of the synchrotron radiation data has also shown mild progressive differences between the epithelial cells lining the bases of the glands. Although when the synchrotron radiation-derived spectral data from these cells were compared by dividing them into three classes, their variability was minimal; when comparing the classes positioned furthest apart, they exhibited significant differences. This points to a progressive alteration in the molecular structure of epithelial cells even in the basalis layer.

All samples in our study were obtained from women that had hysterectomies for reasons other than endometrial pathology or malignancy. Our aim was to perform preliminary spectroscopic investigations on these tissues to evaluate the performance of these techniques in identifying different cellular lineages within the endometrial glands. The variability

between endometrial samples depends on cycle phase, hormonal factors, lifestyle, body mass index, weight, diet and alcohol consumption, parity and several other factors; therefore, obtaining completely homogeneous endometrial samples from different individuals is virtually impossible. Although our sample size was small, all three women included were sampled within the secretory phase of the cycle to remove cycle-dependent variability and all women had histologically confirmed a healthy endometrium without any prior hormonal treatments. The three different samples in our study were used to support information derived when using all specimens together.

Identification of endometrial adult stem cells may be of value in all endometrial proliferative disease including carcinogenesis research as they may be involved in related processes by their alteration into cancer stem cells [15]. Their localisation by means of specific biomarker as extracted by spectroscopic techniques may in the future provide diagnostic and therapeutic avenues for uterine cancer.

Conclusion

With this study, we demonstrate that SR-FTIR microspectroscopy coupled with multivariate computational analysis and wavenumber selection techniques might be able to identify putative stem cells within the population of epithelial cells lining the basalis portion of endometrial glands. Specific identification of the location of these cells remains elusive, but potential stem cells were segregated from their surrounding cells based on their inherent IR spectroscopic signatures. Specific spectral biomarkers responsible for this segregation were identified and corresponded mainly to PO_2^- vibrational modes of DNA and RNA, nucleic acids and amides I and II. In addition, FPA-FTIR spectroscopy coupled with similar chemometric analysis was able to consistently differentiate the epithelial cells lining the basalis portion of the endometrial glands from those lining the functionalis portion. Further future developments in both spectroscopic technologies and related chemometric techniques may be able to track cellular lineages and individual cells within all tissues, therefore potentially revealing physiological and pathological tissue functions based on morphological alterations.

Funding information The authors would like to thank Rosemere Cancer Foundation for financial support and the Lancashire Teaching Hospitals, Liverpool Women's NHS Foundation Trusts and the Diamond Light Source Research Facility for all the support related to this study. CLMM would like to thank CAPES-Brazil (Doutorado Pleno no Exterior, grant 88881.128982/2016-01) for financial support and KMGL thanks CNPq (grant 305962/2014-0) for financial support.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Gargett CE, Chan RW, Schwab KE. Endometrial stem cells. *Curr Opin Obstet Gynecol*. 2007;19:377–83.
2. Hapangama DK, Kamal AM, Bulmer JN. Estrogen receptor β : the guardian of the endometrium. *Hum Reprod Update*. 2015;21:174–93.
3. Noe M, Kunz G, Herbertz M, Mall G, Leyendecker G. The cyclic pattern of the immunocytochemical expression of oestrogen and progesterone receptors in human myometrial and endometrial layers: characterization of the endometrial-subendometrial unit. *Hum Reprod*. 1999;14:190–7.
4. Li L, Xie T. Stem cell niche: structure and function. *Annu Rev Cell Dev Biol*. 2005;21:605–31.
5. Eckfeldt CE, Mendenhall EM, Verfaillie CM. The molecular repertoire of the 'almighty' stem cell. *Nat Rev Mol Cell Biol*. 2005;6:726–37.
6. Gargett CE, Masuda H. Adult stem cells in the endometrium. *Mol Hum Reprod*. 2010;16:818–34.
7. Patel AN, Park E, Kuzman M, Benetti F, Silva FJ, Allickson JG. Multipotent menstrual blood stromal stem cells: isolation, characterization, and differentiation. *Cell Transplant*. 2008;17:303–11.
8. Musina RA, Belyavski AV, Tarusova OV, Solovyova EV, Sukhikh GT. Endometrial mesenchymal stem cells isolated from the menstrual blood. *Bull Exp Biol Med*. 2008;145:539–43.
9. Chan RW, Schwab KE, Gargett CE. Clonogenicity of human endometrial epithelial and stromal cells. *Biol Reprod*. 2004;70:1738–50.
10. Gargett CE, Schwab KE, Zillwood RM, Nguyen HP, Wu D. Isolation and culture of epithelial progenitors and mesenchymal stem cells from human endometrium. *Biol Reprod*. 2009;80:1136–45.
11. Kim JY, Tavare S, Shibata D. Counting human somatic cell replications: methylation mirrors endometrial stem cell divisions. *Proc Natl Acad Sci U S A*. 2005;102:17739–44.
12. Valentijn AJ, Palial K, Al-lamee H, Tempest N, Drury J, Von Zglinicki T, et al. SSEA-1 isolates human endometrial basal glandular epithelial cells: phenotypic and functional characterization and implications in the pathogenesis of endometriosis. *Hum Reprod*. 2013;28:2695–708.
13. Masuda H, Kalka C, Takahashi T, Yoshida M, Wada M, Kobori M, et al. Estrogen-mediated endothelial progenitor cell biology and kinetics for physiological postnatal vasculogenesis. *Circ Res*. 2007;101:598–606.
14. Götte M, Wolf M, Staebler A, Buchweitz O, Kelsch R, Schüring AN, et al. Increased expression of the adult stem cell marker Musashi-1 in endometriosis and endometrial carcinoma. *J Pathol*. 2008;215:317–29.
15. Rosen JM, Jordan CT. The increasing complexity of the cancer stem cell paradigm. *Science*. 2009;324:1670–3.

16. Sasson IE, Taylor HS. Stem cells and the pathogenesis of endometriosis. *Ann N Y Acad Sci.* 2008;1127:106–15.
17. Sourial S, Tempest N, Hapangama DK. Theories on the pathogenesis of endometriosis. *Int J Reprod Med.* 2014;2014:179515.
18. Bokor A, Debrock S, Drijkoningen M, Goossens W, Fülöp V, D'Hooghe T. Quantity and quality of retrograde menstruation: a case control study. *Reprod Biol Endocrinol.* 2009;7:123.
19. Siqueira LFS, Lima KMG. A decade (2004–2014) of FTIR prostate cancer spectroscopy studies: an overview of recent advancements. *Trends Anal Chem.* 2016;82:208–21.
20. Siqueira LFS, Lima KMG. MIR-biospectroscopy coupled with chemometrics in cancer studies. *Analyst.* 2016;141:4833–47.
21. Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, et al. Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc.* 2014;9:1771–91.
22. Taylor SE, Cheung KT, Patel II, Trevisan J, Stringfellow HF, Ashton KM, et al. Infrared spectroscopy with multivariate analysis to interrogate endometrial tissue: a novel and objective diagnostic approach. *Br J Cancer.* 2011;104:790–7.
23. Walsh MJ, Fellous TG, Hammiche A, Lin W-R, Fullwood NJ, Grude O, et al. Fourier transform infrared microspectroscopy identifies symmetric PO_2^- modifications as a marker of the putative stem cell region of human intestinal crypts. *Stem Cells.* 2008;26:108–18.
24. Tobin MJ, Chesters MA, Chalmers JM, Rutten FJ, Fisher SE, Symonds IM, et al. Infrared microscopy of epithelial cancer cells in whole tissues and in tissue culture, using synchrotron radiation. *Faraday Discuss.* 2004;126:27–39.
25. Miller LM, Dumas P. Chemical imaging of biological tissue with synchrotron infrared light. *Biochim Biophys Acta.* 1758;2006:846–57.
26. Movasaghi Z, Rehman S, Ur Rehman I. Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Appl Spectrosc Rev.* 2008;43:134–79.
27. Trevisan J, Angelov PP, Carmichael PL, Scott AD, Martin FL. Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives. *Analyst.* 2012;137:3202–15.
28. Martin FL, Kelly JG, Llabjani V, Martin-Hirsch PL, Patel II, Trevisan J, et al. Distinguishing cell types or populations based on the computational analysis of their infrared spectra. *Nat Protoc.* 2010;5:1748–60.
29. Kennard RW, Stone LA. Computer aided design of experiments. *Technometrics.* 1969;11:137–48.
30. Bro R, Smilde AK. Principal component analysis. *Anal Methods.* 2014;6:2812–31.
31. Morais CLM, Lima KMG. Principal component analysis with linear and quadratic discriminant analysis for identification of cancer samples based on mass spectrometry. *J Braz Chem Soc.* 2018;29:472–81.
32. Keithley RB, Heien ML, Wightman RM. Multivariate concentration determination using principal component regression with residual analysis. *Trends Anal Chem.* 2009;28:1127–36.
33. Soares SFC, Gomes AA, Araujo MCU, Filho ARG, Galvão RKH. The successive projections algorithm. *Trends Anal Chem.* 2013;42:84–98.
34. Theophilou G, Lima KMG, Briggs M, Martin-Hirsch PL, Stringfellow HF, Martin FL. A biospectroscopic analysis of human prostate tissue obtained from different time periods points to a trans-generational alteration in spectral phenotype. *Sci Rep.* 2015;5:13465.
35. Whitley D. A genetic algorithm tutorial. *Stat Comput.* 1994;4:65–85.
36. Siqueira LFS, Araújo Júnior RF, Araújo AA, Morais CLM, Lima KMG. LDA vs. QDA for FT-MIR prostate cancer tissue classification. *Chemom Intell Lab Syst.* 2017;162:123–9.
37. Kelly JG, Nakamura T, Kinoshita S, Fullwood NJ, Martin FL. Evidence for a stem-cell lineage in corneal squamous cell carcinoma using synchrotron-based Fourier-transform infrared microspectroscopy and multivariate analysis. *Analyst.* 2010;135:3120–5.
38. Fogarty SW, Patel II, Trevisan J, Nakamura T, Hirschmugl CJ, Fullwood NJ, et al. Sub-cellular spectrochemical imaging of isolated human corneal cells employing synchrotron radiation-based Fourier-transform infrared microspectroscopy. *Analyst.* 2013;138:240–8.
39. Patel II, Harrison WJ, Kerns JG, Filik J, Wehbe K, Carmichael PL, et al. Isolating stem cells in the inter-follicular epidermis employing synchrotron radiation-based Fourier-transform infrared microspectroscopy and focal plane array imaging. *Anal Bioanal Chem.* 2012;404:1745–58.
40. Walsh MJ, Hammiche A, Fellous TG, Nicholson JM, Cotte M, Susini J, et al. Tracking the cell hierarchy in the human intestine using biochemical signatures derived by mid-infrared microspectroscopy. *Stem Cell Res.* 2009;3:15–27.