

SR-FTIR spectroscopy of renal epithelial carcinoma side population cells displaying stem cell-like characteristics†

Caryn Hughes,^{ab} Matthew Liew,^b Ashwin Sachdeva,^a Paul Bassan,^a Paul Dumas,^c Claire A. Hart,^b Mick D. Brown,^b Noel W. Clarke^b and Peter Gardner^{*a}

Received 26th July 2010, Accepted 5th October 2010

DOI: 10.1039/c0an00574f

It is hypothesized that cells with stem cell-like properties may be influential in carcinogenesis, possessing the ability to self-renew, produce differentiated daughter cells and resist environmental or therapeutic injury. This has led to a surge in interest in identifying and characterizing the tumour initiating or cancer stem cell (CSC) with the aim of discovering novel diagnostic and prognostic markers and of understanding the basic biology with the ultimate aim of generating new therapeutic approaches and biomarkers. However, a major hurdle to this process has been the lack of a truly specific cancer stem cell biomarker allied to the rarity of these cells. This has led to problems in characterising these CSCs by traditional ‘-omic’ techniques. Using a renal carcinoma cell line model, we show that synchrotron radiation-Fourier transform infrared (SR-FTIR) spectroscopy is a suitable tool to measure discrete differences in the biochemistry of small numbers of single-cells. Using the chemometric techniques of Principal Component and Linear Discriminant Analysis (PCA and LDA) for multivariate reduction, biochemical differences between the cells from different sub-populations were evaluated. Results found lipid and phosphodiester vibrations to be particularly good discriminating markers in the spectra of these stem-like cells, relative to the more differentiated, proliferating cells that make up the majority of the cell population.

Introduction

Adult stem cells, originally described in the haemopoietic system, have been described in many normal solid tissues including the brain, liver, muscle, breast, skin, gut, prostate, bladder and kidney.^{1–12} They are rare by comparison with the number of amplifying and terminally differentiated cells of the organ of origin. Stem cells are required for tissue development and are stimulated to differentiate in response to tissue injury in addition to general replacement and repair. Unlike their differentiated progeny, stem cells are long-lived and present throughout the lifetime of the organ. This longevity makes them susceptible to the accumulation of genetic mutations. This has given rise to the cancer stem cell hypothesis, which proposes that within a given tumour there exists a small population of cells with stem-like ability, *i.e.* the capacity to exhibit properties including self-renewal and the ability to undergo asymmetric division leading to the generation of progeny that can differentiate to produce a tumour with a heterogeneous phenotype.

Adult stem cells and putative cancer stem cells are rare and due to the lack of specific stem cell markers isolation and characterisation has been problematic. We have shown previously that the Hoechst 33342 dye efflux assay (Fig. 1), developed by Goodell *et al.* in the haematopoietic system,¹³ can be adapted to solid tissue tumours to isolate a side population (SP) enriched for cells which possess many of the characteristics of stem cells.^{1,4,5}

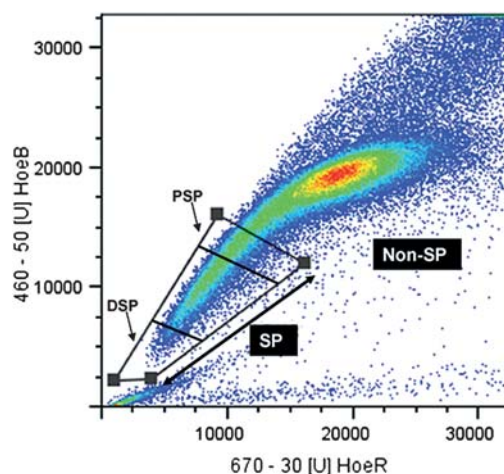


Fig. 1 Hoechst 33342 profile of the 2245R renal clear-cell carcinoma epithelial cell line. A typical Hoechst 33342 red/blue profile of 2245R cells stained with Hoechst 33342 prior to sorting on a FACS Vantage cell sorter. SP gate was defined using the SP inhibitor Verapamil. The SP can be sub-sorted into proximal (PSP) and distal (DSP) side populations. The signals falling below the gate (indicated by the black lines) represent non-viable cellular debris.

^aManchester Interdisciplinary Biocentre, University of Manchester, 131 Princess Street, Manchester, UK M1 7DN

^bGenito-Urinary Cancer Research Group, School of Cancer and Enabling Sciences, Paterson Institute for Cancer Research, University of Manchester, Christie Hospital NHS Trust, Manchester Academic Health Sciences Centre, Manchester, UK M20 4BX

^cSynchrotron Soleil, L'orme des Merisiers, Saint-Aubin, BP48 91192 Gif-sur-Yvette Cedex, France

† This article is part of a themed issue on Optical Diagnosis. This issue includes work presented at SPEC 2010 Shedding Light on Disease: Optical Diagnosis for the New Millennium, which was held in Manchester, UK June 26th–July 1st 2010.

However, although the SP is enriched for stem-like cells, it is still heterogeneous. Within the SP it is possible to isolate further sub-fractionated cells with greater effluxing ability, appearing further down the tail of the SP FACS profile. These 'distal' SP cells (DSP) possess more 'primitive' characteristics hence more closely resembling true stem cells.¹⁴

Characterisation of the SP sub-populations is problematic due to the very low numbers of cells within each sub-population, often below the level required for traditional cellular and '-omic' techniques. Thus, biospectroscopy has considerable potential in this arena as an important tool in stem cell research.¹⁵ Here we show that synchrotron radiation-Fourier transform infrared (SR-FTIR)¹⁶ can be utilised as a non-destructive analytical technique to generate spectral chemical profiles of small cellular sub-populations which enable study providing insight into the nature of small and rare subsets of cells.

Experimental

FACS sorting of cell populations

The 2245R renal carcinoma cell line was cultured in DMEM with FCS (10%) and L-glutamine (1%) at 37 °C in a humidified atmosphere (5% CO₂) until 80% confluent. Hoechst 33342 staining was conducted according to the method of Addla *et al.*¹ Briefly, 1 × 10⁶ 2245R cells in 1 ml Hoechst buffer (Hank's buffered salt solution, 10% FCS, 20 mM HEPES, and 1% D-glucose) were stained with 5 mM Hoechst 33342 for 90 min at 37 °C with continuous agitation (5 Hz). Cells were washed and re-suspended in ice-cold Hoechst buffer prior to FACS analysis. Hoechst 33342 dye efflux SP was defined by 50 mM Verapamil hydrochloride blockade of Hoechst 33342 dye efflux. Hoechst 33342 staining was detected using a Becton Dickinson FACS Vantage SE flow cytometer (FACS), exciting at 357 nm and detecting Hoechst Blue with a 424/44 broad pass (BP) filter and Hoechst Red with a 675/20 BP filter (Omega Optical, Brattleboro, VT) SP, as well as distal- and proximal-SP, and Non-SP cells were 'sorted' based on pre-defined profiles.

Sample preparation for spectroscopy

Following cell sorting, the isolated cell populations were fixed in formalin (4% in PBS). The cells were subsequently spun and re-suspended in 500 µL of PBS. The cells were then mixed to form a homogeneous suspension and 150 µL aliquots were used for deposition with a cell centrifuge (Cytospinner) onto CaF₂ substrates. The cells were initially spun at a gentle 400 g for 5 min, then additionally at 800 g for 5 min. The substrates were left overnight to dry before washing with distilled water to remove residual salt, originating from the PBS.

Data acquisition

SR-FTIR transmission spectra were acquired on the SMIS beamline of the French Synchrotron Facility SOLEIL (Saint-Aubin, France) using a Nicolet Continuum XL microscope equipped with an MCT detector.¹⁶ Spectra were recorded at 4 cm⁻¹ resolution with 256 co-scans and the size of the aperture was adjusted to match the diameter of the cell such that it was

fully illuminated, typically 14 µm × 14 µm. A total of 25 spectra were recorded for each cell type.

Data analysis

The raw spectra were corrected with an ATR Matrigel reference spectrum and 100 iterations of the RMieS-EMSC algorithm.^{17–20}

The corrected spectra were transformed to the second derivative with 7 point Savitsky–Golay smoothing with a polynomial order of 3 in Omnic 7.2 (Thermo Electron Corporation).

Analysis was performed on the second derivative corrected spectra with a range of 1475–1000 cm⁻¹ to achieve maximum chemometric separation between the Non-SP and the SP cells. The corrected spectra were then vector normalised and mean-centred before principal component analysis (PCA), performed in Matlab (The Mathworks Inc.). PC–LDA model was generated using SPSS (SPSS Inc.).

Results and discussion

There are currently no markers available for the isolation of putative stem-like cells from either normal or malignant renal tissue. Addla *et al.* have shown that SP only constituted between 4 and 6% of the epithelial population and expresses a range of markers, none of which is specific for one population or the other.¹ It is therefore necessary to use a functional approach for enriching populations of stem-like cells prior to further characterisation and analysis. A tried and tested protocol for isolating enriched stem cell-like populations is the Hoechst 33342 dye efflux model, which has been shown to isolate populations of haemopoietic stem cells from blood and bone marrow and populations of enriched stem-like cells from various organs including breast, prostate, brain and kidney. It is therefore not possible to conduct a study such as this without the use of a discriminatory protocol to isolate/enrich populations for the cells of a particular phenotype. Although the dye is effluxed out of the cell after several hours, the cells still contain levels of dye. The FACS blot (Fig. 1) clearly shows a SP 'Tail' which contains cells with various concentrations of the dye within. If the cells had no Hoechst 33342 dye within them they would fall within the bottom left hand corner of the plot and would not be collected for this study. The Hoechst 33342 red blue FACS profile generated by 2245R shows a Verapamil sensitive SP containing up to 16% (average 6%) of the renal epithelial cells. This enabled the isolation of 3 sub-populations for analysis: Non-SP containing differentiated cells and 2 SP sub-populations; the PSP containing putative transit amplifying cells, and the DSP which is thought to contain the most primitive (stem) cell types.

Live cells can be obtained from this technique, sorted by FACS and we have shown limited cell culture of the side population phenotype is possible. Fig. 2 compares centrifuge (cyto)spun Non-SP cells onto CaF₂ slides and photomicrographs of the cell types *in vitro*, after 24 h of cell sorting, illustrating the morphological changes and the cellular differentiation that occur if the cells are not fixed and instead cultured. Variance in the morphological development of the cell types is apparent. The Non-SP cells (Fig. 2B) begin to extend their cytoplasmic boundaries, a characteristic morphology of the clear-cell renal carcinoma phenotype. The SP cells remain spherical in

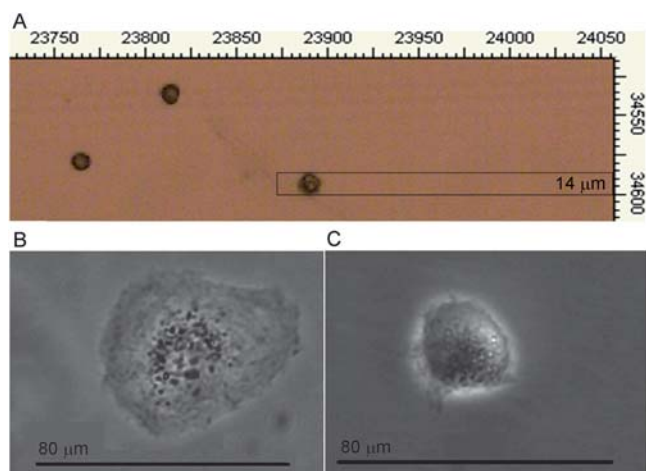


Fig. 2 (A) Isolated FACS sorted cells, chemically fixed immediately after sorting for synchrotron experiment, spun onto CaF_2 slides, (B) *in vitro* photomicrographs of Non-SP cells, and (C) SP cells after 24 h of cell culture post-FACS sorting.

morphology more in keeping with a stem cell phenotype (Fig. 2C). In terms of synchrotron spectroscopy we used cells which had been chemically fixed in formalin (4%) (Fig. 2A). Cells are fixed immediately, and therefore viability issues are minimal. Secondly the fixation preserves the cellular content at the time of sorting so that the cells cannot differentiate further by the time of acquiring their spectroscopic fingerprint. The cells are presented in a 'snapshot' of time, representing a controlled and fair comparison. All cell types analysed remain spherical by plating the cells onto a CaF_2 substrate at a predetermined force so that the cells remain intact whilst attaching to the substrate. This is an

alternative to culturing cells directly onto substrates, resulting in possible extensions of the cellular cytoskeleton and also possible differentiation from the time of establishment onto a 2D surface.

It is also understood that even small differences in cell size may dominate chemometric discrimination, as different sizes of cells would correspond to different resonant Mie scattering profiles which would be detected in chemometric discrimination. This concern can be reduced with use of the recently described RMieS-EMSC correction algorithm.^{17–20}

Fig. 3 presents the mean infrared spectrum with standard deviation of each type of sorted cell. The mean spectra are plotted in bold colour alongside their relative standard deviation plotted as dotted lines. The non-side population displayed the greatest spectral variance, particularly in the $1250\text{--}1000\text{ cm}^{-1}$ region. This is understood to represent the heterogeneity of the non-side population, with single cells more likely than the SPs to adapt different cycle conformations at the time of FACS sorting. Within the cell cycle, changes in intensity of certain molecular vibrations can be detected, as discussed elsewhere.^{21–23}

This is also apparent in the PCA score plot of the total dataset (Fig. 4A). The Non-SP cell spectra are more dispersed than the side population cells of DSP, PSP and SP. There is partial overlap between the Non-SP and SP spectra, as would be expected in a single cell line. However, it is apparent that there is some degree of separation between Non-SP and SP spectra in PC3. The SP contained a mixture of DSP and PSP cells and therefore it was expected that the SP cluster would fall on the same trajectory as the DSP and PSP spectra. The associated loading plot for PC3 is displayed in Fig. 4B. As second derivative peak minima are directly aligned to the centre of the absorbance peak of the original spectra, the loadings can be interpreted as biological references. There are many features that contribute to

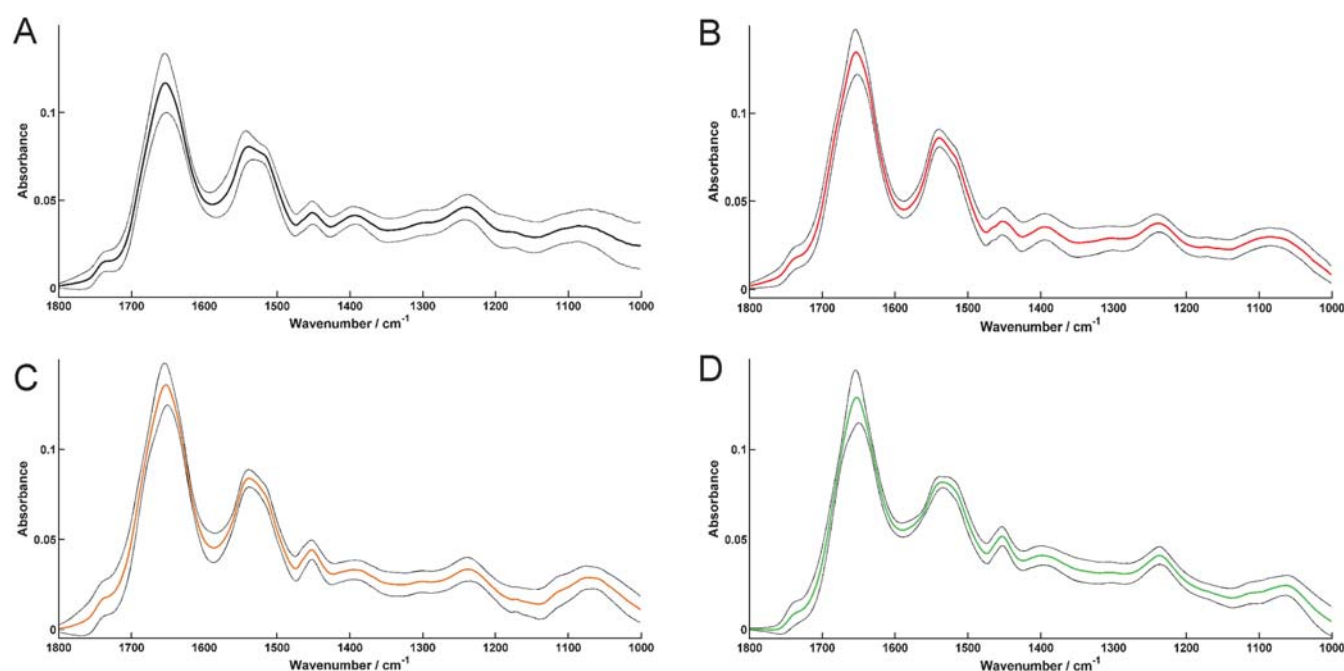


Fig. 3 SR-FTIR mean fingerprint spectra \pm standard deviation from single cells of (A) non-side population (B) side population, (C) sub-side population; proximal, (D) sub-side population; distal. For multivariate analysis the range was reduced to $1475\text{--}1000\text{ cm}^{-1}$ and the spectra transformed to the second derivative.

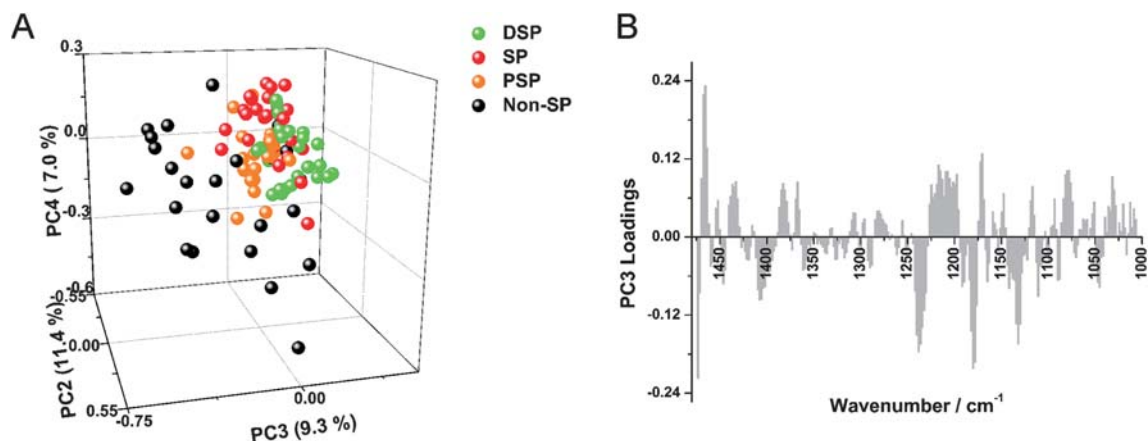


Fig. 4 Total dataset analysis (A) PCA plot of PC2, PC3 and PC4. (B) The loadings of PC3. Discrimination of the Non-SP cells can be seen through principal component 3.

the differences between the non-side populations and side-populations. These include: $\delta_{\text{as}}(\text{CH}_2)$ scissoring vibration of lipids at $\sim 1468 \text{ cm}^{-1}$; $\nu_{\text{s}}\text{C}=\text{O}$ (of COO^-) at $\sim 1400 \text{ cm}^{-1}$; $\delta_{\text{s}}(\text{CH}_3)$ at $\sim 1380 \text{ cm}^{-1}$; $\nu_{\text{as}}\text{P}=\text{O}$ (of PO_2^-) of the phosphodiester group of phospholipids, phosphorylated proteins and nucleic acids (arising from the phosphate sugar backbone of DNA and RNA) at $\sim 1237 \text{ cm}^{-1}$; amide III band protein components at $\sim 1240\text{--}1310 \text{ cm}^{-1}$; $\nu_{\text{as}}(\text{CO}\text{--}\text{O}\text{--}\text{C})$ ring vibrations of carbohydrates at $\sim 1180 \text{ cm}^{-1}$; $\nu_{\text{s}}\text{P}=\text{O}$ (of PO_2^-) of phosphodiesters at $\sim 1080 \text{ cm}^{-1}$; $\nu_{\text{s}}(\text{CO}\text{--}\text{O}\text{--}\text{C})$ of carbohydrates at $\sim 1050 \text{ cm}^{-1}$ and $\nu_{\text{s}}(\text{C}\text{--}\text{O})$ of carbohydrates at $\sim 1030 \text{ cm}^{-1}$.

For clarity, principal component analysis was repeated for the three specific cell types in the dataset, namely, DSP, PSP and

Non-SP. Fig. 5A illustrates the resulting PC score and plot for PCs 1, 2 and 3, and the loading plot of PC1 corresponding to the discrimination of the Non-SP from the SP spectra (Fig. 5B). Here the lipid signals, mentioned in the previous PC analysis, dominate the loadings plot of PC1. However, differences in the phosphodiester stretching absorptions also remain apparent in the loadings of PC2, along with carbohydrate vibrations (Fig. 5C).

Gault *et al.* looked into the effect of the Hoechst 33342 dye on cell FTIR spectra as a negative control when investigating alpha particle cellular irradiation.²⁴ They found that Hoechst 33342 treatment of cells did not damage DNA. They also found that in a qualitative exercise comparing the FSD mean spectra of stained

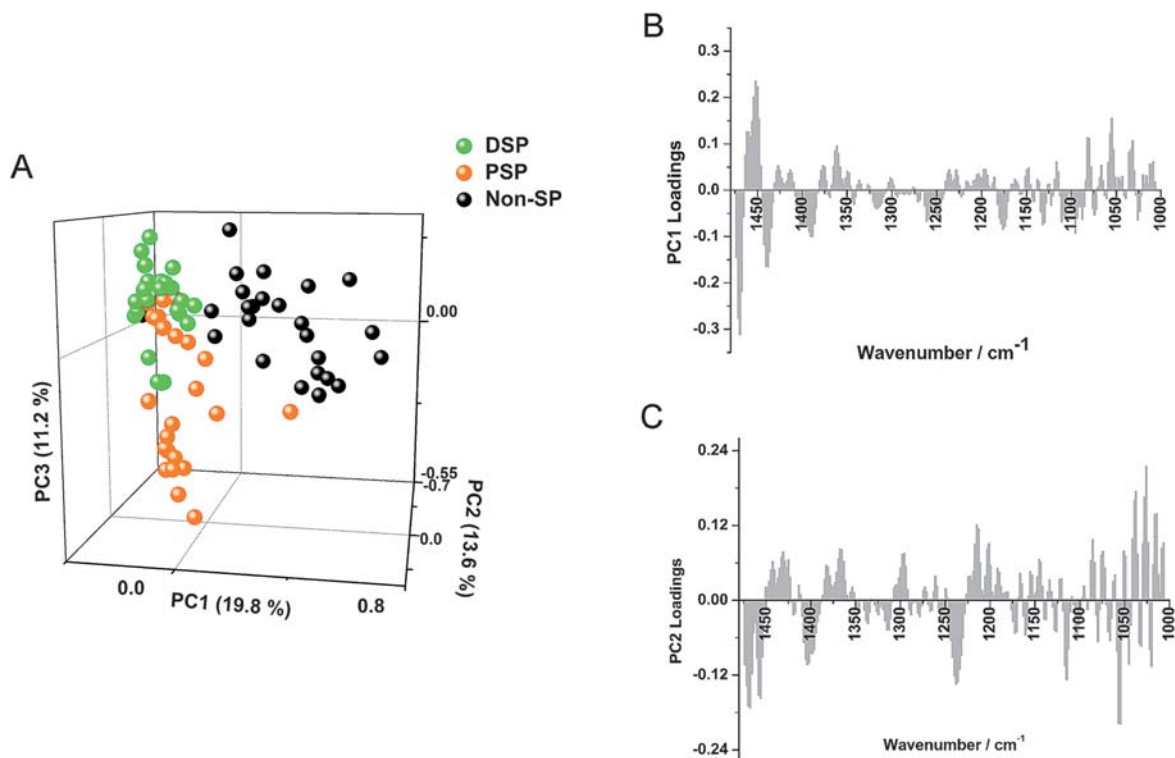


Fig. 5 PC analysis of the sub-side populations DSP and PSP *versus* Non-SP cell spectra. (A) The score plot of PC1, PC2 and PC3. (B) The corresponding loadings of PC1 and (C) PC2. Key biochemical differences are outlined in lipid, phosphodiester and carbohydrate absorption bands.

and unstained cells, binding of Hoechst 33342 did not induce intensity or band shift changes in the IR regions characteristic of proteins and lipids and there was only a subtle change in DNA conformation. (There was a 4 cm^{-1} shift in $\nu_{\text{as}}\text{P}=\text{O}$ (of PO_2^-) centred at $1238\text{--}1242\text{ cm}^{-1}$; however the $\nu_{\text{s}}\text{P}=\text{O}$ (of PO_2^-) centred at 1086 cm^{-1} remained unchanged.²⁴)

As a comparison, the mean second derivative spectra for Non-SP, PSP and DSP were qualitatively assessed (Fig. 6). There is, however, a danger when comparing averaged spectra for a cell type, particularly if one population is more heterogeneous, and the sample population is small. Therefore this is a more subjective approach than analysing PCA loadings which take account all individual cell spectra. The assignments of notable differences in the second derivative spectra were: (1) $\delta_{\text{as}}(\text{CH}_2)$ shift from 1464 to 1468 cm^{-1} between DSP and Non-SP and an increase in absorbance; (2) increased absorbance for DSP/PSP at 1452 cm^{-1} (C–H bending); (3) increase in absorbance for PSP at 1414 cm^{-1} ; (4) $\delta_{\text{s}}(\text{CH}_3)$ shift from 1384 cm^{-1} to 1387 cm^{-1} between DSP and Non-SP; (5) DSP is shifted to 1365 cm^{-1} relative to PSP/Non-SP at 1367 cm^{-1} ; (6,7) notable absorption bands at 1341 cm^{-1} and 1331 cm^{-1} for Non-SP; (8) prominent band in PSP/Non-SP at 1308 cm^{-1} ; (9) carbohydrate marker at 1255 cm^{-1} for DSP/PSP, increasing in absorbance and shift to 1258 cm^{-1} in Non-SP; (10) increased absorbance of $\nu_{\text{as}}\text{P}=\text{O}$ (of PO_2^-) at 1236 cm^{-1} in DSP relative to PSP/Non-SP at 1234 cm^{-1} ; (11) band at 1169 cm^{-1} shifting to 1173 cm^{-1} from DSP to Non-SP, band most prominent in PSP; (12) possible RNA band at 1115 cm^{-1} shifting to 1119 cm^{-1} from DSP to Non-SP; (13) $\nu_{\text{s}}\text{P}=\text{O}$ (of PO_2^-) AT 1085 cm^{-1} in DSP/PSP shifted to 1087 cm^{-1} in Non-SP; (14) glycogen band at 1030 cm^{-1} for DSP/PSP shifted to 1026 cm^{-1} with an increase in absorbance for DSP.

In the assessment into the effect of Hoechst 33342 on the cell spectra, Gault *et al.*²⁴ observed a small shift of 4 cm^{-1} in the antisymmetric $\nu_{\text{as}}\text{P}=\text{O}$ phosphodiester band in mean spectra between stained and unstained cells. For the same band we found spectral shifts of no greater than 2 cm^{-1} between all cell types. The major difference observed, however, was the extent of absorbance: the DSP has a higher $\nu_{\text{as}}\text{P}=\text{O}$ absorbance band relative to the PSP/Non-SP. Gault *et al.* also observed that the $\nu_{\text{s}}\text{P}=\text{O}$ at 1086 cm^{-1} remained unchanged.²⁴ We found a 2 cm^{-1} shift and an increased absorbance for Non-SP relative to the SP

subtypes. Gault *et al.* found no striking changes in the lipid and protein signals between stained and unstained cells. We found in our complete analysis that differences in lipid signals in the wavenumber range studied highly contributed to chemometric separation.

To confirm the residing differences between the three cell types, principal component–linear discriminant analysis (PC–LDA) was performed. LDA optimises inter-group separation whilst minimising intra-group separation.²⁵ Cross-validation may be performed as is highly recommended with a limited-size dataset. The model was validated by using a training data set where the algorithm is supervised with *a priori* knowledge of each data type, followed by the use of a test data set, introduced to the model as an independent validation. It is common practice to input data of reduced dimensionality into the model and therefore the principal components from PCA are invoked. Reduced data are used to overcome issues of over-fitting whereby good classification but poor significance can result. It is reported that over-fitting can be seen when using $(n - g)/3$ components or more, where g is the number of groups and n is the total number of data points.²⁶ For the LDA classifier variables, the principal components were calculated in PCA for the DSP, PSP and Non-SP dataset ($g = 3$, $n = 75$). A total of 8 principal components were used, a number securely far from the over-fitting threshold (24 PCs using the equation ' $(n - g)/3$ '). Eight PCs accounted for 71% of the total variance. It is generally recommended not to use PCs that explain 100% of variance as this will introduce PCs calculated on noise and will cause over-fitting of the LDA data. A training set comprising randomly selected data points from the total data set was constructed. Fifteen spectra per cell type were used for the training set (60%).

Fig. 7A displays the PC–LDA plot for the training set consisting of 45 spectra in total. Relative to the original PCA, the clustering and separation between cell types has greatly improved. Results show that the Non-SP cells are clearly differentiated from the side population cell types in discriminant function 1 (DF1). The proximal 'PSP' cells cluster relatively closer to the Non-SP cells than the distal DSP cells. This is accountable to the fact that the PSP cells share more chemical similarities to the Non-SP than the DSP cells share; the PSP cells are on the frontier of the SP cell gate in the FACS profile. The

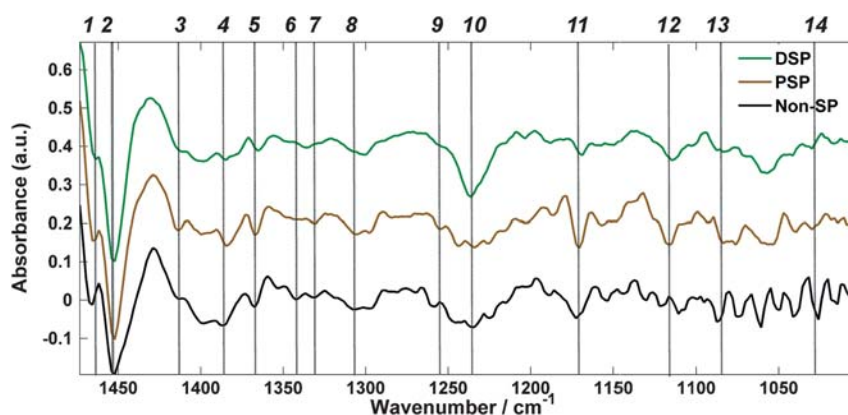


Fig. 6 Qualitative evaluation of the mean second derivative spectrum for DSP, PSP and Non-SP. Differences in spectral shifts are minimal whereas absorbance differences are most distinctive.

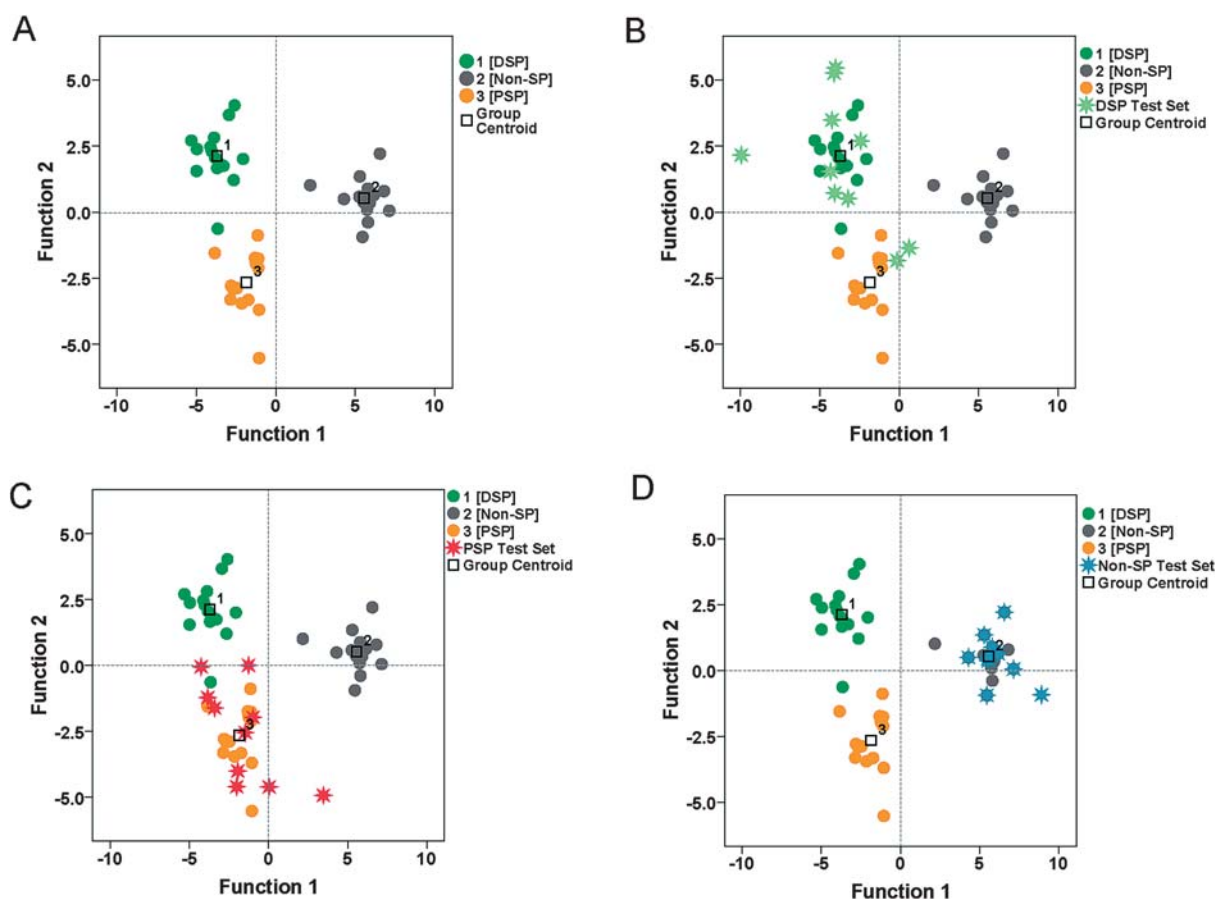


Fig. 7 PC-LDA of DSP and PSP *versus* Non-SP spectra. (A) The training set, $n = 45$. (B) DSP test set. (C) Non-SP test set. (D) PSP test set.

distal SP cells, however, are the most different, separating from the Non-SP in DF1 and even the PSP in DF2. It has been shown previously that the renal SP from human primary cells is a heterogeneous population¹ and that it is possible to sub-fractionate the populations based on the dye efflux status of the cells to generate sub-populations which are increasingly primitive. Here we confirm spectroscopically that populations of cells with differing dye efflux ability have distinct chemical profiles which can be used to further characterise or be utilised as a spectroscopic markers for these cells.

In terms of model classification 93.3% of training spectra correctly classified. Of the total dataset only 3 spectra were misclassified. 6.7% of DSP and 13.3% of Non-SP were misclassified as PSP. In cross-validation using the 'leave-one-out' method, 84.4% of spectra correctly classified. 13.3% of DSP and 20% of Non-SP were misclassified as PSP. 13.3% of PSP were misclassified as DSP.

Fig. 7B, C and D display the same PC-LDA plot but also include the independent validation DSP, PSP and Non-SP datasets respectively. Ten spectra per cell type were used to test the algorithm. The correct classification scores of the independent validation DSP, PSP and Non-SP dataset are 80%, 90% and 100% respectively. In general, the resulting plots and relative scores of each independent validation set reveal that the spectra have been correctly assigned. The similarity in the phenotype of the sub-side populations is a probable explanation where a small percentage of data points have been misclassified.

To get a clearer picture of the distinction of the side population and the non-side population, PC analysis was performed using only the DSP *versus* the Non-SP. The DSP are sorted by their Hoechst profile, which is the greatest distance from the non-side population: therefore their chemistry should be the most distinctive. The PCA score plot (Fig. 8A) of PC1 and PC2 which accounts for 40% of total variance. The distal side population clearly tightly clusters separately from the non-side population. The clustering of the DSP suggests that the cells within the DSP are more chemically/spectrally homogeneous than the Non-SP which appears to form two clusters, separated in PC2. This variation observed in the Non-SP is thought to be due to the actively proliferating cells within this population, which will be at varying differentiation stages. In contrast the DSP, hypothesised to contain the most primitive stem-like cells, is much more likely to be quiescent and more clonal in nature leading to a more uniform spectral profile. The noteworthy differences between the distal and non-side populations are shown in the PC1 loading plot (Fig. 8B). There are significant phosphodiester and carbohydrate signals but are secondary to the biochemical lipids signatures that dominate the loading plot at $\delta_{\text{as}}(\text{CH}_2)$ ~ 1468 cm^{-1} and $\delta_{\text{s}}(\text{CH}_3)$ at ~ 1380 cm^{-1} .

Carbohydrate bands are highly loaded in PC2 where the Non-SP separate into two clusters. Coupled with the symmetric and antisymmetric phosphodiester stretching bands this suggests differences in metabolic activity, probably due to cell cycle effects. The glycogen : phosphate ratio at 1030/1080 cm^{-1} ,

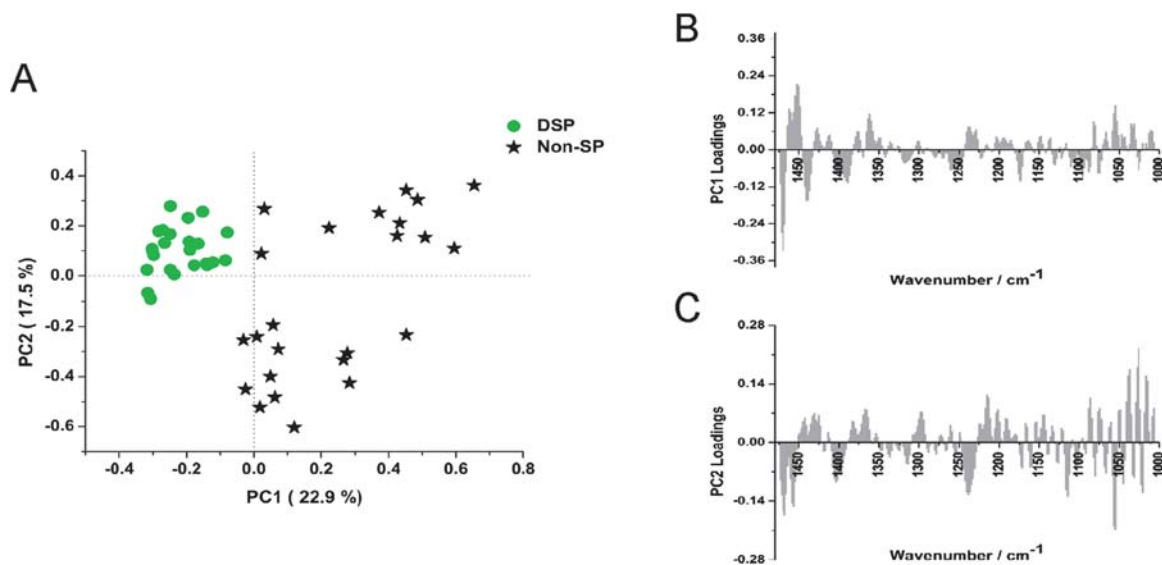


Fig. 8 DSP versus Non-SPs: (A) the PCA plot of PC1 and PC2, and (B) the corresponding loadings.

a known marker of metabolic activity,²⁷ deduced by spectral peak heights, is relatively smaller in the Non-SP cluster. This is thought to correspond to increased metabolic activity, as glycogen stores are broken down at a greater rate. The PC2 loading plot displays the re-occurring symmetric and antisymmetric phosphodiester band features also seen in Fig. 4B and Fig. 5C.

For a final evaluation, the differences in cell chemistry between the DSP and PSP cells were analysed (Fig. 9). Clear separation can be seen between the cell types in PC1 shown in the score plot. In the comparison of DSP and Non-SP spectra, the main differences were highlighted by lipid chemistry. Also highlighted are the phosphodiester stretching bands as a significant discriminatory feature. A possible explanation as to why the $\nu_{\text{as}}\text{P}=\text{O}$ (of PO_2^-) is most apparent in this comparison is that the DSP and PSP cells have greater similarity in terms of their proliferative and morphological phenotype. When possible differences due to cell cycle, or overall cellular heterogeneity are

removed, it is apparent that the phosphodiester stretching bands may be a prominent marker in the distinction of these stem cell-like cells. Differences in the symmetric and asymmetric phosphodiester stretching bands continually re-occur throughout the analysis and are always anticorrelated in the loadings plots. This may be indicative of reported DNA conformational changes where the symmetric phosphodiester modifications at ~ 1080 cm^{-1} were detected as a significant marker of putative stem cells in human intestinal crypts.²⁸ The symmetric and antisymmetric phosphodiester stretching vibrations refer to the central phosphorus atom in a phosphodiester linkage ($-\text{O}-\text{PO}_2-\text{O}-$), seen in DNA, RNA and some phospholipids. In DNA, the negatively charged phosphorus atom is countered by Na^+ ions.²⁹

A recent study³⁰ of other stem-like cell line populations supports this discriminatory band along with a carbohydrate marker at 1054 cm^{-1} . However, rather than looking for a univocal specific marker of stem-like cell discrimination, the authors propose using a change in a global absorption pattern as a better

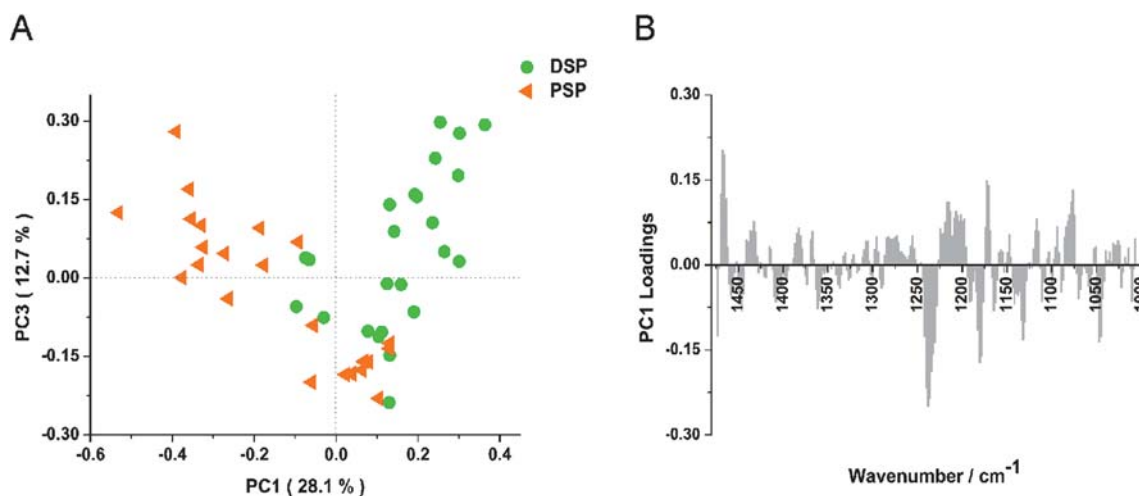


Fig. 9 (A) The PCA score plot of distal (DSP) versus proximal (PSP) second derivative spectra. Sub-side population discrimination is shown in the first principal component (28.1%). (B) The loading plot highlights the $\nu_{\text{as}}\text{P}=\text{O}$ stretch a distinct marker, which has also been present in the previous loadings comparing NSP to the SPs.

approach, particularly monitoring the carbohydrate-rich vibrations at 1000–1200 cm^{-1} , which they report to be a particular stem cell-like discriminator range. Yet spectral changes in carbohydrate bands may be more indicative of cell cycle effects. This would contribute greatly to chemometric discrimination and therefore it is important to bear this in mind when interpreting cell characterisation. The later research characterises the stem-like cells during a live cell suspension. Throughout the time of data acquisition there are a range of variables that may affect the absorption pattern of a single cell. For instance, cell viability may be reduced and dying cells have previously been reported to have different spectral features²² to viable cells. The spectra may be recording the independent cellular response to the non-ideal conditions of their environment for the respective cell types. Also, considerations would have to be taken into account for the cell suspension fluid and the contributions to the spectra. However, if these factors can be monitored and perhaps controlled, a live cell approach would be a very useful application when monitoring side population cells after subsequent passage, or their chemical response to anticancerous agents for instance.

PCA and LDA was used as part of a data-reduction analysis to elucidate chemical differences between the SP and Non-SP cell types. All chemometric analysis was performed on the second derivative corrected spectra with a range of 1475–1000 cm^{-1} to achieve maximum separation between the Non-SP and the SP cells. The full fingerprint region including the amide I and II bands reduced the clarity of separation (Fig. 10). For example, use of the full fingerprint range 1800–1000 cm^{-1} produced a poorer discrimination in PCA in scores for DSP/PSP *versus* Non-SP (Fig. 4). The percentages explained by PC1, PC2 and PC3 were 19.8, 13.6 and 11.2% respectively for range (44.6% of

total variance explained in 3 PCs). If the full fingerprint range was used, the values decreased (15.1, 10.9 and 9.9% respectively, 35.9% in total). Fig. 10A is the score plot for PC1 *versus* PC2 for the full fingerprint range where PC1 displays no separation of the different subtypes. In the loadings plot the cause of the spread of scores is predominantly due to the amide region (Fig. 10B). PC2 (Fig. 10C,D) displays some separation and the PC2 loadings show that distinct phosphate and lipid signals are present, along with the amide I and II which still appear to worsen the separation (relative to omitting the amide region). In PC3 (Fig. 10E) again there is no separation between the SP and the Non-SP subtypes; however, the Non-SP clusters into two groups. In the PC3 loadings it is the amide I and II regions are again largely contributing, along with the carbohydrate absorptions, as noted in the comparison of DSP and Non-SP (Fig. 8).

Conclusion

It has been successfully demonstrated that it is possible to isolate a side population from a human renal cancer cell line as well as from solid tissue. However these 'side population' cells only account for a small proportion of the total cell population (typically <5%), creating difficulties for analysis, which increases with the further sub-fractionation required for increasing primitive cell isolation. It is therefore difficult to characterise cell populations of this type by conventional biological methods.

High brilliance synchrotron-FTIR spectroscopy makes it possible to record the spectral fingerprint of these different cell types. PC and PC-LD analysis confirmed the biochemical differences of these cell types, providing further support that the side population contains a different sub-population of cells and

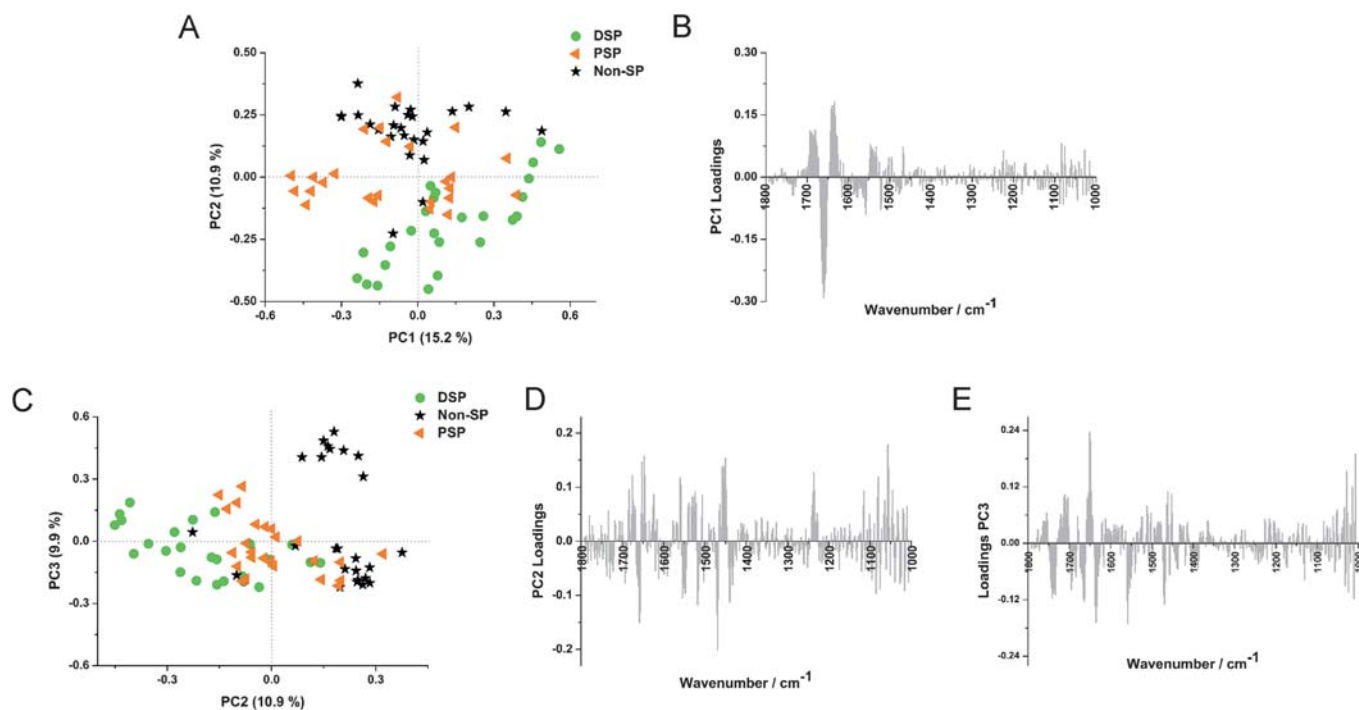


Fig. 10 The PCA analysis of extended spectral range (1800–1000 cm^{-1}) resulted in poorer cell type separation of distal (DSP) *versus* proximal (PSP) and Non-SP cells using second derivative spectra. (A) Score plot of PC1/PC2 and (C) PC2/PC3. Corresponding loading plots of (B) PC1, (D) PC2 and (E) PC3. Prevalent phosphodiester and lipid peaks evaluated previously are apparent in PC2 whereas PC1 and PC3 denote intra-cellular heterogeneity.

that this side population can be further sub-fractionated to give distinct populations based on their cellular biochemistry.

Interpreting the cause of discrimination of the cell types revealed variances within lipid and phosphodiester signals, whereas carbohydrate and protein signals in the loadings plots suggested heterogeneity within the Non-SP cells. It is perhaps the phosphodiester signals that are most revealing, as they may represent conformational changes in the DNA/RNA (as well as some phospholipids). These are indicated in the loadings plots in Fig. 4B, 5C, 8C and most prominently in Fig. 9B. The anti-symmetric phosphodiester vibrations at $\sim 1237\text{ cm}^{-1}$ are partially overlapped by complex amide III vibrations, whilst the symmetric phosphodiester vibration at $\sim 1080\text{ cm}^{-1}$ can be masked by carbohydrate vibrations. However, with the use of the second derivative combined with chemometric data reduction techniques, the phosphodiester bands are distinguishable and seem to be a significant cause of variance in the biochemical fingerprints of these side population cells relative to the general cell population. Another key factor of cellular distinction in the loadings (Fig. 4B, 5B, 8B, 9B, 10D) and in the qualitative mean spectral analysis (Fig. 6) were shifts and a higher absorbance for Non-SP in $\delta_s(\text{CH}_3)$ and $\delta_{\text{as}}(\text{CH}_2)$ and a higher absorbance in for DSP/PSP at 1452 cm^{-1} (C–H bending).

At this point we can only speculate about the results of the differential chemistry surrounding these SP cells which seem to pinpoint possible DNA and lipid differences. However, there may be a link between these features and it is known that DNA and phospholipids share a common motif recognizable by proteins.³¹ Also, it is generally accepted that lipids play both functional and structural roles in the nucleus as well as in the plasma membrane and that lipid metabolism in the nucleus may be regulated separately from that of the plasma membrane³² and perhaps this is a possible route of further investigation into DSP evaluation.

In conclusion, this novel approach offers considerable potential as a tool to distinguish small cell numbers associated with potential cancer stem cells and the SP phenotype and may be of particular use in facilitating the understanding of cancer stem cell biology.

Acknowledgements

We acknowledge the EPSEC-RSC Analytical Science Studentship scheme and also the Paterson Institute of Cancer Research for financial and technical support for C.H. We acknowledge the EU for funding travel to SOLEIL and thank all staff at SOLEIL associated with the SMIS beamline.

References

- 1 S. K. Addla, M. D. Brown, C. A. Hart, V. A. C. Ramani and N. W. Clarke, *Am. J. Physiol.: Renal Physiol.*, 2008, **295**, F680–F687.
- 2 M. Al-Hajj, M. S. Wicha, A. Benito-Hernandez, S. J. Morrison and M. F. Clarke, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6890.
- 3 M. Al-Hajj, M. S. Wicha, A. Benito-Hernandez, S. J. Morrison and M. F. Clarke, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 3983–3988.
- 4 R. I. Bhatt, M. D. Brown, C. A. Hart, P. Gilmore, V. A. C. Ramani, N. J. George and N. W. Clarke, *Cytometry*, 2003, **54a**, 89–99.
- 5 M. D. Brown, P. E. Gilmore, C. A. Hart, J. D. Samuel, V. A. C. Ramani, N. J. George and N. W. Clarke, *Prostate*, 2007, **67**, 1384–1396.
- 6 M. Florek, M. Haase, A. M. Marzesco, D. Freund, G. Ehninger, W. B. Huttner and D. Corbeil, *Cell Tissue Res.*, 2005, **319**, 15–26.
- 7 B. R. Grey, J. E. Oates, M. D. Brown and N. W. Clarke, *BJU Int.*, 2009, **103**, 856–858.
- 8 C. Li, D. G. Heidt, P. Dalerba, C. F. Burant, L. Zhang, V. Adsay, M. Wicha, M. F. Clarke and D. M. Simeone, *Cancer Res.*, 2007, **67**, 1030–1037.
- 9 L. Ricci-Vitiani, D. G. Lombardi, E. Pilozzi, M. Biffoni, M. Todaro, C. Peschle and R. De Maria, *Nature*, 2007, **445**, 111–115.
- 10 S. K. Singh, C. Hawkins, I. D. Clarke, J. A. Squire, J. Bayani, T. Hide, R. M. Henkelman, M. D. Cusimano and P. B. Dirks, *Nature*, 2004, **432**, 396–401.
- 11 S. Yin, J. Li, C. Hu, X. Chen, M. Yao, M. Yan, G. Jiang, C. Ge, H. Xie, D. Wan, S. Yang, S. Zheng and J. Gu, *Int. J. Cancer*, 2007, **120**, 1444–1450.
- 12 J. E. Oates, B. R. Grey, S. K. Addla, J. D. Samuel, C. A. Hart, V. A. C. Ramani, M. D. Brown and N. W. Clarke, *Stem Cells Develop.*, 2009, **18**, 1515–1521.
- 13 M. A. Goodell, K. Brose, G. Paradis, A. S. Conner and R. C. Mulligan, *J. Exp. Med.*, 1996, **183**, 1797–1806.
- 14 M. Brown, B. Grey, J. Oates and N. W. Clarke, *Cancer Stem Cells*, 2010, 107–129.
- 15 P. Heraud and M. J. Tobin, *Stem Cell Res.*, 2009, **3**, 12–14.
- 16 P. Dumas, F. Polack, O. Lagarde, O. Chubar, J. L. Giorgetta and S. Lefrancois, *Infrared Phys. Technol.*, 2006, **49**, 152–160.
- 17 P. Bassan, H. J. Byrne, F. Bonnier, J. Lee, P. Dumas and P. Gardner, *Analyst*, 2009, **134**, 1586–1593.
- 18 P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, P. Dumas, M. Brown, N. Clarke and P. Gardner, *J. Biophotonics*, 2010, **3**, 609–620.
- 19 P. Bassan, H. J. Byrne, J. Lee, F. Bonnier, C. Clarke, P. Dumas, E. Gazi, M. D. Brown, N. W. Clarke and P. Gardner, *Analyst*, 2009, **134**, 1171–1175.
- 20 P. Bassan, A. Kohler, H. Martens, J. Lee, H. J. Byrne, P. Dumas, E. Gazi, M. Brown, N. Clarke and P. Gardner, *Analyst*, 2010, **135**, 268–277.
- 21 E. Goormaghtigh, A. Derenne, A. Bacnard, R. Gasper and V. Raussens, *Spectroscopy*, **24**, 51–54.
- 22 H. Y. N. Holman, M. C. Martin, E. A. Blakely, K. Bjornstad and W. R. McKinney, *Biopolymers*, 2000, **57**, 329–335.
- 23 R. J. Swain and M. M. Stevens, *Biochem. Soc. Trans.*, 2007, **35**, 544–549.
- 24 N. Gault, O. Rigaud, J.-L. Poncey and J.-L. Lefaix, *Radiat. Res.*, 2007, **167**, 551–562.
- 25 T. Fearn, *Discriminate Analysis, Handbook of Vibrational Spectroscopy*, John Wiley & Sons, 2002, vol. 3, pp. 2086–2093.
- 26 M. Defernez and E. K. Kemsley, *TrAC, Trends Anal. Chem.*, 1997, **16**, 216–221.
- 27 E. Gazi, J. Dwyer, P. Gardner, A. Ghanbari-Siahkali, A. P. Wade, J. Miyan, N. P. Lockyer, J. C. Vickermann, N. W. Clarke, J. H. Shanks, L. J. Scott, C. A. Hart and M. Brown, *J. Pathol.*, 2003, **201**, 99–108.
- 28 M. J. Walsh, T. G. Fellous, A. Hammiche, W. R. Lin, N. J. Fullwood, O. Grude, F. Bahrami, J. M. Nicholson, M. Cotte, J. Susini, H. M. Pollock, M. Brittan, P. L. Martin-Hirsch, M. R. Alison and F. L. Martin, *Stem Cells*, 2008, **26**, 108–118.
- 29 R. Salzer and H. W. Siesler, *Infrared and Raman Spectroscopic Imaging*, Wiley-VCH, 2009, p. 183.
- 30 R. Zhao, L. Quaroni and A. G. Casson, *Analyst*, **135**, 53–61.
- 31 K. Sekimizu, *Chem. Phys. Lipids*, 1994, **73**, 223–230.
- 32 K. Tamiya-Koizumi, *J. Biochem.*, 2002, **132**, 13–22.