



Label-free Raman imaging of live osteosarcoma cells with multivariate analysis

Jie Li^{1,2} · Jie Qin³ · Xu Zhang^{1,2} · Rui Wang³ · Zhuowen Liang⁴ · Qingli He² · Zhe Wang⁴ · Kaige Wang^{1,2} · Shuang Wang¹ 

Received: 26 March 2019 / Revised: 15 May 2019 / Accepted: 28 May 2019 / Published online: 22 June 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Confocal Raman microspectral imaging (CRMI) is an advanced cell-imaging method that maps endogenous molecular compositions with their unique spectral fingerprint indicators. The aim of this work was to provide a visualized understanding of subcellular features of live osteosarcoma cells using a 532-nm laser excitation without the use of dyes or molecular probes. Both malignant osteoblast and spindle osteosarcoma cells derived from the BALB/c mouse osteosarcoma cell line K7M2 were investigated in this work. After preprocessing the obtained spectral dataset, K-means cluster analysis (KCA) is employed to reconstruct Raman spectroscopic maps of single biological cells by identifying regions of the cellular membrane, cytoplasm, organelles, and nucleus with their corresponding mean spectra. Principal component analysis (PCA) was further employed to indicate variables of significant influence on the separation of the spectra of each cellular component. The biochemical components of the two cell types were then extracted by showing the spectral and distribution features attributed to proteins, lipids, and DNA. Using this standardized CRMI technique and multivariate analysis approaches, the results obtained could be a sound foundation for a typical Raman imaging protocol of live cellular biomedical analysis.

Keywords Confocal Raman microspectral imaging · Live osteosarcoma cells · Multi-variate analysis · K-means cluster analysis · Principal component analysis

Introduction

Raman spectroscopy contains information on molecular vibrations providing a highly specific fingerprint of the molecular structure and biochemical composition of cells and tissues without any external labels. As one of the most important areas of progress in single-cell Raman spectroscopy, confocal Raman microspectral imaging (CRMI) enables a high-

contrast and high-resolution visualization of intracellular components (Dieing and Hollricher 2008; Dieing et al. 2011). By sequentially obtaining the full Raman spectra point by point, a Raman map or image was then computationally constructed to produce a pseudocolored image according to the relative intensity of Raman spectra at a given wavelength at each pixel, called univariate Raman imaging. Despite the flexibility of the principle, univariate Raman imaging often presents an inherent limitation due to a severely overlapped spectral information. Rather than relying on single band intensities, diverse multivariate algorithms, such as K-means cluster analysis (KCA), principal component analysis (PCA), hierarchical cluster analysis (HCA), and vertex component analysis (VCA), are used to extract the maximum chemical information from the obtained whole spectral dataset without a priori knowledge regarding spectral characteristics (Miljkovic et al. 2010). These methods were applied to reduce the data dimension to elaborate spectral information, and the dataset was partitioned to highlight group differences.

Owing to its high sensitivity and selectivity, CRMI has been recognized as a powerful tool and has been widely used

Jie Li and Jie Qin contributed equally to this work.

✉ Shuang Wang
swang@nwu.edu.cn; wsnwuphy@163.com

¹ Institute of Photonics and Photon-Technology, Northwest University, #229 North Taibai Road, Xi'an 710069, Shaanxi, China

² Department of Physics, Northwest University, Xi'an 710069, Shaanxi, China

³ Department of Orthopedics, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710004, Shaanxi, China

⁴ Department of Orthopaedics, Xijing Hospital, Air Force Medical University, Xi'an 710032, Shaanxi, China

for addressing various biochemical questions related to cell biology (Ashton et al. 2016; Draux et al. 2009; Manen et al. 2005; Nithipatikom et al. 2003). However, the technique and its multivariate analysis approaches were barely employed to reveal the biochemical constitution of osteosarcoma cells by differentiating among organelles, cell membrane, and nucleus. Recently, Pei-san et al. (2013) demonstrated that Raman spectroscopy is an effective biosensor to monitor the production of different mineralized matrices during osteogenic differentiation of mesenchymal stem cells. They further measured hydroxyapatite production in osteosarcoma cells for a rapid and accurate assessment of the malignancy degree in a label-free manner (Chiang et al. 2015). In 2003, Krafft et al. (2003) explored the possibility of using Raman spectroscopy to investigate the chemical composition and molecular structure of human osteogenic sarcoma cells. Slaman et al. demonstrated that Raman spectroscopy in tandem with PCA methods followed by LDA calculations is a useful tool for identifying and differentiating between normal and malignant cells transformed by “murine sarcoma virus” (Salman et al. 2013). However, to the best of our knowledge, the main cellular constituents and biochemical structure of osteosarcoma cells were never evaluated by Raman spectroscopy imaging, which visualizes cells with subcellular resolution based on differences in the spectral features of the various cellular components.

The aim of this work was to provide a direct understanding of subcellular features of live osteosarcoma cells without any fixation and tagged dyes. Malignant osteoblast and spindle osteosarcoma cells, both derived from the BALB/c mouse osteosarcoma cell line K7M2, were cultured in the appropriate medium and seeded on CaF₂ substrate for microspectroscopic analysis. Multivariate analysis algorithms, including KCA and PCA, were performed to obtain the morphology and compositional information of different subcellular structures such as the nucleus, organelles, and cell membrane. By providing a comparison of the efficiency of two multivariate analyses in separating sarcoma cells, the achievements in this work are conducive to establish robust and reliable protocols that are applicable to cell biology.

Experimental

Sample preparation

BALB/c mouse osteosarcoma cell line K7M2 was purchased from the American Tissue Culture Collection (ATCC, Manassas, USA). The frozen K7M2 cells were quickly shaken in a preheated 37 °C water bath and completely thawed in 3 min. The cell suspension was transferred into a 15-mL centrifuge tube, 15 mL Dulbecco's Modified Eagle's Medium (DMEM, Gibco, Waltham, MA, US) was added and mixed

well, and the tube was centrifuged at 1200 rpm for 5 min. The supernatant was discarded; an appropriate amount of DMEM containing 5% fetal bovine serum (FBS, Gibco, Waltham, MA, US) was added. K7M2 cells were gently pipetted into a single cell suspension for counting. Next, cells ($2\sim 4 \times 10^5/\text{cm}^2$) were seeded on a 2.5-cm diameter CaF₂ substrate in a 6-cm diameter cell culture dish. The cells were cultured in an incubator at 37 °C with 5% CO₂. The medium was changed every 3 days. For maintenance of cell line, the ATCC-dictated culture conditions for these cells were met. Cells at passage 3rd to 15th were used for CRMI.

Confocal Raman microspectroscopy

Raman spectra were acquired by an Alpha 500R confocal Raman microscopy system (WITec GmbH, Ulm, Germany) coupled with a 532-nm solid-state laser. The excitation laser beam was collimated into a $\times 63$ water-immersion objective lens (NA = 1, W Plan-APOCHROMAT, Zeiss, Jena, Germany) for spectral measurements. The CaF₂ slide, where cells seeded, was placed on a multi-axis piezo scanning stage (P-524 K081, PI GmbH, Karlsruhe, Germany) for spectral imaging. The spectral signal was recorded by a spectrometer (UHTS300, WITec GmbH, Ulm, Germany) incorporating a 600-mm⁻¹ grating blazed at 500 nm, with a back-illuminated deep-depletion charge-coupled device camera (CCD, Du401A-BR-DD-352, Andor Technology, Belfast, UK) working at -60 °C. Before the experiment, white-light images were captured by using the same objective lens. The wavelength was calibrated by 521-cm⁻¹ silicon as a standard and the spectral response was calibrated by a standard tungsten lamp (RS-3, EG&G Gamma Scientific, San Diego, CA, USA).

Data processing

Pretreatment and multi-variate analysis of the obtained spectra dataset were both performed by using the WITec Project (WITec GmbH, Ulm, Germany) data analysis software. Spectra preprocessing was performed according to our previous protocols (Rutan 2004). Briefly, after cosmic ray removal, all Raman spectra were cropped in the fingerprint region (600–1800 cm⁻¹ and 2800–3100 cm⁻¹) and their background was subtracted with fifth-order polynomial fitting. Spectra were then smoothed using the 10 points Savitzky-Golay (SG) smoothing filter. Then, a 4-D spectral dataset, which consists of the two directions of the point scanning plan (x and y), wavenumber, and spectrum intensity, were ready to be further processed by multi-variate analyses for providing a complete and detailed spectral profile within the scanning area.

KCA, as one type of multivariate analysis algorithm, is an unsupervised method by which the preprocessed Raman dataset was classed into k clusters, such that each spectrum belongs to the cluster with the nearest (closest) mean

(Macqueen 1965). During KCA evaluation, the operator needs to speculate an initial k number of clusters in the dataset as starting centroids, which denoted the center or mean of the cluster. Subsequently, the distances between each spectrum and these centroids are automatically calculated, and each spectrum is assigned to a cluster whose centroid is nearest to the spectrum itself. Once all spectra are sorted into k clusters, the method can also be further applied to form sub-clusters, thus generating a pseudo-hierarchical cluster tree, in which the cluster membership information was plotted as a color-coded image for providing an interpretation on morphological characteristics of the cells. Besides that, the average spectrum of each cluster was adopted to detect variations in the Raman band positions and/or intensities between the different regions for illustrating the biochemical constitution of the cell sample.

PCA is the underlying method for many other multivariate methods, since it is very effective in reducing the dimensionality of the dataset (Baron et al. 2017; Chen et al. 2015). It resolves a dataset into a bilinear model of linear independent variables that are so-called principal components. A variety of results can be extracted following the PCA, such as reduced dataset, loading images, spectra of the principal components, and cross-correlation plots of various components. In the reduced dataset, high loadings are, therefore, indicating variables which are of significant influence on the separation of the spectra from each cellular substance. Each of the spectra used for the analysis can be described by a weighted combination of the principal components. By considering only one component, it will result in one value per image pixel and display this as loading image in a scaled color, which describes the abundance of a certain spectrum (vector) within the dataset. The loading spectra for each component describe the extent to which the variables in the analysis (wavenumber) vary in the same (in case of variables with positive values) or in the opposite direction (variables with negative values) to the PCs (McCreynolds et al. 2017; Woolford et al. 2018; Meade et al. 2010). By plotting a point for each spectrum at a corresponding position in a coordinate system of the first principle component versus the second (or the third, fourth...), the correlation between the components could be visualized as a loading plot or a “score-score” plot.

Results

Spectrum analysis

The measured Raman spectra from malignant osteoblast and spindle osteosarcoma cell are shown in Fig. 1a. The background noise is well controlled by using CaF_2 substrates, so that some major vibrational bands could be identified at 659, 1450, 1578, 1657, and 2922 cm^{-1} in separated cell types. By taking the 1450-cm^{-1} peak as the reference, the achieved

signal-to-noise ratio (SNR) was estimated to be 2.27 in malignant osteoblast and 2.70 in spindle osteosarcoma cell. Figure 1b shows the mean Raman spectra from various locations in the measured cells, in which the standard deviation is marked in gray. Raman spectra of both the two cell types were similar with characterized bands around 659, 902, 1001, 1240, 1321, 1450, 1578, 1657, and 2922 cm^{-1} . Band assignments are summarized in Table 1 according to the Raman studies of pure proteins (Chan et al. 2009; Konorov et al. 2013; Chen et al. 2016), nucleic acids (Pully et al. 2015; Ye et al. 2012; Diem et al. 2013), lipids (Krafft et al. 2003; 24.Diem et al. 2013; Gelder et al. 2007; Hobro et al. 2016; Pascut et al. 2013), and osteosarcoma cells (Chiang et al. 2015; Krafft et al. 2003). Except for lipid content, larger intensity fluctuations for both measured cells are found at 659-, 747-, 1001-, 1240-, and 1657-cm^{-1} bands indicating higher protein content. Further, a slightly higher content in nucleic acids is indicated by the 1578-cm^{-1} Raman band (Short et al. 2005; Xu et al. 1998). Meanwhile, the main spectral differences between malignant osteoblast and spindle osteosarcoma cell are exhibited at 747-cm^{-1} Raman band and spectral deviation centered 1355 cm^{-1} , which are mainly induced by the content variations of tryptophan in cell membrane surfaces. (Hernández et al. 2010; Madzharova et al. 2010; She et al. 2001).

K-means analysis

A Raman imaging dataset from a malignant osteoblast was collected from a $40 \times 40\ \mu\text{m}^2$ area containing 50×50 pixels at a spatial resolution of $0.80\ \mu\text{m}$, while a Raman imaging dataset from a spindle osteosarcoma cell was collected from a $90 \times 40\ \mu\text{m}^2$ area containing 90×50 pixels at a spatial resolution of $1.00\ \mu\text{m}$, which were determined by the morphological characteristics of two types of cells. The integration time for each spectrum was shortened to 1 s; therefore, the whole spectral datasets for each cell type could be obtained around 42 and 75 min respectively with an acceptable spatial resolution.

Pseudo-hierarchical cluster trees for the examined cells were generated by classifying the obtained spectral datasets into five sub-clusters, clearly distinguishing the cell membrane (blue and red), regions that were associated with a high concentration of cell organelle (purple), the nucleus (green), and the cytoplasm (cyan), as shown in Fig. 2. White-light microscopic images of both the two cell types are shown in Fig. 2 a (a) and b (a) as reference images. Mean spectra of each sub-cluster were separated and displayed on the right of each cluster, in which the color of each individual spectrum corresponds to the color of the sub-cluster to which it belonged. Additionally, the transformation of the root spectral image into its corresponding microscopic image to reveal the morphological feature corresponding to the spectral image of the cells used is shown in Fig. 2 a (c) and b (c).

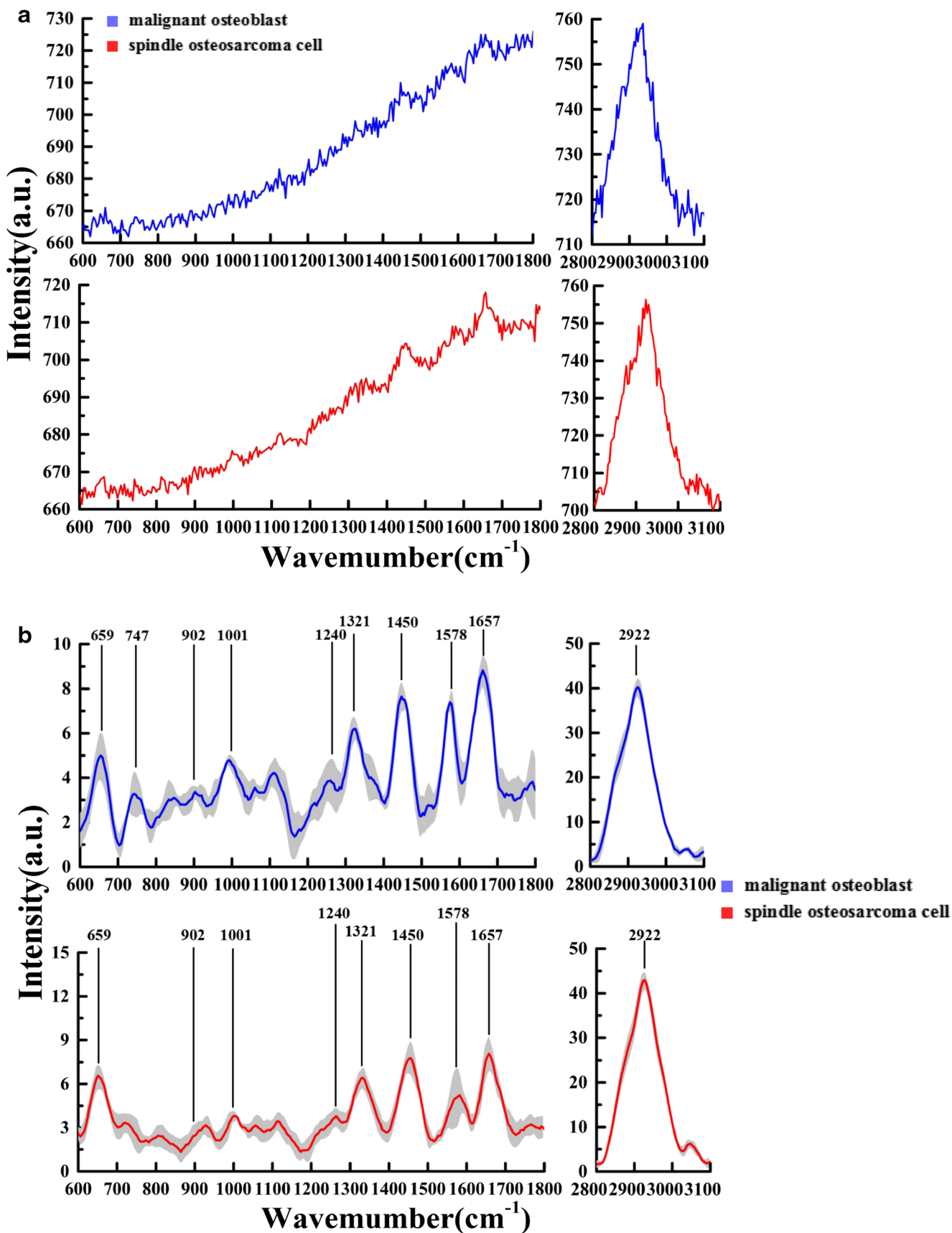


Fig. 1 **a** Measured Raman spectra of malignant osteoblast and spindle osteosarcoma cell for evaluating the quality of the raw data. **b** Processed mean spectra of both the two cell types, in which the standard deviation is marked in gray. All the spectra have been shifted vertically for clarity

Regarding the malignant osteoblast shown in Fig 2a, the blue and its inside red clusters displayed similar morphology and spectral features. Morphologically, both clusters had a triangular body with branched structures, besides an extensive configuration visualized in the blue cluster mentioned above. Spectrally, prominent lipid spectral features were at 902 cm^{-1} (Chan et al. 2009), 1450 cm^{-1} (Krafft et al. 2003; Notingher et al. 2002), 2893 cm^{-1} (Krafft et al. 2003), and 2922 cm^{-1} (Manciu et al. 2017); protein features at 659 cm^{-1} (Su et al. 2017), 747 cm^{-1} (Manciu et al. 2017; Prats et al. 2017; Hernández et al. 2013), 1240 cm^{-1} (Notingher 2007), 1321 cm^{-1} (Manciu et al. 2017; Notingher and Hench 2006), and 1676 cm^{-1} (Notingher and Hench 2006) with moderate intensity; and nucleic acid Raman peaks at 1321 cm^{-1} and 1578 cm^{-1} (Notingher and Hench 2006). Since one of the ubiquitous features of membrane proteins is the preference of tryptophan and tyrosine residues for membrane surfaces (Sun et al. 1998), the L -tryptophan 1355-cm^{-1} peak showed the highest intensity in the outermost blue cluster (Gelder et al. 2007). The mean spectrum of the red cluster exhibited higher spectral intensities than that

of the blue cluster, and has a distinct peak at 2922 cm^{-1} with a shoulder at 2893 cm^{-1} , which is associated with the C–H stretching vibration of phospholipids. In contrast, the spectrum of the blue cluster showed a maximum intensity at 2893 cm^{-1} with a lower peak at 2922 cm^{-1} , which was correlated with the C–H stretching mode of lipids (Hobro et al. 2016; Yu et al. 2006; Naumann 2001). An inconspicuous spectral variation was also present at approximately 2995 cm^{-1} (Brolo et al. 2010), which was also attributed to the lipid component in the cell membrane. The intensity variation of the 2922-cm^{-1} band from the cytoplasm, cell organelles, and the nucleus was more visible, the C–H stretching bands of the spectra corresponding to mitochondrion-rich regions and other regions in the cytoplasm exhibited enhanced intensities between 2850 and 2900 cm^{-1} , which arose from the alkane chains of lipids. All the three subtle but reproducible spectra exhibited distinct protein bands, such as the amide I vibration at 1657 cm^{-1} (Ye et al. 2012; Notingher et al. 2002), the C–H deformation at 1321 cm^{-1} , and the phenylalanine ring-stretching vibration at 1001 cm^{-1} (Pully et al. 2015; Ye et al. 2012), which is not surprising, since the major component in all the three regions was represented by proteins. At approximately 1578 cm^{-1} , the breathing modes of adenine and guanine bases contributed to both the spectra of the nucleus and cell organelles (Meister et al. 2010).

Table 1 Raman peak and tentative assignments of major vibrational bands. *str*, stretching; *sym*, symmetric; *asym*, asymmetric; *def*, deformation; *bk*, backbone; *Tyr*, tyrosine; *Trp*, tryptophan; *Phe*, phenylalanine; *Man*, Mannopyranose; *GlcNac*, *N*-acetylglucose; *GLuA*, glucuronic acid; *Glc*, D-glucose; *A*, adenine; *T*, thymine; *C*, cytosine; *G*, guanine

Raman shifts (cm^{-1})	Assignment			
	Proteins	Carbohydrates	Nucleic acids	Lipids
659	C–C twist, Tyr			
747	Trp sym		T	
902				CH_2
916		Man		
1001	Phe			
1025		GlcNac, Glc GLuA		
1097			bk (PO_2)	
1123	C–N, C–C str			C–C
1177	C–H bend Tyr			
1240	Amide III			
1321	CH def		G	
1355	L -tryptophan			
1375			A, G	
1450	CH_2 def			CH def
1533			G, C	
1578			G, A	
1657	Amide I, α -spiral			C=C str.
1676	Amide I			
2893				CH_2
2922				CH_2
2995				CH

As shown in Fig. 2b, each of the five main clusters in this figure could be assigned to certain cellular structures of the spindle osteosarcoma cell. The green cluster was assigned to the nucleus, the purple cluster to the cell organelles, the cyan cluster to the cytoplasm, and the blue and red clusters to the cell membrane. Each spectrum of the spindle osteosarcoma did not show much difference with the corresponding spectrum of the malignant osteoblast. However, the spectrum from the purple cluster showed two more clear bands assigned to the symmetrical breathing of tryptophan (Prats et al. 2017; Hernández et al. 2013) at 747 cm^{-1} and the C–C skeleton of acyl backbone in lipids at 1123 cm^{-1} (Notingher et al. 2002). Furthermore, another tryptophan Raman band could be identified at 1177 cm^{-1} (Notingher 2007; Notingher and Hench 2006) from the blue cluster due to the C–H vibration of the ring, indicating a different orientation of the ring caused by steric effects (Notingher 2007). Moreover, if we compare the spectra from the cell membrane between two types of cells, it could be found that the Raman band at 1355 cm^{-1} of L -tryptophan (Gelder et al. 2007) presents in the malignant osteoblast (in Fig. 2a (i)) which is not found in the spindle osteosarcoma cell (in Fig. 2b (i)). This is because of the thinner membrane structure that existed in spindle osteosarcoma cell than that in the osteoblast.

Principal component analysis

PCA Raman images were reconstructed for both the two investigated cells with their separated spectral loadings as a function of the wavenumber. Figure 3a shows the PCA image (Fig. 3a (a)) of a malignant osteoblast, which allowed the visualization of each part of the spectrum corresponding to each different structure of the cell, such as the cell membrane (Fig. 3a (b)), cytoplasm (Fig. 3a (c)), and nucleus (Fig. 3a (d)), allowing their discrimination. The loading spectra were offset for clarity, with the dotted line representing the zero level for each principal component (PC). The spectra showed that the loading was composed of different sharp and broad peaks, both positive and negative, corresponding to increased or decreased contributions of specific molecular components to the Raman spectra.

The loading image of PC1 in Fig. 3a (b) exhibited a good correlation with the accumulation of blue and red sub-clusters shown in Fig. 2a. Figure 3a (b) exhibited intense positive loadings for Raman shift characteristics of positive protein (1025 cm^{-1} , 1123 cm^{-1} , 1450 cm^{-1} , and 1676 cm^{-1}) and lipids (1123 cm^{-1} , 1450 cm^{-1} , 2893 cm^{-1}), and presented a wide negative band centered at 747 cm^{-1} , which can be attributed to the symmetrical breathing of tryptophan (Prats et al. 2017; Hernández et al. 2013). The loading plot also showed positive peaks at 1025 cm^{-1} (Vasko et al. 1971) and 916 cm^{-1} (Su et al. 2017), indicating the existence of glucose and its derivatives in the cell membrane with consideration of its

loading image. The loading image of PC2 in Fig. 3a (c) also exhibited a combined morphology of cyan and purple sub-clusters in Fig. 2a. While all average spectra exhibited a relatively flat curve from 600 to 1800 cm^{-1} , the three positive protein bands at 747 , 1321 , and 1657 cm^{-1} in Fig. 3a (c) were less evident, which allowed the differentiation between the cytoplasm and cell membranes. Nevertheless, two negative nucleic acid Raman bands around 1097 cm^{-1} (Krafft et al. 2003; Notingher and Hench 2006) and 1533 cm^{-1} (Xu et al. 1993) highlighted the discrimination between the cytoplasm and the third principal component of the nucleus. However, the strongest variation between them was the negative peak at 2893 cm^{-1} , which was the prominent one in the spectral loading of PC3, indicating significant variations in lipidic content, as well as the negative difference at 747 cm^{-1} indicating low contributions from tryptophan. In Fig. 3a (d), the average spectrum corresponding to PC3 contained positive peaks related to proteins (1001 cm^{-1} , 1240 cm^{-1} , 1676 cm^{-1}) and nucleic acids (1097 cm^{-1} , 1375 cm^{-1} (Pully et al. 2015; Hernández et al. 2013), 1533 cm^{-1}), and a significant negative peak at 2893 cm^{-1} related to lipids. Thus, from the Raman spectra, we concluded that the PC3 region was indicating the nucleus, where genomic DNA and RNA are located. A further supporting argument for DNA/RNA existence was the strong presence of the 1097-cm^{-1} peak in the same spectrum, a feature attributed to the O–P–O stretching mode of DNA/RNA (Krafft et al. 2003).

To further demonstrate the above results, PC1, PC2, and PC3 were calculated for the Raman image of the spindle osteosarcoma cell with their corresponding loading vectors. As in Fig. 3a, the loading plot displayed in Fig. 3b highlighted several details within the cell. In this case, the first loading revealed the constitution of the cell membrane, since the positive peak was related to proteins and lipids, as well as the negative peak at approximately 747 cm^{-1} that indicated the presentation of tryptophan. Interestingly, the PC1 loading image could not outline the overall morphology of the cell membrane as displayed in both Fig. 2 a (e, d) and b (e, d), because of the limited thickness and slight spectra variances at the side of the cell body. Furthermore, the differentiation between the cytoplasm and the nucleus by PC2 vs PC3 was reproducible according to their featured scores, indicating that the loadings were representative of the biochemical structure of osteosarcoma cells.

Discussion

The point-scanning CRMI technique was adopted to address the subcellular features in both malignant osteoblast and spindle osteosarcoma cells. Even though, the disadvantage of CRMI is the weak signal and point scanning methodology which lead to long measurement time. There are several ways

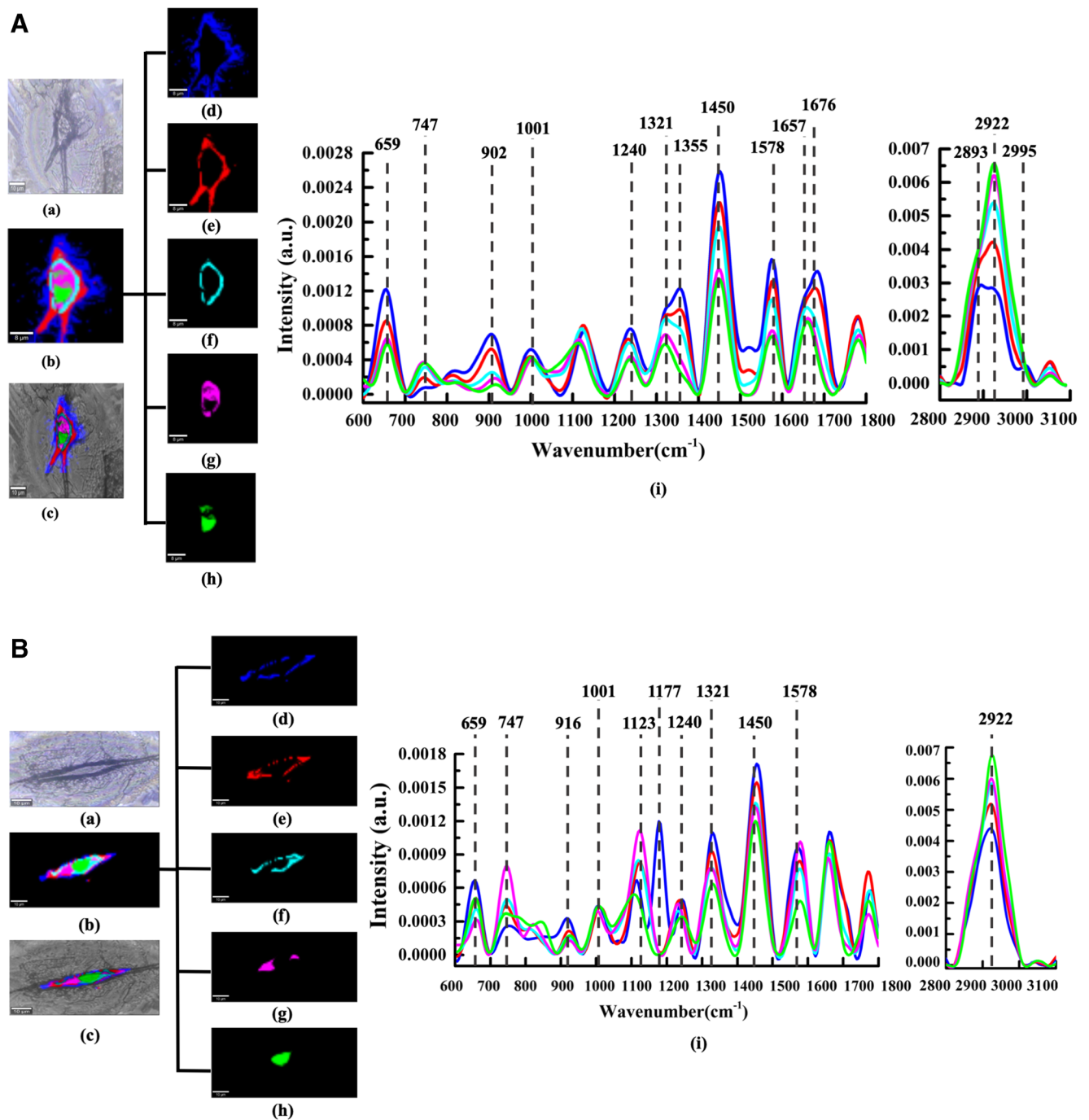


Fig. 2 K-means cluster analysis (KCA) of the acquired Raman datasets from malignant osteoblast (a) and spindle osteosarcoma cell (b). For each sub-figure, image (a) shows the white light micrograph of the studied cell; image (b) shows the roots of a pseudo-stratified clustering tree of malignant osteoblast and spindle osteosarcoma cell; image (c) displays

the spatially transformed image between (a) and (b). Sub-clusters for each group are listed in images (d) to (h) next to its corresponding clustering root. The mean spectrum of each sub-cluster, as shown in image (i), was plotted to match their classification with the corresponding colors

to reduce the measurement time. One is undersampling with a step size of 10 or 100 μm by using the FTIR imaging method (Mavarani et al. 2013). Another possibility is the use of non-linear variants of Raman spectroscopy, such as coherent anti-Stokes Raman spectroscopy and stimulated Raman spectroscopy, which facilitates video-rate, high-resolution images for

in vitro and in vivo cellular analysis (Antonio and Schultz 2014). However, a prohibitive factor in nonlinear Raman spectroscopy remains the cost and complexity associated with synchronization techniques of two pulsed lasers, which make it even more difficult to practice for nonspecialists. The obtained experimental results show that while the adopted

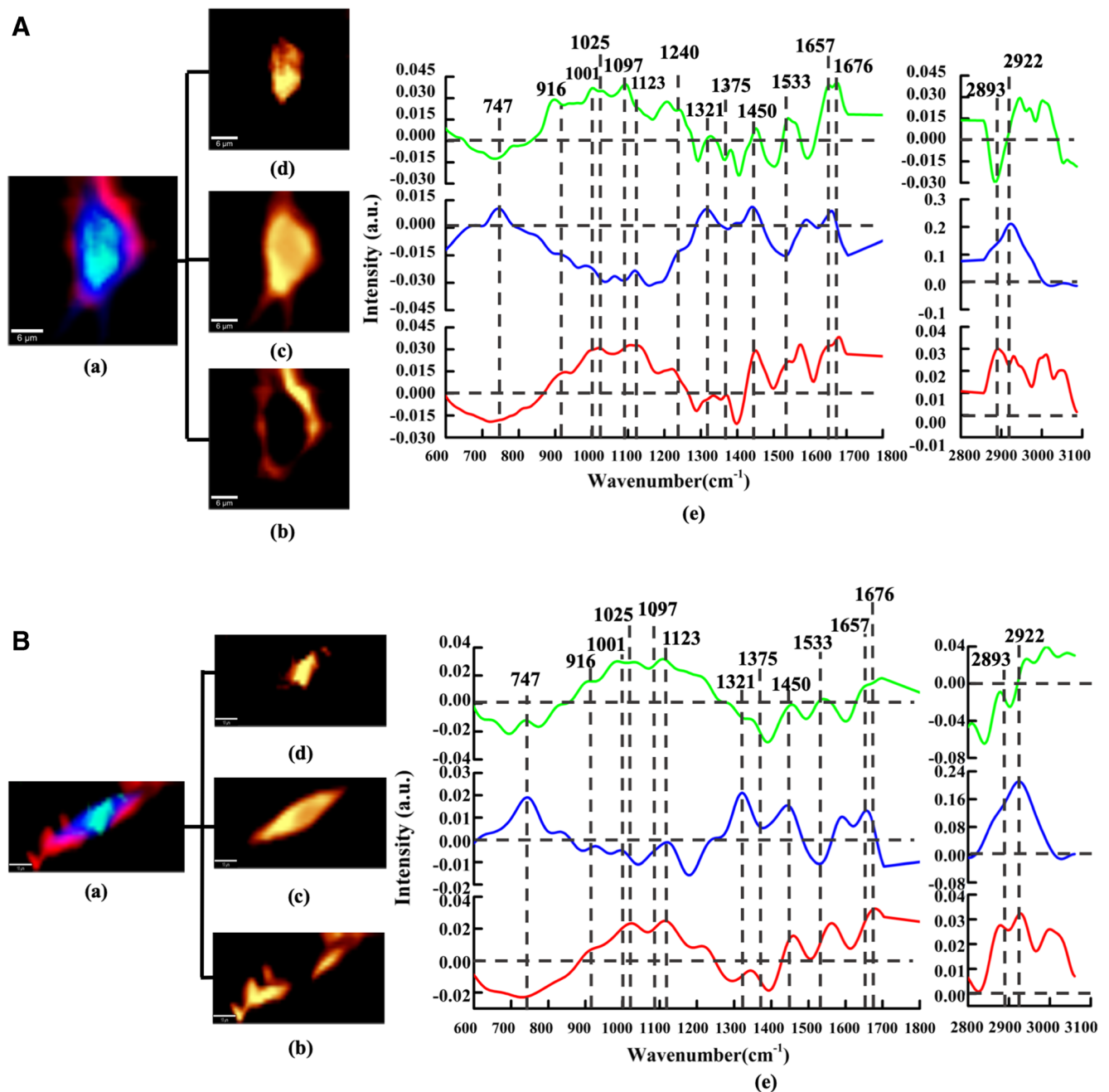


Fig. 3 Loading spectral plots and images for the first three principal components (PCs) resulting from principal component analysis (PCA) of the processed Raman spectra of the malignant osteoblast (a) and

spindle osteosarcoma cell (b). The spatial distribution on the sample of the feature spectra corresponding to each PC is shown on the left. The major bands were labeled with the wavenumber value

method can keep spectral quality of reconstructed Raman image, the spatial detail quality of Raman image can be improved when appropriate pixel density and integration time are chosen, but this should be implemented on the condition that the cell remains alive without any external labels.

Univariate and multivariate Raman spectral maps of individual live cells have been reported previously by several research groups (Krafft et al. 2006; Uzunbajakava et al. 2003; Matthäus et al. 2006; Us Matth et al. 2007). The main

advantage of Raman imaging over other cell microscopic imaging methods is that it provided a fingerprint image of the subcellular substance without the use of molecular labels or dyes. The variations in featured spectra from different regions of an osteosarcoma cell are quite small and unnecessarily interpreted by standard protocols, such as visual inspection and peak assignments. Therefore, the univariate Raman imaging method, displaying band intensity or intensity ratios, cannot provide enough information on the distribution of

subcellular substance. By contrast, a multivariate method interpreted more spectral information by reconstructing a pseudocolor image of the investigated cell by grouping the correlations in the point-scanned hyperspectral datasets. As shown in Fig. 2, the clusters produced by KCA correlate well with the bright-field microscopic image. The mean spectra obtained from each sub-cluster provided information for differentiating the sub-structures of the cells. KCA provided detailed information on the cell morphology and biochemistry by an evident separation of the nucleus from the cytoplasm and perinuclear area. It is possible that the lipid bilayer could be actually depicted by the high lipid contribution to the Raman signal from red and blue clusters in Fig. 2 a and b. However, a large number of clusters are often needed to describe all relevant regions, making the interpretation more difficult. Except for that, in this particular case, the intra-class variances were too slight to allow the identification of subcellular features. This was due to small transitions and slight intensity difference between each cluster. Thus, it shows that a standardized CRMI technique combined with multivariate analysis approaches demonstrates excellent sensitivity for discriminating composition information among the subcellular regions.

Nevertheless, PCA provided information on the molecular composition and biochemical variations of the datasets analyzed, and the results presented were comparable to the difference among spectra that could be calculated by simple subtraction. Figure 4 is a 3-D plot representation of PC1 vs. PC2 vs. PC3 for each spectrum extracted from the spectra dataset obtained from both malignant osteoblast and spindle osteosarcoma cell, respectively. As shown in Fig. 4a, PC1 from malignant osteoblast, which represented 87.07% of the explained variance, allowed good discrimination between the three groups. Notably, the cytoplasm (PC2) in the green color are symmetrically grouped at the positive extremes on PC1 axis, while the nucleus (PC3) in the blue color are grouped at ~ 0 . In PC2, which represents 3.71% of the variance, there is little or no discrimination between the cell membrane (PC1) and the cytoplasm (PC3), since they are both mainly grouped at the negative extremes on the PC2 axis. PC3, representing 0.72% of the explained variance, also exhibited good discrimination between the spectra from the cell membrane and the nucleus, since PC1 mainly distributed in its negative extreme value on the PC3 axis, while PC2 grouped at the negative and positive extremes. A similar observation could be derived from and similar conclusions could be obtained by Fig. 4b related to the spindle osteosarcoma cell, in which the first three components PC1, PC2, and PC3 account for 64.70%, 9.20%, and 1.44% of the total variance in the dataset, respectively. Referenced with Figs. 2 and 3, the tested bio-molecules presented different Raman features, whereas the intra-group variability between two types of cells is very low and the different spectra for each type are grouped with a similar pattern. However, the groups are well

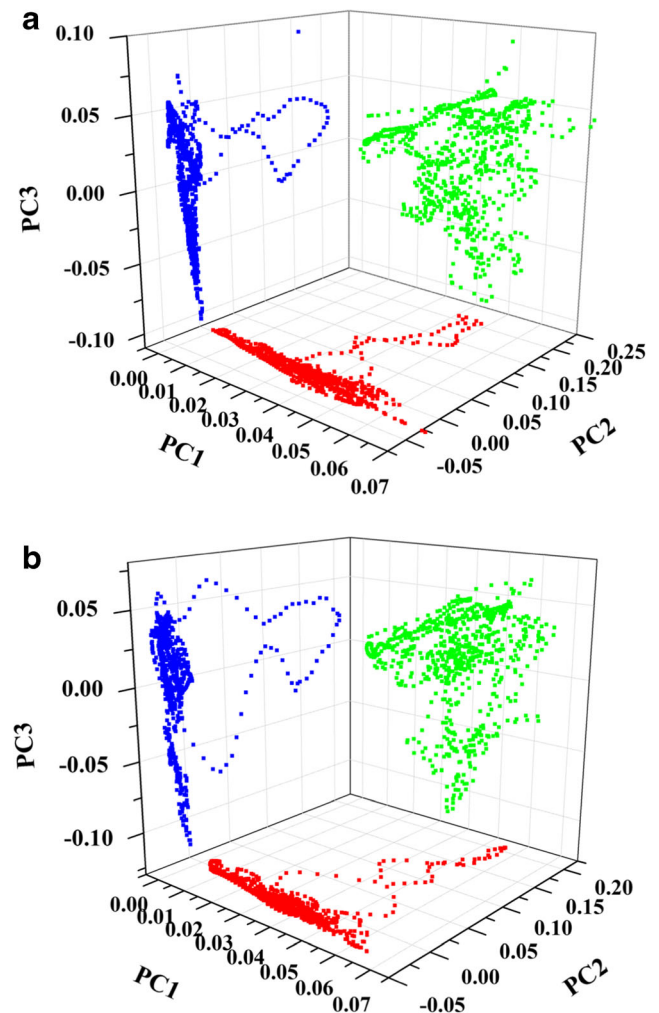


Fig. 4 3-D loading score plots for the first three PCs resulting from PCA of malignant osteoblast (**a**) and spindle osteosarcoma cell (**b**). In both plots, PC1 scatters was plotted in red, PC2 was plotted in green, and PC3 was in blue.

defined and discriminated across the score plot indicating that the PCs, which illustrate different subcellular structure or substance, are a clear representation of intra-group variance in each cell type.

The loadings of the principal components provide a multivariate profile of the biomolecular variations within the cell. Using the scale present for each PC, the negative and positive peaks could be attributed to the different components of the plot and could be a good representation of the spectral variation between the different cell structures tested. However, loadings were complex or inconclusive, and the chemical interpretation was not straightforward, as positive and negative values existed, and the spectral contribution of a chemical feature could be distributed over several principal components. This was observed in PC2 and PC3 for both cell types. As regards the Raman imaging of the single cell, the number of principal components was set to three and corresponded to the expected number of main spectral components. PC1, PC2,

and PC3 describe 91.5% of the spectral information in the malignant osteoblast, meaning that three loading vectors contained almost all chemical differences between the regions, as opposed to KCA algorithms that needed five or more clusters. However, the noise present in the spectra might increase the intragroup variability, thus reducing the specificity of PCA. In such a case, PC1, PC2, and PC3 contributed 75.34% of spectral identities of the spindle osteosarcoma cell. PC3 was rather noisy, and it was difficult to extract specific information from 900 to 1200 cm^{-1} and 2800–3100 cm^{-1} related to biomolecular constituents differentiating the nucleus and its surrounding structures in the score plot and induced a discontinuous structure in the loading image of PC1.

In conclusion, the CRMI approach provides diversified ways in which to address complex compositional and conformational information of live cells in a label-free and non-fixed manner. Furthermore, multivariate methods of analysis such as KCA and PCA were adopted to distinguish the cell membrane, cytoplasm, cell organelles, and nucleus from each other. In contrast to other groups applying Raman spectroscopy for studying osteosarcoma cells (Chiang et al. 2015), we further provide a visualized understanding on the subcellular nature in a label-free and non-fixed manner, since the resulting Raman images, including KCA and PCA, depicted both the morphology and biochemical composition of the nucleus, organelles, and membrane. Moreover, this also provided an integral comparison of two commonly used multivariate analytical techniques with two particular cases. K-means clustering analysis provides a visualized plot of spectrally differentiated regions, but it did not provide enough spectral information to reveal the biochemical differences among the identified regions. PCA loadings combined with K-means cluster analysis provided enough information to allow the discrimination among the identified regions, although it could be a complex mixture of many contributing molecular components. With its optimized experimental methodology including data acquisition, processing, and translation, this work provided a sound basis for future in vitro cell studies in our group, such as the discovery of tumor markers and the monitoring of cell apoptosis and cell responses to drugs.

Funding This research was funded by the Natural Science Research Foundation of Shaanxi Province, China [No. 2018JM6033], the Major Fundamental Research Program of Shaanxi Province, China [No.2016ZDJC-15], the Innovative Team Foundation of Shaanxi Province, China [S2018-ZC-TD-0061], and the Medical research program in “Science and Technology +” projects of Xi’an city, Shaanxi, China [201805096YX4SF30(2)].

Compliance with ethical standards

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

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Antonio KA, Schultz ZD (2014) Advances in biomedical Raman microscopy. *Anal Chem* 86(1):30–46
- Ashton L, Wright KL, Smith R (2016) Raman spectroscopy: An evolving technique for live cell studies. *Analyst* 141:3590
- Baron VO, Chen M, Clark SO, Williams A, Hammond RJ, Dholakia K, Gillespie SH (2017) Label-free optical vibrational spectroscopy to detect the metabolic state of *M. tuberculosis* cells at the site of disease. *Sci Rep* 7:9844
- Brolo AG, Jirasek A, Lum J, Matthews Q, Duan X (2010) Variability in Raman spectra of single human tumor cells cultured in vitro: Correlation with cell cycle and culture confluency. *Appl Spectrosc* 64:871–887
- Chan JW, Lieu DK, Huser T, Li RA (2009) Label-free separation of human embryonic stem cells and their cardiac derivatives using Raman spectroscopy. *Anal Chem* 81:1324–1331
- Chen M, McCreynolds N, Campbell EC, Mazilu M, Barbosa JA, Dholakia K, Powis SJ (2015) The use of wavelength modulated Raman spectroscopy in label-free identification of T lymphocyte subsets, natural killer cells and dendritic cells. *PLoS ONE* 10:e0125158
- Chen F, Flaherty BR, Cohen CE, Peterson DS, Zhao Y (2016) Direct detection of malaria infected red blood cells by surface enhanced Raman spectroscopy. *Nanomed Nanotechnol Biol. Med* 12:1445–1451
- Chiang YH, Wu SH, Kuo YC, Chen HF, Chiou A, Lee OK (2015) Raman spectroscopy for grading of live osteosarcoma cells. *Stem Cell Res Ther* 6:81–85
- Dieing T, Hollricher O (2008) High-resolution, high-speed confocal Raman imaging. *Vib Spectrosc* 48:22–27
- Dieing T, Hollricher O, Toporski J (2011) Confocal Raman Microscopy. *Microsc. Microanal* 15:1084–1085
- Diem M, Mazur A, Lenau K, Schubert J, Bird B, Miljkovič M, Krafft C, Popp J (2013) Molecular pathology via IR and Raman spectral imaging. *J Biophoton* 6:855–886
- Draux F, Jeannesson P, Beljebbar A, Tfayli A, Fourre N, Manfai M (2009) Raman spectral imaging of single living cancer cells: A preliminary study. *Analyst* 134:542–548
- Gelder JD, Gussem KD, Vandenabeele P, Moens L (2007) Reference database of Raman spectra of biological molecules. *J Raman Spectrosc* 38:1133–1147
- Hernández B, Pflüger F, Adenier A, Kruglik SG, Ghomi M (2010) Vibrational Analysis of Amino Acids and Short Peptides in Hydrated Media. Viii. Amino Acids with Aromatic Side Chains: L-Phenylalanine, L-Tyrosine, and L-Tryptophan. *J Phys Chem. B* 114:15319–15330
- Hernández B, Pflüger F, Kruglik SG, Ghomi M (2013) Characteristic Raman lines of phenylalanine analyzed by a multiconformational approach. *J. Raman. Spectrosc* 44:827–833
- Hobro AJ, Kumagai Y, Akira S, Smith NI (2016) Raman spectroscopy as a tool for label-free lymphocyte cell line discrimination. *Analyst* 141:3756–3764
- Konorov SO, Schulze HG, Piret JM, Blades MW, Turner RFB (2013) Label-free determination of the cell cycle phase in human embryonic stem cells by Raman microspectroscopy. *Anal. Chem* 85:8996–9002
- Krafft C, Knetschke T, Siegner A, Funk RHW, Salzer R (2003) Mapping of single cells by near infrared Raman microspectroscopy. *Vib. Spectrosc* 32:75–83
- Krafft C, Knetschke T, Funk RHW, Salzer R (2006) Studies on stress-induced changes at the subcellular level by Raman microspectroscopic mapping. *Anal. Chem* 78:4424–4429

- Macqueen J (1965) Some Methods for Classification and Analysis of MultiVariate Observations. Proc Berkeley Symp Math Stat Probab 1:281–297
- Madzharova F, Heiner Z, Kneipp J (2017) Surface enhanced hyper-Raman scattering of the amino acids tryptophan, histidine, phenylalanine, and tyrosine. J Phys Chem C 121:1235–1242
- Manciu FS, Ciubuc JD, Parra K, Manciu M, Bennet KE, Valenzuela P, Sundin EM, Durrer WG, Reza L, Francia G (2017) Label-free Raman imaging to monitor breast tumor signatures. Technol Cancer Res Treat 16:461–469
- Manen HJV, Kraan YM, Roos D, Otto C (2005) Single-cell Raman and fluorescence microscopy reveal the association of lipid bodies with phagosomes in leukocytes. Proc Natl Acad Sci USA 102:10159–10164
- Matthäus C, Boydston-White S, Miljković M, Romeo M, Diem M (2006) Raman and infrared microspectral imaging of mitotic cells. Appl Spectrosc 60:1–8
- Mavarani L, Petersen D, El-Mashtoly SF, Mosig A, Tannapfel A, Kotting C, Gerwert K (2013) Spectral histopathology of colon cancer tissue sections by Raman imaging with 532 nm excitation provides label free annotation of lymphocytes, erythrocytes and proliferating nuclei of cancer cells. The Analyst 138(14):4035–4039
- Mcreynolds N, Cooke FG, Chen M, Powis SJ, Dholakia K (2017) Multimodal discrimination of immune cells using a combination of Raman spectroscopy and digital holographic microscopy. Sci Rep 7:43631
- Meade AD, Clarke C, Draux F, Sockalingum GD, Manfait M, Lyng FM, Byrne HJ (2010) Studies of chemical fixation effects in human cell lines using Raman microspectroscopy. Anal. Bioanal. Chem 396:1781–1791
- Meister K, Schmidt DA, Brundermann E, Havenith M (2010) Confocal Raman microspectroscopy as an analytical tool to assess the mitochondrial status in human spermatozoa. Analyst 135:1370–1374
- Miljkovic M, Chernenko T, Romeo MJ, Bird B, Matthaus C, Diem M (2010) Label-free imaging of human cells: Algorithms for image reconstruction of Raman hyperspectral datasets. Analyst 135:2002–2013
- Naumann D (2001) Ft-infrared and ft-Raman spectroscopy in biomedical research. Appl. Spectrosc. Rev 36:239–298
- Nithipatikom K, Mccoy MJ, Hawi SR, Nakamoto K, Adar F, Campbell WB (2003) Characterization and application of Raman labels for confocal Raman microspectroscopic detection of cellular proteins in single cells. Anal. Biochem 322:198–207
- Notingher I (2007) Raman Spectroscopy Cell-based Biosensors. Sensors 7:1343–1358
- Notingher I, Hench LL (2006) Raman microspectroscopy: A noninvasive tool for studies of individual living cells in vitro. Expert Rev Med Dev 3:215–234
- Notingher I, Verrier S, Romanska H, Bishop AE, Polak JM, Hench LL (2002) In situ characterisation of living cells by Raman spectroscopy. J Spectrosc 16:43–51
- Pascut FC, Kalra S, George V, Welch N, Denning C, Notingher I (2013) Non-invasive label-free monitoring the cardiac differentiation of human embryonic stem cells in-vitro by Raman spectroscopy. Biochim Biophys 1830:3517–3524
- Pei-San H, Yi-Chun K, He-Guei C, Kenny HH, Kuang-Sheng LO, Ram SS (2013) Detection of osteogenic differentiation by differential mineralized matrix production in mesenchymal stromal cells by Raman spectroscopy. PLoS ONE 8:e65438–e65445
- Prats MB, Harreither E, Schosserer M, Puxbaum V, Gludovacz E, Borth N, Gierlinger N, Grillari J (2017) Label-free live cell imaging by confocal Raman microscopy identifies CHO host and producer cell lines. Biotechnol J 12:1600037
- Pully VV, Lenferink ATM, Otto C (2015) Time-lapse Raman imaging of single live lymphocytes. J Raman Spectrosc 42:167–173
- Rutan SC (2004) Chemometrics in Analytical Spectroscopy, 2nd edn. Adams MJ, Ed.; RMIT University, Melbourne, pp 14678–14679
- Salman A, Shufan E, Zeiri L, Huleihel M (2013) Detection and identification of cancerous murine fibroblasts, transformed by murine sarcoma virus in culture, using Raman spectroscopy; and advanced statistical methods. Biochim. Biophys. Acta 1830:2720–2727
- She CY, Dinh ND, Tu AT (2001) Laser Raman scattering of glucosamine N-acetylglucosamine, and glucuronic acid. Biochim Biophys 372:345–357
- Short KW, Carpenter S, Freyer JP, Mourant JR (2005) Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures. Biophys. J 88:4274–4288
- Sun H, Greathouse DV, Andersen OS, Ii REK (1998) The preference of tryptophan for membrane interfaces: insights from n-methylation of tryptophans in gramicidin channels*. Biochemistry 37(42):14713–14718
- Su X, Fang S, Zhang D, Zhang Q, Lu X, Tian J, Fan J, Li YZ (2017) Raman spectrum reveals mesenchymal stem cells inhibiting h160 cells growth. Spectrochim Acta A Mol Biomol Spectrosc 177:15–19
- Us Matth C, Chernenko T, Newmark JA, Warner CM, Diem M (2007) Label-free detection of mitochondrial distribution in cells by non-resonant Raman microspectroscopy. Biophys J 93:668–673
- Uzunbajakava N, Lenferink A, Kraan Y, Volokhina E, Otto C (2003) Nonresonant confocal Raman imaging of dna and protein distribution in apoptotic cells. Biophys J 84:3968–3981
- Vasko PD, Blackwell J, Koenig JL (1971) Infrared and Raman spectroscopy of carbohydrates: Part I: Identification of OH and CH-related vibrational modes for D-glucose, maltose, cellobiose, and dextran by deuterium-substitution methods. Carbohydr Res 19:297–310
- Woolford L, Chen M, Dholakia K, Herrington CS (2018) Towards automated cancer screening: Label-free classification of fixed cell samples using wavelength modulated Raman spectroscopy. J. Biophoton 11:e201700244
- Xu Y, Zhang Z, Zhang H (1998) Raman spectroscopic study of DNA after photosensitive damage caused by *hypocrellins* a and b. Sci China Ser C 41:360–366
- Xu YM, Zhang ZY, Zhao KJ, Zhang ZL, Liu CX, Wang DH, Xu GR, Zheng YZ (1993) Molecular mechanism of high energy proton radiation-Raman spectroscopic character of microcosmic damage in the space structure of DNA. Sci China B 36:1325–1332
- Ye YH, Chen Y, Li YZ, Feng SY, Su Y, Zou CY (2012) Discrimination of nasopharyngeal carcinoma and normal nasopharyngeal cell lines based on confocal Raman microspectroscopy. Chin J Lasers 39:0504003–0504005
- Yu C, Gestl E, Eckert K, Allara D, Irudayaraj J (2006) Characterization of human breast epithelial cells by confocal Raman microspectroscopy. Cancer Detect. Prev 30:515–522

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.