

Journal of Biomedical Optics

SPIDigitalLibrary.org/jbo

Toward label-free Raman-activated cell sorting of cardiomyocytes derived from human embryonic stem cells

Flavius C. Pascut
Huey T. Goh
Vinoj George
Chris Denning
Ioan Notingher

Toward label-free Raman-activated cell sorting of cardiomyocytes derived from human embryonic stem cells

Flavius C. Pascut,^a Huey T. Goh,^b Vinoj George,^b Chris Denning,^b and Ioan Notingher^a

^aUniversity of Nottingham, School of Physics and Astronomy, Nottingham NG7 2RD, United Kingdom

^bUniversity of Nottingham, Wolfson Centre for Stem Cells, Tissue Engineering and Modelling, Centre for Biomolecular Sciences, University Park, Nottingham NG7 2RD, United Kingdom

Abstract. Raman micro-spectroscopy (RMS) has been recently proposed for label-free phenotypic identification of human embryonic stem cells (hESC)-derived cardiomyocytes. However, the methods used for measuring the Raman spectra led to acquisition times of minutes per cell, which is prohibitive for rapid cell sorting applications. In this study we evaluated two measurement strategies that could reduce the measurement time by a factor of more than 100. We show that sampling individual cells with a laser beam focused to a line could eliminate the need of cell raster scanning and achieve high prediction accuracies (>95% specificity and >96% sensitivity) with acquisition times ~5 seconds per cell. However, the use of commercially-available higher power lasers could potentially lead to sorting speeds of ~10 cells per s. This would start to progress RMS to the field of cell sorting for applications such as enrichment and purification of hESC-derived cardiomyocytes. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3570302]

Keywords: label-free cell sorting; Raman spectroscopy; embryonic stem cells; cardiomyocytes.

Paper 11034LR received Jan. 20, 2011; revised manuscript received Mar. 2, 2011; accepted for publication Mar. 4, 2011; published online Apr. 21, 2011.

The pluripotency and ability to proliferate indefinitely make human embryonic stem cells (hESCs) attractive to many applications, including regenerative medicine, tissue engineering, drug screening, and even cancer treatments. More recently, the first clinical trials with hESC derivatives have started in humans to evaluate the safety issues and tolerance for human patients with thoracic spinal cord injuries.¹ Despite these advances, high efficiency differentiation and/or enrichment to produce homogeneous populations of the desired cell types still remains a challenge. This is of considerable concern since any contaminating undifferentiated hESCs have the potential to give rise to tumor-like growths after transplantation. Indeed, it is well known that undifferentiated hESCs form teratomas when injected into immuno-compromised mice.² In addition, the presence of unwanted cell types may compromise the suitability of the transplants. For example, transplant of skeletal muscle cells instead of cardiomyocytes into the left ventricular wall of the heart can cause harmful arrhythmias.³ These concerns highlight the immediate need for noninvasive techniques capable of phenotypic identification and sorting of hESC progeny. Such label-free techniques could deliver high-purity cell populations with a desired phenotype suitable for therapy.

Fluorescence and magnetic activated cell sorting (FACS, MACS) are the standard tools of choice in stem cell sorting application. However, if the cells are to be used for therapeutic purposes, the applications of these techniques are restricted to cells that exhibit lineage-specific surface markers.

For many cell types, such as cardiomyocytes (CMs), there are no appropriate surface markers, therefore FACS and MACS are not suitable for providing viable cells as required for clinical applications.

A unique feature of Raman micro-spectroscopy (RMS) is the ability to measure biochemical properties and provide functional imaging of live cells without requiring labeling or sample preparation.⁴⁻⁶ Compared to lasers in the visible range, using near-infrared laser (e.g., 785 nm) with powers below 150 mW, allows measurements of Raman spectra from live cells with reduced probability of inducing photodamage. Thus, cells can be investigated over extended periods of times under sterile physiological conditions (culture medium, temperature, etc).⁷ RMS was used to discriminate between undifferentiated hESCs and hESC-derived CMs with 66% accuracy.⁸ More recently, it was also demonstrated that RMS can identify spectral markers which can be used for identification of CMs within highly-heterogeneous cell populations as derived during differentiation of hESCs toward cardiac phenotype.⁹ The Raman spectral markers consisted of bands attributed mainly to glycogen and myofibril proteins, enabling classification of hESC-derived CMs with >97% specificity and >96% sensitivity against the immunofluorescence staining gold-standard (alpha-actinin).⁹

However, the main limitation of the RMS technique is the low efficiency of Raman scattering effect. Acquisition times of 2 to 10 min per cell have been reported,^{8,9} which are indeed not practical for cell sorting applications. In this paper we investigated the feasibility of two measurement strategies that could dramatically reduce the acquisition time. These strategies are based on two key factors: intracellular spatial distribution

Address all correspondence to: Ioan Notingher, University of Nottingham School of Physics and Astronomy, Nottingham NG7 2RD, United Kingdom. Tel: 004411595151572; E-mail: ioan.notingher@nottingham.ac.uk.

of the Raman spectral markers and the level of uncorrelated measurement noise in the Raman spectra. These factors affect both the measurement time and the prediction accuracy of the RMS. The current analysis is performed on a database of Raman spectra collected from 50 CMs and 40 non-CMs derived from hESCs and maintained under physiological conditions during the measurements.⁹ The measurements were carried out on more than 20 cell culture flasks and both CMs and non-CMs were measured for each flask.

The first objective of this paper was to carry out an analysis that quantifies the effect of noise in the Raman spectra of individual cells on the prediction accuracy for the cardiac phenotype. For this purpose, the principal component analysis (PCA) model described previously⁹ was modified to determine the effect of acquisition time for Raman spectra on the classification sensitivity and specificity. For each cell included in the model, the Raman spectrum was calculated by averaging point spectra measured by raster scanning the cell through the laser focus at $2\ \mu\text{m}$ steps (total of 625 data points) with an integration time of 1 s per point. These Raman spectra had a high signal-to-noise ratio (SNR) but required a total acquisition time of 10 min per cell. However, for prediction purposes, the sorting time could be reduced considerably if the Raman spectra of the sorted cells were acquired at the lowest acquisition time (lowest SNR) which would still provide the required classification accuracy (e.g., 95% sensitivity/specificity). Such a model is valid in a cell sorting scenario, where the prediction model is constructed well in advance using a database of high SNR spectra, while the sorting is performed at a high speed on low SNR Raman spectra.

To evaluate the effect of the noise in the Raman spectra of sorted cells, a 5-fold cross-validation (CV) algorithm was applied on the modified PCA model: 80% of Raman spectra (10 min per cell acquisition time) were used for building a PCA model while the remaining 20% of Raman spectra corresponding to acquisition times of 1 to 10 s per cell were used for prediction. Raman spectra corresponding to a typical CM and a non-CM obtained are shown in Fig. 1(a), alongside the spectrum corresponding to the Raman spectral marker of CMs identified by the PCA model.⁹ Figure 1(b) shows the Raman spectrum of the same CM but with various signal-to-noise ratio values to simulate acquisition times between 1 to 20 s per cell. The main Raman bands at 860, 938, 1084, and $1123\ \text{cm}^{-1}$ can be visually identified even in the spectrum with the shortest acquisition time of ~ 1 s per cell.

The CV was performed for various acquisition times. The specificity and sensitivity for discrimination CMs at a target 100% specificity are shown in Fig. 2. The gray shaded areas represent the \pm standard deviation of the sensitivity/specificity as calculated using different partitions for the CV. The specificity of 97.54% and sensitivity of 96.3% depicted as horizontal lines were obtained when the acquisition time of the Raman spectra for the prediction cells was similar to the acquisition time for the Raman spectra of the cells used for building the classification model. Despite imposing a highly specific regime with target 100% specificity as required for regenerative medicine (achieve cell populations with high levels of phenotypic purity), the spectral model shows a high resilience to noise up to the point corresponding to 6 s per cell acquisition time, after which the average predicted specificity drops below the 95% mark.

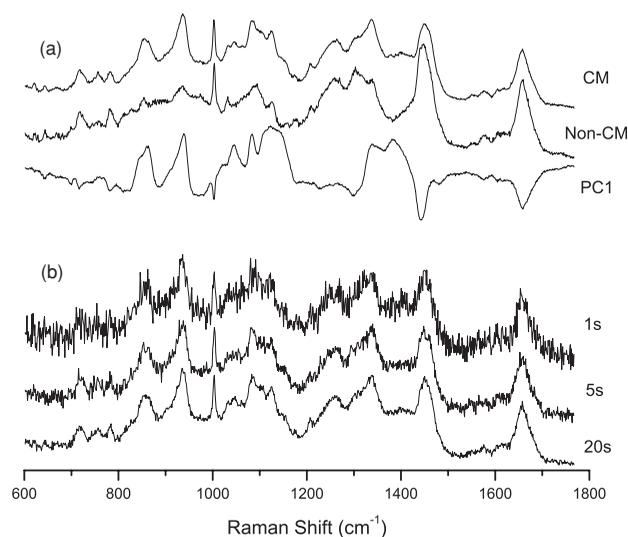


Fig. 1 (a) Comparison between a typical Raman spectrum of an individual CM, non-cardiomyocyte (non-CM), and the Raman spectral marker as identified by the PCA model (PC1). (b) Raman spectra of the same CM with added noise corresponding to an acquisition time of 1, 5, and 20 s.

However, the prediction sensitivity remains mostly unchanged regardless of the noise level.

Apart from acquisition time, the intracellular distribution of the Raman spectral markers can also affect the predictive performance of the classification model. Ideally, the measurement system should be optimized such as the sampling volume to match the volume of the cells ($\sim 20\ \mu\text{m}$). However, such conditions are difficult to achieve in practice. A common approach is to use a low numerical aperture objective; however since the depth of the laser focus becomes larger than the thickness of the cell, the laser power density interacting with the

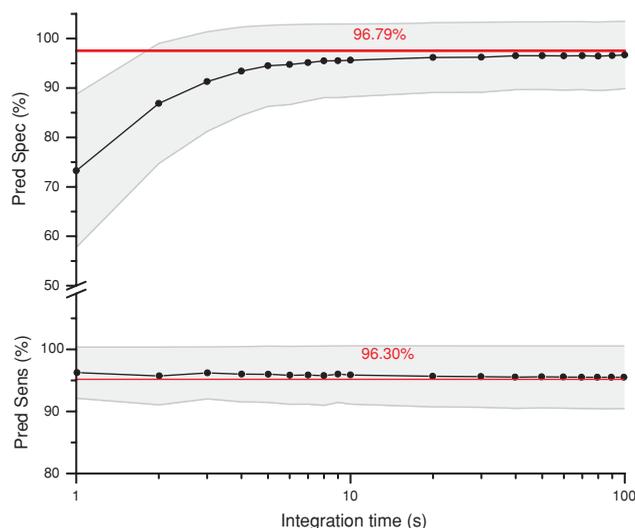


Fig. 2 Prediction sensitivity and specificity as a function of the acquisition time of Raman spectra of individual cells. Horizontal lines show the values that are obtained when data corresponding to 625 s integration time for each cell is used.

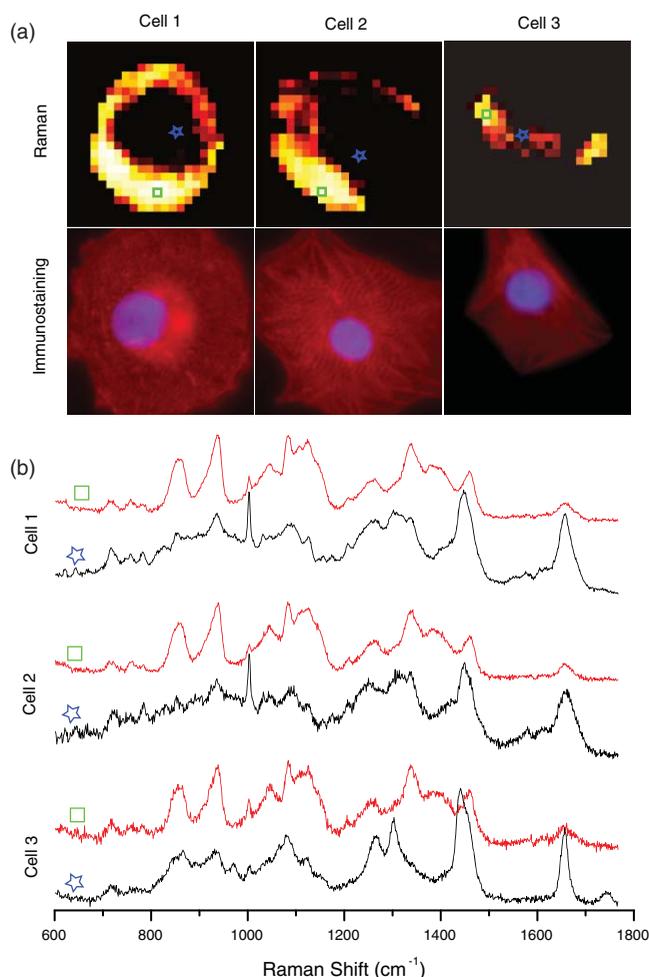


Fig. 3 (a) Raman spectral maps corresponding to the cardiac Raman spectral marker for three typical CMs and fluorescence staining for alpha-actinin (myofibrils) and DAPI (cell nuclei). (b) Typical point Raman spectra taken from various points inside cells, corresponding to regions rich in myofibrils, nucleus, and lipid bodies.

cell molecules decreases thus the signal to noise of the Raman spectra becomes lower. To avoid under-sampling when using high numerical aperture objectives, a common practice involves raster scanning the cells through the laser focus at step sizes similar to the laser spot ($2\ \mu\text{m}$), requiring ~ 600 points to ensure full sampling. However, if the Raman spectral markers are distributed inside the cells in regions larger than the laser spot, the number of steps and the acquisition time could be reduced.

A surface angle mapping algorithm,¹⁰ was used to determine the spatial distribution of the Raman spectral marker for individual CMs. The maps corresponding to the Raman spectral marker show three typical spatial distributions (Fig. 3). Approximately 50% of CMs showed a circular distribution of the Raman marker around a focal center. For the remaining CMs, the Raman spectral marker was either concentrated in one place at one extremity of the cell ($\sim 30\%$ of CMs) or the distribution was split up in several smaller regions of higher concentrations toward the edge of the cells. Point Raman spectra from selected positions inside the three cells are also shown in Fig. 3.

Raman bands corresponding to proteins ($853\ \text{cm}^{-1}$ assigned to C–C in proline and ring-breathing mode in tyrosine), and $936\ \text{cm}^{-1}$ associated with C–C stretching of protein backbones) can be identified at cell positions where the alpha-actinin staining indicates a high concentration of myofibrils. Raman bands corresponding to DNA (O–P–O phosphodiester bonds at $788\ \text{cm}^{-1}$) can also be identified in the point Raman spectra corresponding to the cell nuclei (DAPI staining). Certain cells also show regions for which Raman spectra have strong bands associated to lipids (C=C stretching at $1659\ \text{cm}^{-1}$, CH₂ deformation $1441\ \text{cm}^{-1}$, CH₂ twisting $1303\ \text{cm}^{-1}$), cytoplasmic regions likely to correspond to lipid bodies.

Based on the Raman spectral marker distributions in Fig. 3, focusing the laser beam in a line across the cell and focusing the Raman scattered photons on the entrance-slit of the spectrometer^{6–11} would significantly reduce acquisition time as no raster scanning of the cell or laser spot would be required. This approach is also better suited for cells passing continuously through micro-fluidic channels as in flow cytometry. To determine the feasibility of this method, laser-line sampling was simulated by selecting only 10 Raman spectra corresponding to a single row in the original raster scans (the equivalent acquisition time for such line Raman spectra is 10 s). The PCA model was modified and a 5-fold CV was carried out to evaluate the classification accuracy if each cell in the classification group were scanned only in a single line across the cell at a randomly chosen position. The CV model suggests that this sampling method would enable identification of CMs with $93.62 \pm 3.18\%$ specificity and $90.36 \pm 3.22\%$ sensitivity at a target specificity of 99%.

As it appears that glycogen could be used for discrimination of hESC-derived CMs, measurements in the lower frequency spectral range have been carried where glycogen has a strong band at $482\ \text{cm}^{-1}$ while proteins show no contributions. Raman spectral map of PC1 and the map corresponding to the $482\ \text{cm}^{-1}$ band area are very similar in hESC-derived CMs Figs. 4(a)–4(c). In addition, this band is not observed in the Raman spectra of non-CMs. Figure 4(d) shows that the $482\ \text{cm}^{-1}$ can be visually identified in a Raman spectrum measured in a line scan (cell scanned through the laser spot) in only 1 s. Considering that this band is isolated from the other major Raman bands of the cells, it may be used in a single wavelength detection setup, which could greatly simplify and speed up the technique.

The results in this paper suggest that high classification accuracy ($>95\%$ specificity and sensitivity) for hESCs-derived cardiomyocytes can be achieved by single measurements of ~ 5 s acquisition time per cell if the laser beam was focused to a $\sim 2\ \mu\text{m} \times 20\ \mu\text{m}$ line. This sampling model ensures detection of the Raman spectral markers specific to hESC-derived CMs and avoids the time-consuming raster scanning. However, in the future, this acquisition time could be reduced even further if higher power lasers were used. For example, continuous wave Ti:sapphire lasers with output power of >5 W at 785 nm are commercially available (e.g., 3900S Spectra-Physics, UK). The output power of such lasers is ~ 20 times higher than in the current study, thus the laser beam could be expanded into a line of $1\ \mu\text{m} \times 800\ \mu\text{m}$ while maintaining the same laser power density as in the current study. The focused laser line could cover ~ 40 parallel sorting flow columns enabling sorting speeds of

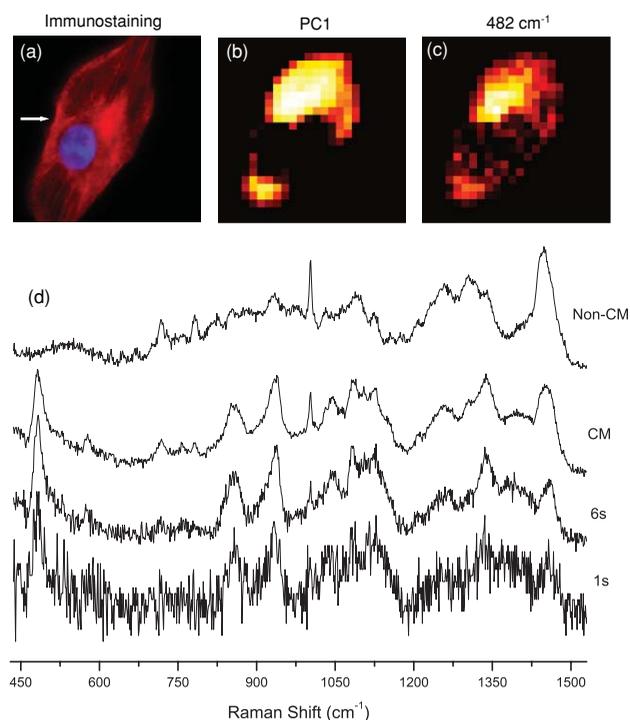


Fig. 4 Immuno-fluorescence for α -actinin (a), Raman map of the PC1 (b), and Raman map of the 482 cm^{-1} band attributed to glycogen (c) for a typical CM derived from hESC. (d) Average Raman spectrum for the same CM compared to typical average spectrum for a non-CM. Line scans over the CM for acquisition times 1 to 6 second recorded by scanning the cell through the laser spot in the horizontal direction at the position indicated by the arrow.

approximately ~ 8 cells per s. Such sorting speeds compare favorably with the speeds of FACS during its developmental phase.¹² In addition, the continuous flow of cells through the laser line would also provide an increase in sampling volume, particularly for CMs for which the molecular markers are concentrated in one or several smaller regions at the edge of the cells, thus increasing the discrimination accuracy.

Acknowledgments

The authors acknowledge the financial support from the Biotechnology and Biological Sciences Research Council UK (BB/G010285/1).

References

1. J. Apler, "Geron gets green light for human trial of ES cell-derived product," *Nat. Biotechnol.* **27**, 213–214 (2009).
2. B. E. Reubinoff, M. F. Pera, C. Y. Fong, A. Trounson, and A. Bongso, "Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro," *Nat. Biotechnol.* **18**(4), 399–404 (2000).
3. P. Menasché, A. A. Hagège, J. T. Vilquin, M. Desnos, E. Abergel, B. Pouzet, A. Bel, S. Sarateanu, M. Scorsin, K. Schwartz, P. Bruneval, M. Benbunan, J. P. Marolleau and D. Duboc, "Autologous skeletal myoblast transplantation for severe postinfarction left ventricular dysfunction," *J. Am. Coll. Cardiol.* **41**(7), 1078–1083 (2003).
4. G. J. Puppels, F. F. de Mul, C. Otto, J. Greve, M. Robert-Nicoud, D. J. Arndt-Jovin, and T. M. Jovin, "Studying single living cells and chromosomes by confocal Raman microspectroscopy," *Nature (London)* **347**(6290), 301–303 (1990).
5. I. Notingher and L. L. Hench, "Raman microspectroscopy: a noninvasive tool for studies of individual living cells in vitro," *Expert Rev. Med. Devices* **3**(2), 215–234 (2006).
6. K. Hamada, K. Fujita, N. I. Smith, M. Kobayashi, Y. Inouye and S. Kawata, "Raman microscopy for dynamic molecular imaging of living cells," *J. Biomed. Opt.* **13**(4), 044027 (2008).
7. A. Zoladek, F. C. Pascut, P. Patel, and I. Notingher, "Non-invasive time-course imaging of apoptotic cells by confocal Raman microspectroscopy," *J. Raman Spectrosc.* (in press).
8. J. W. Chan, D. K. Lieu, T. Huser and R. A. Li, "Label-free separation of human embryonic stem cells and their cardiac derivatives using Raman spectroscopy," *Anal. Chem.* **81**(4), 1324–1331 (2009).
9. F. C. Pascut, H. T. Goh, N. Welch, L. D. Buttery, C. Denning and I. Notingher, "Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells," *Biophys. J.* **100**(1), 251–259 (2011).
10. N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Trans. Geoscience and Remote Sensing* **42**(7), 1552–1565 (2004).
11. D. K. Veirs, J. W. Ager III, E. T. Loucks, and G. M. Rosenblatt, "Mapping materials properties with Raman spectroscopy utilizing a 2-D detector," *Appl. Opt.* **29**(33), 4969–4980 (1990).
12. H. P. Mansberg, A. M. Saunders, and W. Groner, "The Hemalog D white cell differential system," *J. Histochem. Cytochem.* **22**(7), 711–724 (1974).