Label–Free Separation of Human Embryonic Stem Cells and Their Cardiac Derivatives Using Raman Spectroscopy

4 AUTHORS, INCLUDING:

James Chan

University of California, Davis

61 PUBLICATIONS **1,809** CITATIONS

SEE PROFILE

Deborah Lieu

University of California, Davis

42 PUBLICATIONS **1,031** CITATIONS

SEE PROFILE

Thomas R Huser

Bielefeld University

162 PUBLICATIONS **4,479** CITATIONS

SEE PROFILE

Label-free separation of human embryonic stem cells (hESCs) and their cardiac derivatives using Raman spectroscopy

James W. Chan[1,2], **Deborah K. Lieu**[2,4], **Thomas Huser**[2,5], and **Ronald A. Li**[3,4,6]

[1]*Applied Physics and Biophysics Division, Physical Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA*

[2]*NSF Center for Biophotonics Science and Technology, University of California, Davis, CA*

[3]*Stem Cell Program, University of California, Davis, CA*

[4]*Department of Cell Biology and Human Anatomy, University of California, Davis, CA*

[5]*Department of Internal Medicine, University of California, Davis, CA*

[6]*Institute of Pediatric Regenerative Medicine, Shriners Hospital for Children of North America, Sacramento, CA*

Abstract

Self-renewable, pluripotent human embryonic stem cells (hESCs) can be differentiated into cardiomyocytes (CMs), providing an unlimited source of cells for transplantation therapies. However, unlike certain cell lineages such as hematopoietic cells, CMs lack specific surface markers for convenient identification, physical separation, and enrichment. Identification by immunostaining of cardiac-specific proteins such as troponin requires permeabilization, which renders the cells unviable and non-recoverable. Ectopic expression of a reporter protein under the transcriptional control of a heart-specific promoter for identifying hESC-derived CMs (hESC-CMs) is useful for research but complicates potential clinical applications. The practical detection and removal of undifferentiated hESCs in a graft, which may lead to tumors, is also critical. Here, we demonstrate a *non-destructive, label-free* optical method based on Raman scattering to interrogate the intrinsic biochemical signatures of individual hESCs and their cardiac derivatives, allowing cells to be identified and classified. By combining the Raman spectroscopic data with multivariate statistical analysis, our results indicate that hESCs, human fetal left ventricular CMs, and hESC-CMs can be identified by their intrinsic biochemical characteristics with an accuracy of 96%, 98% and 66%, respectively. The present study lays the groundwork for developing a systematic and automated method for the non-invasive and label-free sorting of i) high-quality hESC*s* for expansion, and ii) *ex vivo* CMs (derived from embryonic or adult stem cells) for cell-based heart therapies.

Keywords

Raman spectroscopy; human embryonic stem cells; myocardial repair; regeneration; cardiomyocytes; single cell; label-free

Correspondence to: James W. Chan.

CORRESPONDING AUTHOR FOOTNOTE: James W. Chan, Applied Physics and Biophysics Division, Lawrence Livermore National Laboratory, 7000 East Ave, L-211, Livermore, CA 94550, Tel: 925-423-3565, Fax: 925-424-2778, Email: chan19@llnl.gov.
EMAIL ADDRESS: chan19@llnl.gov

INTRODUCTION

Cardiomyocytes (CMs) are normally non-regenerative and their permanent loss due to aging or disease compromises cardiac functions. Heart transplantation is typically the last resort for end-stage heart failure patients. However, this option is hampered by a severe shortage of donor organs[1-3]. Human embryonic stem cells (hESCs) derived from the inner cell mass of blastocysts can self-renew while maintaining their normal karyotypes and pluripotency to differentiate into virtually all cell types[4], including CMs[5-10]. Indeed, hESC-derived CMs (hESC-CMs) display structural and functional properties of early-stage CMs[6,7], and can functionally integrate with[10] or even electrically pace[5] the recipient heart after transplantation *in vivo*. Thus, hESCs have the potential to act as an unlimited *ex vivo* source of CMs for cell-based heart therapies. Although hESCs offer unprecedented hopes for myocardial repair, there are presently numerous technical hurdles. For instance, *in vitro* differentiation, typically by forming three-dimensional aggregates termed as embryoid bodies, non-specifically generates all three germ layers (i.e. endoderm, mesoderm and ectoderm) and their corresponding lineages. Therefore, it is necessary to purify CMs for clinical applications. Additionally, the presence of contaminated, undifferentiated hESCs in a graft may lead to the formation of tumors after transplantation. Unlike many other lineages, CMs lack specific surface markers for convenient physical separation or enrichment (e.g., magnetic bead sorting of CD34+ hematopoietic cells). Immunostaining of cardiac-specific proteins such as troponin requires permeabilization, which renders the cells unviable and non-recoverable. Ectopic expression of a reporter protein under the transcriptional control of a heart-specific promoter for identifying hESC-CMs[11] is useful for research but complicates potential clinical applications.

Isolation techniques are also needed to maintain high quality and purity, pluripotent hESC colonies. Pluripotent hESCs are cultured as colonies and tend to spontaneously differentiate even under the best culturing conditions. Conventional enzymatic methods for propagation involve the digestion of all colonies[4,5] virtually without selection and thus compromise the culture quality over time (e.g., by accumulating karyotypic abnormalities). For quality control, viable cells need to be sacrificed for non-recoverable analytical procedures such as karyotyping and immunostaining for pluripotency markers. The mechanical dissection method[3] allows experienced users to select the most pluripotent cells for propagation; although this labor-intensive technique generally improves the culture quality, it still lacks the systematic objectivity required for high-throughput, high-quality cell culture maintenance and the eventual clinical applications. Similar arguments can be made for the isolation of hESC-CMs by physical dissection of the beating areas. Although a useful research technique for isolating these cells, it is not sufficient for clinical use because these areas may still contain a wide range of cells, both cardiac and non-cardiac, as well as cells in different maturation stages.

An objective, label-free and noninvasive approach is clearly needed for systematic identification, isolation and purification of hESCs and their derived cardiomyocytes. Micro-Raman spectroscopy is a laser-based, label-free, and noninvasive method that measures the inelastic scattering of incident photons by intrinsic molecular bonds[12,13]. Scattered photons that are shifted in wavelength from that of the incident photon reflect the underlying biomolecular composition and structural conformations of macromolecules in living cells. DNA, RNA, proteins, lipids and carbohydrates exhibit multiple unique spectral markers that can be detected as vibrational Raman frequencies (see Table 1 for a list of representative Raman peak frequencies and their corresponding assignments). Puppels and colleagues[12] first demonstrated the use of confocal Raman microspectroscopy on single eukaryotic cells. This method has since evolved and been tested as a potential diagnostic tool for atherosclerosis[14,15] and cancer detection[16-20]. For example, the detection of single leukemia cells[18] and the diagnosis of breast cancer[17] have been demonstrated using Raman spectroscopy. Similarly,

different grades of squamous dysplasia[20], a precursor to cervical cancer, can be classified based on their Raman spectra.

However, the potential of Raman spectroscopy for stem cell characterization is in its infancy. Only a handful of studies[21-25] using vibrational spectroscopy to characterize mouse embryonic stem cells and mesenchymal stem cells have been performed but, to date, no studies have specifically characterized hESCs and their cardiac derivatives. Here, we demonstrate, using micro-Raman spectroscopy in combination with multivariate statistical methods for classification, that hESCs and hESC-CMs carry specific Raman signals that have the potential to function as "optical fingerprints" for their identification, separation, and enrichment. These results lay the foundation for the development of single cell Raman spectroscopy as a systematic method for sorting ex-vivo CMs derived from reprogrammed, embryonic, or adult stem cells for future cell-based heart therapies.

EXPERIMENTAL SECTION

Culturing and differentiation of hESCs

The HES2 line (ESI, Singapore; NIH code ES02) used in this study was cultured and differentiated as previously described[3,7]. Briefly, HES2 cells were grown on mitomycin C-inactivated mouse embryonic feeders (mEFs). HES2 culture medium consisted of DMEM (Invitrogen, Carlsbad, CA) containing 2 mM L-glutamine, 1% insulin-transferrin-selenium, 1% non-essential amino acids, 90 μM β-mercaptoethanol, 20 U/ml penicillin, 20 μg/ml streptomycin, and 20% fetal bovine serum (Hyclone, Logan, UT). HES2 cells were passaged by manually cutting colonies into pieces and removing them from the mEFs using 8 mg/mL dispase (Invitrogen, Carlsbad, CA). To induce cardiac differentiation, HES2 cells were co-cultured with mitomycin C-treated END2 cells, a visceral endoderm-like cell line, at 100% confluence for ~2-3 weeks in HES2 medium without serum. To isolate hESC-CMs, beating outgrowths were microsurgically dissected from 18-24-day old embryoid bodies by a glass knife.

Isolation of human fetal left ventricular cardiomyocytes (FLV-CMs)

Human FLV-CMs were isolated according to protocols approved by the UC Davis IUPAC and IRB (Protocol #200614787-1). Briefly, fetal hearts (16-18 weeks; Advanced Bioscience Resources, INC. Alameda, CA) were perfused with 200U/ml collagenase II, 0.4mg/ml protease, and 1.2mg/ml bovine serum albumin in Tyrode solution without Ca^{2+} for 30min at 37°C with constant oxygenation using a customized Langendorff apparatus, followed by mincing of the left ventricle to collect FLV-CMs in high K^+ solution. Ca^{2+} was slowly reconstituted back into the FLV-CM suspension by slowly replacing the high K^+ solution with Medium 199 (Invitrogen) containing: 5 mM carnitine (Sigma-Aldrich Corp. St. Louis, MO), 5 mM creatine, 5 mM taurine, 100 μg ml^{-1} penicillin-streptomyocin and 10% fetal bovine serum.

Micro-Raman spectroscopy system

A schematic of the confocal Raman microscope used in the present study is shown in Figure 1. The setup consisted of a fiber-coupled 785 nm continuous wave (CW) laser (Crystalaser, Reno, NV) that operated at 70 mW. The output from the fiber was collimated with a telescope to achieve a beam diameter of roughly 6 mm. This beam was delivered into the side port of an inverted microscope (Olympus, Center Valley, PA) and delivered through a 60x, 1.2 NA water immersion objective using a longpass dichroic filter. A fused silica coverslip containing a ~ 100 μL cell solution was placed on the microscope stage. Fused silica was specifically chosen because of its low autofluorescence when excited with 785 nm to minimize the interference with the cell spectra. Furthermore, its dominant Raman signals are located at lower

wavenumbers (<600 cm^{-1}), which does not overlap with any major biological Raman signatures that primarily occur in the 700-2000 cm^{-1} region, making it straightforward to remove from the cell spectra. White light images obtained by a charge-coupled device (CCD) camera mounted on the microscope functioned as visual guides to locate the cells for Raman interrogation. Raman signals generated at the sample were epi-detected using the same objective and focused through a 100 μm pinhole confocal to the laser focus to reject out-of-focus background light. A 785 nm notch filter was used to suppress residual excitation light, and the beam was focused into a spectrometer (PI Acton, Trenton, NJ) equipped with a 600 grooves/mm grating that dispersed the Raman signals onto a thermoelectrically cooled 100 × 1340 pixel CCD camera (PI Acton). The microscope stage was translated until a cell overlapped the focused laser beam, as observed by white light imaging with the CCD camera. Spectra were acquired by the Winspec (PI Acton) software program and the data converted for analysis using MATLAB (The MathWorks, Natick, MA).

Preparation of samples for Raman analysis

Raman spectroscopy was performed on individual cells for all three cell types (i.e. hESCs, hESC-CMs, and FLV-CMs). Figure 2 shows white light images representing the three different cell types that were prepared for Raman analysis. The laser focus was positioned in the center of a cell and the spectra were acquired for 2 minutes. The laser power at the focus was roughly 25 mW. While the FLV-CMs were prepared as individual isolated cells, the hESCs and hESC-CMs cells were analyzed while remaining intact in a cell colony in order to maintain the same physiological conditions used in the normal passaging of these cells. Individual cells within the colonies can still be delineated, especially on the outer edges of the cluster. Furthermore, cells within the clusters that were closest to the glass substrate were chosen. Given the tight focusing conditions we are using (~ 1 μm^3 volume), individual cells can be selected and probed with minimal interference from neighboring cells. For the hESC-CMs, clusters were allowed to stop spontaneously beating before Raman analysis was performed on the cells. On average, cells for each cell type were prepared and analyzed over 2-3 separate days. All cell samples were immersed in medium at room temperature during the Raman analysis. Approximately 10-15 cells were sampled on any particular day; therefore, the total time required to analyze the cells under the microscope typically lasted 40 minutes to 1 hour. We restricted the time that the cells were on the microscope to be equivalent to the average time it takes for routine passaging of these cells at room temperature to ensure cell viability and pluripotency and reduce the effects of temperature and CO_2 variability.

Spectral processing and multivariate analysis

Routines were written in MATLAB for processing and analyzing the raw data. Each spectrum was background-subtracted using a modified polynomial fitting method[26]. In this method, the initial spectrum was fit to a fourth order polynomial curve. This curve was then modified such that all the data points in the generated curve with a higher intensity value than the original input spectrum were replaced by their initial values from the original spectrum. The resulting curve was fit to another fourth order polynomial and the same modification was performed on the new polynomial fit. This process was iterated approximately 100 times until a final smooth curve was generated to represent the overall background feature of the spectrum. This final background polynomial curve was then subtracted from the original spectrum. This approach provides an objective, automated method for correcting background contributions to each spectrum. All spectra were then normalized to the peak intensity of the 1450 cm^{-1} peak, which is predominantly related to protein and lipid, to enable direct comparison of the data. Given the lack of a known internal spectral standard for normalization, the 1450 cm^{-1} peak was chosen because of previous knowledge[21,22] that the largest spectral variations in the Raman spectra of stem cells and differentiated cells primarily involved Raman bands associated with DNA and RNA. It should be noted that the cell spectra exhibited minimal contribution from water,

which we attribute to the tight focusing condition that enabled the majority of the signal to be collected from the cellular components. Any water spectral contributions that are present are consistent for all cell spectra, and would primarily overlap with the 1660 cm^{-1} protein peak.

Principal component analysis (PCA)[27] was first applied to the spectra for reducing dimensionality in the data set and extracting combinations of the original spectral markers (i.e. the 1340 Raman channels in a spectrum) that maximize variances in the data. The result was the representation of the original data with a new set of variables, or principal components (PCs), which are a linear combination of the original variables. The principal components (PCs) were used in a supervised classification model, linear discriminant analysis (LDA), for automated group discrimination and data classification. LDA optimizes data classification and separation by maximizing the between-group separation variance while minimizing the within-group variance. Data reduction with PCA prior to LDA is essential to adhere to the LDA criterion that the number of variables must be less than the number of observations. This combination of PCA-LDA has previously been demonstrated in Raman spectral analysis[16, 28-30], yielding excellent data separation. The first criterion for selecting the number of PC scores for LDA analysis was to restrict it to at least two-fold smaller than the number of spectra in the smallest data group in order to prevent any overfitting of the data[16]. In addition, an increasing number of PCs was used until the use of additional PCs did not improve the separation and classification of the data. A total of 6 PC scores were used, each capturing greater than 3% of the data variance.

Blind samples were classified based on their position within the classification model. Given the low number of spectral samples available in this study, a leave-one-out cross-validation method[28,29] was used to evaluate the predictive accuracy of this algorithm. This involved removing one spectrum from the data set, using the remaining n-1 spectra as the training set to define the PC values and the LDA classification model, and then adding the omitted `test' spectrum into the PCA/LDA algorithm to classify this spectrum. This process was repeated for each spectrum (i.e. each spectrum was individually removed and used as a `test' spectrum) and the overall sensitivity values were determined. Only the first six principal components, which account for the majority of the data variance, were used to avoid potential over-fitting. Performing the classification in this manner ensures that the data to be classified does not bias the algorithm.

RESULTS AND DISCUSSION

Characterization of cell identity and intracellular spectral variability

For all three cell types, cells were routinely divided into two populations, one for Raman analysis and the other for conformational analysis of their biological identity. The pluripotency and cardiac identities of the cells recorded were confirmed by positive immunostaining of markers specific for undifferentiated hESCs (SSEA1, Oct4 and TRA-1-60) and hESC-CMs (myosin heavy chain, troponin and tropomyosin)[3,31]. For example, Figure 2 shows an image of a hESC-CM cell cluster that was stained for tropomyosin, a cardiomyocyte marker, to confirm its identity. The identity of hESC-CMs was also confirmed by patch-clamp recordings of cardiac action potentials[3,7] (data not shown).

Since the physical extension of our laser beam (theoretical spot size and confocal parameter are roughly 800 nm and 1.3 um, respectively) is considerably smaller than the cells, experiments were first performed to assess the spectral variability within individual cells. Figure 3 shows representative mean spectra (*black lines*) +/- one SD (*gray lines*) of two hESCs, in which four to five different areas were randomly probed within the cell. This was repeated for ten cells. Figure 3 also shows the average spectra for all ten cells and the overall standard deviation. The low spatial variability of the spectra within a single cell supports our use of a

single spectral measurement per cell for the following comparisons between different cell groups.

The low intracellular spectral variability is somewhat surprising because even though the embryonic cells are small (~ 10-15 μm), the focal volume of the laser beam interrogates only a fraction of the cell. The compartmentalization of biochemical components in a cell should render different Raman spectra when different regions are probed as has been demonstrated in previous Raman imaging studies. To explain our observation, we suggest that multiple factors may be responsible for the overall homogeneity in the spectra. The laser powers that are used are sufficient to enable optical trapping of organelles within a cell. Due to the small size of the cell, consistent trapping[32] of the optically densest part of the cell, the nucleus, is possible. This is supported by the fact that we observe DNA signatures in all the spectra. Although the cells are settled onto the glass substrate, they are not physically immobilized or chemically fixed, which are two common sample preparation techniques done for Raman imaging of cells, so that slight movements of these live cells as they reorient within the laser beam is possible. Also, the long interrogation time needed to obtain spectra with high signal-to-noise ratios may permit the diffusion of cytoplasmic molecules and even organelles through the probe volume. These differences between our current study and the previous Raman imaging studies may explain the spatial homogeneity of our spectral data.

Comparison of Raman spectra of hESCs, hESC-CMs, and FLV-CMs

Figure 4a shows the averaged Raman spectra (*black lines*) +/- one standard deviation (SD) (*gray lines*) recorded from pluripotent hESCs, hESC-CMs and primary FLV-CMs. Each spectrum shown represents the average of 40-50 independent recordings from different cells of the same cell type. For a direct comparison, subtraction of the hESC spectrum from the hESC-CM and FLV-CM spectra were performed to yield the difference spectra shown in Figure 4b. In both cases, reproducible spectral differences were observed. For example, the spectra of FLV-CMs consistently showed lower Raman peak intensities at 785, 811, 854, 1090, 1128, 1320, and 1578 cm^{-1} while having slightly higher intensities in the 937 cm^{-1} region relative to those of hESCs. Based on the peak assignments shown in Table 1, these results indicate that overall, hESCs had higher DNA and RNA content relative to FLV-CMs. For example, the 785 cm^{-1} peak is a clear marker of the cytosine ring vibration of DNA and the 1090 cm^{-1} peak is assigned to the PO^{2-} stretch of the DNA phosphate backbone. The 811 cm^{-1} peak is a well known marker of the O-P-O stretching mode in RNA. The 937 cm^{-1} region is commonly assigned to the protein α-helix carbon backbone stretch. The difference spectra comparing hESCs and hESC-CMs exhibit similar differences, particularly in the Raman bands associated with DNA/RNA, although overall they appeared less pronounced such as at 785, 811, 1090, and 1320 cm^{-1}. In addition, several features such as the 930-980 and the 1320-1450 cm^{-1} regions that are associated with a combination of protein, carbohydrate, and lipid vibrations exhibited slightly different trends in both difference spectra. An additional difference spectrum between the hESC-CMs and the FLV-CMs is shown in Figure 4b to highlight the differences between these two groups.

Of note is that the magnitude of the within-group variability, as reflected in the standard deviation plots in Figure 4a was smaller than the between-group differences for all comparisons, and that the spectral regions where the greatest within-group variability is observed in figure 4a did not match the spectral regions of greatest difference in figure 4b. Similarly, the magnitude of the intracellular signal variability, as shown in Figure 3, was smaller than the between-group differences in Figure 4b. To highlight this further, Figure 4b also contains a difference spectrum comparing two different regions within a single cell to illustrate the low intracellular signal variations, especially when compared to the differences between cell groups. These comparisons emphasize the significance of the between-group

spectral differences that are observed, and substantiate that these differences are not merely artifacts related to cell-to-cell variability.

Consistent with previous studies on differentiated murine ESCs, the spectral differences between self-renewable hESCs and non-proliferative human FLV-CMs observed in this study could result from their different active cell cycles and protein synthesis[21,22]. For example, the, 811 cm[-1] (phosphodiester bond) Raman peak, whose intensity is stronger in hESCs, suggests an elevated mRNA level due to the low translational rate of mRNA in these undifferentiated cells during the early stages of embryogenesis. The lower DNA signals (e.g. 785, 1090 cm[-1]) in the differentiated cells may be indicative of the lower proliferation of the se cells after they have developed into a more matured phenotype.

Separation and classification of Raman data using PCA and LDA

To develop a systematic un-biased method for separating hESCs and hESC-CMs, we next analyzed and compared the individual Raman spectra of hESCs and FLV-CMs to quantify and maximize the differences and separation between the two groups using PCA and LDA. PCA enabled simplification of the spectra by representing the data using only a few new principal component (PC) variables that capture the maximum variance. The first three and six PCs account for approximately 75% and 90% of the total variance in the Raman data, respectively.

To visualize the degree of group separation, the first six PC variables were implemented into a linear discriminant analysis (LDA) to generate a scatter plot (similar to flow cytometry scattergrams) based on the first and second discriminant functions. Figure 5a shows that a clear separation of hESCs and FLVCMs was achieved, primarily by the first discriminant function. Using the same PC axes as defined by the hESC and FLV-CM data, we then calculated the PC scores of hESC-CMs and implemented them into the previously defined LDA routine. By performing the analysis in this manner, hESC-CMs were treated as a blind sample with no prior influence on the definition of the PCA-LDA algorithm. Figure 5b shows that hESC-CMs clearly fell into a region largely separate from hESCs. Similarly, a separation between hESC-CMs and FLV-CMs could also be observed, although some overlap also existed presumably due to their many similar functional aspects[6,7].

To assess the accuracy with which Raman spectroscopy and PCA-LDA can distinguish between hESCs and hESC-CMs, a leave-one-out cross-validation (LOOCV) method was employed to predict the cell identities. The tabulated results in Table 2 show that the sensitivities for identifying hESCs and FLV-CMs were 98% and 96%, respectively. As for hESC-CMs, 66% of the cells tested could be accurately identified as cardiac derivatives with the remainder being classified as undifferentiated hESCs.

The combination of PCA and LDA and the use of the LOOCV sufficed to optimally separate and classify the cell groups based on the Raman differences. Analysis and representation of the data in LDA scatter plots is highly useful for comparing individual cell data and visualizing the group separation. This approach leads to the formation of defined group clusters for separating hESCs, hESC-CMs and FLV-CMs from each other. Based on their Raman spectra, the identities of the three cell types can be predicted with high statistical confidences. Almost 100% was achieved for hESCs and FLV-CMs. The relatively lower accuracy in identifying hESC-CMs as fetal CMs is not surprising considering the heterogeneity and the embryonic nature of hESC-CMs compared to fetal CMs. It is known that the beating areas can include cardiomyocytes at different stages of maturity[6]. Nevertheless, the HES2 line was chosen for this study because it has previously been shown to favor the derivation of ventricular CMs with a yield of >80% (compared to 18% and 2% for atrial and pacemaker CMs, respectively[7]). In this study, only those regions with consistent beating properties were chosen for Raman analysis to ensure that a high percentage of the cells in the hESC-CMs are indeed CMs and not

undifferentiated hESCs or other types of differentiated cells. In the future, more stringent separation criteria will need to be used in order to ensure that the selected hESC-CMs are free from hESCs. Interestingly, hESC-CMs nicely fell between hESCs and FLV-CMs, hinting that the first discriminant function may be a useful indicator for assessing the extent of CM maturation.

Assessing cell viability

A major advantage of Raman spectroscopy is that it can potentially allow cells that have been identified and selected to remain viable and be used post-analysis. In our study, we observed no deleterious effect of the laser exposure to the cells using 785 nm light based on several observations. First, no degradation of the Raman signal intensity was observed even after several minutes of laser exposure (data not shown), which suggests that there is no cellular photodamage that is occurring. Previous studies using 514.5 nm light to irradiate cells have shown a decrease in intensity of several DNA Raman peaks[33], which is a good indicator of laser damage to the cell. Secondly, the laser exposure did not appear to perturb the beating of the hESC-CM clusters since no change in the beating rhythm was visually observed under white light microscopy. Thirdly, cells were stained with trypan blue following exposure to the 785 nm beam for more than five minutes. No permeation of the dye into the exposed cells was observed, indicating that the cell retained its membrane integrity. These results indicate that the Raman spectra that have been acquired from these cells are signatures that accurately represent the biochemical composition and state of a live, viable cell and that overall there are no drastic biological changes to the laser exposed cells. Whether or not the laser exposure affects the long-term pluripotency of hESCs or the development of the hESC-CM lineages remains to be seen, which would require additional growth studies, gene expression assays, or functional assays.

CONCLUSION

A major pragmatic implication of this current work is the separation of hESCs and their cardiac derivatives by *intrinsic* Raman signatures, eliminating the need for extrinsic genetic or biochemical interventions (e.g., ectopic expression of a reporter protein, antibody labeling, etc). As such, the cells remain viable and essentially unaltered even after their analysis, and can be readily recovered by physical manipulation and sorting of the cells, i.e. by using laser tweezers. Such a non-labeling spectroscopic approach is particularly useful for cells such as CMs that lack lineage-specific surface markers for physical separation by conventional means (e.g., magnetic sorting or column-based purification). Thus, our results provide proof-of-concept data that Raman spectroscopy and multivariate statistical methods, when further developed and optimized, can lead to a systematic method for identifying the most pluripotent hESCs for high-quality propagation, for ridding the potential tumorigenicity of hESC-derived grafts, and for enriching hESC-CMs. Based on the data presented here, we conclude that Raman-based spectroscopic methods present a label-free approach for separating undifferentiated hESCs and their cardiac derivatives based on objective biophysical criteria. Although Raman sorting techniques remain in their infancy, several recent studies have demonstrated that Raman activated cell sorting is feasible and can be practically implemented. These studies have based their technology using various Raman schemes, including laser tweezers Raman spectroscopy (LTRS) for analysis and physical sorting of cells[34] and coherent anti-Stokes Raman scattering (CARS) for faster analysis[35]. These technologies are expected to be critical for future applications in label-free Raman sorting of stem cells and their lineages.

Future studies are planned to possibly improve the identification and separation of hESCs and their derived cardiomyocytes. Currently, Raman spectroscopy suffers from low signal intensities inherent to the Raman scattering process. Improvements to the signal-to-noise ratio,

such as by using nonlinear Raman processes such as CARS to increase the signal[36,37], may improve the detection of subtle Raman differences between cell types. Additionally, improvements can be made to the classification of the Raman data by directly correlating the Raman spectra to fluorescence staining to confirm cell phenotype of individual cells. These studies are currently planned for the near future. Also, a recent study by Inya-Agha et al.[38] used Raman spectroscopy to monitor the dynamic changes of single cardiac myocytes during their contraction cycle using Raman spectroscopy. This opens up tremendous possibilities to use dynamic cellular function as an additional parameter to further characterize, compare, and discriminate between hESC-derived and mature cardiomyocytes.

ACKNOWLEDGMENT

REFERENCES

(1). Li RA, Moore JC, Tarasova YS, Boheler KR. Journal of Stem Cells 2006;1:109–124.

(2). Siu CW, Moore JC, Li RA. Cardiovasc Hematol Disord Drug Targets 2007;7:145–152. [PubMed: 17584049]

(3). Moore JC, van Laake LW, Braam SR, Xue T, Tsang SY, Ward D, Passier R, Tertoolen LL, Li RA, Mummery CL. Reprod Toxicol 2005;20:377–391. [PubMed: 15967632]

(4). Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Science 1998;282:1145–1147. [PubMed: 9804556]

(5). Xue T, Cho HC, Akar FG, Tsang SY, Jones SP, Marban E, Tomaselli GF, Li RA. Circulation 2005;111:11–20. [PubMed: 15611367]

(6). He JQ, Ma Y, Lee Y, Thomson JA, Kamp TJ. Circ Res 2003;93:32–39. [PubMed: 12791707]

(7). Mummery C, Ward-van Oostwaard D, Doevendans P, Spijker R, van den Brink S, Hassink R, van der Heyden M, Opthof T, Pera M, de la Riviere AB, Passier R, Tertoolen L. Circulation 2003;107:2733–2740. [PubMed: 12742992]

(8). Kehat I, Kenyagin-Karsenti D, Snir M, Segev H, Amit M, Gepstein A, Livne E, Binah O, Itskovitz-Eldor J, Gepstein L. J Clin Invest 2001;108:407–414. [PubMed: 11489934]

(9). Xu C, Police S, Rao N, Carpenter MK. Circ Res 2002;91:501–508. [PubMed: 12242268]

(10). Kehat I, Khimovich L, Caspi O, Gepstein A, Shofti R, Arbel G, Huber I, Satin J, Itskovitz-Eldor J, Gepstein L. Nature Biotechnology 2004;22:1282–1289.

(11). Huber I, Itzhaki I, Caspi O, Arbel G, Tzukerman M, Gepstein A, Habib M, Yankelson L, Kehat I, Gepstein L. Faseb Journal 2007;21:2551–2563. [PubMed: 17435178]

(12). Puppels GJ, Demul FFM, Otto C, Greve J, Robertnicoud M, Arndtjovin DJ, Jovin TM. Nature 1990;347:301–303. [PubMed: 2205805]

(13). Baena JR, Lendl B. Current Opinion in Chemical Biology 2004;8:534–539. [PubMed: 15450497]

(14). Motz JT, Fitzmaurice M, Miller A, Gandhi SJ, Haka AS, Galindo LH, Dasari RR, Kramer JR, Feld MS. Journal of Biomedical Optics 2006;11

(15). van de Poll SWE, Romer TJ, Puppels GJ, van der Laarse A. Journal of Cardiovascular Risk 2002;9:255–261. [PubMed: 12394318]

(16). de Jong BWD, Bakker TC, Maquelin K, van der Kwast T, Bangma CH, Kok DJ, Puppels GJ. Analytical Chemistry 2006;78:7761–7769. [PubMed: 17105169]

(17). Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, Feld MS. Proceedings of the National Academy of Sciences of the United States of America 2005;102:12371–12376. [PubMed: 16116095]

(18). Chan JW, Taylor DS, Lane SM, Zwerdling T, Tuscano J, Huser T. Analytical Chemistry 2008;80:2180–2187. [PubMed: 18260656]

(19). Schut TCB, Witjes MJH, Sterenborg HJCM, Speelman OC, Roodenburg JLN, Marple ET, Bruining HA, Puppels GJ. Analytical Chemistry 2000;72:6010–6018. [PubMed: 11140770]

(20). Utzinger U, Heintzelman DL, Mahadevan-Jansen A, Malpica A, Follen M, Richards-Kortum R. Applied Spectroscopy 2001;55:955–959.

(21). Notingher L, Bisson I, Polak JM, Hench LL. Vibrational Spectroscopy 2004;35:199–203.

(22). Notingher I, Bisson I, Bishop AE, Randle WL, Polak JMP, Hench LL. Analytical Chemistry 2004;76:3185–3193. [PubMed: 15167800]

(23). Krafft C, Salzer R, Seitz S, Ern C, Schieker M. Analyst 2007;132:647–653. [PubMed: 17592583]

(24). Konorov SO, Glover CH, Piret JM, Bryan J, Schulze HG, Blades MW, Turner RFB. Analytical Chemistry 2007;79:7221–7225. [PubMed: 17691751]

(25). Kim BS, Lee CCI, Christensen JE, Huser TR, Chan JW, Tarantal AF. Stem Cells and Development 2008;17:185–198. [PubMed: 18211228]

(26). Lieber CA, Mahadevan-Jansen A. Applied Spectroscopy 2003;57:1363–1367. [PubMed: 14658149]

(27). Wold S, Esbensen K, Geladi P. Chemometrics and Intelligent Laboratory Systems 1987;2:37–52.

(28). Molckovsky A, Song LMWK, Shim MG, Marcon NE, Wilson BC. Gastrointestinal Endoscopy 2003;57:396–402. [PubMed: 12612529]

(29). Notingher I, Jell G, Lohbauer U, Salih V, Hench LL. Journal of Cellular Biochemistry 2004;92:1180–1192. [PubMed: 15258901]

(30). Notingher L, Jell G, Notingher PL, Bisson I, Tsigkou O, Polak JM, Stevens MM, Hench LL. Journal of Molecular Structure 2005;744:179–185.

(31). Wang K, Xue T, Tsang SY, Van Huizen R, Wong CW, Lai KW, Ye ZH, Cheng LZ, Au KW, Zhang J, Li GR, Lau CP, Tse HF, Li RA. Stem Cells 2005;23:1526–1534. [PubMed: 16091557]

(32). Ashkin A, Dziedzic JM, Bjorkholm JE, Chu S. Optics Letters 1986;11:288–290.

(33). Puppels GJ, Olminkhof JHF, Segersnolten GMJ, Otto C, Demul FFM, Greve J. Experimental Cell Research 1991;195:361–367. [PubMed: 2070819]

(34). Lau AY, Lee LP, Chan JW. Lab on a Chip 2008;8:1116–1120. [PubMed: 18584087]

(35). Wang HW, Bao N, Le TT, Lu C, Cheng JX. Optics Express 2008;16:5782–5789. [PubMed: 18542688]

(36). Cheng JX, Xie XS. Journal of Physical Chemistry B 2004;108:827–840.

(37). Shi KB, Li P, Liu ZW. Applied Physics Letters 2007;90

(38). Inya-Agha O, Klauke N, Davies T, Smith G, Cooper JM. Analytical Chemistry 2007;79:4581–4587. [PubMed: 17511420]
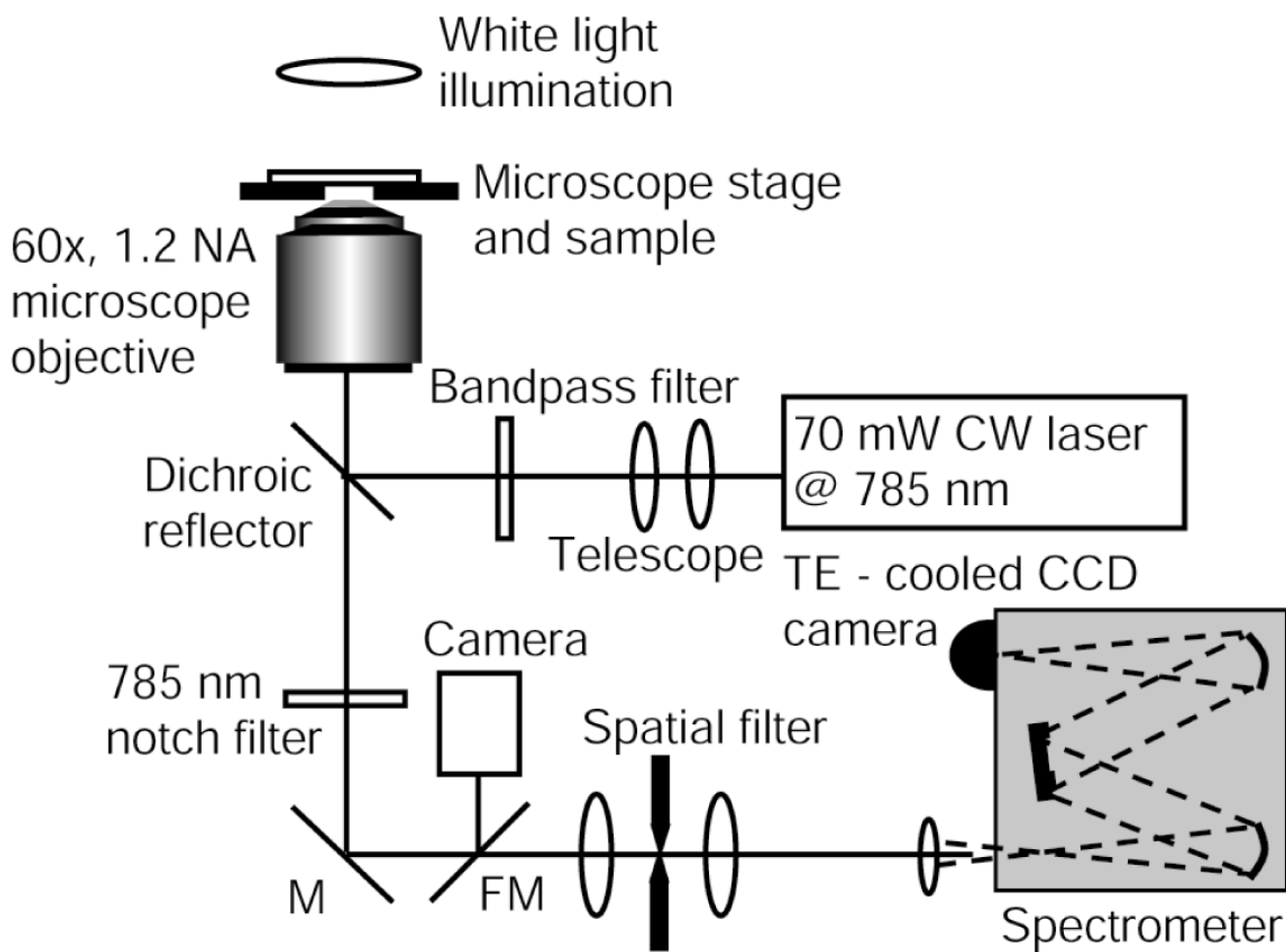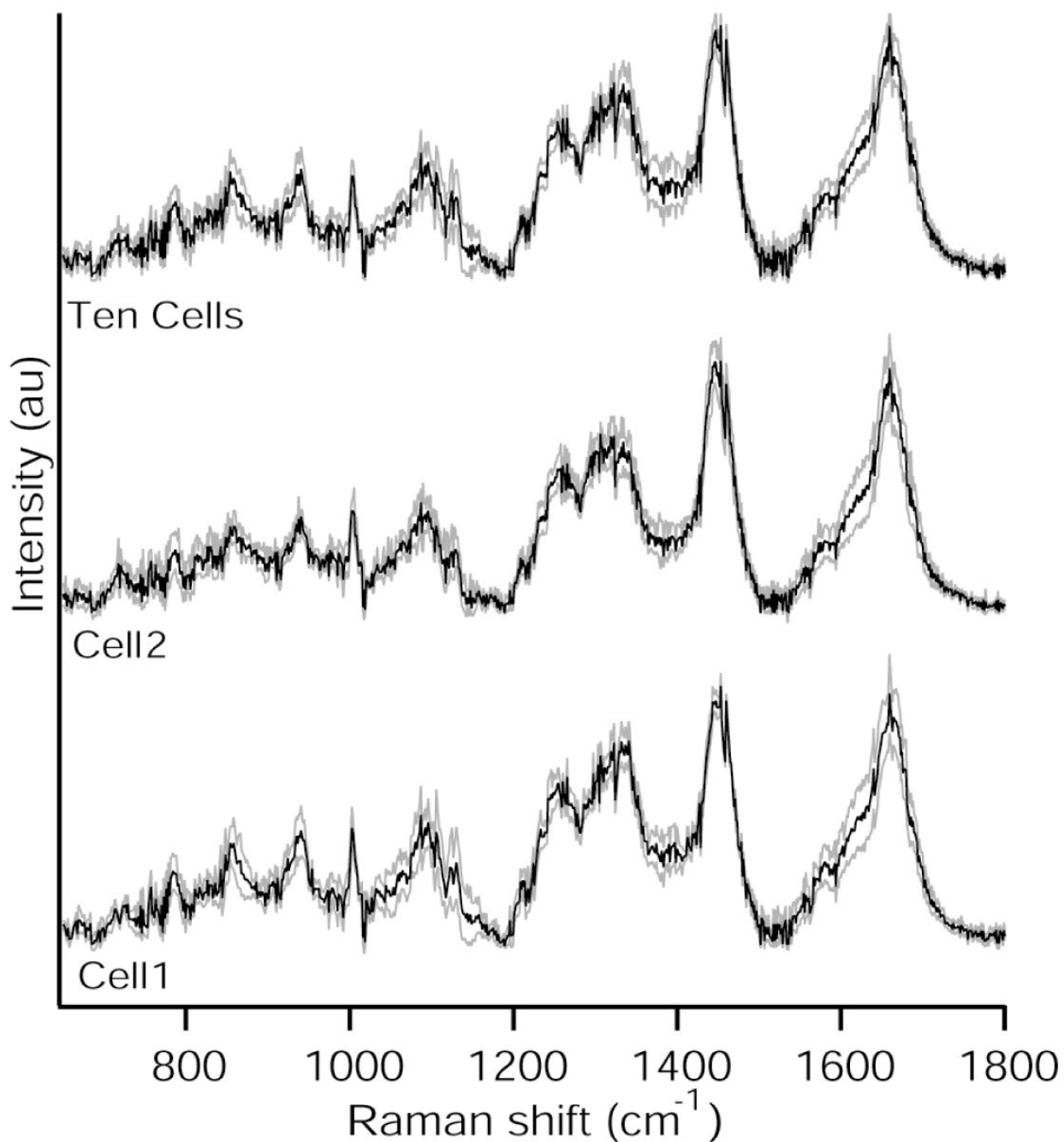
Figure 1.
Schematic of the Raman confocal microscope system. A 785 nm CW laser beam delivered to an inverted microscope is used for Raman analysis of cells. Raman spectra are acquired by epi-detection of the signals, sent through a confocal pinhole, and delivered into a dispersive spectrometer and CCD detector. (M - mirror, FM - flip mirror)
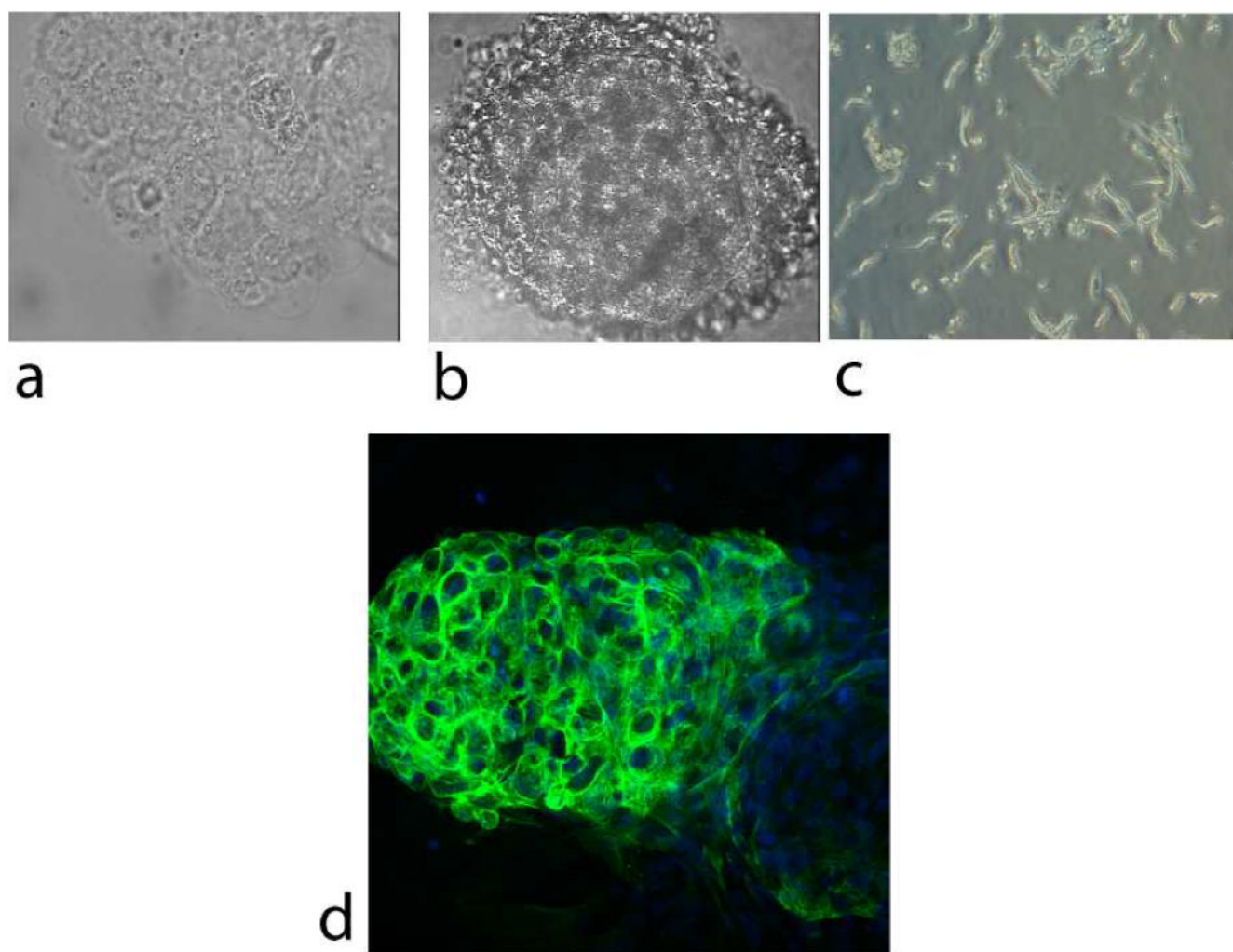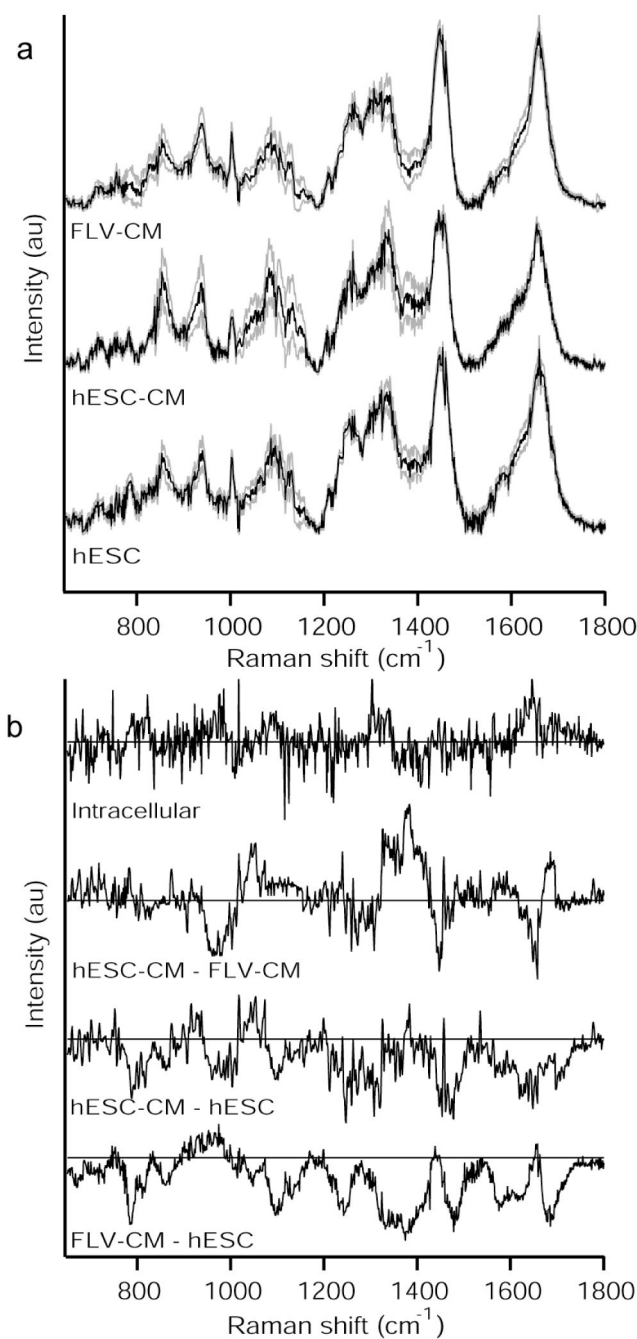
Figure 2.
White light images showing representative samples of (a) hESCs, (b) hESC-CMs and (c) FLV-CMs that were analyzed by Raman spectroscopy. hESC-CMs were visually identified as beating areas. To further confirm their identity, areas were dissected and stained for tropomyosin (green), a cardiomyocyte marker, and Hoechst nuclear counterstain (blue). (d) Cells in the beating cluster stained positive for tropomyosin. Image was taken using a laser scanning confocal microscope.
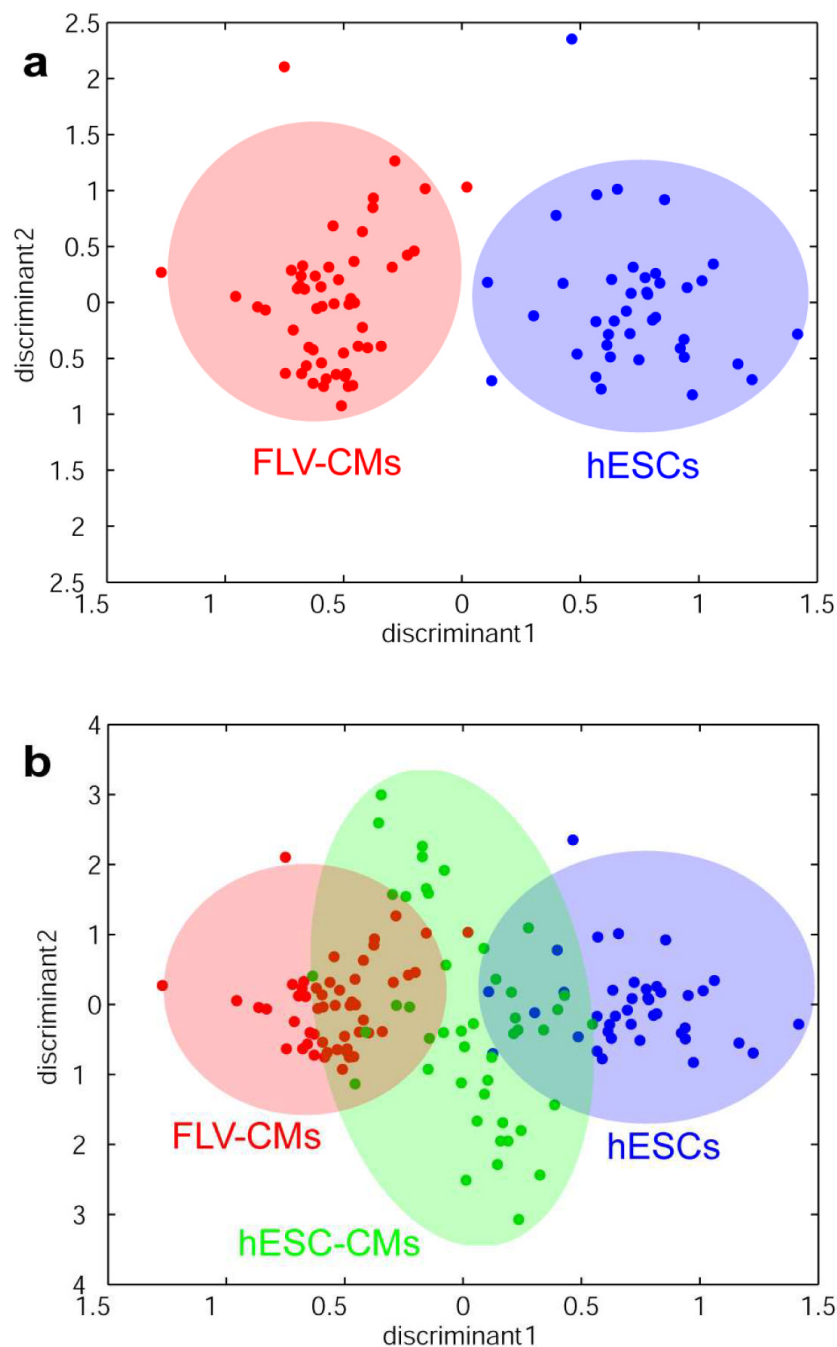
Figure 3.
Mean Raman spectra obtained from individual hESCs by probing five different locations randomly within the cell and averaging the five spectra. Results from two individual cells are shown. 10 different cells were probed in this same manner and the fifty spectra were averaged to yield the mean Raman spectra. Black lines indicate the average spectra and the gray lines delineate one standard deviation. The results indicate that spatial variability of the Raman spectra is relatively low, in particular when compared to the magnitude of the difference spectra. Spectra are offset in the vertical axis for clarity.

Figure 4.
(a) Mean Raman spectra of hESC, hESC-CM, and FLV-CM cells. Black lines indicate the average spectra and the gray lines delineate one standard deviation. Each spectrum is an average of 20 to 40 different cells. (b) Difference spectra obtained by subtracting the mean spectra of hESC from FLV-CM and hESC-CM show similar spectral differences for several peaks (see text for details) but less pronounced for the hESC-CM spectra. Also shown is a difference spectra between hESC-CM and FLV-CM to highlight the spectral differences observed between these two cell groups. For comparison, a difference spectrum between two points within an individual hESC cell is plotted to highlight the low intracellular variability compared to the differences between cell groups. Spectra are offset in the vertical axis for clarity.

Figure 5.
(a) LDA plot comparing FLV-CMs and hESCs shows a clear separation of the two groups. Cells can be classified with a ~100% accuracy. (b) LDA plot of the position of the hESC-CMs relative to the FLV-CMs and hESCs. hESC-CMs fall in between, possibly indicating a larger heterogeneity and the embryonic nature of this sample. Roughly 70% of these cells are classified as CMs. Shaded regions overlaid on the data points serve as visual guides to the eye.

Table 1

Raman peak frequencies and their assignments

Raman frequency (cm^{-1})	Assignment[18,22]
717	CN$^+$(CH$_3$)$_3$ str (l)
729	A
760	Trp ring breath. (p)
785	U,C,T ring breath., O-P-O str.
828	O-P-O asym. str., Tyr ring breath. (p)
854	Tyr ring breath. (p)
937	C-C bk str α-helix (p), C-O-C glycos (c)
980	C-C bk str β-sheet (p), =CH bend (l)
1004	Phe sym. ring breath. (p)
1033	Phe C-H in-plane (p)
1093	PO$_2^-$ str., chain C-C str. (l), C-O str. (c), C-C str. (c)
1128	C-N str. (p)
1158	C-C/C-N str. (p)
1209	Phe C-C$_6$H$_5$ str. (p), Trp (p)
1257	A,T, amide III (p), =CH bend (l)
1320	G, CH def(p)
1342	A,G, CH def (p), CH def (c)
1450	G,A, CH def (d), (p), (l), (c)
1578	G,A
1610	Tyr C=C, Trp C=C, Phe C=C (p)
1658	Amide I (p), C=C str (l)

Abbreviations: p: protein, l: lipid, d: DNA/RNA, A,G,T,C,U: ring breathing modes of DNA/RNA, str: stretching, breath: breathing, Phe: phenylalanine, Tyr: tyrosine, Trp: tryptophan, bk: backbone

Table 2

Cell classification results based on leave-one-out cross validation

Cell type	Predicted cell type based on PCA/LDA of Raman spectra		Sensitivity
	hESCs	CMs	
hESCs	39	1	98%
FLV-CMs	2	51	96%
hESC-CMs	16	31	66%