Supplementary Text S1. Why is Background Subtraction Needed in Cellular Raman Spectroscopy

Raman spectrum measurement of living cells is often practically low in reproducibility in technical replicate due to slight deviation in optical adjustment and/or photodamage by long and strong laser irradiation. In addition, the scattering light detected by the sensor contains not only Raman scattering light from cells but also autofluorescence of the cells, the Raman scattering light from the glass substrate and the medium, and multiple scattering between intracellular micro-structures and/or between the cell and the substrate, which are undesired background signals. These problems disturb the investigation of cellular heterogeneity. To solve them, the background subtraction process to obtain a pure cell-derived Raman spectrum from the observed spectrum have been developed, such as first/second-order derivatives, frequency-domain filtering, polynomial fitting, wavelet transformation, and so on (s1). These are primarily methods for autofluorescence removal and are not intended to remove substrate/medium-derived background spectra. Because the background spectrum can be reduced by selecting quartz as the substrate and a medium containing less substances, it is not often discussed. However, since hiPSCs are sensitive to surrounding environment, such that optimizations for observation may perturb a hiPSC condition and degrade its quality. It is needed to discuss the removal of substrate/medium-derived spectrum from observed spectrum to achieve the aim.

In general, the background signal can be measured before or after in the absence of cells. The cell-derived spectrum is obtained by subtracting the pre-measured background spectrum from the overall observed spectrum (Figure S1A). At this time, the background spectrum should be appropriately scaled due to the various optical phenomena, and the automation of the scaling is proposed (s2). This method assumes full technical reproducibility of focal position control. Practically, some technical deviation remains, such as the parallelism of the sample stage and the reproducibility of the focal position control between the background observation ant the cell observation. Also, the throughput of spectrum acquisition is reduced by half. It is preferable that the background spectrum is acquired at the same time as the cell observation. The average background signal can be measured from the signals emitted from substrates other than the cells in field of view (Figure S1B). The cell-derived spectrum is acquired correctly if under less optical aberration, such as using stage-scanning type confocal system. The background signal in the field of view is not uniform because there are optical aberrations in many cases. Hence, the background spectrum is estimated from signals in the vicinity of the cells (Figure S1C). Still, it is difficult to completely estimate the background spectrum because the multi-scattering effects are unpredictable. The resulting cell-derived spectra are slightly biased depending on the sample dish, experiment date, and analyst.

Supplementary Text S2. Procedure of Background Prediction for Raman Spectrum Measurement

A line-illumination was formed on 125 μm corresponding to 400 pixels as X-axis with a cylindrical lens, and was scanned only 10 steps of 3.35 μm/step as Y-axis (see Figure 1A in the main text). Notably, the spectral intensity depended on the position due to optical aberration by the cylindrical lens: The X-position dependence was observed that the center of the field of view was irradiated more strongly than both ends in the line illumination (Figure S2A, *p1, p2 and p3*). Plotting average intensity on the Y-axis at each wavenumber as a function of the X-position, ridges depending on the presence or absence of cells were

confirmed at the cell-specific wavenumbers (Figure S2B, *red, 755 cm^{-1} and green, 1653 cm^{-1}*), whereas the substrate-specific wavenumber showed a Gauss-like distribution derived from the radiant intensity profile of the used laser (Figure S2BC, *blue, 495 cm^{-1}*). First, for segmentation between cell-area and noncell-area, cell-specific peaks were crudely extracted by subtracting the 41-point averaged spectrum from the 5-point averaged spectrum (Figure S2C, *upper*). We used the cell-specific peak at 1438 cm^{-1}, attributed to CH2 scissoring vibration mainly in lipids, among several indicative peaks for the segmentation (Figure S2D, *lower*). The 1006 cm^{-1} peak, attributed to phenyl-ring breathing vibration mainly in proteins, was also applicable as the indicator. The threshold values for the segmentation were set separately for cell area and non-cell area, and the threshold for cell area should always be higher than the other (Figure S2E). This two thresholds setup contributed to prevent the contamination of cell area signal in the non-cell area and to eliminate too low signal intensity area (Figure S2E, *upper, double asterisks*). Single cells were separately labeled by a watershed algorithm (s3) and then confirmed or corrected by the human eye (Figure S2E, *upper, #1–#6*). Then, we obtained the averaged spectrum within each cell area was obtained (Figure S2E, *lower*).

Raman spectra obtained with the procedure mentioned above were a mixture of cell-derived and substrate/medium-derived spectra. The underlying concept of the conventional methods shown in Figure S1 has been to subtract the signal acquired in an area without cells from the signal with cells. Following this concept basically, we devised an algorism to predict spectrum derived from non-cell contents on a cell-exciting area. The first step is to plot the averaged intensity on the Y-axis at all each wavenumber of the non-cell area segmented with the threshold shown in Figure S2D (Figure S2F). The second is to estimate the background signal intensity at a given wavenumber as a function of X-position by approximating the X-position dependence with a polynomial function at the wavenumber (Figure S2F, *solid lines*). By doing this process at all wavelengths, the background spectrum on the cell areas can be predicted (Figure S2G). Finally, the cell-derived spectrum could be extracted by subtracting the predicted back-ground spectrum from the spectrum obtained on the cell area (Figure S2H). Herein, there was almost no effective aberration on Y-axis because the scanning range on Y-axis was limited to 33.5 μm. Therefore, we used the average spectrum along Y-axis on the X-position dependence at each wavenumber. Even if the aberration is observed on long ranged Y-axis, the background spectrum can be similarly predicted by two-dimensionally fitting the non-cell area with a polynomial plane. Either way, our background subtraction method requires image information that includes non-cellular regions, causing degradation of throughput of data collection. It is essential to improve the throughput to measure the cell heterogeneity more quantitatively.

Supplementary Text S3. Procedure of Autofluorescence Subtraction and the Later

The cell-derived spectrum obtained with the present procedure contains Raman spectrum and autofluorescence spectrum. Of the proposed methods (s1), polynomial fitting is the most common autofluorescence removal technique in a wide range of bioapplications (s4). The autofluorescence spectrum was determined as the lower envelop predicted by a least-squares approximation with a polynomial function, for example, quintic function (Figure S3A). First, the whole spectrum is fitted with

a polynomial function. The data whose intensity is lower than the fitted curve is collected as a new data. Next the new data is again fitted with a different polynomial function. The 25~200 repeat of this process gradually eliminates the higher-frequency Raman peaks from the initial spectrum (s4). Thus, The Raman spectrum was estimated by subtracting the predicted envelop from the initial spectrum (Figure S3B). The number of repetitions is not deterministic but case-by-case and relies on analyst intervention. Automation of the polynomial fitting has been developed (s5, s6) but has been practically difficult to uniquely or automatically determine the number of repetitions in condition of low signal to noise ratio such as in cell observation. Instead, though the automation of the fitting using multiresolution wavelet transformation was proposed (s7), the resulted spectrum depends on the wavelet threshold and the spectra resolution, generating a new issue of the selection of the appropriate threshold and resolution. Here, we applied rolling-circle algorism, because the only parameters to be set are diameter and ellipticity, eliminating necessary of fine-tune for each data (s8). Assuming that a circle rolls on an uneven surface, the trajectory of the ball surface is the envelope (Figure S3C). Thereby, the envelope can be deterministic independent of analyst or case. In this study, the diameter and ellipticity were set so that the sharp Raman peaks would be enhanced (Figure S3D).

The Raman spectroscopy needs wavenumber calibration. Usually, the obtained spectra were calibrated with the peak wavenumbers of pure materials, such as ethanol. To avoid the risk that error occurs between the measurements of ethanol and cell samples, we calibrated the all spectra with the peak wavenumbers of cell specific peaks, 755 cm^{-1}, 1131 cm^{-1}, and 1585 cm^{-1}, of themselves. Finally, the spectrum data was interpolated from 500 to 1800 cm^{-1} with 1 cm^{-1} interval and processed with a 5-points moving average to reduce random noise.

Supplementary Text S4. *Pre-Investigation of Reproducibility in Technical Replicate*

Whichever the method shown in Figure S1 is used, a threshold setting is needed to segment cell-area from non-cell area. We investigated dependency of the threshold values on the background subtraction. Three sets of threshold pair for cell and for non-cell area were prepared (Figure S4A). The basic threshold values were set based on standard deviation of pixel intensities of non-cell area segmented by Otsu's image segmentation algorithm (s9). In the *Set #1*, the cell area and the non-cell area were segmented with the same threshold, three-folds of the standard deviation. In the *Set #2*, the cell area was segmented with the threshold of three-folds of the standard deviation and the non-cell area was with the threshold of two-folds of the standard deviation. In the *Set #3*, the cell area and the non-cell area were segmented with the same threshold, two-folds of the standard deviation. The different threshold sets resulted in different spectral shape in the conventional method shown in Figure S1C (Figure S4B). The present analysis procedure successfully reduced the threshold dependence between the Set #1 and the Set #2 (Figure S4C, *red and blue*). It is thought that this is because the background spectrum prediction with the numerical approximation was not so easily affected by the threshold for non-cell area. The threshold for cell area affected the spectral shape, but the effect was still smaller than the conventional method (Figure S4C, *blue and green*).

We collected 52 single-cell spectra of 253G1 hiPSCs in four culture dishes, each dish included seven field of views, each FOV include one to six cells, and asked two different analysts to independently measure the average spectrum for each dish

with the present procedure (Figure S5A). The threshold for cell division was not clearly different between dishes or between analysts, even though it depended on the analyst's senses (Figure S5B). The variation among the analysts was smaller than the variation among the culture dishes (Figure S5C) Visualization of the difference from the overall average spectrum as a heat map clearly exhibited the independency of analyst of the present procedure (Figure S5D).

One more important origin of the low reproducibility in cellular Raman spectroscopy has been optical aberrations. Since optical aberrations directly affected obtained spectrum in Raman microscopy, causing the deviation between experimental day, precise optical adjustments are necessary for each experiment to ensure the reproducibility. Especially when using a home-built microscope, the precision and accuracy of the adjustment de-pends on skill of microscopist because the optical parts are set and adjusted by his/her hands and eyes. Moreover, deviation in the optical adjustment may be accidentally generated due to relaxation of each screw after fine adjustment even under the control of room temperature and humidity. Such an accidental optical deviations are often overlooked be-cause microscopists and biological experimenters are different in most cases, which is one of the main causes of poor reproducibility in biological Raman measurements. Here, we intentionally shifted the optical axis adjustment to investigate its effect on the analysis. The misalignment of the optical axis affected not only the intensity distribution but also the wavenumber characteristics (Figure S6A). Though the X-position dependency of intensity would be identically approximated with a Gaussian distribution, the optical aberration distorted the distribution (Figure S6A, *right*). That was why we used polynomial function as a fitting function instead of Gaussian function. The present procedure could extract almost the same cell-derived spectrum regardless of the optical adjustment difference, be-cause of the prediction of the background signal at each wavenumber (Figure S6B). The average spectra for each dish were almost identical between the two optical aberration conditions (Figure S6C), indicating that the present procedure eliminated the effect of the aberration.

Supplementary References

[s1] Schulze, G., Jirasek, A., Yu, M. M. L., Lim, A., Turner, R. F. B., Blades M. W. Investigation of selected baseline removal techniques as candidates for automated implementation. Appl. Spectrosc. 59, 545–574 (2005). https://doi.org/10.1366/0003702053945985

[s2] Sugawara, T., Yang, Q., Nakabayashi, T., Morita, S. I. A proposal for automated background removal of Bio-Raman data. Anal. Sci. 33, 1323–1325 (2017). https://doi.org/10.2116/analsci.33.1323

[s3] Malpica, N., de Solorzano, C. O., Vaquero, J. J., Santos, A., Vallcorba, I., Garcia-Sagredo, J. M., et al. Applying watershed algorithms to the segmentation of clustered nuclei. Cytometry 28, 289–297 (1997). https://doi.org/10.1002/(sici)1097-0320(19970801)28:4<289::aid-cyto3>3.0.co;2-7

[s4] Lieber, C.A., Mahadevan-Jansen, A. Automated method for subtraction of fluorescence from biological Raman spectra. Appl. Spectrosc. 57, 1363–1367 (2003). https://doi.org/10.1366/000370203322554518

[s5] Zhao, J., Lui, H., McLean, D. I., Zeng, H. Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy. Appl. Spectrosc. 61, 1225–1232 (2007). https://doi.org/10.1366/000370207782597003

[s6] Brennan, J. F., Wang, Y., Dasari, R. R., Feld, M. S. Near-infrared raman spectrometer systems for human tissue studies. Appl. Spectrosc. 51, 201–208 (1997). https://doi.org/10.1366/0003702971940134

[s7] Cai, T. T., Zhang, D., Ben-Amotz, D. Enhanced chemical classification of Raman images using multiresolution wavelet transformation. Appl. Spectrosc. 55, 1124–1130 (2001). https://doi.org/10.1366/0003702011953289

[s8] Brandt, N. N., Brovko, O. O., Chikishev, A. Y., Paraschuk, O. D. Optimization of the rolling-circle filter for Raman background subtraction. Appl. Spectrosc. 60, 288–293 (2006). https://doi.org/10.1366/000370206776342553

[s9] Otsu, N. A threshold selection method from gray level histogram. IEEE Trans. Syst. Man Cybern. 9, 62–66 (1979). https://doi.org/10.1109/TSMC.1979.4310076
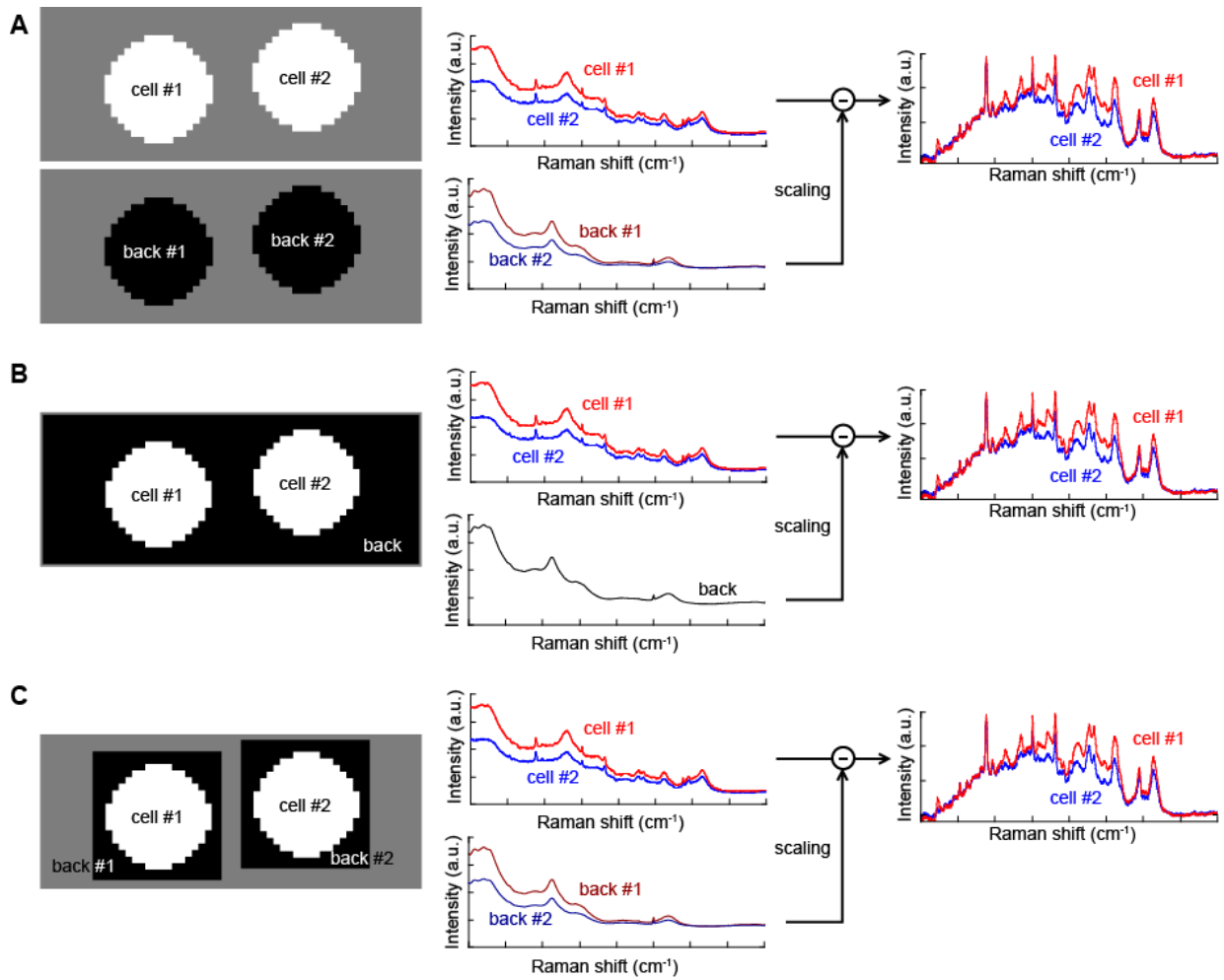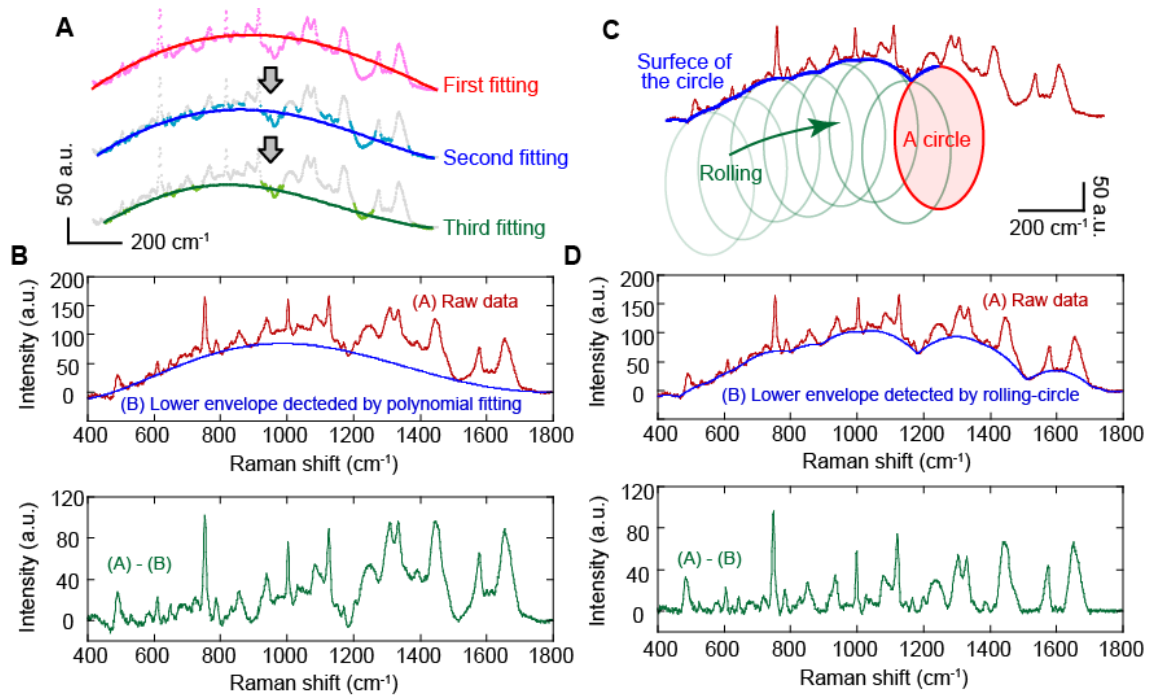
Supplementary Figure S1. Explanatory diagram of the current background subtraction methods. (**A**) A method for obtaining the average spectra of a cell area and the same area without the cell. (**B**) A method of obtaining the average spectra in each of a cell area and a non-cell area. (**C**) A method of obtaining the average spectra of each of a cell area and a non-cell area in the vicinity of the cell. In many cases, the background signal contained in the spectrum obtained in the cell area is larger than in the absence of cells. Therefore, the background spectrum is linearly scaled to zero intensity in the silent region when subtracted.
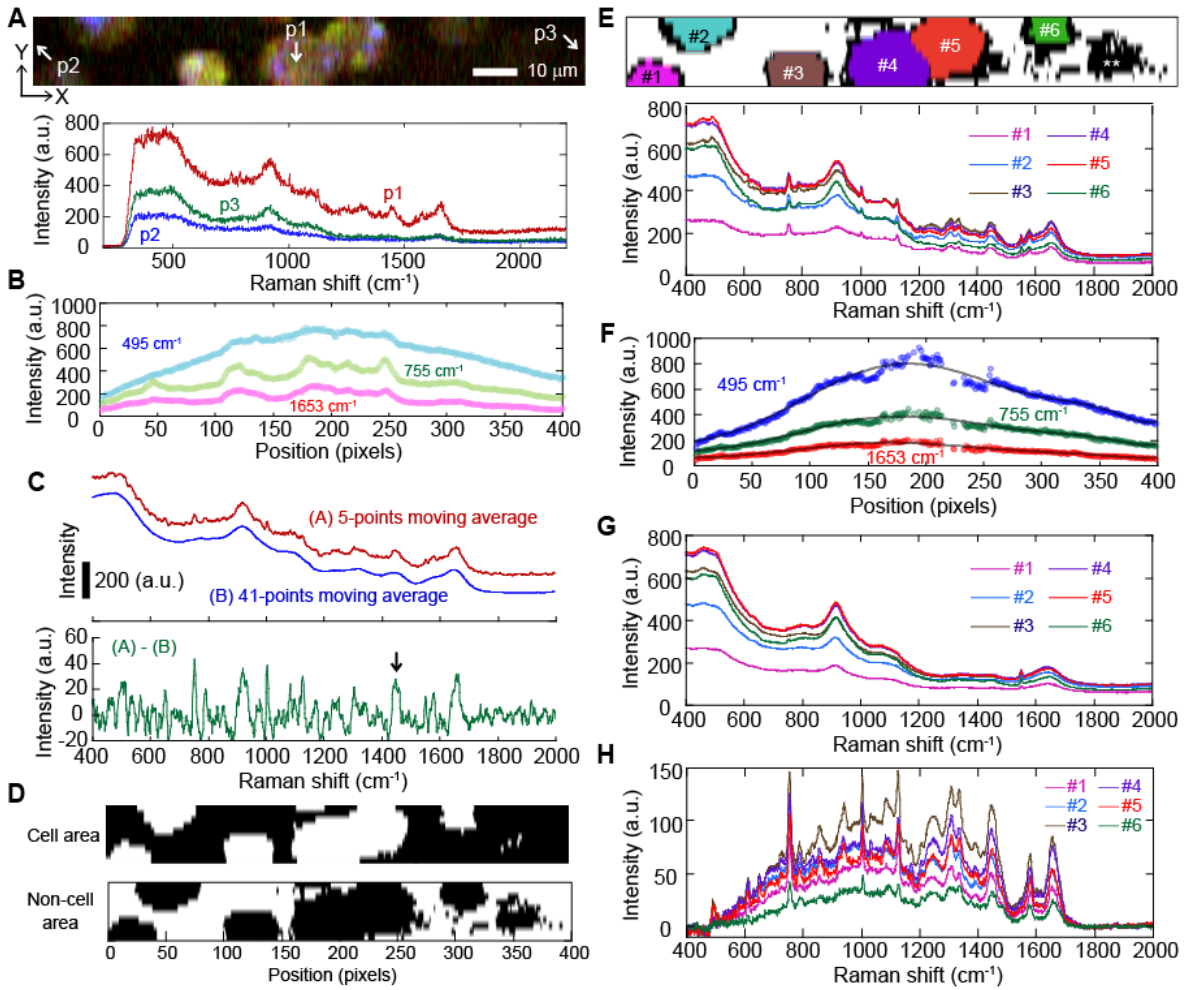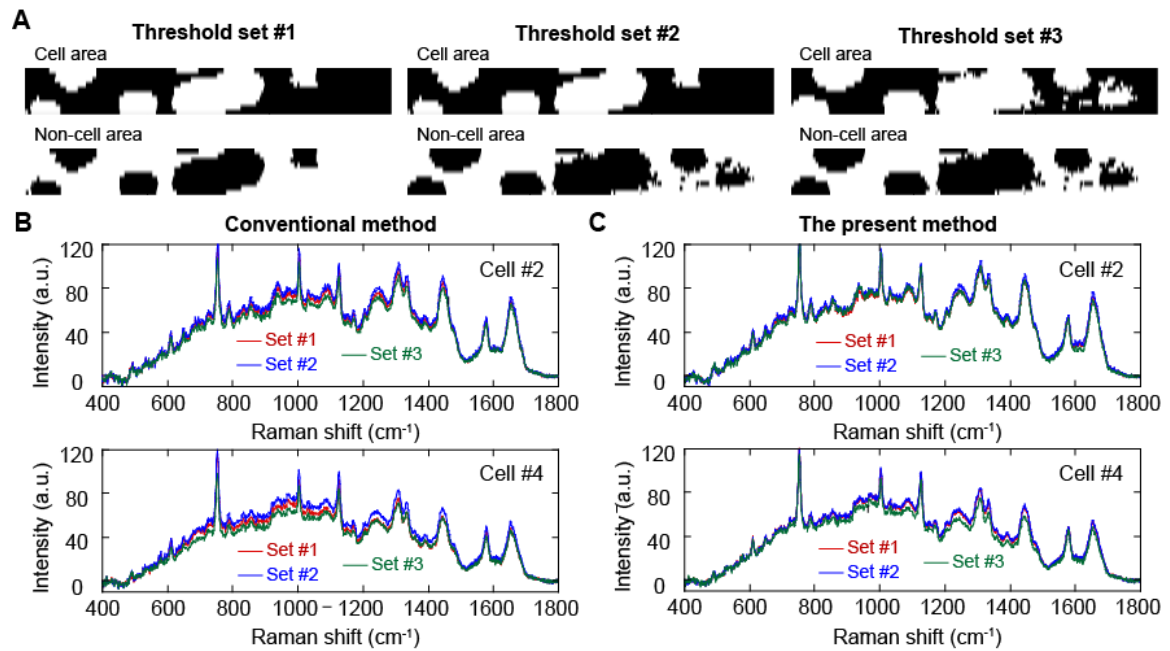
Supplementary Figure S2. Extraction of cell-derived spectrum. (**A**) A pseudo-color image with cytochrome-C (755 cm^{-1}) colored in blue, lipids (1653 cm^{-1}) in green, and proteins (1438 cm^{-1}) in red (*upper*) and typical Raman Spectra of each pixel in a cell area (*p1*) and non-cell areas at the both ends (*p2* and *p3*)(*lower*). The axis direction of the line-formed illumination is the X-axis, and the scanning axis is the Y-axis. The scale bar is 10 μm. (**B**) X-position dependence of averaged intensity in a Y-axis at 495 cm^{-1} (*blue*), 755 cm^{-1} (*green*) and 1653 cm^{-1} (*red*) in the hyperspectral image shown in **a**. (**C**) Moving averaged spectrum in the range of 5 points (*A, red*) and 41 points (*B, blue*) of *p1* in **a** (*upper*) and the subtracting between the two (*green*). (**D**) A segmentation result of cell and non-cell are with threshold of intensity at 1438 cm^{-1}. (**E**) A labeling result (*upper*) and Raman spectra averaged within each labeled cell area (*lower*). The number and the color are corresponding in upper and lower. (**F**) X-position dependence of averaged intensity in non-cell area in a Y-axis at 495 cm^{-1} (*blue*), 755 cm^{-1} (*green*) and 1653 cm^{-1} (*red*) in the hyperspectral image shown in **A**. (**G**) predicted background spectrum on each cell area. The number and the color correspond to **E**. (**H**) Subtractions of **E** and **G**.

Supplementary Figure S3. Two kinds of the current autofluorescence subtraction methods. (**A**) Explanatory diagram of the current autofluorescence subtraction method with polynomial fitting alogism. (**B**) A typical result of the polynomial fitting alogism. Raw data, *red* (*A*); Lower envelope obtained by the alogism, *blue* (*B*); the subtraction result of *A* and *B*, *green*. (**C**) Explanatory diagram of the current autofluorescence subtraction method rolling-circle alogism. (**D**) A typical result of the rolling-circle alogism. Raw data, *red* (*A*); Lower envelope obtained by the alogism, *blue* (*B*); the subtraction result of (*A*) and (*B*), *green*.
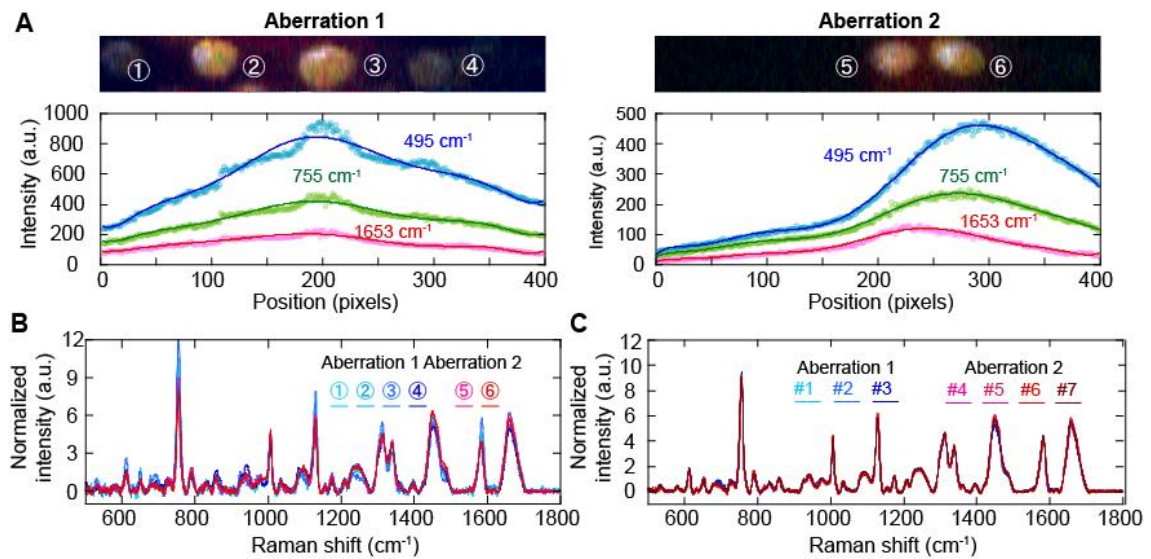
Supplementary Figure S4. Threshold dependence of background subtraction. (**A**) Segmentation results with distinct threshold sets. The difference between threshold *set #1* and *#2* was the threshold for non-cell area, and the difference between threshold set *#2* and *#3* was the threshold for cell area. (**B, C**) Background-subtracted Raman spectra of the cell *#2* (*upper*) and *#4* (*lower*) in Figure S2E by the conventional method (**B**) and the present method (**C**) with threshold *sets #1* (*red*), *#2* (*blue*) and *#3* (*green*).
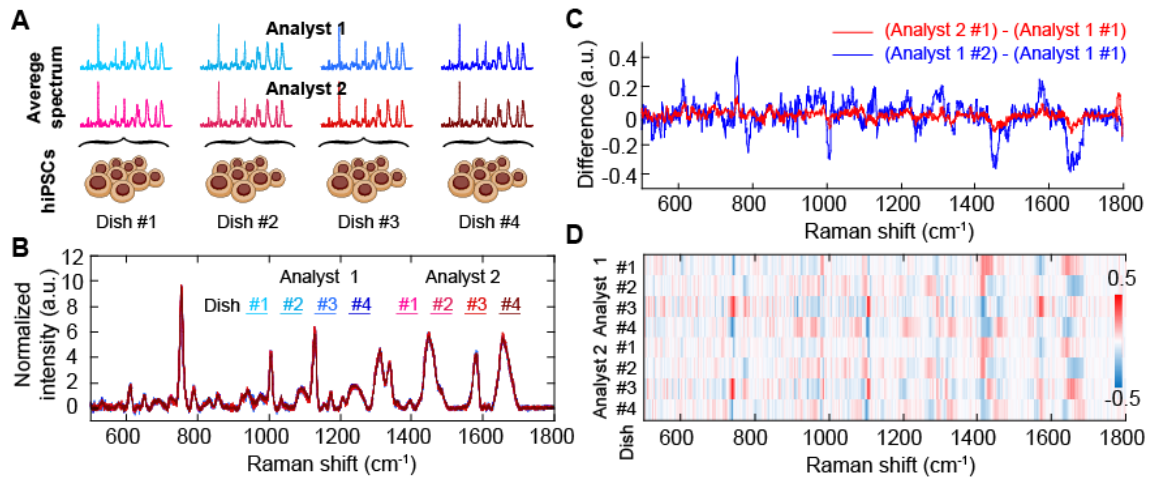
Supplementary Figure S5. Investigation analyst-dependence in the present background subtraction. (**A**) Explanative cartoon of analyst-dependence investigation. (**B**) Raman spectra averaged in each dish independently analyzed by two analysts (*analyst #1*, *blue*; *analyst #2*, *red*). (**C**) Difference spectra in *dish #1* between the *analyst #1* and *#2* (*red*), and in the analyst#1 between #1 and #2 (*blue*). (**D**) A heat map of the differences from the total averaged spectrum among four dishes and the two analysts. The color indicates the difference of mean spectrum of each dish and the average spectrum of all eight collections.

Supplementary Figure S6. Aberration dependence of background subtraction. (**A**) A pseudo-color image with cytochrome-C (755 cm^{-1}) colored in blue, lipids (1653 cm^{-1}) in green, and proteins (1438 cm^{-1}) in red (*upper*) and X-position dependence of averaged intensity in a Y-axis at 495 cm^{-1} (*blue*), 755 cm^{-1} (*green*) and 1653 cm^{-1} (*red*) of two cases of aberration. The image size is 125 μm × 33.5 μm. (**B**) Average spectra of each cell in A and B. The number is corresponding between **A** and **B**. (**C**) Mean Raman spectra of 6 dishes obtained in the two cases of aberration. Different color, different dish. Blueness, aberration 1; Redness, aberration 2. The number *#1* to *#7* is each dish.
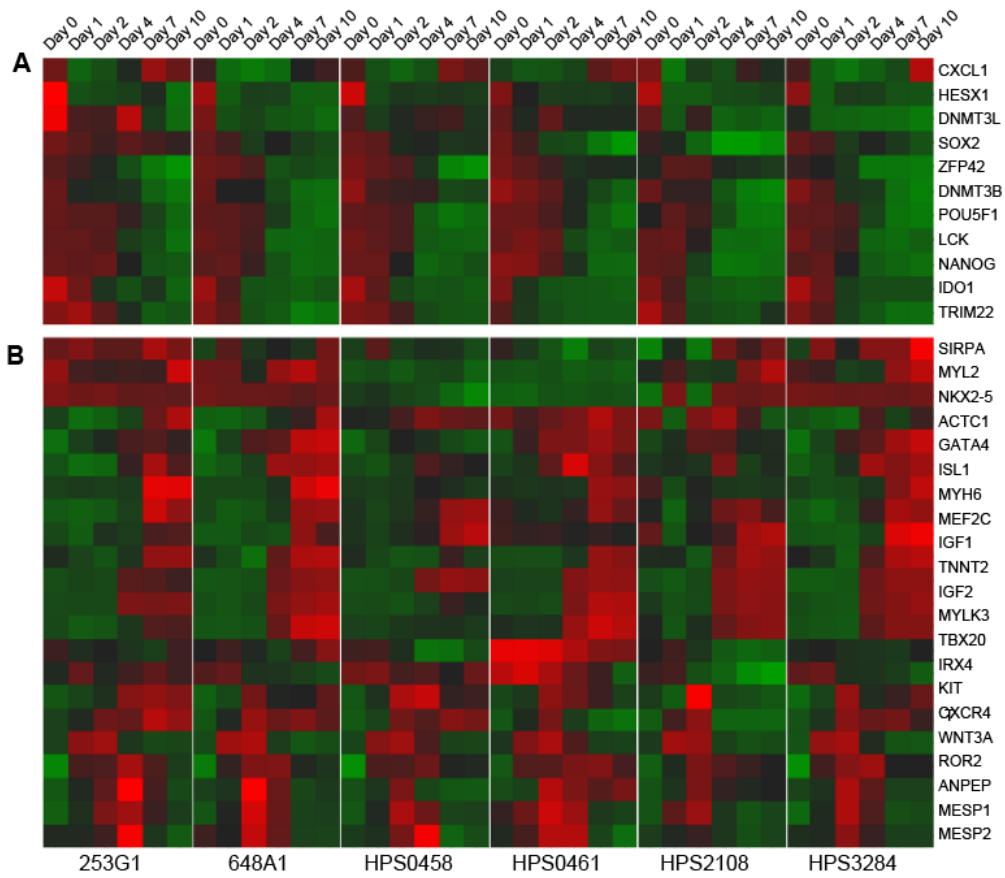
Supplementary Figure S7. Transcriptomic analysis of six hiPSC-lines during the myocardial differentiation. (**A**) Heat map representing color-coded expression of pluripotency related genes during the differentiation. (**B**) Heat map representing color-coded expression of cardiomyogenesis related genes during the differentiation.
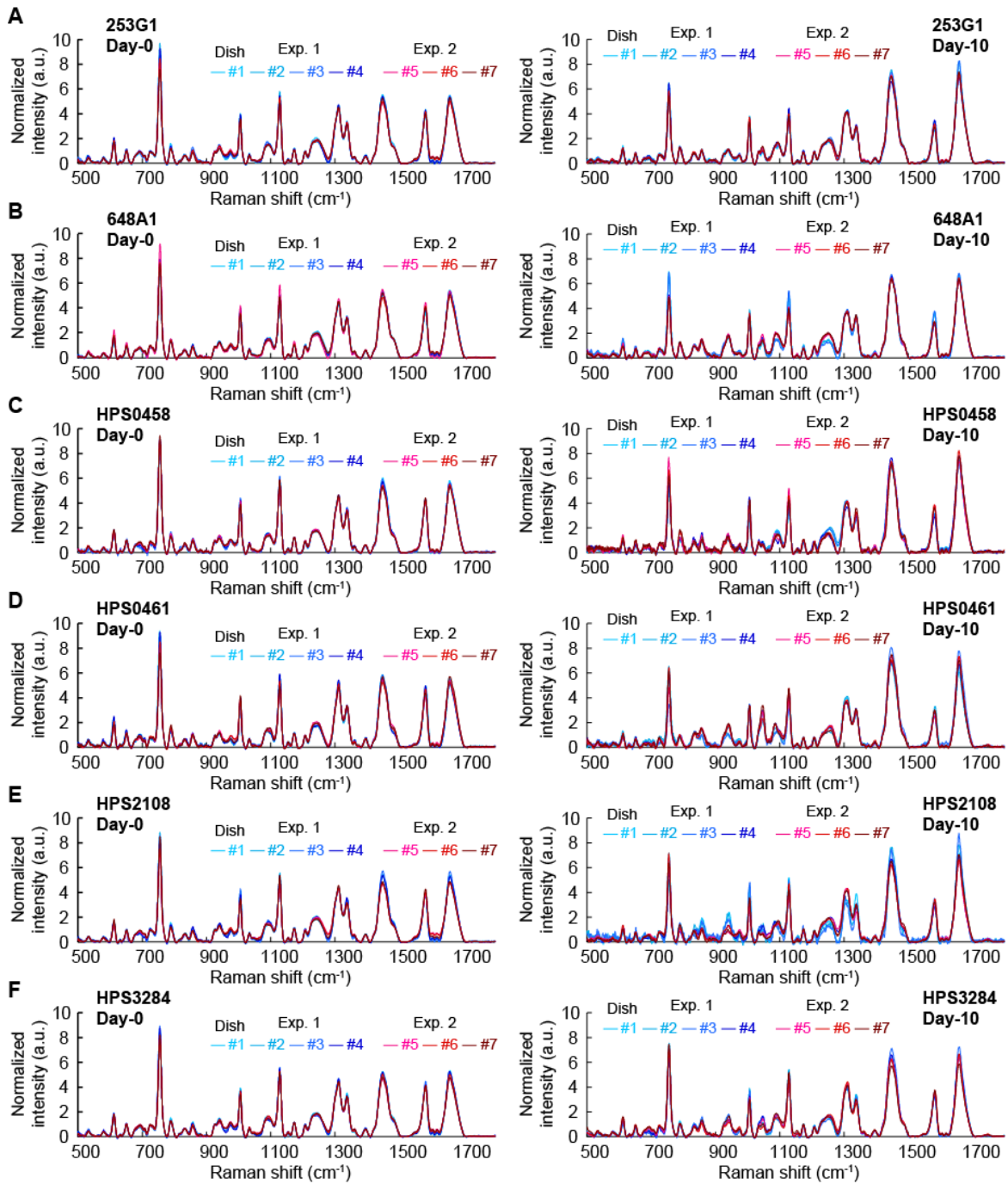
Supplementary Figure S8. Mean Raman spectrum of six hiPSCs before and on 10th day after differentiation induction. Averaged Raman spectra of 4 individual dishes of 253G1 (**A**), 648A1 (**B**), HPS0458 (**C**), HPS0461 (**D**), HPS2108 (**E**) and HPS3284 (**F**) before (*left*) and on 10th day after differentiation induction (*right*). Color indicates each spectrum averaged.
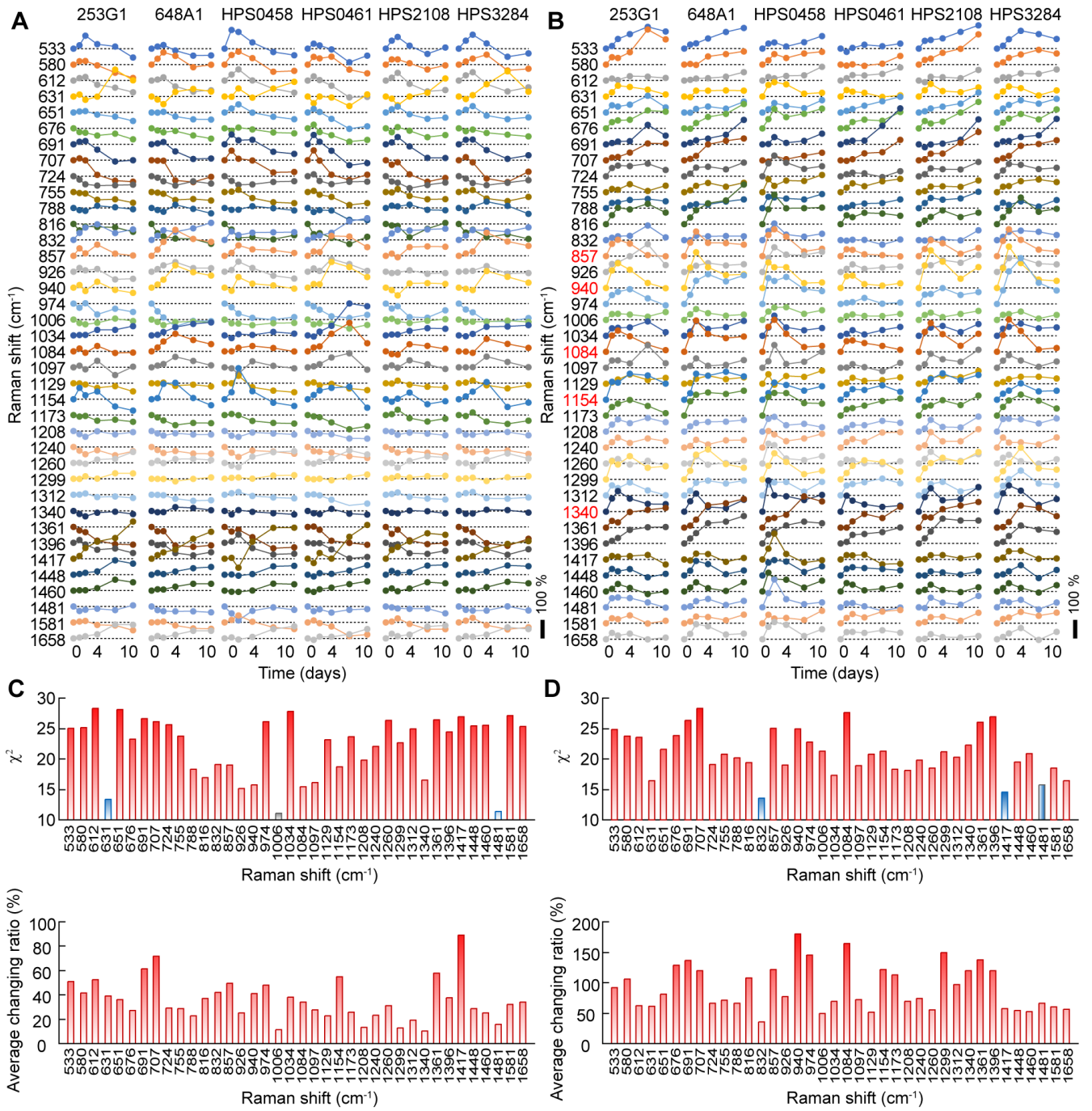
Supplementary Figure S9. Time development of main peaks on Raman spectra during the myocardial differentiation.
(**A, B**) Percentage change from before differentiation induction in mean intensity (**A**) or coefficient of variance (**B**) in a dish.
Red characters in **B** indicate the peaks whose standard deviation increased more than 2-folds immediately after the

differentiation induction and decreased in the latter half in all hiPSC-lines. (**C, D**) Chi squared (χ^2) values between the 6 hiPSC-lines obtained using Kendall's coefficient of concordance for the mean intensity (**C**) or the coefficient of variance (**D**) in a dish. Average changing rate was calculated by integrating the amount of change in the intensity or CV between day 0 and day 1~10 in all 6 cell-lines. Red color indicates p<0.01, blue indicates p<0.05, respectively.

Supplementary Table S1. Number of cells measured

Cell-line	Day 0	Day 1	Day 2	Day 4	Day 7	Day 10	Day 14
253G1	107	120	135	151	96	112	81
648A1	100	93	108	132	135	142	86
HPS0458	93	131	138	189	137	93	74
HPS0461	151	139	141	144	112	72	78
HPS2108	83	83	125	140	108	77	66
HPS3284	116	130	164	146	111	114	87