Electronic Supplementary Material (ESI) for ChemComm.

Vibrational phenomics decoding of stem cell stepwise aging process at single-cell resolution

Yue Wang[1,2,4#], Yadi Wang[2, #], Xueling Li[3], and Junhong Lü[2*]

[1]Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China

[2] School of Pharmacy, Binzhou Medical University, Yantai 264003, China

[3] Shanghai University of Medicine & Health Sciences, Shanghai 201318, China

[4] University of Chinese Academy of Sciences, Beijing 100049, China

[#] These authors contributed equally to this study.

*Correspondence: lvjunhong@jnl.ac.cn (J.H.L.)

Materials and Methods

1. Cell culture

Human umbilical cord mesenchymal stem cells (hMSCs, purchased from iCell, Shanghai) were cultured (at a density of $1\times10^4/cm^2$ at each passage) in low glucose DMEM (Procell, Wuhan) supplemented with 10% fetal bovine serum (Gibco), 100 U/ml penicillin and 100 mg/ml streptomycin at 37°C with 5% CO_2 saturation in a humidified atmosphere. Cell media was replaced every three days, cells were passaged when grown to about 70%-80% confluence.

2. SA-β-galactosidase detection assay

The cells were passaged to 4 and 14 generations, respectively. The cells were washed three times with PBS, then transferred to staining using SA-β-Gal kit (Beyotime, Shanghai). Briefly, the cells were fixed with fixative solution for 15min, after which they were washed with PBS three times for 5 min each time. The cells were incubated with staining solution overnight for 16-24 h at 37 ℃ without CO_2 and observed under a microscope.

3. Sample preparation

The cells were passaged to 4, 6, 8, 11, 14 generations. For single spectrum collection, the cells were then washed three times with PBS, scraped with cell scrapers and transferred into centrifuge tubes. After centrifugation, the cells were resuspended and fixed with 4% paraformaldehyde for 30 minutes. The fixed cells were washed three times with ultrapure water and resuspended with 100 μL of ultrapure water. Then, 10 μL of cell suspension was dropped onto BaF_2 windows and dried at room temperature for at least 30 minutes until completely dry. For single-cell imaging, cells were grown on CaF_2 windows in 6-well culture plates. The cells were also fixed with 4% paraformaldehyde for 30 minutes after washing three times with PBS. After fixation, the cells were washed three times with ultrapure water, and the CaF_2 windows were dried at room temperature. All samples were stored at 4 ℃ prior to experiments.

4. SR-FTIR microspectroscopy experiment

SR-FTIR was performed at beamline BL01B and BL06B of Shanghai Synchrotron Radiation Facility (SSRF). Single spectra of individual cells on BaF_2 windows were collected, and the aperture size was set to 25×25 μm, all the spectra were obtained within the spectral range 4000-650 cm^{-1} with a spectral resolution of 4 cm^{-1}, 64 co-added scans per spectrum. Infrared mapping images of single cells on CaF_2 windows were collected and the aperture size was set to 12×12 μm with a step size of 6×6 μm.

5. Data analysis

5.1 Spectra pre-processing

For single-cell spectral data, the original spectra were smoothed (9-point) and automatic baseline corrected on OMNIC (Thermo Fisher Scientific Inc.). The Mie scatter of all the spectra were corrected using RMieS-EMSC correction [1]. After that, the corresponding second-derivative spectra were calculated on OMNIC.

Mapping data were analyzed on CytoSpec. The original spectral was cut and the region 4000-1000 cm^{-1} was retained, then the mapping spectra was smoothed (15-point) and baseline corrected by AsLs method. The chemical images were acquired based on the area of absorption band regions

listed in Table S1.

For whole spectral region, the range 3000-2800 cm^{-1} together with 1800-900 cm^{-1}; fatty acid component, 3000-2800, 1480-1300 cm^{-1}; proteins, 1800-1480 cm^{-1}; carbohydrates, 1200-900 cm^{-1} were selected, respectively.

5.2 Intercellular Euclidean distance calculation

The similarity of infrared spectra between cells were evaluated by Euclidean distance: $D_{Eu}(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$, where p and q are two single cells represented by n dimensional vectors ("n" is corresponding to the number of points of infrared spectra). The value of $D_{Eu}(p,q)$ is negative correlation with intercellular similarity.

The intercellular Euclidean distance matrix was constructed by calculating all the cell-to-cell distances in a group (in-group) or between two groups (inter-group). For a group with "x" number of cells, a total of $C_x^2 = x(x-1)/2$ Euclidean distances can be acquired. For two groups (group A and B with "x" and "y" numbers, respectively), a total of x×y Euclidean distances can be acquired.

Intercellular Euclidean distance matrixes were calculated on MATLAB R2014a. The intercellular Euclidean distances heatmaps and frequency distributions were drawn using Origin Pro 2016.

5.3 Principle component analysis (PCA) and hierarchical cluster analysis (HCA)

PCA was carried out on the second-derivative spectra of single-cells. The first two principle components (PCs) were used to construct the score plots using OriginPro 2016, the corresponding loading plots of PC1 and PC2 were also acquired.

HCA was also carried out on the second-derivative spectra, "Euclidean distance" calculation and "Ward" linkage were chosen.

Both PCA and HCA were carried out on MATLAB R2014a.

Supplementary figure

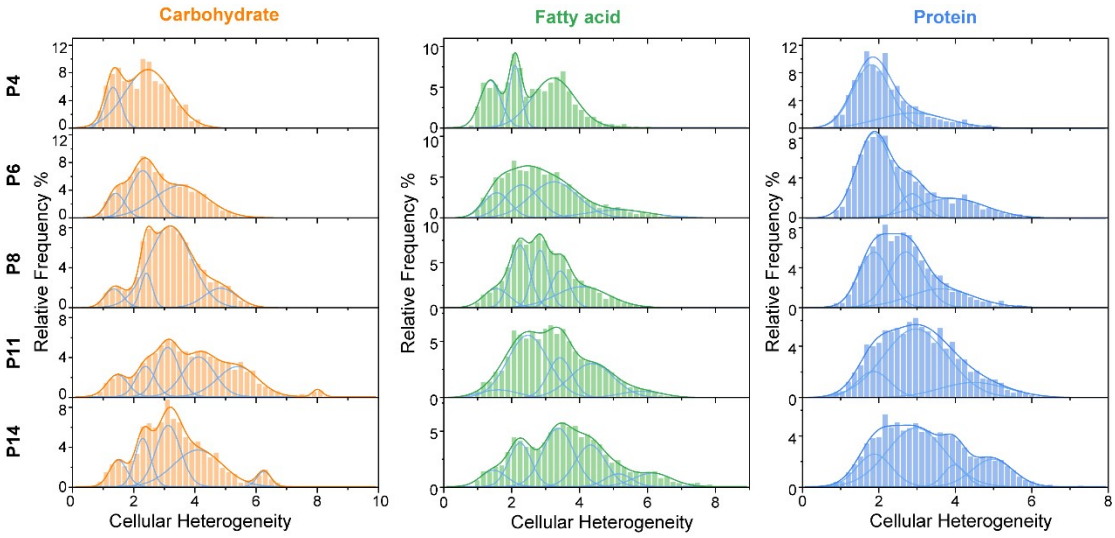

Fig. S1 Heterogeneity of biochemical components during cell aging. Heterogeneity of P4, P6, P8, P11 and P14 cells during aging was evaluated by carbohydrate (orange, left), fatty acid (green, middle) and protein (blue, right) components, respectively. The curves were fitted with the Gaussian function.