System requirements

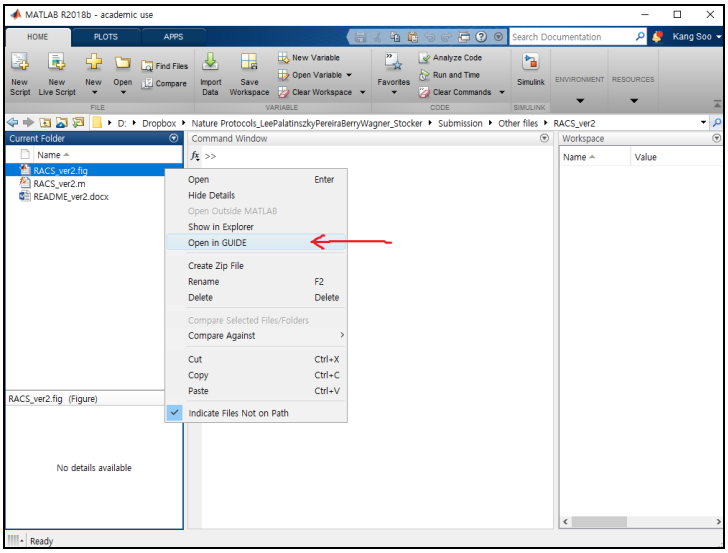
- Commercial Raman microspectroscope: LabRAM HR800, Horiba Scientific
- LabSpec: to control the Raman microspectroscope (provided by the manufacturer, Horiba Scientific)
- Desktop/laptop that has MATLAB at any versions that provides GUIDE (graphical user interface development environment) module.

  Note: the code built in-house allows the Raman-activated cell sorting (RACS) to be run within a commercial Raman microspectroscope (LabRAM HR800), and thus several commands are specific for the Horiba system. We include a footnote to those commands (marked by 'Horiba-specific') so that users can modify them as appropriate to the system available. The two other major Raman manufacturers (Renishaw and Bruker) have confirmed that their systems are compatible with control using third-party software (e.g., MATLAB as in this protocol, or Python). Our code can also be adopted with suitable modifications for other Raman systems (including those built in-house) if they allow an application programming interface (API).

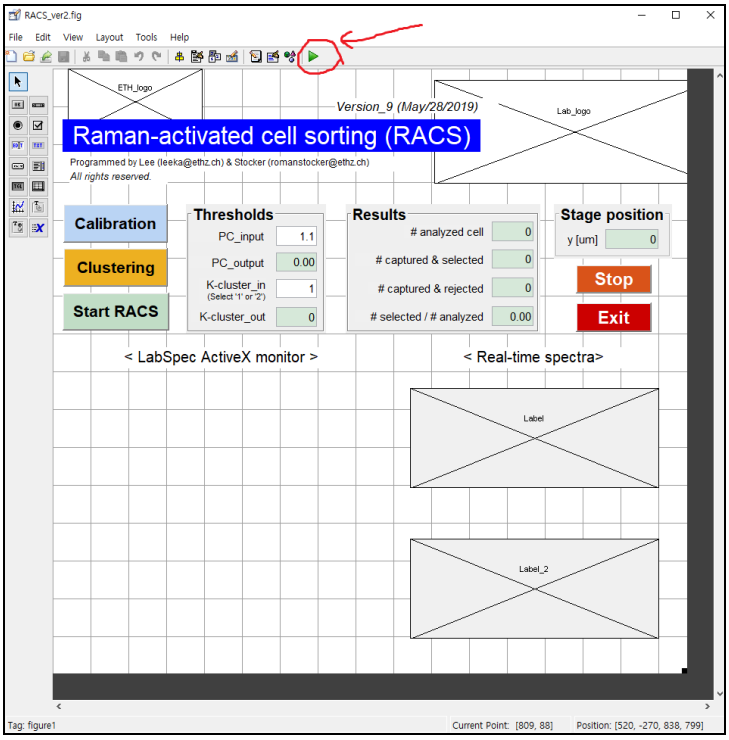
Instructions for use

Note: this version is for sorting of cells with respect to cytochrome *c*. To operate the software platform with respect to other chemical fingerprints (e.g., carotenoids), we have included a footnote to commands where modification is required to adjust the spectral region of interest.

1. In Step 53 (in main text), right-click the 'RACS_ver2.fig' file. Click 'Open in GUIDE'.

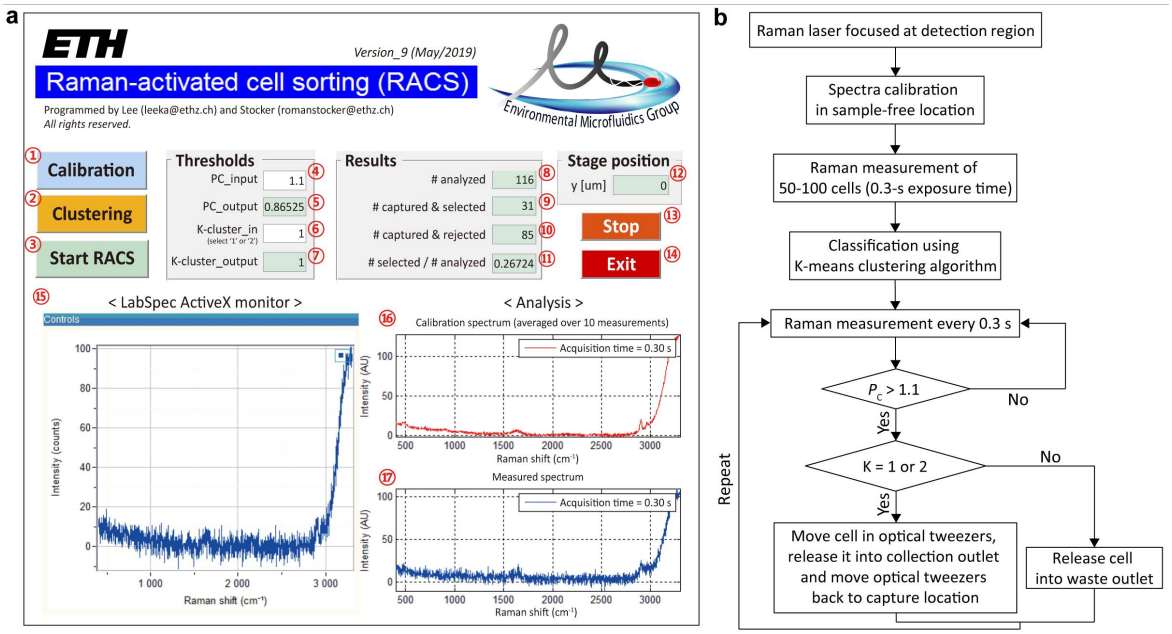


2. Click the green **Run** button to start the platform.



3. Click the ‘Calibration’ button (1) (panel **a** in figure below). The stage moves to the sample-free region (70 µm away from the sample stream) and measure the Raman spectrum of the working fluid (averaged over 10 measurements). One output file ‘Calibration_1981-10-07 19_02_07.l6s’ (named using the date and time ‘1981-10-07 19_02_07’; this will vary depending on the timing of the operation) that contains a Raman spectrum measured during the calibration is created.

   Note: ‘.l6s’ is the LabSpec-specific format. Boxes with white (4,6) and green (5,7–12) colours represent input (adjustable by the user) and output, respectively.



4. Click the ‘Clustering’ button (2) to begin the initial determination of the clusters, in which the software evaluates the Raman spectral characteristics of the sample: when a cell is captured in the optical tweezers, it is measured and immediately released without sorting.
5. Click the ‘Stop’ button (13) when the number of cells analysed (displayed in box 8) meets the user’s requirements. The software classifies the measured data into several clusters using the K-means clustering algorithm (the number of cells to be analysed during this procedure depends on the relative abundance of taxa of interest within the sample). Three output files are created:

- Clustering_Spectra_1981-10-07 19_02_07.txt: Raman spectra of cells used in K-means clustering
- Clustering_PC_1981-10-07 19_02_07.txt: P_C values of cells used in the clustering
- Clustering_summary_1981-10-07 19_02_07.txt: cluster numbers of each cell (1 and 2 indicate the uppermost and second uppermost clusters, respectively; e.g., clusters 1 and 2 in **Fig. 8j,k** in main text)

6. Click the 'Start RACS' button (3) to start the RACS process. The software will ask the '.txt' file that had been obtained during the 'Clustering' step. Select the 'Clustering_Spectra_1981-10-07 19_02_07.txt' file. The software then begins the procedures described in panel **b** of the figure above.
   - Boxes (4) & (5) display the threshold set by the user for P_C and the P_C value measured in real time, respectively.
   - Box (6) displays the value of the clusters from which cells are to be collected (1 and 2 mean the uppermost cluster and the uppermost and second uppermost clusters, respectively; e.g., cluster 1 alone and clusters 1 and 2 in **Fig. 8j,k** in main text); Box (7) displays real time output of the cluster value of the cell currently being analysed ('N/A' indicates clusters other than those to be sorted – e.g., clusters 3–5 if the software is classifying the data into 5 clusters and clusters 1 and 2 are being collected; see **Fig. 8j** in main text for the selection criterion for the number of clusters)
   - Boxes (8)–(11) display the number of cells analysed, captured and selected, and captured and rejected, and the proportion of selected cells among those analysed, respectively.
   - Box (12) displays the current position in the *y*-direction of the microscope stage (perpendicular to the flow direction)
   - Panel (15) displays the Raman spectrum measured in real time.
   - Panel (16) displays the calibration spectrum (used to calculate P_C).
   - Panel (17) displays the latest captured and selected cell.
7. When the number of sorted cells meets the user's requirements, click the 'Stop' button (13) to stop the RACS process. Then, click the 'Exit' button (14) to close the software platform.
8. Six output files are created:

- RACS_Spectra_Selected_1981-10-07 19_02_07.txt: contains Raman spectra of sorted cells
- RACS_Spectra_Rejected_1981-10-07 19_02_07.txt: contains Raman spectra of discarded cells
- RACS_PC_Selected_1981-10-07 19_02_07.txt: contains P_C values of sorted cells
- RACS_PC_Rejected_1981-10-07 19_02_07.txt: contains P_C values of discarded cells
- RACS_Spectra_Clustering_1981-10-07 19_02_07.txt: Raman spectra of cells used in K-means clustering
- RACS_Summary_Clustering_1981-10-07 19_02_07.txt: cluster number of each cell (1 and 2 indicate the uppermost and second uppermost clusters, respectively; other cluster numbers are not ordered by their physical domain)