

Label-Free Identification of Exosomes using Raman Spectroscopy and Machine Learning

Ugur Parlatan, Mehmet Ozgun Ozen, Ibrahim Kecoglu, Batuhan Koyuncu, Hulya Torun, Davod Khalafkhany, Irem Loc, Mehmet Giray Ogut, Fatih Inci, Demir Akin, Ihsan Solaroglu, Nesrin Ozoren, Mehmet Burcin Unlu,* and Utkan Demirci*

Exosomes, nano-sized extracellular vesicles (EVs) secreted from cells, carry various cargo molecules reflecting their cells of origin. As EV content, structure, and size are highly heterogeneous, their classification via cargo molecules by determining their origin is challenging. Here, a method is presented combining surface-enhanced Raman spectroscopy (SERS) with machine learning algorithms to employ the classification of EVs derived from five different cell lines to reveal their cellular origins. Using an artificial neural network algorithm, it is shown that the label-free Raman spectroscopy method's prediction ratio correlates with the ratio of HT-1080 exosomes in the mixture. This machine learning-assisted SERS method enables a new direction through label-free investigation of EV preparations by differentiating cancer cell-derived exosomes from those of healthy. This approach will potentially open up new avenues of research for early detection and monitoring of various diseases, including cancer.

1. Introduction

Exosomes are nano-sized extracellular vesicles (EVs) that are released to the extracellular space and circulation establishing a channel for communication between cells.^[1] Recently, EVs were found to be useful for broad applications in medicine, molecular biology, and medical physics.^[2] They are critical, especially in cancer biology, as tumor cells release EVs providing structural information and diagnostic potential.^[3] According to the International Society for Extracellular Vesicles guidelines,^[4] EVs are characterized using several techniques, including immunoblotting,^[5] transmission electron microscopy (TEM),^[6–8] flow cytometry,^[9] and optical imaging systems such as fluorescence microscopy.^[10,11]

U. Parlatan, I. Kecoglu, I. Loc, M. B. Unlu
Department of Physics
Bogazici University
Istanbul 34342, Turkey
E-mail: burcin.unlu@boun.edu.tr

U. Parlatan, M. O. Ozen, M. G. Ogut, D. Akin, U. Demirci
Department of Radiology Stanford School of Medicine
BioAcoustic MEMS in Medicine Lab (BAMM)
Canary Center at Stanford for Cancer Early Detection
Palo Alto, CA 94304, USA
E-mail: utkan@stanford.edu

B. Koyuncu
Department of Computer Engineering
Bogazici University
Istanbul 34342, Turkey

H. Torun
Koc University Graduate School of Sciences and Engineering
Istanbul 34450, Turkey

H. Torun, I. Solaroglu
Koc University Research Center for Translational Medicine (KUTTAM)
Istanbul 34450, Turkey

D. Khalafkhany, N. Ozoren
Department of Molecular Biology and Genetics
Center for Life Sciences and Technologies
Apoptosis and Cancer Immunology Laboratory (AKiL)
Bogazici University
Istanbul 34342, Turkey

F. Inci
UNAM-National Nanotechnology Research Center
Bilkent University
Ankara 06800, Turkey

F. Inci
Institute of Materials Science and Nanotechnology
Bilkent University
Ankara 06800, Turkey

I. Solaroglu
School of Medicine
Koc University
Istanbul 34450, Turkey

M. B. Unlu
Faculty of Engineering
Hokkaido University
North-13 West-8, Kita-ku, Sapporo, Hokkaido 060–8628, Japan

M. B. Unlu
Global Center for Biomedical Science and Engineering Quantum
Medical Science and Engineering (GI-CoRE Cooperating Hub)
Faculty of Medicine
Hokkaido University
Sapporo 060–8638, Japan

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/smll.202205519>.

DOI: 10.1002/smll.202205519

Among optical methods, fluorescence microscopy is one of the most frequently used, often require lengthy and repetitive protocols, and antibodies utilized may not bind to the desired regions and provide limited information about surface proteins of the EVs in some cases due to the size of the vesicles.^[12] Besides, autofluorescence or photobleaching are commonly encountered problems that make it difficult to differentiate the fluorescence signal of the sample from the background.^[13]

Raman spectroscopy is a sample preparation-free vibrational spectroscopic method that provides structural information specific to the examined sample without staining or related structural and conformational damage.^[14] It has been highlighted in the literature that fundamental differences in the disease states and molecular profiles of cancerous cells and tissues,^[15–17] bacteria,^[18] viruses,^[19] red blood cells,^[20] immune cells^[21] can be identified by investigating their Raman signals.

Recently, there has been increasing attention to using Raman spectroscopy and multivariate analysis/machine learning methods to characterize EVs.^[22–30] Most of these studies applied

traditional spectral analysis methods or statistical data exploration/reduction methods like principal components analysis (PCA) and bayesian classification algorithms like linear discriminant analysis (LDA).^[31] The emergence of machine learning algorithms has dramatically improved spectral analysis methods, and they found application in EV investigations.^[32–36] Artificial neural networks (ANN) algorithm (also known as deep learning) has been used with Raman spectroscopy to classify exosomes derived from healthy plasma samples from those taken from lung cancer patients.^[32] A similar approach was applied to investigate urogenital cancers^[33] and breast cancers^[34] recently. Besides cancers, pathogen-derived EVs were identified at Gram±, species, strain, and physiological levels.^[35]

Here, we presented identification of the EVs derived from five cell lines in respect to their chemical signatures using Raman spectroscopy and machine learning algorithms. We prepared a mixture of EVs originated from two different cell lines to evaluate our detection platform and algorithms (Figure 1a). Using the collected data from different points

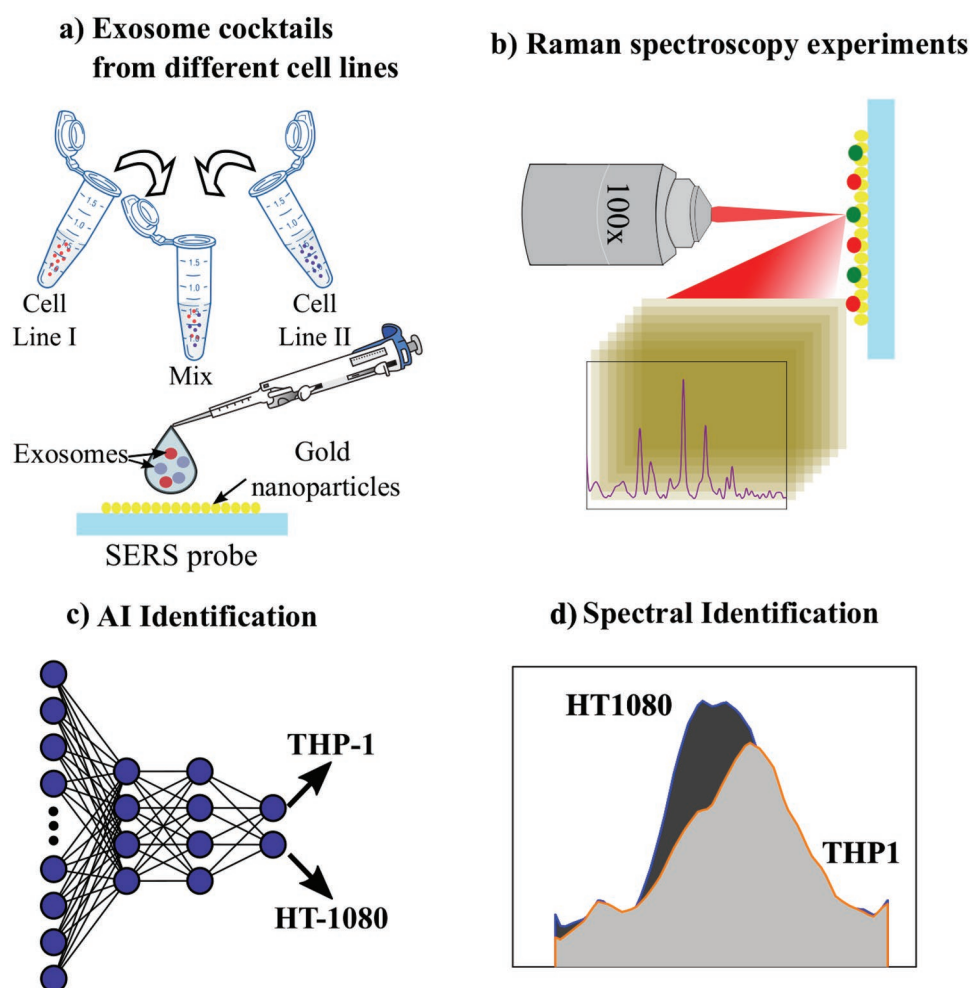


Figure 1. Sample preparation and spectral analysis of EVs to identify their cellular origin. a) We mixed different cell line-derived EVs in a separate solution and dried a drop (15 μ L) of this solution on a SERS probe. Vectorial drawing credit: <https://doi.org/10.5281/zenodo.3925953>, <https://doi.org/10.5281/zenodo.3925967>. b) The dried sample on the SERS probe was measured in the custom-built Raman spectroscopy setup available in our laboratory. A motorized stage was used to scan SERS probes to investigate EVs site-by-site by measuring the spectra. c) Measured Raman spectra were used to identify the cell origins of the EVs via machine learning models. d) Same spectra were analyzed using band component analysis to extract spectral information under broad peaks in the spectra which can potentially lead to classification / identification.

from surfaces and different SERS surfaces (Figure 1b), we utilized an artificial neural network algorithm (ANN) to profile the EV mixtures prepared in different ratios quantitatively (Figure 1c). We also analyzed the Raman spectra of EVs using a band component analysis to reveal hidden patterns, bringing a chemical perspective to our classification method (Figure 1d). This study demonstrates that pre-trained machine learning algorithms can detect the cellular origins of cancer cell-derived EVs using label-free surface-enhanced Raman spectroscopy (SERS) / ANN analysis to identify exosomal signatures in a mixture. These signatures can identify cancer cell-derived EVs in liquid biopsy samples for cancer early detection and prognosis.

2. Experimental Section

2.1. Sample Preparation

All the media and supplements for cell culture were purchased from Gibco BRL (NY, US). FBS used in these experiments was heat-inactivated at 60 °C and filtered by 0.2 µm filters.

HEC-1A cells (ATCC, Germany) maintained in McCoy's 5A medium supplemented with 10% fetal bovine serum (FBS), 1 mM Non-Essential Amino Acids (MEM-NEAA), and 1 mM Penicillin/Streptomycin. THP-1 cells were maintained in RPMI-1640 medium supplemented with 10% FBS, 1 mM MEM-NEAA, 0.05 mM 2-mercaptoethanol, and 1 mM Penicillin/Streptomycin. HT-1080 cell line (HT-1080 Xeno ATCC CRL-12011) was kindly provided by Prof. Maria Soengas, and HeLa cells were kindly provided by Prof. Cemalettin Bekpen. HT-1080, HeLa, and MEF cell cultures were maintained in DMEM medium supplemented with 10% FBS, 1 mM Non-Essential Amino Acids (MEM-NEAA), and 1 mM Penicillin/Streptomycin.

All the cells were grown at 37 °C with 5% CO₂, culture media was refreshed every 2–3 days, and cultures were passaged after the cells reached 90% confluency. For EV collection, passaged cells were cultured until they reached confluency with the growth media mentioned above. Then the cells were washed with PBS (3 times) and maintained in the media prepared in 5% exosome-depleted FBS for another 48 h. After 48 h, culture media were centrifuged at 1500 RPM (730 g) to separate cells and cell debris from the culture media. Then the supernatant was used for EV isolation after filtering with a 220 nm injector-type PES filter.

2.2. EV Isolation

The EVs were isolated from the supernatants of the HT-1080, HEC-1-A, THP-1, HeLa, and MEF cell lines using an in-house developed exosome isolation platform, Exosome Total Isolation Chip-ExoTIC. ExoTIC is a membrane-based filtration device designed to isolate EVs from multiple sample types, including cell culture media.^[37–43] The filtration allows particles whose pore sizes are between 30 and 200 nm. The isolated EVs were stored at –80 °C until the measurement and they were measured on the same day that they were taken from the freezer.

The selected EV tubes were taken from the freezer on the measurement day and let them reach room temperature.

2.3. Characterization of EVs

Characterization experiments of EVs were performed via Nanosight Nanoparticle Tracking Analysis (NTA), Transmission Electron Microscopy (TEM), flow cytometry, and western blotting by applying the protocols given in Supplementary Information.

2.4. SERS Characterization

To enhance the weak signals from smaller vesicles, the help of plasmon resonances was used from metallic surfaces called SERS probes. Among several SERS sampling methods, a method was used that utilizes drying the sample on the probe. 15 µL of EV preparation was dropped on the SERS probe and the sample was air-dried for 10 mins. The probes were kept enclosed to minimize contamination from the environment for the drying period. A new probe was used in each measurement.

Raman spectra of EVs were taken with 2 s of integration time. All measurements were performed by scanning the gold surface (raster) to reveal the site-to-site differences. Scans were employed on an XY motorized stage with a 20 µm step size and 1.4 mm X 1.4 mm scan area. A boxcar averaging was applied during the measurements with a boxcar width of one.

2.5. Data Analysis

For baseline correction, a discrete profile was manually determined by combining fitted linear lines that pass over local minimum points we selected on the spectrum. The number of selected points was kept as few as possible. The purpose of doing this was that when we fit a continuous polynomial curve on a spectrum, the relative intensity of the weak peaks decreases or increases artificially after the subtraction, creating artificial results during classification. Our linear baseline approach minimized this kind of classification error. Normalization: Each data point was divided in a spectrum by the vector norm of the spectrum. This type of normalization is generally called vector normalization. The combination of polynomial curve subtraction and vector normalization was found to be the most efficient pre-process method among other standard methods.^[44,45]

2.5.1. Outlier Exclusion

We used z-scores which are calculated as in the equation below:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Here, x , μ , and σ are the individual measurement array, mean of all measurements, and standard deviation of all measurements, respectively.

2.5.2. Random Forest Algorithm

100 estimators were utilized, with the criterion of Gini impurity, and were built in scikit-learn.^[46] Gini impurity denotes the probability of classifying a randomly selected sample in a node incorrectly, where the sample is labeled according to the class distributions in the node. The best Gini impurity score is 0, which is also the lowest score, and it denotes that every sample in the node shares the same class label.

2.5.3. ANN Model

Five dense layers were used with hidden channels of 32, 16, 8, 4, and 2. The activation function was selected as a rectified linear unit (ReLU) except for the last layer, which utilized the softmax activation function. The ANN was trained with categorical cross-entropy for 150 epochs with early stopping. ADAM optimizer^[47] was used with a batch size of 10 and an initial learning rate of 10⁻⁴. The number of layers was selected based on the performance, and increasing the depth of networks did not affect the model performance. The ANN was built with Keras backend.^[48] We compared the predicted count of HT-1080 with the estimated concentrations provided by the simulation.

2.5.4. Gaussian Curve Fitting

This analysis was performed using Grams AI 8 software. Initially, the peak locations were found and the number of peaks was found by calculating the second derivative of the spectra. The locations of the sharp turning points gave us the Raman

shifts of the sub-components. Initial Gaussian peaks were placed with widths slightly wider than the narrowest feature in the region. The algorithm iteratively modified the peak heights, widths, and locations to find the best fitted reconstructed curve. The parameters, such as peak location and area, were our indicators for EV investigation.

Data analysis steps were performed either on MATLAB (The MathWorks, Inc., Natick) or Python, using our codes. Figures were created using Inkscape^[49] and shaded plots were created using a script written by Rob Campbell (Shaded error bar, Mathworks, 2009).

3. Results and Discussion

3.1. Characterization of EVs via Transmission Electron Microscopy (TEM), Flow Cytometry and Nanoparticle Tracking Analysis (NTA), Western Blotting

To characterize EVs derived from HT-1080 and THP-1 cell lines, we performed Nanoparticle Tracking Analysis (NTA), TEM, Flow Cytometry, and Western Blot (WB) experiments. The protocols for each method were summarized in Supplementary Information.

The TEM characterization results showed that the diameter of the vesicles was between 130 and 170 nm (Figure 2a). These results were then further confirmed with NTA analysis (Figure 2b). Mode diameters were 172.4 and 174.7 nm for HT-1080 and THP-1, respectively. Moreover, we detected the tetraspanin marker CD63 on exosomes (Figure 2c) and counted the CD63-positive particles using the EV-marker antibody (FITC CD63 anti-human antibody, BioLegend). We observed

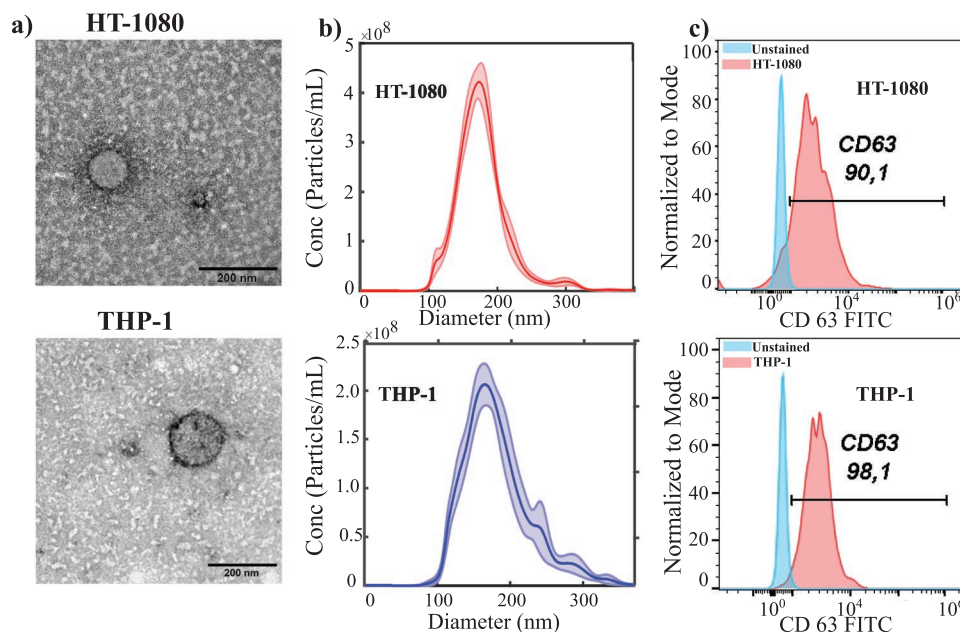


Figure 2. Characterization of EVs derived from THP-1 and HT-1080 cell lines. a) Transmission electron microscopy (TEM) images of EVs demonstrated approximately 130–170 nm average diameter. b) Nanoparticle Tracking Analysis (NTA) of EVs exhibiting mode diameters of 172.4 and 174.7 nm mode diameters for HT-1080 and THP-1, respectively. c) Flow cytometry analysis showed that 90.1 and 98.1% of molecules were stained with FITC-CD63 antibody (blue: unstained, red: HT-1080 or THP-1, respectively).

that 90.1 and 98.1% of molecules were stained for HT-1080 and THP-1 (Figure 2c and Figure S1, Supporting Information), respectively. We also confirmed these results with WB experiments for CD63 exosome marker protein (see Figure S2, Supporting Information). We obtained staining pattern spanning between 30–63 kDa for HT-1080 and 50–75 kDa for THP-1 (due to the differentiations in posttranslational modification), which are consistent with previous studies.^[50–52]

3.2. Measurement and Classification of EVs From Five Different Cell Lines

We measured EVs derived from five different cell lines (He-La, Hec-1-A, THP-1, mouse embryo fibroblast (MEF), HT-1080) using the setup described and illustrated in Figure S3 (Supporting Information). In Figure 3a, the raw and the baseline-corrected spectra of the average of all spectra of EVs isolated from He-La cells were given as an example to demonstrate the pre-processing (see Data Analysis for pre-process). In Figure 3b, the baseline-corrected spectra obtained from all cell lines were given. Additionally, the shaded standard deviation lines covering the region $\pm 1\sigma$ from the average line were overlaid with the average spectra. We added arbitrary offset levels between spectra to increase visibility in the graph. The averaged normalized spectra from measured EV types were given in Figure 3c, where the essential tentative band assignments were also written on the corresponding peaks.

When we plotted the map graphs corresponding to each cell line type, we found that some parts of the surfaces did not respond intensely (Figure S4, Supporting Information), which

indicated low or no amount of EVs. Therefore, we excluded these data using the z-score elimination method and employed principal component analysis (PCA) and random forest analysis to reveal the variances between the EVs from different cell types. Figure 3d shows that the first two principle components can separate the population into five classes. The coefficients plot (Figure 3e) compares EV types' similarities (or differences). Thus, we labeled the peaks with high intensities in the graph, which show the significant Raman shift values that led to the differentiation of the origins of EVs. For example, peaks at 1003 and 1030 cm^{-1} were tentatively assigned to benzene ring breathing (mainly in phenylalanine amino acid), and their relative intensity was negatively correlated with the peak at 1321 cm^{-1} . This negative correlation was seen in Figure 3c, where the highest peaks at 1321 and 1030 cm^{-1} correspond to MEF and THP-1 cell lines, respectively. We reduced the number of measurements by utilizing a z-score outlier elimination (see Data Analysis section) method and randomly selected 80% of all data for training input in a random forest classifier. We showed the confusion table calculated using this classifier on the test set in Figure 3f. Our classification model showed that the Raman spectroscopy and the random forest classifier model identified EVs secreted from different cell lines with high accuracy (96.8%).

3.3. Classification of the Signals From Exosome Mixtures With Different Concentrations

To quantify the predictive ability of Raman spectroscopy to reveal the origin of a certain EV type from a mixture preparation, we

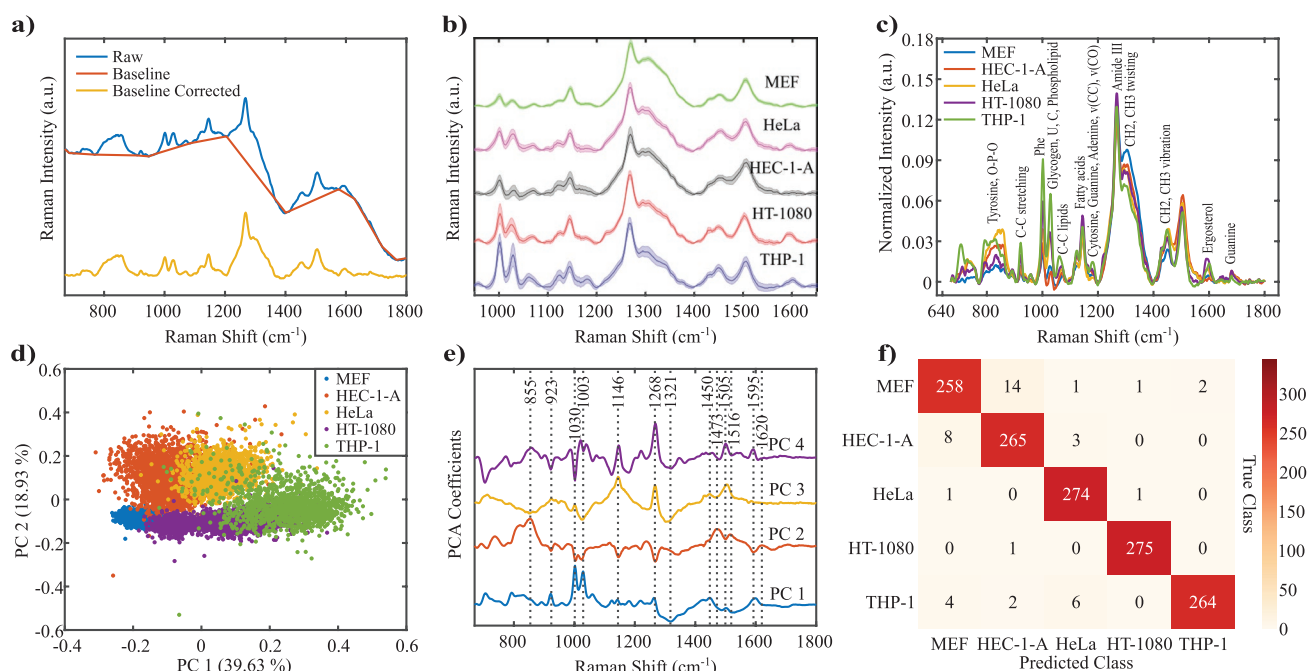


Figure 3. Random forest classifier results of five types of EVs after pre-processing steps. a) A four-step pre-processing algorithm was applied using MATLAB to obtain the baseline corrected spectrum. b) Average baseline-corrected spectra from five different cell lines overlaid with shaded standard deviation lines ($\pm 1\sigma$). c) Average normalized Raman spectra from all cell lines with the tentative assignments. d) PCA score plots of the first two components. e) PCA coefficients plot of the first four components with tentative assignments. f) Confusion table of the test set that includes 20% of the outlier-excluded data. See the methods section for the elimination algorithm.

studied individual EV preparations and different mixtures with increasing ratios of HT-1080. To build the data set, we first measured the EV samples derived from both HT-1080 and THP-1 without mixing. We followed a serial dilution protocol to predict the origin of the EVs analyzed from the mixture. We measured four different mixtures prepared by using EVs isolated from THP-1 and HT-1080 cell lines. We kept the THP-1 EV concentration constant at 50% and prepared the mixture using concentrations of HT-1080 EVs as 1%, 10%, 25%, and 50%. The EV concentrations were measured by nanoparticle tracking analysis (NTA) as 1.97×10^{10} particles/mL for THP-1 cell line and 3.07×10^{10} particles/mL for HT-1080 cell line. To obtain the average measurements, we repeated the measurements with three different SERS substrates (from the same batch) for each concentration and combined them in a single dataset. **Figure 4a** shows the portion of the spectra corresponding to these samples (Whole spectra are displayed in Figure S5, Supporting Information). Each proportion of individual concentration level was used as a class label in our machine learning process. The labels correspond to different levels of proportions. After the pre-processing, the number of measurements was 2398 for each class of EV mixture signals. We utilized a random forest classifier, which was trained using 1918 samples of each class. The classifier had 100 estimators, used the criterion Gini impurity, and required two samples to split a node. The classifier was then tested with the remaining 480 samples, which yielded the confusion matrix in **Figure 4b**. We quantified the classifier performance with micro and macro F1 scores. F1 score is the harmonic mean of precision and recall. Micro averaging (micro-F1) is a global metric that calculates precision and recall using total true positives, false positives, and false negatives, independent of class labels. Macro F1 score, on the other hand, is the mean of class-wise F1 scores. Both micro and macro F1 scores take values ranging from 0 to 1, while the higher values refer to better performance. As a result, the classifier had a score of 99.58% for both micro-F1 and macro-F1.^[53]

3.4. Predicting the Density of a Particular EV Type in a Mixture Using Artificial Neural Networks (ANN)

Since we classified different types of isolated EVs and EV preparations at various concentrations, our next step was to predict the overall density of a particular EV type in a mixed sample. According to the simulation results (see Supporting Information), the overall ratio of the number of given EV types (HT-1080 and THP-1) to all the EVs in the mixed preparation was within the 1% error range. We used the same mixture measurements to predict the outcomes from each mixture. First, we trained an ANN model by creating a dataset that consists of 2398 unmixed HT-1080 and THP-1 EV measurements. Then, we calculated the confusion matrix of the ANN model on the test set of unmixed EVs and displayed it in **Figure 5e**. After we tested the models for different spectral windows, we found the 880–1200 cm^{-1} region as the most specific spectral window to give the most accurate predictions.

As shown in **Figure 5a–d**, the probability of predicting HT-1080 cell line-derived EVs increased as their concentration was increased in the EV mixture. In **Figure 5f**, we presented an increasing trend between the number of predicted HT-1080 labels by ANN and the simulation outputs.

We repeated the measurements in **Figures 4,5** with a different batch of SERS probes. We summarized the results from these measurements in the Supporting Information section (see **Figure S6–S10**, Supporting Information). We found that the probes with different localized surface plasmon resonance (LSPR) wavelengths responded differently and amplify different spectral regions even if the wavenumbers of the peaks were the same for both probes, which indicated that comparison of the intensities of the probes was only achievable if the SERS probes are from the same batch. Otherwise, we can only get wavenumber information. Thus, comparing the intensity was

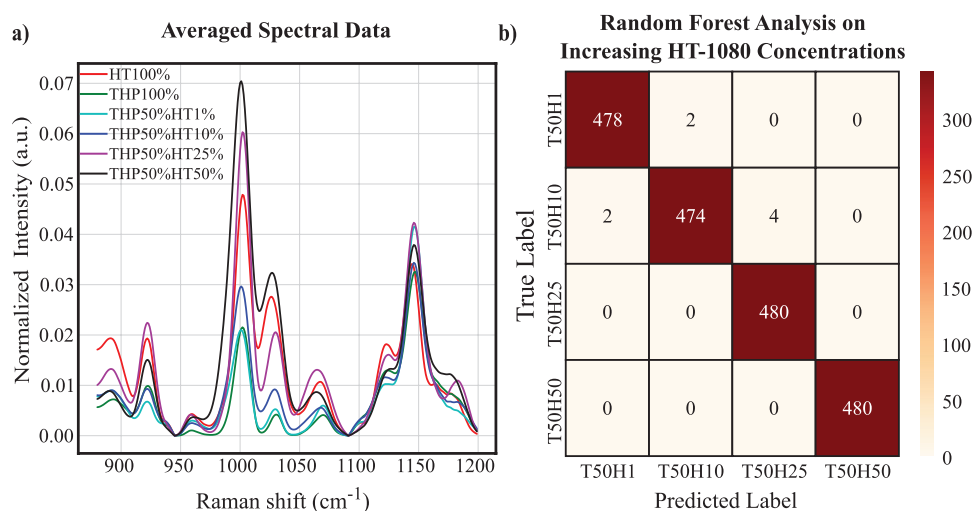


Figure 4. Random forest classifier results for the gradually changing concentrations of EVs derived from HT-1080 cell line while the number of EVs derived from THP-1 cell was kept constant. In the context of this figure, HT refers to HT-1080, THP refers to THP-1 cell lines and numbers after cell line codes referred to with which concentration they are mixed. For example, HT100% means HT-1080 sample with no THP-1 cell line-derived EVs. THP50%HT10% means a mix of THP-1 (diluted to 50%) and HT-1080 (diluted to 10%). a) Averaged spectra of the THP-1 and HT-1080 EVs (without mixing) were plotted overlaid with the spectra from mixture samples that include constant particles/mL of THP-1 EVs, and varying concentrations of HT-1080 EVs whose concentration values were increasing from 1% to 50%. b) Confusion matrix of the random forest classifier on the classes that correspond to the mixture spectra in panel (a).

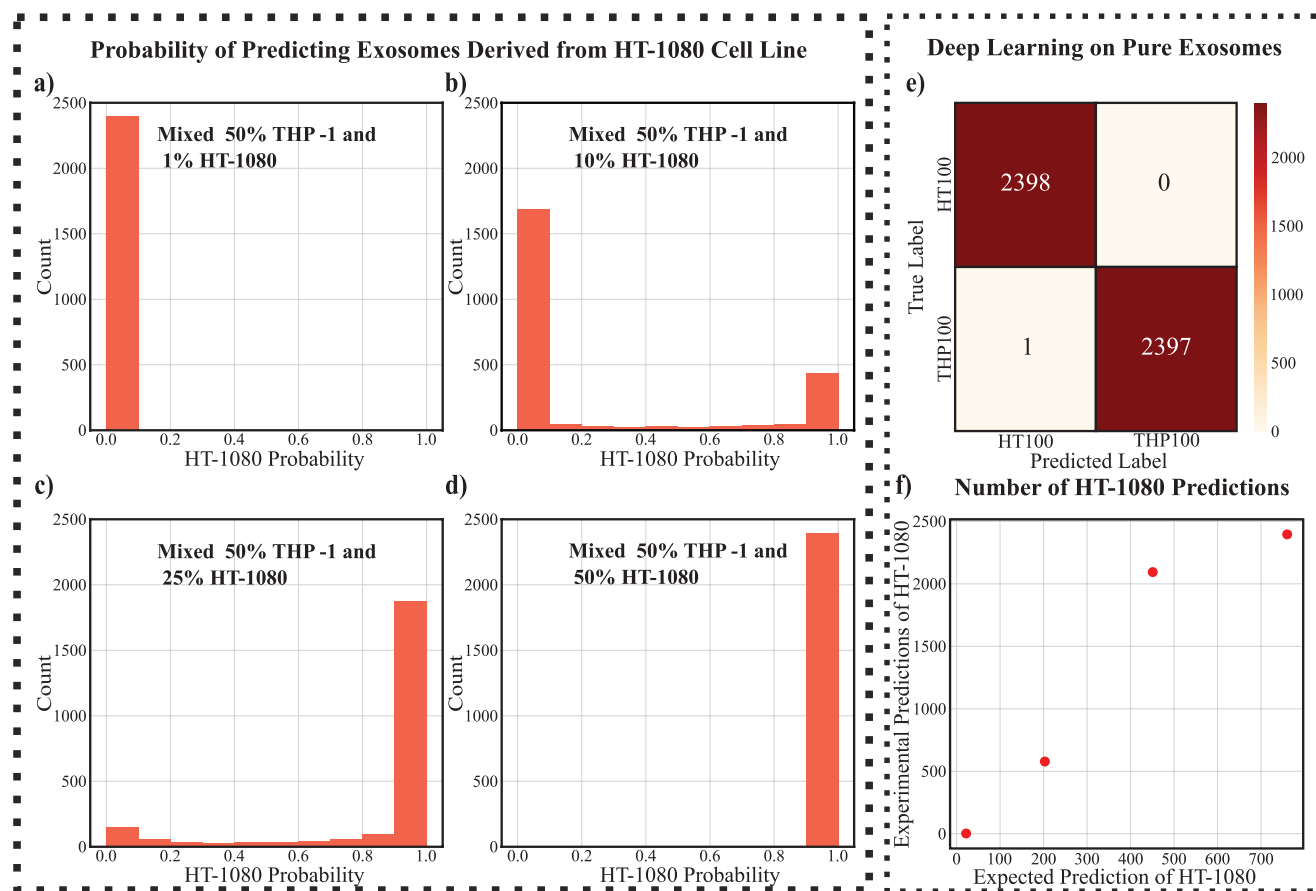


Figure 5. Neural network results in mixed spectra. a–d) ANN results for predicted HT-1080 EV counts. Bar graphs for each concentration show the probability distribution (softmax outputs) of the predicted labels corresponding to the HT-1080 EV. e) Confusion matrix for the ANN model. f) Scatter plot shows the increasing trend of the predicted number of HT-1080 EVs (from ANN) in the solutions against the number of predictions obtained from simulation.

only valid for the measurements with the same kind or batch of SERS probes. Otherwise, we can only get wavenumber information from the measurements from different probes. We showed in Figure S7 (Supporting Information) that the spectral frequencies at 922, 1001, 1126, 1224, 1254, 1305, 1380, 1447, 1595, 1660 cm^{-1} were shared for the responses from both probes. This result shows that a standardized surface is critical in SERS measurements to avoid false classification predictions. Additionally, the ANN results are similar to those obtained from the first probe (Figure S10, Supporting Information).

3.5. Spectral Differences Between Different Concentrations of EV Types

Figure 4a shows the average spectra of the EV mixtures where the concentration of THP-1 cell line-derived EVs was kept constant at 50% and the concentration of HT-1080 cell line-derived EVs were increased from 1 to 50%. Then, we analyzed spectral peaks that varied in average in Figure 4a (shown in Table 1). Peak intensities: Peak intensities at 996, 1026, 1066, 1120, 1183, and 1597 cm^{-1} increased in line with the number of HT-1080 EVs, while their adjacent peaks at 1002, 1029, 1070,

1123, 1180, and 1600 cm^{-1} decreased. These lower intensity peaks were tentatively assigned to the benzene ring of phenylalanine, CH₂-CH₃ bending of phospholipids, C-O stretching of phospholipids, C-C stretching of phospholipids, C-C (or C-O) stretching of lipids, and ergosterol. The intensity of CH deformations at 1344 cm^{-1} was higher.

3.6. Band Component Analysis: Spectral Analysis Using Gaussian Curve Fitting

We used a band component analysis to investigate the chemical composition of EVs that can lead to the differences observed in the ANN results. In this analysis, we precisely calculated the Raman shifts given in Table 1 using a Gaussian curve fitting algorithm. The decomposition of spectral band components revealed the unresolvable structures under the peaks that broaden due to the well-known spectral broadening schemes such as natural, Doppler, and pressure broadening. Therefore, we focused on those regions after finding the peaks that changed in either intensity or Raman shift. In Figure 6a–d, we compared the regions 970–1100 and 1390–1690 cm^{-1} for THP-1 and HT-1080 cell line-derived EVs, respectively. We employed

Table 1. Comparison of the varying peaks in the averaged spectra is given in Figure 4a. The comments in the change column are made according to the peak locations and intensities in the averaged spectrum that includes the least number of HT-1080 EVs.

Peak [cm ⁻¹]	Change	Assignment	Ref
722	Increase	Adenine, DNA	[59]
740	Increase	Tyminine, Ring breathing mode	[60]
796	Decrease	C-H ring out-of-plane deformation	[61]
816	Increase	PO2 stretch of nucleic acids	[62]
835	Shift	835: Asymmetric O-P-O stretching, tyrosine@@@ 840: Saccharide/amines	[60,63]
947	Increase	Proline, Polysaccharides	[64]
1002	Shift	996: C-O Ribose@@@ 1002: Benzene ring breathing, Phenylalanine	[57,65]
1025	Shift	1025: Carbohydrates peak@@@ 1030: C-H in-plane bending, Phe, Uracil, Phospholipids	[32,63,66,67]
1070	Shift	1066: Chain C-C stretching in lipids@@@ 1071: C-O stretching, Phospholipids	[63,68–71]
1180	Shift	1180: $\nu(\text{CC})$, $\nu(\text{CO})$ (e.g., phospholipids)@@@ 1183: Cytosine, adenine, guanine	[32,68]
1243	Increase	PO2 asymmetric stretching, pyrimidines	[62]
1305	Increase	CH3 CH2 twisting (e.g., lipid)	[32]
1370	Increase	Saccharides, T, A, G ring breathing	[60,72]
1450	Shift	1447: C-H vibration (proteins, lipids)@@@ 1450: CH2, CH3 def (proteins, lipids)	[63,73]
1578	Increase	Adenine, Guanine	[73]
1600	Shift	1600: Ergosterol@@@ 1595: C = O stretch in proteins	[55,56]
1662	Increase	Nucleic acids (purines)	[74]

ν : Symmetric stretching, δ : deformation

the Raman shift values of the decomposed peaks in the first region for different cell lines. For the EVs derived from HEC-1-A, HeLa, THP-1, MEF, and HT-1080, the center wavenumbers around 1030 cm⁻¹ were varied from 1033 to 1025 cm⁻¹ (Figure 6i). There were three peaks under the broad Raman band around 1030 cm⁻¹. The center of mass of this band depends on the dominance of the peak components. The centers of the components were around 1023, 1029, and 1034 cm⁻¹. Since the most dominant peak related to HT-1080 EVs was between 1026 and 1028 cm⁻¹ in this region, the center of mass of the broad peak converges toward these values for the spectra taken from HT-1080 EVs. Similarly, the most dominant peak for an HEC-1-A EV is around 1034 cm⁻¹ which led to a shoulder when HEC-1-A EVs are measured. These peak wavenumber shifts were the factors that cause a broad peak to be centered around different wavenumbers for each kind of EV.

We inserted dashed lines in Figure 6a–d, pointing out the differences in frequencies between THP-1 and HT-1080 cell-derived EVs. To evaluate whether these differences were prominent in the altering concentrations of HT-1080 EVs, we performed the same analysis on the spectra that have already been employed to predict the cellular origins of the analyzed EVs. When the regions with different profiles in the average spectra were examined, we found that as the concentration of HT-1080 increased, the peak center positions moved to lower Raman shift values. We validated this initial observation with our findings from the curve fitting results from the bands at 1025, 1070, 1450, and 1600 cm⁻¹, as seen in Figures 6e–h. These results were parallel to the band shift trends observed in Figures 6a–d. To quantify the shifting trends, we applied a power fit on the mean values of the peak center of the selected band components. R-square values were between 0.95 and 0.98 for the fittings.

The band center positions and areas for the most characteristic peaks were summarized in Table 2. In this table, we obtained the first two column values by curve fitting the averaged spectra of the THP-1 and HT-1080 EV preparations. As we showed in Table 2, the change in the Raman shift values observed in the unmixed exosomal spectra is similar to those in the mixed spectral data. As we increased the number of HT-1080-originated vesicles in the EV preparation, Raman shift values moved to lower wavenumbers for the peaks around 1025, 1070, 1450, and 1600 cm⁻¹. These band components were statistically significant according to the Mann-Whitney U non-parametric test ($p < 0.01$).

We developed a method that characterizes EVs as label-free and demonstrated that Raman spectroscopy can predict individual EV types secreted from different cells when mixed in preparations. We used a Raman spectroscopy platform, a machine learning algorithm, and defined Raman markers, to identify EVs originating from different cells. Further, we applied spectral analysis to investigate the Raman peaks that have prediction potential in the identification of the source of the EVs.

Our measurements and analysis with EVs derived from five different cell lines proved the recognition potential of the surface-enhanced Raman system. In addition, we obtained 96.8% accuracy from random forest analysis, proving that our model could also identify a mixture's cell origins. However, since we lacked class labels for a supervised analysis, we followed another path to validate recognizing different types of EVs. By training only unmixed EV samples, we estimated the concentrations of different EV mixtures. Furthermore, based on our simulation results, we demonstrated that we could obtain the prediction performance by comparing the EV ratio (in solution) to the predicted rate.

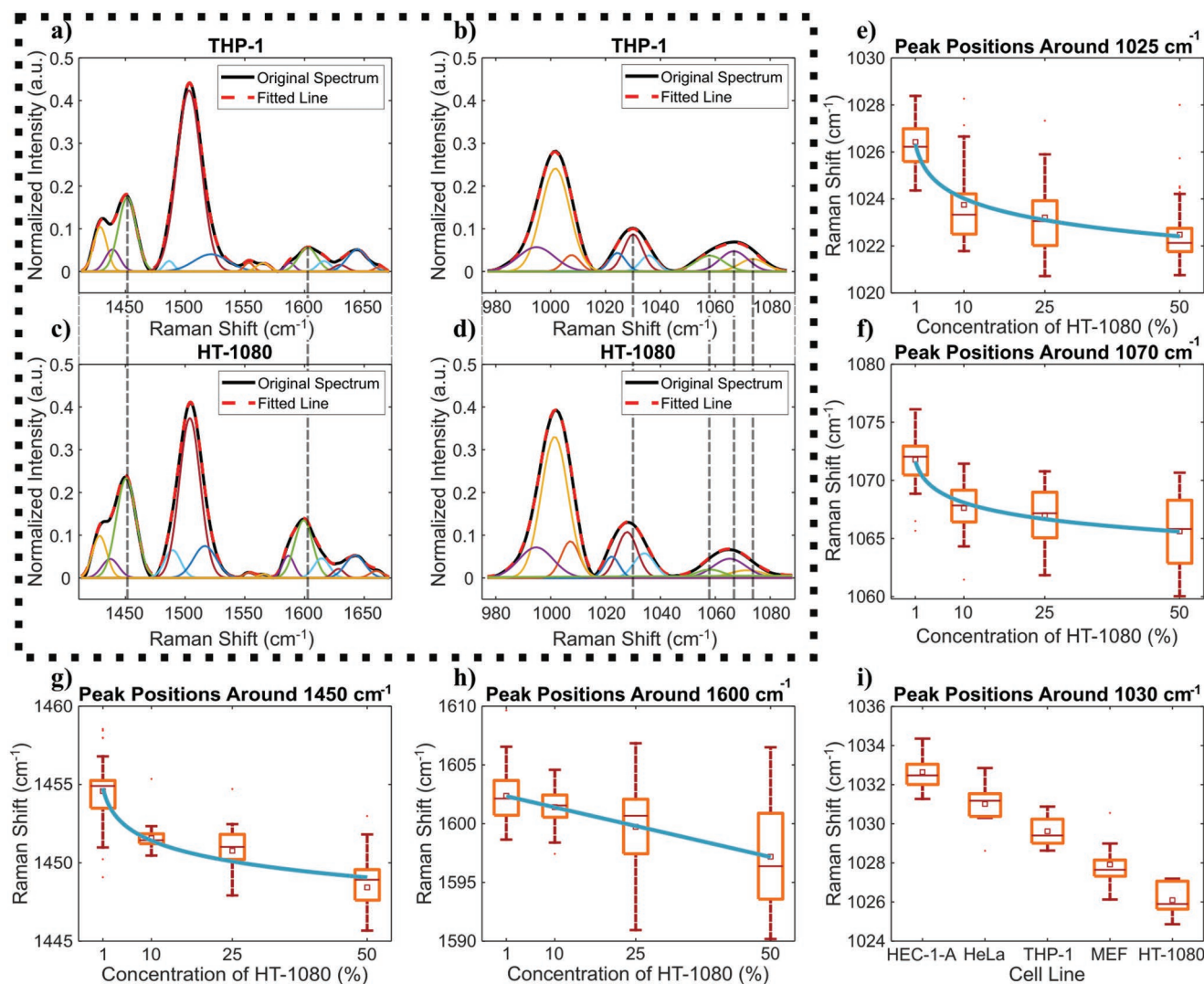


Figure 6. Band component analysis results. We performed a band component analysis on the measured and pre-processed spectra. We decomposed the sub-spectral components of the regions between 970–1100 cm^{-1} b,e) and 1390–1690 cm^{-1} a,d). The fit results in pictures (a) and (b) corresponded to the averaged spectrum of THP-1, while those in (d) and (e) corresponded to the averaged spectrum HT-1080. Graphs in (c), (f), (g), and (h) showed the differences between the two cell lines by utilizing Raman peak centers around 1025, 1070, 1450, and 1600 cm^{-1} . i) The peak center locations were around 1030 cm^{-1} for all cell lines.

When we examined the spectra of the prepared mixtures, we realized that the general behavior of the spectral intensities was not linear, as seen in the measurements of unmixed EVs. This non-linearity limited our ability to estimate the EV concentra-

tions that were not trained. To overcome this, we first tested our ability to predict known concentrations in the data for mixture measurements. We utilized a random forest algorithm to predict four classes of EV mixture preparation. Here, we could

Table 2. Summary of the most distinctive peaks obtained by band component analysis applied to HT-1080 and THP-1 samples.

THP-1 [Avg]		HT-1080 [Avg]		HT-1080 1%		HT-1080 50%	
RS [cm^{-1}]	A	RS [cm^{-1}]	A	RS [cm^{-1}]	A	RS [cm^{-1}]	A
1024.34	0.32	1022.27	0.34	1026.45	1.79 ± 0.01	1022.42	0.66 ± 0.01
1066.78	0.60	1065.23	0.68	1071.82	0.57 ± 0.01	1066.55	0.22 ± 0.01
1451.63	3.25	1450.62	4.82	1454.58	3.04 ± 1.12	1448.48	3.81 ± 1.10
1602.75	1.09	1599.29	2.57	1603.78	1.14 ± 0.17	1597.19	2.23 ± 0.07

RS: Raman Shift, A: Area

classify the samples even with a low number of variables. The confusion matrix given in Figure 4b represents a 99.58% accuracy in predicting the concentrations of the mixtures of different EVs. Although the model accuracy for the first probe is too high; we refrained from overfitting by using 100 estimators. Besides, we expect the model to give such outputs since shifts in the wavenumber axes create distinctive variables to differentiate the classes. As a result, the model gives results with a meaningful trend. Furthermore, the model confuses the mixtures with 1% and 10% HT-1080 EVs since there are (probably) not enough HT-1080 EVs to differentiate.

Next, we developed an ANN algorithm trained with unmixed EV data and tested it on the mixture data. We used the algorithm to determine the type of EV that the mixture signal mostly resembles. We aimed to capture the increase of HT-1080 concentration in the mixtures. We tried several regions to input the spectra. We found that the region between 880 and 1200 cm^{-1} has the most specific features since we obtained the best accuracy by employing the variables here. The bar graphs in Figure 5a–d show that as the concentration of HT-1080 EV type increases, there was an increase in the number of HT-1080 predictions.

On the other hand, as we showed in Figure 5c, the increase in predictions was proportional to the number of particles in the solution (data from HT-1080 with 25% concentration deviates from linearity). However, the distribution of predicted probabilities (between 0 and 1) was relatively sparse. Thus, the predominance of the predictions of THP-1 EVs on HT-1080 EVs rapidly changes direction. Besides, when the concentration of the HT-1080 is the highest, the model predicts no THP-1 EVs. This graph emphasizes that we need to consider the mixture behavior before using a statistical model in a clinical device used to diagnose cancer. In Figure 5, as the concentration of exosomes derived from the HT-1080 cell line increased, the predicted probability of HT-1080 classification also increased. However, the predicted probability of THP-1 decreased rapidly at the proportion of 50:50 due to the statistical dominance effects leading to misclassifications. In this case, the predicted probability of HT-1080 was 1.0, indicating that the likelihood of THP-1 being present in the solution is zero. Overall, nonlinear behaviors are expected since the laser's spot size was larger than the average size of the EVs, which allows a few EVs in the focal volume when the concentration is high.

The chemical components corresponding to the factors that lead to the identification were determined by interpreting the band component analysis results. As we examined the peak intensity changes in Table 1, we concluded that as the concentration of HT-1080 EVs increased, the RNA/DNA-related peaks became more abundant. Furthermore, as the peaks red/blueshift, it signifies the increase of a new chemical in the field of interest since the center of mass of the broad peak is shifting. As we demonstrated in Table 2, after applying the BCA for the regions 970–1100 and 1390–1690 cm^{-1} , we found out that, lipid/phospholipid (1066, 1071, 1180, 1309, 1450, 1600 cm^{-1}), and nucleic acid (1030, 1183, 996 cm^{-1}) cargo on/in EVs predominantly contribute to the recognition. The region between 970 and 1100 cm^{-1} includes some of the most important spectral components for exosomal identification. They have contributions from phospholipids, nucleic acids, and benzene-con-

taining amino acids such as phenylalanine. The bands at 1030 and 1060 cm^{-1} correspond to phospholipids which constitute the core structure of the vesicles. Our results showed that as the ratio of HT-1080 to THP-1 EVs increases, the band around 1600 cm^{-1} slowly decreases in intensity (band component analysis also proves this), and the peak around 1597 cm^{-1} increases. Kruglik et al. proposed that the band around 1600 cm^{-1} corresponds to ergosterols^[55] while the band around 1597 was tentatively assigned to C = O stretching in proteins.^[56] In the repetition experiments performed with another batch of SERS probes (same brand, given in Figure S6–S10, Supporting Information), we used the region between 1155 and 1660 cm^{-1} , which gave the best prediction accuracy. However, considering the intensity distribution and the most predictive spectral regions, we think both surfaces have different LSPR profiles. Thus, we did not compare their intensity and used the data from both batches in different datasets.

One of the most important advantages of our method is its ability to determine the cellular origin of the EVs detected. This advantage allows seeking cancer-cell-derived EVs in an individual's circulation without performing time-consuming and costly medical imaging. Previous studies that focus on characterizing EVs using Raman spectroscopy provided smart strategies to understand the role of EVs in cancer diagnosis. However, most of these studies utilized EV preparations from single cell-line types in reported works. Hence, the information with mixed populations is missing.^[28,31] Some studies succeeded in detecting tumor-derived EVs obtained from blood samples, but these studies used isolated samples during characterization, as well.^[32,53] On the other hand, our study identified cancerous EVs in both unmixed and mixed states. Thus, our approach is promising in non-invasive cancer diagnosis since it proposes to predict any previously learned cancerous EV type by examining it as a liquid biopsy method. On the other hand, this method can also be used to monitor metastatic EVs by self-applied daily measurements with portable Raman/SERS platforms.

In Figure 4, all of the experiments have been done by mixing two different cell lines. However, given how complex the human plasma is and to represent it more accurately, this method needs to be further studied in mixtures of four or more different cell lines so that the approach can be used with plasma. Further studies will also use the same basic principles to even test to predict random mixing concentrations of multiple cell-line derived exosomes in a standardized plasma background. These studies are currently planned/ongoing and will be reported later when they are complete. In the future, the current platform will be improved toward enhancing the resolution of Raman spectroscopy down to sub-100 nm. Thus, we will be able to eliminate the factors caused by signal collection from multiple EVs. On the other hand, a size-controlled experiment will help understand the size effect, improving the statistical models used in a clinical-level device. Furthermore, the EV interaction can be reduced by measuring less dilute samples, which requires a more sensitive detection system. Finally, to boost the performance of AI models, we plan to increase the number of mixtures with different ratios. This improvement will also enable us to train ANN models to perform a regression task.

4. Conclusion

We demonstrated that EVs derived from different cell lines mixed in the same sample could be ascribed to their original source using SERS and ANN-based classification. Molecular identification is essential since many EV characterization methods are size-based, and there are still fundamental questions about how the size distributions of the EVs are determined by the cell and what makes the difference between healthy and diseased states. Characterizing EV preparations and comparing individually characterized EVs are essential tasks already demonstrated by others^[23,24,27] and were achieved in our work. However, there are remaining challenges in the characterization of mixed preparations of EVs whose Raman spectra are similar. Furthermore, the EVs may interact with each other, further confounding the appropriate classification. We showed that this is achievable by SERS with the help of deep learning models. A limitation of this method was the lack of ground truth since optical visualizations of individual EVs stemming from are still challenging because of the diffraction limit. However, with the help of a numerical Monte Carlo simulation of the expected number of predictions of the sources of the EV origins, we found out that we successfully predicted the origin of source measurements by SERS analysis. On the other hand, we utilized a Gaussian curve fitting on the spectra to find the chemical origins that led to these successful predictions of EV classes.

We were challenged by the lack of repeatability stemming from the SERS probe heterogeneity in the surfaces we used. To overcome this problem, we compared the measurements taken with the SERS probes from the same batch.

Our approach brings the EV field one step closer to the practical applications of cancer diagnostics. From a clinical standpoint, the label-free characterization of exosomes will offer new avenues for cancer detection, diagnosis, and clinical management.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors thank Ahmet Melek for his suggestions during the machine learning analysis and Batuhan Govce for his help in the additional experiments. The authors thank Prof Dr. Ahmet Gul, who provided the THP-1 cells. The authors gratefully acknowledge the use of the services and facilities of the Koc University Research Center for Translational Medicine (KUTTAM). The authors also thank the technical support of Dr. Ozgur Albayrak (Flow cytometry experiments) and Zafer Eroglu (TEM experiments) from Koc University. BU acknowledges support from TUBITAK (Project number:18S1113) and the Republic of Turkey Ministry of Industry and Technology (Project number:2009K120520). UD acknowledges Canary foundation Seed Grant and support from Precision Health and Integrated Diagnostics (PHIND) Center. [Correction added after publication 1 March 2023: Figure 2 was updated.]

Conflict of Interest

Prof. Utkan Demirci (UD) is a founder of and has an equity interest in: (i) DxNow Inc., a company that is developing microfluidic IVF tools and

imaging technologies, (ii) Koek Biotech, a company that is developing microfluidic technologies for clinical solutions, (iii) Levitas Inc., a company focusing on developing microfluidic sorters using magnetic levitation, (iv) Hillel Inc., and Hilla Biotech, companies bringing microfluidic cell phone tools for spermograms to home settings, and (v) Mercury Biosciences, a company developing vesicle isolation technologies. UD's interests were viewed and managed in accordance with the conflict-of-interest policies.

Data Availability Statement

The data that support the findings of this study are openly available in zenodo at <https://doi.org/10.5281/zenodo.7011380>, ref. [58] 7011379.

Keywords

exosome, extracellular vesicles, neural networks, Raman spectroscopy

Received: September 6, 2022

Revised: December 2, 2022

Published online: January 15, 2023

- [1] C. Théry, S. Amigorena, G. Raposo, A. Clayton, *Curr. Protoc. Cell Biol.* **2006**, 30, 3.22.1.
- [2] R. Kalluri, V. S. Lebleu, *Science* **2020**, 367, aau6977.
- [3] B. Zhou, K. Xu, X. Zheng, T. Chen, J. Wang, Y. Song, Y. Shao, S. Zheng, *Signal Transduction Targeted Ther.* **2020**, 5, 1.
- [4] C. Théry, K. W. Witwer, E. Aikawa, M. J. Alcaraz, J. D. Anderson, R. Andriantsitohaina, A. Antoniou, T. Arab, F. Archer, G. K. Atkin-Smith, D. C. Ayre, J.-M. Bach, D. Bachurski, H. Baharvand, L. Balaj, S. Baldacchino, N. N. Bauer, A. A. Baxter, M. Bebawy, C. Beckham, A. Bedina Zavec, A. Benmoussa, A. C. Berardi, P. Bergese, E. Bielska, C. Blenkiron, S. Bobis-Wozowicz, E. Boilard, W. Boireau, A. Bongiovanni, J. *Extracell. Vesicles* **2018**, 7, 1535750.
- [5] J. Kowal, G. Arras, M. Colombo, M. Jouve, J. P. Morath, B. Primdal-Bengtson, F. Dingli, D. Loew, M. Tkach, C. Théry, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, 113, E968.
- [6] M. K. Jung, J. Y. Mun, *JoVE (Journal of Visualized Experiments)* **2018**, e56482.
- [7] S. Sharma, H. I. Rasool, V. Palanisamy, C. Mathisen, M. Schmidt, D. T. Wong, J. K. Gimzewski, *ACS Nano* **2010**, 4, 1921.
- [8] V. Sokolova, A.-K. Ludwig, S. Hornung, O. Rotan, P. A. Horn, M. Eppe, B. Giebel, *Colloids Surf., B* **2011**, 87, 146.
- [9] S. A. Melo, L. B. Luecke, C. Kahlert, A. F. Fernandez, S. T. Gammon, J. Kaye, V. S. Lebleu, E. A. Mittendorf, J. Weitz, N. Rahbari, C. Reissfelder, C. Pilarsky, M. F. Fraga, D. Piwnicka-Worms, R. Kalluri, *Nature* **2015**, 523, 177.
- [10] K. Laulagnier, H. Vincent-Schneider, S. Hamdi, C. Subra, D. Lankar, M. Record, *Blood Cells, Mol., Dis.* **2005**, 35, 116.
- [11] J. N. Higginbotham, Q. Zhang, D. K. Jeppesen, A. M. Scott, H. C. Manning, J. Ochieng, J. L. Franklin, R. J. Coffey, *J. Extracell. Vesicles* **2016**, 5, 29254.
- [12] A. Radbruch, *Flow cytometry and cell sorting*, Springer, Berlin **2000**.
- [13] N. Baumgarth, M. Roederer, *J. Immunol. Methods* **2000**, 243, 77.
- [14] D. A. Long *Raman spectroscopy*. New York **1977**, pp. 1–12.
- [15] A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari, M. S. Feld, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 12371.
- [16] W. Lu, A. K. Singh, S. A. Khan, D. Senapati, H. Yu, P. C. Ray, *J. Am. Chem. Soc.* **2010**, 132, 18103.
- [17] A. Pallaoro, M. R. Hoonejani, G. B. Braun, C. D. Meinhardt, M. Moskovits, *ACS Nano* **2015**, 9, 4328.

- [18] R. M. Jarvis, R. Goodacre, *Anal. Chem.* **2004**, 76, 40.
- [19] S. Shanmukh, L. Jones, Y.-P. Zhao, J. D. Driskell, R. A. Tripp, R. A. Dluhy, *Anal. Bioanal. Chem.* **2008**, 390, 1551.
- [20] B. R. Wood, P. Caspers, G. J. Puppels, S. Pandiancherri, D. McNaughton, *Anal. Bioanal. Chem.* **2007**, 387, 1691.
- [21] T. J. Moritz, D. S. Taylor, D. M. Krol, J. Fritch, J. W. Chan, *Biomed. Opt. Express* **2010**, 1, 1138.
- [22] N. M. Culum, T. T. Cooper, G. I. Bell, D. A. Hess, F. Lagugné-Labarthe, *Anal. Bioanal. Chem.* **2021**, 413, 5013.
- [23] M. Dash, K. Palaniyandi, S. Ramalingam, S. Sahabudeen, N. S. Raja, *Biochim. Biophys. Acta Biomembr.* **2021**, 1863, 183490.
- [24] A. Al Mannai, Y. Haik, A. Elmel, S. Qadri, K. M. Saud, *Emergent Mater.* **2021**, 1579.
- [25] J. Carmicheal, C. Hayashi, X. Huang, L. Liu, Y. Lu, A. Krasnoslobodtsev, A. Lushnikov, P. G. Kshirsagar, A. Patel, M. Jain, Y. L. Lyubchenko, Y. Lu, S. K. Batra, S. Kaur, *Nanomedicine* **2019**, 16, 88.
- [26] A. Merdalimova, V. Chernyshev, D. Nozdriukhin, P. Rudakovskaya, D. Gorin, A. Yashchenok, *Appl. Sci.* **2019**, 9, 1135.
- [27] S. Stremersch, M. Marro, B.-E. Pinchasik, P. Baatsen, A. Hendrix, S. C. De Smedt, P. Loza-Alvarez, A. G. Skirtach, K. Raemdonck, K. Braeckmans, *Small* **2016**, 12, 3292.
- [28] Z. Yan, S. Dutta, Z. Liu, X. Yu, N. Mesgarzadeh, F. Ji, G. Bitan, Y.-H. Xie, *ACS Sens.* **2019**, 4, 488.
- [29] P. Zhang, L. Wang, Y. Fang, D. Zheng, T. Lin, H. Wang, *Molecules* **2019**, 24, 2947.
- [30] J. Park, M. Hwang, B. Choi, H. Jeong, J.-h. Jung, H. K. Kim, S. Hong, J.-h. Park, Y. Choi, *Anal. Chem.* **2017**, 89, 6695.
- [31] P. Xanthopoulos, P. M. Pardalos, T. B. Trafalis, in *Robust data mining*, Springer, New York **2013**, pp. 27–33.
- [32] H. Shin, S. Oh, S. Hong, M. Kang, D. Kang, Y.-g. Ji, B. H. Choi, K.-W. Kang, H. Jeong, Y. Park, S. Hong, H. K. Kim, Y. Choi, *ACS Nano* **2020**, 14, 5435.
- [33] H. Qian, X. Shao, H. Zhang, Y. Wang, S. Liu, J. Pan, W. Xue, *Spectrochim. Acta, Part A* **2022**, 281, 121603.
- [34] Y. Xie, X. Su, Y. Wen, C. Zheng, M. Li, *Nano Lett.* **2022**, 22, 7910.
- [35] Y.-F. Qin, X.-Y. Lu, Z. Shi, Q.-S. Huang, X. Wang, B. Ren, L. Cui, *Anal. Chem.* **2022**, 94, 12416.
- [36] L. Guerrini, E. Garcia-Rico, A. O'loghlen, V. Giannini, R. A. Alvarez-Puebla, *Cancers* **2021**, 13, 2179.
- [37] R. E. Reggiardo, S. V. Maroli, H. Halasz, M. Ozen, D. Carrillo, E. LaMontagne, L. Whitehead, E. Kim, S. Malik, J. Fernandes, G. Marinov, E. Collisson, U. Demirci, D. H. Kim, *bioRxiv* **2020**.
- [38] F. Liu, U. Demirci, S. S. Gambhir, *US Patent 11,073,511*, **2019**.
- [39] U. Aygun, A. Y. Ozkumur, N. G. Durmus, U. Demirci, H. Urey, Label-free imaging of exosomes using depth scanning correlation (DSC) interferometric microscopy. Label-free Biomedical Imaging and Sensing (LBIS) **2020**, p. 112510Q.
- [40] K. Yuan, E. A. Shamskhou, M. E. Orcholski, A. Nathan, S. Reddy, H. Honda, V. Mani, Y. Zeng, M. O. Ozen, L. Wang, U. Demirci, W. Tian, M. R. Nicolls, V. A. De Jesus Perez, *Circulation* **2019**, 139, 1710.
- [41] C. G. O'Brien, M. O. Ozen, E. Vaskova, M. Santoso, C. Wahlquist, J. H. Jung, G. Ikeda, L. Shi, U. Demirci, M. Mercola, P. C. Yang, *Circulation* **2019**, 140, A12922.
- [42] A. Mataji-Kojouri, M. O. Ozen, M. Shahabadi, F. Inci, U. Demirci, *ACS Nano* **2020**, 14, 8518.
- [43] M. Chandy, J.-W. Rhee, M. O. Ozen, D. R. Williams, L. Pepic, C. Liu, H. Zhang, J. Malisa, E. Lau, U. Demirci, J. C. Wu, *Circulation* **2020**, 142, 1794.
- [44] P. Heraud, B. R. Wood, J. Beardall, D. McNaughton, *J. Chemom.* **2006**, 20, 193.
- [45] N. K. Afseth, V. H. Segtnan, J. P. Wold, *Appl. Spectrosc.* **2006**, 60, 1358.
- [46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, *J. Mach. Learn. Res.* **2011**, 12, 2825.
- [47] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.
- [48] F. Chollet, keras. **2015**.
- [49] The Inkscape Project, Inkscape, <https://inkscape.org> (accessed: May 2019).
- [50] F. Ito, K. Kato, I. Yanatori, T. Murohara, S. Toyokuni, *Redox Biol.* **2021**, 47, 102174.
- [51] J. M. Carnino, K. Ni, Y. Jin, *Front. Immunol.* **2020**, 11, 948.
- [52] N. Tominaga, K. Hagiwara, N. Kosaka, K. Honma, H. Nakagama, T. Ochiya, *Mol. Cancer* **2014**, 13, 1.
- [53] G. Tsoumakas, I. Katakis, *Int. J. Data Warehous. Min.* **2007**, 3, 1.
- [54] L. Qiao, S. Hu, K. Huang, T. Su, Z. Li, A. Vandergriff, J. Cores, P.-U. Dinh, T. Allen, D. Shen, H. Liang, Y. Li, K. Cheng, *Theranostics* **2020**, 10, 3474.
- [55] S. G. Kruglik, F. Royo, J.-M. Guigner, L. Palomo, O. Seksek, P.-Y. Turpin, I. Tatischeff, J. M. Falcón-Pérez, *Nanoscale* **2019**, 11, 1661.
- [56] W. Zheng, W. Huang, J. Chen, J. Zhou, X. Xu, G. Xu, R. Chen, G. Wang, S. Zhang, *Int. J. Clin. Exp. Med.* **2018**, 11, 12991.
- [57] A. Gualerzi, S. Picciolini, C. Carlomagno, F. Terenzi, S. Ramat, S. Sorbi, M. Bedoni, *Nanomedicine* **2019**, 22, 102097.
- [58] U. Parlattan, M. O. Ozen, I. Kecoglu, B. Koyuncu, H. Torun, D. Khalafkhan, I. Loc, M. G. Ogut, F. Inci, D. Akin, I. Solaroglu, N. Ozoren, M. B. Unlu, U. Demirci, Label-Free Identification of Exosomes Using Raman Spectroscopy and Machine Learning, **2022**.
- [59] J. Binoy, J. P. Abraham, I. H. Joe, V. S. Jayakumar, G. R. Pettit, O. F. Nielsen, *J. Raman Spectrosc.* **2004**, 35, 939.
- [60] J. W. Chan, D. S. Taylor, T. Zwerdling, S. M. Lane, K. Ihara, T. Huser, *Biophys. J.* **2006**, 90, 648.
- [61] J. D. Driskell, A. G. Seto, L. P. Jones, S. Jokela, R. A. Dluhy, Y.-P. Zhao, R. A. Tripp, *Biosens. Bioelectron.* **2008**, 24, 917.
- [62] W.-T. Cheng, M.-T. Liu, H.-N. Liu, S.-Y. Lin, *Microsc. Res. Tech.* **2005**, 68, 75.
- [63] Z. J. Smith, C. Lee, T. Rojalin, R. P. Carney, S. Hazari, A. Knudson, K. Lam, H. Saari, E. L. Ibañez, T. Viitala, T. Laaksonen, M. Yliperttula, S. Wachsmann-Hogiu, *J. Extracell. Vesicles* **2015**, 4, 28533.
- [64] M. Gniadecka, H. C. Wulf, N. Nymark Mortensen, O. Faurskov Nielsen, D. H. Christensen, *J. Raman Spectrosc.* **1997**, 28, 125.
- [65] J. C. Fraire, S. Stremersch, D. Bouckaert, T. Monteyne, T. De Beer, P. Wuytens, R. De Rycke, A. G. Skirtach, K. Raemdonck, S. De Smedt, K. Braeckmans, *ACS Appl. Mater. Interfaces* **2019**, 11, 39424.
- [66] S. E. J. Bell, N. M. S. Sirimuthu, *J. Am. Chem. Soc.* **2006**, 128, 15580.
- [67] A. Jaworska, L. E. Jamieson, K. Malek, C. J. Campbell, J. Choo, S. Chlopicki, M. Baranska, *Analyst* **2015**, 140, 2321.
- [68] N. Stone, C. Kendall, J. Smith, P. Crow, H. Barr, *Faraday Discuss.* **2004**, 126, 141.
- [69] L. Chiriboga, P. Xie, H. Yee, V. Vigorita, D. Zarou, D. Zakim, M. Diem, *Biopolymers* **1998**, 4, 47.
- [70] L. Silveira, S. Sathiaiah, R. A. Zangaro, M. T. T. Pacheco, M. C. Chavantes, C. A. G. Pasqualucci, *Lasers Surg. Med.* **2002**, 30, 290.
- [71] U. Ladiwala, A. Bankapur, B. Thakur, C. Santhosh, D. Mathur, *arXiv:1401.7497* **2014**.
- [72] C. Krafft, L. Neudert, T. Simat, R. Salzer, *Spectrochim. Acta, Part A* **2005**, 61, 1529.
- [73] I. Nottingher, C. Green, C. Dyer, E. Perkins, N. Hopkins, C. Lindsay, L. L. Hench, *J. R. Soc., Interface* **2004**, 1, 79.
- [74] A. Mahadevan-Jansen, R. Richards-Kortum Raman spectroscopy for cancer detection: a review. *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 'Magnificent Milestones and Emerging Opportunities in Medical Engineering'* (Cat. No. 97CH36136). **1997**, pp 2722–2728.