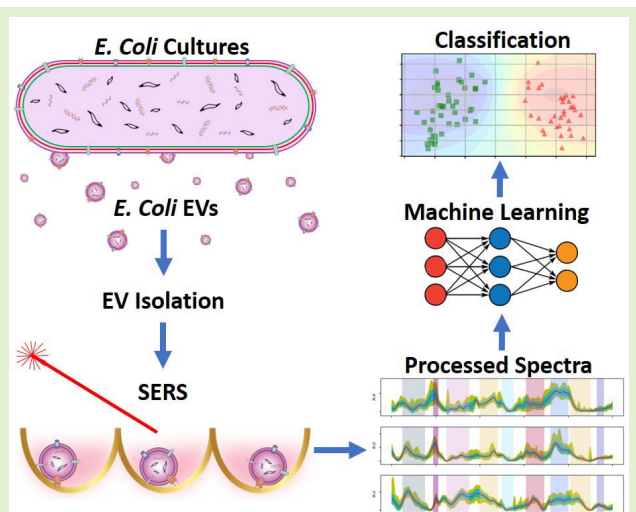# Label-Free Classification of Bacterial Extracellular Vesicles by Combining Nanoplasmonic Sensors With Machine Learning

Mohammadrahim Kazemzadeh[iD], Colin L. Hisey[iD], Priscila Dauros-Singorenko, Simon Swift[iD], Kamran Zargar-Shoshtari, Weiliang Xu[iD], *Senior Member, IEEE*, and Neil G. R. Broderick

***Abstract*—Bacterial extracellular vesicles (EVs) are nano-scale lipid-enclosed packages that are released by bacteria cells and shuttle various biomolecules between bacteria or host cells. They are implicated in playing several important roles, from infectious disease progression to maintaining proper gut health, however the tools available to characterise and classify them are limited and impractical for many applications. Surface-enhanced Raman Spectroscopy (SERS) provides a promising means of rapidly fingerprinting bacterial EVs in a label-free manner by taking advantage of plasmonic resonances that occur on nanopatterned surfaces, effectively amplifying the inelastic scattering of incident light. In this study, we demonstrate that by applying machine learning algorithms to bacterial EV SERS spectra, EVs from cultures of the same bacterial species (*Escherichia coli*) can be classified by strain, culture conditions, and purification method. While these EVs are highly purified and homogeneous compared to complex samples, the ability to classify them from a single species demonstrates the incredible power of SERS when combined with machine learning, and the importance of considering these parameters in future applications. We anticipate that these findings will play a crucial role in developing the laboratory and clinical utility of bacterial EVs, such as the label-free, noninvasive, and rapid diagnosis of infections without the need to culture samples from blood, urine, or other fluids.**

***Index Terms*— Plasmonic, SERS, biosensor, extracellular vesicles, Raman spectroscopy, nonlinear optics, *Escherichia coli*, outer membrane vesicles, exosomes.**

I. INTRODUCTION

RAMAN spectroscopy is a unique and powerful tool for noninvasively fingerprinting the chemical structure of materials by measuring the inelastic scattering of incident light. This technique has been used extensively for biological applications, from cancer diagnosis and classification [1]–[3] to viral and bacterial identification [4]–[6]. The investigation of *Escherichia coli* using Raman spectroscopy is of particular interest, given this bacterium's complex role in both maintaining and adversely affecting human health; with applications pursued to detect it in food and drinking water [7], [8], investigate its resistance to antibiotics, and classify it based on its strains or pathogenic potential [9]. The need for label-free, rapid, and highly sensitive identification is especially evident when considering that some *E. coli* strains can lead to high human health risk even at very low doses [10], [11]. However, due to the bacteria's small sizes

and rarity in complex environmental and human bodily fluids, it can be challenging to effectively identify or classify them using conventional Raman spectroscopy.

To overcome the limitations of conventional Raman spectroscopy, techniques such as Surface-enhanced Raman Spectroscopy (SERS) [12] and laser tweezers Raman spectroscopy (LTRS) [13] have been applied to characterising bacteria cells. SERS benefits from strong plasmonic interactions of incident laser light with the electrons of noble metals on a substrate's surface [14], [15], creating strong localised electric fields at specific hot spots by confining the light in the nanoscale geometric surface features. In doing so, any nonlinear light-material interactions, including Stokes and anti-Stokes inelastic scattering, are substantially amplified. This technique is capable of identifying biomolecules at extremely low concentrations and is even capable of single molecule detection. Plasmonic nanoparticles conjugated with antibodies have also been used in immunoassays for the selective detection of *E. coli* from complex fluid samples [16]. In addition, SERS has been used for the label-free detection of *E. coli* strains by the direct acquisition of the Raman signal of cells and subsequent multivariate analysis techniques [17].

Recently, there has been growing interest in using SERS for both labeled [18]–[23] and label-free [24]–[29] characterisation of mammalian cell-derived extracellular vesicles (EVs). Bacteria, as their eukaryotic counterparts, also release EVs, which are lipid-enclosed nanoscale packages that are derived from the membranes of parental cells and can harbour a diverse molecular cargo such as proteins, DNA, RNA, glycolipids, organic small molecules, etc [30]. Therefore, analysing bacterial EVs using SERS could provide valuable information related to a range of biomolecules and help identify the condition of the parental bacteria cells. Although EV isolation techniques have become more rapid, less laborious, and require less expensive or specialised equipment, leading to the development of novel downstream characterisation tools including SERS, bacterial EV research itself still faces a multitude of experimental-dependent factors. Most bacterial EV research has been performed in vitro, where hundreds of different culture conditions (e.g. media, growth mode or stage, temperature, oxygen, $+/-$ stimulus) along with different EV isolation methods can produce significantly different EV compositions, mainly quantified based on protein [31]–[33] and RNA [34]–[36] content.

Recently, SERS has proven to be a promising method for the early detection and classification of various diseases such as cancer [37], [38]. As all bacteria release EVs into their surrounding space, translating into numerous biological roles in their host or their own life cycles [39]–[50], circulating bacterial EVs may also hold potential as a substrate for infection detection and identification. Importantly, the rapid growth pace of bacteria allows cells to release EVs in quick response to changes in environmental cues, resulting in EV molecular compositions that may represent specific temporal or environmental conditions [32]. The bacterial EV field has focused mainly on conventional proteomics-based methods of bacterial EV characterisation, as bacterial EV proteins

have shown to be diverse, abundant, highly responsive to environmental conditions, and informative about biogenesis mechanisms or potential virulence roles [51]. More recently, RNA-seq has also provided the ability to characterise bacterial EV RNA content. However, RNA's variable presence or low abundance in EVs, combined with time-consuming methods and advanced equipment or skills needed have restricted efficient progress. Clearly, there is a need for a platform which enables analysis and profiling across all the molecular cargo present in bacterial EVs, and SERS could provide a powerful, comprehensive, and sensitive tool to characterise and classify them, or even classify their parental bacteria cells' identity or condition.

To date, only one recent study demonstrated the use of conventional Raman spectroscopy for characterising bacterial EVs, where changes in the chemical profiles of *Pseudomonas chlororaphis* O6 (PcO6) cells and resultant EVs were identified according to different abiotic stressors. In addition, they showed that the bacterial EVs had higher relative concentrations of proteins, lipids, and nucleic acids than PcO6 cells [52]. However, to the best of our knowledge, no studies to date have examined whether bacterial EVs could be successfully fingerprinted or classified using SERS.

In this study, we demonstrate for the first time, that by combining label-free SERS with machine learning, *E. coli* EVs can be classified based on differences in strain, culture condition, and purification method. Probiotic (Nissle), uropathogenic (UPEC), and laboratory model (K12) *E. coli* strains were classified in both pairwise and collective approaches. In addition, manifold machine learning proved to be a viable method of dimension reduction for EV SERS spectra visualisation. We believe that these results will encourage and support other researchers towards utilising SERS for bacterial EV characterisation in a multitude of laboratory and clinical applications.

II. Materials and Methods

A. Bacteria Culture and EV Isolation

Three *Escherichia coli* strains were used for this study: Uropathogenic *E. coli* (UPEC) strain 536 (O6:K15:H31) [53], probiotic Nissle 1917 [54] and laboratory model strain MG1655 (K-12, ATCC® 47076) [55]. Culture and EV isolation methods have been previously published in detail [32]. Briefly, bacterial cells were grown in either of two iron conditions: iron restricted, in plain RPMI 1640 medium (R) (Thermo Fisher Scientific) or iron sufficient, in RPMI medium supplemented with 10 μM iron(III) chloride (RF) to better reflect physiological conditions. At the desired incubation time, bacterial cells were removed from the culture by centrifugation and filtration. Cell-free EV-containing supernatants were concentrated to smaller volumes with 100 kDa Vivaflow 200 cassettes (Sartorius AG) and EVs were pelleted by ultracentrifugation at 75,000 $\times g$ for 2.5 h at 4 °C (Beckman Avanti J-30I), then resuspended in PBS. EVs were then further purified with either of two well-established purification methods: Density
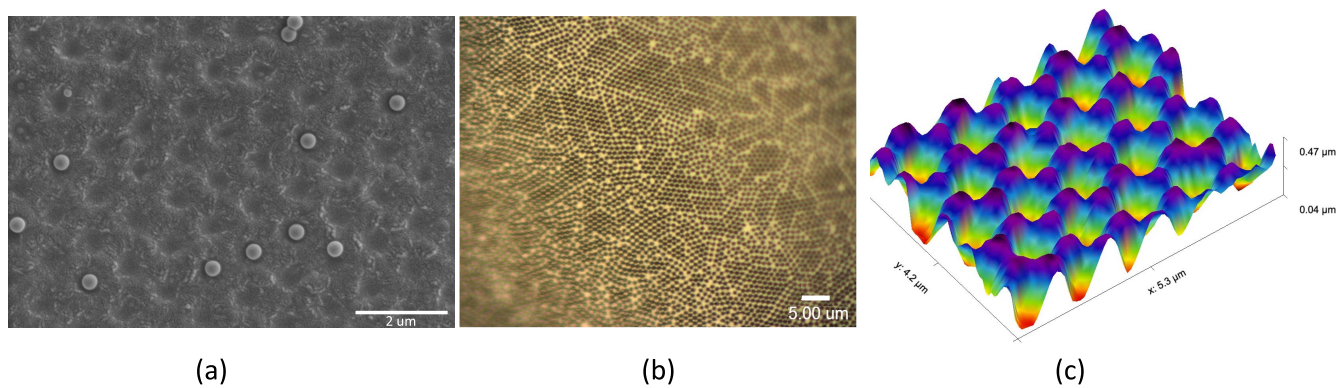
Fig. 1. Nanocup SERS substrate characterisation showing (a) SEM of 300 nm beads trapped in the nanocup SERS hot spots (scale bar = 4 μm), (b) optical microscopy image of tightly packed nanocup structures across the entire surface (scale bars =5 μm), and (c) AFM of the nanocup SERS demonstrating the desired structure.

Gradient Centrifugation (DG) using an iodixanol (Optiprep, SigmaAldrich) gradient or Size Exclusion Chromatography (SEC) using a qEV Original column (70 nm, Izon Science) [32]. For the remainder of this study, EV samples are referred to using the notation "Strain-Culture Medium-Purification Method". For example, EVs from UPEC cells grown with iron supplementation and purified with size exclusion chromatography are referred to as "UPEC-RF-SEC".

B. EV Characterisation

EV-rich fractions from both SEC and DG purification methods were determined by protein (Pierce BCA Protein Assay, ThermoFisher) and particle quantification using nanoparticle tracking analysis (NTA) using an NS300 Nanosight (Malvern Panalytical), then pooled for analysis. Once EV-rich fractions were pooled, they were diluted at a 1:250 ratio in PBS and three 30 second videos were taken under low flow conditions (Screen gain:2, Camera level:14) and characterised using the Nanosight 3.4 software (Screen gain:10, Detection threshold:6) to calculate mean and mode particle diameters, concentration, and size distributions. Prior to TEM or SERS, EV samples were transferred from PBS buffer to ultrapure water by loading 200 μl of the purified EVs into a Vivaspin 500 (Sartorius AG) centrifugal concentrator with a 100 kDa cutoff and centrifuging at 10,000 $\times g$ until most of the PBS had flowed through the filter (roughly 10 minutes). Ultrapure water (ThermoFisher), 450 μl, was then added to the filter and the centrifugation process was repeated twice more before finally suspending the EVs in 100 μl ultrapure water. Negative staining TEM of purified EVs was conducted by adsorption onto Formvar-coated copper grids (Electron Microscopy Sciences) for 10 minutes. Excess liquid was removed with filter paper (Whatman) and the copper grid was then transferred to 20 μL of 2% filtered uranyl acetate for 2 minutes. Excess liquid was again removed with filter paper and the grid was allowed to dry under a lamp for 10 minutes. Grids were visualised on Tecnai G2 Spirit TWIN (FEI, Hillsboro, OR, USA) transmission electron microscope (TEM) at 120 kV accelerating voltage. Images were captured using a Morada digital camera (SIS GmbH, Munster, Germany).

C. SERS Preparation, Characterisation, and Measurements

A gold plasmonic microresonator was fabricated as previously reported [56]–[59]. Briefly, 1 μm polystyrene beads (Sigma) were carefully spread onto glass slides and baked, then a PDMS mold was taken of the resulting bead monolayer. Following demolding, the PDMS was sonicated in acetone and isopropyl alcohol to remove residual beads and baked overnight. Gold was deposited by sputter coating (Q150R S, Quorum) with the rate of 0.5 nm per minute until a 12.5 nm thickness was achieved, then the deposition rate was increased to 6 nm per minute until a final thickness of 50 nm. Moreover, the reproducibility of this SERS for chemical species and EV samples of similar size have been experimentally validated previously [58]. This SERS design was found to have more than one million times Raman enhancement and works effectively to trap EVs in its hotspot areas [24]. Scanning electron microscopy (SEM) of 300 nm diameter polystyrene beads released on the SERS substrate was performed using a Hitachi SU-70 SEM to illustrate its trapping capability for EV-sized particles (Fig. 1 (a)) and optical microscopy demonstrates the consistent and tightly packed patterning of cup-shaped structures across the entire surface (Fig. 1 (b)). Atomic force microscopy (AFM) was also performed to further validate the nanocup structure of the fabricated SERS substrate using an Asylum Cypher ES AFM system.

For all SERS measurements, a total volume of 20 μl, resulting in roughly 1 μl of EVs (in an ultrapure water suspension) per square millimeter with the concentration of $\approx 1 \times 10^{10}$ EVs per millilitre, were dropped on the SERS surface and dried quickly in a 40 °C oven (20 square millimeters of EV coating total). This was done to increase the homogeneity of the EVs' distribution on the SERS surface by minimizing the coffee ring effect, which would be worse if the sample was left to air dry at ambient conditions [24]. Raman measurements were carried out using a Horiba LabRAM HR Evolution confocal Raman microscope by using 785 nm laser and 50× microscope objective. 50 SERS spectra were taken for each EV sample from 800-1800 cm^{-1}, with a minimum distance

of one laser spot size between acquisition locations. The laser power at the surface of SERS was controlled using neutral density filter and set to 10 percent of the maximum power (100 mW). A 10 sec acquisition time for the detector was chosen per measurement. Then, the baseline was established and noise was removed automatically using previously established asymmetric least squares smoothing [60].

D. Machine Learning Analyses

All machine learning and multivariate analyses were performed using the established Python library, Sklearn [61]. Some rationale behind the classification methods are explained within results and discussion section for added clarity.

III. RESULTS AND DISCUSSION

A. EV Characterisation

NTA data demonstrated that all UPEC EVs exhibited similar sizes, with an average mean diameter of 115.4 +/− 6.5 nm and mode of 89.7 +/− 4.5 nm (Fig. 2 (a)). Interestingly, the NTA size distribution for Nissle-R-SEC EVs showed noticeably larger particles with a mean diameter of 135.8 nm and mode of 101.2 nm, while K12-RF-SEC yielded smaller EVs mean diameter of 71.1 nm and a mode of 62.9 nm. Negative staining TEM imaging demonstrated similar EV morphologies for all samples, and representative TEM images for EVs from each *E. coli* strain are shown in Fig. 2 (b). However, the slight differences in particle size measured by NTA were not apparent in the TEM images, potentially due to the random adsorption to the copper grids.

The post-processed SERS spectra for all *E. coli* EVs investigated in this work are shown in Fig. 2 (c). These spectra were normalised using euclidean normalization and thus any information related to the amplitude of SERS spectra were removed. The mean, maximum/minimum and standard deviation around the mean was calculated and presented with red, yellow and cyan colour, respectively. The correspondence of Raman bands to the known types of biomolecules are shown using vertical colour bands [62]. As most EV spectral peaks are the result of contributions from the chemical bonds of many different biochemical species in combination, only general indications about their contents are noted. For both EV size distributions and SERS spectra, mean, min/max, and standard deviation are represented as red, yellow, and cyan, respectively.

B. Multivariate Analysis and Classification

Samples were classified using standard machine learning algorithms in two different scenarios. In the first scenario, data from the three relevant pairwise comparisons based on strain, culture medium, or purification method were used for training and classification. This analysis was performed first to determine the effect of each individual parameter at a time. In the second scenario, the machine learning algorithms were trained to identify the EVs from all 6 subtypes simultaneously, such that the spectra of all EVs were used for training.

1) Paired Subset Classification: For the purpose of strain classification, paired subsets {Nissle-R-SEC, UPEC-R-SEC} and {K12-RF-SEC, UPEC-RF-SEC} were used as they are different strains but were grown in identical culture medium and purified identically. Similarly, the paired subsets of {UPEC-R-DG, UPEC-RF-DG} and {UPEC-R-SEC, UPEC-RF-SEC} were used for the investigation of culture medium as they are the same strain and purified identically, but grown in different culture media. Lastly, {UPEC-R-SEC, UPEC-R-DG} and {UPEC-RF-SEC, UPEC-RF-DG} were used for the evaluation of purification methods as they are the same strain and grown in identical culture medium, but purified using different methods.

To effectively present the variance of the obtained SERS spectra we investigated all the components of each spectra. Each of the obtained spectra consists of 1512 data points between Raman shifts of 800-1800 cm^{-1}. These are, in fact, an array of 1512 dimensions and the vast number of dimensions severely limits the visualisation of the variance within the data. Principle Component Analysis (PCA) was used to transform the obtained spectra in a way which reduces the spectral dimensions while preserving the maximum variance between the data after transformation. This was done firstly for the sake of the visualisation and secondly as means of classification in lower dimensions between samples.

To demonstrate the possibility of the automatic classification, we used four different types of established machine learning algorithms, including Linear Discriminant Analysis (LDA), Gaussian Process Classifier (GPC), K-Nearest Neighbour (KNN), and Support Vector Classifier (SVC) with RBF kernel. All the machine learning algorithms were employed over the PCA transformed data to calculate the probability of each point in the PCA plane to be classified as the correct type within each subgroup. The average SERS spectra for K12-RF-SEC and UPEC-RF-SEC, PCA transformation, and probability distribution obtained using machine learning algorithms, as well as first and second PCA scores are depicted in Fig. 3. Similar results of classification and PCA transformation for the other pairwise strain comparison, culture medium comparisons, and purification comparisons are available in Supplementary Fig. 1-5.

For the paired comparison of K12-RF-SEC and UPEC-RF-SEC (Fig. 3), clear differences in several peaks can be visually discerned, resulting in an expected and highly efficient classification by all the machine learning algorithms tested. By comparison, relatively subtle differences can be seen in the averaged SERS spectra for Nissle-R-SEC and UPEC-R-SEC (Supplementary Fig 1), and while the classification is not as efficient, all the machine learning algorithms tested were still largely effective. For pairwise comparisons between plain and iron supplemented culture media (Supplementary Fig 2-3), clear differences in several peaks can be easily discerned, resulting in highly efficient classification in both cases. Lastly, differences between SEC and DG-purified UPEC EVs from both types of culture media can also be effectively classified (Supplementary Fig 4-5), although not quite to the level of efficiency as when comparing different culture media.
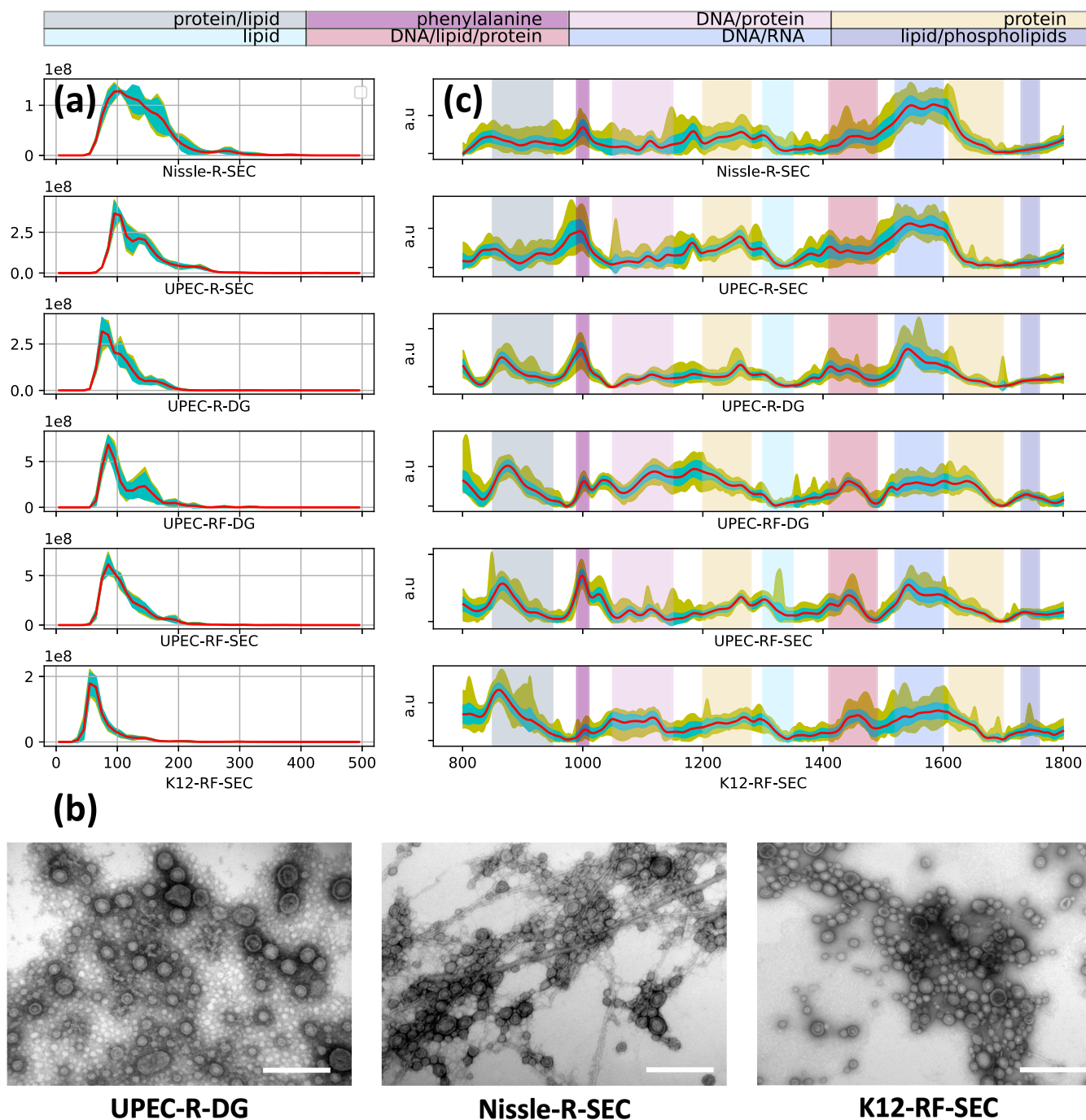
Fig. 2.   EV characterisation showing (a) size distributions of all bacterial EV samples from NTA, (b) representative TEM of UPEC-R-DG, Nissle-R-SEC, and K12-RF-SEC *E. coli* EVs (scale bars = 200 nm), and (c) acquired SERS spectra of all *E. coli* EV samples after baseline establishment and denoising with mean, min/max, and standard deviation represented as red, yellow, and cyan, respectively.

2) Classification of All Samples: Based on the initial success of classification using simple pairwise comparisons, a more challenging investigation of classifying all the samples simultaneously was explored. For the purpose of whole data classification, we first trained all the machine learning algorithms over each of the normalised spectra used to produce Fig. 2 (c) without PCA transformation. The confusion matrices produced when 60% of the data was used for training and the other 40%

was used as a testing set are shown in Fig. 4 for the all chosen machine learning algorithms used previously, in addition Artificial Neural Network (ANN) and Random Forest (RF). Of the six classification methods tested, five produced efficiencies above 90%, with GPC, SVC, and ANN all resulting in greater than 95%. LDA performed the worst, but still produced a classification efficiency of 86.8%. Collectively, these results demonstrate the power of combining SERS with machine
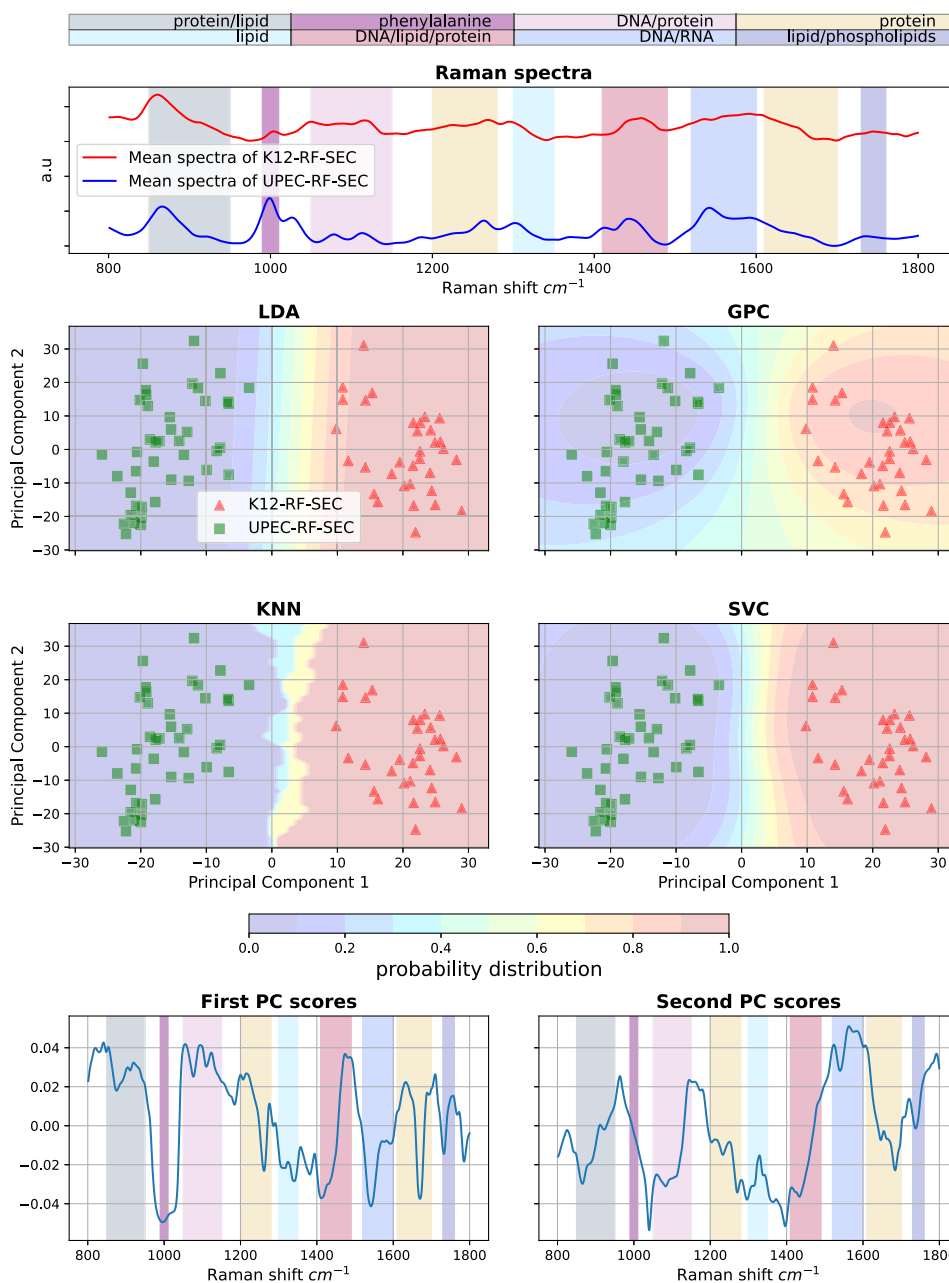
Fig. 3. Machine learning results for classification by strain between K12-RF-SEC and UPEC-RF-SEC showing (top) averaged EV SERS spectra, (middle) Linear Discriminant Analysis (LDA), Gaussian Process Classifier (GPC), K-nearest neighbour (KNN), and Support Vector Classifier (SVC), and (bottom) first and second PC scores.

learning, as even subtle differences between normalised bacterial EV fingerprints can enable highly efficient classification, even from the same species of bacteria.

We also performed the same algorithms on the PCA-transformed data. Cumulative explained variance and the first, second, and 100th PCA scores are depicted in Fig. 5 (a), (b), (c) and (d), respectively. The accuracy of SVC and ANN machine learning algorithms when different PCA dimensions were considered for training are shown in Fig. 5 (e) and (f), demonstrating that the acceptable accuracy is achieved when few (5-10) PCs are used. Interestingly, increasing the number of components beyond a certain point does not lead to better classification result. One of the main reasons for this phenomena may be the natural variance in

the *E. coli* EV biochemical contents within each subtype and the noise level due to weak Raman intensity of EVs. In fact, the 100th PC depicted in Fig. 5 (d) contains little biochemical information about the EVs and thus adding more PCs for classification purposes can actually lead to weaker classification results.

3) Effects of Culture Medium on SERS Spectra: To further investigate the effect of culture medium on the SERS spectra, which is known to significantly affect the proteome of *E. coli* EVs according to a previous study [32], two dimensional PCA transformation of the whole data set is plotted in Fig. 6 (a) with the plain (R) and iron(III) chloride-supplemented (RF) medium marked by {red, green, blue} and {purple, orange, brown}, respectively. Clearly, the first principal component
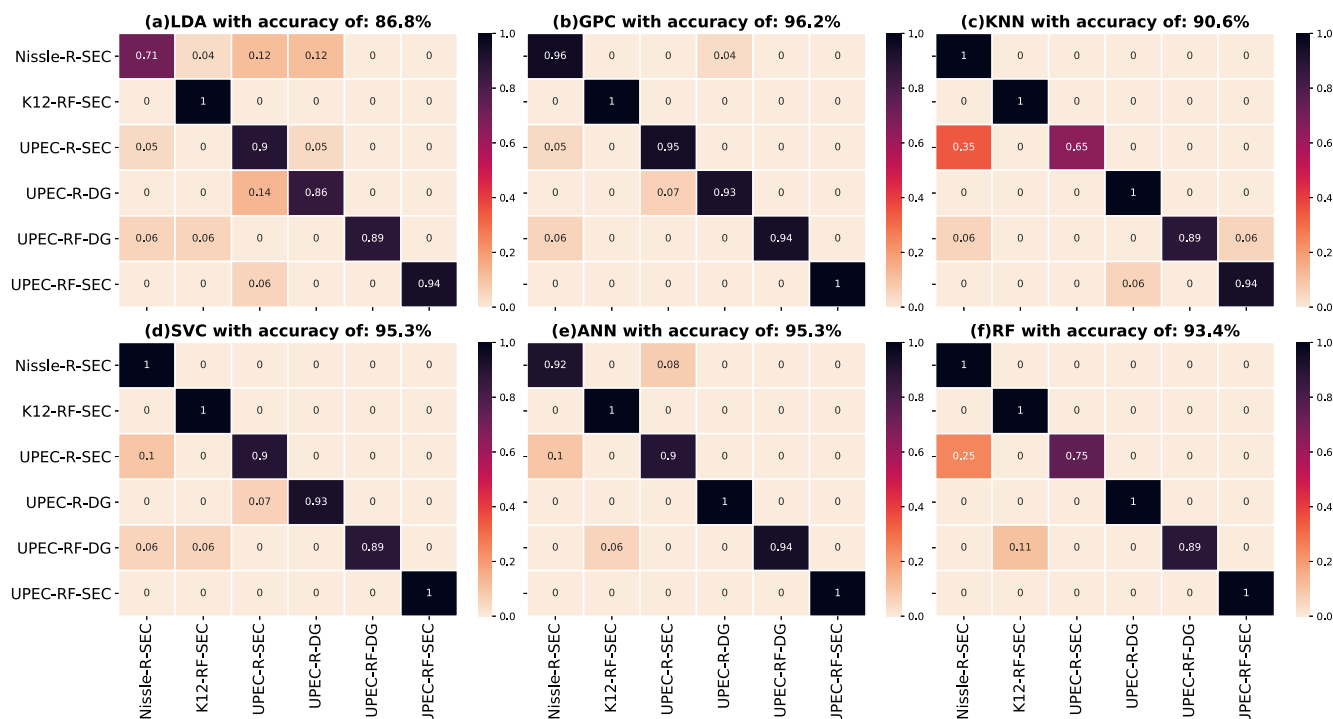
Fig. 4. Confusion matrices showing classification efficiency for all *E. coli* EVs using different machine learning methods, including (a) LDA, (b) GPC, (c) KNN, (d) SVC, (e) ANN and (f) RF.
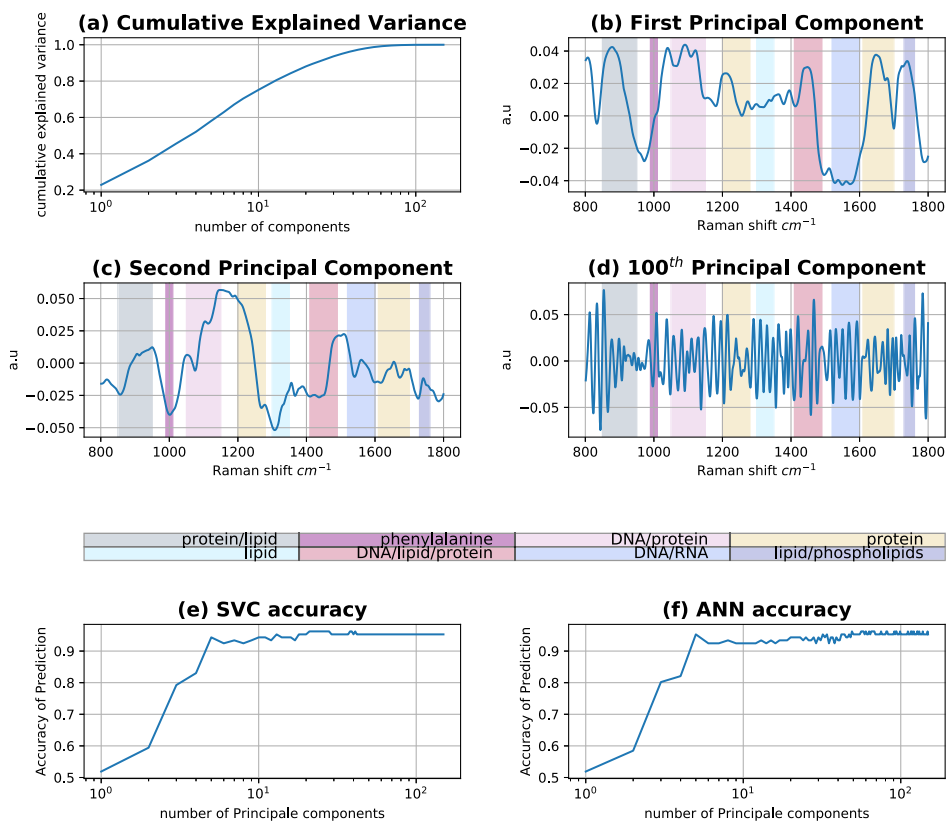
Fig. 5. All sample classification following PCA transformation showing (a) cumulative explained variance vs the number of principal components, scores of (b) first, (c) second, and (d) 100th principal components, (e) SVC and (f) ANN accuracy when different number of principal components are considered.

can effectively separate the data of *E. coli* EVs produced in the different media. This is likely due to the fact that the

SERS spectra of EVs is mostly determined by the surface biochemical makeup of the EVs, which appears to be heavily

Fig. 6. Result of the manifold machine learning approaches comparing (a) conventional PCA to (b) t-SNE and (c) UMAP for all *E. coli* EV SERS spectra.

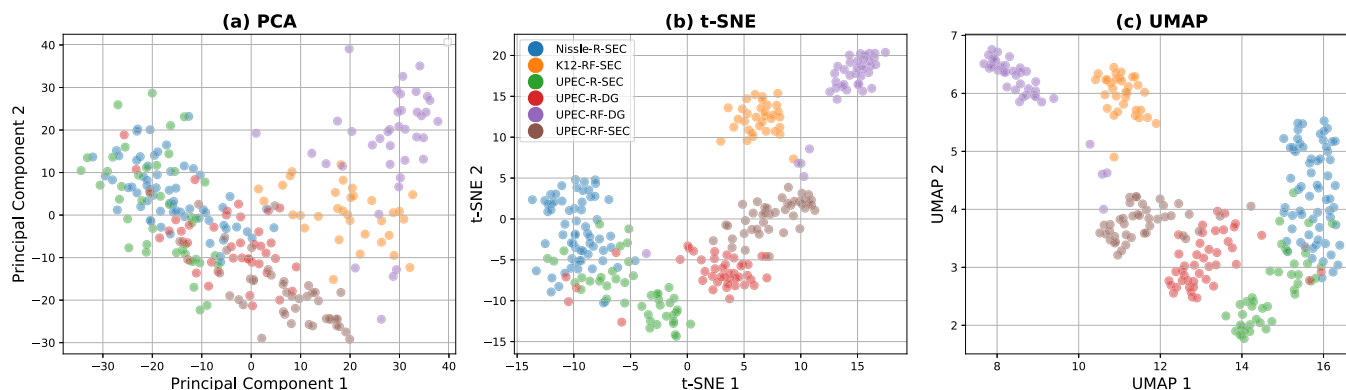influenced by the culture medium. In this study, the spectra below a Raman shift of 800 cm^{-1}, which would contain peaks related to iron, were not acquired to focus on the EV biochemical contents rather than any free iron bound to the EV surfaces, while also increasing the speed of spectral acquisitions. This strong influence of culture media on the SERS spectra not only corroborates the previous study [32], but indicates that the type of fluid in which bacterial cells are cultured may need to be carefully considered in future EV SERS comparisons. This could be particularly important in clinical applications, as different bodily fluids may confound specific species or strain identification.

C. Manifold Machine Learning for Improved Data Visualisation

As shown in Fig. 6 (a), 2 dimensional PCA transformation can not effectively illustrate the variance of the data as the data points of each subgroup are mixed with no clear spacing between them. From the results of the classification algorithms using PCA-transformed data in Fig. 5 (e) and (f), we can see that the 5 dimensional PCA transformation, however, results in a very good classification accuracy. In other words, in 5 dimensional PCA space, the data corresponding to each sub-type of *E. coli* EVs are effectively separated with clear boundaries between them. Given the challenges conceptualizing anything beyond 3 dimensions, a better dimensional reduction technique could ideally reduce the dimensional complexity. To achieve this, we employed t-SNE [63] and UMAP [64] as unsupervised manifold machine learning and non-linear dimensional reduction techniques, with the results presented in Fig. 6 (b) and (c), respectively. In clear contrast to the standard PCA transformation (Fig. 6 (a)), the variance and differences between the spectra of each *E. coli* EV subtype are much more clear in two dimensional UMAP and t-SNE planes. These results encourage the future use of manifold machine learning and non-linear dimensional reduction techniques for EV SERS studies.

IV. CONCLUSION

In this study we demonstrate, for the first time, that EVs from bacteria can be effectively characterised using SERS and

that the acquired spectra have the potential to be used for classification purposes. Several parameters appear to influence the SERS spectra of *E. coli* EVs including strain, purification method, and culture medium, which can be classified using standard machine learning approaches. In addition, manifold machine learning algorithms t-SNE and UMAP, which to the best of our knowledge have not yet been applied to EV SERS analysis, appear to be very effective for dimension reduction and visualisation of all the obtained SERS spectra. Collectively, these findings establish the incredible sensitivity and potential utility of SERS for bacterial EV analysis, as classification-enabling differences were seen in each sample subtype, despite the fact that all of the samples tested were from the same species of bacteria.

Future SERS experiments will include EVs from other species of gram-negative bacteria, as well as gram-positive bacteria and mycobacteria. As each have different cell and EV membrane compositions and proposed EV biogenesis mechanisms, the different EV envelope and contents will likely affect the SERS spectra, enabling rapid and label-free classification approaches. In addition, the isolation of bacteria-specific EVs from more complex samples like host bodily fluids could be performed by targeting appropriate bacterial EV surface antigens. As the field of bacterial EVs continues to grow, we anticipate that these findings will enable and encourage other researchers to examine bacterial EVs using this highly sensitive technique for a nearly limitless range of laboratory or clinical applications.

Colin L. Hisey is with the Hub for Extracellular Vesicle Investigations, Department of Obstetrics and Gynaecology, The University of Auckland, Auckland 1010, New Zealand (e-mail: colin.hisey@auckland.ac.nz).

Priscila Dauros-Singorenko is with the Department of Molecular Medicine and Pathology, School of Biological Sciences, The University of Auckland, Auckland 1010, New Zealand (e-mail: p.dauros@auckland.ac.nz).

Simon Swift is with the Department of Molecular Medicine and Pathology, The University of Auckland, Auckland 1010, New Zealand (e-mail: s.swift@auckland.ac.nz).

Kamran Zargar-Shoshtari is with the Department of Surgery, The University of Auckland, Auckland 1010, New Zealand (e-mail: kamran.zargar@auckland.ac.nz).

Weiliang Xu is with the Department of Mechanical Engineering, The University of Auckland, Auckland 1010, New Zealand (e-mail: p.xu@auckland.ac.nz).

Neil G. R. Broderick is with the Dodd-Walls Centre for Photonic and Quantum Technologies, Department of Physics, The University of Auckland, Auckland 1010, New Zealand (e-mail: n.broderick@auckland.ac.nz).

REFERENCES

[1] A. T. Tu, *Raman Spectroscopy in Biology: Principles and Applications*. Hoboken, NJ, USA: Wiley, 1982.

[2] M. Jermyn *et al.*, "Intraoperative brain cancer detection with Raman spectroscopy in humans," *Sci. Transl. Med.*, vol. 7, no. 274, 2015, Art. no. 274ra19.

[3] H. Chen *et al.*, "Identification and characterization of bladder cancer by low-resolution fiber-optic Raman spectroscopy," *J. Biophoton.*, vol. 11, no. 9, Sep. 2018, Art. no. e201800016.

[4] D. Klein, R. Breuch, S. von der Mark, C. Wickleder, and P. Kaul, "Detection of spoilage associated bacteria using Raman-microspectroscopy combined with multivariate statistical analysis," *Talanta*, vol. 196, pp. 325–328, May 2019.

[5] M. Saleem *et al.*, "Optical diagnosis of hepatitis B virus infection in blood plasma using Raman spectroscopy and chemometric techniques," *J. Raman Spectrosc.*, vol. 51, no. 7, pp. 1067–1077, Jul. 2020.

[6] C.-S. Ho *et al.*, "Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning," *Nature Commun.*, vol. 10, no. 1, pp. 1–8, 2019.

[7] J. van de Vossenberg *et al.*, "Identification of bacteria in drinking water with Raman spectroscopy," *Anal. Methods*, vol. 5, no. 11, pp. 2679–2687, 2013.

[8] D. Kusić, B. Kampe, P. Rösch, and J. Popp, "Identification of water pathogens by Raman microspectroscopy," *Water Res.*, vol. 48, pp. 179–189, Jan. 2014.

[9] K. Hamasha, Q. I. Mohaidat, R. A. Putnam, R. C. Woodman, S. Palchaudhuri, and J. S. Rehse, "Sensitive and specific discrimination of pathogenic and nonpathogenic *Escherichia coli* using Raman spectroscopy—A comparison of two multivariate analysis techniques," *Biomed. Opt. Exp.*, vol. 4, no. 4, pp. 481–489, 2013.

[10] K. Bradley *et al.*, "Epidemiology of a large restaurant-associated outbreak of Shiga toxin-producing *Escherichia coli* O111: NM," *Epidemiol. Infection*, vol. 140, no. 9, pp. 1644–1654, Sep. 2012.

[11] C. N. Berger *et al.*, "Fresh fruit and vegetables as vehicles for the transmission of human pathogens," *Environ. Microbiol.*, vol. 12, no. 9, pp. 2385–2397, Sep. 2010.

[12] P. L. Stiles, J. A. Dieringer, N. C. Shah, and R. P. Van Duyne, "Surface-enhanced Raman spectroscopy," *Annu. Rev. Anal. Chem.*, vol. 1, no. 1, pp. 601–626, Jul. 2008.

[13] C. Xie *et al.*, "Identification of single bacterial cells in aqueous solution using confocal laser tweezers Raman spectroscopy," *Anal. Chem.*, vol. 77, no. 14, pp. 4390–4397, Jul. 2005.

[14] F. J. García-Vidal and J. Pendry, "Collective theory for surface enhanced Raman scattering," *Phys. Rev. Lett.*, vol. 77, no. 6, p. 1163, 1996.

[15] J. Pendry, A. Aubry, D. Smith, and S. Maier, "Transformation optics and subwavelength control of light," *Science*, vol. 337, pp. 549–552, Aug. 2012.

[16] E. Temur, İ. Boyacı, U. Tamer, H. Unsal, and N. Aydogan, "A highly sensitive detection platform based on surface-enhanced Raman scattering for *Escherichia coli* enumeration," *Anal. Bioanal. Chem.*, vol. 397, no. 4, pp. 1595–1604, Jun. 2010.

[17] D. Yang, H. Zhou, C. Haisch, R. Niessner, and Y. Ying, "Reproducible *E. coli* detection based on label-free SERS and mapping," *Talanta*, vol. 146, pp. 457–463, Jan. 2016.

[18] S. Zong *et al.*, "Facile detection of tumor-derived exosomes using magnetic nanobeads and SERS nanoprobes," *Anal. Methods*, vol. 8, no. 25, pp. 5001–5008, 2016.

[19] C. Lee, R. Carney, K. Lam, and J. W. Chan, "SERS analysis of selectively captured exosomes using an integrin-specific peptide ligand," *J. Raman Spectrosc.*, vol. 48, no. 12, pp. 1771–1776, Dec. 2017.

[20] Y.-F. Tian, C.-F. Ning, F. He, B.-C. Yin, and B.-C. Ye, "Highly sensitive detection of exosomes by SERS using gold nanostar@Raman reporter@nanoshell structures modified with a bivalent cholesterol-labeled DNA anchor," *Analyst*, vol. 143, no. 20, pp. 4915–4922, 2018.

[21] Z. Wang *et al.*, "Screening and multiple detection of cancer exosomes using an SERS-based method," *Nanoscale*, vol. 10, no. 19, pp. 9053–9062, 2018.

[22] E. A. Kwizera *et al.*, "Molecular detection and analysis of exosomes using surface-enhanced Raman scattering gold nanorods and a miniaturized device," *Theranostics*, vol. 8, no. 10, pp. 2722–2738, 2018.

[23] T.-D. Li *et al.*, "An ultrasensitive polydopamine bi-functionalized SERS immunoassay for exosome-based diagnosis and classification of pancreatic cancer," *Chem. Sci.*, vol. 9, no. 24, pp. 5372–5382, 2018.

[24] C. Lee *et al.*, "3D plasmonic nanobowl platform for the study of exosomes in solution," *Nanoscale*, vol. 7, no. 20, pp. 9290–9297, 2015.

[25] K. Sivashanmugam *et al.*, "Bimetallic nanoplasmonic gap-mode SERS substrate for lung normal and cancer-derived exosomes detection," *J. Taiwan Inst. Chem. Eng.*, vol. 80, pp. 149–155, Nov. 2017.

[26] L. Tirinato *et al.*, "SERS analysis on exosomes using super-hydrophobic surfaces," *Microelectron. Eng.*, vol. 97, pp. 337–340, Sep. 2012.

[27] J. Park *et al.*, "Exosome classification by pattern analysis of surface-enhanced Raman spectroscopy data for lung cancer diagnosis," *Anal. Chem.*, vol. 89, no. 12, pp. 6695–6701, Jun. 2017.

[28] H. Shin, H. Jeong, J. Park, S. Hong, and Y. Choi, "Correlation between cancerous exosomes and protein markers based on surface-enhanced Raman spectroscopy (SERS) and principal component analysis (PCA)," *ACS Sens.*, vol. 3, no. 12, pp. 2637–2643, Dec. 2018.

[29] Z. Yan *et al.*, "A label-free platform for identification of exosomes from different sources," *ACS Sens.*, vol. 4, no. 2, pp. 488–497, Feb. 2019.

[30] T. Nagakubo, N. Nomura, and M. Toyofuku, "Cracking open bacterial membrane vesicles," *Frontiers Microbiol.*, vol. 10, p. 3026, Jan. 2020.

[31] Y. Chen, L. Liu, H. Fu, C. Wei, and Q. Jin, "Comparative proteomic analysis of outer membrane vesicles from *Shigella flexneri* under different culture conditions," *Biochem. Biophys. Res. Commun.*, vol. 453, no. 4, pp. 696–702, Oct. 2014.

[32] J. Hong *et al.*, "Analysis of the *Escherichia coli* extracellular vesicle proteome identifies markers of purity and culture conditions," *J. Extracellular Vesicles*, vol. 8, no. 1, Dec. 2019, Art. no. 1632099.

[33] C.-H. Chiu *et al.*, "Bacterial membrane vesicles from *Acinetobacter baumannii* induced by ceftazidime are more virulent than those induced by imipenem," *Virulence*, vol. 11, no. 1, pp. 145–158, Dec. 2020.

[34] A. Malabirade *et al.*, "The RNA complement of outer membrane vesicles from *Salmonella enterica* serovar Typhimurium under distinct culture conditions," *Frontiers Microbiol.*, vol. 9, p. 2015, Aug. 2018.

[35] C. Pérez-Cruz, F. Briansó, E. Sonnleitner, U. Bläsi, and E. Mercadé, "RNA release via membrane vesicles in *Pseudomonas aeruginosa* PAO1 is associated with the growth phase," *Environ. Microbiol.*, vol. 23, no. 9, pp. 5030–5041, Sep. 2021.

[36] B. S. R. D. Luz *et al.*, "Environmental plasticity of the RNA content of *Staphylococcus aureus* extracellular vesicles," *Frontiers Microbiol.*, vol. 12, p. 482, Mar. 2021.

[37] T. Rojalin, B. Phong, H. J. Koster, and R. P. Carney, "Nanoplasmonic approaches for sensitive detection and molecular characterization of extracellular vesicles," *Frontiers Chem.*, vol. 7, p. 279, May 2019.

[38] H. Shin, D. Seo, and Y. Choi, "Extracellular vesicle identification using label-free surface-enhanced Raman spectroscopy: Detection and signal analysis strategies," *Molecules*, vol. 25, no. 21, p. 5209, Nov. 2020.

[39] J. R. White, P. Dauros-Singorenko, J. Hong, F. Vanholsbeeck, A. Phillips, and S. Swift, "The complex, bidirectional role of extracellular vesicles in infection," *Biochem. Soc. Trans.*, vol. 49, no. 2, pp. 881–891, Apr. 2021.

[40] V. M. Chernov *et al.*, "Extracellular membrane vesicles secreted by mycoplasma *Acholeplasma laidlawii* PG8 are enriched in virulence proteins," *J. Proteomics*, vol. 110, pp. 117–128, Oct. 2014.

[41] Y. Liu, K. A. Y. Defourny, E. J. Smid, and T. Abee, "Gram-positive bacterial extracellular vesicles and their impact on health and disease," *Frontiers Microbiol.*, vol. 9, p. 1502, Jul. 2018.

[42] P. Dauros-Singorenko, J. Hong, S. Swift, A. Phillips, and C. Blenkiron, "Effect of the extracellular vesicle RNA cargo from uropathogenic *Escherichia coli* on bladder cells," *Frontiers Mol. Biosci.*, vol. 7, p. 269, Sep. 2020.

[43] P. Briaud and R. K. Carroll, "Extracellular vesicle biogenesis and functions in gram-positive bacteria," *Infection Immunity*, vol. 88, no. 12, 2020, Art. no. e00433.

[44] C.-S. Kang *et al.*, "Extracellular vesicles derived from gut microbiota, especially *Akkermansia muciniphila*, protect the progression of dextran sulfate sodium-induced colitis," *PLoS ONE*, vol. 8, no. 10, pp. 1–11, 2013.

[45] J. C. Caruana and S. A. Walper, "Bacterial membrane vesicles as mediators of microbe–microbe and microbe–host community interactions," *Frontiers Microbiol.*, vol. 11, p. 432, Mar. 2020.

[46] C. Coelho and A. Casadevall, "Answers to naysayers regarding microbial extracellular vesicles," *Biochem. Soc. Trans.*, vol. 47, no. 4, pp. 1005–1012, Aug. 2019.

[47] L. Brown, J. M. Wolf, R. Prados-Rosales, and A. Casadevall, "Through the wall: Extracellular vesicles in gram-positive bacteria, mycobacteria and fungi," *Nature Rev. Microbiol.*, vol. 13, no. 10, pp. 620–630, Oct. 2015.

[48] C. Schwechheimer and M. J. Kuehn, "Outer-membrane vesicles from gram-negative bacteria: Biogenesis and functions," *Nature Rev. Microbiol.*, vol. 13, no. 10, pp. 605–619, Oct. 2015.

[49] S. Erdmann, B. Tschitschko, L. Zhong, M. J. Raftery, and R. Cavicchioli, "A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells," *Nature Microbiol.*, vol. 2, no. 10, pp. 1446–1455, Oct. 2017.

[50] C. Coelho *et al.*, "Listeria monocytogenes virulence factors, including listeriolysin O, are secreted in biologically active extracellular vesicles," *J. Biol. Chem.*, vol. 294, no. 4, pp. 1202–1217, Jan. 2019.

[51] J. Lee, O. Y. Kim, and Y. S. Gho, "Proteomic profiling of Gram-negative bacterial outer membrane vesicles: Current perspectives," *Proteomics-Clin. Appl.*, vol. 10, nos. 9–10, pp. 897–909, Oct. 2016.

[52] M. Potter, C. Hanson, A. J. Anderson, E. Vargis, and D. W. Britt, "Abiotic stressors impact outer membrane vesicle composition in a beneficial rhizobacterium: Raman spectroscopy characterization," *Sci. Rep.*, vol. 10, no. 1, p. 21289, Dec. 2020.

[53] S. Knapp, J. Hacker, T. Jarchau, and W. Goebel, "Large, unstable inserts in the chromosome affect virulence properties of uropathogenic *Escherichia coli* O6 strain 536," *J. Bacteriol.*, vol. 168, no. 1, pp. 22–30, Oct. 1986.

[54] A. Nissle, "Die antagonistische behandlung chronischer darmstörungen mit colibakterien," *Med. Klin*, vol. 2, pp. 29–33, 1918.

[55] G. Sezonov, D. Joseleau-Petit, and R. D'Ari, "*Escherichia coli* physiology in Luria–Bertani broth," *J. Bacteriol.*, vol. 189, no. 23, pp. 8746–8749, Dec. 2007. [Online]. Available: https://jb.asm.org/content/189/23/8746

[56] P. Mao *et al.*, "Broadband single molecule SERS detection designed by warped optical spaces," *Nature Commun.*, vol. 9, no. 1, pp. 1–8, Dec. 2018.

[57] M. Kazemzadeh, N. Broderick, W. Xu, and K. Z. Shoshtari, "Design and optimization of broadband and transparent SERS based on transformation optics," *Proc. SPIE*, vol. 11201, Dec. 2019, Art. no. 112011L.

[58] M. Kazemzadeh *et al.*, "Space curvature-inspired nanoplasmonic sensor for breast cancer extracellular vesicle fingerprinting and machine learning classification," *Biomed. Opt. Exp.*, vol. 12, no. 7, pp. 3965–3981, 2021.

[59] M. Kahraman, P. Daggumati, O. Kurtulus, E. Seker, and S. Wachsmann-Hogiu, "Fabrication and characterization of flexible and tunable plasmonic nanostructures," *Sci. Rep.*, vol. 3, no. 1, pp. 1–9, Dec. 2013.

[60] P. H. Eilers and H. F. Boelens, "Baseline correction with asymmetric least squares smoothing," *Leiden Univ. Med. Centre Rep.*, vol. 1, no. 1, p. 5, 2005.

[61] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.

[62] Z. Movasaghi, S. Rehman, and I. U. Rehman, "Raman spectroscopy of biological tissues," *Appl. Spectrosc. Rev.*, vol. 42, no. 5, pp. 493–541, Sep. 2007.

[63] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.

[64] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform manifold approximation and projection for dimension reduction," 2018, *arXiv:1802.03426*.