

Specific Intracellular Signature of SARS-CoV-2 Infection Using Confocal Raman Microscopy

Hamideh Salehi (✉ s_hamideh@yahoo.com)

university of Montpellier

Anuradha Ramoji

Friedrich-Schiller university <https://orcid.org/0000-0003-2723-6614>

Said Mougari

Lyon university

Peggy Merida

university of Montpellier

Aymeric Neyret

university of Montpellier

Juergen Popp

Leibniz Institute of Photonic Technology

Delphine Muriaux

university of Montpellier

Branka Horvat

University of Claude Bernard Lyon

frederic cuisinier

university of Montpellier

Article

Keywords:

Posted Date: March 25th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1439575/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Communications Chemistry on July 25th, 2022. See the published version at <https://doi.org/10.1038/s42004-022-00702-7>.

Abstract

SARS-CoV-2 infection remains to spread worldwide and requires a better understanding of virus-host interactions. Here, we analyzed biochemical modifications due to SARS-CoV-2 infection in cells by confocal Raman microscopy. Obtained results were compared with the infection with another RNA virus, the measles virus. Our results have demonstrated a virus-specific Raman hallmark of molecular signature, reflecting intracellular modification during each infection. Advanced data analysis has been used to distinguish non-infected versus infected cells for two RNA viruses. Further, classification between non-infected and SARS-CoV-2 and measles virus-infected cells yielded an accuracy of 98.9 and 97.2 respectively, with a significant increase of the essential amino-acid tryptophan in SARS-CoV-2-infected cells. These results present proof of concept for the application of Raman spectroscopy to study virus-host interaction and to identify factors that contribute to the efficient SARS-CoV-2 infection and may thus provide novel insights on viral pathogenesis, targets of therapeutic intervention and development of new COVID-19 biomarkers.

Introduction

Coronavirus disease 2019 (COVID-19) is an ongoing pandemic infection caused by the positive-sense RNA virus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)¹, provoking untold disruption throughout the world. Symptoms of COVID-19 such as dyspnea, fever, cough, fatigue could be followed by numerous complications including pneumonia, myocarditis, kidney injury, resulting in some cases with death². Understanding the mechanisms of virus-induced cell modifications is critical for the development of rapid diagnostics of infection and efficient antiviral treatment.

Sensitive viral identification could be obtained using currently available methods, including immunological and molecular tests, ELISA (enzyme-linked immunosorbent assay) and PCR (polymerase chain reaction)³⁻⁶, however, these analyses require previous information on the infectious agent identity⁷. Other techniques such as Next-Generation Sequencing are very sensitive, although low virus quantity may be challenging^{8,9}. Globally, sixty percent of medical care is engaged by viral infections¹⁰, and the requirement of accurate biomarkers of viral infection presents a big challenge. Due to the significant heterogeneity of COVID-19 disease profiles, biomarkers that allow either the identification of patients at high risk for developing severe forms of COVID-19 and its long-term complications, or guide personalized treatment options, are scarce. Therefore, new approaches leading to the identification of COVID-19-related biomarkers are urgently required for the development of precision medicine-based therapeutic strategies.

Cells could respond to different steps of viral infection, including virus entry, trafficking through the cell organelles for replication and egress of viral particles or their entry into lysosomes for degradation of viral proteins^{11,12}. As Raman microscopy allows the analysis of on the single-cell level¹³⁻¹⁵, the biochemical modifications due to the cell reaction to the virus could be detected in the early stage. The conventional diagnostic methods are applicable after symptoms appearance which might be too late to avoid serious consequences. As Raman identifies the chemical modification, by investigation of spectral changes induced by a virus in any environment, there is no requirement for any genetic or proteomic information of the virus,^{8, 16-18} presenting an additional advantage for the Raman spectroscopy in viral detection.

Biomedical applications of Raman spectroscopy have been reported previously due to its noninvasive examination of a single live cell without labeling^{15, 19-21}. Raman spectroscopy imaging is known as a well-established method for chemical analysis of intra/extracellular environment^{22,23}. Each confocal Raman voxel contains the biochemical fingerprint signature with a dimension of 300nm*300nm*900nm^{20,21}. Data analysis on this information reveals precise cell compartments identifications as well as slight modifications. Usual data analysis on Raman spectra is categorized into two groups of supervised and unsupervised approaches¹⁵. The unsupervised methods only depend on Raman spectra to classify into different categories. By reducing the number of variables, Principal Component Analysis (PCA) is considered as an unsupervised method. As a supervised method to differentiate between various cell organelles (including the nucleus, nucleolus, mitochondria, Golgi and cytoplasm), k-means cluster analysis (KMCA) could be realized on more than twenty thousand spectra of each Raman

scan^{24,25}. Organelle's separation enables distinctive extraction of single spectra for further hyperspectral data analysis such as PCA or Support vector machine (SVM). Viral infection with different incubations time has been monitored in previous studies^{26,27}. Recent studies reported the detection of SARS-CoV-2 using Surface Enhanced Raman Scattering (SERS)^{2,28-33}. SERS substrate design dictates sensitivity and accuracy of the method and presents an essential key towards a successful result, but is also associated to a rather high cost of analysis^{30-32,34}.

In the current study, spontaneous Raman spectroscopy has been used to avoid complex substrate design. Biomolecular modification of cellular organelles after SARS-CoV-2 infections has been compared to another RNA virus, the measles virus (MeV) to analyze virus-specific intracellular modifications. The list of obtained biochemical changes is provided to help in better understanding of the cellular signature of virus infection. The obtained results show tryptophan traces in SARS-CoV-2-infected cells which were observed neither in non-infected cells nor in MeV-infected cells. These results suggest a novel easily detectable viral signature in an early stage of infection and potentially new biomarker of COVID-19. Understanding the intracellular biochemical modification following the virus infection of the host cells should provide new information about SARS-CoV-2-induced intracellular modifications and may reveal novel specific markers for personalized treatment of COVID-19 patients.

Results

Raman imaging of virus-infected Vero E6 cells

SARS-CoV-2-induced modifications in different sub-cellular compartments were followed in Vero E6 cells, a widely used cell line for the amplification of different viruses. Cells were initially infected with SARS-CoV-2 or MeV for 24h and analyzed by Raman spectrometry. This time point was determined in the initial analysis as optimal to observe the effect of viral infection on cells, at MOI = 0.1. In parallel, the sample of SARS-CoV-2 infected cells was analyzed by electron microscopy, to follow the virus infection and the visible intracellular modifications due to the virus, 24h post infection (pi) (Figure S1). Transmission electron microscopy (TEM) images showed viral particles at the cell plasma membrane signing viral production (Figure S1A) and/or further virus attachment for cell entry. In addition, intracellular organelles filled with viral factories were recognizable inside the cytoplasm, near the nucleus and mitochondria, most probably reflecting the double-membrane vesicles or spherules replicative factories of SARS-CoV-2 derived from the ER-Golgi apparatus (Figure S1B) as previously observed and described^{35,36}. These TEM images confirmed that virus particles were produced by the cells and that they can be visualized at the cell plasma membrane and internalized 24h post infection.

We then proceeded with Raman microscopy to follow virus-induced intracellular biochemical modifications in traceable in sub-cellular organelles using a label free method. For Raman imaging of virus infected cells and measurements the central wavelength of gating was set at 610 nm (which presents the common setting for Raman cell imaging), showing the "fingerprint region" from 400 to 1800 cm^{-1} , the C-H peak at 2800–3100 cm^{-1} . Each Raman scan contains more than twenty thousand individual spectra. A Raman reconstructed image is based on the selected C-H bands to represent the intracellular proteins and lipids. During the scan, with the nanometric scale movement of the piezoelectric table, each pixel is registered with its relevant spectrum. Molecular structures information is gathered in Raman spectra with individual bands associated with the biomolecules under the laser spot. The average spectral shape of K-means cluster analysis (KMCA) clusters was obtained to represent the cellular organelles, such as the nucleus, cytoplasm, mitochondria, and Golgi. To assure the virus specificity of observed effects the results obtained with SARS-CoV-2 were compared with those found after the infection of Vero E6 cells with another RNA virus, the measles virus (MeV).

Figure 1 illustrates scans of SARS-CoV-2- and MeV- infected for 24h and non-infected Vero E6 cells followed with the relevant KMCA analysis. Cellular organelles separation is followed after KMCA analysis to extract the single spectrum of each cluster. Figure 1A-C-E illustrates reconstructed C-H Raman images of infected versus non-infected cells, with light yellow as highest intensities and dark hues for lowest intensities of C-H peak, corresponding to the intracellular content of lipids and proteins. Part 1B-D-F belongs to the relevant KMCA analysis of 1A-C-E. This analysis allows to mark the cellular organelles to separate

single spectra of each organelle (nucleus, nucleolus, Golgi-mitochondria together, lipid droplets) as well as the cytoplasm, for further principal component analysis (PCA) and support vector machine (SVM) analysis. For a better comparison, the same color code has been used for all KMCA parts as indicated in the legend.

Chemical profile of SARS-CoV-2 and MeV-infected Vero E6 cells

The mean of the Raman spectra extracted from different intracellular components, namely, cytoplasm, Golgi-mitochondria bodies, and nucleus region of non-infected and SARS-CoV-2/MeV infected Vero E6 cells has been displayed in Fig. 2A-2C, respectively. For visualizing the difference between intracellular components as a result of the viral infection, difference Raman spectra were generated by subtracting the mean Raman spectra of the non-infected cells from the mean Raman spectra of the cells infected with virus (the Raman spectra from the cells with MeV infection and cells with SARS-CoV-2 infection were used together in the infected group), for individual intracellular components, as shown in Fig. 2D-2F. Further, for a deeper insight into the biochemical difference between MeV and SARS-CoV-2 infected cells, difference Raman spectra were generated for different intracellular components by subtracting Raman spectra of the SARS-CoV-2 infected cells from the MeV infected cells (Fig. 2G-2H).

In the difference spectra shown in Fig. 2D-2F between non-infected (Control) and the virus infection (MeV and SARS-CoV-2 infection considered together), the positive Raman peaks belong to the control group, whereas the negative Raman peaks belong to the virus-infected cells. The difference Raman spectra (Fig. 2D-2I) highlights the intracellular changes occurring within the cell as a result of viral infection. The differences between infected and non-infected cells in the cytoplasm, Golgi-mitochondria bodies, and nucleus can be observed mainly in the Raman peak profile of the C-H stretching region ($2800\text{--}3050\text{ cm}^{-1}$) (Fig. 2D-2F). In the difference spectra of cytoplasm and Golgi-mitochondria bodies shown in Fig. 2D and 2E, phosphate backbone vibration is visible at 788 cm^{-1} and 790 cm^{-1} respectively³⁷, possibly due to the presence of viral RNA³⁸. The Raman spectra extracted from the non-infected cells' Golgi-mitochondria bodies show Raman spectral signature of lipid at 1732 cm^{-1} (Fig. 2E). The difference in Raman spectrum between the two types of viral infections, MeV and SARS-CoV-2, indicates chemical changes that are specific to the type of the virus causing the infection (Fig. 2G-2I). The positive Raman peaks in the presented spectra belong to the SARS-CoV-2 infected cells and the negative Raman peaks belong to the MeV infected cells. Within all the three intracellular compartments, SARS-CoV-2 infected cells show the presence of symmetric CH_2 vibrations of lipids, (1732 cm^{-1} and $2847/2864\text{ cm}^{-1}$)^{40,41}, ring breathing modes of the DNA/RNA bases (1429 cm^{-1})⁴², strong C-O band of ribose serves as a marker of RNA (1125 cm^{-1})³⁹, tryptophan vibrations (566 cm^{-1} and 749 cm^{-1})⁴³⁻⁴⁶, C-C twisting mode of phenylalanine (619 cm^{-1})⁴⁵, whereas MeV-infected cells show dominating presence of proteins ($2933/2970\text{ cm}^{-1}$)⁴¹. The assignments of the Raman peaks have been described in supplement Table 1 and S1.

To visualize the possibility to differentiate between the non-infected, MeV infected, and the SARS-CoV-2 infected Vero E6 cells, principal component analysis (PCA) was performed. The PCA 3D-score plot along with the principal component (PC) loadings are shown in Fig. 3 for the three different intracellular components. The PCA analysis was performed for each experimental batch separately to evaluate the Raman spectral data. In supplement Figure S2, PCA analysis for the remaining two batches has been shown. In the PCA 3D-score plots shown for cytoplasm (Fig. 3A), Golgi-mitochondria bodies (Fig. 3C), and nucleus (Fig. 3E), the Raman spectra of non-infected Vero E6 cells are well separated from the Raman spectra of the infected cells. This indicates a distinct chemical profile between infected and non-infected cells as shown in Table 1 and S1. The Raman spectra extracted from the nucleus region of the MeV and SARS-CoV-2 infected cells (Fig. 3E) are well separated from each other, compared to the Raman spectra extracted from the cytoplasm and Golgi-mitochondria bodies of infected Vero E6 cells. This indicates the nucleus of MeV-infected Vero E6 cells has a different chemical profile compared to the nucleus of the SARS-CoV-2 infected Vero E6 cells. The PC loadings shown in Fig. 3B, 3D, and 3F show the respective contributing Raman peaks. The PC1 is mainly responsible for separating infected and non-infected Vero E6 cells. The PC2 and PC3 jointly contribute to the separation of the MeV- and SARS-CoV-2-infected Vero E6 cells. Further, in the cytoplasm of MeV infected cells presence of tyrosine (815 cm^{-1} and 1629 cm^{-1}) is specifically observed compared to the SARS-CoV-2 infected cells where tryptophan peaks (755 cm^{-1} , 1201 cm^{-1} , 1547 cm^{-1} , 1607 cm^{-1}) are prominent (Table 1).

Raman Model To Differentiate Sars-cov-2 And Mev-infected Vero E6 Cells

The unsupervised PCA was unable to resolve the infected and non-infected Vero E6 cells when Raman spectral data from all three experimental batches were pooled together for the analysis. Hence, a supervised machine learning algorithm-SVM was used for the classification of the non-infected, MeV-infected and SARS-CoV-2 infected Vero E6 cells. Two-class SVM models were built as shown in Fig. 4A-4C to visualize the chemical differences. Further, Raman difference spectra were generated as shown in Fig. 4D-4F. In Fig. 4A, the SVM model separates between non-infected and MeV-infected Vero E6 cells with a 10-fold cross-validated accuracy of 99%. Within the SVM plot, the separation of Raman spectra extracted from the different intracellular components of infected and non-infected cells can be visualized from the shape of the legend used (Square = cytoplasm, circle = golgi-mitochondria bodies, and triangle = nucleus region). The difference spectrum (Fig. 4D) generated by subtracting the mean Raman spectrum of non-infected cells (positive Raman peaks) from the mean Raman spectrum of MeV-infected cells (negative Raman peaks) indicate major chemical differences observed in the infected cells are lipids (2860 cm^{-1} , 2885 cm^{-1}) and proteins (1599 cm^{-1} , 2929 cm^{-1} , 2947 cm^{-1}), nucleic acids (793 cm^{-1} , 1255 cm^{-1}) and carbohydrates (1464 cm^{-1}). Similarly, the SVM model classifies the non-infected and SARS-CoV-2 infected cells with a 10-fold cross-validated accuracy of 97% (Fig. 4B). The changes observed in the Raman difference spectrum (Fig. 4E) are similar as described above for the MeV-infected cells. In Fig. 4C, the two-class SVM classification model shows differentiation between the MeV-infected cells and SARS-CoV-2 infected cells. All the three intracellular components of the MeV and SARS-CoV-2-infected cells were well separated. The total 10-fold cross-validated accuracy was 98%. The difference spectrum generated between MeV and SARS-CoV-2-infected cells show differences in proteins (1612 cm^{-1} in SARS-CoV-2 and 1656 cm^{-1} in MeV-infected cells), higher lipids in SARS-CoV-2 (2856 cm^{-1} , 2885 cm^{-1}), changes in the nucleic acids (1091 cm^{-1} and 1255 cm^{-1} in MeV) and tryptophan vibrations at 749 cm^{-1} in SARS-CoV-2 infected cells).

A three-class SVM model was built (supplement Figure S3) to classify non-infected and MeV infected and SARS-CoV-2 infected cells. The non-infected cells are well separated, further, the separation between the two types of viral-infected cells can also be visualized.

PCs used 13, 10 fold CV other parameters same as below)

Table 1

Significant Raman peaks observed in PC loadings contributing to the separation between SARS-CoV-2 and MeV infected Vero E6 cells

Raman peaks in cm^{-1}					
Cytoplasm		Golgi-mitochondria-		Nucleus	
MeV	SARS-CoV-2	MeV	SARS-CoV-2	MeV	SARS-CoV-2
586 (sym str PO_4^{3-})	459 (DNA/Glycogen)	1258 (Ade/Cyt)	566 (Trp/Cyt/Gua)	733 (Ade)	566 (Cyt/Gua)
815 (Tyr)	566 (Trp/Cyt/Gua/Kyn)	1470 (lipids)	619 (Cys)	1011 (Str C-O Ribose)	675 (Gua)
1099 (str C-N)	749 (Trp)	1592 (Str C = C and C = N protein)	749 (Trp)	1258 (Ade/Cyt)	749 (Tyr)
1250 (Gua/Cyt)	842 (Trp/Kyn)	1661 (α -helix protein)	1125 (Trp/Str C-C lipids)	1279 (Nucleic acids/Phosphates)	839 (Pro/hydroxyproline, Tyr/glycogen)
1315 (Gua)	1125(Str C-C lipids/glucose/polysacahhride)	2935 (CH_3 sym str proteins)	1429 (Deoxyribose)	1289 (Cyt)	931 (Gua)
1487 (Gua)	1201 (Trp)		1732 (lipids)	1345 (Gua)	1125* C-O ribose
1575 (Gua/Ade)	1429 (Fatty acids)		2847 (CH_3 sym str lipids)	1470* (lipids)	1235* PO_2 mode
1629 (Tyr)	1510 (Cyt)			1641* (proteins)	1378 (Ade/Gua)
1661 (α -helix protein)	1547 (Trp)			1686* (proteins)	1429 (Deoxyribose)
2939 (lipids)	1607 (Trp)			2935* (proteins)	1477 (Gua)
	1732 (lipids)			2970* (lipids)	1575 (Gua/Ade)
	2849 (CH_3 sym str lipids)				1732* (lipids)
	3026 (long-chain fatty acid)				2864* (lipids)
					3026 (long-chain fatty acid)

Trp : Trptophan, Cys : cysteine, Tyr : Tyrosine, Cyt : Cytosine, Gua : Guanine, Pro : Proline, Sym :Symmetric, Str : Stretching, Ade : Adenine, Kyn : Kynurenine

*The contribution of these Raman bands arise due to the extraction of nucliec acid spectra from the Raman spectroscopic image of the cells.

Discussion

Understanding the virus-induced biochemical modifications in host cells is of high interest to a better comprehension of virus-host interactions and developing novel diagnostic biomarkers and therapeutic approaches. As Raman spectroscopy can detect these intracellular molecular metabolic changes without any marker or sample preparation^{47,48} or need for previous

knowledge of virus strains, this method has the potential to be used as a precise label free detection technique with a low cost of analysis^{49,50}.

In this study, two highly contagious airborne RNA viruses were used to infect Vero E6 cells and compared to non-infected cells using the Raman spectroscopy approach. The initial observation of SARS-CoV-2 infection by TEM microscopy confirmed the abundant virus replication at 24h post infection. Raman spectroscopy imaging using 532 nm laser excitation, enables along with visualization of the virions also the intracellular molecular modifications due to the effect of SARS-CoV-2 or MeV infection of the cells. Multivariate analysis is indispensable to realize major dissimilarities within the individual spectrum, hence, to understand the chemical changes occurring in the infected cells chemometric methods, namely, KMCA, PCA, and SVM were employed. The KMCA of Raman spectral data allows visualization of intracellular organelles in the micrometer ranges enabling sensing of clustered virion particles or replicative viral factories in vesicles. For precise co-localization of viruses or viral replicative organelles within the cells, immuno-fluorescence staining method could be added on for further characterization, since cells are intact after Raman spectroscopy imaging as it is a non-destructive method.

The study aimed at the identification of virus-specific intracellular biochemical changes occurring during the infection with two different RNA viruses. The difference Raman spectrum generated between non-infected/infected cells and between SARS-CoV-2/MeV-infected cells for different organelles (Fig. 2) demonstrates the unique host cell response to of SARS-CoV-2. The contributing Raman peaks as observed in Fig. 2D to 2I and the biomolecules giving rise to these peaks (Table 1 and S1) provide hints about biochemical changes occurring during viral infection. Especially, the prominent presence of the tryptophan as a significant consequence during SARS-CoV-2 infection is observed. This observation corroborates our analysis of virus genomes, revealing that SARS-CoV-2 proteome contains approximately three times higher tryptophan levels compared to MeV proteome (159 versus 57 aa, respectively, found in each virus), which probably leads to a higher content of intracellular tryptophan during SARS-CoV-2 replication. Tryptophan metabolism presents an important biochemical pathway, implicated in the inflammation, generation of immune tolerance and in mental health⁵¹. Interestingly, in accord with our results on the intracellular tryptophan changes based on Raman spectroscopy, the modifications in the tryptophan ratio in the sera of COVID-19 patients were observed in the previous studies^{52,53} and the assessment of tryptophan was recently suggested to be the prognostic marker in COVID-19 patients⁵². Our results suggest that accumulation of this essential amino acid in SARS-CoV-2-infected cells may lead to a decrease in the serum level of tryptophan in infected patients, associated with aggravation of their clinical symptoms, underlying the potential of tryptophan to serve as a biomarker for COVID-19. Similar observations have been made in a recent Raman spectroscopy study on Epstein-Barr virus infection of glial cells where tryptophan along with other biomolecules such as cholesterol, glucose polysaccharides are associated with the viral infection⁵⁰.

Further, SARS-CoV-2 triggers different chemical reactions within the cells compared to another RNA virus MeV, which could be evidenced from the simple unsupervised PCA analysis (Fig. 3). The PCA analysis indicates that SARS-CoV-2-infected cells have different biochemical makeup compared to MeV-infected cells, demonstrating the specificity of the Raman approach. In this study only Vero E6 cell line was used for the viral infection, to obtain the first insight in the possible application of the Raman approach that has been demonstrated to be powerful in virus discrimination. However, the use of the other cell types known as hosts for the COVID-19 infection in patients should be performed in the future and may reveal the additional fingerprints of the SARS-CoV-2 biochemical modifications in specific cell types.

For robust analysis, the experiments were repeated a minimum of three times, and the Raman spectral data were pooled together. A higher-end chemometric method was required as PCA was unable to handle the heterogeneity within the data set normally present due to the biological and experimental repetitions. However, to avoid overfitting, the variance captured by the first 13 PCs was used as input for the SVM analysis. The differentiation of non-infected and infected Vero E6 cells visualized in the SVM decision value plot indicates the possibility to apply Raman firstly to detect the presence of infection. Secondly, the differentiation of the SARS-CoV-2 and the MeV-infected cells highlights the possibility to classify the type of virus causing the infection. The simple difference spectrum generated between SARS-CoV-2 and MeV-infected cells (Fig. 4F) for spectral interpretation of the SVM differentiation (Fig. 4C) indicates changes in the protein profile based on the Raman signature of

different proteins observed in SARS-CoV-2 and MeV-infected cells. This difference in the protein profile of SARS-CoV-2-infected and MeV-infected cells indicates the differential transcriptomic signature of these two viral infections.

Huge pressure on medical establishments due to the epidemic of COVID-19, leads to the development of new detection methods which require fewer material resources and background knowledge, specifically in the case of new viral infections and their different variants, and requires the identification of new biomarkers which could allow better stratification of COVID-19 patients and allow their personalized treatment. Virus-cell interaction brings new insight into this field, based on Raman spectroscopy approaches. Developing a new data-reading method called Raman barcode, shows the precision of Raman signature to detect changes in isolated virus strain⁵⁴. The study of spike RBD receptor interaction based on protein antibody recognition has been possible by lowering the detection limit using SERS⁵⁵. Development of label-free, low-cost detection method based on the significant Raman peak, indicated in this study, the testing method could be available in medical centers. The results are obtained immediately and no previous knowledge of the virus strain is needed. The accuracy, sensitivity, and specificity are high enough in the early interaction of the virus with host cells and could be used in the case of suspected viral infection and search for the identification of novel biomarkers of viral infection. Finally, Raman spectroscopy allows the identification of novel biochemical markers of viral infection, presenting thus a promising technique in the analysis of viral infections. The presented study demonstrates the proof of concept for the application of confocal Raman microscopy combined with advanced data analysis, to study SARS-CoV-2 infection in cell culture and possibly develop novel COVID-19-related biomarkers. The spectral changes were followed after KMCA analysis on different cellular organelles, and separation of relevant single spectra for hyperspectral analysis, opening novel perspectives for further research in the field.

Methods

Virus, infection and cell culture

Vero E6 cells (African green monkey kidney cells) were obtained from ECACC and maintained in Dulbecco's minimal essential medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) at 37°C with 5% CO₂. SARS-CoV-2 isolate BetaCoV/France/IDF0372/2020, was supplied by the National Reference Center for Respiratory Viruses hosted by Institute Pasteur (Paris, France). The human sample from which this strain was isolated has been provided from the Bichat Hospital, Paris, France. Moreover, the BetaCoV/France/IDF0372/2020 strain was supplied through the European Virus Archive goes Global (EVAg) platform, a project that has received funding from the European Union's Horizon 2020 research and innovation program under the grant agreement No 653316. The virus was propagated in Vero E6 cells in DMEM medium containing 2.5% FBS at 37°C with 5% CO₂ and harvested 72 h post-inoculation. Vero E6 cells were plated on Raman glass slides for 24 h and then infected with 10⁶ PFU of SARS-CoV-2 for 24 h multiplicity of infection (MOI) of 0.1. Cells were then fixed with 4% PFA in PBS at 24 h and 48 h post-infection, washed several times in PBS before proceeding for Raman confocal microscopy.

A similar procedure was followed for the infection of Vero E6 cells with MeV virus (Schwartz strain), in 6 well plates using an MOI of 0,1, as described previously⁵⁶. At 90 min post-infection, the virus was aspirated, and a fresh medium was added. MeV-infected Vero E6 cells were then fixed at 24 h post-infection with 4% PFA in PBS, washed several times in PBS before proceeding for Raman confocal microscopy. All the infection assays were performed in triplicate conditions.

Electron Microscopy

Vero E6 cells were infected with 1.10⁶ PFU of SARS-CoV-2 for 24h. Cells were fixed with 2,5% (v/v) glutaraldehyde in PHEM buffer and postfixed in osmium tetroxide 1% / K₄Fe(CN)₆ 0,8%, at room temperature for 1h for each treatment. The samples were then dehydrated in successive ethanol baths (50/70/90/100%) and infiltrated with propylene oxide/ EMbed812 mixes before embedding. 70 nm ultrathin cuts were made on a PTXL ultramicrotome (RMC, France), stained with OTE/lead citrate, and observed on a Tecnai G2 F20 (200kV, FEG) TEM at the Electron Microscopy Facility MRI-COMET, INM, Plate-Forme Montpellier RIO Imaging, Biocampus, Montpellier.

Raman Spectroscopy Imaging

WITec Confocal Raman Microscope Alpha 300 R (WITec Inc., Ulm, Germany) is used to collect Raman spectra. A frequency-doubled Nd:YAG laser (Newport, Evry, France) with 532nm wavelength and 50mW power provided sample excitation. A 60x NIKON water immersion objective (numerical aperture of NA = 1.0.) focused laser beam on PBS immersed cells. An electron-multiplying charge-coupled device (EMCCD) camera (DU 970 N-BV353, Andor, Hartford, USA) captured the scattered signals. Using the formula $r_{\text{lateral}} = 1.22 \cdot \lambda_{\text{laser}} / 2 \cdot \text{NA}$ gives the spatial resolution of the system 325 nm. For the axial resolution, $r_{\text{axial}} = 1.4 \cdot \lambda_{\text{laser}} \cdot n / \text{NA}^2$ (where n is the index of refraction 1.33 for the water-based objective) gives 991 nm. WITec Image Plus software performed data acquisition and processing. Calcium fluoride (CaF₂) substrate was employed due to its characteristic Raman peak at 320cm⁻¹ to avoid extra Raman signal interfering with cells signature. Each Raman scan contains more than twenty thousand single spectra. The recorded Raman spectra collected from each voxel (300nm·300nm·900nm) contain the sample biochemical fingerprint under the laser spot of 1µm.

Raman Spectral Data Analysis

Reconstructed Raman images using KMCA

To generate a false color Raman image, advanced data analysis based on two steps was applied. The first method is collecting the integrated Raman intensities of the C-H stretching mode to visualize the cell image. The second method *k*-mean clusters analysis (KMCA) contains the post-processing of the recorded spectra. KMCA is a simple algorithm to divide all collected spectra into *k* mutual groups. It considers each data set -Raman spectrum- as a spot in multidimensional space. KMCA splits these spots in groups within each, spots -Raman spectra- are as close to each other as possible, and as far from spots in other groups as possible. Based on the presence of specific peaks or their relative intensities, the *k* clusters are used to identify Raman information or separation of certain cellular organelles (including the nucleus, nucleolus, mitochondria, Golgi, and cytoplasm). KMCA is done using WITec Project Plus software (Ulm, Germany).

Raman Spectra Preprocessing

Raman spectra were extracted from the different intracellular components, namely nucleus, Golgi-mitochondria bodies, and cytoplasm. The infection experiments were performed in triplicate and from each experimental batch, 40 spectra/condition were extracted from each intracellular component. The Raman spectra were preprocessed using GNU R platform⁵⁷ with an in-house built script. Before preprocessing, Raman spectra were truncated from 360 cm⁻¹ to 3050 cm⁻¹, and background was subtracted using the Sensitive Nonlinear Iterative Peak (SNIP)^{58,59} algorithm. The background-subtracted spectra were further truncated and spectral range from 400 cm⁻¹ to 1800 cm⁻¹ and 2700 cm⁻¹ to 3050 cm⁻¹ and the spectra were vector normalized. The difference Raman spectra were generated and plotted using OriginPro v2020 software (OriginLab Corporation, Northampton, MA, USA).

Principal Component Analysis (Pca)

The PCA was performed on the normalized Raman spectra and principal component (PC) score plots and loadings coefficients were generated. Raman spectra along with PCA score plots and loadings were plotted using Origin software.

Support Vector Machine (Svm)

The support vector machine (SVM) algorithm was used to build the Raman models. For the SVM analysis, Raman spectral data from three biological replication experiments were used. Firstly, PCA analysis was performed, separately, on the Raman

spectra extracted from the nucleus, Golgi-mitochondria, and cytoplasm. The first 13 principal components (PCs) obtained from PCA analysis were used as an input for the SVM analysis. Three separate, two-class SVM models were built to differentiate between 1) non-infected and MeV-infected Vero E6 cells, 2) non-infected and SARS-CoV-2 infected Vero E6 cells and 3) MeV and SARS-CoV-2 infected Vero E6 cells. The parameters chosen for SVM analysis are as follow: $\gamma = 1e-4$, $\text{cost} = 200$, $\text{kernel} = \text{"radial"}$, $\text{type} = \text{"C-classification"}$. The SVM analysis was cross-validated using 10-fold cross-validation. The SVM decision value plots were generated using Origin software.

Declarations

Author contributions:

HS performed Raman measurements and KMCA data analysis, wrote the manuscript and AR performed the PCA, SVM data analysis and contributed to draft preparation. PM, DM and SM performed viral stock, cell culture and infection; AN performed TEM sample preparation and imaging; BH and DM edited the manuscript and raised funding. JP and FC planned the study and reviewed the manuscript.

Acknowledgments:

Raman microscopy analysis during this study was realized using the EDMOS platform which was created with the financial support of the Region Occitanie (France) and the European Regional Development Fund (ERDF).

References

1. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.***382**, 727–733 (2020).
2. Saviñon-Flores, F. *et al.* A Review on SERS-Based Detection of Human Virus Infections: Influenza and Coronavirus. *Biosensors***11**, 66 (2021).
3. Toh, S. Y., Citartan, M., Gopinath, S. C. B. & Tang, T.-H. Aptamers as a replacement for antibodies in enzyme-linked immunosorbent assay. *Biosens. Bioelectron.***64**, 392–403 (2015).
4. Griffiths, C., Drews, S. J. & Marchant, D. J. Respiratory Syncytial Virus: Infection, Detection, and New Options for Prevention and Treatment. *Clin. Microbiol. Rev.***30**, 277–319 (2017).
5. Lu, H. A Longitudinal Study of a Novel Dot-Enzyme-Linked Immunosorbent Assay for Detection of Avian Influenza Virus. *Avian Dis.***47**, 361–369 (2003).
6. Spackman, E. *et al.* Development of a Real-Time Reverse Transcriptase PCR Assay for Type A Influenza Virus and the Avian H5 and H7 Hemagglutinin Subtypes. *J. Clin. Microbiol.***40**, 3256–3260 (2002).
7. Ramoji, A. *et al.* Understanding viruses and viral infections by biophotonic methods. *Transl. Biophotonics* (2022) doi:10.1002/tbio.202100008.
8. Gardy, J. L. & Loman, N. J. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat. Rev. Genet.***19**, 9–20 (2018).
9. Quick, J. *et al.* Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.***12**, 1261–1276 (2017).
10. Dronina, J., Samukaite-Bubniene, U. & Ramanavicius, A. Advances and insights in the diagnosis of viral infections. *J. Nanobiotechnology***19**, 348 (2021).
11. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.***23**, 3–20 (2022).
12. Bracquemond, D. & Muriaux, D. Betacoronavirus Assembly: Clues and Perspectives for Elucidating SARS-CoV-2 Particle Formation and Egress. *mBio***12**, e02371-21 (2021).

13. Pistiki, A. *et al.* Biochemical Analysis of Leukocytes after In Vitro and In Vivo Activation with Bacterial and Fungal Pathogens Using Raman Spectroscopy. *Int. J. Mol. Sci.***22**, 10481 (2021).
14. Elumalai, S., Managó, S. & De Luca, A. C. Raman Microscopy: Progress in Research on Cancer Cell Sensing. *Sensors***20**, 5525 (2020).
15. Gulka, M. *et al.* Simultaneous label-free live imaging of cell nucleus and luminescent nanodiamonds. *Sci. Rep.***10**, 9791 (2020).
16. Santos, M. C. D., Nascimento, Y. M., Araújo, J. M. G. & Lima, K. M. G. ATR-FTIR spectroscopy coupled with multivariate analysis techniques for the identification of DENV-3 in different concentrations in blood and serum: a new approach. *RSC Adv.***7**, 25640–25649 (2017).
17. Silva, L. G. *et al.* ATR-FTIR spectroscopy in blood plasma combined with multivariate analysis to detect HIV infection in pregnant women. *Sci. Rep.***10**, 20156 (2020).
18. Huang, J. *et al.* Vibrational spectroscopy as a powerful tool for follow-up immunoadsorption therapy treatment of dilated cardiomyopathy – a case report. *The Analyst***145**, 486–496 (2020).
19. Rauwel, E. *et al.* Assessing Cobalt Metal Nanoparticles Uptake by Cancer Cells Using Live Raman Spectroscopy. *Int. J. Nanomedicine***15**, 7051–7062 (2020).
20. Salehi, H. *et al.* Confocal Raman data analysis enables identifying apoptosis of MCF-7 cells caused by anticancer drug paclitaxel. *J. Biomed. Opt.***18**, 056010 (2013).
21. Salehi, H. *et al.* Label-free detection of anticancer drug paclitaxel in living cells by confocal Raman microscopy. *Appl. Phys. Lett.***102**, 113701 (2013).
22. Ramoji, A. *et al.* Raman Spectroscopy Follows Time-Dependent Changes in T Lymphocytes Isolated from Spleen of Endotoxemic Mice. *ImmunoHorizons***3**, 45–60 (2019).
23. Ramoji, A. *et al.* Leukocyte Activation Profile Assessed by Raman Spectroscopy Helps Diagnosing Infection and Sepsis. *Crit. Care Explor.***3**, e0394 (2021).
24. Vian, R. *et al.* Adsorption of proteins on TiO₂ particles influences their aggregation and cell penetration. *Food Chem.***360**, 130003 (2021).
25. Salehi, H. *et al.* Confocal Raman data analysis enables identifying apoptosis of MCF-7 cells caused by anticancer drug paclitaxel. *J. Biomed. Opt.***18**, 056010 (2013).
26. Early detection of virus infection in live human cells using Raman spectroscopy.
27. Moor, K. *et al.* Noninvasive and label-free determination of virus infected cells by Raman spectroscopy. *J. Biomed. Opt.***19**, 067003 (2014).
28. Taha, B. A., Al Mashhadany, Y., Hafiz Mokhtar, M. H., Dzulkefly Bin Zan, M. S. & Arsad, N. An Analysis Review of Detection Coronavirus Disease 2019 (COVID-19) Based on Biosensor Application. *Sensors***20**, 6764 (2020).
29. Birech, Z., Mwangi, P. W., Bukachi, F. & Mandela, K. M. Application of Raman spectroscopy in type 2 diabetes screening in blood using leucine and isoleucine amino-acids as biomarkers and in comparative anti-diabetic drugs efficacy studies. *PLOS ONE***12**, e0185130 (2017).
30. Carlomagno, C. *et al.* COVID-19 salivary Raman fingerprint: innovative approach for the detection of current and past SARS-CoV-2 infections. *Sci. Rep.***11**, 4943 (2021).
31. Sitjar, J. *et al.* Challenges of SERS technology as a non-nucleic acid or -antigen detection method for SARS-CoV-2 virus and its variants. *Biosens. Bioelectron.***181**, 113153 (2021).
32. Cialla, D. *et al.* Surface-enhanced Raman spectroscopy (SERS): progress and trends. *Anal. Bioanal. Chem.***403**, 27–54 (2012).
33. Scholtz, A. *et al.* COVID-19 Diagnostics: Past, Present, and Future. *ACS Photonics***8**, 2827–2838 (2021).
34. Saviñon-Flores, F. *et al.* A Review on SERS-Based Detection of Human Virus Infections: Influenza and Coronavirus. *Biosensors***11**, 66 (2021).

35. Snijder, E. J. *et al.* A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. *PLoS Biol.***18**, e3000715 (2020).
36. Mendonça, L. *et al.* Correlative multi-scale cryo-imaging unveils SARS-CoV-2 assembly and egress. *Nat. Commun.***12**, 4629 (2021).
37. Notingher, I. *et al.* Discrimination between ricin and sulphur mustard toxicity *in vitro* using Raman spectroscopy. *J. R. Soc. Interface***1**, 79–90 (2004).
38. Chen, H., Wurm, T., Britton, P., Brooks, G. & Hiscox, J. A. Interaction of the Coronavirus Nucleoprotein with Nucleolar Antigens and the Host Cell. *J. Virol.***76**, 5233–5250 (2002).
39. Old, O. J. *et al.* Vibrational spectroscopy for cancer diagnostics. *Anal. Methods***6**, 3901 (2014).
40. Shetty, G., Kendall, C., Shepherd, N., Stone, N. & Barr, H. Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus. *Br. J. Cancer***94**, 1460–1464 (2006).
41. Koljenović, S., Schut, T. B., Vincent, A., Kros, J. M. & Puppels, G. J. Detection of Meningioma in Dura Mater by Raman Spectroscopy. *Anal. Chem.***77**, 7958–7965 (2005).
42. Chan, J. W. *et al.* Micro-Raman Spectroscopy Detects Individual Neoplastic and Normal Hematopoietic Cells. *Biophys. J.***90**, 648–656 (2006).
43. Stone, N., Kendall, C., Smith, J., Crow, P. & Barr, H. Raman spectroscopy for identification of epithelial cancers. *Faraday Discuss.***126**, 141 (2004).
44. Cheng, W.-T., Liu, M.-T., Liu, H.-N. & Lin, S.-Y. Micro-Raman spectroscopy used to identify and grade human skin pilomatrixoma. *Microsc. Res. Tech.***68**, 75–79 (2005).
45. Stone, N., Kendall, C., Shepherd, N., Crow, P. & Barr, H. Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *J. Raman Spectrosc.***33**, 564–573 (2002).
46. Huang, Z. *et al.* Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *Int. J. Cancer***107**, 1047–1052 (2003).
47. Wakisaka, Y. *et al.* Probing the metabolic heterogeneity of live *Euglena gracilis* with stimulated Raman scattering microscopy. *Nat. Microbiol.***1**, 16124 (2016).
48. Lee, K. S. *et al.* An automated Raman-based platform for the sorting of live cells by functional properties. *Nat. Microbiol.***4**, 1035–1048 (2019).
49. Huleihel, M., Shufan, E., Zeiri, L. & Salman, A. Detection of Vero Cells Infected with Herpes Simplex Types 1 and 2 and Varicella Zoster Viruses Using Raman Spectroscopy and Advanced Statistical Methods. *PLOS ONE***11**, e0153599 (2016).
50. Tiwari, D., Jakhmola, S., Pathak, D. K., Kumar, R. & Jha, H. C. Temporal *In Vitro* Raman Spectroscopy for Monitoring Replication Kinetics of Epstein–Barr Virus Infection in Glial Cells. *ACS Omega***5**, 29547–29560 (2020).
51. Cervenka, I., Agudelo, L. Z. & Ruas, J. L. Kynurenines: Tryptophan’s metabolites in exercise, inflammation, and mental health. *Science***357**, eaaf9794 (2017).
52. Michaelis, S. *et al.* Assessment of tryptophan and kynurenine as prognostic markers in patients with SARS-CoV-2. *Clin. Chim. Acta***525**, 29–33 (2022).
53. Lionetto, L. *et al.* Increased kynurenine-to-tryptophan ratio in the serum of patients infected with SARS-CoV2: An observational cohort study. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.***1867**, 166042 (2021).
54. Pezzotti, G. *et al.* Raman Molecular Fingerprints of SARS-CoV-2 British Variant and the Concept of *Raman Barcode*. *Adv. Sci.* 2103287 (2021) doi:10.1002/advs.202103287.
55. Awada, C., Abdullah, M. M. B., Traboulsi, H., Dab, C. & Alshoabi, A. SARS-CoV-2 Receptor Binding Domain as a Stable-Potential Target for SARS-CoV-2 Detection by Surface–Enhanced Raman Spectroscopy. *Sensors***21**, 4617 (2021).
56. Horvat, B. *et al.* Transgenic mice expressing human measles virus (MV) receptor CD46 provide cells exhibiting different permissivities to MV infections. *J. Virol.***70**, 6673–6681 (1996).

57. Silagadze, Z. K. A new algorithm for automatic photopeak searches. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.***376**, 451–454 (1996).
58. Ryan, C. G., Clayton, E., Griffin, W. L., Sie, S. H. & Cousens, D. R. SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.***34**, 396–402 (1988).
59. *R-Project*. (R Foundation for Statistical Computing).

Figures

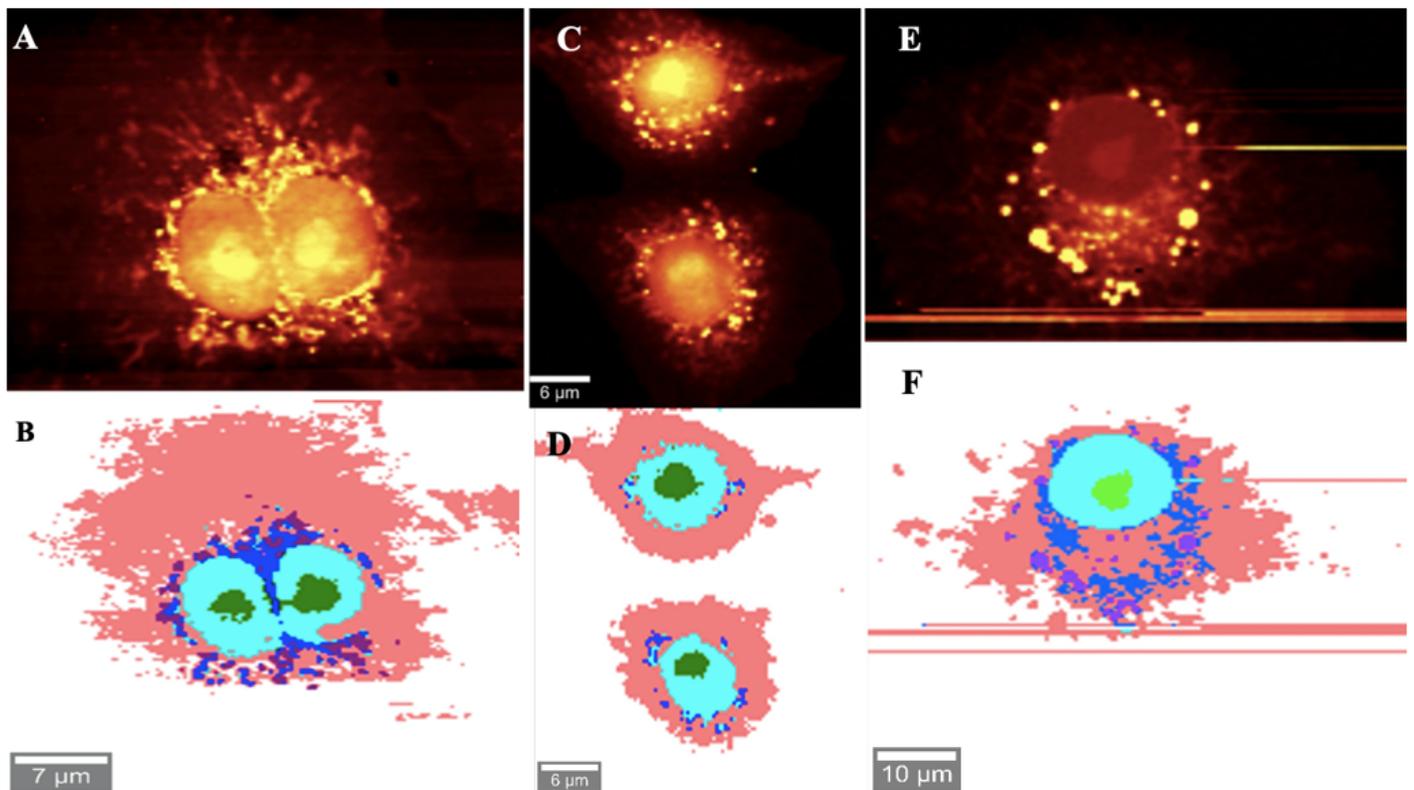


Figure 1

Raman confocal images of SARS-CoV-2 and MeV-infected (24h pi) and non-infected Vero E6 cells. **A:** Raman images of non-infected cells. Light yellow color corresponds to highest intensities of lipids and proteins and dark shadow for lowest intensities. **B:** KMCA results for the non-infected cells of part A. Nucleolus is marked in dark green, the nucleus in turquoise, mitochondria, and golgi are in blue, cytoplasm in pink and lipid droplet in purple. The same color code has been used for all analyses. **C:** SARS-CoV-2-infected Raman C-H images and **D:** relevant KMCA illustrations for C part. **E:** MeV-infected Raman C-H images and **F:** relevant KMCA illustrations for E part. Scale bars are variable between 6-10 μm as indicated.

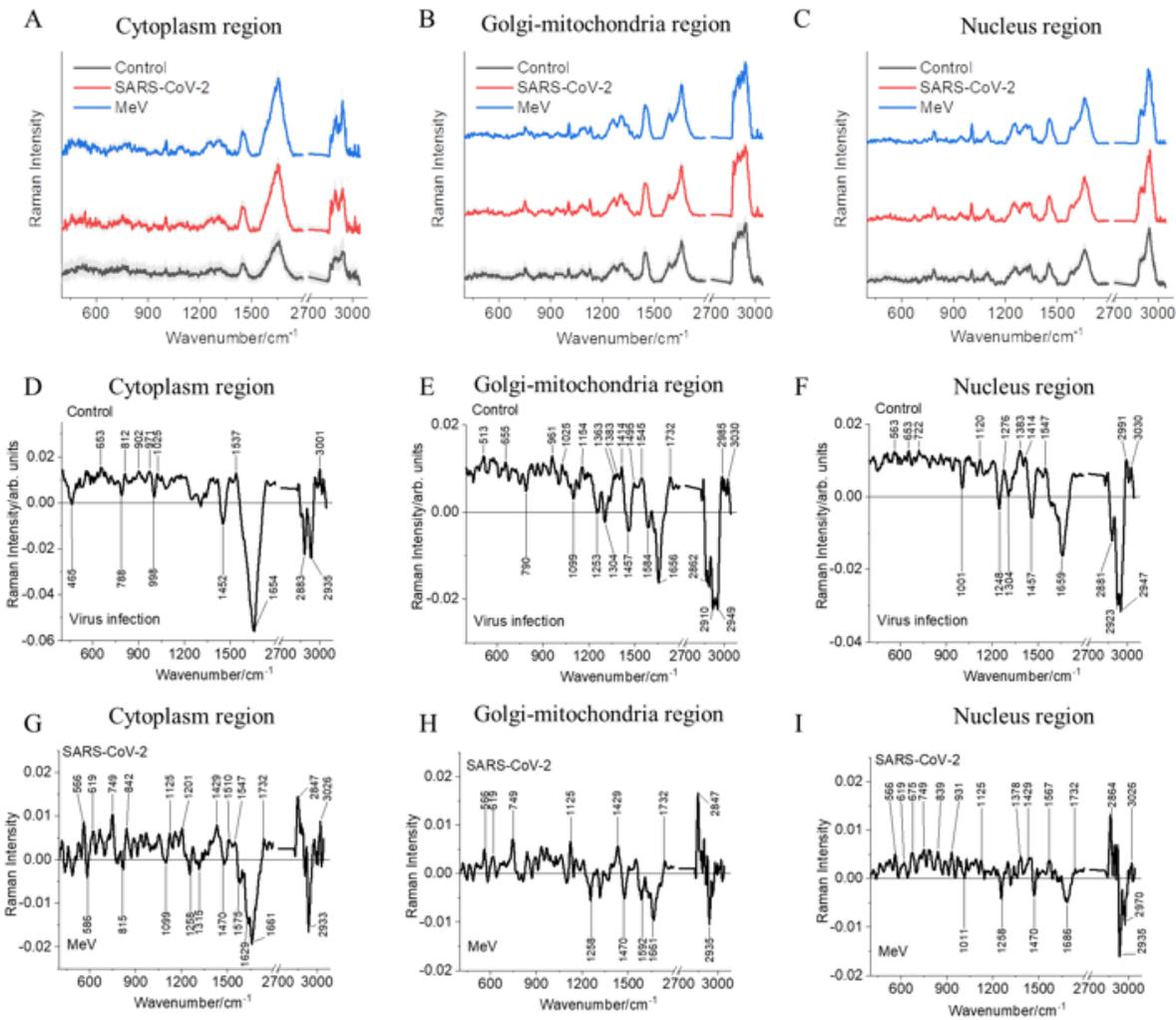


Figure 2

Mean Raman spectra from the cytoplasm, Golgi-mitochondria bodies, and nucleus region shown for non-infected (Control) and infected Vero E6 cells (SARS-CoV-2 and MeV) along with the standard deviation. A) Raman spectra extracted from cytoplasm region, B) Raman spectra extracted from Golgi-mitochondria bodies, and C) Raman spectra extracted from nucleus region. D to F) Difference Raman spectra calculated between non-infected (Control) and virus infected cells (MeV and SARS-CoV-2 taken together) for D) cytoplasm, E) Golgi-mitochondria and F) nucleus region of the Vero E6 cells. G to I) Difference Raman spectra calculated between the SARS-CoV-2 infected cells and MeV infected cells and for G) cytoplasm, H) Golgi-mitochondria and I) nucleus region of the Vero E6 cells. The difference Raman spectra have been smoothed using the Savitzky-Golay method with points of window 15 and polynomial order 2. The Raman spectra are shifted on the y-axis for sake of clarity.

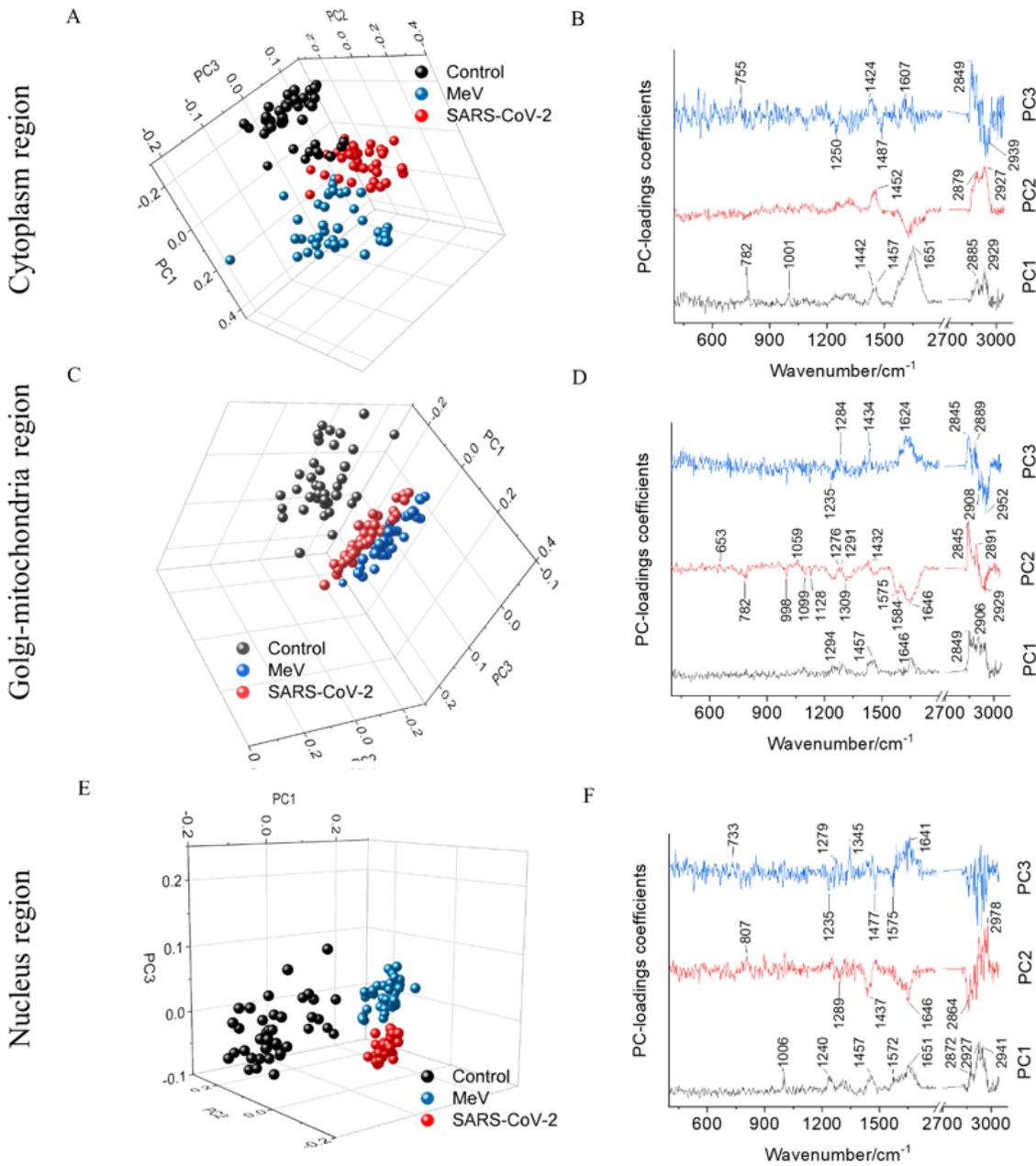


Figure 3

Principal component analysis (PCA) of Raman spectra extracted from intracellular components of SARS-CoV-2 and MeV infected and non-infected (Control) Vero E6 cells. 3D PCA score plot of Raman spectra from **A**) cells' cytoplasm and **B**) corresponding PC loadings, **C**) cells' Golgi-mitochondria bodies, **D**) corresponding PC loadings, **E**) cells' nucleus, and **F**) corresponding PC loadings. The loading coefficients are shifted on the y-axis for sake of clarity.

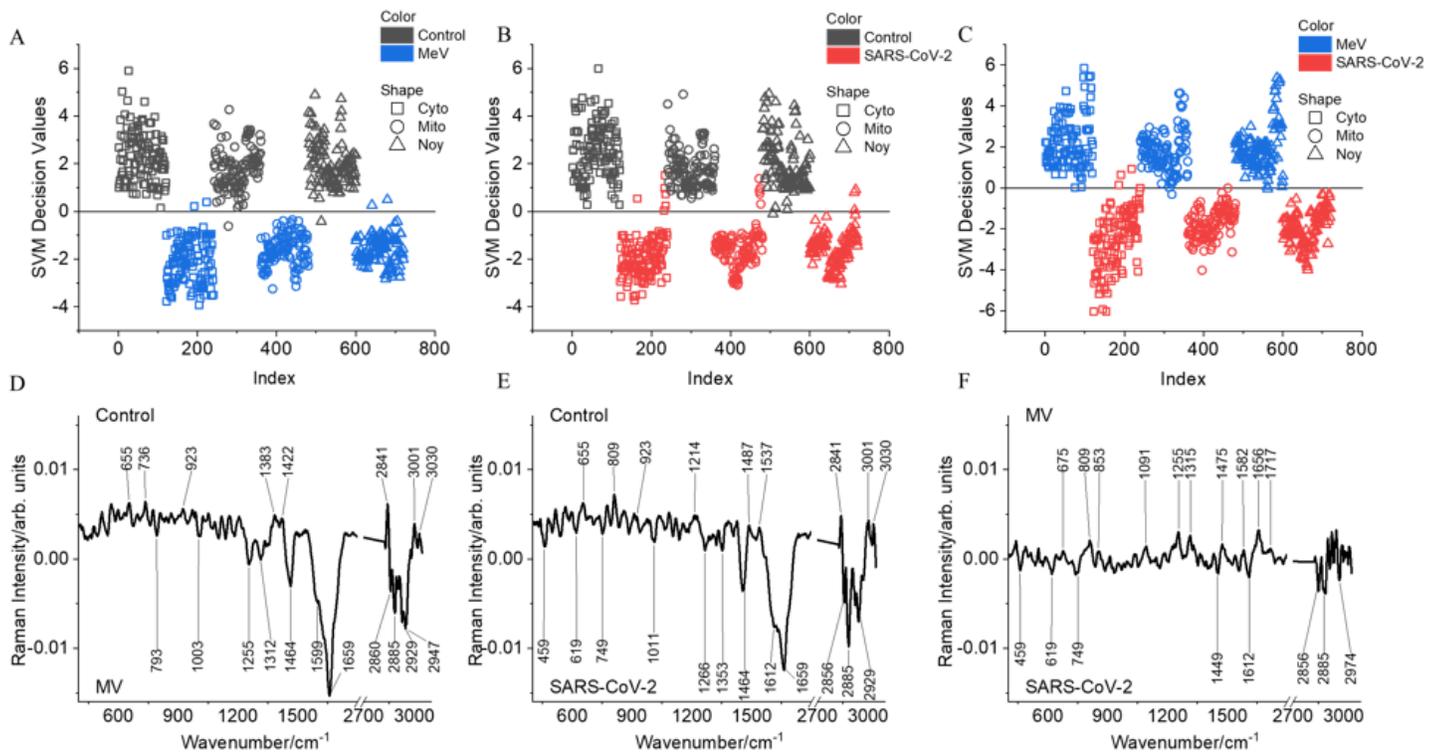


Figure 4

Raman model generated using support vector machine (SVM) algorithm to differentiate A) non-infected (Control, black) from measles virus (MeV, blue) infected Vero E6 cells, B) non-infected (Control, black) from SARS-CoV-2 (red) infected Vero E6 cells and C) MeV and SARS-CoV-2 infected Vero E6 cells. The model was generated using Raman spectra extracted from intracellular components: cytoplasm (Cyto, square), Golgi-mitochondria bodies (Mito, circle), and nucleus (Noy, triangle) of the Vero E6 cells (Total accuracy A) 98.89%, B) 97.22%, C) 97.78%

PCs used 13, 10 fold CV other parameters same as below)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarydata10MARCH2022.docx](#)