Analyst

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Check for updates

Cite this: *Analyst*, 2024, **149**, 5535

Liquid saliva-based Raman spectroscopy device with on-board machine learning detects COVID-19 infection in real-time†

Katherine J. I. Ember,[a,b] Nassim Ksantini,[a,b] Frédérick Dallaire,[a,b] Guillaume Sheehy,[a,b] Trang Tran,[a,b] Mathieu Dehaes,[c,d,e] Madeleine Durand,[b,f] Dominique Trudel[b,g] and Frédéric Leblond*[a,b,g]

Received 22nd May 2024,
Accepted 28th August 2024

DOI: 10.1039/d4an00729h

rsc.li/analyst

With greater population density, the likelihood of viral outbreaks achieving pandemic status is increasing. However, current viral screening techniques use specific reagents, and as viruses mutate, test accuracy decreases. Here, we present the first real-time, reagent-free, portable analysis platform for viral detection in liquid saliva, using COVID-19 as a proof-of-concept. We show that vibrational molecular spectroscopy and machine learning (ML) detect biomolecular changes consistent with the presence of viral infection. Saliva samples were collected from 470 individuals, including 65 that were infected with COVID-19 (28 from hospitalized patients and 37 from a walk-in testing clinic) and 251 that had a negative polymerase chain reaction (PCR) test. A further 154 were collected from healthy volunteers. Saliva measurements were achieved in 6 minutes or less and led to machine learning models predicting COVID-19 infection with sensitivity and specificity reaching 90%, depending on volunteer symptoms and disease severity. Machine learning models were based on linear support vector machines (SVM). This platform could be deployed to manage future pandemics using the same hardware but using a tunable machine learning model that could be rapidly updated as new viral strains emerge.

Introduction

COVID-19 has brought to light the destructive power of pandemics. The SARS-CoV-2 virus killed millions, infected hundreds of millions more, cost jobs and disrupted food supply chains worldwide.[1,2] With climate change, greater population density, international travel and increased contact between humans and animals, pandemics and epidemics are likely to increase in frequency. COVID-19 demonstrated the importance of accessible, affordable, and rapid viral testing in pandemic control. It also provided a testing environment for new detection technologies.

Current viral testing techniques are limited. For COVID-19, the average polymerase chain reaction (PCR) tests are the gold standard. However, these are time-consuming, require trained personnel and are expensive. The average PCR test costs $ 127 (ref. 3) and takes hours from sample acquisition to diagnosis if a PCR machine is on site,[4] which is often not the case. Meanwhile, rapid antigen tests are limited in accuracy with recent studies suggesting that 90% of asymptomatic individuals go undetected.[5] Furthermore, both PCR and rapid antigen tests rely on tailored biochemical reagents which must potentially be re-adapted for new viral strains. Indeed, the U.S. Food and Drug Administration (FDA) has warned that PCR tests may be less effective at detecting variants, a situation that has already been observed with influenza.[6] Management of future pandemics require low-cost, high-sensitivity tests allowing frequent screening regardless of variants.

Raman spectroscopy (RS) is a label-free analytical tool using laser light to yield molecular information about samples.[7] The Raman scattering phenomenon was predicted in 1928 by Smekal[8] and experimentally proven in 1928 by Raman and Krishnan.[9] When light is shone at a sample, the light inelastically scattered back from the sample gives information about molecular structure and bonding, visualized as a Raman spec-

[a]Department of Engineering Physics, Polytechnique Montréal, Montreal, Quebec, Canada. E-mail: frederic.leblond@polymtl.ca
[b]Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, Quebec, Canada
[c]Department of Radiology, Radio-oncology and Nuclear Medicine, Université de Montréal, Montreal, Canada
[d]Institute of Biomedical Engineering, Université de Montréal, Montreal, Canada
[e]Centre de Recherche du Centre Hospitalier Universitaire Sainte-Justine (CRCHUSJ), Montreal, Canada
[f]Internal Medicine service, Centre Hospitalier de l'Univsersité de Montréal (CHUM), Montreal, Quebec, Canada
[g]Institut du cancer de Montréal, Montreal, Quebec, Canada

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4an00729h

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

trum. The position of a peak on the *x*-axis of a Raman spectrum gives information about the identity of the molecule responsible for the peak. Meanwhile, the *y*-axis is related to the concentration of the molecules in the sample.[7,10,11] RS has proven sensitive to disease and metabolic state of tissues and biofluids, rendering it a possible screening tool for cancer,[12–14] organ health,[15] and pathogenic infections.[16,17] We developed a high-speed, low-cost, reagent-free, portable instrument using RS to detect the biomolecular changes associated with COVID-19 infection in saliva. The switch from normal cellular metabolism to viral synthesis,[18] the death of host cells,[19] and the activation of the immune response[20] all bring about biomolecular changes which may affect biofluid composition. In previous studies, vibrational spectroscopy has been used to detect COVID-19 infection in saliva from senior, hospitalized volunteers.[21,22] Ember *et al.* (including authors of this manuscript) extended this to asymptomatic and symptomatic volunteers at a walk-in COVID testing clinic.[23] This achieved a sensitivity of 79% and a specificity of 75% in males, and a sensitivity of 84% and a specificity of 64% in females. However, these results were obtained using an expensive, commercial, slow RS instrument unsuitable for widespread testing.

An RS-based rapid COVID test would not rely on chemical reagents which can be costly, require refrigeration, have limited shelf life, and can lose their specificity as the virus mutates. This can limit deployment in low- and middle-income countries where rapid on-site screening with minimal sample preparation and operator involvement can be required. Furthermore, an RS-based system can integrate a user-friend graphical user interface into a portable device, rendering it more widely applicable than PCR tests. Here, a study is presented that was designed to evaluate Raman spectroscopy for its potential to detect COVID-19 infected individuals based on their saliva, specifically in the supernatant. The dataset presented was collected during the COVID-19 pandemic in 2020 and is the first demonstration of label-free, optical COVID-19 infection detection using liquid saliva samples. It is also the first case in which a system suitable for point-of-care use has been employed. As of 2023, SARS-CoV-2 symptoms have rapidly declined in severity; however, it remains in widespread circulation and may mutate into more severe variants. Infection is associated with lung damage,[24] brain damage[25] and long-term exhaustion.[26]

Outbreaks in the workplace continue to directly impact the economy whilst outbreaks in care homes and hospitals often prove lethal. Detection of the virus, therefore, remains paramount.

Experimental

Study design and saliva sample acquisition

Measurements from 470 saliva supernatant samples were taken. Of these, 251 samples of non-stimulated saliva were obtained from individuals that had a negative PCR test who had come to be tested for COVID-19 at the Point St Charles COVID testing clinic between December 2020 and April 2021. A further 37 were collected from individuals with a positive COVID-19 PCR test from the same clinic. Further details on saliva collection are mentioned in Ember *et al.*[23] Volunteers did not necessarily have symptoms consistent with a viral infection (*e.g.*, cough, runny nose, difficulty breathing), but may have come to be tested in the case of exposure to someone with COVID-19. To assess the capacity to detect COVID-19 in hospitalized patients, 28 samples were acquired from patients that were hospitalized with COVID-19 symptoms at the Centre hospitalier de l'universite de Montreal (CHUM) between January 2021 and May 2021. These patients were considered contagious by the hospital staff and confined to the "hot zone" of the hospital. Finally, 154 samples were also acquired from healthy individuals with no COVID symptoms and no suspicion of COVID-19 infection (collected from workers at the Centre hospitalier de l'université de Montreal (CHUM) and Centre de Recherche du CHUM (CRCHUM) in October 2020). This latter cohort also had not travelled abroad within the previous two weeks.

A saliva preparation protocol was developed aiming at minimizing the presence of confounding saliva constituents during the spectroscopy interrogation process (*e.g.*, food products) which is described in detail in a paper by Ember *et al.*[23] The protocol allows the sample to be collected at any moment during the day and includes using rinsing with water, and optimized centrifugation cycles. This separates samples into pellet (containing food debris, not used) and supernatant component used for analyses. Samples were aliquoted into 4 tubes and stored in a −80 °C freezer for research. Supernatant samples were thawed and a 10 μL drop pipetted and deposited on a low-Raman background aluminum holder, and Raman spectra were immediately acquired from the liquid sample. Each spectrum was correlated with the correspondent demographics and infection status characteristics (ESI Table S1†). For dried saliva supernatant samples, a 10 μL drop pipetted and deposited on a low-Raman background aluminum slide, and dried for 45 minutes.

Raman spectroscopy system and data pre-processing

The single-point system was in *epi*-illumination mode and used a 785 nm laser source (Model IO785MM1500M4S, Innovative Photonic Solutions, USA) with an output of 1.5 W and a spectral bandwidth <2 nm (Fig. 1A). The other main constituent of the system was a spectrometer (HT model, EmVision, USA) composed of a diffraction grating and a charge-coupled device (CCD) camera (Newton 920, Oxford Instruments, USA) resulting in a spectral resolution <8.7 cm^{-1}. A source fiber guided the light beam to a collimating lens and a band-pass filter (Semrock, USA) before intersecting with a dichroic mirror reflecting light to a focusing lens onto the sample. The light re-emitted from the sample went through another focusing lens, the dichroic mirror, a high-pass filter (Semrock, USA), a second collimating lens and a fiber optics bundle. The latter was connected to the spectrometer and consisted of 9 optical fibers with a core diameter of 300 μm and a

 This journal is © The Royal Society of Chemistry 2024

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

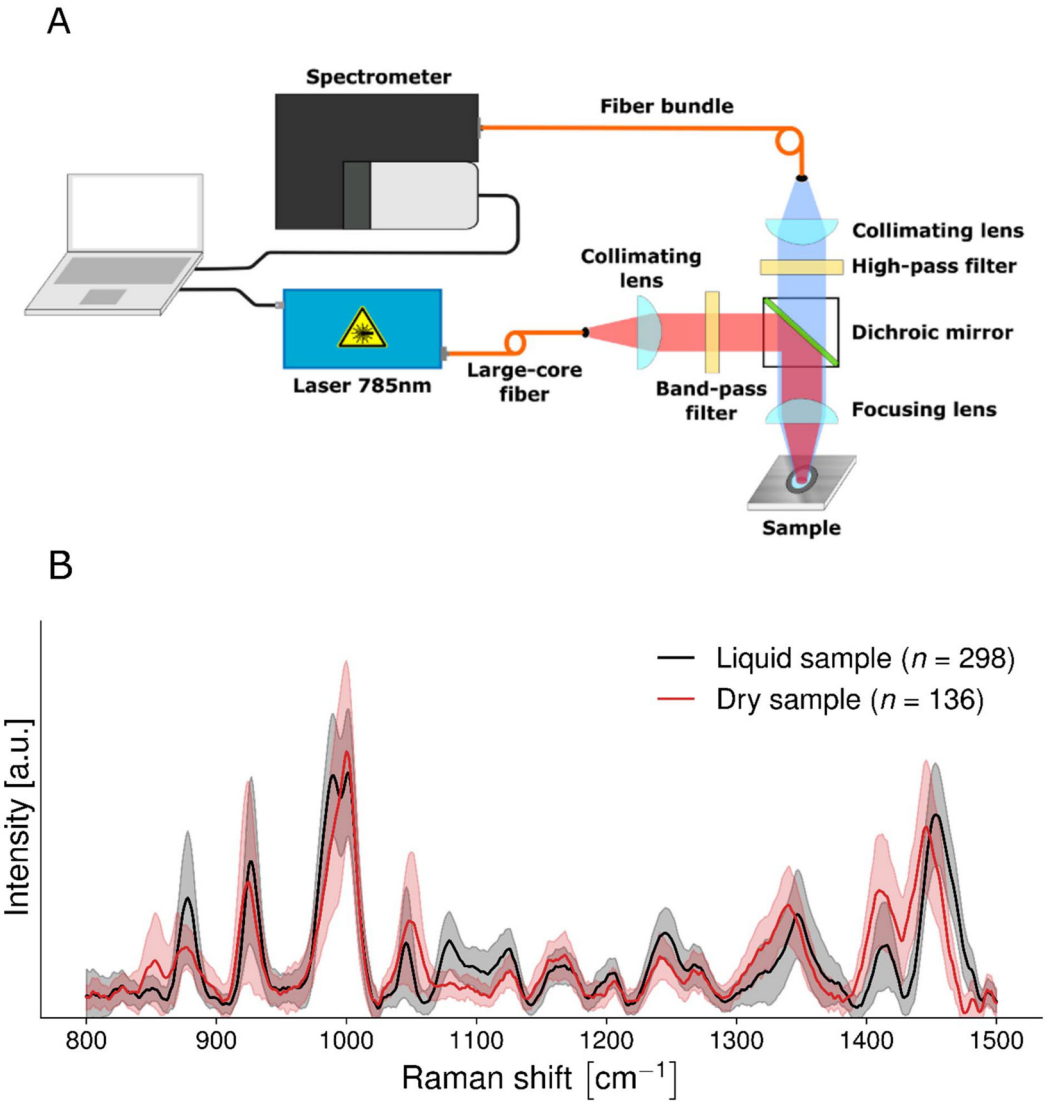

Fig. 1 (A) Schematic representation of the single-point Raman spectroscopic biofluid imaging system. The system is in *epi*-illumination with a collection path formed from a source fiber, a collimating lens, a band-pass filter, a dichroic mirror and a focusing lens, which illuminates the sample in free space. The collection path includes a high-pass filter, a collimating lens and a fiber bundle which guides the scattered light to the spectrometer. (B) Mean measurement computed from Raman spectra from liquid saliva supernatant samples (n = 298, black line) and dried saliva supernatant samples (n = 136, red line). Spectra were taken using a custom single-point system. Differences in Raman spectra were likely due to changes in confirmation of biomolecules during the drying process. The standard deviation is shown as translucent shading.

numerical aperture of 0.22. The focusing optics consisted of a 25 mm diameter lens with a focal length of 50 mm. This resulted in a depth of focus of about 12.56 mm at a laser wavelength of 785 nm. The system also included a 3D motorized stage, a white light source and a brightfield camera. The system was controlled through an in-house custom acquisition MATLAB software (MathWorks, MA, USA) and a microcontroller (Arduino).

Before each measurement the system CCD sensor was cooled to −80 °C. Calibration of the x-axis (Raman shift) was determined from a spectrum acquired using acetaminophen powder (Tylenol®) prior to each measurement. The system response was characterized using the fluorescence spectrum of a standard reference material (SRM 2214, National Institute of Standards and Technology, NIST, USA). For each sample, a dark count measurement was taken with the laser off (integration time of 1600 ± 800 ms). Then, a series of 200 repeat spectroscopy measurements were acquired. The laser power was kept at a fixed output value of 890 mW at the surface of the biofluid. Automated exposure control was used to optimize the overall photon counts while ensuring no camera saturation resulted.[27] The laser spot size was approximately 1 mm^2. Several processing steps were applied to each raw spectrum to isolate the vibrational spectroscopy contribution using a custom software.[20] These included: averaging of the 200 repeat spectra, subtraction of dark count spectrum acquired with the

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

laser turned off, normalization with the NIST standard to correct for the instrument response, *x*-axis (wavenumber shift) calibration and interpolation, baseline removal of low-frequency background signals using the BubbleFill algorithm,[28] and standard normal variate (SNV) normalization.[28] The spectral range is of the instrument is 350 cm^{-1}–2100 cm^{-1}. Furthermore, spectra are cropped at 1500 cm^{-1} because of the very large water peak between ~1520 and 1720 cm^{-1}, which completely covers the Amide I peak.

A few metrics were computed to assess of the signal quality of each measured spectrum. Prior to the SNV normalization of the Raman spectra, two metrics were computed: (i) the signal-to-noise ratio (SNR), which is the ratio between the total number of photonic counts in a Raman spectrum and the square root of the total signal (Raman + baseline + dark count), and (ii) the signal-to-background ratio (SBR), namely the ratio between the total number of photonic counts in the Raman spectrum and in the baseline. Then a quality factor (QF) value was computed on the SNV-normalized, which was used to separate the dataset into low and high quality (ESI Fig. S1†). The QF metric ranges between 0 and 1; low QF values usually means higher stochastic noise, lower inelastic scattering photonic counts, and poorly defined Raman peaks. A random signal would have a QF close to 0.[28]

In each SNV-normalized spectrum, any peak with an intensity higher than 0.5 was fitted with a Gaussian to extract its position, height, and width, with a 2 cm^{-1} tolerance on the position. Only peaks present in at least 50% of the dataset were considered for training machine learning models.[29]

To compare liquid and dried saliva supernatant samples, measurements were taken from 136 dried samples and 298 liquid samples.

COVID-19 detection machine learning models

Spectral fingerprints were collected from 470 individuals (ESI Fig. S2†). Nine machine learning models were developed for the detection of infection (Table 1). Models were built to classify COVID negative samples from COVID positive samples. In terms of a numbering system, the number gives information about the negative sample set and the letter gives information about the positive sample set. Models 1a–1c relied on PCR-confirmed COVID negative samples as the negative group, whereas Models 2a–2c additionally incorporated samples from healthy volunteers with no suspicion of COVID-19 infection, no contact with a COVID-19 positive case, and no COVID-19 symptoms. Models 3a–3c used only this group of healthy volunteers as a negative sample set. All "a" models relied on PCR-confirmed COVID positive samples as the positive group. All "b" models relied on COVID-19 hospitalized patients as the positive group. All "c" models relied on both PCR-confirmed COVID positive samples and COVID-19 hospitalized patients as a positive group.

Prior to machine learning model training and validation, the feature pool set consists of 700 individual intensity bins and 45 peak features (15 peaks × 3 features). The number of features was reduced to include only those that contributed the most to the variance between the categories. This was accomplished using linear support vector machines (SVM) with L1 regularization (regularization parameter between 0.05 and 0.5). Machine learning model training from the dimensionally reduced features set was then done using linear SVM with the regularization parameter C. Each time a model was trained, hyperparameters (number of features, C) were selected by carrying out a grid search across all possible combinations. The regularization parameter C was varied between 0.05 and 1, the number of individual bands was varied between 10 and 40 for the individual intensities and between 5 and 15 for the peak features. Because we wanted our models to account for age and sex, these two variables were added to the training pool set as extra features, making the total number of features for a given model between 17 and 57. This was done as sex and age had been revealed to be a potential confounding variable in other analysis.[23] Only spectra with QF > 0.4 were considered for classification and outliers were also removed (spectra with an intensity at 1002 cm^{-1} lower than 1). For each combination, performance was assessed using leave-one-out 5-fold cross validation based on the number of false/true positives and false/true negatives, by comparing the model predic-

Table 1 Description of machine learning (ML) models created for COVID-19 detection. Numbers outside of brackets indicate the number of samples used in the machine learning models before the quality factor cutoff. Numbers in brackets indicate the number of samples used in the machine learning models after the quality factor cutoff. The groups were: PCR-confirmed COVID negative from testing clinic (PCR negative), PCR-confirmed COVID positive from testing clinic (PCR positive), COVID-19 hospitalized patients (hospitalized), healthy volunteers with very low COVID risk (healthy)

	Negative group size (after cutoff)	Negative group description	Positive group size (after cutoff)	Positive group description
Model 1a	251 (174)	PCR negatives	37 (30)	PCR positives
Model 1b	251 (174)	PCR negatives	28 (23)	Hospitalized
Model 1c	251 (174)	PCR negatives	65 (53)	37 PCR positives + 28 hospitalized
Model 2a	405 (293)	251 PCR negatives + 154 healthy	37 (30)	PCR positives
Model 2b	405 (293)	251 PCR negatives + 154 healthy	28 (23)	Hospitalized
Model 2c	405 (293)	251 PCR negatives + 154 healthy	65 (53)	37 PCR positives + 28 hospitalized
Model 3a	154 (124)	Healthy	37 (30)	PCR positives
Model 3b	154 (124)	Healthy	28 (23)	Hospitalized
Model 3c	154 (124)	Healthy	65 (53)	37 PCR positives + 28 hospitalized

 This journal is © The Royal Society of Chemistry 2024

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

tion with the assigned PCR test result. Accuracy, sensitivity, and specificity were calculated from a receiver-operating-characteristic (ROC) curve analysis and the area under curve (AUC) was reported. The features that were retained by the model for machine learning model building are detailed in Table 2.

Results

Signal-to-noise ratio of Raman spectra from liquid saliva supernatant samples is comparable to that of dried samples

Saliva spectra were obtained from liquid samples with peaks that were comparable with saliva spectra from dry samples (Fig. 1B). The main Raman peaks that were detected in this study were consistent with those from our previous work studying samples using Raman micro-spectroscopy.[23] Peaks include those due to proteins, amino acids, glucose, salts, fatty acids, and lipids.[23,30–35]

Time from supernatant being obtained to spectral acquisition was 51 minutes for dried samples and 6 minutes for wet samples. The mean signal-to-noise (SNR) for dried saliva supernatant was 15.31 ± 7.66 whilst the mean SNR for liquid saliva supernatant was 15.73 ± 4.98. We are therefore able to obtain slightly more stable and Raman-rich data from wet samples than from dried samples. The signal to background ratio (SBR) for dried saliva supernatant was 0.0048 ± 0.0016 whilst the mean SNR for liquid saliva supernatant was 0.0044 ± 0.0023.

Spectral shifts in some peaks were apparent when comparing wet samples with dried samples. This is consistent with numerous studies indicating that Raman spectra of biomolecules change depending on whether they are interrogated in solid or liquid state.[36] For example, peaks at 877, 989 and 1079 cm^{-1} were clearly visible in the liquid saliva (Table 2) and were not as apparent in the dried saliva. The peaks at 877 and 1079 cm^{-1} correspond to the peaks of dihydrogen phosphate (H_2PO_4) in solution and also to weak peaks of monohydrogen phosphate (HPO_4) in solution, and the peak at 989 corresponds with the very strong peak of monohydrogen phosphate in solution (ESI Fig. S3†).[30]

It is also worth noting that the saliva supernatant spectra from dried samples taken using a Raman microscope and those taken using our device are comparable in terms of relative peak intensity and position (ESI Fig. S4†).

Raman spectra from COVID-19 hospitalised individuals exhibit changes compared to COVID-negative individuals

The mean Raman spectra for each COVID status group were plotted in Fig. 2A. This included COVID-19 negative PCR confirmed, COVID-19 positive PCR confirmed, healthy, and COVID-19 hospitalized individuals. The spectra from COVID-19 hospitalized individuals exhibited the greatest differences in comparison to any other spectra (red line, Fig. 2A). Compared to the COVID negative samples from the testing clinic (black solid line, Fig. 2A), peaks at 989 cm^{-1} (HPO_4) and 927 cm^{-1} (protein N–C–C) were lower, whilst peaks at 876 and 1080 were higher (corresponding to (H_2PO_4)).[23,30,31] Both *hypo*- and hyperphosphatemia (low and high levels of phosphate in the blood respectively) have been associated with increased COVID-19 infection severity,[37–40] and this may be reflected in saliva.

Raman spectra from the COVID positive PCR samples taken from the COVID testing clinic (blue line, Fig. 2A) exhibited greater overlap with the COVID negative PCR samples. However, there were still spectral differences at 876 and 989 cm^{-1}, associated with phosphates,[23,30] and at 1347 cm^{-1}, associated with multiple amino acid side chains.[23,31,36] This latter change suggests that COVID-19 infection is associated with either a difference in the composition of free amino acids

Table 2 List of features used in model building and their assignments as determined using compounds in model saliva (model saliva correspondence) and from literature (biomolecular assignment). The full list of models are shown in Table 1. The peak center indicates the centre of the Raman band in wavenumbers (cm^{-1}). Amino acids listed next to the protein may either be bound within proteins or present free in saliva

Peak center [cm^{-1}]	Biomolecular assignement	Models								
		1a	2a	3a	1b	2b	3b	1c	2c	3c
878	Phosphate (dihydrogen phosphate)[23,31]				√	√	√	√		√
927	Protein (N–C–C)[23,32]						√			
990	Phosphate (monohydrogen phosphate)[23,31]									
1001	Protein (phenylalanine, tryptophan),[23,30,32] carotenoids,[33,34] urea[23]		√		√	√				
1046	Nitrate,[23,31] protein (phenylalanine)[23,30,32]									
1080	Phosphate (dihydrogen phosphate, monohydrogen phosphate),[23,31] lipid (C–C)[35]									
1090		√		√					√	
1126	Fatty acid (C–C),[35] protein (C–N, serine),[23,30,32] glucose[33,34,36]	√	√	√	√		√	√	√	√
1163	Fatty acid (C–C)[35]				√			√	√	√
1205	Protein (tyrosine)[23,30,32]			√				√	√	
1245	Protein (amide III)[23,32]				√	√	√		√	
1267	Protein (amide III, histidine, valine),[30,32] glucose,[33,34,36] lipid (=CH)[35]	√		√	√		√	√		√
1347	Protein (histidine, leucine, lysine, methionine, serine, threonine)[23,30,32]	√					√			
1416	Lipid (beta, CH_2),[35] protein (alanine, cystine, glycine, lysine, methionine, proline, serine, threonine)[23,30,32]	√		√						
1453	Protein (C–H, glycine, isoleucine, lysine, valine)[23,30,32] lipid (CH_2/CH_3)[35]		√							

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

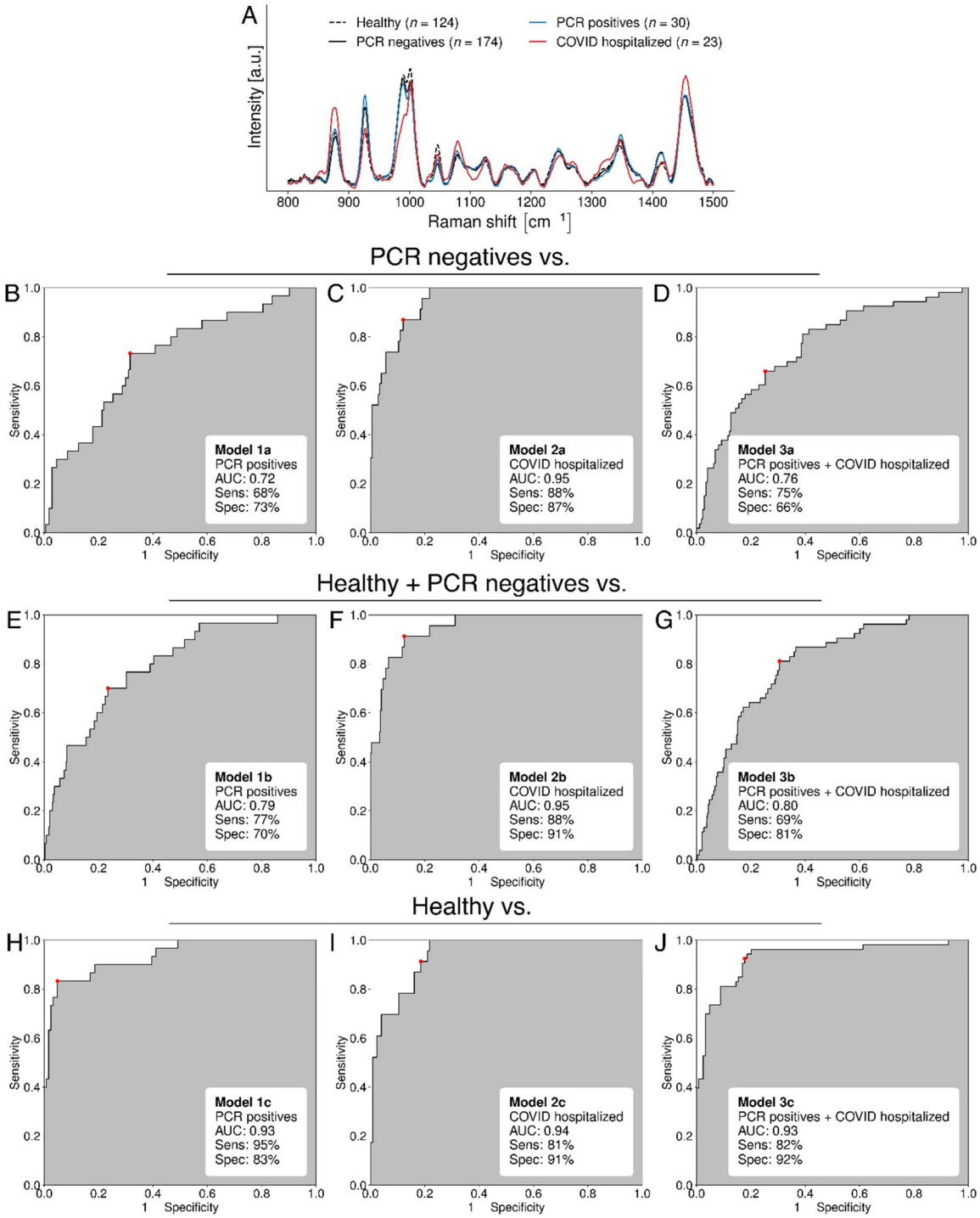

Fig. 2 Raman spectra and associated machine learning models classifying liquid saliva supernatant samples from PCR negative volunteers at a COVID-19 testing clinic, PCR positives from a COVID-19 testing clinic, healthy volunteers and hospitalized COVID patients. (A) Mean spectra were taken from PCR negatives from a COVID-19 testing clinic (n = 174 black line), PCR positives from a COVID-19 testing clinic (n = 30, blue line), healthy volunteers (n = 124, black dashed line) and hospitalized COVID patients (n = 23, red line). Spectra were taken using a custom single-point system with <6 min acquisition time. (B–J) Receiver operating characteristic (ROC) curves for machine learning models discriminating between saliva supernatant spectral fingerprints. Negative groups are (B–D) PCR-confirmed COVID-19 negative volunteers from a testing clinic, (E–G) PCR-confirmed COVID-19 negatives and healthy volunteers. Positive groups are (B, E and H), (C, F and I) and (D, G and J) (B) PCR-confirmed COVID-19 positives from a testing clinic, (C) patients hospitalized with COVID-19, (D) COVID-19 individuals from a testing clinic and hospitalized patients. Panels on the left-hand side of the figure (A, C and E) show the The red dot on each ROC represents the point at which the accuracy, sensitivity, and specificity are calculated *i.e.* the point with the minimal distance to the upper left corner.

 This journal is © The Royal Society of Chemistry 2024

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

in saliva, or a difference in the types of proteins found in saliva. Indeed, there are metabolomics studies with nuclear magnetic resonance (NMR) spectroscopy showing that alanine, glutamine, histidine, leucine, lysine, phenylalanine, and proline were all downregulated in COVID-19 PCR positives *vs.* PCR negatives.[41]

The greater overlap of the COVID-negative samples with the positives from the testing clinic is likely because testing clinic volunteers exhibited fewer symptoms and less severe pathology than the hospitalized patients. The metabolism of testing clinic volunteers was therefore less likely to be perturbed by the SARS-CoV-2 virus, and the saliva less likely to contain immune cells, cytokines, metabolic by-products, and cellular debris associated with COVID-19 infection. Multiple studies show that changes in metabolism such as ceramide metabolism, tryptophan degradation, lipoproteins and cholesterol.[42,43] Notably, one saliva sample from a hospitalized COVID patient had visible tissue (possibly lung tissue) present in the sample whilst this was not the case with any testing clinic samples. Saliva from hospitalized patients was also much stickier than from other individuals.

Machine learning models detect COVID-19 in both testing clinic individuals and hospitalized individuals

Machine learning (ML) models were trained to discriminate between PCR-confirmed COVID-19 negative PCR-tested individuals from a testing clinic and COVID-19 positives (Fig. 2B–D). ML was first applied to discriminate between PCR-confirmed COVID-19 negatives and positives from volunteers at the testing clinic. This yielded a receiver operating characteristic (ROC) curve with and area-under-curve (AUC) of 0.72, a sensitivity of 68% and a specificity of 73% (Model 1a, Fig. 2B, and Table 2). The sensitivity of lateral flow tests is 25.6%–99.1% and specificity is 92.4–100%.[44,45] For COVID-19 detection in the testing clinic, our sensitivity is therefore greater than the median lateral flow test sensitivity (62%) and our specificity is lower than that of lateral flow tests. This model used multiple Raman peaks associated with protein, one associated with lipid, and two peaks associated with glucose, suggesting that these metabolites change with COVID-19 infection.

An ML model for discrimination between PCR confirmed COVID-19 negative individuals and hospitalized individuals resulted in an AUC of 0.95, corresponding to a sensitivity of 88% and a specificity of 87% (Model 2a, Fig. 2C, and Table 2). This sensitivity is 26% higher than the median lateral flow test sensitivity, and the specificity is only slightly lower, exhibiting an increase of 14% compared to detection of those in the testing clinic. As stated earlier, the improvement in performance when detecting COVID-19 in hospitalized individuals compared to testing clinic volunteers is to be expected, as disease severity can impact the viral load and the metabolic signature of COVID-19 in biofluids.[46] The three key features used were 1001, 1126 and 1453 cm^{-1}, all of which can be found in proteins, but also have contributions from glucose and lipids. The peak at 1001 cm^{-1} is assigned to phenylalanine which may be present within proteins or as a free amino acid. A study using NMR spectroscopy shows that phenylalanine in saliva is reduced in COVID-19 patients compared to PCR negative controls.[41]

Finally, a model was created to discriminate between PCR-confirmed COVID-19 negative individuals and the whole COVID positive dataset (hospitalized and testing clinic individuals). This led to an ROC curve with an AUC of 0.76, corresponding to a sensitivity of 75% and a specificity of 66% (Model 3a, Fig. 2D, and Table 2). All features overlapped with those of model 1A except there is one extra feature at 1205 cm^{-1} associated with tyrosine.[23,31,36]

Increasing the number of samples increases test sensitivity and specificity. COVID-19 negative testing clinic individuals may have had symptoms of a respiratory virus. To address this, we included 154 samples from volunteers without COVID-19 symptoms who had no suspicion of COVID-19 infection (Fig. 2E–G). These saliva samples were not from a testing clinic and therefore had not been PCR confirmed as negative, so were not included in the initial analysis. However, they had a very low likelihood of COVID-19 infection and, at the time, only 0.52% of the Canadian population tested positive for COVID-19.[47] Including these samples increased our presumed COVID-19 negative dataset by 124 (number of spectra after quality factor cutoff), specifically from the number of healthy people with no symptoms of COVID-19.

An ML model trained to discriminate between total COVID-19 negative individuals (healthy + PCR negatives) and PCR positives resulted in an AUC of 0.79, corresponding to a sensitivity of 77% and a specificity of 70% (Model 1b, Fig. 2E, and Table 2). This was a slightly higher AUC than that of Model 1a. The ML model trained to classify total COVID-19 negative individuals and COVID-19 hospitalized patients showed an almost identical AUC, sensitivity, and specificity (0.95, 88% and 91% respectively) to the model produced with PCR-confirmed negative samples only (Model 2b, Fig. 2F, and Table 2). The model trained from the whole dataset (hospitalized and testing clinic individuals) led to a ROC curve with a greater AUC of 0.80, corresponding to a sensitivity of 69% and a specificity of 81% (Fig. 2G). Overall, these increases in accuracy may be because the "B" machine learning models have more examples of COVID-19 negative spectra than the "A" machine learning models. Therefore, they can distinguish features that are particular to COVID-19 negative spectra more easily than when using PCR-negative cases alone.

Individuals with no COVID-19 symptoms may be easier to detect than COVID-19 negative individuals who have respiratory symptoms. To give an indication as to whether respiratory symptoms may affect our ability to detect COVID-19, ML models were created to discriminate between those who had no COVID-19 symptoms and those who had COVID-19. This negative group in these group "c" models were healthy individuals with no suspicion of COVID-19 infection and replaced the PCR-negative COVID testing clinic individuals in the group "a" models. The COVID-19 testing clinic volunteers may have had other respiratory symptoms, leading to biomolecular changes in their saliva.

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Model 1C discriminated between the healthy group and COVID-19 positive volunteers from a testing clinic, and had an AUC of 0.93, a sensitivity of 95% and a specificity of 83% (Model 1c, Fig. 2H, and Table 2). Both sensitivity and specificity were greatly improved compared to Model 1A. The sensitivity is 33% higher than median lateral flow test sensitivity and the specificity is 13% lower. These parameters would be highly useful in a case where the need for detection of infected individuals is greater than the detection of non-infected individuals, such as choosing which individuals could visit a clinic, hospital or care home.

Model 2c discriminated between the healthy group and COVID-19 hospitalized individuals with an AUC of 0.94, sensitivity of 81% and specificity of 91%, comparable to both "a" and "b" models.

Finally, Model 3c, which was developed to classify healthy and all COVID positive samples, achieved an AUC of 0.93, a sensitivity of 82% and a specificity of 92%. Both sensitivity and specificity are comparable to those of lateral flow tests.

All "c" models used the peak at 1163 cm^{-1} which was not present in the "a" models. This was associated with fatty acids, which are implicated in respiratory diseases. Metabolomics using liquid chromatography and high-resolution mass spectrometry have shown that the lipid profile of human sputum changes with different respiratory viruses (*e.g.* influenza H3, rhinovirus).[48] Furthermore, Pérez-Torres *et al.* suggest that SARS-CoV-2 may alter fatty acid metabolism as total non-esterified fatty acids are reduced in plasma compared to healthy subjects.[49] It may be that changes in fatty acids could indicate the presence of respiratory viruses. The peak at 1126 cm^{-1} is the most consistent feature between all models, although it was not a major feature in the Model 2B for discriminating between all negatives and COVID-19 hospitalised patients (Table 2). This peak is the major peak due to aqueous glucose. Glucose dysregulation is strongly linked with COVID-19 severity.[50–52] The peak at 1126 cm^{-1} also has contributions from proteins and unsaturated lipids (discussed earlier).

Conclusions

The aim of this study was to determine whether Raman spectroscopy using a single-point portable device could show sufficient sensitivity and specificity for identification of COVID-19 infected individuals. The mean acquisition time per sample was less than 6 minutes and the classification detected infected individuals with sensitivities and specificities of 68–95% and 66–92% respectively, with varying degrees of accuracy for hospitalized and non-hospitalized individuals. We found that AUCs of the machine learning models increased when the COVID-negative dataset included more samples from those without any respiratory symptoms ("c" models) and when the COVID-positive dataset included more samples from those with severe symptoms (models with hospitalized patients). In the future, we must verify to what extent we can use RS to detect COVID-19 specifically, our sensitivity and specificity when accounting for different symptoms, and to what extent we can detect general respiratory viruses *e.g.* influenza, rhinovirus, respiratory syncytial virus (RSV).

The COVID-19 positive cohort analysed in the Ember *et al.* 2022 (see Table 1 of that article) were almost identical to the COVID-19 positive testing clinic cohort analysed in this study. In terms of the COVID-19 negative testing clinic cohort, we were able to analyse over six times more samples using our rapid single-point system than with the microscope system. We matched the positive and negative samples for as many demographic characteristics as possible. Finally, the addition of healthy volunteers who had no suspicion of respiratory infection allowed us to determine that there is an increase in accuracy when using COVID negatives who have no respiratory infection symptoms or chance of respiratory infection in ML models.

Overall, the preliminary results shown in this article provide enticing evidence that Raman spectroscopy and machine learning could be utilized for rapid biochemical analysis and biofluid classification for infectious disease screening. The true value of the platform applied to infectious disease characterization will come from machine learning models developed from larger scale datasets with high spectral quality. For example, our power studies suggest that at least 150 COVID-19 positive, 150 RSV positive and 150 influenza positive patients would allow us to use many features from the Raman spectra to develop truly generalizable models tested on independent hold-out datasets. Studies will be carefully designed to ensure that models can be trained from data fully capturing the heterogeneity of the general population and the disease pathology. For COVID-19 hospitalized patients, it is possible that the treatment for COVID-19 may itself influence the Raman spectrum of saliva through metabolic changes. Therefore, in a future clinical trial of such a device, controls would need to be taken from patients without COVID-19 but the same course of treatment.

The system used in this study is the size of a microwave on top of a cart, allowing it to be wheeled into a pharmacy, doctor's office or testing clinic. However, we have since developed a suitcase-based system for greater portability. The total cost of the system is five times cheaper than commercial Raman microscopes.

The new system will be adaptable to other biofluids, *e.g.*, blood, urine and tears, and the detection of other diseases, *e.g.*, seasonal influenza which kills 500 000 people each year, measles which is one of the most infectious human viruses, and the early stages of cancer. In future projects, we also aim to investigate the effects of using RS to monitor COVID-19 progression in terms of the pre-symptomatic, symptomatic, and immunogenic periods of the disease. To this end, SERS nanoparticles or surfaces[53] could be functionalized using ligands against viral proteins and/or antibodies,[54–56] allowing selective enhancement of the Raman signal.

In conclusion, our spontaneous Raman-based saliva assay allows detection of biomolecular changes associated with viral infection in real time.

 This journal is © The Royal Society of Chemistry 2024

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Data availability

To obtain anonymized samples, images, or processed Raman spectra, please contact Frederic Leblond directly. Code repository for model training, analysis and validation is publicly available in the paper[14] and also on Github (**https://github.com/mr-sheg/orpl**).

Conflicts of interest

Frederic Leblond and Nassim Ksantini are co-founders of Exclaro-Tridan formed in 2015 to commercialize a Raman spectroscopy system for infectious disease testing applications. He has ownership and patents in the company. Katherine Ember and Frederick Dallaire are both scientific advisors to Exclaro-Tridan.

Acknowledgements

This work is supported by the Canadian Foundation for Innovation (Exceptional Opportunities Fund: COVID-19), the Natural Sciences and Engineering Research Council of Canada (NSERC) Alliance grant program, the NSERC Discovery Grant program, the Fonds de Recherche du Québec – Santé (FRQS), the TransMedTech Institute and Institut de Valorisation des Données (IVADO). The lead author was funded by a Fonds de Recherche du Québec (FRQNT) grant during much of this research. Thank you to Myriam Mahfoud, Esmat Zamani Ahmad, Gabriel Beaudoin, Pascale Arlotto, and Marc Messier-Peet for assistance in sample collection.

References

1 P. Jha, P. E. Brown and R. Ansumana, *Lancet*, 2022, **399**, 1937–1938.
2 J. Moosavi, A. M. Fathollahi-Fard and M. A. Dulebenets, *Int. J. Disaster Risk Reduct.*, 2022, **75**, 102983.
3 John Hopkins Coronavirus Resource Center, **https://coronavirus.jhu.edu/from-our-experts/q-and-a-how-much-does-it-cost-to-get-a-covid-19-test-it-depends**.
4 C. H. Chau, J. D. Strope and W. D. Figg, *Pharmacotherapy*, 2020, **40**, 857–868.
5 A. Soni, C. Herbert, H. Lin, Y. Yan, C. Pretz, P. Stamegna, B. Wang, T. Orwig, C. Wright, S. Tarrant, S. Behar, T. Suvarna, S. Schrader, E. Harman, C. Nowak, V. Kheterpal, L. V. Rao, L. Cashman, E. Orvek, D. Ayturk, L. Gibson, A. Zai, S. Wong, P. Lazar, Z. Wang, A. Filippaios, B. Barton, C. J. Achenbach, R. L. Murphy, M. L. Robinson, Y. C. Manabe, S. Pandey, A. Colubri, L. O'Connor, S. C. Lemon, N. Fahey, K. L. Luzuriaga, N. Hafer, K. Roth, T. Lowe, T. Stenzel, W. Heetderks, J. Broach and D. D. McManus, *Ann. Intern. Med.*, 2023, **176**, 975–982.
6 K. A. Stellrecht, *J. Clin. Microbiol.*, 2018, **56**, e01531–e01517.
7 E. Smith and G. Dent, *Modern Raman S pectroscopy: A Practical Approach*, John Wiley & Sons Ltd, 2005.
8 A. Smekal, *Die Naturwissensch.*, 1928, **16**, 612–613.
9 C. V. Raman and K. S. Krishnan, *Nature*, 1928, **121**, 501–502.
10 D. W. Shipp, F. Sinjab and I. Notingher, *Adv. Opt. Photonics*, 2017, **9**, 315.
11 D. A. Long, *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*, Wiley, 2002.
12 M. Jermyn, K. Mok, J. Mercier, J. Desroches, J. Pichette, K. Saint-Arnaud, L. Bernstein, M. C. Guiot, K. Petrecca and F. Leblond, *Sci. Transl. Med.*, 2015, **7**, 1–10.
13 A. A. Grosset, F. Dallaire, T. Nguyen, M. Birlea, J. Wong, F. Daoust, N. Roy, A. Kougioumoutzakis, F. Azzi, K. Aubertin, S. Kadoury, M. Latour, R. Albadine, S. Prendeville, P. Boutros, M. Fraser, R. G. Bristow, T. van der Kwast, M. Orain, H. Brisson, N. Benzerdjeb, H. Hovington, A. Bergeron, Y. Fradet, B. Têtu, F. Saad, F. Leblond and D. Trudel, *PLoS Med.*, 2020, **17**, 1–20.
14 S. David, T. Tran, F. Dallaire, G. Sheehy, F. Azzi, D. Trudel, F. Tremblay, A. Omeroglu, F. Leblond and S. Meterissian, *J. Biomed. Opt.*, 2023, **28**, 036009.
15 K. J. I. Ember, F. Hunt, L. E. Jamieson, J. M. Hallett, H. Esser, T. J. Kendall, R. E. Clutton, R. Gregson, K. Faulds, S. J. Forbes, G. C. Oniscu and C. J. Campbell, *Hepatology*, 2021, 1–16.
16 S. A. Camacho, R. G. Sobral-Filho, P. H. B. Aoki, C. J. L. Constantino and A. G. Brolo, *ACS Sens.*, 2018, **3**, 587–594.
17 D. Sebba, A. G. Lastovich, M. Kuroda, E. Fallows, J. Johnson, A. Ahouidi, A. N. Honko, H. Fu, R. Nielson, E. Carruthers, C. Diédhiou, D. Ahmadou, B. Soropogui, J. Ruedas, K. Peters, M. Bartkowiak, M. N'Faly, S. Mboup, Y. B. Amor, J. H. Connor and K. Weidemaier, *Sci. Transl. Med.*, 2018, **10**(471), eaat0944.
18 D. Walsh and I. Mohr, *Nat. Rev. Microbiol.*, 2011, **9**, 860–875.
19 C. Yuan, Z. Ma, J. Xie, W. Li, L. Su, G. Zhang, J. Xu, Y. Wu, M. Zhang and W. Liu, *Signal Transduction Targeted Ther.*, 2023, **8**, 357.
20 A. P. D. Iwamura, M. R. Tavares Da Silva, A. L. Hümmelgen, P. V. Soeiro Pereira, A. Falcai, A. S. Grumach, E. Goudouris, A. C. Neto and C. Prando, *Rev. Med. Virol.*, 2021, **31**, e2199.
21 C. Carlomagno, D. Bertazioli, A. Gualerzi, S. Picciolini, P. I. Banfi, A. Lax, E. Messina, J. Navarro, L. Bianchi, A. Caronni, F. Marenco, S. Monteleone, C. Arienti and M. Bedoni, *Sci. Rep.*, 2021, **11**, 1–13.
22 B. R. Wood, K. Kochan, D. E. Bedolla, N. Salazar-Quiroz, S. Grimley, D. Perez-Guaita, M. J. Baker, J. Vongsvivut, M. Tobin, K. Bambery, D. Christensen, S. Pasricha, A. K. Eden, A. Mclean, S. Roy, J. Roberts, J. Druce, D. A. Williamson, J. McAuley, M. Catton, D. Purcell, D. Godfrey and P. Heruad, *Angew. Chem., Int. Ed.*, 2021, **60**(31), 17102–17107.
23 K. Ember, F. Daoust, M. Mahfoud, F. Dallaire, E. Z. Ahmad, T. Tran, A. Plante, M.-K. Diop, T. Nguyen, A. St-Georges-

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Robillard, N. Ksantini, J. Lanthier, A. Filiatrault, G. Sheehy, G. Beaudoin, C. Quach, D. Trudel and F. Leblond, *J. Biomed. Opt.*, 2022, **27**, 25002.
24 F. K. F. Kommoss, C. Schwab, L. Tavernar, J. Schreck, W. L. Wagner, U. Merle, D. Jonigk, P. Schirmacher and T. Longerich, *Dtsch. Arztebl. Int.*, 2020, **117**(29–30), 500–506.
25 O. A. Gomazkov, *Biol. Bull. Rev.*, 2022, **12**, 131–139.
26 R. Twomey, J. DeMars, K. Franklin, S. N. Culos-Reed, J. Weatherald and J. G. Wrightson, *Phys. Ther.*, 2022, **102**, pzac005.
27 F. Dallaire, F. Picot, J.-P. Tremblay, G. Sheehy, É. Lemoine, R. Agarwal, S. Kadoury, D. Trudel, F. Lesage, K. Petrecca and F. Leblond, *J. Biomed. Opt.*, 2020, **25**, 1.
28 G. Sheehy, F. Picot, F. Dallaire, K. Ember, T. Nguyen, K. Petrecca, D. Trudel and F. Leblond, *J. Biomed. Opt.*, 2023, **28**(2), 025002.
29 A. Plante, F. Dallaire, A.-A. Grosset, T. Nguyen, M. Birlea, J. Wong, F. Daoust, N. Roy, A. Kougioumoutzakis, F. Azzi, K. Aubertin, S. Kadoury, M. Latour, R. Albadine, S. Prendeville, P. Boutros, M. Fraser, R. G. Bristow, T. Van Der Kwast, M. Orain, H. Brisson, N. Benzerdjeb, H. Hovington, A. Bergeron, Y. Fradet, B. Têtu, F. Saad, D. Trudel and F. Leblond, *J. Biomed. Opt.*, 2021, **26**(11), 116501.
30 M. D. Fontana, K. Ben Mabrouk and T. H. Kauffmann, in *Spectroscopic Properties of Inorganic and Organometallic Compounds*, ed. J. Yarwood, R. Douthwaite and S. Duckett, Royal Society of Chemistry, Cambridge, 2013, vol. 44, pp. 40–67.
31 A. Rygula, K. Majzner, K. M. Marzec, A. Kaczor, M. Pilarczyk and M. Baranska, *J. Raman Spectrosc.*, 2013, **44**, 1061–1076.
32 C. G. Atkins, K. Buckley, M. W. Blades and R. F. B. Turner, *Appl. Spectrosc.*, 2017, **71**, 767–793.
33 J. Y. Qu, B. C. Wilson and D. Suria, *Appl. Opt.*, 1999, **38**, 5491.
34 K. Czamara, K. Majzner, M. Z. Pacia, K. Kochan, A. Kaczor and M. Baranska, *J. Raman Spectrosc.*, 2015, **46**, 4–20.
35 A. J. Berger, T.-W. Koo, I. Itzkan, G. Horowitz and M. S. Feld, *Appl. Opt.*, 1999, **38**, 2916.
36 G. Zhu, X. Zhu, Q. Fan and X. Wan, *Spectrochim. Acta, Part A*, 2011, **78**, 1187–1195.
37 J. Malinowska, M. Małecka-Giełdowska, D. Bańkowska, K. Borecka and O. Ciepiela, *Int. J. Infect. Dis.*, 2022, **122**, 543–549.
38 H. L. Corrêa, L. A. Deus, T. B. Araújo, A. L. Reis, C. E. N. Amorim, A. B. Gadelha, R. L. Santos, F. S. Honorato, D. Motta-Santos, C. Tzanno-Martins, R. V. P. Neves and T. S. Rosa, *Front. Immunol.*, 2022, **13**, 1006076.
39 M. Hadavi, F. Taghinezhad, E. Shafiei, S. H. Babakr, S. Bastaminejad, M. Kaffashian, I. Ahmadi and A. Mozafari, *Int. J. Endocrinol. Metab.*, 2022, **20**(3), e126386.
40 T. A. T. G. Van Kempen and E. Deixler, *Am. J. Physiol.: Endocrinol. Metab.*, 2021, **320**, E2–E6.
41 T. K. Da Silva Fidalgo, L. B. Freitas-Fernandes, B. B. F. Marques, C. S. De Araújo, B. J. Da Silva, T. C. Guimarães, R. G. Fischer, E. M. B. Tinoco and A. P. Valente, *Metabolites*, 2023, **13**, 263.
42 J. Marín-Corral, J. Rodríguez-Morató, A. Gomez-Gomez, S. Pascual-Guardia, R. Muñoz-Bermúdez, A. Salazar-Degracia, P. Pérez-Terán, M. I. Restrepo, O. Khymenets, N. Haro, J. R. Masclans and O. J. Pozo, *Int. J. Mol. Sci.*, 2021, **22**, 4794.
43 Y. Li, Y. Zhang, R. Lu, M. Dai, M. Shen, J. Zhang, Y. Cui, B. Liu, F. Lin, L. Chen, D. Han, Y. Fan, Y. Zeng, W. Li, S. Li, X. Chen, H. Li and P. Pan, *Clin. Chim. Acta*, 2021, **517**, 66–73.
44 D. A. Mistry, J. Y. Wang, M.-E. Moeser, T. Starkey and L. Y. W. Lee, *BMC Infect. Dis.*, 2021, **21**, 828.
45 C. Hogg, S. Boots, D. Howorth, C. Williams, M. Heginbothom, J. Salmon and R. Howe, *PLoS One*, 2023, **18**, e0290406.
46 S. Krishnan, H. Nordqvist, A. T. Ambikan, S. Gupta, M. Sperk, S. Svensson-Akusjärvi, F. Mikaeloff, R. Benfeitas, E. Saccon, S. M. Ponnan, J. E. Rodriguez, N. Nikouyan, A. Odeh, G. Ahlén, M. Asghar, M. Sällberg, J. Vesterbacka, P. Nowak, Á. Végvári, A. Sönnerborg, C. J. Treutiger and U. Neogi, *Mol. Cell. Proteomics*, 2021, **20**, 100159.
47 COVID-19 epidemiology update: Current situation, **https://health-infobase.canada.ca/covid-19/current-situation.html**.
48 S. T. Humes, N. Iovine, C. Prins, T. J. Garrett, J. A. Lednicky, E. S. Coker and T. Sabo-Attwood, *Respir. Res.*, 2022, **23**, 177.
49 I. Pérez-Torres, V. Guarner-Lans, E. Soria-Castro, L. Manzano-Pech, A. Palacios-Chavarría, R. R. Valdez-Vázquez, J. G. Domínguez-Cherit, H. Herrera-Bello, H. Castillejos-Suastegui, L. Moreno-Castañeda, G. Alanís-Estrada, F. Hernández, O. González-Marcos, R. Márquez-Velasco and M. E. Soto, *Front. Physiol.*, 2021, **12**, 667024.
50 S. Mirabella, S. Gomez-Paz, E. Lam, L. Gonzalez-Mosquera, J. Fogel and S. Rubinstein, *Diabetes Metab. Syndr.*, 2022, **16**, 102439.
51 L. Montefusco, M. Ben Nasr, F. D'Addio, C. Loretelli, A. Rossi, I. Pastore, G. Daniele, A. Abdelsalam, A. Maestroni, M. Dell'Acqua, E. Ippolito, E. Assi, V. Usuelli, A. J. Seelam, R. M. Fiorina, E. Chebat, P. Morpurgo, M. E. Lunati, A. M. Bolla, G. Finzi, R. Abdi, J. V. Bonventre, S. Rusconi, A. Riva, D. Corradi, P. Santus, M. Nebuloni, F. Folli, G. V. Zuccotti, M. Galli and P. Fiorina, *Nat. Metab.*, 2021, **3**, 774–785.
52 E. A. Barreto, A. S. Cruz, F. P. Veras, R. Martins, R. S. Bernardelli, I. M. Paiva, T. M. Lima, Y. Singh, R. C. Guimarães, S. Damasceno, N. Pereira, J. M. Alves, T. T. Gonçalves, J. Forato, S. P. Muraro, G. F. Souza, S. S. Batah, J. L. Proenca-Modena, M. A. Mori, F. Q. Cunha, P. Louzada-Junior, T. M. Cunha, H. I. Nakaya, A. Fabro, R. D. R. De Oliveira, E. Arruda, R. Réa, Á. Réa Neto, M. M. Fernandes Da Silva and L. O. Leiria, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**, e2217119120.

 This journal is © The Royal Society of Chemistry 2024

Open Access Article. Published on 21 October 2024. Downloaded on 11/19/2024 5:22:24 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

53 E. Zamani, N. Ksantini, G. Sheehy, K. J. I. Ember, B. Baloukas, O. Zabeida, T. Trang, M. Mahfoud, J. Sapieha, L. Martinu and F. Leblond, *Lasers Surg. Med.*, 2024, **56**, 206–217.

54 A. T. Bido, K. J. I. Ember, D. Trudel, M. Durand, F. Leblond and A. G. Brolo, *Anal. Methods*, 2023, **15**, 3955–3966.

55 P.-C. Guan, H. Zhang, Z.-Y. Li, S.-S. Xu, M. Sun, X.-M. Tian, Z. Ma, J.-S. Lin, M.-M. Gu, H. Wen, F.-L. Zhang, Y.-J. Zhang, G.-J. Yu, C. Yang, Z.-X. Wang, Y. Song and J.-F. Li, *Anal. Chem.*, 2022, **94**, 17795–17802.

56 M. Mohammadi, D. Antoine, M. Vitt, J. M. Dickie, S. Sultana Jyoti, J. G. Wall, P. A. Johnson and K. E. Wawrousek, *Anal. Chim. Acta*, 2022, **1229**, 340290.