

# Integrated Approach for Rapid and Accurate Identification of Emerging Viruses: Microfluidics, Surface-Enhanced Raman Spectroscopy, and Machine Learning

Na Zhang<sup>1</sup>, Yin-Ting Yeh<sup>1</sup>, Jiarong Ye<sup>2</sup>, Yuan Xue<sup>3</sup>, Kunyan Zhang<sup>4</sup>, Julian Arman Mintz<sup>1</sup>, Peng jin<sup>2</sup>, RyeAnne Ricker<sup>5</sup>, Nestor Perea Lopez<sup>1</sup>, Elodie Ghedin<sup>5</sup>, Shengxi Huang<sup>4</sup>, Sharon Xiaolei Huang<sup>2</sup>, Mauricio Terrones<sup>1,\*</sup>

<sup>1</sup> Department of Physics, The Pennsylvania State University, University Park, PA 16802.

<sup>2</sup> College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802

<sup>3</sup> Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218

<sup>4</sup> Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77005

<sup>5</sup> Systems Genomics Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20894

We proposed a portable microfluidic platform featuring carbon nanotube arrays with differential filtration porosity, designed for the rapid enrichment and Raman identification of viruses. <sup>1</sup> This platform, validated using avian influenza A viruses and human samples with respiratory infections, demonstrates a 70-fold enrichment enhancement with a virus specificity of 90%. The machine learning (ML) approach was applied to recognize the viruses based on the rapidly obtained Raman dataset, with remarkable accuracy for classifying different subtypes and even strains of viruses which was achieved by a convolutional neural network (CNN) classifier. <sup>2 3</sup> The ML model effectively recognizes the Raman signatures of proteins, lipids, and other functional groups present in different viruses and uses a weighted combination of these signatures to identify viruses. Besides, to address the pressing issue of the ongoing COVID-19 pandemic, the optical identification of the receptor-binding domain (RBD) of the SARS-CoV-2 spike protein was explored employing SERS with gold nanoparticles (AuNPs) for the capture of the Raman modes of the SARS-CoV-2 RBD. <sup>4</sup> This was also followed by the ML algorithms, which provide a method for distinguishing SARS-CoV-2 from MERS-CoV, enabling rapid analysis and discrimination of complex proteins in infectious viruses and other biomolecules. In order to further improve the sensitivity and bio-comparability of the SERS substrate, gold nanostars with better resonance under excitation laser were combined with the classical gold nanofilm, exhibiting better Raman signal-to-noise ratio and more characteristic peaks of the viruses for rapid analysis and identification.

This integrated approach, encompassing microfluidic platforms, label-free Raman spectroscopy, and machine learning, presents a powerful toolkit for the swift and accurate identification, characterization, and surveillance of emerging viruses, ultimately contributing to enhanced global public health preparedness.

## References:

1. Yeh, Y.-T.; Gulino, K.; Zhang, Y.; Sabestien, A.; Chou, T.-W.; Zhou, B.; Lin, Z.; Albert, I.; Lu, H.; Swaminathan, V.; Ghedin, E.; Terrones, M., A rapid and label-free platform for virus capture and identification from clinical samples. *Proceedings of the National Academy of Sciences* **2020**, *117* (2), 895-901.
2. Ye, J.; Yeh, Y.-T.; Xue, Y.; Wang, Z.; Zhang, N.; Liu, H.; Zhang, K.; Ricker, R.; Yu, Z.; Roder, A.; Perea Lopez, N.; Organtini, L.; Greene, W.; Hafenstein, S.; Lu, H.; Ghedin, E.; Terrones, M.; Huang, S.; Huang, S. X., Accurate virus

identification with interpretable Raman signatures by machine learning. *Proceedings of the National Academy of Sciences* **2022**, *119* (23), e2118836119.

3. Jin, P.; Yeh, Y. T.; Ye, J.; Wang, Z.; Xue, Y.; Zhang, N.; Huang, S.; Ghedin, E.; Lu, H.; Schmitt, A.; Huang, S. X.; Terrones, M. In *Strain-Level Identification and Analysis of Avian Coronavirus Using Raman Spectroscopy and Interpretable Machine Learning*, 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), 18-21 April 2023; 2023; pp 1-5.
4. Zhang, K.; Wang, Z.; Liu, H.; Perea-López, N.; Ranasinghe, J. C.; Bepete, G.; Minns, A. M.; Rossi, R. M.; Lindner, S. E.; Huang, S. X.; Terrones, M.; Huang, S., Understanding the Excitation Wavelength Dependence and Thermal Stability of the SARS-CoV-2 Receptor-Binding Domain Using Surface-Enhanced Raman Scattering and Machine Learning. *ACS Photonics* **2022**, *9* (9), 2963-2972.