

# The Supporting Information contains Supplementary Figures and Tables

## **High-speed diagnosis of bacterial pathogens at single cell level by Raman microspectroscopy with machine learning filters and denoising autoencoders**

Jiabao Xu<sup>1</sup>, Xiaofei Yi<sup>2,3</sup>, Guilan Jin<sup>2,3</sup>, Di Peng<sup>2,3</sup>, Gaoya Fan<sup>2,3</sup>, Xiaogang Xu<sup>4,5</sup>, Xin Chen<sup>4,5</sup>,  
Huabing Yin<sup>6</sup>, Jon M. Cooper<sup>6</sup>, Wei E. Huang<sup>1\*</sup>

<sup>1</sup>Department of Engineering Science, University of Oxford; Oxford, Oxford, OX1 3PJ, U.K.

<sup>2</sup>Shanghai Hesun Biotechnology Co., Ltd; Shanghai, 201802, China.

<sup>3</sup>Shanghai D-band Medical Instrument Co., Ltd; Shanghai, 201802, China.

<sup>4</sup>Institute of Antibiotics, Huashan Hospital, Fudan University; Shanghai, 200040, China.

<sup>5</sup>National Clinical Research Center for Aging and Medicine, Huashan Hospital, Fudan University; Shanghai, 200040, China.

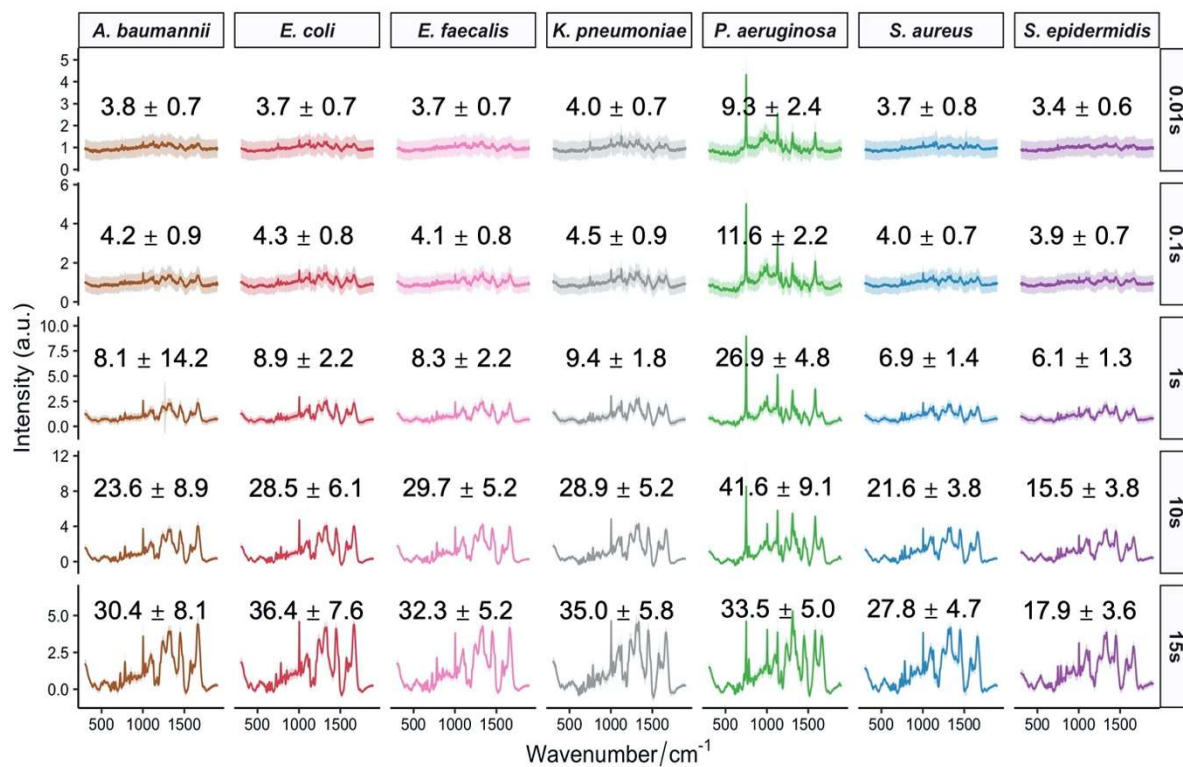
<sup>6</sup>James Watt School of Engineering, University of Glasgow, Glasgow, G12 8LT, U.K.

\*Corresponding author. Wei E. Huang. Email: [wei.huang@eng.ox.ac.uk](mailto:wei.huang@eng.ox.ac.uk)

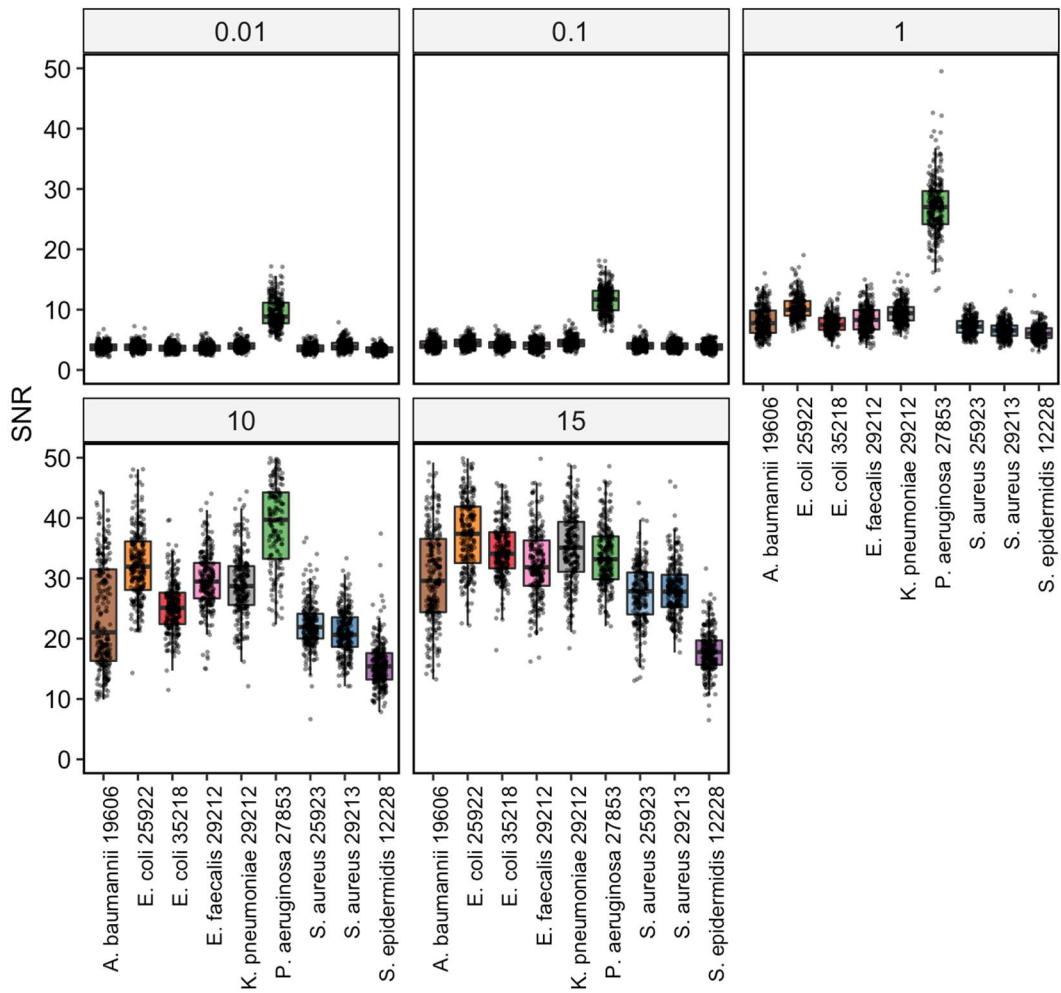
### **This PDF file includes:**

Figures S1 to S3

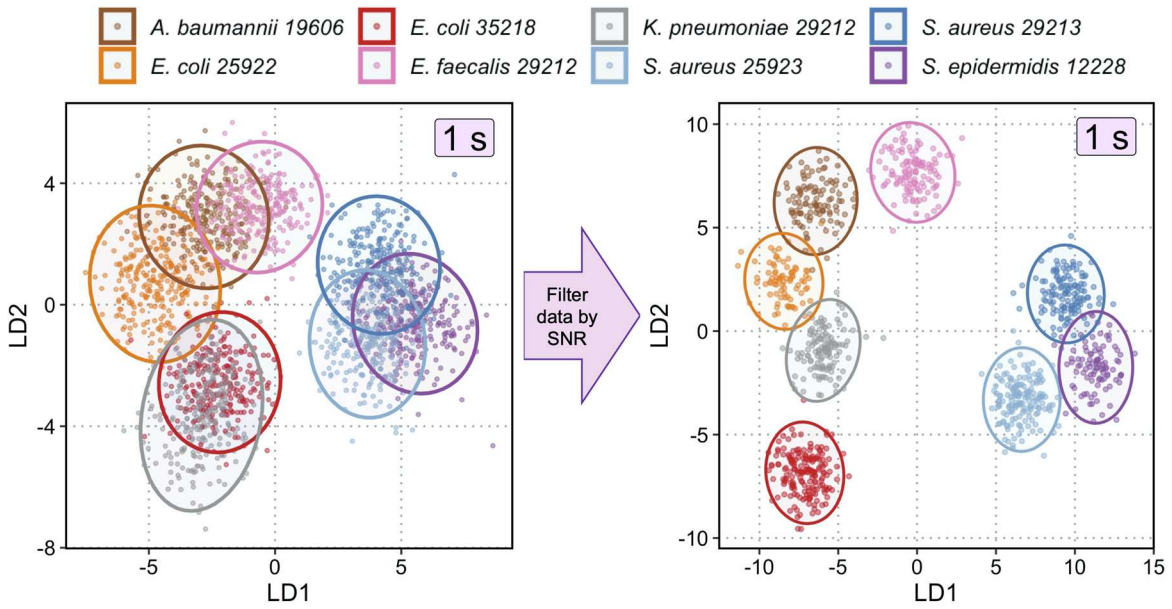
Tables S1



**Figure S1.** Averaged Raman spectra of 7 bacterial species with acquisition time of 0.01, 0.1, 1, 10 or 15 s. The shaded area represents standard deviation from single-cell measurements. The numbers represent averaged SNR and its standard deviation at each measurement condition.



**Figure S2.** Box-whisker plots with data points of SNR in each strain and at each acquisition time to illustrate the distribution of SNRs.



**Figure S3.** LDA clustering shows much better predictive power at the strain level after applying the SNR filter which removes spectra with high (quantile > 0.75) or low SNR (quantile < 0.25) for each strain.

**Table S1.** Averaged SNR of Raman spectra from seven bacterial species at acquisition time of 0.01, 0.1, 1, 10 or 15 seconds.

	<i>Acinetobacter baumannii</i>	<i>Escherichia coli</i>	<i>Enterococcus faecalis</i>	<i>Klebsiella pneumoniae</i>	<i>Pseudomonas aeruginosa</i>	<i>Staphylococcus aureus</i>	<i>Staphylococcus epidermidis</i>
0.01s	3.8 ± 0.7	3.7 ± 0.7	3.6 ± 0.7	4.0 ± 0.7	9.3 ± 2.4	3.7 ± 0.8	3.4 ± 0.6
0.1s	4.2 ± 0.9	4.3 ± 0.8	4.1 ± 0.8	4.5 ± 0.9	11.6 ± 2.2	4.0 ± 0.7	3.9 ± 0.7
1s	8.1 ± 14.2	8.9 ± 2.2	8.3 ± 2.2	9.4 ± 1.8	26.9 ± 4.8	6.9 ± 1.4	6.1 ± 1.3
10s	23.6 ± 8.9	28.5 ± 6.1	29.7 ± 5.2	28.9 ± 5.2	41.6 ± 9.1	21.6 ± 3.8	15.5 ± 3.8
15s	30.4 ± 8.1	36.4 ± 7.6	32.3 ± 5.2	35.0 ± 5.8	33.5 ± 5.0	27.8 ± 4.7	17.9 ± 3.6