Supplementary information

Discrimination of Mycoplasma Infection Using Machine Learning Models Trained on Autofluorescence Signatures of Host Cells

Kenzo Bamba,[†a] Kyosuke Takabe,[†b] Hiroaki Daitoku,[c] Yoshikazu Tanaka,[a] Azusa Ohtani,[a] Midori Ozawa,[a] Akiyoshi Fukamizu,[c] Nobuhiko Nomura,[bd] Arihiro Kohara,[*a] and Tatsuki Kunoh[*b]

[a]JCRB Cell Bank, Center for Drug Discovery Resources Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan.
[b]Faculty of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan.
[c]Center for Tsukuba Advanced Research Alliance, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan.
[d]Microbiology Research Center for Sustainability, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan.

[†]Equal contribution to this study.

Corresponding Authors: Arihiro Kohara (kohara@nibiohn.go.jp); Tatsuki Kunoh: (kunou.tatsuki@aist.go.jp)

Present addresses: Department of Bacteriology I, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo, Japan 162-8640 (KT); Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Higashi, Tsukuba 305-8566, Japan (TK).

Supplementary table

Conditions	Analyzed cell number	
	Mycoplasma (−)	Mycoplasma (+)
Trypsinized	90	166
Adhered		
0 day	50	51
1 day	16	29
2 days	16	29
3 days	36	21
Total	118	130

Table S1. Analyzed cell numbers for machine learning models.
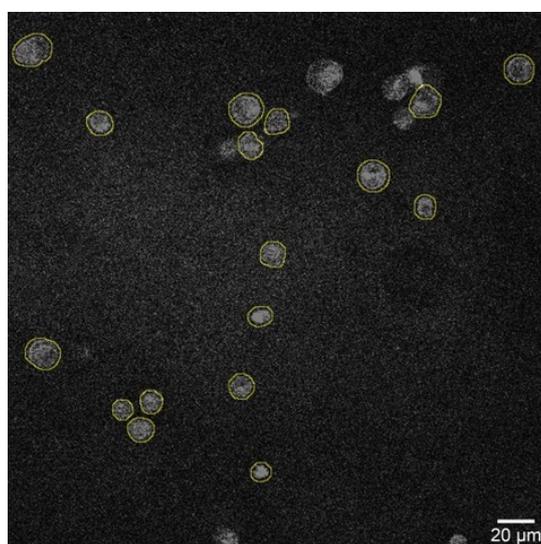
Supplementary figures

Figure S1. Representative image of segmentation showing cell boundaries of individual cells by yellow lines.
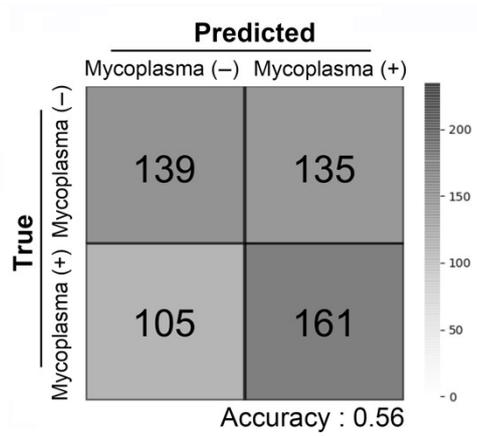
Figure S2. Discrimination of mycoplasma infection of dissociated VERO cells using an SVM-supported machine learning model.
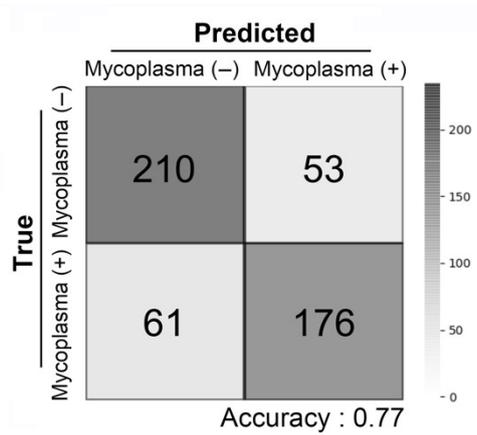

Figure S3. Discrimination of mycoplasma infection of adhered VERO cells using an SVM-supported machine learning model.
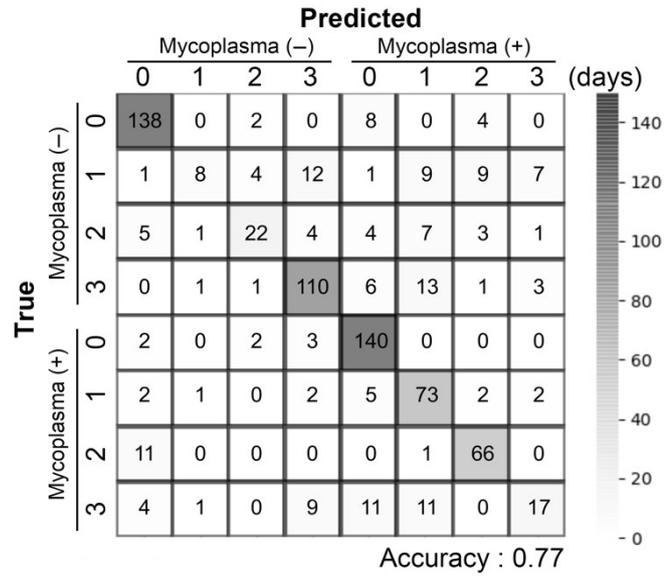
Predicted

		Mycoplasma (−)				Mycoplasma (+)				
		0	1	2	3	0	1	2	3	(days)
Mycoplasma (−) 0		138	0	2	0	8	0	4	0	
1		1	8	4	12	1	9	9	7	
2		5	1	22	4	4	7	3	1	
3		0	1	1	110	6	13	1	3	
Mycoplasma (+) 0		2	0	2	3	140	0	0	0	
1		2	1	0	2	5	73	2	2	
2		11	0	0	0	0	1	66	0	
3		4	1	0	9	11	11	0	17	

(True on vertical axis)

Accuracy : 0.77

Figure S4. Classifying cells across different culture stages. Our neural network–supported machine learning model achieved an accuracy of ~77%.
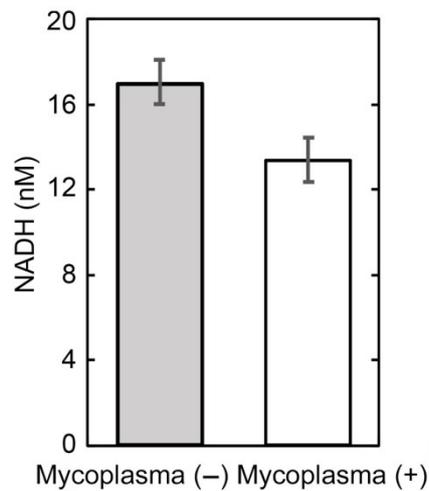

Figure S5. Attenuation of cellular NADH levels in host cells by mycoplasma infection. Cellular NADH levels are lower in mycoplasma (+) cells compared to mycoplasma (−) cells ($p <$ 0.015).