Supplementary Information

High-Precision Rapid Testing of Omicron SARS-CoV-2 Variants in Clinical Samples Using AI-Nanopore

Kaoru Murakami[1,2], Shimpei Kubota[1,2], Kumiko Tanaka[1], Keiichiroh Akabane[1], Rigel Suzuki[3], Yuta Shinohara[1], Hiroyasu Takei[4], Shigeru Hashimoto[1], Yuki Tanaka[2], Shintaro Hojyo[1,2], Osamu Sakamoto[4], Norihiko Naono[4], Takayui Takaai[5], Kazuki Sato[1,6], Yuichi Kojima[1,6], Toshiyuki Harada[7], Takeshi Hattori[8], Satoshi Fuke[9], Isao Yokota[10], Satoshi Konno[6], Takashi Washio[5], Takasuke Fukuhara[3], Takanori Teshima[11], Masateru Taniguchi[5*], and Masaaki Murakami[1,2,12,13*]

The Supplementary Information includes:

Table S1 Number of waveforms obtained from six variants and negative control.

Virus	Number of nanopores in the measurements	Number of extracted waveforms
Wuhan	2	8308
α	5	3028
β	3	3274
γ	6	2967
δ	6	3174
o	3	11385
NTC	11	3243

Pulses were extracted automatically using Aipore-ONE^{TM} extraction software.

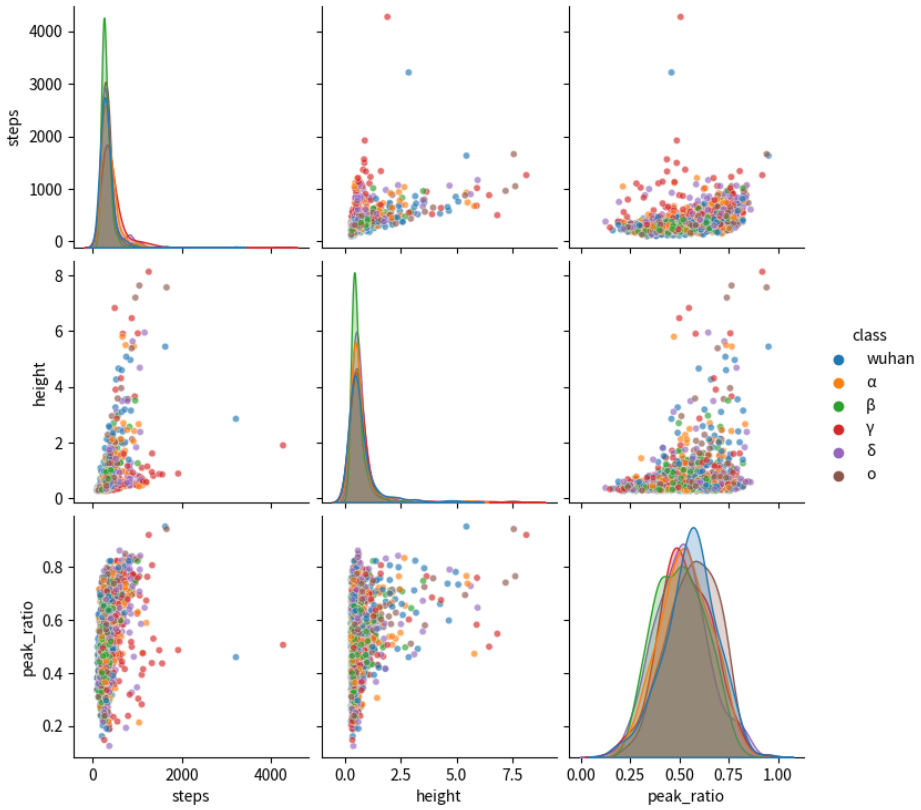


Figure S1 Histograms of height (I_p), steps (t_d), and peak position ratios for six cultured coronaviruses.

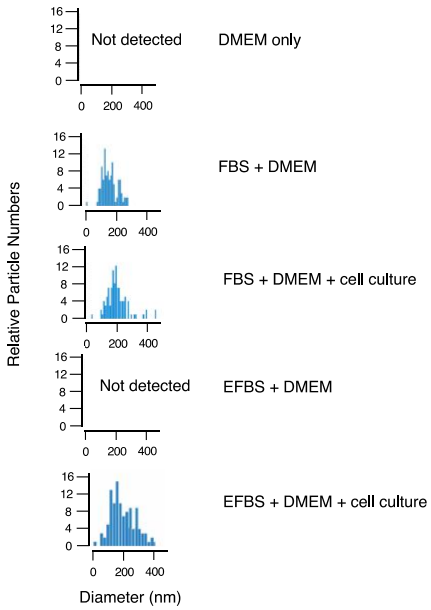


Figure S2 Interference light microscopy (VideoDrop) analysis of control medium. Nanoparticles were detected in medium plus FBS (FBS + DMEM) but not in medium alone (DMEM only). A 48-h cell culture increased the number of bigger nanoparticles in FBS + DMEM + cell culture. No nanoparticles in medium plus exosome free-FBS (EFBS + DMEM) were detected, but the number of bigger nanoparticles appeared after a 48-h cell culture (EFBS + DMEM + cell culture).

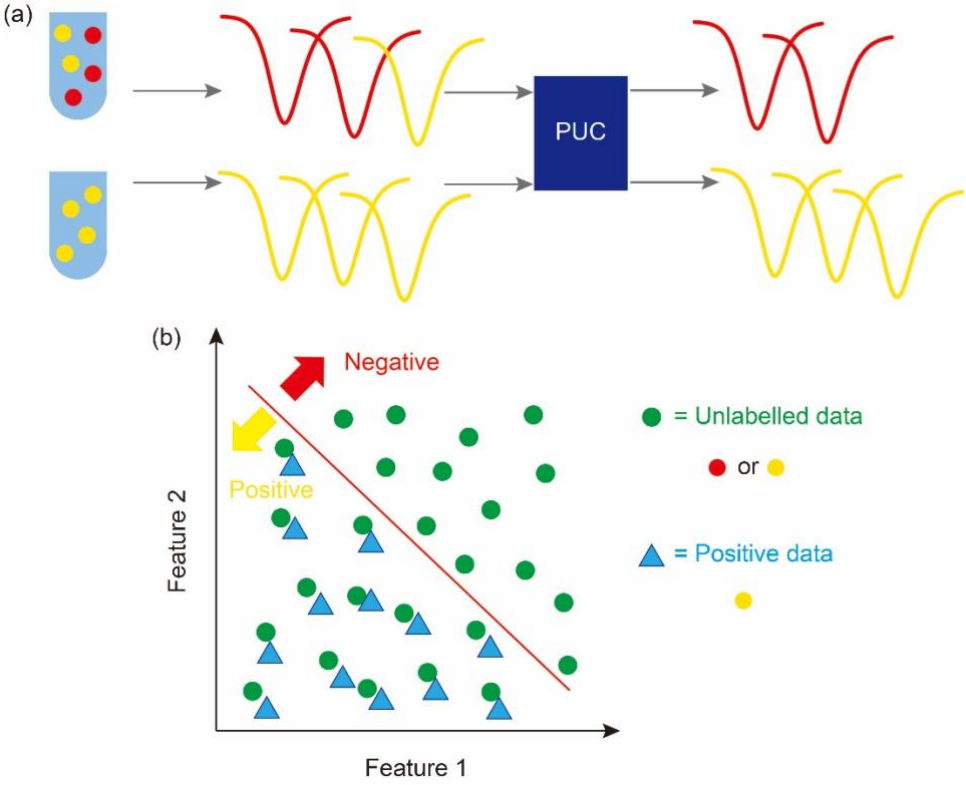


Figure S3 Positive unlabelled classifier. (a) In a solution of cultured virus, virus (●) and exosomes (●) are present. In a culture supernatant, only exosomes (●) are present. The PUC extracts the viral waveform from the mixed viral and exosomal waveform. (b) A conceptual diagram of the PUC in a feature space. Exosome waveforms are defined as positive, and virus waveforms as negative. Waveforms obtained from cultured virus measurements are either viral or exosomal and therefore unlabelled. The data above and below the red line indicate viral and exosomal waveforms, respectively.

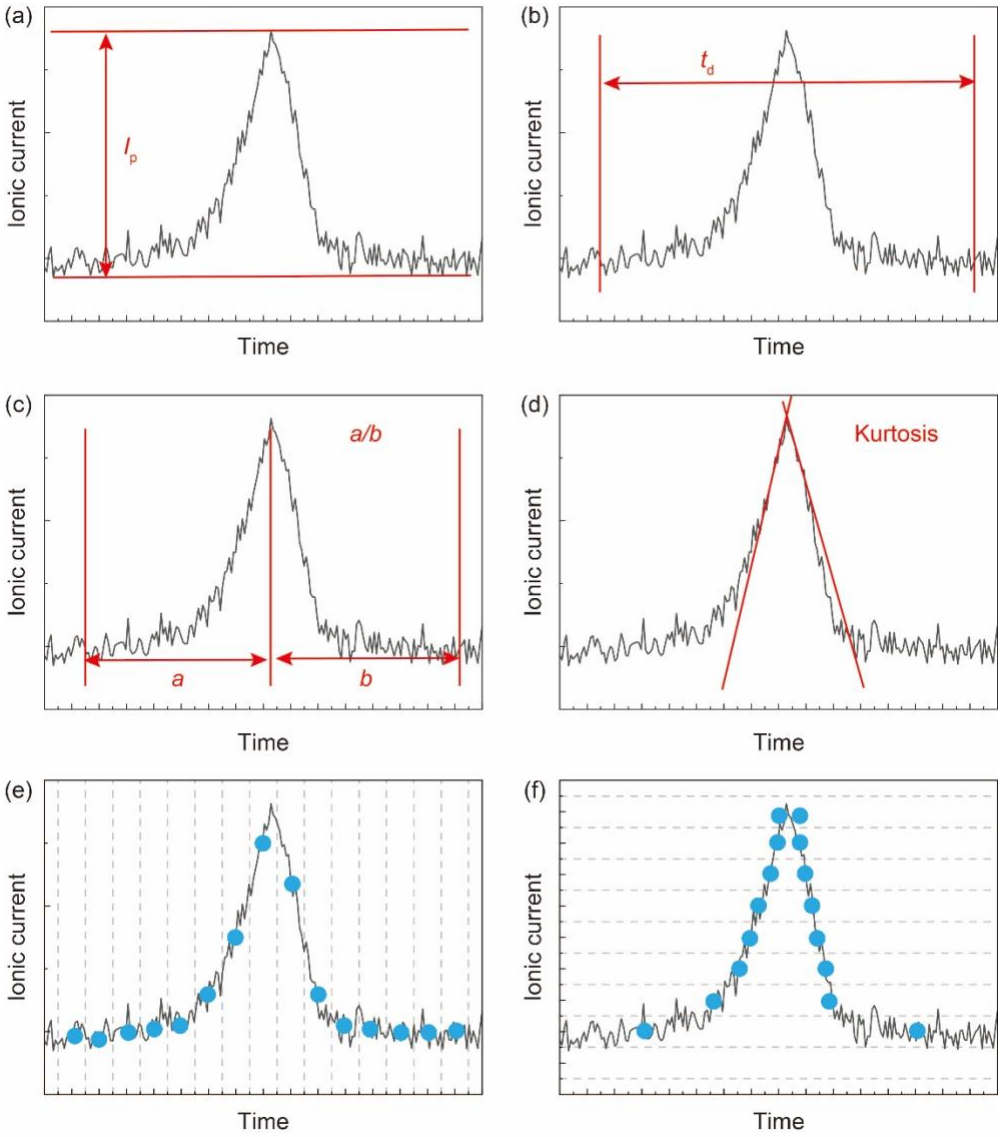


Figure S4 Features used in machine learning. The (a) maximum current value and (b) current duration were defined as I_P and t_d, respectively. The (c) symmetry and (d) kurtosis of one ionic current-time waveform. The vector of one ionic current-time waveform that was divided into 10 equal parts in the (e) time and (f) current directions. These features and their combination patterns were used in the machine learning.

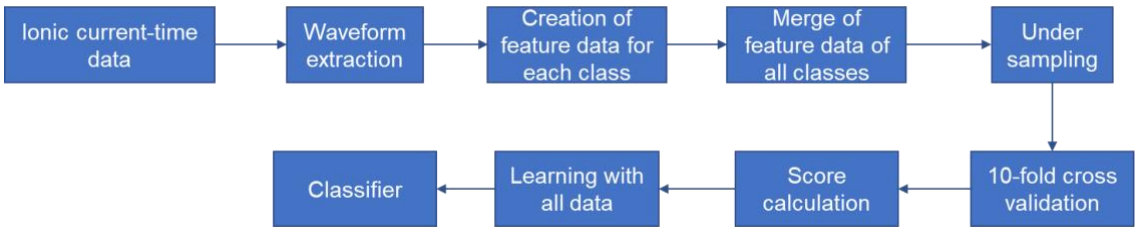


Figure S5 Machine learning algorithm implemented on cultured coronaviruses. Each waveform was automatically extracted from the measured data using the waveform extraction software installed in Aipore-ONE^{TM}. The features of each cultured coronavirus were generated. The features of the six cultured viruses were merged. The number of waveforms for each coronavirus was adjusted to the minimum number of waveforms to avoid overfitting due to the nonuniform data volume. This method is an under-sampling method. Subsequently, the waveform data were randomly divided into 10 parts: 9 data groups were used for learning and 1 data group for testing. A 10-fold iteration was implemented, and the average score of the *F*-measure was calculated. To improve the accuracy, a machine learning procedure was performed using all the under-sampled waveforms. The classifier that afforded the highest *F*-measure was considered the most effective in identifying the cultured coronaviruses.

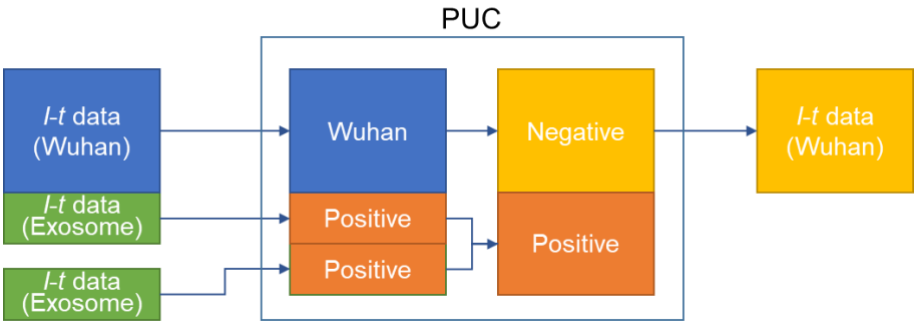


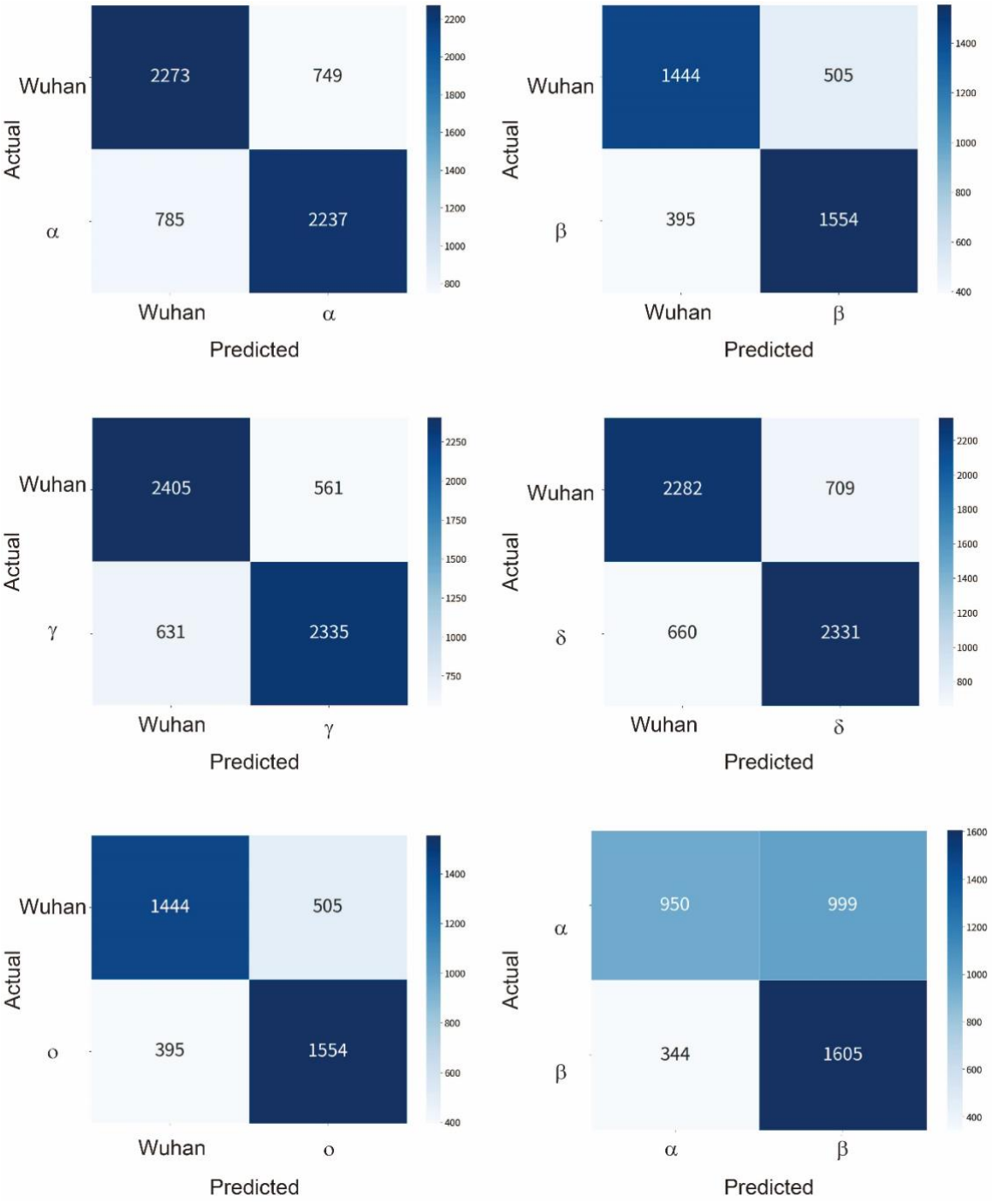
Figure S6 Algorithm of PUC. The cultured Wuhan coronavirus solutions contained two types of particles: coronaviruses and exosomes. The culture medium contained exosomes only. The waveform of the exosomes was learned by learning the ionic current–time waveform of the culture medium. By learning the ionic current–time waveform of the cultured Wuhan coronavirus with PUC, the waveform was classified into the virus waveform and the exosome waveform.

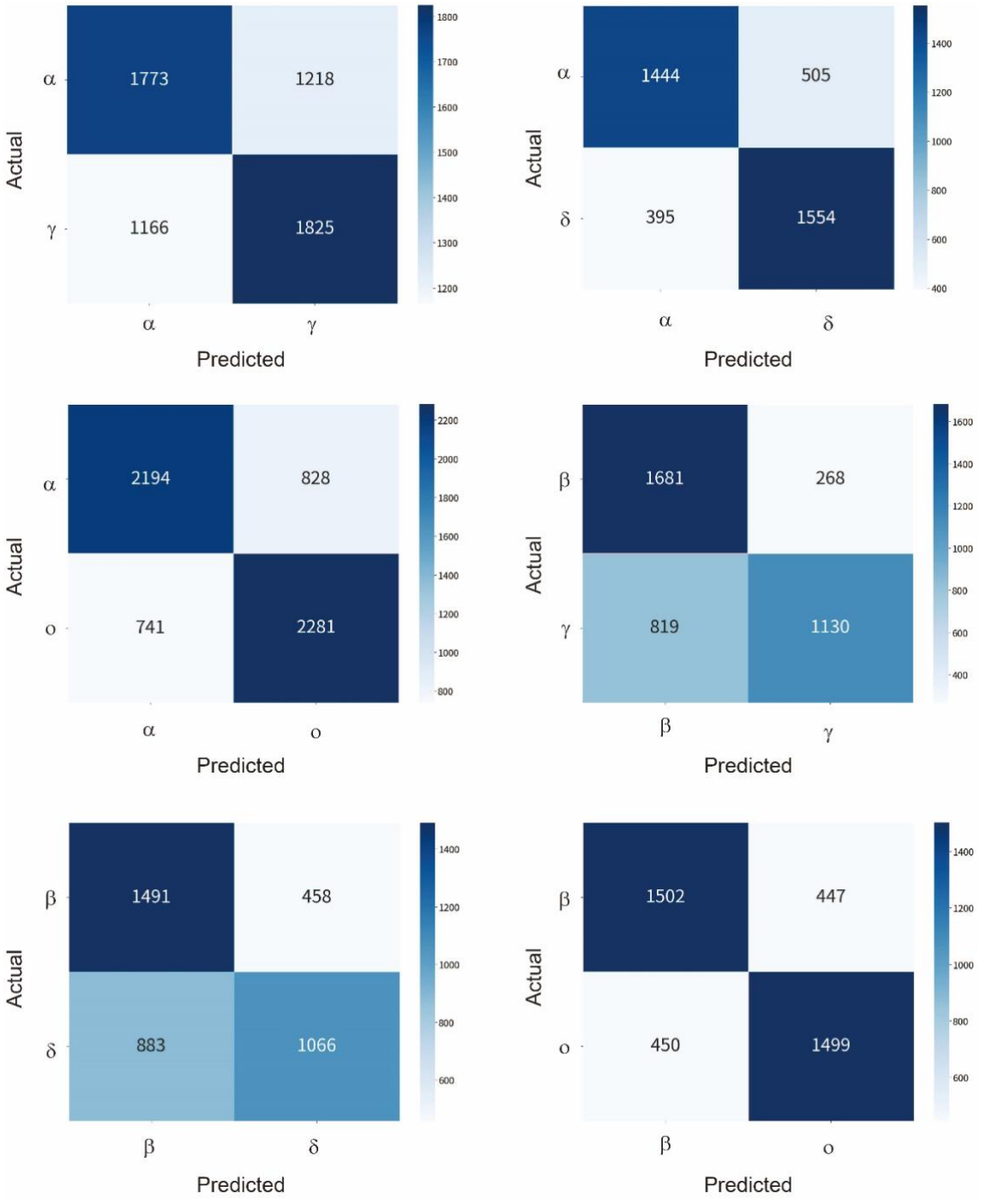
Table S2 Classifiers used in machine learning.

No.	Classifier
1	naive_bayes.BernoulliNB
2	tree.DecisionTreeClassifier
3	tree.ExtraTreeClassifier
4	ensemble.ExtraTreesClassifier
5	naive_bayes.GaussianNB
6	neighbors.KNeighborsClassifier
7	semi_supervised.LabelPropagation
8	semi_supervised.LabelSpreading
9	discriminant_analysis.LinearDiscriminantAnalysis
10	discriminant_analysis.QuadraticDiscriminantAnalysis
11	svm.LinearSVC (setting multi_class=”crammer_singer”)
12	svm.LinearSVC (setting multi_class=”ovr”)
13	linear_model.LogisticRegression (setting multi_class=”multinomial”)
14	linear_model.LogisticRegressionCV (setting multi_class=”multinomial”)
15	linear_model.LogisticRegression (setting multi_class=”ovr”)
16	linear_model.LogisticRegressionCV (setting multi_class=”ovr”)
17	neural_network.MLPClassifier
18	neighbors.NearestCentroid
19	neighbors.RadiusNeighborsClassifier
20	ensemble.RandomForestClassifier
21	linear_model.RidgeClassifier
22	linear_model.RidgeClassifierCV
23	svm.NuSVC
24	svm.SVC.
25	gaussian_process.GaussianProcessClassifier (setting multi_class = “one_vs_one”)
26	gaussian_process.GaussianProcessClassifier (setting multi_class = “one_vs_rest”)
27	ensemble.GradientBoostingClassifier
28	linear_model.SGDClassifier
29	linear_model.Perceptron
30	linear_model.PassiveAggressiveClassifier
31	Light Gradient Boosting Machine
32	eXtreme Gradient Boosting

Table S3 Top5 classifiers that contribute to the high *F*-measures.

Rank	Classifier	*F*-measure
1	neural_network.MLP	0.3489
2	eXtreme Gradient Boosting	0.3334
3	Light Gradient Boosting Machine	0.3333
4	ensemble.RandomForest	0.3300
5	svm.SVC (multi class)	0.3222


Wuhan
α
Actual
2273
749
785
2237
Wuhan
α
Predicted
2200
2000
1800
1600
1400
1200
1000
800
Wuhan
β
Actual
1444
505
395
1554
Wuhan
β
Predicted
1400
1200
1000
800
600
400
Wuhan
γ
Actual
2405
561
631
2335
Wuhan
γ
Predicted
2250
2000
1750
1500
1250
1000
750
Wuhan
δ
Actual
2282
709
660
2331
Wuhan
δ
Predicted
2200
2000
1800
1600
1400
1200
1000
800
Wuhan
ο
Actual
1444
505
395
1554
Wuhan
ο
Predicted
1400
1200
1000
800
600
400
α
Actual
950
999
344
1605
α
β
Predicted
1600
1400
1200
1000
800
600
400


α
γ
Actual
1773
1218
1166
1825
α
γ
Predicted
1800
1700
1600
1500
1400
1300
1200
α
δ
Actual
1444
505
395
1554
α
δ
Predicted
1400
1200
1000
800
600
400
α
o
Actual
2194
828
741
2281
α
o
Predicted
2200
2000
1800
1600
1400
1200
1000
800
β
γ
Actual
1681
268
819
1130
β
γ
Predicted
1600
1400
1200
1000
800
600
400
β
δ
Actual
1491
458
883
1066
β
δ
Predicted
1400
1200
1000
800
600
β
o
Actual
1502
447
450
1499
β
o
Predicted
1400
1200
1000
800
600

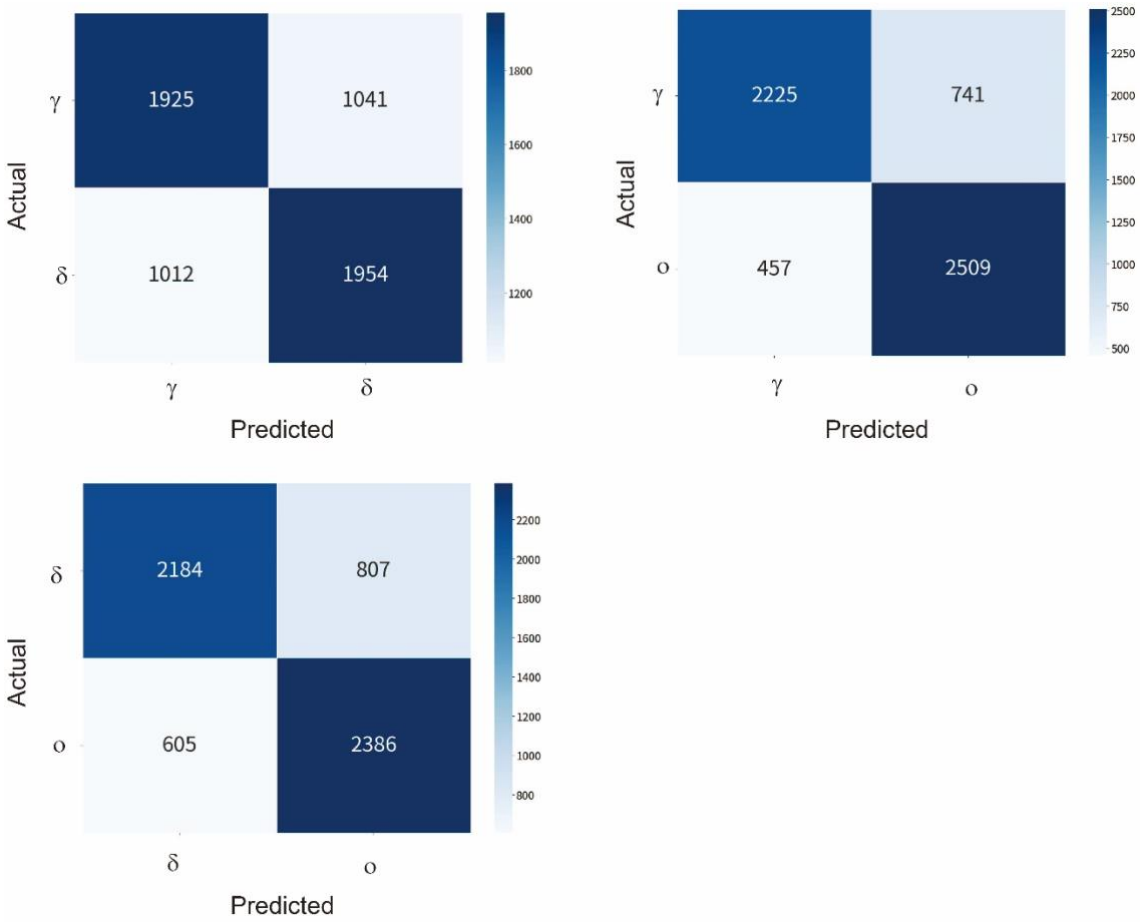
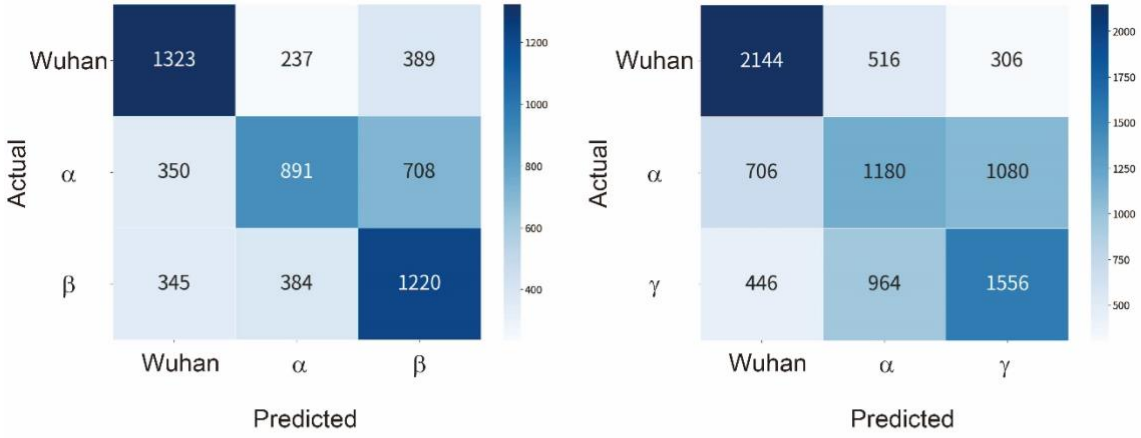
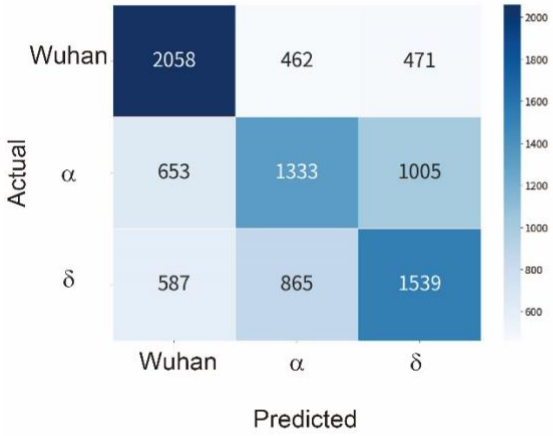
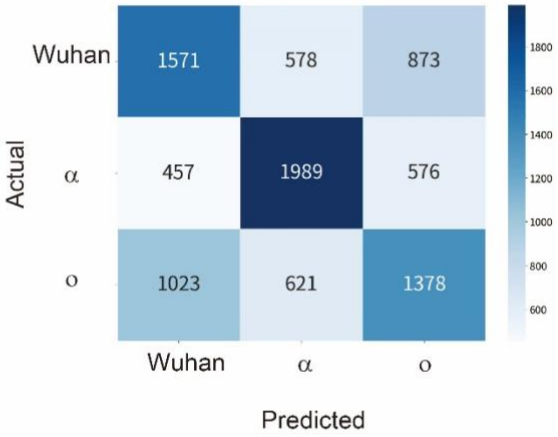


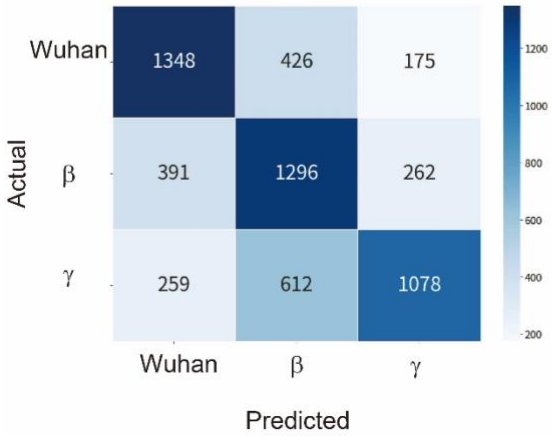
Figure S7 Confusion matrixes for comparisons of two types of cultured SARS-CoV-2 variants. The number in each matrix element indicates the number of waveforms that were obtained by the measurements. The darker the color, the higher the identification accuracy. The *F*-measure was calculated from the confusion matrixes.

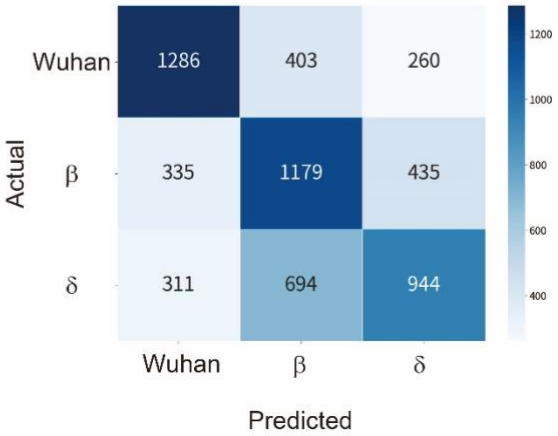

Wuhan
α
β
Actual
1323
237
389
350
891
708
345
384
1220
Wuhan
α
β
Predicted
1200
1000
800
600
400

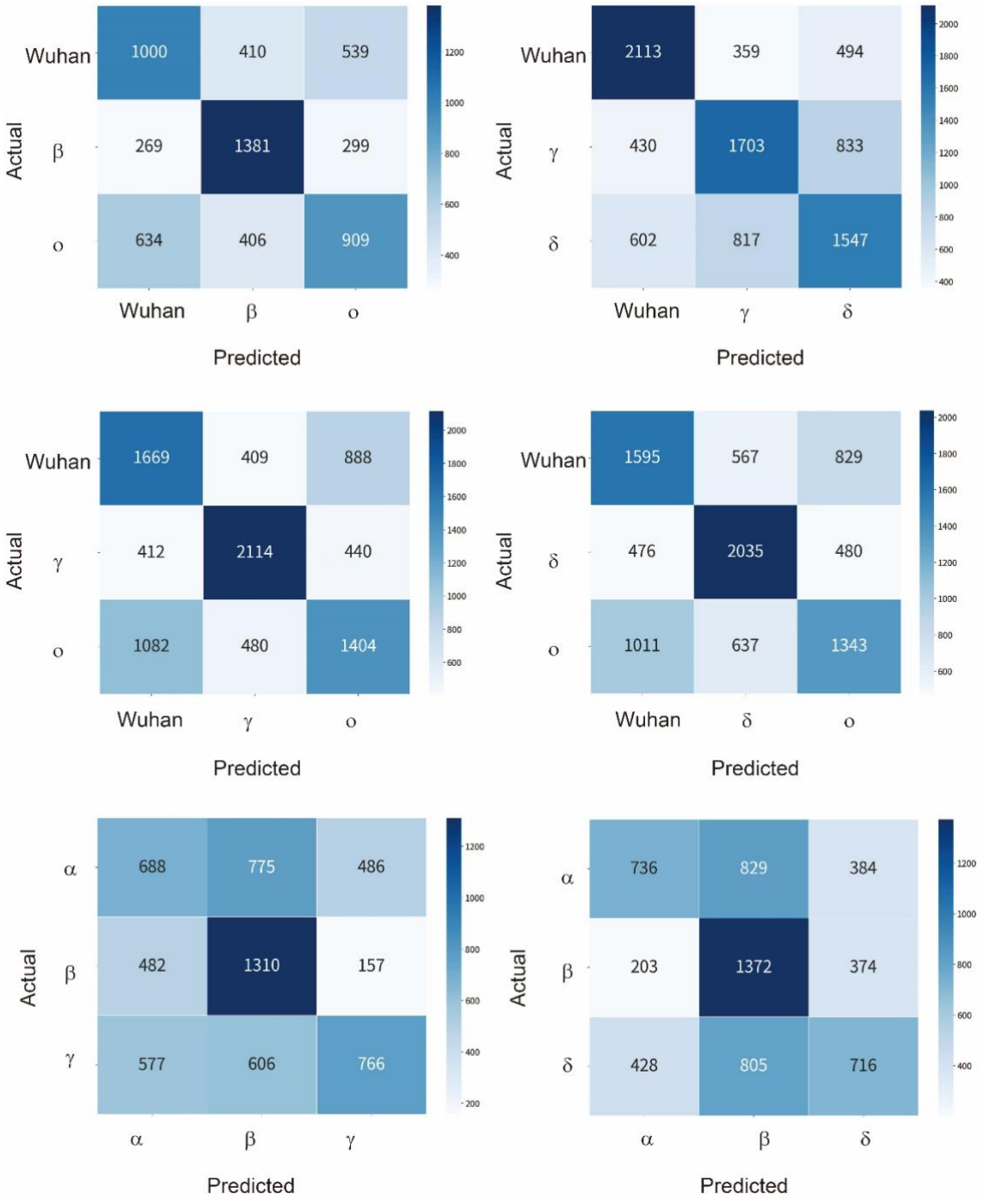
Wuhan
α
γ
Actual
2144
516
306
706
1180
1080
446
964
1556
Wuhan
α
γ
Predicted
2000
1750
1500
1250
1000
750
500

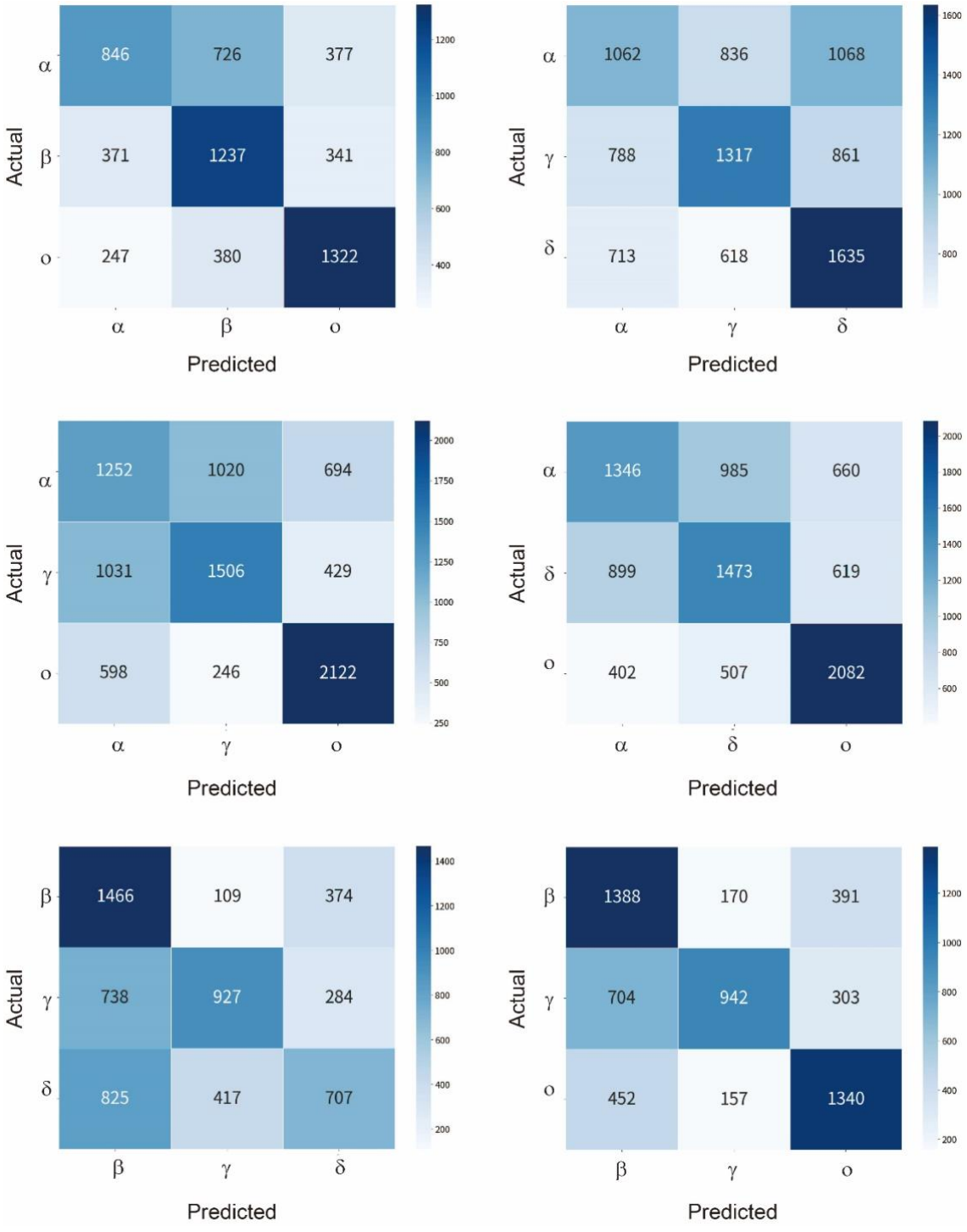

Wuhan
α
δ
Actual
2058
462
471
653
1333
1005
587
865
1539
Wuhan
α
δ
Predicted
2000
1800
1600
1400
1200
1000
800
600


Wuhan
α
o
Actual
1571
578
873
457
1989
576
1023
621
1378
Wuhan
α
o
Predicted
1800
1600
1400
1200
1000
800
600


Wuhan
β
γ
Actual
1348
426
175
391
1296
262
259
612
1078
Wuhan
β
γ
Predicted
1200
1000
800
600
400
200


Wuhan
β
δ
Actual
1286
403
260
335
1179
435
311
694
944
Wuhan
β
δ
Predicted
1200
1000
800
600
400


Wuhan
β
ο
Actual
1000
410
539
269
1381
299
634
406
909
Wuhan
β
ο
Predicted
Wuhan
γ
δ
Actual
2113
359
494
430
1703
833
602
817
1547
Wuhan
γ
δ
Predicted
Wuhan
γ
ο
Actual
1669
409
888
412
2114
440
1082
480
1404
Wuhan
γ
ο
Predicted
Wuhan
δ
ο
Actual
1595
567
829
476
2035
480
1011
637
1343
Wuhan
δ
ο
Predicted
α
β
γ
Actual
688
775
486
482
1310
157
577
606
766
α
β
γ
Predicted
α
β
δ
Actual
736
829
384
203
1372
374
428
805
716
α
β
δ
Predicted


Actual
Predicted
α
β
ο
846
726
377
371
1237
341
247
380
1322
1200
1000
800
600
400
Actual
Predicted
α
γ
δ
1062
836
1068
788
1317
861
713
618
1635
1600
1400
1200
1000
800
Actual
Predicted
α
γ
ο
1252
1020
694
1031
1506
429
598
246
2122
2000
1750
1500
1250
1000
750
500
250
Actual
Predicted
α
δ
ο
1346
985
660
899
1473
619
402
507
2082
2000
1800
1600
1400
1200
1000
800
600
Actual
Predicted
β
γ
δ
1466
109
374
738
927
284
825
417
707
1400
1200
1000
800
600
400
200
Actual
Predicted
β
γ
ο
1388
170
391
704
942
303
452
157
1340
1200
1000
800
600
400
200

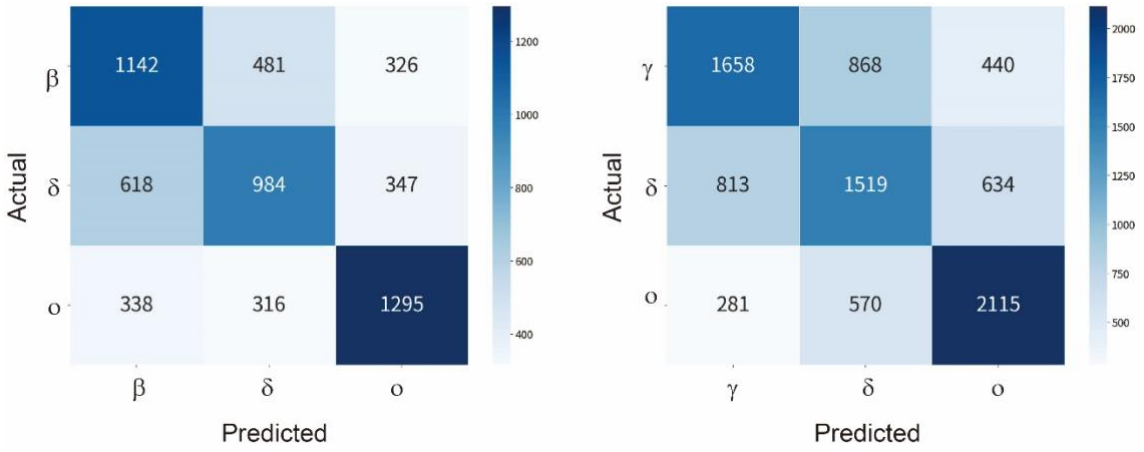
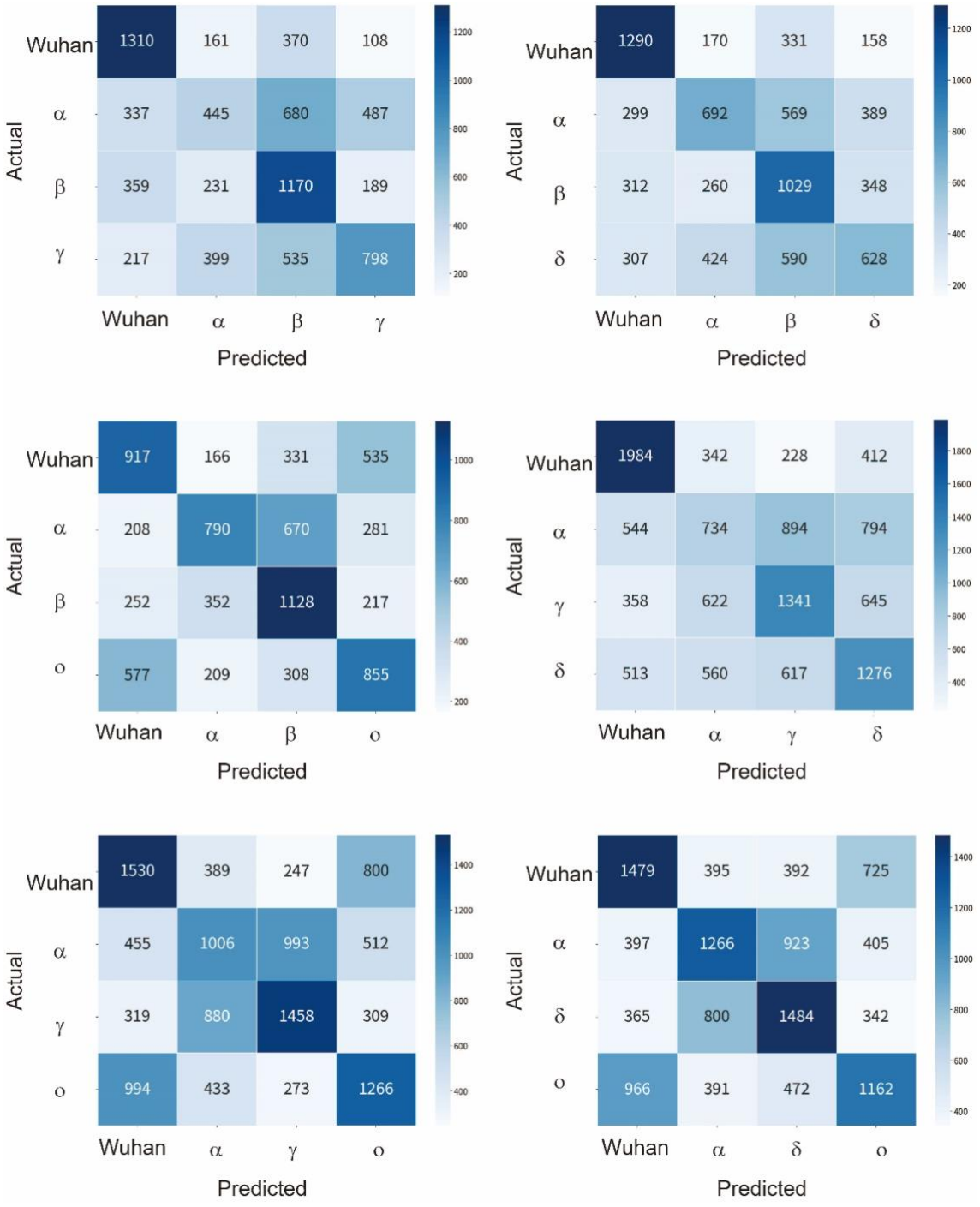
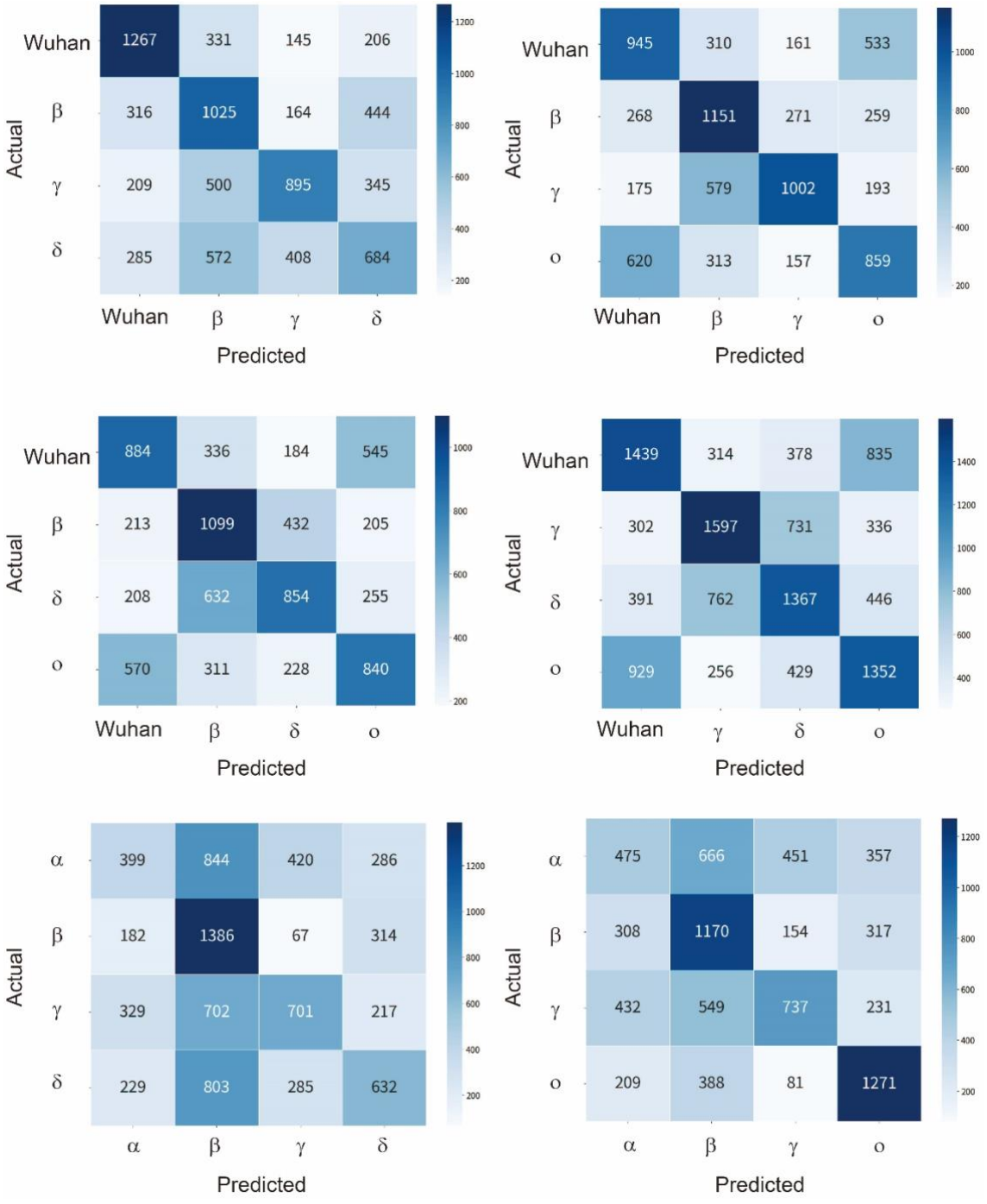


Figure S8 Confusion matrixes for comparisons of three types of cultured SARS-CoV-2 variants. The number in each matrix element indicates the number of waveforms that were obtained by the measurements. The darker the color, the higher the identification accuracy. The *F*-measure was calculated from the confusion matrixes.


Wuhan
α
β
γ
δ
ο
Actual
Predicted
1310 161 370 108
337 445 680 487
359 231 1170 189
217 399 535 798
1290 170 331 158
299 692 569 389
312 260 1029 348
307 424 590 628
917 166 331 535
208 790 670 281
252 352 1128 217
577 209 308 855
1984 342 228 412
544 734 894 794
358 622 1341 645
513 560 617 1276
1530 389 247 800
455 1006 993 512
319 880 1458 309
994 433 273 1266
1479 395 392 725
397 1266 923 405
365 800 1484 342
966 391 472 1162


Wuhan
β
γ
δ
Actual
Predicted
1267 331 145 206
316 1025 164 444
209 500 895 345
285 572 408 684
945 310 161 533
268 1151 271 259
175 579 1002 193
620 313 157 859
884 336 184 545
213 1099 432 205
208 632 854 255
570 311 228 840
1439 314 378 835
302 1597 731 336
391 762 1367 446
929 256 429 1352
α
399 844 420 286
182 1386 67 314
329 702 701 217
229 803 285 632
475 666 451 357
308 1170 154 317
432 549 737 231
209 388 81 1271

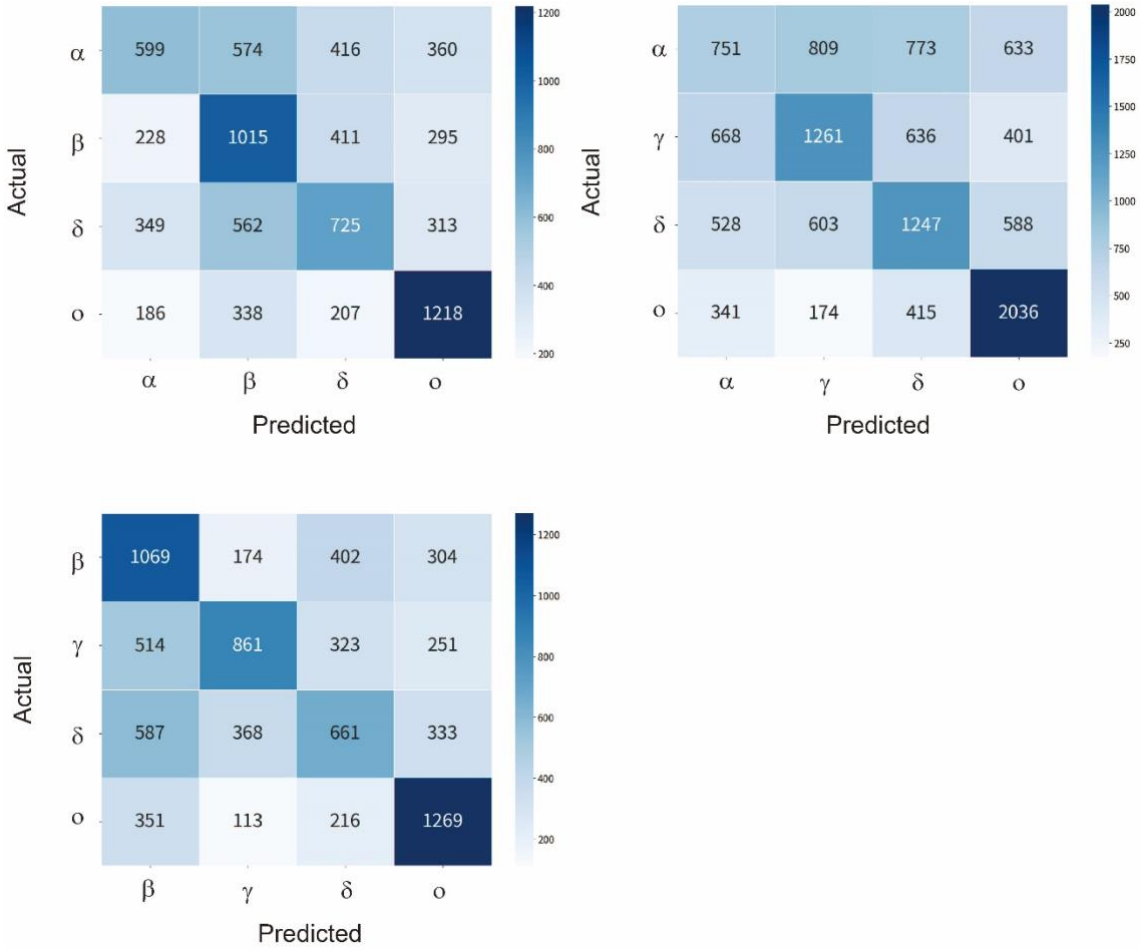


Figure S9 Confusion matrixes for comparisons of four types of cultured SARS-CoV-2 variants. The number in each matrix element indicates the number of waveforms that were obtained by the measurements. The darker the color, the higher the identification accuracy. The *F*-measure was calculated from the confusion matrixes.

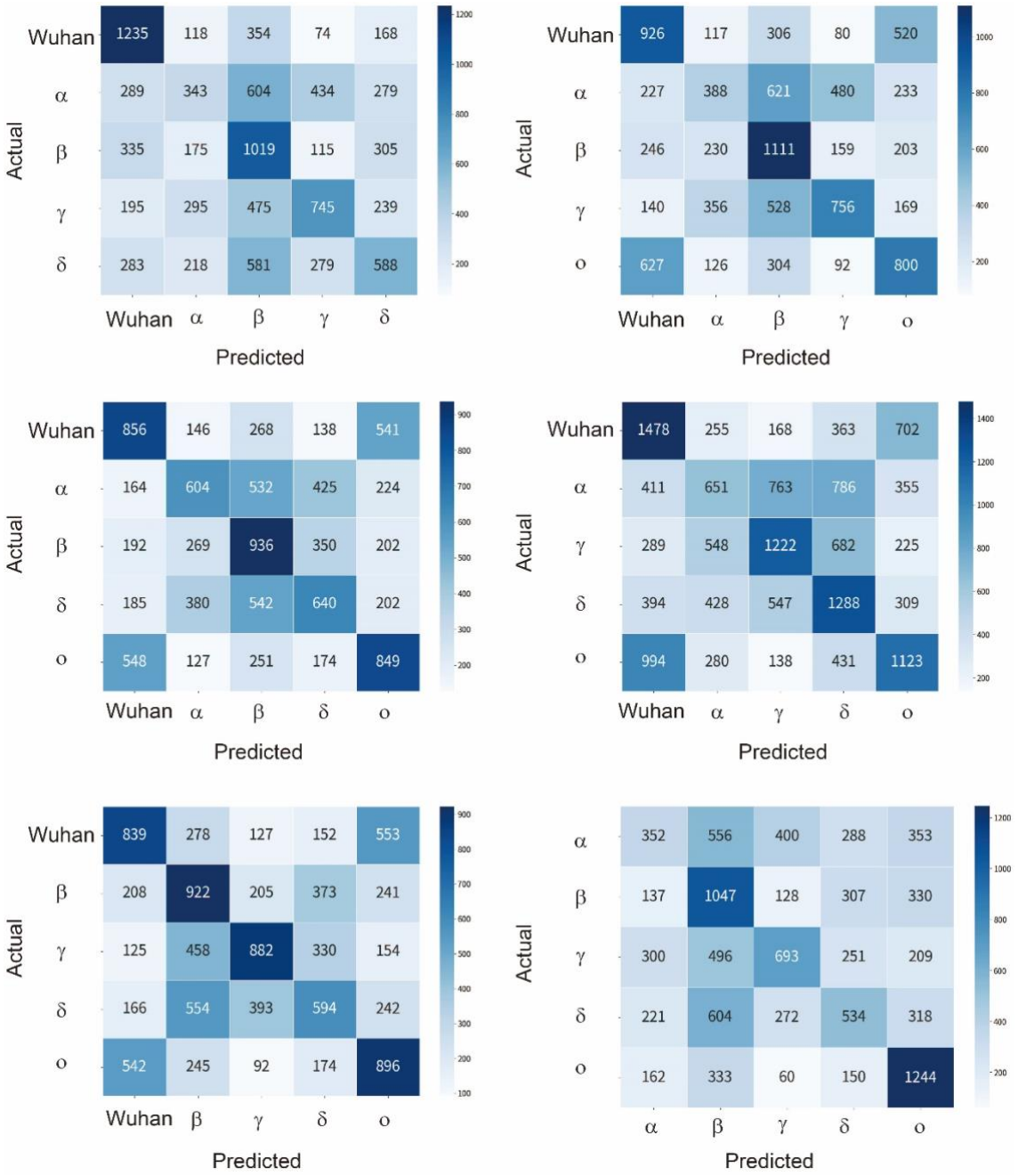


Figure S10 Confusion matrixes for comparisons of five types of cultured SARS-CoV-2 variants. The number in each matrix element indicates the number of waveforms that were obtained by the measurements. The darker the color, the higher the identification accuracy. The *F*-measure was calculated from the confusion matrixes.

Table S4 Number of waveforms obtained from three coronavirus samples and used in machine learning.

Virus	Number of nanopores for measurements	Number of extracted waveforms
Wuhan	2	8308
Wuhan-δ	3	8216
Wuhan-o	3	9859

Pulses were extracted automatically using Aipore-ONE^{TM} extraction software.

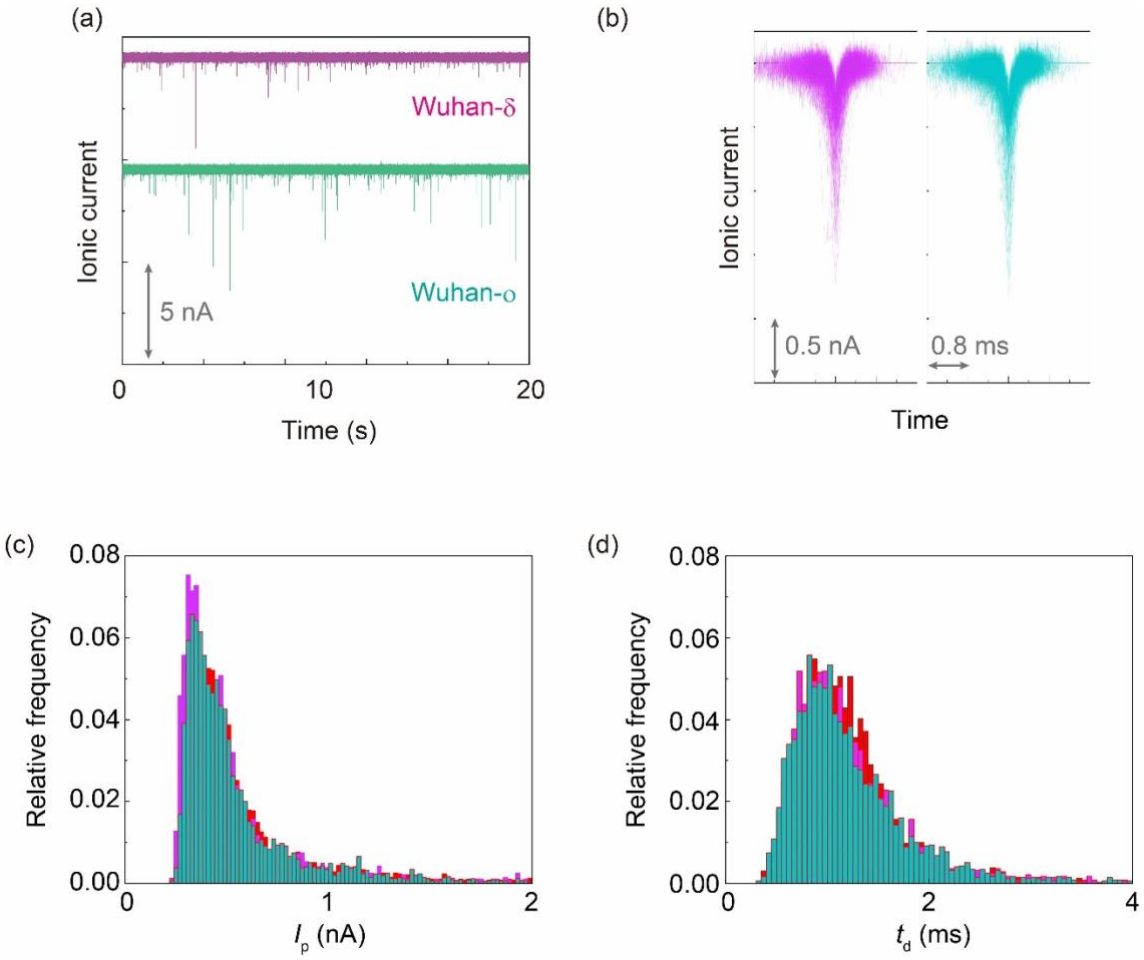


Figure S11 Ionic current-time profiles obtained for Wuhan-type SARS-CoV-2 with δ- and o-type spike proteins. (a) Ionic current-time profiles of Wuhan-δ and Wuhan-o viruses. (b) One hundred superimposed ionic current-time waveforms of the culture supernatant (Wuhan-δ and Wuhan-o). Histograms of the (c) maximum current value (I_p) and (d) current duration (t_d). Red, purple, and emerald green correspond to Wuhan, Wuhan-δ, and Wuhan-o, respectively.

Table S5 Number of waveforms obtained from clinical specimens and used in machine learning.

Clinical specimen	Number of specimens for measurements	Number of extracted waveforms
PCR positive	132	99980
PCR negative	109	75701

Pulses were extracted automatically using Aipore-ONE^{TM} extraction software.

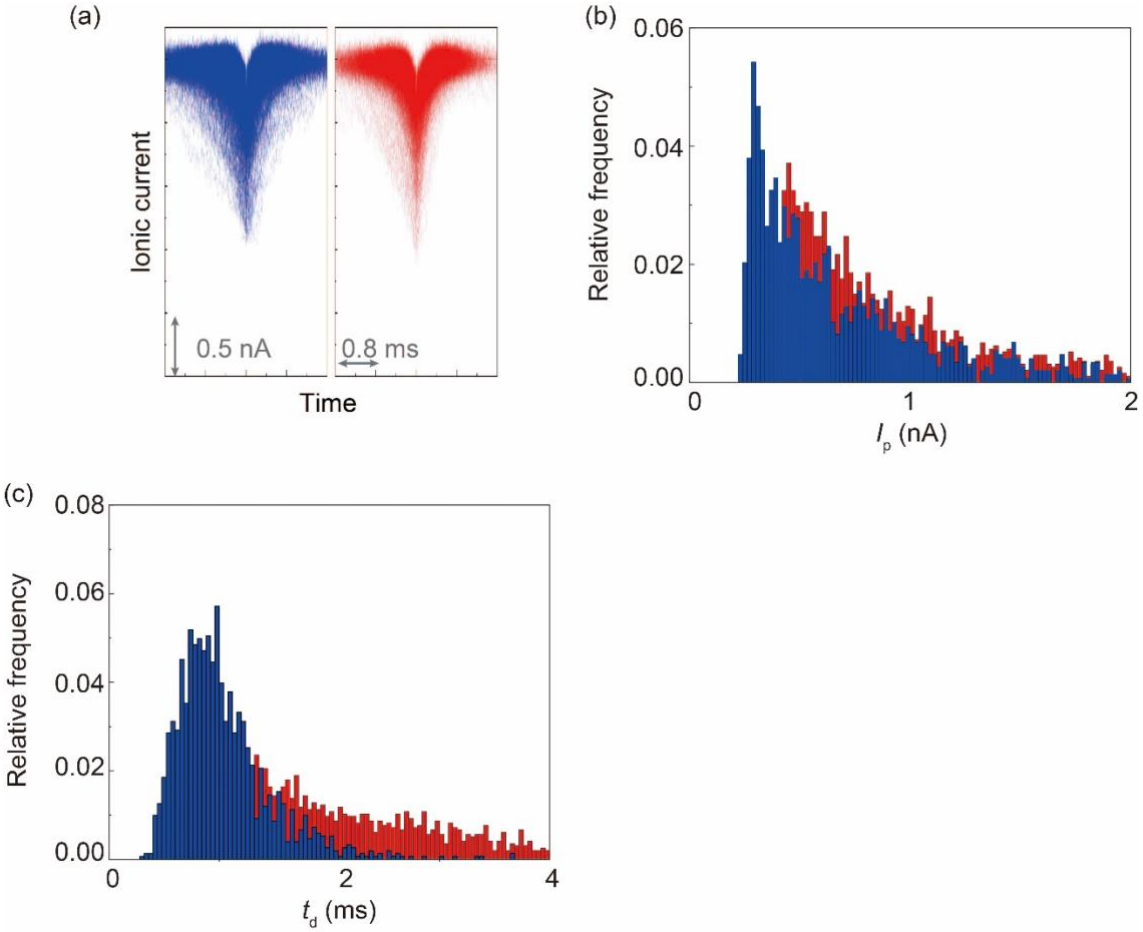


Figure S12 Ionic current-time profiles obtained from clinical specimens. (a) One hundred superimposed ionic current-time waveforms of Wuhan-δ and Wuhan-o. Histograms of the (c) maximum current value (I_p) and (d) current duration (t_d). Red and blue correspond to PCR-positive and PCR-negative specimens, respectively.

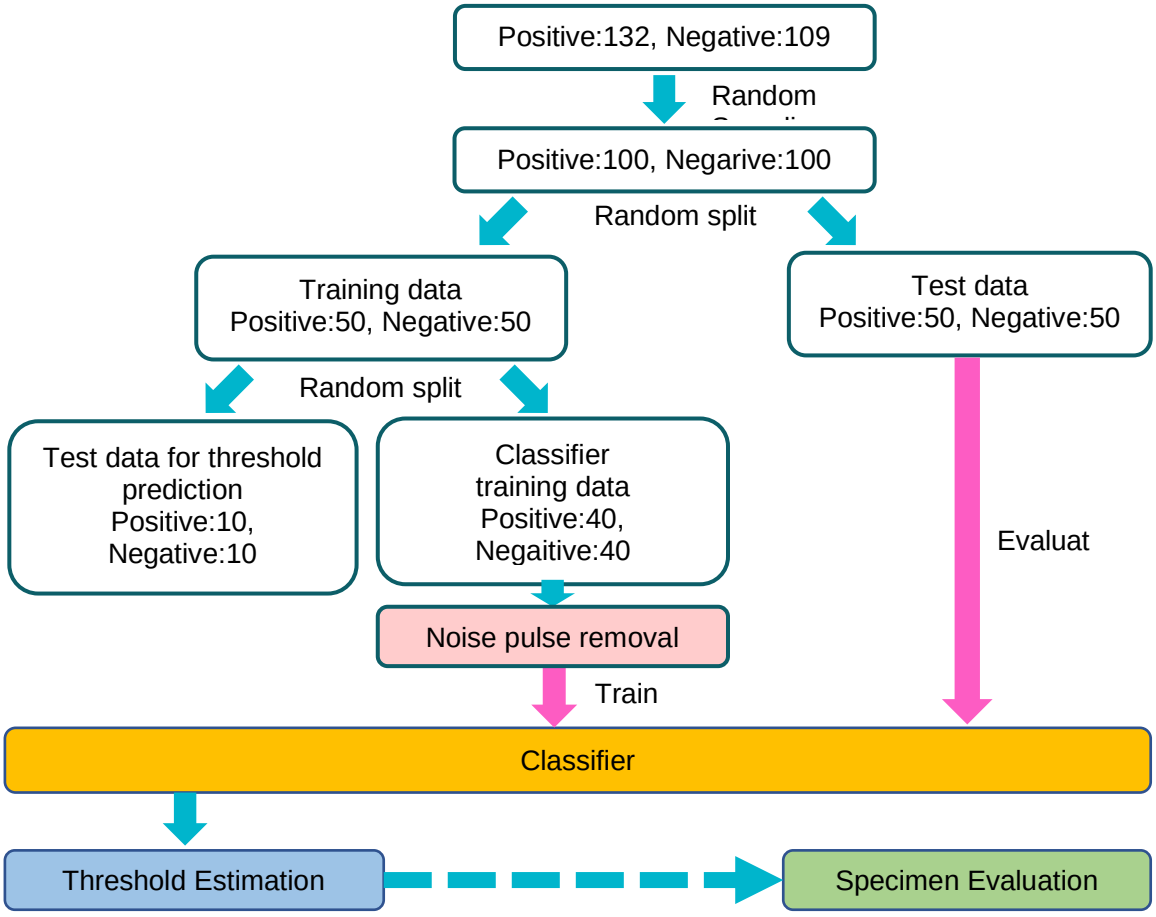


Figure S13 Flowchart of machine learning of clinical specimens.

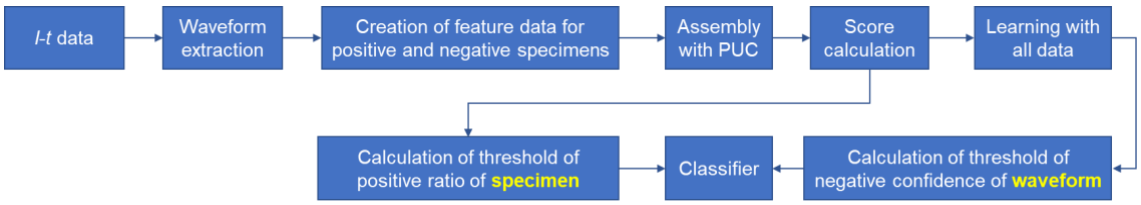


Figure S14 Detailed algorithm of the learning process of clinical specimens. The positive ratio is the ratio of the number of positive waveforms in a specimen unit, and it is defined as the ratio between the number of waveforms judged to be positive in one specimen to the total number of waveforms in the specimen sample. A target specimen is classified as positive for the coronavirus when the positive ratio exceeds a threshold value. The negative confidence threshold is the limiting value used for each waveform, and on this basis, the negative waveform was determined by the following equation:

negative waveform = negative confidence of waveform > negative confidence threshold.

Conversely, the positive waveform was determined by the following equation:

positive waveform = negative confidence of waveform < negative confidence threshold.

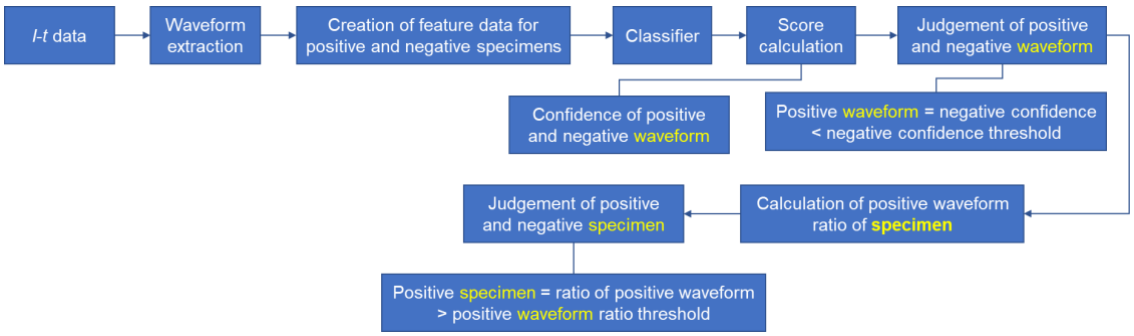


Figure S15 Detailed algorithm of the diagnosis process of clinical specimens. The positive ratio is the ratio of the number of positive waveforms in a specimen unit, and it is defined by (the number of waveforms judged to be positive in one specimen)/(the total number of waveforms in the specimen sample). The threshold value of the positive ratio is employed to classify the target specimen as positive when the positive ratio exceeds this value. The negative confidence threshold is the confidence threshold used for each waveform. For example, the negative waveform was determined by the following logic: negative waveform = negative confidence of waveform > negative confidence threshold. In contrast, the positive waveform was determined by the following logic:
positive waveform = negative confidence of waveform < negative confidence threshold.

Table S6 Top 20 features used in machine learning for clinical specimens.

No.	Feature
1	matrix_profile__feature_"75"__threshold_0.98
2	matrix_profile__feature_"max"__threshold_0.98
3	length
4	fft_coefficient__attr_"abs"__coeff_99
5	change_quantiles__f_agg_"var"__isabs_False__qh_1.0__ql_0.0
6	change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.0
7	friedrich_coefficients__coeff_0__m_3__r_30
8	variation_coefficient
9	agg_linear_trend__attr_"intercept"__chunk_len_50__f_agg_"var"
10	change_quantiles__f_agg_"mean"__isabs_True__qh_1.0__ql_0.2
11	skewness
12	agg_autocorrelation__f_agg_"var"__maxlag_40
13	change_quantiles__f_agg_"var"__isabs_False__qh_1.0__ql_0.2
14	change_quantiles__f_agg_"var"__isabs_True__qh_1.0__ql_0.0
15	change_quantiles__f_agg_"mean"__isabs_True__qh_0.8__ql_0.0
16	fft_coefficient__attr_"abs"__coeff_98
17	ar_coefficient__coeff_1__k_10
18	ar_coefficient__coeff_2__k_10
19	matrix_profile__feature_"25"__threshold_0.98
20	matrix_profile__feature_"min"__threshold_0.98

Table S7 PCR results and Ct values of clinical specimens.

	Learning (L)/Testing (T)	Specimen ID	PCR	C_t
1	T	0314-02+	positive	unknown
2	T	0314-72+	positive	unknown
3	T	0314-73+	positive	unknown
4	L	0318K-01+	positive	unknown
5	L	0318J-26+	positive	unknown
6	L	0318J-29+	positive	unknown
7	T	0324J-30+	positive	unknown
8	L	0324J-31+	positive	unknown
9	L	0324J-32+	positive	unknown
10		0324J-33+	positive	unknown
11	L	0124-8930+	Positive	unknown
12	T	0218-39+	positive	unknown
13		0219-4034+	positive	unknown
14	T	0314J-24+	positive	ND
15	T	0212-10+	positive	37.4
16	L	0224-5279+	positive	36.1
17	T	0311H-18+	positive	35.2
18	L	0228-64+	positive	34.59
19	L	0317-46+	positive	34.4
20	T	0307J-15+	positive	33.8
21	T	0307J-16+	positive	33.8
22	L	0125-27+	Positive	33.71
23	T	0310J-20+	positive	33.7
24	T	0214-90+	positive	33.614
25	T	0124-89+	Positive	33.3
26	T	0204H-02+	positive	33.2988
27	T	0209H-09+	positive	33.2
28		0218J-03+	positive	33.2
29		0228J-13+	positive	33
30	T	0314J-22+	positive	33
31	T	0302-11+	positive	32.9
32	L	0225J-09+	positive	32.8
33	L	0204-49+	positive	32.676
34	L	0307-08+	positive	32.59

35	L	0314J-21+	positive	32.1
36		0214-74+	positive	31.992
37	L	0207H-07+	positive	31.9
38	L	0209-12+	positive	31.87
39		0219J-04+	positive	31.7
40		0225J-11+	positive	31.5
41	T	0131-5989+	positive	31.4
42	L	0228-40+	positive	31.33
43		0318H-19+	positive	30.7
44	L	0309-45+	positive	30.67
45	L	0202-47+	positive	30.621
46	T	0124-77+	Positive	30.25
47		0125-28+	Positive	29.99
48	T	0217-11+	positive	29.924
49		0228J-12+	positive	29.9
50	T	0126-50+	Positive	29.63
51		0309H-17+	positive	29.5
52	L	0309-81+	positive	29.5
53	L	0210-1079+	positive	29.3
54	T	0303-49+	positive	29.02
55	T	0228J-11+	positive	29
56	T	0322-56+	positive	28.89
57	T	0128-39+	positive	28.51
58	T	0204H-01+	positive	28.4384
59		0126-3676+	Positive	28.3
60	L	0210-93+	positive	27.857
61	L	0204H-03+	positive	27.8179
62	T	0302J-14+	positive	27.7
63	L	0124-56+	Positive	27.69
64		0318J-27+	positive	27.6
65		0207-102+	positive	27.57
66	L	0221-91+	positive	27.5
67	T	0307J-17+	positive	27.5
68	T	0202-49+	positive	27.481
69		0224J-08+	positive	27.4
70		0316-85+	positive	27.4

71	L	0210-23+	positive	27.33
72	T	0126-22+	Positive	27.32
73	T	0124-50+	Positive	27.25
74	T	0307-60+	positive	27.06
75	T	0202-42+	positive	26.981
76	L	0203-7423+	positive	26.9
77	L	0222-60+	positive	26.9
78		0201-52+	positive	26.81
79		0222-83+	positive	26.8
80	L	0209H-08+	positive	26.7
81	T	0125-85+	Positive	26.59
82	T	0310J-19+	positive	26.5
83	L	0124-49+	Positive	26.4
84	L	0318K-03+	positive	26.4
85	L	0207-8532+	positive	26.1
86	L	0212-11+	positive	26.1
87		0314J-23+	positive	26
88		0209-65+	positive	25.979
89		0223-03+	positive	25.93
90	T	0211H-12+	positive	25.9
91	L	0318J-25+	positive	25.9
92	L	0303-50+	positive	25.88
93	L	0218-3381+	positive	25.5
94	L	0302-87+	positive	25.5
95	L	0221J-06+	positive	25.3
96	T	0211H-11+	positive	25.2
97		0322-57+	positive	25.19
98	L	0128-41+	positive	25.1
99		0201-55+	positive	24.51
100	T	0307-57+	positive	24.13
101	T	0318J-28+	positive	24.1
102		0318-6+	positive	24
103	L	0214-75+	positive	23.902
104	L	0214-77+	positive	23.066
105	T	0221J-07+	positive	23
106		0202-85+	positive	22.898

107	L	0207-60+	positive	22.76
108	L	0322-50+	positive	22.53
109	T	0124-45+	Positive	22.52
110	L	0307-86+	positive	22.49
111	L	0204H-05+	positive	22.2744
112	T	0204H-04+	positive	21.8555
113		0225-5703+	positive	21.5
114	T	0127-21＋	Positive	21.19
115		0207-64+	positive	21.13
116	T	0209H-10+	positive	20.1
117	T	0207-58+	positive	20.09
118	T	0208-85＋	positive	19.508
119	T	0310J-18+	positive	19.3
120	L	0208-76＋	positive	19.258
121		0207H-06+	positive	19
122	L	0207-103+	positive	18.83
123	T	0203-29＋	positive	18.673
124	L	0126-52+	Positive	18.1
125	T	0217-10+	positive	17.796
126	L	0318K-02+	positive	17.7
127	L	0218J-02+	positive	17.1
128		0221-02+	positive	16.8
129		0209-66+	positive	16.747
130	T	0209-60+	positive	16.053
131	T	0121-6493+	Positive	15.32
132	L	0221H-14+	positive	14.8
133		0218H-14+	positive	14.4
134	T	0122MM	Negative	
135	L	0122MW	Negative	
136	L	0122SK	Negative	
137		0124YT	Negative	
138	T	0124TS	Negative	
139	T	0124SO	Negative	
140	T	0125-44	Negative	
141		0125-47	Negative	
142	T	0125-48	Negative	

143	L	0125-53	Negative
144	T	0125-54	Negative
145	T	0126-29	Negative
146	T	0126-25	Negative
147		0126-31	Negative
148	T	0126-53	Negative
149	T	0126-55	Negative
150	T	0127-2	Negative
151	T	0127-3	Negative
152	T	0127-4	Negative
153	L	0127-9	Negative
154		0127-10	Negative
155	T	0128-4	Negative
156	T	0128-14	Negative
157	L	0128-21	Negative
158	T	0128-40	Negative
159	T	0128-44	Negative
160	T	0131-3	negative
161	L	0131-6	negative
162	T	0131-33	negative
163		0131-37	negative
164	T	0131-63	negative
165	L	0203-5	negative
166	L	0203-6	negative
167	T	0203-7	negative
168		0203-15	negative
169	L	0203-16	negative
170	T	0204-39	negative
171	L	0204-42	negative
172	L	0204-45	negative
173	L	0204-46	negative
174	T	0204-54	negative
175	L	0208-11	negative
176	T	0208-22	negative
177		0208-39	negative
178	T	0208-40	negative

179	T	0208-45	negative
180	L	0210-24	negative
181	L	0210-32	negative
182	T	0210-55	negative
183	T	0210-73	negative
184	L	0210-87	negative
185		0214-14	negative
186	T	0214-15	negative
187	L	0214-31	negative
188	L	0215-24	negative
189	T	0215-96	negative
190	L	0215-26	negative
191	L	0215-29	negative
192		0215-30	negative
193	L	0215-25	negative
194	T	0216-25	negative
195	L	0216-27	negative
196	L	0216-50	negative
197	T	0216-51	negative
198	L	0216-53	negative
199	T	0217-12	negative
200	T	0217-13	negative
201	L	0217-14	negative
202	T	0217-17	negative
203	L	0217-18	negative
204	L	0217-36	negative
205	L	0217-44	negative
206	L	0217-54	negative
207	L	0217-55	negative
208		0217-56	negative
209	L	0218H-13+	negative
210	L	0218J-01+	negative
211	T	0219-4021+	negative
212	L	0219J-05+	negative
213	L	0221-06	negative
214	L	0221-04	negative

215	L	0221-09	negative
216	T	0221-11	negative
217	L	0221-20	negative
218	T	0221-46	negative
219	T	0221-48	negative
220	L	0221H-15+	negative
221	T	0222-16	negative
222	L	0222-23	negative
223	L	0222-26	negative
224	L	0222-28	negative
225	T	0222-31	negative
226	T	0222-32	negative
227	T	0222-37	negative
228	L	0222-47	negative
229	T	0222-48	negative
230	L	0222-50	negative
231	L	0224-24	negative
232	L	0224-05	negative
233	T	0224-06	negative
234	T	0224-09	negative
235	L	0224-10	negative
236	L	0224-14	negative
237	T	0224-17	negative
238	T	0224-22	negative
239	T	0224-34	negative
240	L	0224-49	negative
241	L	0125-28+2	Positive

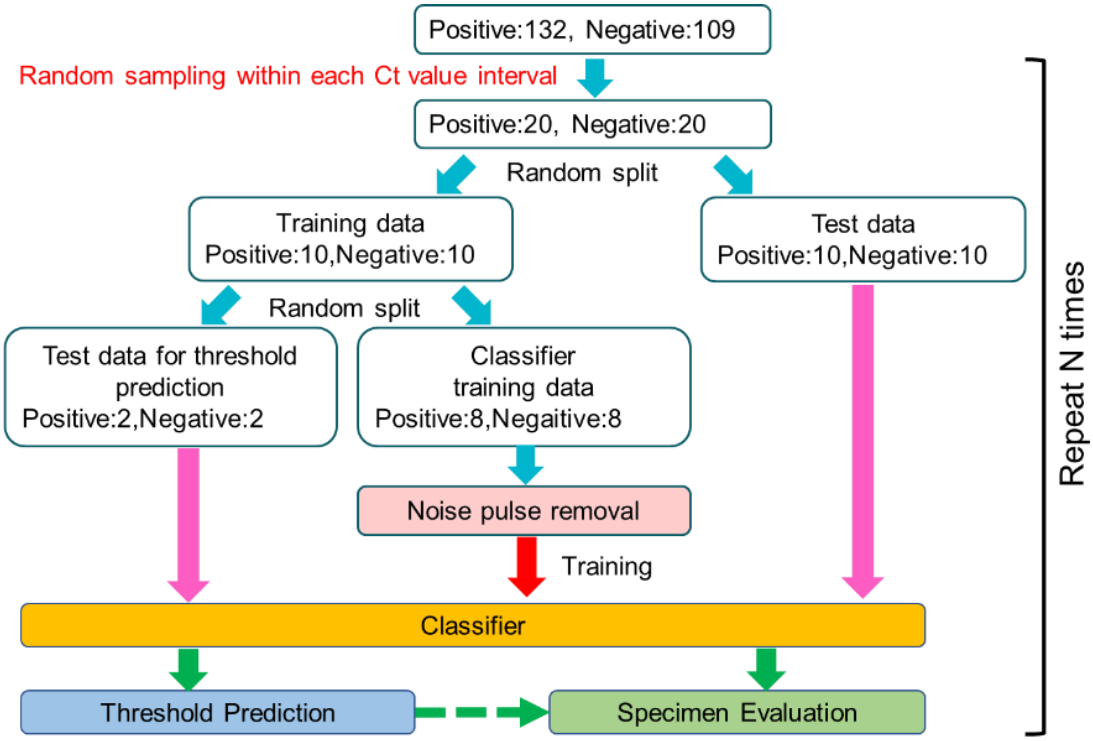
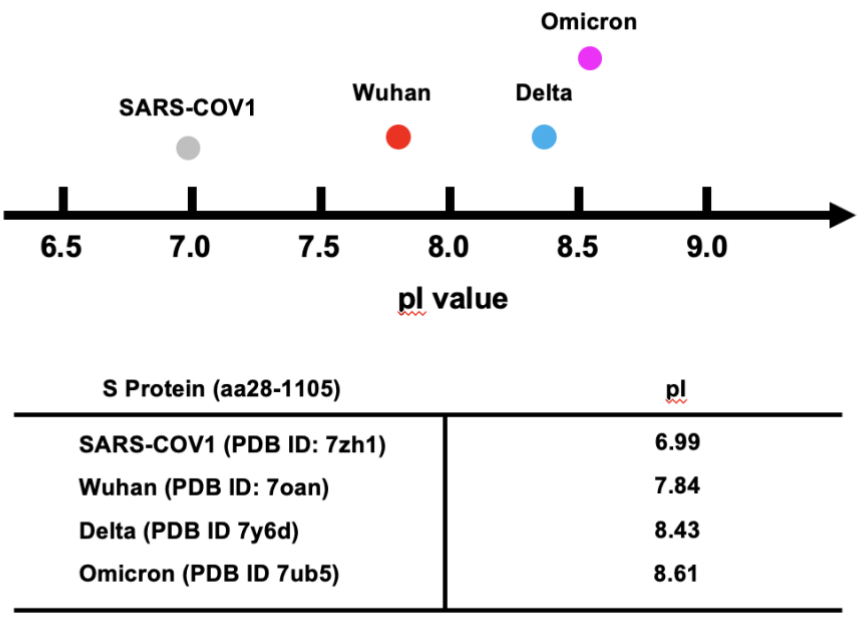


Figure S16 Analysis procedure for the C_t dependence of the *F*-measures, sensitivity, and specificity. Separate specimens for each C_t value were used to train and evaluate the classifiers.



S Protein (aa28-1105)	pI
SARS-COV1 (PDB ID: 7zh1)	6.99
Wuhan (PDB ID: 7oan)	7.84
Delta (PDB ID 7y6d)	8.43
Omicron (PDB ID 7ub5)	8.61

Figure S17 Isoelectric points of S-proteins on different coronaviruses.

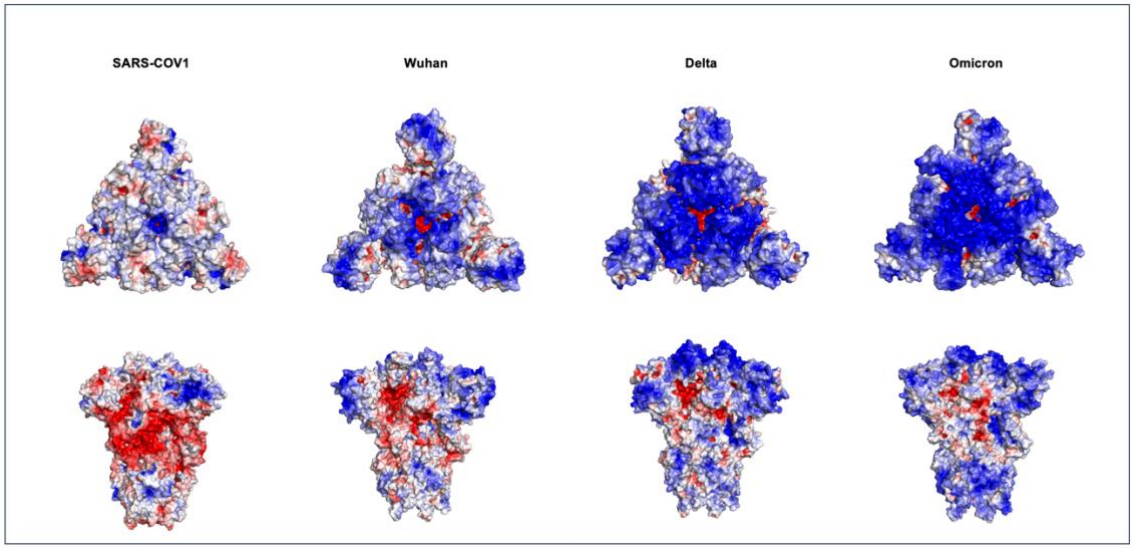


Figure S18 Electrostatic potentials of S-proteins on different coronaviruses.