

## SUPPLEMENTAL FILE

### Identification of driver genes for severe forms of COVID-19 in a deeply phenotyped young patient cohort

Raphael Carapito<sup>1,2,3,†</sup>, Richard Li<sup>4†</sup>, Julie Helms<sup>1,3,5</sup>, Christine Carapito<sup>3,6</sup>, Sharvari Gujja<sup>4</sup>, Véronique Rolli<sup>1,2,3</sup>, Raony Guimaraes<sup>4</sup>, Jose Malagon-Lopez<sup>4</sup>, Perrine Spinnhirny<sup>1,3</sup>, Razieh Mohseninia<sup>7</sup>, Aurélie Hirschler<sup>3,6</sup>, Leslie Muller<sup>3,6</sup>, Paul Bastard<sup>8,9,10</sup>, Adrian Gervais<sup>9,10</sup>, Qian Zhang<sup>8,9,10</sup>, François Danion<sup>3,11</sup>, Yvon Ruch<sup>3,11</sup>, Maleka Schenck-Dhif<sup>3,12</sup>, Olivier Collange<sup>3,13</sup>, Thiên-Nga Chamaraux-Tran<sup>3,14</sup>, Anne Molitor<sup>1,3</sup>, Angélique Pichot<sup>1,3</sup>, Alice Bernard<sup>1,3</sup>, Ouria Tahar<sup>2,3</sup>, Sabrina Bibi-Triki<sup>1,3</sup>, Haiguo Wu<sup>4</sup>, Nicodème Paul<sup>1,3</sup>, Sylvain Mayeur<sup>1,3</sup>, Annabel Larnicol<sup>1,3</sup>, Géraldine Laumond<sup>1,3</sup>, Julia Frappier<sup>1,3</sup>, Sylvie Schmidt<sup>1,3</sup>, Antoine Hanauer<sup>1,3</sup>, Cécile Macquin<sup>1,3</sup>, Tristan Stemmelen<sup>1,2,3</sup>, Michael Simons<sup>15</sup>, Xavier Mariette<sup>16,17</sup>, Olivier Hermine<sup>10,18</sup>, Samira Fafi-Kremer<sup>1,3,19</sup>, Bernard Goichot<sup>3,20</sup>, Bernard Drenou<sup>21</sup>, Khaldoun Kuteifan<sup>22</sup>, Julien Pottecher<sup>3,14</sup>, Paul-Michel Mertes<sup>3,13</sup>, Shweta Kailasan<sup>23</sup>, M. Javad Aman<sup>23</sup>, Elisa Pin<sup>24</sup>, Peter Nilsson<sup>24</sup>, Anne Thomas<sup>25</sup>, Alain Viari<sup>25</sup>, Damien Sanlaville<sup>25</sup>, Francis Schneider<sup>3,12</sup>, Jean Sibilia<sup>1,3,26</sup>, Pierre-Louis Tharaux<sup>27</sup>, Jean-Laurent Casanova<sup>8,9,10,28</sup>, Yves Hansmann<sup>3,11</sup>, Daniel Lidar<sup>7,29</sup>, Mirjana Radosavljevic<sup>1,2,3</sup>, Jeffrey R. Gulcher<sup>4</sup>, Ferhat Meziani<sup>3,5</sup>, Christiane Moog<sup>1,3</sup>, Thomas W. Chittenden<sup>4,30\*</sup>, Seiamak Bahram<sup>1,2,3\*</sup>

## SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

**Table S1. Critical patients with detectable viremia and the corresponding FPKM values of SARS-CoV-2 gene**

Sample ID	FPKM* mean	ORF1ab	ORF1a b	S	N	ORF3a	M	ORF8	ORF7a	E	ORF6	ORF7 b	ORF1 0
P14	0.0008333	0	0.01	0	0	0	0	0	0	0	0	0	0
P27	0.0008333	0	0.01	0	0	0	0	0	0	0	0	0	0
P31	0.0125	0	0	0.01	0	0	0	0	0.14	0	0	0	0
P32	0.0025	0	0	0	0.03	0	0	0	0	0	0	0	0
P37	0.2683333	0.14	0	0.18	0.41	0.08	0.52	0.13	0.13	0.35	1.28	0	0
P39	0.0175	0	0.01	0.03	0	0	0.05	0	0.12	0	0	0	0
P43	0.0066667	0.01	0	0	0.07	0	0	0	0	0	0	0	0
P46	0.02	0.02	0	0.04	0.15	0.03	0	0	0	0	0	0	0

\*FPKM: fragments per kilo per million

**Table S2. Performance metrics on the training and testing sets of each algorithm in the ensemble computational intelligence approach**

	LASSO	Ridge	SVM	qSVM	XGB	RF	DANN
<b>Accuracy (Train/Test)</b>	0.9991 ±	1.0000 ±	1.0000 ±	0.9245 ±	0.9952 ±	1.0000 ±	1.0000 ±
	0.0004 /	0.0000 /	0.0000 /	0.0028 /	0.0008 /	0.0000 /	0.0000 /
	0.9677 ±	0.9169 ±	0.9223 ±	0.8677 ±	0.9146 ±	0.9254 ±	0.9131 ±
	0.0050	0.0072	0.0075	0.0121	0.0076	0.0072	0.0083
<b>Balanced Acc. (Train/Test)</b>	0.9987 ±	1.0000 ±	1.0000 ±	0.9189 ±	0.9930 ±	1.0000 ±	1.0000 ±
	0.0006 /	0.0000 /	0.0000 /	0.0039 /	0.0012 /	0.0000 /	0.0000 /
	0.9503 ±	0.8990 ±	0.9068 ±	0.8607 ±	0.8932 ±	0.9072 ±	0.9032 ±
	0.0078	0.0094	0.0092	0.0118	0.0100	0.0094	0.0097
<b>AUROC (Train/Test)</b>	1.0000 ±	1.0000 ±	1.0000 ±	0.9667 ±	0.9999 ±	1.0000 ±	1.0000 ±
	0.0000 /	0.0000 /	0.0000 /	0.0029 /	0.0000 /	0.0000 /	0.0000 /
	0.9908 ±	0.9547 ±	0.9633 ±	0.9386 ±	0.9443 ±	0.9360 ±	0.9435 ±
	0.0036	0.0075	0.0070	0.0081	0.0079	0.0091	0.0081
<b>F1 (Train/Test)</b>	0.9993 ±	1.0000 ±	1.0000 ±	0.9426 ±	0.9964 ±	1.0000 ±	1.0000 ±
	0.0003 /	0.0000 /	0.0000 /	0.0020 /	0.0006 /	0.0000 /	0.0000 /
	0.9780 ±	0.9404 ±	0.9487 ±	0.9095 ±	0.9391 ±	0.9467 ±	0.9359 ±
	0.0034	0.0052	0.0049	0.0071	0.0054	0.0052	0.0062
<b>MCC (Train/Test)</b>	0.9980 ±	1.0000 ±	1.0000 ±	0.8339 ±	0.9893 ±	1.0000 ±	1.0000 ±
	0.0009 /	0.0000 /	0.0000 /	0.0065 /	0.0018 /	0.0000 /	0.0000 /
	0.9251 ±	0.8128 ±	0.8364 ±	0.7398 ±	0.8061 ±	0.8308 ±	0.8091 ±
	0.0118	0.0169	0.0161	0.0198	0.0181	0.0168	0.0185

**Table S3. eQTLs identified in three driver genes by MatrixeQTL**

SNP*	rs number	gene	beta	t-stat	P-value	FDR
chr8:38996464-C-A	rs7840270	<i>ADAM9</i>	-0.560481	-4.461647	0.000034	0.038072
chr8:38997543-G-T	rs7831735	<i>ADAM9</i>	-0.565580	-4.359599	0.000049	0.046521
chr19:7742229-G-A	rs11465401	<i>MCEMP1</i>	1.912424	4.333792	0.000054	0.048775
chr19:7742364-G-A	rs11465397	<i>MCEMP1</i>	1.912424	4.333792	0.000054	0.048775
chr11:60510522-C-T	rs189755275	<i>MS4A4A</i>	2.328040	4.358676	0.000049	0.046648
chr11:60547398-G-A	rs76847438	<i>MS4A4A</i>	2.328040	4.358676	0.000049	0.046648
chr11:60582964-G-A	rs10736707	<i>MS4A4A</i>	-2.328040	-4.358676	0.000049	0.046648
chr11:60623519-G-A	rs10792287	<i>MS4A4A</i>	-2.328040	-4.358676	0.000049	0.046648

\*The positions refer to GRCh38

### Figure S1. Type I interferon response

**A.** Interferon stimulated genes (ISG) scores based on the mean normalized expression of six genes (*IFI44L*, *IFI27*, *RSAD2*, *SIGLEC1*, *IFIT1*, and *ISG15*) obtained from RNAseq data. **B.** Heatmap showing the expression of type I IFN-related genes obtained from RNAseq data. Upregulated proteins are shown in red, and downregulated proteins are shown in light blue. **C.** IFN $\alpha$ 2a (pg/ml) concentration evaluated using an ultra-sensitive S-PLEX Human IFN $\alpha$ 2a Kit (Mesoscale Discovery). **D.** Time-dependent IFN $\alpha$ 2a concentration in the critical group. **E.** Quantification of plasmacytoid dendritic cells as a percentage of PBMCs. The *P*-values were determined with the Kruskal-Wallis test followed by Dunn's posttest for multiple group comparisons; \**P* < 0.05, \*\* *P* < 0.01, \*\*\* *P* < 0.001, \*\*\*\* *P* < 0.0001.

### Figure S2. Additional immune profiling of healthy individuals, noncritical and critical COVID-19 patients by mass cytometry

Proportions of modified lymphocyte subsets from COVID-19 patients and healthy controls as determined by mass cytometry. The proportions of dendritic cell subsets (**A**), monocyte subsets (**B**), NK cell subsets (**C**), NKT cells (**D**),  $\gamma\delta$  T cells (**E**) and granulocyte subsets (traces), including neutrophils (**F**), are shown. Each dot represents a single patient. The *P*-values were determined with the Kruskal-Wallis test, followed by Dunn's posttest for multiple group comparisons; \* *P* < 0.05, \*\* *P* < 0.01, \*\*\* *P* < 0.001, \*\*\*\* *P* < 0.0001.

### Figure S3. SARS-CoV-2 immunization in healthy individuals, non-critical and critical COVID-19 patients

**A.** Antibody reactivities against three viral antigens expressed in mean fluorescence intensities (MFI). *P*-values were determined with the Kruskal-Wallis test, followed by Dunn's post-test for multiple group comparison; \* *P* < 0.05, \*\* *P* < 0.01, \*\*\* *P* < 0.001, \*\*\*\* *P* < 0.0001. **B.** Half-maximal neutralizing titer (NT50) values of SARS-CoV-2 pseudo-viruses. *P*-values were determined with the Mann-Whitney's test; \* *P* < 0.05, \*\* *P* < 0.01, \*\*\* *P* < 0.001, \*\*\*\* *P* < 0.0001. **C.** Neutralization activity of type IFN autoantibodies was assessed by

measuring the IFN stimulated response after treatment of HEK293T cells with patient or control serum. Relative luciferase activity is shown after stimulation with 10 ng/ml of IFN $\alpha$  or IFN $\omega$ . ISRE, IFN stimulation response element; RLU, relative light units. The dotted line indicates the cutoff for positivity. For individuals in which different time points were measured, the lowest RLU value is shown. **D.** Time-course evaluation of the anti-SARS-CoV-2 neutralizing activity (left axis) and the anti IFN type I autoantibodies (right axis) in two critical patients with anti-type I autoantibodies. The farmost right points correspond to the time points represented in panel C.

#### **Figure S4. Results of *in silico* perturbation experiments**

Left: change in BIC (Bayesian Information Criterion) when perturbing each gene individually. Genes are ordered by the number of ancestors minus the number of descendants for the DAG shown in Figure 5B: the top five driver genes are the five leftmost points, and the top five response genes are the five rightmost points. Right: Change in the BIC of a random sample of five genes from the left. The mean BIC of the top five driver genes is shown in red.

#### **Figure S5. Validation of the RNA-seq signature-based classification performance of critical and recovered critical COVID-19 patients**

**A.** ROCs on the train and test sets for critical vs recovered critical groups comparison in the replication cohort using the 600 gene signature identified from the initial cohort. All methods performed similarly. **B.** Classification metrics. **C.** Box plots showing the normalized gene counts of the five driver genes in critical and recovered critical patients. The indicated values correspond to the FDR.

#### **Figure S6. *ADAM9* expression in publicly available data**

Box plots showing the normalized gene counts of *ADAM9* in healthy (n=17), severe (n=8) and ICU (n=3) patients in the GSE152418 dataset reported by Arunachalam et al., in Science (DOI:10.1126/science.abc6261). The indicated values correspond to the FDR.

**Figure S7. Validation of *ADAM9* silencing**

**A.** Quantitative RT-PCR of the *ADAM9* transcript in Vero 76 or A549-ACE2 cells silenced with a control siRNA or an *ADAM9*-specific siRNA. The average silencing achieved was 66% and 93% for Vero 76 and A549-ACE2 cells, respectively (mean of three representative experiments). **B.** Western blot of Vero 76 and A549-ACE2 cells that were not transfected (NT), were silenced with a control siRNA (ctl) or were silenced with an *ADAM9*-specific siRNA (sil.).

Figure S1.

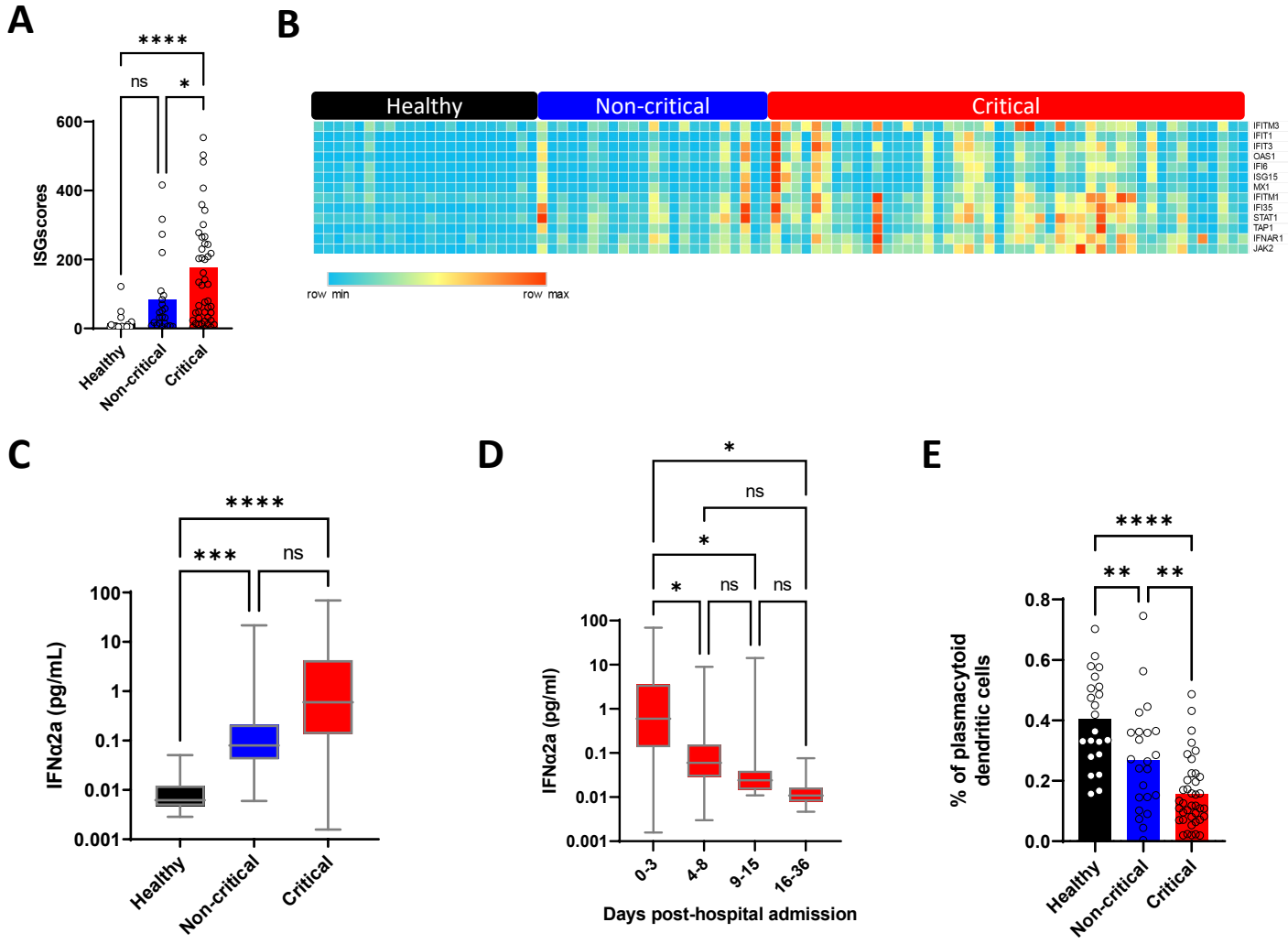




Figure S2.

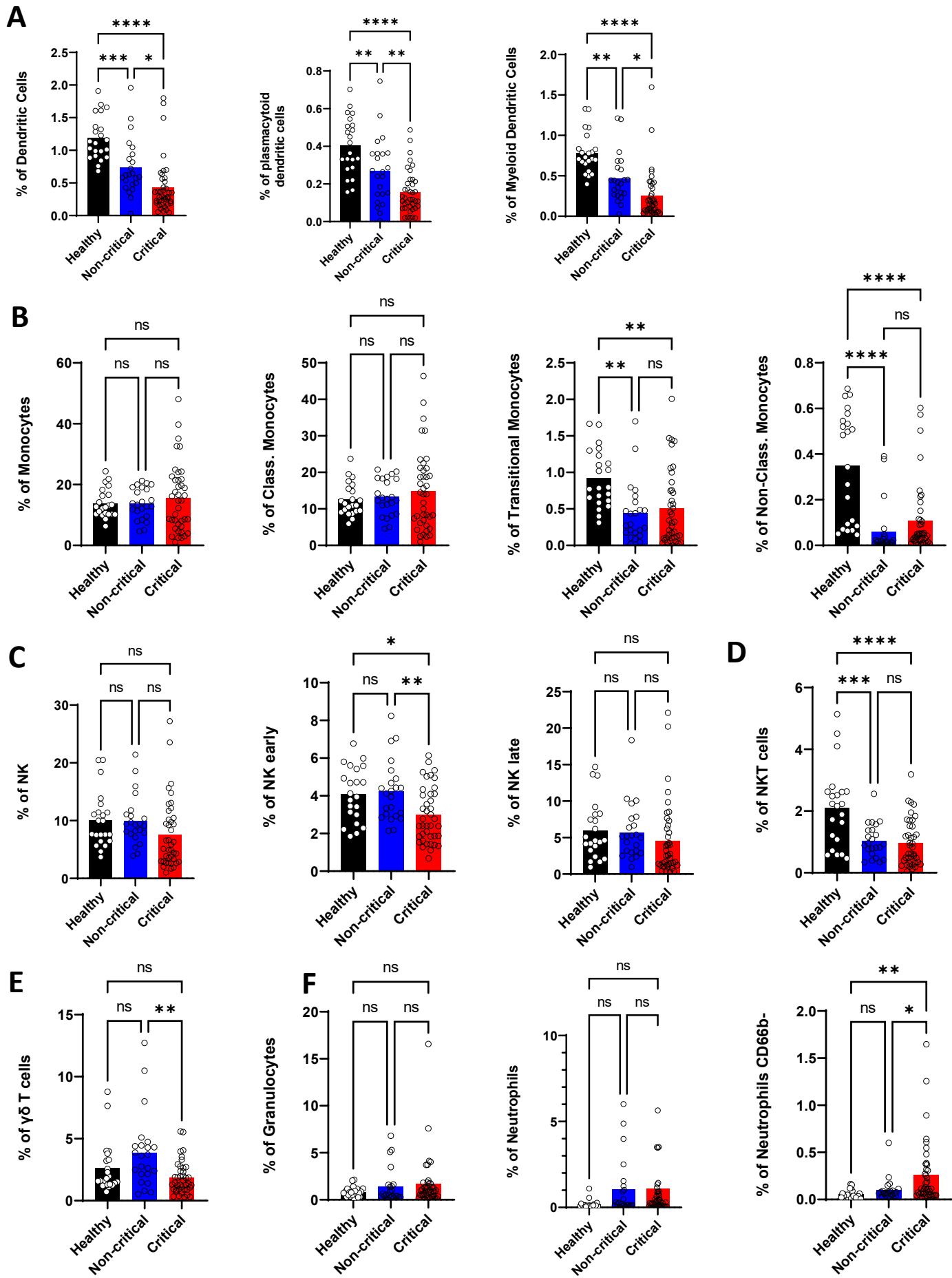


Figure S3.

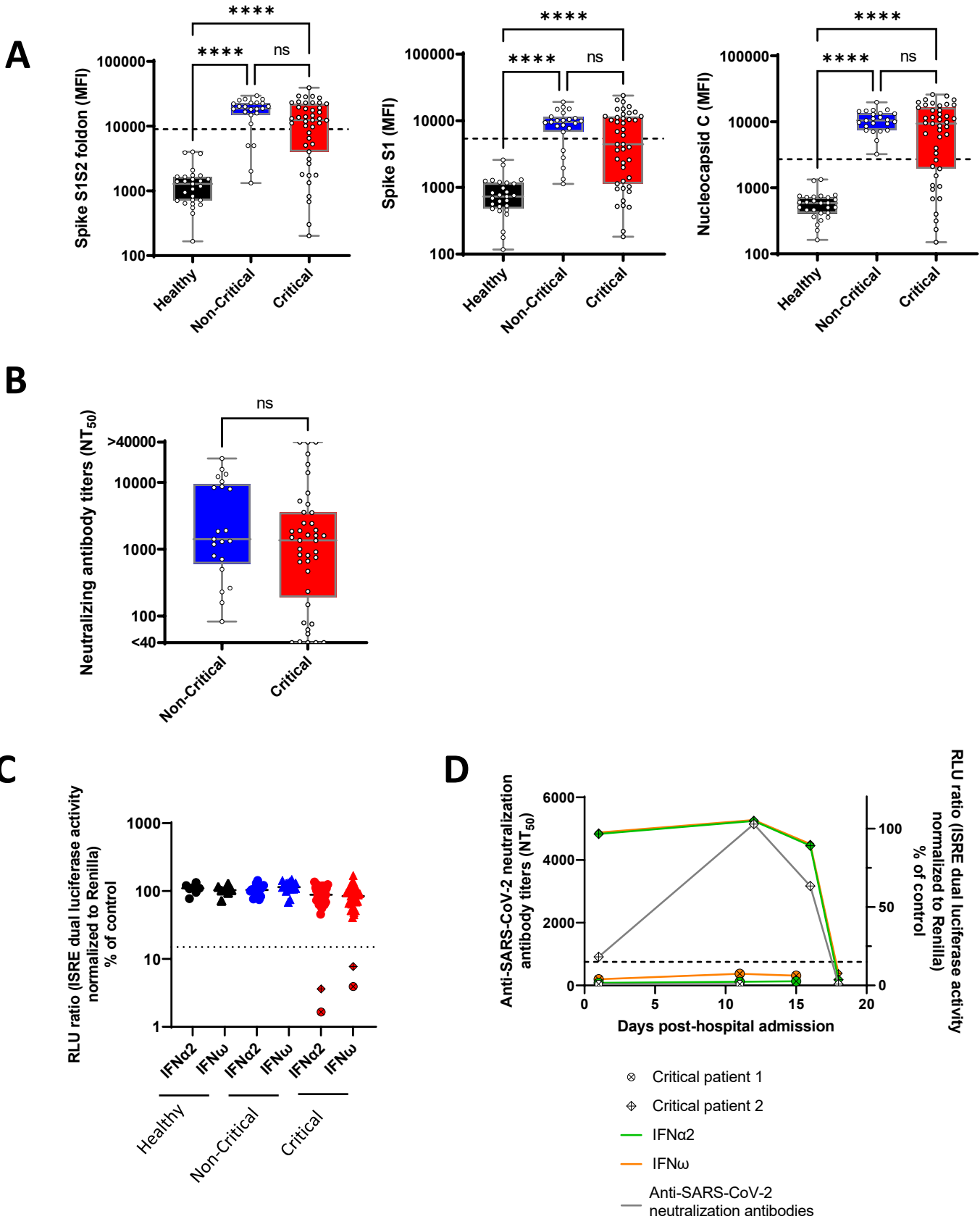


Figure S4.

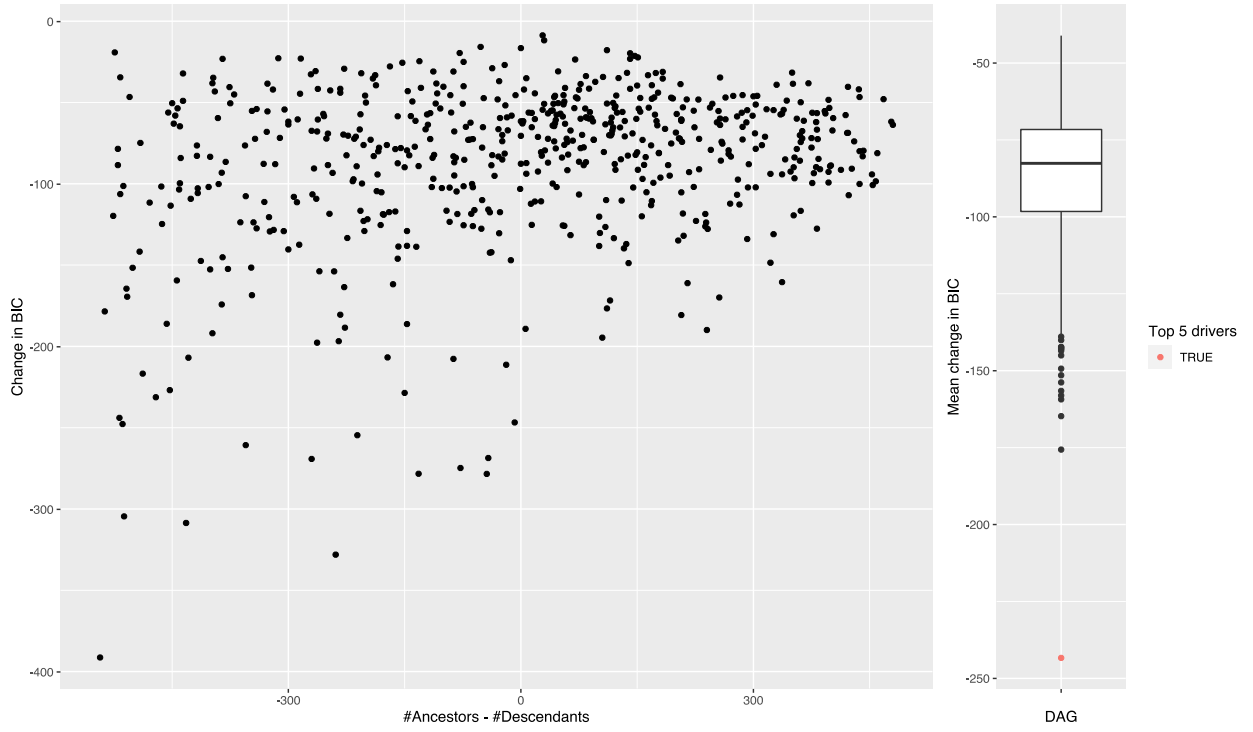
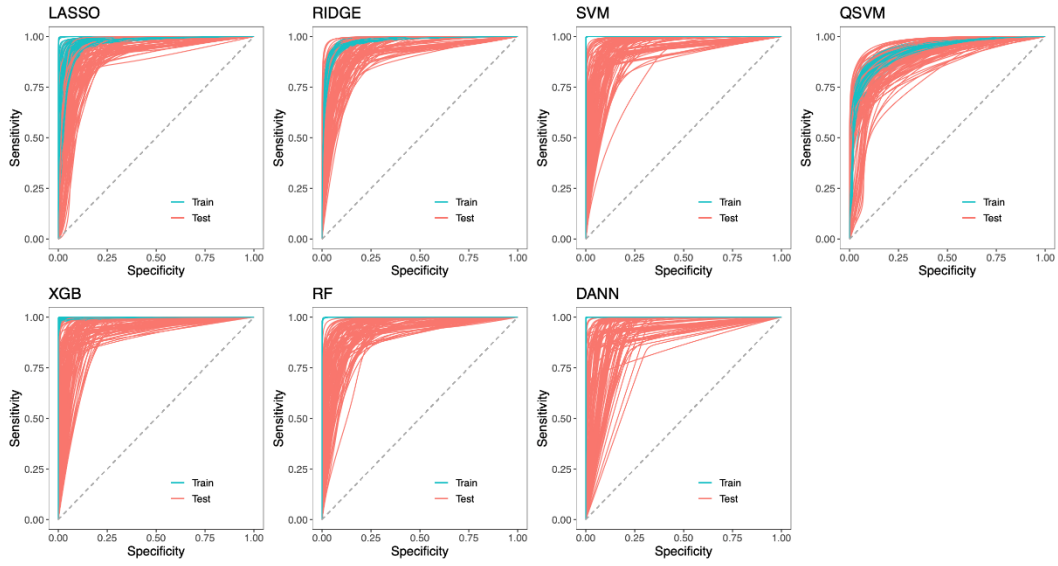


Figure S5.

A



B

	LASSO	Ridge	SVM	αSVM	XGB	RF	DANN
Accuracy (Train/Test)	0.9535 ± 0.0016 / 0.9233 ± 0.0044	0.9453 ± 0.0010 / 0.9123 ± 0.0055	1.0000 ± 0.0000 / 0.9160 ± 0.0047	0.9077 ± 0.0019 / 0.8873 ± 0.0064	0.9983 ± 0.0004 / 0.9223 ± 0.0045	1.0000 ± 0.0000 / 0.9180 ± 0.0049	1.0000 ± 0.0000 / 0.9130 ± 0.0051
Balanced Acc. (Train/Test)	0.9544 ± 0.0016 / 0.9262 ± 0.0042	0.9463 ± 0.0010 / 0.9154 ± 0.0053	1.0000 ± 0.0000 / 0.9184 ± 0.0045	0.9090 ± 0.0018 / 0.8894 ± 0.0062	0.9983 ± 0.0004 / 0.9242 ± 0.0044	1.0000 ± 0.0000 / 0.9197 ± 0.0048	1.0000 ± 0.0000 / 0.9156 ± 0.0049
AUROC (Train/Test)	0.9849 ± 0.0009 / 0.9628 ± 0.0031	0.9879 ± 0.0003 / 0.9684 ± 0.0029	1.0000 ± 0.0000 / 0.9671 ± 0.0029	0.9548 ± 0.0011 / 0.9486 ± 0.0041	0.9999 ± 0.0000 / 0.9789 ± 0.0023	1.0000 ± 0.0000 / 0.9644 ± 0.0032	1.0000 ± 0.0000 / 0.9643 ± 0.0033
F1 (Train/Test)	0.9547 ± 0.0016 / 0.9233 ± 0.0046	0.9466 ± 0.0010 / 0.9115 ± 0.0058	1.0000 ± 0.0000 / 0.9164 ± 0.0050	0.9061 ± 0.0019 / 0.8849 ± 0.0063	0.9984 ± 0.0004 / 0.9235 ± 0.0046	1.0000 ± 0.0000 / 0.9199 ± 0.0049	1.0000 ± 0.0000 / 0.9133 ± 0.0053
MCC (Train/Test)	0.9080 ± 0.0032 / 0.8547 ± 0.0081	0.8917 ± 0.0021 / 0.8346 ± 0.0102	1.0000 ± 0.0000 / 0.8406 ± 0.0087	0.8179 ± 0.0036 / 0.7839 ± 0.0121	0.9966 ± 0.0008 / 0.8515 ± 0.0085	1.0000 ± 0.0000 / 0.8418 ± 0.0094	1.0000 ± 0.0000 / 0.8343 ± 0.0096

C

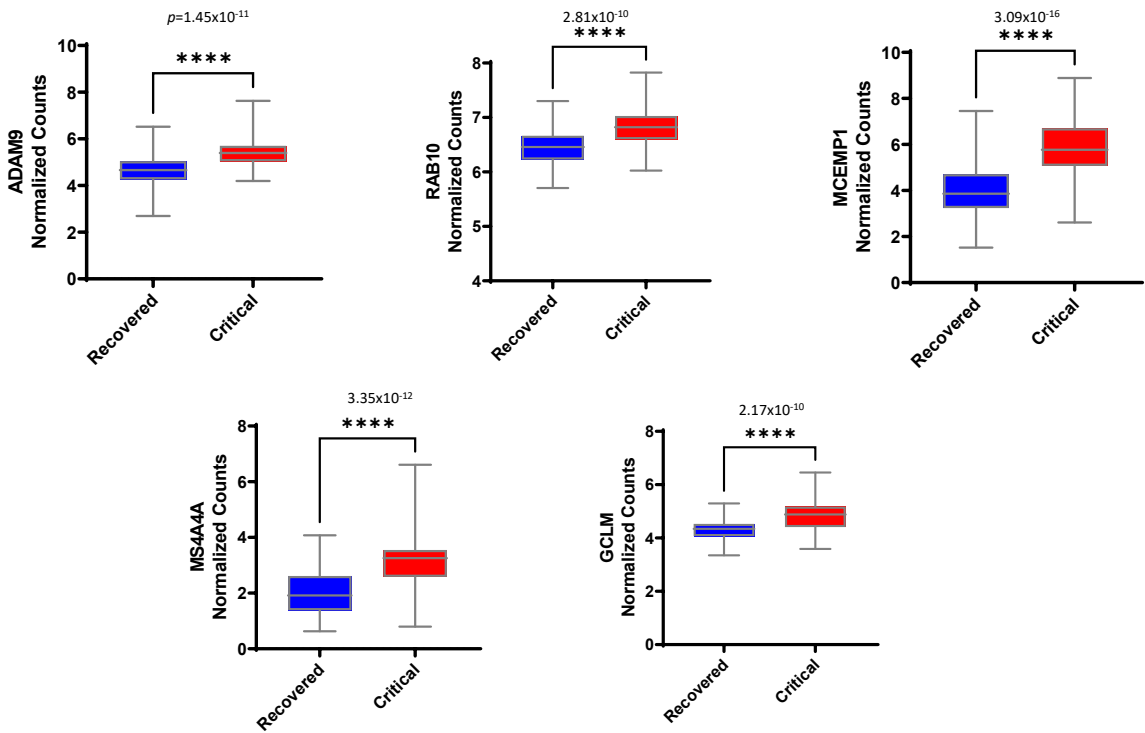


Figure S6.

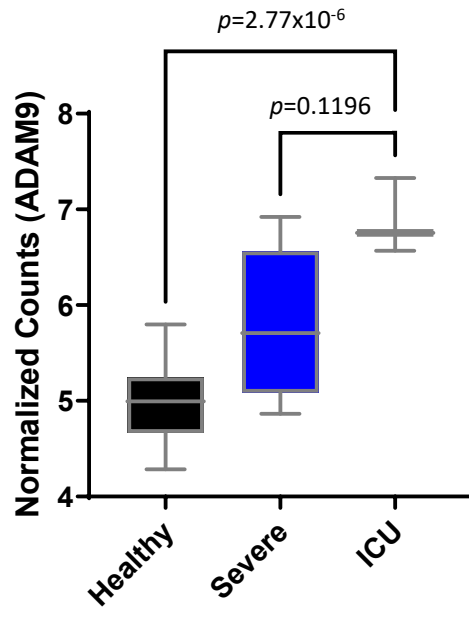
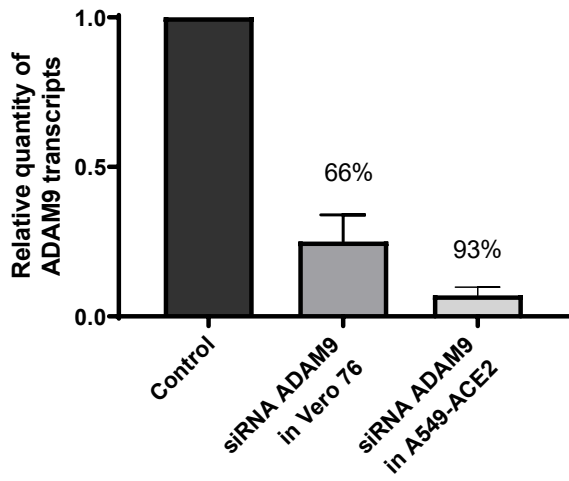


Figure S7.

**A**



**B**

