

1 ***In vivo* antiviral host response to SARS-CoV-2 by viral load, sex, and age**

2

3 *Short Title: Antiviral host response to SARS-CoV-2*

4

5 Nicole A. P. Lieberman¹, Vikas Peddu¹, Hong Xie¹, Lasata Shrestha¹, Meei-Li Huang¹, Megan C.

6 Mears^{2,3}, Maria N. Cajimat^{2,3}, Dennis A. Bente^{2,4}, Pei-Yong Shi^{2,5}, Francesca Bovier⁶, Pavitra

7 Roychoudhury^{1,7}, Keith R. Jerome^{1,7}, Anne Moscona^{6,8,9,10}, Matteo Porotto^{6,8,11}, Alexander L.

8 Greninger^{1,7*}

9

10 ¹Department of Laboratory Medicine, University of Washington School of Medicine, Seattle,

11 WA, USA

12 ²Galveston National Laboratory, University of Texas Medical Branch, Galveston, Texas, 77550

13 ³Department of Experimental Pathology, University of Texas Medical Branch, Galveston, Texas,

14 77550

15 ⁴Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston,

16 Texas, 77550

17 ⁵Department of Biochemistry and Molecular Biology, University of Texas Medical Branch,

18 Galveston, Texas, 77550

19 ⁶Center for Host–Pathogen Interaction, Columbia University Medical Center, New York, New

20 York, 10032, USA

21 ⁷ Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA

22 ⁸Department of Pediatrics, Columbia University Medical Center, New York, New York, USA

23 ⁹Department of Microbiology & Immunology, Columbia University Medical Center, New York,
24 New York, 10032, USA

25 ¹⁰Department of Physiology & Cellular Biophysics, Columbia University Medical Center, New
26 York, New York, 10032, USA

27 ¹¹Department of Experimental Medicine, University of Campania “Luigi Vanvitelli”, 81100
28 Caserta, Italy

29 *Corresponding Author

30

31 **Correspondence:** Alexander L. Greninger, agrening@uw.edu, 1616 Eastlake Ave E, Suite 320,
32 Seattle, WA 98102, Phone: 415 439 3448, Fax: 206 616 4340

33

34 **Abstract**

35 Despite limited genomic diversity, SARS-CoV-2 has shown a wide range of clinical
36 manifestations in different patient populations. The mechanisms behind these host differences
37 are still unclear. Here, we examined host response gene expression across infection status, viral
38 load, age, and sex among shotgun RNA-sequencing profiles of nasopharyngeal swabs from 430
39 individuals with PCR-confirmed SARS-CoV-2 and 54 negative controls. SARS-CoV-2 induced a
40 strong antiviral response with upregulation of antiviral factors such as *OAS1-3* and *IFIT1-3*, and
41 Th1 chemokines *CXCL9/10/11*, as well as a reduction in transcription of ribosomal proteins.
42 SARS-CoV-2 culture in human airway epithelial cultures replicated the *in vivo* antiviral host
43 response. Patient-matched longitudinal specimens (mean elapsed time = 6.3 days)
44 demonstrated reduction in interferon-induced transcription, recovery of transcription of
45 ribosomal proteins, and initiation of wound healing and humoral immune responses. Expression
46 of interferon-responsive genes, including *ACE2*, increased as a function of viral load, while
47 transcripts for B cell-specific proteins and neutrophil chemokines were elevated in patients with
48 lower viral load. Older individuals had reduced expression of Th1 chemokines *CXCL9/10/11* and
49 their cognate receptor, *CXCR3*, as well as CD8A and granzyme B, suggesting deficiencies in
50 trafficking and/or function of cytotoxic T cells and natural killer (NK) cells. Relative to females,
51 males had reduced B and NK cell-specific transcripts and an increase in inhibitors of NF- κ B
52 signaling, possibly inappropriately throttling antiviral responses. Collectively, our data
53 demonstrate that host responses to SARS-CoV-2 are dependent on viral load and infection time
54 course, with observed differences due to age and sex that may contribute to disease severity.

55

56 Keywords: SARS-CoV-2, COVID-19, RNAseq, interferon, immune response, ribosomal proteins,

57 *ACE2*

58

59

60 **Introduction**

61 The novel coronavirus SARS-CoV-2 that emerged in late 2019 from Wuhan, China, has
62 rapidly spread throughout the world, causing more than 6 million cases and 400,000 deaths
63 globally as of June 2020. COVID-19 morbidity and mortality has been overwhelmingly
64 concentrated in elderly individuals and those with preexisting comorbidities (1). In older
65 individuals, immunosenescence and dysregulated antiviral responses due to viral chronic low-
66 grade age-related inflammation may play an important role (2), as has been proposed for
67 influenza (3). Males are known to be generally more susceptible to infectious disease than
68 females (4) and SARS-CoV-infected male mice had increased infiltration of inflammatory
69 macrophages into their lungs, leading to a deleterious inflammatory response (5). Accordingly,
70 systemic inflammatory markers such as neutrophil-to-lymphocyte ratio and C-reactive protein
71 were elevated in men who died of SARS-CoV-2 (6). However, the mechanisms behind increased
72 mortality among older adults and males with COVID-19 remain speculative.

73 Entry of SARS-CoV-2 into host cells depends on binding to the receptor ACE2 (7),
74 expressed at a high level in the nasal epithelium (8), then further induced upon exposure to
75 interferon (9), suggesting a mechanism by which SARS-CoV-2 exploits host antiviral responses.
76 SARS-CoV antagonizes initial viral detection and interferon responses by an as-yet unknown
77 mechanism (10,11). SARS-CoV-2 may employ similar mechanisms, as low MOI infections of
78 bronchial epithelial cells do not result in extensive transcription of interferon-stimulated genes
79 (ISGs) at 24 hours post infection (12). An important consequence of these observations is that
80 SARS-CoV-2 viral load and transmissibility peaks at the time of symptom onset (13,14). The
81 temporal relationship between viral load and host gene expression has not been fully explored.

82 In the United States, diagnostic testing is generally performed on nasopharyngeal (NP)
83 swabs, from which SARS-CoV-2 RNA can be recovered. Shotgun RNA sequencing of this material
84 allows for simultaneous recovery of viral genomes for transmission tracking as well as
85 understanding of *in situ* host response (15). Since the first detection of SARS-CoV-2 in the USA
86 in WA State, the University of Washington Virology Laboratory has performed shotgun RNA
87 sequencing to recover more than 1,000 viral genomes to understand the evolution and
88 molecular epidemiology of the virus (16,17). Here, we examine host specific gene expression
89 differences by SARS-CoV-2 infection status, host age, sex, and viral load in nasopharyngeal
90 swabs from 430 SARS-CoV-2 infected individuals and 54 negative controls.

91

92 **Results**

93 Since early March, 2020, the University of Washington Virology Laboratory has tested
94 more than 100,000 samples, primarily NP swabs, for infection with SARS-CoV-2 (18). Thousands
95 of SARS-CoV-2 positive samples, as well as negative controls, have subsequently been
96 metagenomically sequenced, contributing to a detailed understanding of the phylogeny and
97 molecular epidemiology of the virus (16,19). In this study, we selected a subset of sequenced
98 samples that had sufficient reads (>500,000) pseudoaligned to the human transcriptome to
99 examine gene expression changes as a result of RT-PCR-confirmed SARS-CoV-2 infection.
100 Patient demographics are summarized in Table 1.

101 We first characterized the genes most differentially expressed (DE) in the nasopharynx
102 as a result of SARS-CoV-2 infection (n=430 positive, 54 negative). After correcting for batch

103 effects, we found 83 differentially expressed genes ($\text{padj} < 0.1$ and absolute $\log_2\text{FoldChange} > 1$)
104 between SARS-CoV-2 positive and negative samples, comprising 41 upregulated genes and 42
105 downregulated genes (Supplementary Table 1). Clustering of samples by the 50 most significant
106 DE genes reveals multiple gene expression clusters among SARS-CoV-2 positive samples, while
107 most negative samples cluster together (Figure 1A). Consistent with results from Butler et al
108 (20), SARS-CoV-2 infection induces an interferon-driven antiviral response in the nasopharynx,
109 upregulating transcripts encoding viral sensors (*DDX60L*), chemokines that attract effector T
110 cells and NK cells (*CXCL9*, *10*, *11*), and direct inhibitors of viral replication and function (*MX2*,
111 *RSAD2*, *HERC5*), highlighted in Figure 1B

112 To interrogate the global regulatory and signaling programs induced by SARS-CoV-2
113 infection, we employed Gene Set Enrichment Analysis (GSEA) (21,22) of the 50 Hallmark Gene
114 Sets of the Molecular Signatures Database (23). Sets with a significant ($\text{FDR} < 0.05$) positive
115 enrichment score included Interferon Alpha, Interferon Gamma, and Inflammatory Responses
116 (Figure 1C, Supplementary Figure 1A). Interestingly, we also found several metabolic pathways
117 negatively enriched, including both Oxidative Phosphorylation and Glycolysis, suggesting a
118 global reduction in production of proteins related to cellular energy production (Figure 1C,
119 Supplementary Figure 1B). Broad downregulation of transcripts encoding metabolic machinery
120 may represent either an antiviral response or viral-mediated disruption of host transcripts. We
121 also performed a statistical enrichment test against the Biological Processes Gene Ontology
122 (24,25). The most enriched processes (Supplementary Figure 1C-E) are related to either
123 immune responses or translation. In addition to upregulation of innate antiviral transcripts

124 (Figure 1B, Supplementary Figure 1D), we also found a consistent downregulation of transcripts
125 encoding ribosomal proteins (Supplementary Figure 1E).

126 The SARS-CoV-2 receptor ACE2 is an interferon-regulated gene and is upregulated in
127 response to SARS-CoV-2 infection (8). We examined the relationship between viral load,
128 defined by the cycle threshold (Ct) of the N1 target during diagnostic PCR, and *ACE2*. We found
129 that *ACE2* expression was associated with increased viral load: median counts of negative, low
130 viral load (N1 ct > 24), medium viral load (N1 ct 24-19), and high viral load (N1 ct < 19) were 0,
131 1.93, 3.45, and 7.82, respectively ($p = 7.46e^{-13}$, by Kruskal-Wallis one-way ANOVA; Figure 2A). A
132 similar trend was found for other interferon-induced genes, a subset of which is shown in
133 Figure 2A including those significantly upregulated in SARS-CoV-2 infection (*CXCL9*, *OASL*, *MX1*),
134 negative regulators of inflammation (*CD274/PD-L1*, *USP18*), monocyte chemoattractant
135 protein-1 (*CCL2*) (26). Conversely, the protease required for viral entry, *TMPRSS2*, was reduced
136 upon viral infection but was not modulated by viral dose, nor were ribosomal proteins (*RPL4*,
137 *RPS6*).

138 We next specifically examined gene expression differences between high (N1 ct < 19,
139 n=108) relative to low (N1 ct > 24, n=99) viral load samples. Figure 2B highlights the 15 most
140 upregulated and 15 most downregulated of the 363 total differentially expressed genes
141 (adjusted pvalue < 0.1, Supplementary Table 2). While genes upregulated in high viral load
142 samples were dominated by proinflammatory and/or interferon-induced factors such as
143 *CXCL9/10*, *IDO1*, and *CD80*, genes with higher expression in low viral load samples included
144 chemokines for neutrophils (*CXCL8*, *S100A9*), and B cell-specific transcripts (*FCRL2*, *IGHG1*,
145 *IGHM*, *IGLL5*, *IGHG2*, *CD22*). Because this suggested differences in immune infiltration as a

146 result of viral load, we performed in silico cell sorting of immune cells using CIBERSORTx (27)
147 and found a higher proportion naïve B and T cells, neutrophils, and M2-polarized macrophages
148 in low viral load samples (3.5, 2.2, 1.6, and 1.8 fold increased, respectively), while high viral load
149 samples contained a larger proportion of M1 macrophages, activated NK cells, and activated
150 dendritic cells (2.5, 1.6, and 1.6 fold upregulated, respectively; Figure 2C). Levels of transcripts
151 encoding B cell proteins and neutrophil chemokines varied by viral load (Figure 2D).

152 Detection of differential infiltration of antigen-presenting cells and lymphocytes to the
153 nasopharynx in high vs low viral load samples highlights the role immune cells play in the host
154 response to SARS-CoV-2. To understand whether in vivo infection could be adequately
155 modeled in vitro, we examined gene expression differences in human airway epithelial (HAE)
156 cells 3 and 7 days post infection with SARS-CoV-2 and compared the DE genes at day 7 to those
157 from SARS-CoV-2 positive vs negative (Figure 1) and high vs low viral load SARS-CoV-2 positive
158 samples (Figure 2), resulting in a consensus set of 19 upregulated genes that define cell-intrinsic
159 host antiviral responses to SARS-CoV-2 infection (Figure 3A). When this consensus set was
160 tested for statistical enrichment in the DisGeNET (28) of disease ontologies, we found a high
161 degree of overlap with influenza signature genes (Figure 3B), including a number of interferon-
162 induced genes that mediate the acute antiviral response in the respiratory tract (Figure 3C).
163 Notably, in the HAE cells, there was no sign of induction of an interferon response at 3 days
164 post infection in spite of a 10-fold higher infectious dose of virus used and virus making up 0.3%
165 of reads. At 7 days post infection, SARS-CoV-2 comprised 5.3% of reads.

166 Observed heterogeneity in host response to SARS-CoV-2 infection (Figure 1A) may be a
167 result of co-infection or composition of the nasal flora. Proportion of reads assigned to virus,

168 bacteria, or human in Supplemental Figure 2A shows a range of bacterial:human ratio among
169 negative samples, while SARS-CoV-2 reads predominate at lower Ct values. Consistent with a
170 dramatic reduction in respiratory virus transmission (29), presumably due to physical distancing
171 measures enacted due to the SARS-CoV-2 pandemic, we found viral coinfections in only 14 of
172 430 SARS-CoV-2 positive samples (3.25%), and a single SARS-CoV-2 negative patient with a viral
173 infection (2.5%) (Supplementary Figure 2B). We also found a number of samples enriched for
174 potentially pathogenic bacterial components of the nasal flora, particularly *Moraxella*
175 *catarrhalis* (RPM>100 in 3/37 (8.1%) SARS-CoV-2 negative, and 58/413 (14.0%) SARS-CoV-2
176 positive), although the clinical significance of this observation is uncertain. Even after SARS-
177 CoV-2 reads were subtracted from each sample, we found that high viral load samples had a
178 significantly lower burden of bacteria than mid or low viral load, or SARS-CoV-2 negative
179 samples (Mann Whitney p = 0.0014, 0.0067, 0.00028, respectively; Supplementary Figure 2C).

180 Infection time course may account for the observed differences in immune related
181 genes in high vs low viral load samples: Patients receiving repeat SARS-CoV-2 testing have a
182 reduced viral load over time (13,30,31). Although we do not have the ability to tie our data
183 from our large set of positive samples back to the onset of symptoms, we have seen gene
184 expression changes in a small set (n=3) of matched longitudinal samples with a mean elapsed
185 time of 6.3 days between collections and a mean increase in Ct value of 5.29, representing a 39-
186 fold reduction in viral load (Figure 4A). Enriched GO terms include those related to translation
187 and immune regulation (Figure 4B). Notably, in the second sample collected, we saw increases
188 in genes such as *C1QA*, *-B*, and *-C* and *HLA-DQB1* that drive humoral immune responses and
189 those involved in wound healing (*APOE*, *CD36*, *RHOC*), as well as reductions in negative

190 regulators of each process (ie, *TREM1*, *TFPI*). We also tested the 19 SARS-CoV-2 signature genes
191 (Figure 3A), and found reductions in most at the second collection timepoint, although only
192 *RSAD2*, *IFIT2*, and *HERC5* decreases were statistically significant with three samples. We also
193 saw a recovery in the expression of ribosomal proteins over time (Figure 4C). Analysis of data
194 from additional patients with more extensive

195 Clinically, COVID-19 cases tend to be more severe for older adults and males (1). No
196 significant difference in N1 Ct was observed based on age or sex (Figure 5A). To understand the
197 differences in host response to SARS-CoV-2 infection, we tested the interaction between
198 infection and age (greater than 60), controlling for non-infection-related age differences with
199 SARS-CoV-2 negative samples. We found only two genes altered as a result of the interaction
200 between age and SARS-CoV-2 infection: a 30-fold reduction in production of CXCL11 (Figure
201 5B), an interferon-induced chemokine for natural killer and CD8+ T cells, and a 17-fold
202 reduction in polycomb group factor 6 (*PCGF6*) (Supplementary Figure 3A), a polycomb
203 repressor complex protein known to play a role in repression of dendritic cell activation (32).
204 Although we did not find additional genes altered specifically as a result of age in SARS-CoV-2
205 infection, we did find that *CXCL9* and *CXCL10* are not induced as strongly in SARS-CoV-2 positive
206 patients age 60 or higher. We also found reduced expression of the receptor for *CXCL9/10/11*,
207 *CXCR3*, the apoptosis-inducing factor *GZMB* secreted by NK and T cells, and the effector T cell
208 marker, *CD8A*. This data suggests that age-related T and NK cell dysfunction (33,34) may play a
209 role in SARS-CoV-2 pathogenesis in older individuals.

210 We performed a similar analysis to evaluate sex differences in SARS-CoV-2 infection and
211 found a total of 19 genes for which the differences in expression based on sex could be

212 attributed to SARS-CoV-2 infection. Supplementary Figure 3B highlights the top 10 non-
213 redundant enriched GO categories, most of which are related to immune function. In men, we
214 found downregulation of B cell-specific transcripts (*IGHG1* and *MSA4A/CD20*), downregulation
215 of the NK-activating receptor *SLAMF6*, and an upregulation of several inhibitors of NF κ B
216 signaling (*NDRG1*, *ARRB2*, *CD300A*, and *NFKBID*) (Figure 5D). The downregulation of B cell-
217 specific markers suggests differences in lymphocyte composition and/or trafficking in males.
218 Furthermore, the reduction in NK cell activating receptors and upregulation of negative
219 regulators of immune effector function, and resultant throttling of effector function, is
220 consistent with a more severe manifestation of COVID-19 in males.

221

222 **Discussion**

223 One of the hallmarks of COVID-19 is a dysregulated antiviral immune response. Studies
224 of SARS-CoV, which also employs *ACE2/TMPRSS2*-mediated entry, have demonstrated infection
225 does not always result in production of interferon- β in macrophages and dendritic cell (35), and
226 significantly delayed expression of type II or III interferon in lung cells (10). Moreover, infection
227 of BALB/c mice with SARS-CoV did not result in detectable IFN β until 24 hours, at which point
228 viral titers had nearly reached a peak; lung damage resulting from the subsequent massive
229 infiltration of inflammatory macrophages could be abrogated by pre-treatment with type I
230 interferons (11). Similar viral kinetics have been observed in SARS-CoV-2-infected patients
231 (13,14), and ferrets (12). Collectively, these results support a common mechanism by which

232 SARS-CoV and SARS-CoV-2 suppress intracellular viral detection and subsequent interferon
233 induction long enough for viral replication to occur.

234 Our transcriptomic analysis of nasopharyngeal swabs reveals the robust induction of an
235 interferon response by SARS-CoV-2 infection (Figure 1), similar to that observed by Butler et al
236 (20). The highest levels of individual interferon-responsive genes were seen in samples with the
237 highest viral load (Figure 2A) and enriched for transcripts associated with inflammatory
238 macrophages and activated DCs and NK cells, three primary sources of type I and II interferons
239 (Figure 2B). When repeat swabs were taken from patients with an average 6.3 day time period
240 between sampling, the interferon response had waned, as had viral load (Figure 4C). Notably, in
241 contrast to the robust expression of interferon-regulated transcripts in HAE seven days after
242 infection with SARS-CoV-2, there was limited evidence of induction of an interferon response
243 after only three days, consistent with a SARS-CoV-like functional repression of interferon
244 signaling.

245 COVID-19 patients frequently develop interleukin-6-driven cytokine release syndrome
246 (CRS), and elevated serum IL-6 correlates with respiratory failure and poor clinical outcomes
247 (36). Treatment with the IL-6 receptor blocking antibody tocilizumab has effectively treated
248 COVID-19 symptoms in some patients (37). We did not see a significant difference in expression
249 of *IL-6*, nor of other CRS-associated factors such as *TNF* or *VEGF*, when we analyzed SARS-CoV-2
250 positive samples relative to negative, nor in high vs low viral load SARS-CoV-2 positive samples.
251 This could be attributed to the nasopharynx not being a particularly sensitive anatomic location
252 to probe markers of systemic inflammation compared to serum or lower respiratory sites.
253 Additionally, our choice to use a large number of samples at relatively low sequencing depth

254 likely reduced our sensitivity to detect differences in low-abundance and short-lived transcripts
255 like cytokines.

256 One of the more striking patterns we observed is the marked downregulation of
257 transcription of ribosomal proteins upon SARS-CoV-2 infection (Figure 1B), and the recovery of
258 expression during disease progression (Figure 4C). Global inhibition of host transcription is a
259 strategy employed by many viruses via diverse mechanisms such as disrupting transcriptional
260 pre-initiation complex assembly (38,39) or cleavage of TATA-binding protein (40). MERS-CoV
261 and SARS-CoV nsp1 both cause decay of host mRNA (41,42); in MERS-CoV, host mRNA
262 degradation results from an endonucleolytic function of nsp1 itself (43). Nsp1 from SARS-CoV
263 and SARS-CoV-2 share 84% amino acid identity, therefore it is likely that SARS-CoV-2 nsp1 can
264 also function directly or indirectly to promote host RNA degradation. Global downregulation of
265 host transcription may also be driven in part by SARS-CoV-2 ORF6 protein, which binds to the
266 mRNA export factor RAE1 and nuclear pore protein Nup98 in a similar manner as the VSV M
267 protein (44), or the ORF7a protein, which binds to proteins involved in ribosomal assembly and
268 nuclear export (44).

269 Metagenomic analysis of SARS-CoV-2 positive samples revealed a low rate of viral co-
270 infection (3.25%), consistent with 3.2% reported by Butler et al in New York City (20). This is
271 likely due to a dramatic reduction in circulating respiratory viruses in March and April 2020
272 caused by physical distancing measures across the country. Among the samples from which we
273 could obtain high quality metagenomic data, we found 46 of 413 SARS-CoV positive samples
274 but only 1 of 37 SARS CoV negative samples extensively colonized (RPM>1000) by potentially
275 pathogenic bacteria (Supplementary Figure 2B), although no pattern of infection was found

276 based on viral load. More work is required to understand how these normalized read count
277 thresholds for resident nasopharyngeal microbiota correlate with transitions from colonization
278 to mild illness to potentially invasive disease.

279 Finally, understanding age- and sex-related differences in responses to SARS-CoV-2
280 infection is of critical importance as approximately 90% of SARS-CoV-2 deaths in Washington
281 State have been seen in individuals over 60. Our data show that in individuals over 60,
282 expression of interferon-induced chemokines is reduced, possibly contributing to a reduction in
283 transcripts for cytotoxic T and NK cells. Immune dysfunctions in older individuals are well-
284 characterized (2,3,33,34), and likely contribute to poorer COVID-19 outcomes; results from
285 clinical trials of type I and III interferons in severely ill patients are likely to further define the
286 role of interferon signaling in older adults (45–47).

287 Differences in immune responses in males and females are due to a variety of factors,
288 including the effects of sex hormones and the X-linked nature of many immune genes (48). The
289 bias towards expression of B cell transcripts in females in our study is consistent with higher
290 levels of B cells in females regardless of age (49). Females also tend to have increased
291 inflammation in response to viral infections (4). The observed increased expression of
292 inhibitors of NFkB in males with SARS-CoV-2 may represent either inappropriate throttling of
293 the antiviral immune response or an adaptive mechanism to reduce deleterious inflammation,
294 a hallmark of COVID-19 pathogenesis.

295 Collectively, we demonstrate induction of an antiviral response characterized by type I
296 and II interferon induction, which wanes with time and is correlated with viral load. We also
297 find evidence of transcriptional repression by SARS-CoV-2. Lastly, we show that differences in

298 immune responses may underlie disparities in outcomes for two higher risk groups, males and
299 the elderly.

300

301 **Methods**

302 **IRB Approval:** Sequencing of excess clinical samples was approved by the University of
303 Washington IRB (STUDY00000408).

304 **Sample Collection, RNA extraction, and qPCR:** NP swabs of patients with suspected SARS-CoV-
305 2 infection were collected in 3 mL viral transport medium (VTM). Total RNA was extracted from
306 200 or 140 μ L of VTM using either the Roche MagNAPure or Qiagen BioRobot automated
307 platforms, respectively (50). Quantitative PCR for the SARS-CoV2 N1 target was performed on
308 the Applied Biosystems 7500 real time PCR instrument (51,52).

309 **Library preparation and sequencing:** Metagenomic next-generation sequencing (mNGS) was
310 performed as previously described (17,53). Briefly, 18 μ L of extracted RNA was treated with
311 Turbo DNase (ThermoFisher). First strand cDNA synthesis was completed using SuperScript IV
312 (ThermoFisher) and random hexamers (Invitrogen) followed by second strand synthesis by
313 Sequenase V2.0 (ThermoFisher). The resulting cDNA was purified using either the DNA Clean &
314 Concentrator kit (Zymo) or 1.6x volumes of AMPure XP beads (Beckman Coulter). Library
315 preparation was performed using the Nextera XT Kit (Illumina). Libraries were cleaned with 0.7x
316 or 0.75x volumes of Ampure beads (Beckman Coulter), quantified using either the Qubit dsDNA
317 HS assay (ThermoFisher) or Quant-iT dsDNA HS assay (ThermoFisher), quality checked by

318 Bioanalyzer or TapeStation (Agilent), pooled, and sequenced on 1 x 75 bp runs on an Illumina
319 NextSeq or 1 x 101 bp runs on an Illumina NovaSeq.

320 **Pseudoalignment:** Raw FASTQ files were adapter and quality trimmed by Trimmomatic v0.39
321 (54) using the call “leading 3 trailing 3 slidingwindow:4:15 minlen 20”. Trimmed reads were
322 pseudoaligned to the Ensembl v96 human transcriptome using Kallisto v0.46 (55) assuming an
323 average library size of 300+/-100 base pairs. Only samples with more than 500,000
324 pseudoaligned reads were used for RNAseq analysis.

325 **Differential Expression:** Pseudoaligned reads were pre-filtered to remove any genes with
326 average expression of less than one read per sample, then normalized and differential
327 expression calculated with the R package DESeq2 v1.28.1 (56). Correction for batch effects was
328 incorporated into the design formula and modeling performed using the Wald test with outlier
329 replacement. Results were deemed significant at a Benjamini-Hochberg adjusted pvalue <0.1.
330 Gene expression differences attributable to sex or age were incorporated into the design
331 formula as interaction terms.

332 **Gene Set Enrichment Analysis (GSEA):** GSEA was performed on normalized counts on GSEA
333 Software version 4.0.3 (21,22). Gene ranking was generated with the Signal2Noise metric and
334 analyzed against the mSigDB Hallmarks v7.1 gene sets (23).

335 **Metagenomics:** Metagenomic analysis of the RNA sequence was performed using CLOMP
336 v0.1.4 (17) with the default options and visualized using the Pavian metagenomic explorer (57).
337 Viral species level taxonomical classifications with an RPM greater than 25 were confirmed via
338 BLAST v2.10.1 (e-value 1e-5).

339 **Human Airway Epithelial (HAE) cultures.** The EpiAirway AIR-100 system (MatTek Corporation)
340 consists of normal human-derived tracheo/bronchial epithelial cells that have been cultured to
341 form a pseudostratified, highly differentiated mucociliary epithelium closely resembling that of
342 epithelial tissue in vivo. Upon receipt from the manufacturer, HAE cultures were transferred to
343 6-well plates containing 1.0 ml EpiAirway medium per well (basolateral feeding, with the apical
344 surface remaining exposed to air) and acclimated at 37°C in 5% CO₂ for 24 hours prior to
345 experimentation.

346 **Viral growth in HAE.** HAE cultures were infected by applying 200 µl of EpiAirway phosphate-
347 buffered saline (MatTek TEER Buffer) containing 2,000 PFU or 20,000 PFU of infectious clone-
348 derived SARS-CoV-2 expressing a stable mNeonGreen reporter gene (icSARS-CoV-2-mNG) (58)
349 to the apical surface for 90 min at 37°C. At 90 min, the medium containing the inoculum was
350 removed, the apical surface was washed with 200 µl of TEER buffer, and cultures were placed at
351 37°C. Cultures were fed every other day with 1.0 ml medium via the basolateral side. Media
352 was removed, and cultures were lysed with TRIzol Reagent (ThermoFisher) at three days post
353 infection (20,000 PFU challenge) and at 7 days post infection (2,000 PFU challenge). Bam files of
354 viral sequence are deposited in the sequence read archive, NCBI Bioproject PRJNA634194.

355 **HAE RNAseq and analysis:** RNA from uninfected and infected HAE was extracted using Direct-
356 zol RNA MicroPrep (Zymo). Libraries were generated using the TruSeq Stranded mRNA kit
357 (Illumina) and 2x100bp paired-end reads sequenced on a Novaseq. Pseudoalignment using
358 Kallisto v0.44 and differential expression analysis was performed as above.

359 **Statistics and visualization:** All calculations were performed in R v4.0.0. Statistical enrichment
360 tests of Gene Ontology (24,25) and DisGeNET (28) pathways were performed in the

361 clusterProfiler R package (59). Images were generated using packages including DOSE (60),
362 ggplot2, pheatmap, and VennDiagram.

363 **Data Availability:** Raw counts and metadata for each nasopharyngeal sample is deposited in
364 the NCBI Gene Expression Omnibus GSE152075.

365

366 **Acknowledgements:**

367 The authors thank Amin Addetia and Joshua Lieberman for their thoughtful comments to
368 improve the manuscript. This work was supported by funding from the National Institutes of
369 Health (AI146980, AI121349, and NS091263 to MP) and the Department of Laboratory
370 Medicine at the University of Washington School of Medicine.

371 **References**

- 372 1. OpenSAFELY: factors associated with COVID-19-related hospital death in the linked
373 electronic health records of 17 million adult NHS patients. | medRxiv [Internet]. [cited
374 2020 Jun 5]. Available from:
375 <https://www.medrxiv.org/content/10.1101/2020.05.06.20092999v1>
- 376 2. Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older
377 people? *Aging*. 2020 May 29;12(10):9959–81.
- 378 3. McElhaney JE, Verschoor CP, Andrew MK, Haynes L, Kuchel GA, Pawelec G. The immune
379 response to influenza in older humans: beyond immune senescence. *Immun Ageing*. 2020
380 May 7;17(1):10.
- 381 4. vom Steeg LG, Klein SL. Sex Matters in Infectious Disease Pathogenesis. *PLoS Pathog*. 2016
382 Feb;12(2):e1005374.
- 383 5. Channappanavar R, Fett C, Mack M, Ten Eyck PP, Meyerholz DK, Perlman S. Sex-based
384 differences in susceptibility to SARS-CoV infection. *J Immunol Baltim Md 1950*. 2017 May
385 15;198(10):4046–53.
- 386 6. Meng Y, Wu P, Lu W, Liu K, Ma K, Huang L, et al. Sex-specific clinical characteristics and
387 prognosis of coronavirus disease-19 infection in Wuhan, China: A retrospective study of
388 168 severe patients. *PLOS Pathog*. 2020 Apr 28;16(4):e1008520.

- 389 7. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-
390 2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease
391 Inhibitor. *Cell*. 2020 Apr 16;181(2):271-280.e8.
- 392 8. Ziegler CGK, Allon SJ, Nyquist SK, Mbanjo IM, Miao VN, Tzouanas CN, et al. SARS-CoV-2
393 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is
394 Detected in Specific Cell Subsets across Tissues. *Cell*. 2020 May;181(5):1016-1035.e19.
- 395 9. Hou YJ, Okuda K, Edwards CE, Martinez DR, Asakura T, Dinnon KH, et al. SARS-CoV-2
396 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell*. 2020
397 May;S0092867420306759.
- 398 10. Menachery VD, Einfeld AJ, Schäfer A, Josset L, Sims AC, Proll S, et al. Pathogenic influenza
399 viruses and coronaviruses utilize similar and contrasting approaches to control interferon-
400 stimulated gene responses. *mBio*. 2014 May 20;5(3):e01174-01114.
- 401 11. Channappanavar R, Fehr AR, Vijay R, Mack M, Zhao J, Meyerholz DK, et al. Dysregulated
402 Type I Interferon and Inflammatory Monocyte-Macrophage Responses Cause Lethal
403 Pneumonia in SARS-CoV-Infected Mice. *Cell Host Microbe*. 2016 Feb 10;19(2):181–93.
- 404 12. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, Uhl S, Hoagland D, Møller R, et al. Imbalanced
405 Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. 2020
406 28;181(5):1036-1045.e9.
- 407 13. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and
408 transmissibility of COVID-19. *Nat Med*. 2020 May;26(5):672–5.

- 409 14. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 Viral Load in Upper
410 Respiratory Specimens of Infected Patients. *N Engl J Med*. 2020 Mar 19;382(12):1177–9.
- 411 15. Greninger AL, Chen EC, Sittler T, Scheinerman A, Roubinian N, Yu G, et al. A metagenomic
412 analysis of pandemic influenza A (2009 H1N1) infection in patients from North America.
413 *PloS One*. 2010 Oct 18;5(10):e13381.
- 414 16. Fauver JR, Petrone ME, Hodcroft EB, Shioda K, Ehrlich HY, Watts AG, et al. Coast-to-Coast
415 Spread of SARS-CoV-2 during the Early Epidemic in the United States. *Cell*. 2020
416 28;181(5):990-996.e5.
- 417 17. Peddu V, Shean RC, Xie H, Shrestha L, Perchetti GA, Minot SS, et al. Metagenomic analysis
418 reveals clinical SARS-CoV-2 infection and bacterial or viral superinfection and colonization.
419 *Clin Chem*. 2020 May 7;
- 420 18. Randhawa AK, Fisher LH, Greninger AL, Li SS, Andriesen J, Corey L, et al. Changes in SARS-
421 CoV-2 Positivity Rate in Outpatients in Seattle and Washington State, March 1-April 16,
422 2020. *JAMA*. 2020 May 8;
- 423 19. Bedford T, Greninger AL, Roychoudhury P, Starita LM, Famulare M, Huang M-L, et al. Cryptic
424 transmission of SARS-CoV-2 in Washington State. *medRxiv*. 2020 Apr
425 16;2020.04.02.20051417.
- 426 20. Butler DJ, Mozsary C, Meydan C, Danko D, Foon J, Rosiene J, et al. Host, Viral, and
427 Environmental Transcriptome Profiles of the Severe Acute Respiratory Syndrome
428 Coronavirus 2 (SARS-CoV-2). *bioRxiv*. 2020 Apr 20;2020.04.20.048066.

- 429 21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set
430 enrichment analysis: a knowledge-based approach for interpreting genome-wide
431 expression profiles. *Proc Natl Acad Sci U S A*. 2005 Oct 25;102(43):15545–50.
- 432 22. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-
433 responsive genes involved in oxidative phosphorylation are coordinately downregulated in
434 human diabetes. *Nat Genet*. 2003 Jul;34(3):267–73.
- 435 23. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular
436 Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015 Dec
437 23;1(6):417–25.
- 438 24. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool
439 for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000
440 May;25(1):25–9.
- 441 25. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing
442 strong. *Nucleic Acids Res*. 2019 08;47(D1):D330–8.
- 443 26. Hubel P, Urban C, Bergant V, Schneider WM, Knauer B, Stukalov A, et al. A protein-
444 interaction network of interferon-stimulated genes extends the innate immune system
445 landscape. *Nat Immunol*. 2019 Apr;20(4):493–502.
- 446 27. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell
447 type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*.
448 2019 Jul;37(7):773–82.

- 449 28. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The
450 DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* 2020
451 08;48(D1):D845–55.
- 452 29. Meyers L, Ginocchio CC, Faucett AN, Nolte FS, Gesteland PH, Leber A, et al. Automated
453 Real-Time Collection of Pathogen-Specific Diagnostic Data: Syndromic Infectious Disease
454 Epidemiology. *JMIR Public Health Surveill.* 2018 Jul 6;4(3):e59.
- 455 30. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological
456 assessment of hospitalized patients with COVID-2019. *Nature.* 2020;581(7809):465–9.
- 457 31. Liu Y, Yan L-M, Wan L, Xiang T-X, Le A, Liu J-M, et al. Viral dynamics in mild and severe cases
458 of COVID-19. *Lancet Infect Dis.* 2020;20(6):656–7.
- 459 32. Boukhaled GM, Cordeiro B, Deblois G, Dimitrov V, Bailey SD, Holowka T, et al. The
460 Transcriptional Repressor Polycomb Group Factor 6, PCGF6, Negatively Regulates
461 Dendritic Cell Activation and Promotes Quiescence. *Cell Rep.* 2016 16;16(7):1829–37.
- 462 33. Hazeldine J, Lord JM. The impact of ageing on natural killer cell function and potential
463 consequences for health in older adults. *Ageing Res Rev.* 2013 Sep;12(4):1069–78.
- 464 34. Li M, Yao D, Zeng X, Kasakovski D, Zhang Y, Chen S, et al. Age related human T cell subset
465 evolution and senescence. *Immun Ageing.* 2019 Sep 11;16(1):24.
- 466 35. Frieman M, Baric R. Mechanisms of Severe Acute Respiratory Syndrome Pathogenesis and
467 Innate Immunomodulation. *Microbiol Mol Biol Rev MMBR.* 2008 Dec;72(4):672–85.

- 468 36. Chen G, Wu D, Guo W, Cao Y, Huang D, Wang H, et al. Clinical and immunological features
469 of severe and moderate coronavirus disease 2019. *J Clin Invest*. 2020 May 1;130(5):2620–
470 9.
- 471 37. Xu X, Han M, Li T, Sun W, Wang D, Fu B, et al. Effective treatment of severe COVID-19
472 patients with tocilizumab. *Proc Natl Acad Sci*. 2020 May 19;117(20):10970–5.
- 473 38. Dasgupta A, Scovell WM. TFIIA abrogates the effects of inhibition by HMGB1 but not E1A
474 during the early stages of assembly of the transcriptional preinitiation complex. *Biochim*
475 *Biophys Acta*. 2003 Jun 19;1627(2–3):101–10.
- 476 39. Di Valentin E, Bontems S, Habran L, Jolois O, Markine-Goriaynoff N, Vanderplasschen A, et
477 al. Varicella-zoster virus IE63 protein represses the basal transcription machinery by
478 disorganizing the pre-initiation complex. *Biol Chem*. 2005 Mar;386(3):255–67.
- 479 40. Kundu P, Raychaudhuri S, Tsai W, Dasgupta A. Shutoff of RNA polymerase II transcription by
480 poliovirus involves 3C protease-mediated cleavage of the TATA-binding protein at an
481 alternative site: incomplete shutoff of transcription interferes with efficient viral
482 replication. *J Virol*. 2005 Aug;79(15):9702–13.
- 483 41. Lokugamage KG, Narayanan K, Nakagawa K, Terasaki K, Ramirez SI, Tseng C-TK, et al. Middle
484 East Respiratory Syndrome Coronavirus nsp1 Inhibits Host Gene Expression by Selectively
485 Targeting mRNAs Transcribed in the Nucleus while Sparing mRNAs of Cytoplasmic Origin. *J*
486 *Virol*. 2015 Nov 1;89(21):10970–81.

- 487 42. Kamitani W, Narayanan K, Huang C, Lokugamage K, Ikegami T, Ito N, et al. Severe acute
488 respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by
489 promoting host mRNA degradation. *Proc Natl Acad Sci*. 2006 Aug 22;103(34):12885–90.
- 490 43. Nakagawa K, Narayanan K, Wada M, Popov VL, Cajimat M, Baric RS, et al. The
491 Endonucleolytic RNA Cleavage Function of nsp1 of Middle East Respiratory Syndrome
492 Coronavirus Promotes the Production of Infectious Virus Particles in Specific Human Cell
493 Lines. *J Virol*. 2018 01;92(21).
- 494 44. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2
495 protein interaction map reveals targets for drug repurposing. *Nature*. 2020 Apr 30;1–13.
- 496 45. Park A, Iwasaki A. Type I and Type III Interferons – Induction, Signaling, Evasion, and
497 Application to Combat COVID-19. *Cell Host Microbe [Internet]*. 2020 May 27 [cited 2020
498 Jun 7]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7255347/>
- 499 46. Fragkou PC, Belhadi D, Peiffer-Smadja N, Moschopoulos CD, Lescure F-X, Janocha H, et al.
500 Review of trials currently testing treatment and prevention of COVID-19. *Clin Microbiol*
501 *Infect [Internet]*. 2020 May 23 [cited 2020 Jun 7]; Available from:
502 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7245266/>
- 503 47. Hung IF-N, Lung K-C, Tso EY-K, Liu R, Chung TW-H, Chu M-Y, et al. Triple combination of
504 interferon beta-1b, lopinavir–ritonavir, and ribavirin in the treatment of patients admitted
505 to hospital with COVID-19: an open-label, randomised, phase 2 trial. *The Lancet*. 2020 May
506 30;395(10238):1695–704.

- 507 48. Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol*.
508 2016;16(10):626–38.
- 509 49. Teixeira D, Longo-Maugeri IM, Santos JLF, Duarte YAO, Lebrão ML, Bueno V. Evaluation of
510 lymphocyte levels in a random sample of 218 elderly individuals from São Paulo city. *Rev*
511 *Bras Hematol E Hemoter*. 2011;33(5):367–71.
- 512 50. Lieberman JA, Pepper G, Naccache SN, Huang M-L, Jerome KR, Greninger AL. Comparison of
513 Commercially Available and Laboratory Developed Assays for in vitro Detection of SARS-
514 CoV-2 in Clinical Laboratories. *J Clin Microbiol*. 2020 Apr 29;
- 515 51. Nalla AK, Casto AM, Huang M-LW, Perchetti GA, Sampoleo R, Shrestha L, et al. Comparative
516 Performance of SARS-CoV-2 Detection Assays Using Seven Different Primer-Probe Sets and
517 One Assay Kit. *J Clin Microbiol*. 2020 26;58(6).
- 518 52. Perchetti GA, Nalla AK, Huang M-L, Zhu H, Wei Y, Stensland L, et al. Validation of SARS-CoV-
519 2 detection across multiple specimen types. *J Clin Virol Off Publ Pan Am Soc Clin Virol*.
520 2020 13;128:104438.
- 521 53. Greninger AL, Zerr DM, Qin X, Adler AL, Sampoleo R, Kuypers JM, et al. Rapid Metagenomic
522 Next-Generation Sequencing during an Investigation of Hospital-Acquired Human
523 Parainfluenza Virus 3 Infections. *J Clin Microbiol*. 2017;55(1):177–82.
- 524 54. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
525 *Bioinforma Oxf Engl*. 2014 Aug 1;30(15):2114–20.

- 526 55. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq
527 quantification. *Nat Biotechnol.* 2016;34(5):525–7.
- 528 56. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
529 seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
- 530 57. Breitwieser FP, Salzberg SL. Pavian: interactive analysis of metagenomics data for
531 microbiome studies and pathogen identification. *Bioinforma Oxf Engl.* 2020 Feb
532 15;36(4):1303–4.
- 533 58. Xie X, Muruato A, Lokugamage KG, Narayanan K, Zhang X, Zou J, et al. An Infectious cDNA
534 Clone of SARS-CoV-2. *Cell Host Microbe.* 2020 13;27(5):841-848.e3.
- 535 59. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological
536 themes among gene clusters. *Omics J Integr Biol.* 2012 May;16(5):284–7.
- 537 60. Yu G, Wang L-G, Yan G-R, He Q-Y. DOSE: an R/Bioconductor package for disease ontology
538 semantic and enrichment analysis. *Bioinforma Oxf Engl.* 2015 Feb 15;31(4):608–9.
- 539
- 540

541 **Figure Legends**

542

543 **Figure 1: Differentially expressed genes in SARS-CoV-2 nasopharyngeal swabs.** A) Clustering of
544 samples based on 50 genes with the lowest adjusted pvalue. Log2 fold changes relative to gene
545 mean are displayed by color. B) Volcano plot of 15 most upregulated and 15 most
546 downregulated genes in SARS-CoV-2 positive samples relative to negative by log2 fold change.
547 Red color indicates genes with log2 fold change $> |1.5|$ and adjusted pvalue < 0.05 . C).
548 Significant (FDR < 0.05) pathways affected by SARS-CoV-2 infection identified by Gene Set
549 Enrichment Analysis.

550 **Figure 2: Differences in gene expression by SARS-CoV-2 viral load.** A) Violin plots of select
551 genes by viral load. Statistical significance between low and high viral load calculated by Mann
552 Whitney U test, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. B) Volcano plot of 15
553 most upregulated and 15 most downregulated genes in SARS-CoV-2 high viral load samples
554 relative to low viral load by log2 fold change. Red color indicates genes with log2 fold change $>$
555 $|1.5|$ and adjusted pvalue < 0.05 . C) Proportion of cell types as a total of all immune cells, by
556 CIBERSORTx. Significant differences in proportion of each cell type determined by T test, $*p <$
557 0.05 , $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. D) Violin plots of B cell transcripts and
558 neutrophil chemokine transcripts by viral load. Statistical significance between low and high
559 viral load calculated by Mann Whitney U test, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p <$
560 0.0001 .

561 **Figure 3: Consensus genes induced upon SARS-CoV-2 expression.** A) Venn diagram of DE genes
562 in SARS-CoV-2 positive vs negative, high vs low viral load, and top 100 genes with the highest

563 absolute log₂ fold change in infected vs uninfected HAE. Consensus set of 19 genes DE in all
564 three analyses are shown, with log₂ fold change values relative to uninfected HAE (for day 3
565 and day 7 post infection), SARS-CoV-2 negative NP swabs (for SARS-CoV-2 positive NP swabs),
566 or low viral load (for high SARS-CoV-2 viral load samples). SARS-CoV-2 reads at day 3 and 7 post
567 infection were 0.3% and 5.3%, respectively. B) Top 20 DisGeNET terms for which SARS-CoV-2
568 cell-intrinsic antiviral response consensus genes are overrepresented. Number Enriched is the
569 number of SARS-CoV-2 consensus genes that belong to each disease term. C) Interaction
570 network of SARS-CoV-2 consensus genes for top 5 most similar diseases identified in B. Size of
571 disease node represents the number of genes enriched, and fold change is the log₂ fold change
572 seen in SARS-CoV-2 positive vs negative NP swabs.

573 **Figure 4: Differentially expressed genes in patient-matched longitudinal samples.** A) Patient
574 demographics information for longitudinal samples. B) Top 20 Biological Process Gene Ontology
575 terms for which longitudinal DE genes are overrepresented. Number Enriched is the number of
576 DE genes that belong to each GO Term. C) Log₂ fold changes for DE genes in Humoral Immune
577 Response and Wound Healing GO Terms, consensus antiviral SARS-CoV-2 genes, and ribosomal
578 proteins. Grey bars: padj < 0.1, white bars: padj > 0.1.

579 **Figure 5: Age and sex cause differences in gene expression upon SARS-CoV-2 infection.** A) N1
580 Ct values by age group. No significant differences between were observed by Kruskal-Wallis
581 ANOVA. B) N1 Ct values by sex. No significant difference between groups was observed by t-
582 test. C) Gene expression differences by age and viral load. Significance by Mann Whitney U test
583 between SARS-CoV-2 positive samples age >60 and <60 is shown, *p < 0.05, **p < 0.01, ***p <
584 0.001, ****p < 0.0001. D) Sex-modulated DE genes (padj < 0.1) upon SARS-CoV-2 infection.

585 Genes elevated in females are shown as negative log2 fold changes, and those elevated in
586 males as positive log2 fold changes.

587 **Supplementary Figure 1: Differentially expressed Gene Sets and Gene Ontology Biological**

588 **Process Terms in SARS-CoV-2 nasopharyngeal swabs.** A) Enrichment plots of gene sets
589 significantly (FDR<0.05) positively enriched in SARS-CoV-2 samples. B) Enrichment plots of gene
590 sets significantly (FDR<0.05) negatively enriched in SARS-CoV-2 samples. C) Top 20 Biological
591 Process Gene Ontology terms for which differentially expressed genes in SARS-CoV-2 samples
592 are overrepresented. Number Enriched is the number of SARS-CoV-2 differentially expressed
593 genes that belong to each GO Term D) Fold change of genes belonging to GO Term “defense
594 response to virus”. E) Fold change of genes belonging to GO Term “SRP-dependent
595 cotranslational protein targeting to membrane”.

596 **Supplementary Figure 2: Metagenomic analyses of SARS-CoV-2 positive and negative**

597 **samples.** A) Loess-smoothed area plot showing the proportion of human, viral, and bacterial
598 reads for each sample. Positive samples are arranged in reverse order of N1 Ct. B) Colonization
599 and co-infection of non-SARS-CoV-2 respiratory viruses and clinically relevant bacterial species.
600 C) Violin plot of bacterial RPM after correcting by subtraction of SARS-CoV-2 reads.

601 **Supplementary Figure 3: Age and sex differences in gene expression upon SARS-CoV-2**

602 **expression.** A) Gene expression differences by age and viral load. Significance by Mann Whitney
603 U test between SARS-CoV-2 positive samples age >60 and <60 is shown, *p < 0.05, **p < 0.01,
604 ***p < 0.001, ****p < 0.0001. B) Top 10 Biological Process Gene Ontology terms in which genes
605 defining the male vs female response to virus are overrepresented.

606

607 **Supplementary Table 1:** Differentially expressed genes in SARS-CoV-2 positive samples relative
608 to negative.

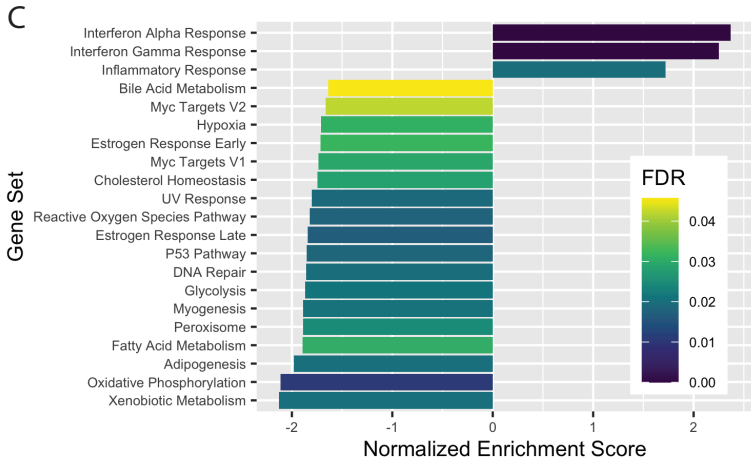
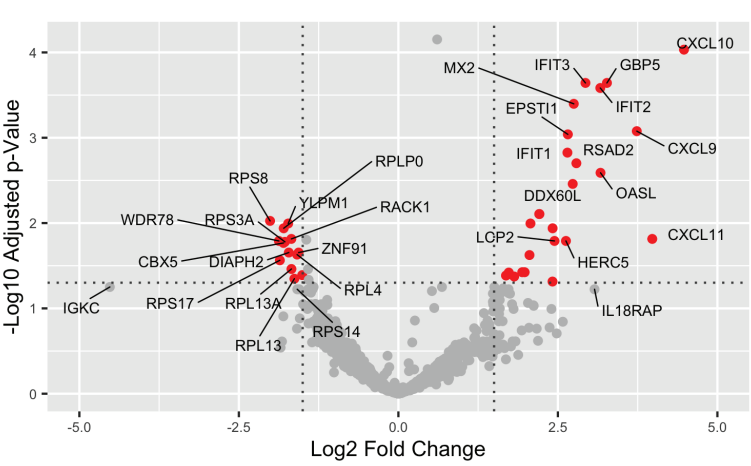
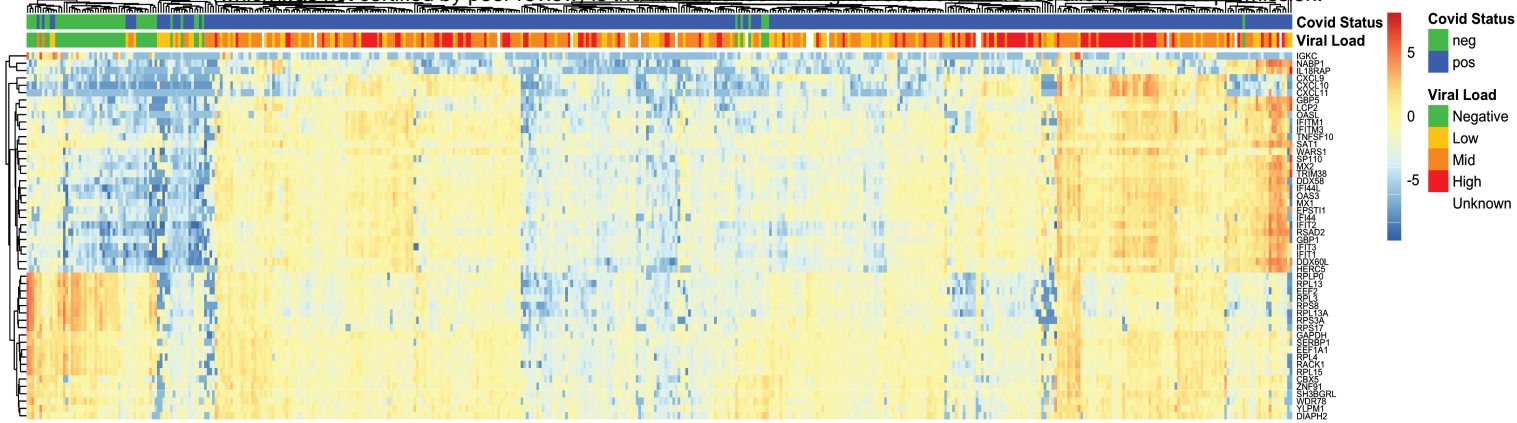
609 **Supplementary Table 2:** Differentially expressed genes in SARS-CoV-2 high viral load samples
610 relative to low viral load.

611

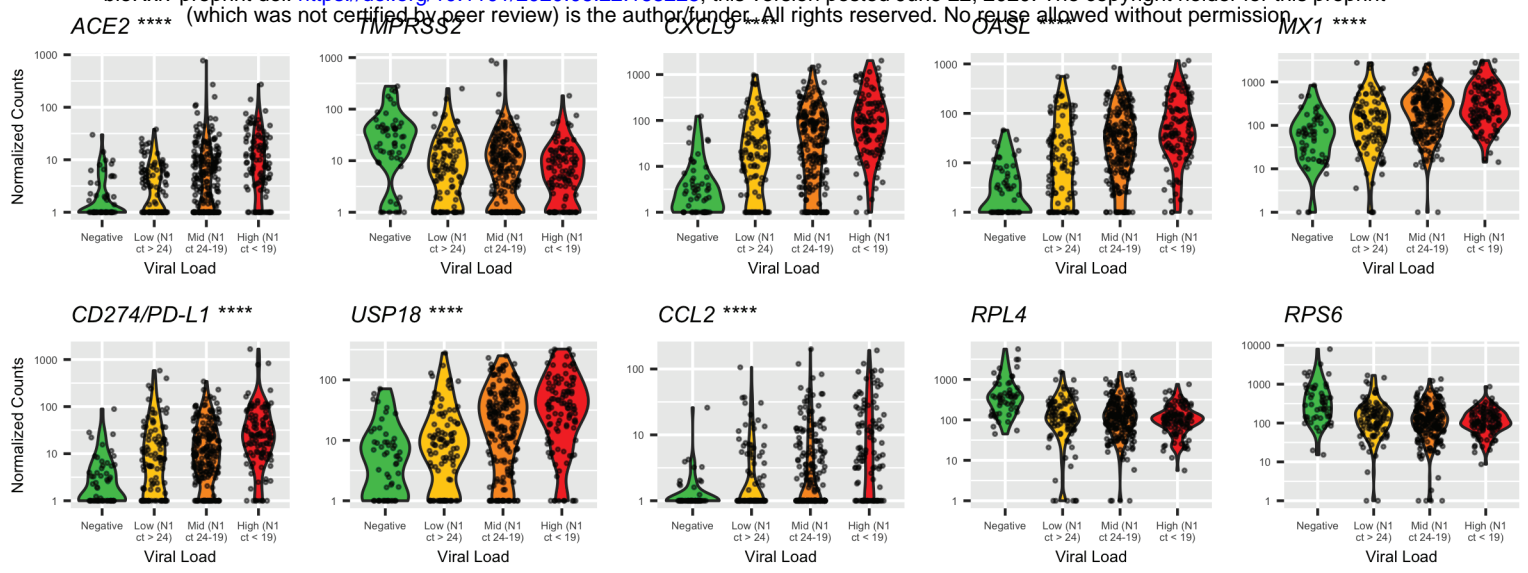
Table 1. Patient demographics of SARS-CoV-2 positive and negative samples

| SARS-CoV-2 Status | Viral Load | Total Number | Sex | | | Age (yr) | | N1 Ct | |
|-------------------|----------------|--------------|------|--------|---------|----------|-------|-------|-------------|
| | | | Male | Female | Unknown | Median | Range | Mean | Range |
| Positive | | 430 | 176 | 201 | 53 | 54 | 2-98 | 21.21 | 12.32-30.54 |
| | Low | 99 | 45 | 50 | 4 | 59 | 2-96 | 25.64 | 24.00-30.54 |
| | Medium | 206 | 90 | 99 | 17 | 56 | 12-98 | 21.33 | 19.08-23.99 |
| | High | 108 | 41 | 52 | 15 | 52 | 16-97 | 16.92 | 12.32-18.93 |
| | Unknown | 17 | Unk | Unk | 17 | Unk | Unk | Unk | Unk |
| Negative | n/a | 54 | 30 | 24 | 0 | 46.5 | 12-90 | n/a | n/a |

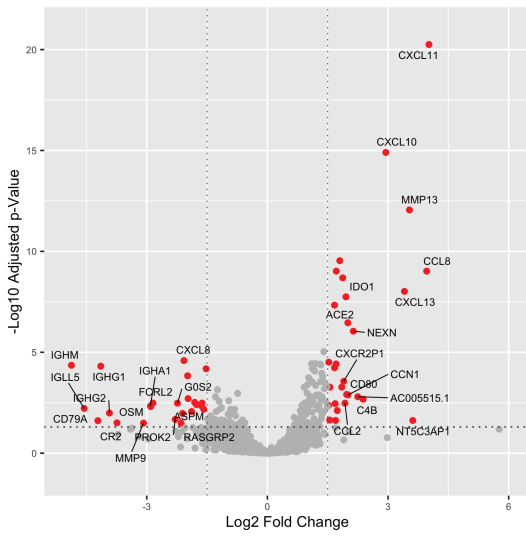
Unk: Unknown, n/a: not applicable



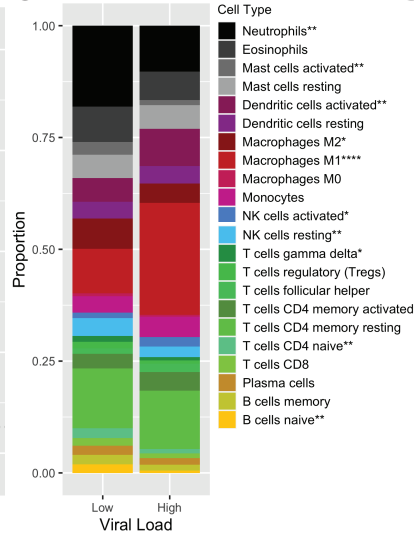
A



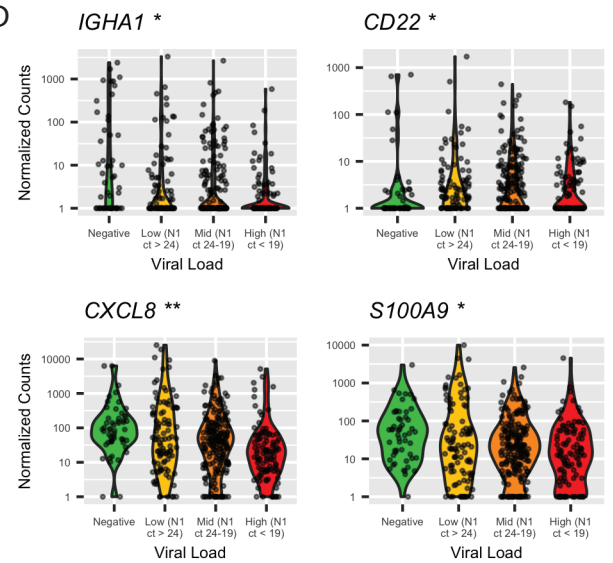
B



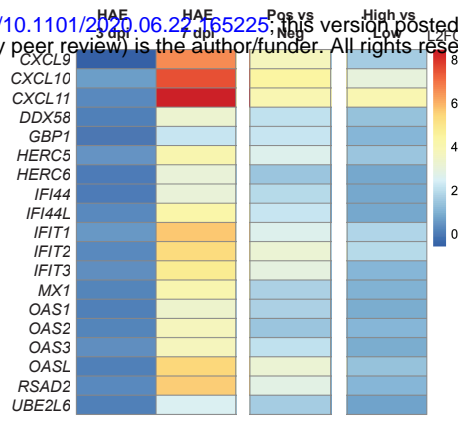
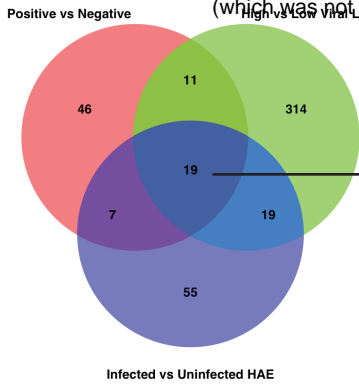
C



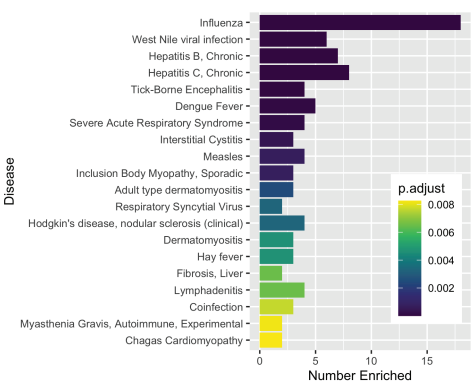
D



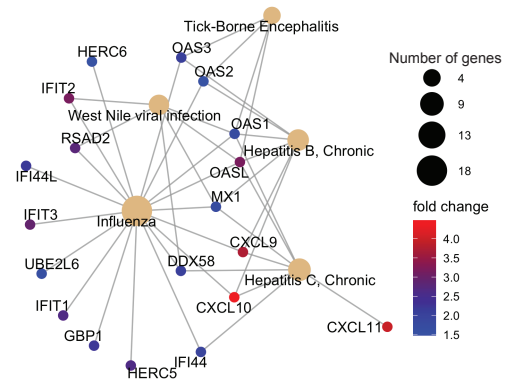
A



B



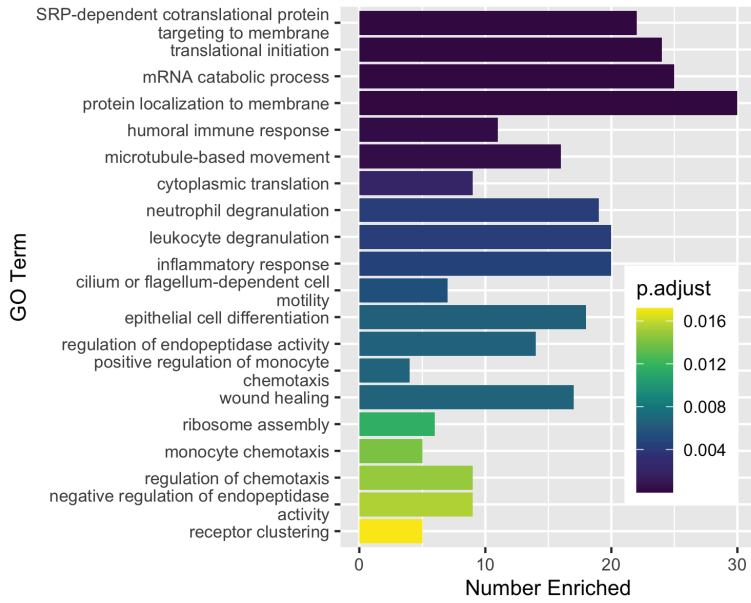
C



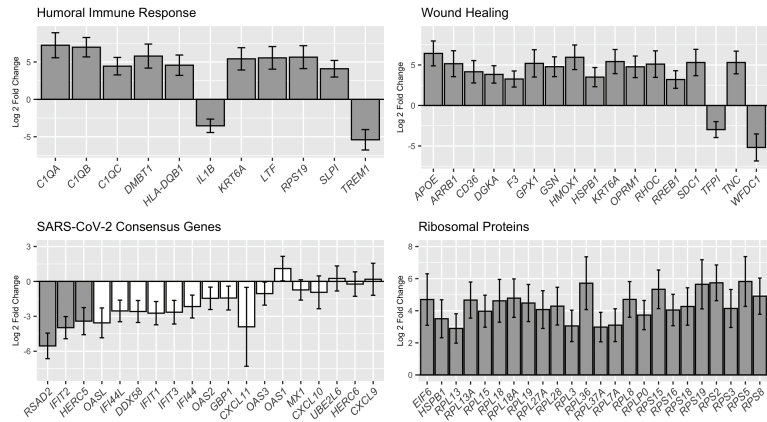
Patient Age Sex N1 Ct #1 N1 Ct #2 Days Elapsed

| Patient | Age | Sex | N1 Ct #1 | N1 Ct #2 | Days Elapsed |
|---------|-----|-----|----------|----------|--------------|
| 1 | 49 | M | 17.7 | 26.4 | 5 |
| 2 | 51 | M | 16.7 | 24.8 | 6 |
| 3 | 78 | M | 25.4 | 24.6 | 8 |

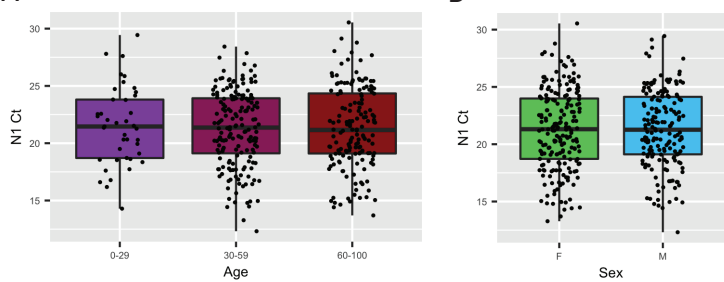
B



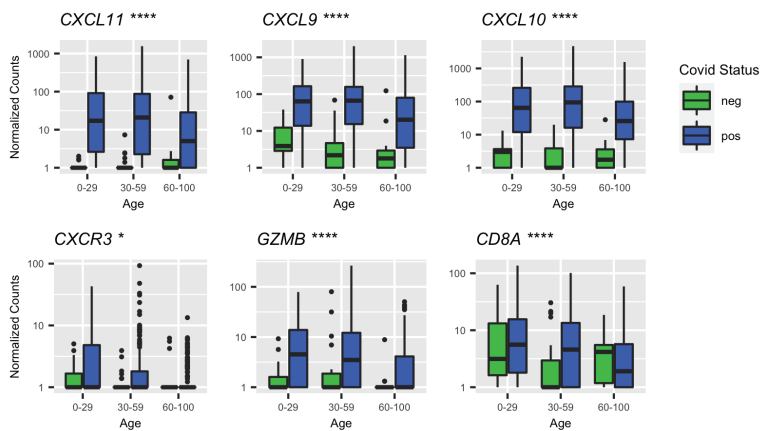
C



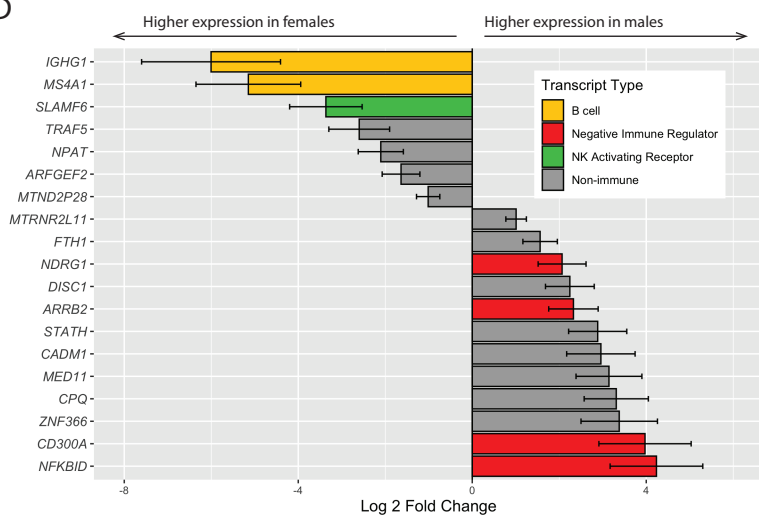
A

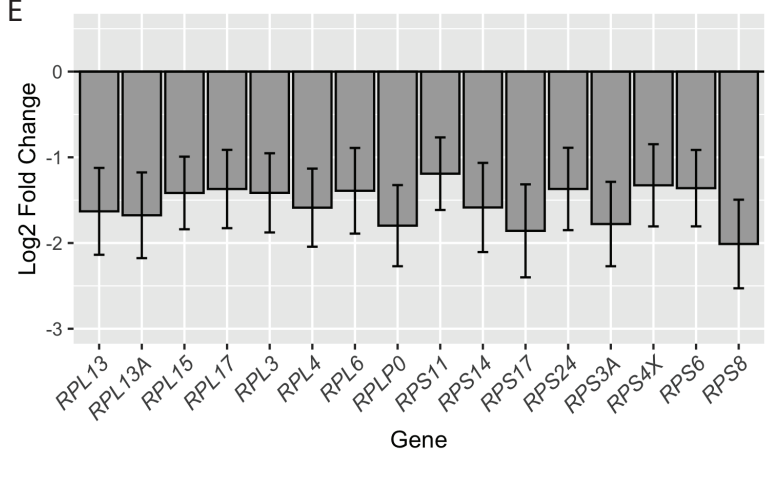
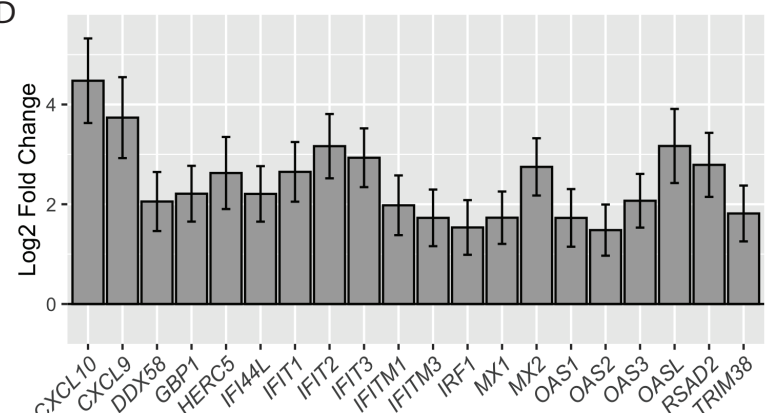
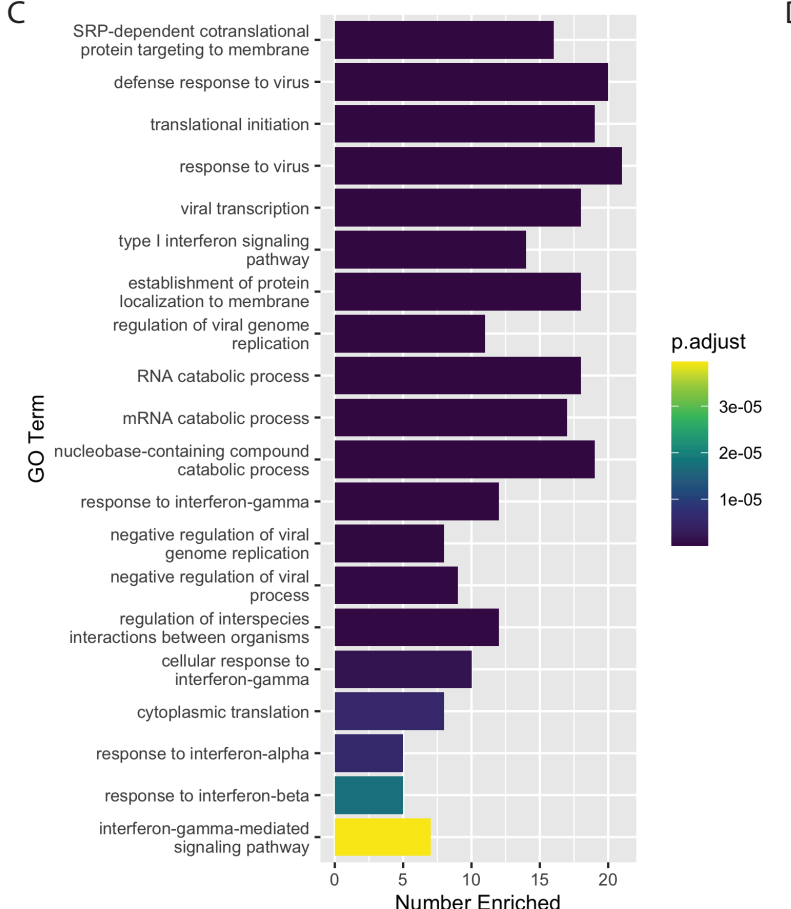
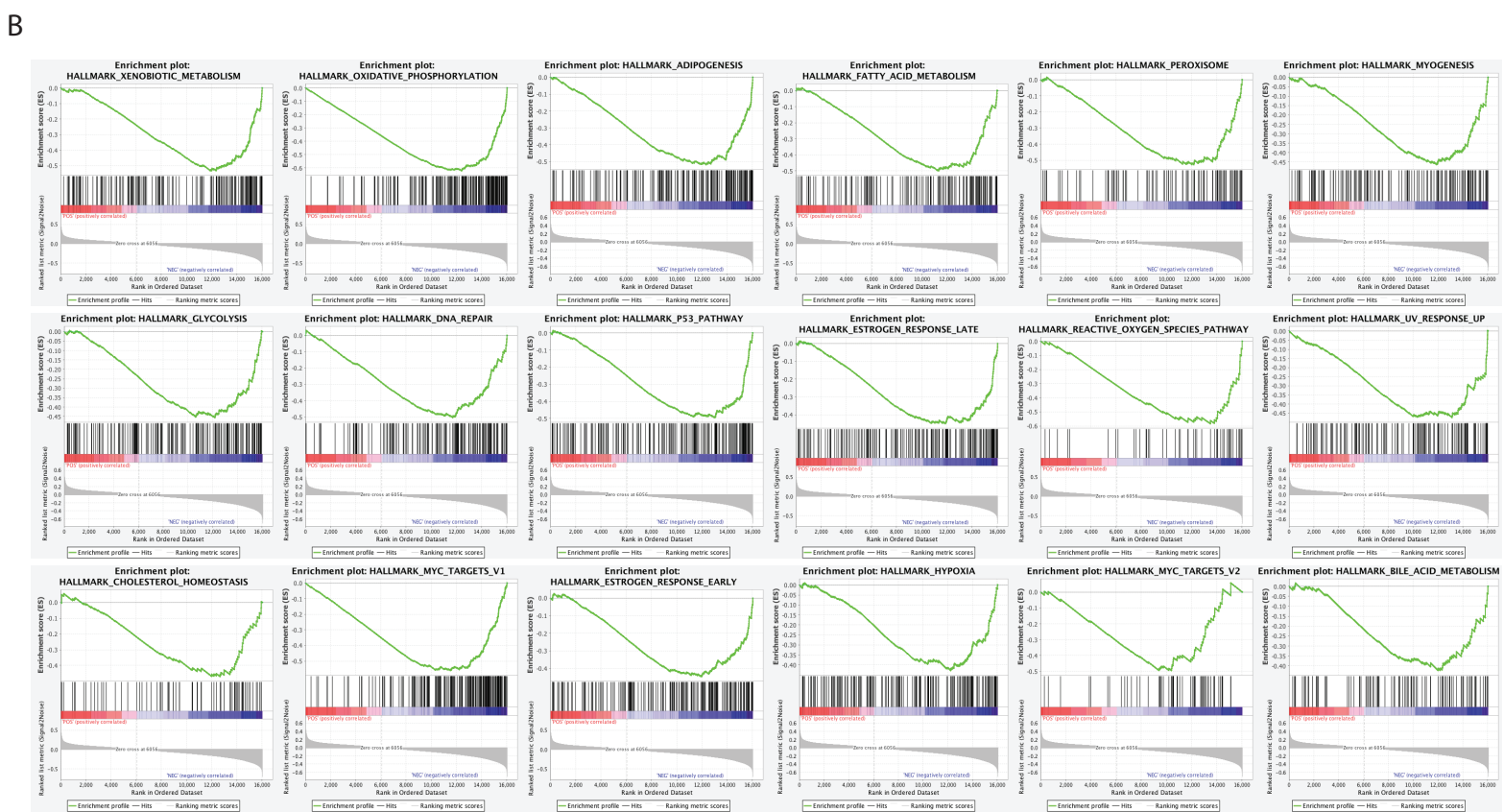
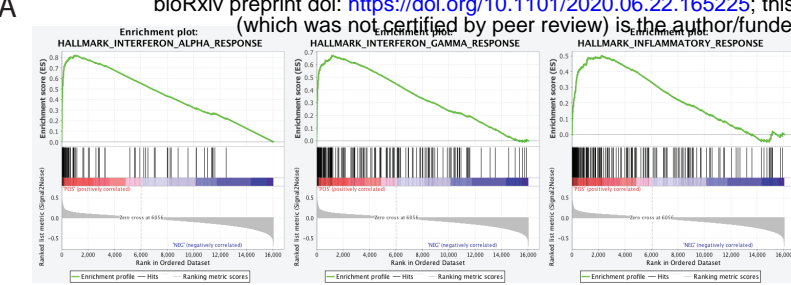


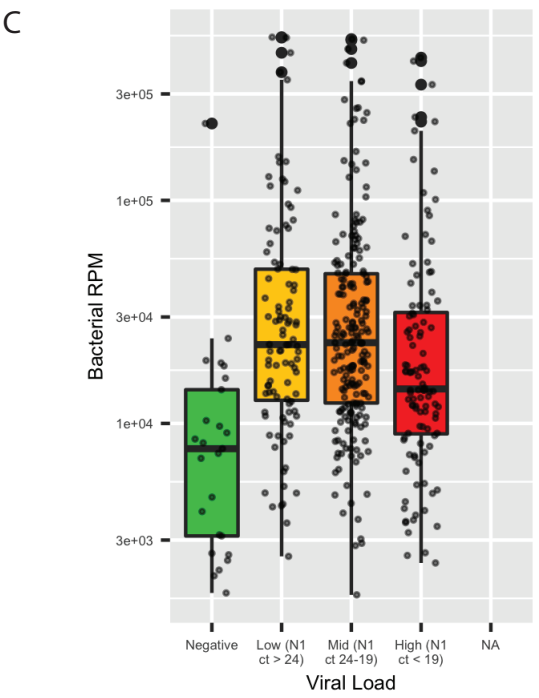
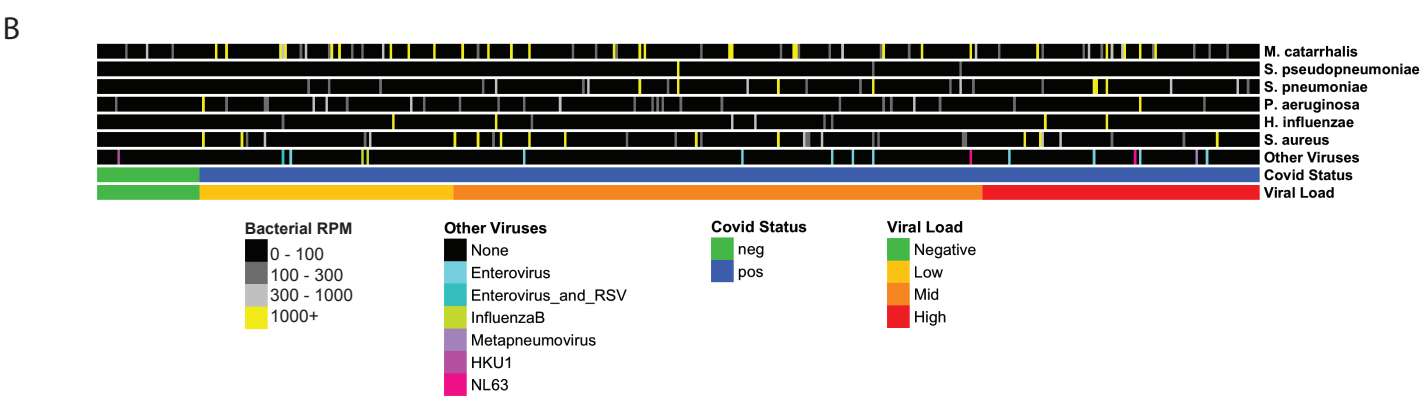
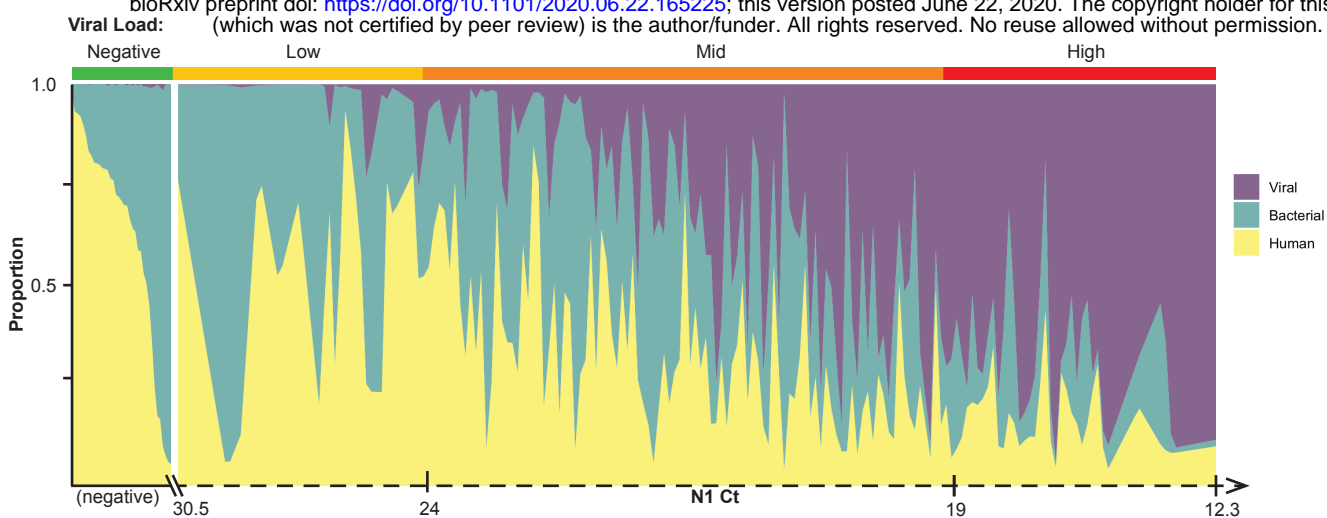
C



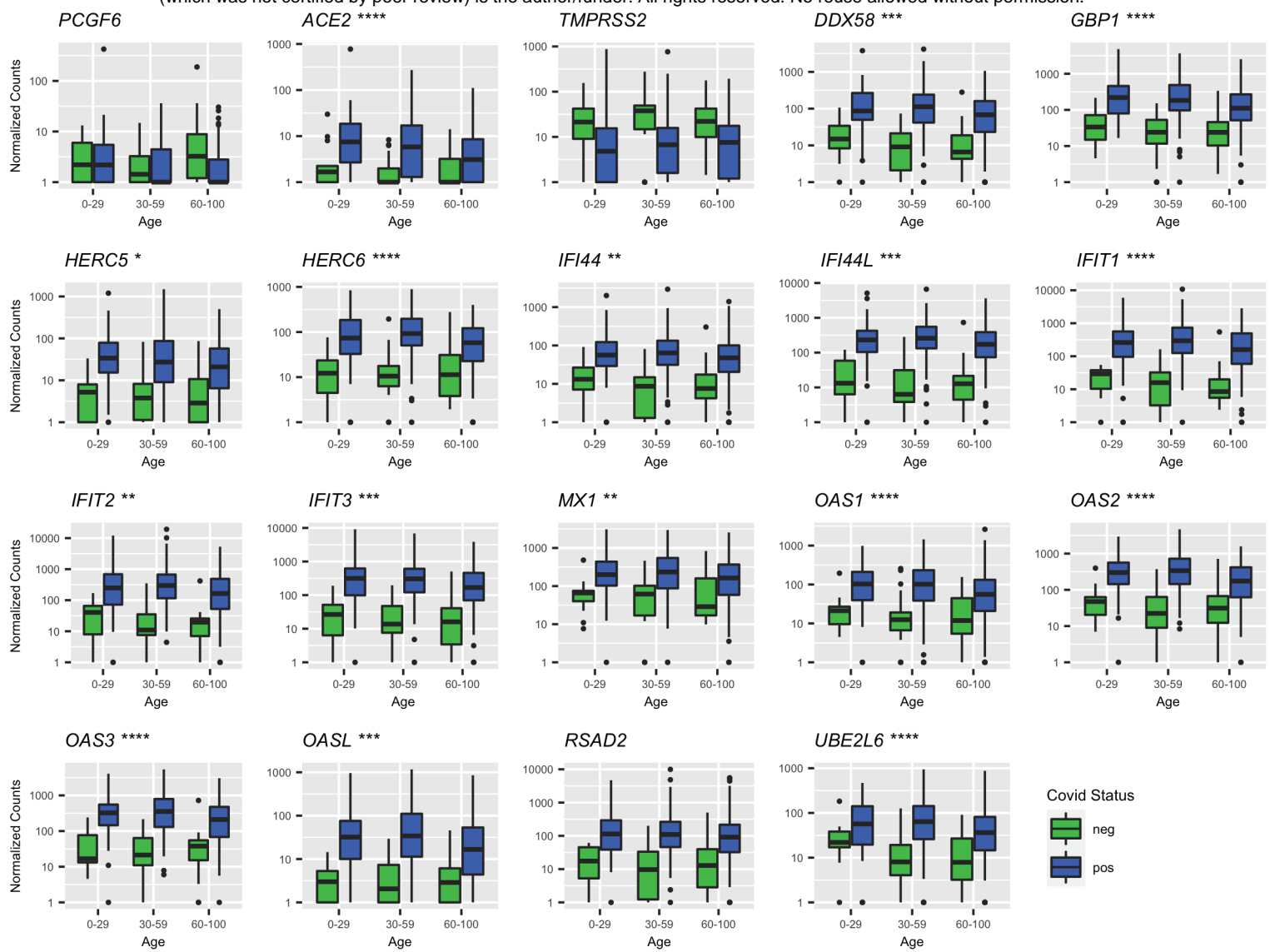
D







A



B

