

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61

Abstract

Coronaviruses remodel intracellular membranes to form specialized viral replication compartments, such as double-membrane vesicles where viral RNA genome replication takes place. Understanding the factors affecting host response is instrumental to design of therapeutics to prevent or ameliorate the course of infection.

As part of explorative tests in hospitalized patients with confirmed COVID-19 infection participating in ODYSSEY trial, we obtained samples for whole genome sequencing analysis as well as for viral genome sequencing. Based on our data, we confirm one of the strongest severity susceptibility locus thus far reported in association with severe COVID-19: 3p21.31 locus with lead variant rs73064425. We further examine the associated region. Interestingly based on LD analysis we report 3 coding mutations within one gene in the region of FYVE and Coiled-Coil Domain Autophagy Adaptor 1 (*FYCO1*). We specifically focus on the role of *FYCO1* modifiers and gain-of-function variants. We report the associations between the region and clinical characteristics in this severe set of COVID-19 patients.

We next analyzed expression profiles of *FYCO1* across all 466 compounds tested. We selected only those results that showed a significant reduction of expression of *FYCO1*. The most significant candidate was indomethacin – an anti-inflammatory that could potentially downregulate *FYCO1*. We hypothesize that via its direct effects on efficiency of viral egress, it may serve as a potent therapeutic decreasing the replication and infectivity of the virus. Clinical studies will be needed to examine the therapeutic utility of indomethacin and other compounds downregulating *FYCO1* in COVID-19 infection and other strains of betacoronaviruses.

Keywords: COVID-19, therapeutics, whole genome sequencing, viral genetics, blocking viral egress

62 **Introduction**

63 The emergence of highly pathogenic zoonotic viruses calls for further research in order to
64 understand, treat and ideally prevent such infections ultimately circumventing epidemics.

65 Currently accurate prognoses are confounded by our inability to predict disease severity and
66 virulence across space and time. Virus-human interactions have driven around 30% of the human
67 genome evolution since divergence from chimpanzee. Different variants, genes affect severity to
68 a range of infections, for example: *IFITM3*, *IRF7* in influenza, *SFPA/D* and *VTR* in RSV, *CCR5*
69 in HIV, *IFNG* in HTVL-1 and *TNF* in HPV to name a few. The complex interplay between
70 environmental, viral, and host genetic factors drives differences in individual disease
71 progression, severity, and outcome results. Understanding resilience and susceptibility factors
72 should enable one to design prophylactics and therapeutics.

73

74 SARS-CoV-2 is a novel coronavirus that has led to a worldwide pandemic¹. There is a
75 high mortality rate associated with the virus and despite measures implemented to contain the
76 spread of the virus, over 2.1 million deaths were reported worldwide (1/26/21). The ACE2 binding
77 affinity with RBD on the Spike protein of SARS-CoV-2 is 10-20x higher than that of the SARS-
78 CoV. The clinical spectrum of this disease is heterogeneous, and infected individuals viral RNA
79 shedding pattern differ across individuals at baseline. Similarly to other positive-sense RNA
80 viruses, coronaviruses remodel intracellular membranes to form specialized viral replication
81 compartments. Depending on the time course of the infection, the structure, composition, and
82 formation of virus replication organelles appear to be varied and dynamic. Understanding the
83 mode of action of this virus in the context of its interaction with the host genome is fundamental
84 to the design of optimal therapeutic strategies.

85 In the early months of the epidemic it became apparent that advanced age and
86 comorbidities are associated with higher risk for severe infection yet none of these fully explain
87 the heterogeneous course of the infection amongst individuals². To that end several sequencing
88 projects have been commenced since the beginning of the epidemic. Initial results from COVID-
89 19 initiative suggest a locus on chromosome 3p21.31, the peak association with severity of the
90 disease, signal that has been observed by the Severe Covid-19 GWAS Group and further
91 replicated by other host genome sequencing consortia^{3,4,5} and the UKbiobank⁶. The genetic
92 variants on chromosome 3 that are most associated with severe COVID-19 are all in high linkage

93 disequilibrium (LD). Furthermore recent phylogenetic analysis showed that the risk haplotype
94 thus entered the modern human population from Neanderthals (Vindija 33.19 Neanderthal)⁷.
95 Specifically, we focused on this region is because the 3p21.31 regional severity association was
96 the strongest one replicated by multiple consortia including the COVID-19 initiative and latest
97 Nature report by Casteneira et al., (rs73064425, OR=2.14, discovery $p=4.77 \times 10^{-30}$) (top
98 variants can be accessed in the supplementary material)

99
100 As part of explorative tests in hospitalized patients with confirmed SARS-CoV-2 positive
101 infection participating in ODYSSEY trial, we obtained blood samples for whole genome
102 sequencing analysis as well as nasopharyngeal swabs for further viral genome sequencing. We
103 have conducted whole genome sequencing on first cohort participants as well as deep clinical
104 phenotyping. Based on our data, we confirm one of the strongest susceptibility locus reported
105 and now replicated with lead variant being rs73064425. We further examine the associated
106 region, with specific focus on the role of *FYCO1* gene. We further discuss the associations
107 between the region and clinical characteristics in this severe set of COVID-19 patients.

108
109
110

111 **Results**

112 ***Regional association***

113 We replicate the strongest susceptibility locus reported and now replicated is within, *3p21.31*
114 *locus*, rs73064425 Leucine Zipper Transcription Factor Like 1 (LZTFL1) reported by Casteneira
115 et al., as the strongest association with severity (rs73064425, OR=2.14, discovery $p=4.77 \times 10^{-}$
116 30). We replicate this association in our cohort of hospitalized patients (OR 3.2, CI
117 1.8728 to 5.4637, p -value<0.0001). The variant has a global gnomAD MAF of 0.05. We report
118 an increased MAF 0.19 in ODYSSEY, and consistently with gnomAD, we report a MAF of 0.05
119 in Vanda 2000 controls. The highest allelic frequency for this variant globally is present in
120 Ashkenazi Jewish (MAF 0.12) and the lowest frequency is present amongst East and South
121 Asian populations (MAF 0.0006). The GWAS signal covers a cluster of six genes (*SLC6A20*,
122 *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, and *XCRI*), several of which with functions that could be
123 potentially relevant to Covid-19. The variant is a strong eQTL for CXCR6 and SLC6A20 in
124 GTEX⁸. The locus itself is within a highly regulatory region, as an enhancer for Primary B cells
125 from peripheral blood and modifier H3K4me1 in accordance with Haploreg. Furthermore the
126 locus is associated with monocyte percentage (UK Biobank field: 30190-0.0) in Gene Atlas.
127

128 Next we reconstructed LD to test if the signal (non-coding, top SNP by p -value) from
129 other large studies can be reproduced and extended to nearby genes, Where are the “limits” of
130 the signal on Odyssey data, Are there any coding variants .Interestingly there are 3 coding
131 mutations within one gene in the region (*FYCO1*) causing 2 amino acid substitutions (Table1;
132 two mutations within same codon, 1001, resulting in one missense mutation, third in strong LD).
133 All 3 coding variants are in strong LD with top SNPs published previously within the region and
134 seem to be the strongest signal for our Odyssey samples. Having ~2000 control genomes we can
135 see that all 3 coding variants form a strong LD. Quintessentially the risk is conveyed by the
136 extent of existing LD structure in that region as shown in Supplementary Table1. Specifically,
137 rs33910087 is an eQTL for CXCR6 and based on GeneAtlas, the genotyped variant is highly
138 significant modifier of monocyte percentage ($p=3.8479e-46$).

139
140 The association of the locus with severity of infection in interaction with SARS-CoV-2
141 remains to be examined. Functionally *FYCO1* encodes a protein involved in vesicle transport and

142 autophagy. It has been suggested as a key mediator linking ER-derived double membrane
 143 vesicles, the primary replication site for coronaviruses, with the microtubule network⁹. *FYCO1*
 144 through its LC3-interacting region (LIR) motif was shown to be important for the fusion of
 145 autophagosomes with lysosomes. *FYCO1* dimerizes via the CC region, interacts with PI3P via
 146 its FYVE domain, and forms a complex with Rab7 via a part of the CC region located in front of
 147 the FYVE domain⁹. Specifically, *FYCO1* was shown to act as a Rab7 effector that binds to LC3
 148 and PI3P to mediate microtubule plus end-directed vesicle transport¹⁰. The depletion of Rab7
 149 inhibits maturation of late endosomes/MVBs and leads to reduced lysosome numbers in cells¹¹.
 150 *FYCO1* also mediates clearance of α -synuclein aggregates through a Rab7-dependent
 151 mechanism, overexpression reducing number of cell with α -synuclein aggregates¹². In another
 152 recent report authors perform a genome-scale CRISPR loss-of-function screen in human alveolar
 153 basal epithelial carcinoma cells to identify genes whose loss enabled resistance to SARS-CoV-2
 154 viral infection¹³. Loss of *RAB7A* reduces viral entry/egress by sequestering the ACE2 receptor
 155 inside cells¹³. Furthermore, depletion of *FYCO1* or antibodies against the N-terminus of LC3
 156 blocks the subcellular redistribution of autophagosomes¹⁴. Indeed, rare *FYCO1* variants some of
 157 which contain missense mutations in the LIR domain have recently been associated with
 158 inclusion body myositis a disease characterized by impaired autophagic degradation.

159

CHR	BP	SNP	A1	MAF.Covid19	MAF.Controls	Pval	OR	Ref	Alt	Gene	Function	AAChange	RSID
3	46007825	3:46007825[b37]C,T	C	22.3%	8.3%	7.01E-06	3.168	T	C	FYCO1	nonsynonymous	c.A3001G:p.N1001D	rs13059238
3	46007823	3:46007823[b37]G,T	T	21.4%	7.8%	7.84E-06	3.22	G	T	FYCO1	nonsynonymous	c.C3003A:p.N1001K	rs13079478
3	46009487	3:46009487[b37]A,G	A	21.4%	8.1%	1.34E-05	3.104	G	A	FYCO1	nonsynonymous	c.C1339T:p.R447C	rs33910087
3	45901089	3:45901089[b37]C,T	T	17.0%	5.3%	1.19E-05	3.628	C	T	LZTFL1	intronic	.	rs73064425

160

161 *Table1. Top coding variants in association with severity of COVID19 infection.*

162 We stipulate gain of function variants in *FYCO1* confer higher risk of severe course of
 163 COVID-19 induced infection and likely other betacoronaviruses. Temporarily downregulating
 164 *FYCO1* seems to be protective, hence the locus is likely crucial to designing therapeutic
 165 strategies for this and other strain of betacoronaviruses.

166

167 ***FYCO1 and drug-gene expression***

168 We have conducted a high throughput drug screen gene expression analysis to identify
169 compounds that would downregulate the expression of *FYCO1*. To discover potential,
170 pharmaceutical agents capable of affecting transcriptional expression levels of *FYCO1*
171 implicated in SARS-CoV and SARS-CoV2 patho-physiology, we have screened 466 compounds
172 belonging to 14 different therapeutic classes. Screening was conducted using human retinal
173 pigment epithelia cell line (ARPE-19) and gene expression changes were collected across 12,490
174 genes. We analyzed the expression profiles of *FYCO1* across all 466 compounds tested. In order
175 to find positive hits we selected only those results that showed a reduction of expression of
176 *FYCO1* (1.5 -fold difference). The top significant candidate indomethacin is (displayed in **Table**
177 **2**) a drug that has both anti-inflammatory and antiviral actions. Indomethacin was previously
178 shown to have potent direct antiviral activity against the coronaviruses SARS-CoV and CCoV.
179 Its potent antiviral activity (>1,000-fold reduction in virus yield) was previously confirmed in
180 vivo in CCoV-infected dogs (Amici). These screens warrant further confirmatory tests
181 nevertheless the strategy to weaken viral egress could be potentially useful for strains beyond
182 SARS-CoV-2

183

<i>Compound</i>	FYCO1 Log2(treated)	FYCO1 Log2(Ctrl)	FYCO1 Log2(diff)
<i>Indomethacin</i>	4.94	6.48	1.53
<i>Primidone</i>	6.44	7.94	1.49
<i>TriprolidineHydrochloride</i>	6.49	7.98	1.49
<i>Baclofen</i>	6.72	7.94	1.21

184

185 **Table 2. Top significant candidates**

186

187 **Clinical Characteristics of carriers**

188 Both heterozygotes and homozygotes were severe hospitalized patients treated for COVID-19.
189 The signal is one of severity however we wanted to check if the variant carriers would have any
190 particular clinical characteristics. To that end we tested systemic cytokine panel in addition to
191 standard lab tests. Clinical characteristics across genotypes are displayed in **Supplementary**
192 **Material**. Noticeable is the linear distribution with higher sodium and chloride seen in
193 homozygous individuals. All study patients met inclusion criteria included: age 18-90; confirmed

194 laboratory COVID-19 infection; confirmed pneumonia by chest radiograph or computed
195 tomography; fever defined as temperature ≥ 36.6 °C armpit, ≥ 37.2 °C oral, or ≥ 37.8 °C rectal
196 since admission or the use of antipyretics; PaO₂ / FiO₂ ≤ 300 ; 6. in-patient hospitalization. We
197 furthermore sequenced the viral genome sequence data in conjunction with host sequencing. We
198 report coding mutations in SARS-CoV-2 based CoV-GLUE annotations¹⁵. All known variants
199 reported in CoV-Glue¹⁵ (until 11/23/2020) variants were annotated. As the viral strains evolved,
200 the 614th aa position of the S protein, aspartate became replaced by glycine ultimately becoming
201 the ubiquitous strain, hence making open conformations more likely. All patients are carriers of
202 the S:D614G variant - variant in the spike protein D614G that rapidly became dominant strain
203 according to latest GISAID data^{16,17}.

204

205 Discussion

206 Coronaviruses remodel intracellular membranes to form specialized viral replication
207 compartments, such as double-membrane vesicles where viral RNA genome replication takes
208 place. Understanding the formation and operation of these vesicles is instrumental to optimal
209 design of therapeutics to prevent or ameliorate the course of infection. Here via findings from
210 whole genome sequencing and association with severity of COVID-19, we were led to a locus
211 that may in fact be functionally relevant in predisposing individuals to severe course of infection
212 via its direct effects on efficiency of viral egress.

213

214 *FYCO1* has a direct link to formation of vesicles and autophagy. Pathogenic mutations in
215 *FYCO1* can affect intracellular transport of autophagocytic vesicles. Depletion of *FYCO1* has
216 been shown to block distribution of autophagosomes. *FYCO1* also mediates clearance of
217 α -synuclein aggregates through a Rab7-dependent mechanism, overexpression reducing
218 number of cell with α -synuclein aggregates. Recently using single cell RNAseq authors
219 identified a group of genes (*ATP6AP1*, *ATP6V1A*, *NPC1*, *RAB7A*, *CCDC22*, and *PIK3C3*)
220 whose knockout induced shared transcriptional changes in cholesterol biosynthesis pathway¹³.
221 They showed how the actual perturbation of the cholesterol biosynthesis pathway reduced viral
222 infection. That furthermore reveals the links via RAB7. We hypothesize fewer pLOFs in *NPC1*
223 to be present in cohorts of severe hospitalized patients with COVID-19. Niemann-Pick disease
224 type C1 lipid storage disorder offers resistance to Ebola in cell line experiments and in

225 homozygous recessive mice ((Npc1-/-))^{18,19}. Furthermore, bat species show selective sensitivity
226 to Ebola versus Marburg viruses²⁰. Our hypothesis is that people with certain lysosomal storage
227 diseases may be resistant to one of these viruses. Along these lines there is suggestive evidence
228 for this to be exactly the case. This is in line with a report by Sturley et al., authors propose that
229 50 loci that in the homozygous state cause a lysosomal storage disorder, may account for
230 divergent clinical outcomes in COVID-19.

231
232 Blocking or perturbing viral egress whether via Rab7, FYCO1 or other components may
233 be a viable strategy to stop the proliferation of the virus reducing viral load. In the case of
234 betacoronaviruses unconventional egress can be blocked by the Rab7 GTPase competitive
235 inhibitor CID1067700¹¹ - compound potentially decreasing viral egress in a dose-dependent
236 manner. Here we focused on downregulation of FYCO1: based on cell line expression with
237 resulting potential candidates including indomethacin - a drug that has both anti-inflammatory
238 and antiviral actions. Slowing viral spread by targeting regulators of lysosomal trafficking, by
239 focusing on egress and double membrane vesicles may offer a therapeutics strategy for novel
240 strain of betacoronaviruses.

241
242
243 The GWAS locus has led for further understanding of not only the more efficient egress
244 leading to successful propagation of the viral particles but also to potential protective
245 mechanisms and greater understanding of the heterogeneous clinical outcomes. We hence
246 stipulate gain of function variants in *FYCO1* confer higher risk of infection with SARS and other
247 betacoronaviruses. Downregulating or temporarily limiting, *FYCO1* activity may be protective.
248 This susceptibility locus in fact highlights the potentially druggable cellular aspects of the host.

249
250 **Methods**
251 **Clinical Trial Information**
252 ODYSSEY is a double-blinded Phase 3 study with a planned randomization of a total of 300
253 hospitalized severely ill COVID-19 patients to receive either tradipitant 85 mg bid or placebo for
254 a total of up to 14 days or discharge (CONSORT Flowchart in S. File 1). The randomization is
255 stratified by site with a block size of four. Inclusion criteria for the study comprised of: 1. Adults

256 aged 18-90; 2. confirmed laboratory COVID-19 infection; 3. confirmed pneumonia by chest
257 radiograph or computed tomography; 4. fever defined as temperature ≥ 36.6 °C armpit, ≥ 37.2 °C
258 oral, or ≥ 37.8 °C rectal since admission or the use of antipyretics; 5. PaO₂ / FiO₂ ≤ 300 ; 6. in-
259 patient hospitalization. Patients were to be followed for up to 28 days to record clinical
260 outcomes. Patients' clinical progress was recorded on a 7 point clinical status ordinal scale
261 defined as follows: 1- Death; 2- Hospitalized on mechanical ventilation or ECMO; 3-
262 Hospitalized on non-invasive ventilation or high-flow oxygen supplementation; 4- Hospitalized
263 requiring supplemental oxygen; 5- Hospitalized not requiring supplemental oxygen, requiring
264 continued medical care; 6- Hospitalized not requiring supplemental oxygen, not requiring
265 continued medical care; 7- Not hospitalized.

266

267 **Human whole genome sequencing**

268 **Vanda Pharmaceuticals Inc. COVID-19 Cohort:**

269 WGS cohort consisted of 80 COVID-19 hospitalized patient samples and 1876 WGS controls.
270 DNA was extracted from 200 μ l of whole blood using QIAamp blood mini kit and eluted into
271 100 μ l volume in order to obtain 500 ng of DNA. Libraries were prepared by random
272 fragmentation of DNA followed by adapter ligation using Illumina TruSeq adapters. Paired-end
273 libraries 2x150 bp were sequenced on an Illumina NovaSeq6000 S4 with a target of 90Gb raw
274 read depth per sample. The cohort was multi-ethnic and the age ranged from 35-87 for severe
275 hospitalized COVID-19 cases (68% male, 32% female). Specifically the PC defined ancestry
276 was as follows (African 9.8%, European 46.3%, Asian 4.9%, Hispanic 22% and other 17.1%).
277 Furthermore, 16% of the cohort were patients hospitalized on mechanical ventilation or ECMO,
278 13% deceased.

279 Both cases and controls were processed with the same bioinformatic pipeline for variant calling.
280 Paired-end 150 \times bp reads were aligned to the GRCh37 human reference (BWA-MEM v0.7.8)
281 followed by Picard (MarkDuplicates) and processed with GATK best-practices workflow
282 (GATK v3.5.0). The mean coverage was 35.8x. All high quality variants obtained from GATK
283 were annotated. Annotations include variant effect predictions using VEP; allele frequencies

284 from Gnomad;dbSNP 150 rsIDs; conservation scores from PhyloP, GERP, PhastCons; damaging
285 effect predictions from Polyphen 2, SIFT and clinically relevant information from ClinVar. We
286 removed samples with a discordance between self-declared and sequence-derived gender, we
287 used KING for detection of related individuals and performed PCA for ancestry adjustment. The
288 lead variant was in addition verified with genotyping (Supplementary Material). Logistic
289 regression was conducted in PLINK, MATLAB. The analysis was corrected for PC 1-15, age
290 and sex.

291 **Viral sequencing and genome assembly**

292 Sequencing was attempted on all samples with a positive RT-PCR assay result that had Ct <32
293 using either a metagenomic approach described previously²¹, via IDT probe-capture²², or by
294 amplicon sequencing with SWIFT library preparation. Libraries were sequenced on Illumina
295 MiSeq or NextSeq instruments using 1x150 or 1x75 runs respectively. Consensus sequences
296 were assembled using a custom bioinformatics pipeline [<https://github.com/proychou/hCoV19>]
297 adapted for SARS-CoV-2 from previous work^{22, 23}. Briefly, raw reads were trimmed to remove
298 adapters and low quality regions using BBDuk and a k-mer based filter was used to pull out
299 reads matching the reference sequence NC_045512. Filtered reads were *de novo* assembled using
300 SPAdes²⁴ and contigs were ordered against the reference using BWA-MEM. Gaps were filled by
301 remapping reads against the assembled scaffold and a consensus sequence was called from this
302 alignment using a custom script in R/Bioconductor.

303

304 **Cell culture and drug treatment**

305 Drugs screening was carried out, the same one as applied in our previous study²⁵. The retinal
306 pigment epithelia cell line, ARPE-19/HPV-16, was chosen to establish a database of drug
307 profiles because of its non-cancerous, human origin, with a normal karyotype. It can also be
308 easily grown as monolayer in 96-well plates. Compounds were obtained from Sigma (St. Louis,
309 MO) or Vanda Pharmaceuticals (Washington, DC). Cells were aliquoted on 96-well plates
310 (~2×10⁵ cells/well) and incubated for 24 h prior to providing fresh media with a drug, or the
311 drug vehicle (water, dimethyl sulfoxide, ethanol, methanol, or phosphate-buffered saline
312 solution). Drugs were diluted 1000 fold in buffered in Dulbecco's Modified Eagle Medium:
313 Nutrient Mixture F-12 (D-MEM/F-12) culture medium (Invitrogen, Carlsbad, CA) containing

314 nonessential amino acids and 110 mg/L sodium pyruvate. In these conditions, no significant
315 changes of pH were expected, which was confirmed by the monitoring of the pH indicator
316 present in the medium. A final 10 μ M drug concentration was chosen because it is believed to fit
317 in the range of physiological conditions²⁵. Microscopic inspection of each well was conducted at
318 the end of the treatment to discard any samples where cells had morphological changes
319 consistent with apoptosis. We also verified that the drug had not precipitated in the culture
320 medium.

321

322 **Gene expression**

323 Cells were harvested 24 h after treatment and RNA was extracted using the RNeasy 96 protocol
324 (Qiagen, Valencia, CA). Gene expression for 22,238 probe sets of 12,490 genes was generated
325 with U133A2.0 microarrays following the manufacturer's instructions (Affymetrix, Santa Clara,
326 CA). Drugs were profiled in duplicate or triplicate, with multiple vehicle controls on each plate.
327 A total of 708 microarrays were analyzed including 74 for the 18 antipsychotics, 499 for the
328 other 448 compounds, and 135 for vehicle controls. The raw scan data were first converted to
329 average difference values using MAS 5.0 (Affymetrix). The average difference values of both
330 treatment and control data were set to a minimum of 50 or lower. For each treatment category,
331 all probe sets were then ranked based on their amplitude, or level of expression relative to the
332 vehicle control (or the average of controls was selected when more than one was used).
333 Amplitude was defined as the ratio of expression $(t-v) / [(t+v) / 2]$ where t corresponds to
334 treatment instance and v to vehicle instance. Each drug group profile was created using our novel
335 Weighted Influence Model, Rank of Ranks (WIMRR) method which underscores the rank of
336 each probe set across the entire gene expression profile rather than the specific change in
337 expression level. WIMRR takes the average rank of each probe set across all of the members of
338 the group and then reranks the probe sets from smallest average rank to largest average rank. A
339 gene-set enrichment metric based on the Kolmogorov– Smirnov (KS) statistic. Specifically, for a
340 given set of probes, the KS score gives a measure of how up (positive) or down (negative) the set
341 of probes occurs within the profile of another treatment instance.

342

343 **Declarations**

344 **Funding:** Vanda Pharmaceuticals Inc.

345 **Competing Interests:** SPS,BPP, CP. VP MHP are employees of Vanda Pharmaceuticals.

346 **Ethical Approval:** Reviewed and approved by Advarra IRB; Pro00043096

347

348

349

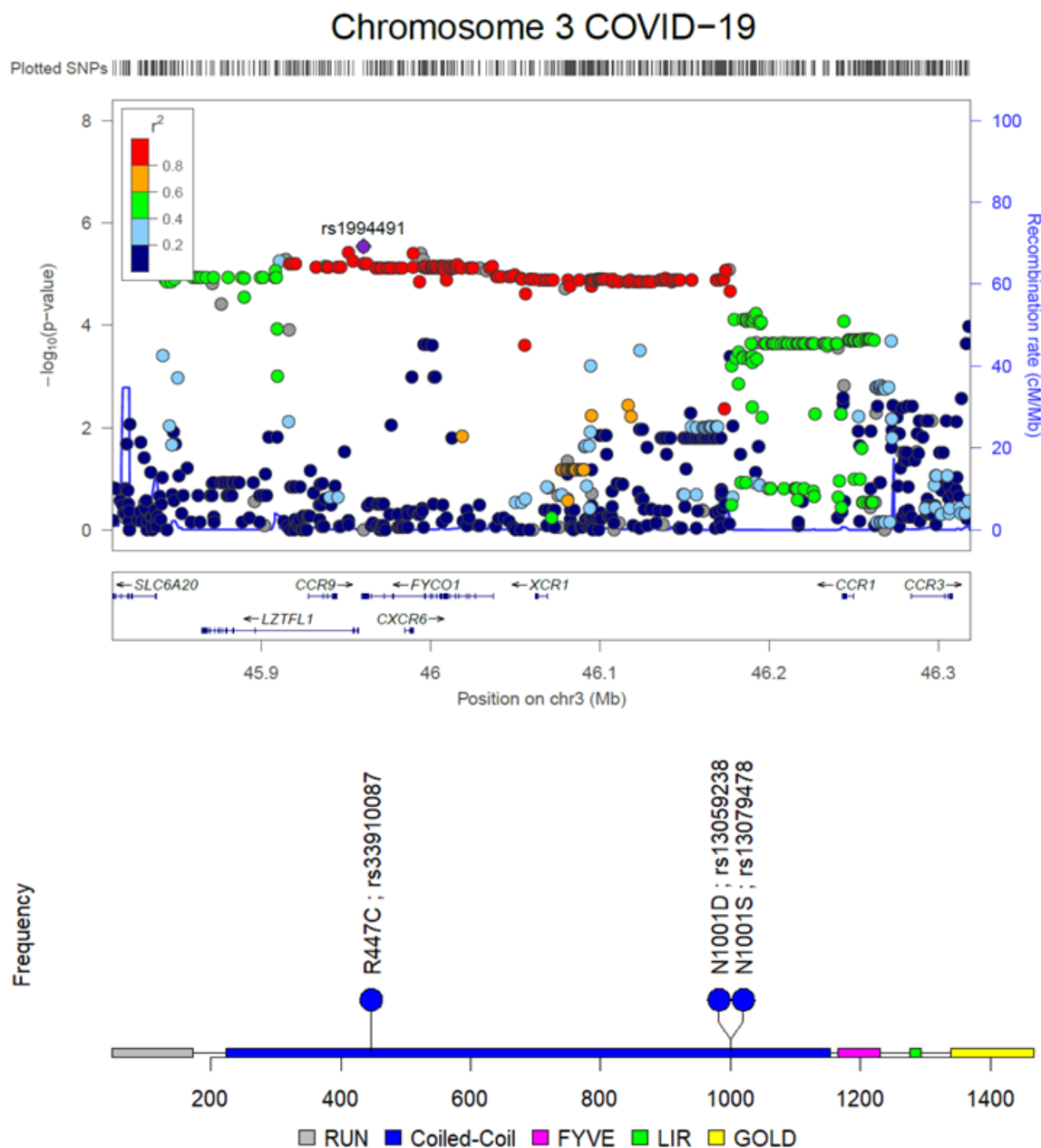
350 **References**

- 351 1. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New*
352 *England Journal of Medicine* **382**, 727–733 (2020).
- 353 2. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with
354 COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* **395**, 1054–1062
355 (2020).
- 356 3. Ganna, A., Unit, T. G. & General, M. The COVID-19 Host Genetics Initiative , a global
357 initiative to elucidate the role of host genetic factors in susceptibility and severity of the
358 SARS-CoV-2 virus pandemic. 715–718 (2020). doi:10.1038/s41431-020-0636-6
- 359 4. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *New*
360 *England Journal of Medicine* **383**, 1522–1534 (2020).
- 361 5. Pairo-Castineira, E. *et al.* Genetic mechanisms of critical illness in Covid-19. *medRxiv* **17**,
362 25 (2020).
- 363 6. Karczewski, K. J. *et al.* Variation across 141,456 human exomes and genomes reveals the
364 spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*
365 531210 (2019). doi:10.1101/531210
- 366 7. Zeberg, H. & Pääbo, S. The major genetic risk factor for severe COVID-19 is inherited
367 from Neanderthals. *Nature* (2020). doi:10.1038/s41586-020-2818-3
- 368 8. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**,
369 580–585 (2013).
- 370 9. Zhang, J., Lan, Y. & Sanyal, S. Membrane heist□: Coronavirus host membrane
371 remodeling during replication. (2020).
- 372 10. Pankiv, S. *et al.* *FYCO1* is a Rab7 effector that binds to LC3 and PI3P to mediate
373 microtubule plus end - Directed vesicle transport. *Journal of Cell Biology* **188**, 253–269
374 (2010).
- 375 11. Ghoush Amit Kumar Mandal , Paulami Dam , Octavio L. Franco , Hanen Sellami ,
376 Sukhendu Mandal , Gulden Can Sezgin , Kinkar Biswas , Partha Sarathi Nandi, I. O. b-
377 Coronaviruses Use Lysosomes for Egress Instead of the Biosynthetic Secretory Pathway.
378 *Cell* 19–21 (2020).
- 379 12. Saridaki, T. *et al.* *FYCO1* mediates clearance of α -synuclein aggregates through a Rab7-
380 dependent mechanism. *Journal of Neurochemistry* **146**, 474–492 (2018).

- 381 13. Daniloski, Z. *et al.* Identification of required host factors for SARS-CoV-2 infection in
382 human cells. *Cell* 1–14 (2020). doi:10.1016/j.cell.2020.10.030
- 383 14. Reggiori, F., de Haan, C. A. M. & Molinari, M. Unconventional use of LC3 by
384 coronaviruses through the alleged subversion of the ERAD tuning pathway. *Viruses* **3**,
385 1610–1623 (2011).
- 386 15. Singer, J., Gifford, R., Cotten, M. & Robertson, D. CoV-GLUE: A Web Application for
387 Tracking SARS-CoV-2 Genomic Variation. *Preprints* 2020060225 (2020).
388 doi:10.20944/PREPRINTS202006.0225.V1
- 389 16. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID’s innovative
390 contribution to global health. *Global Challenges* **1**, 33–46 (2017).
- 391 17. Grubaugh, N. D., Hanage, W. P. & Rasmussen, A. L. Making Sense of Mutation: What
392 D614G Means for the COVID-19 Pandemic Remains Unclear. *Cell* **182**, 794–795 (2020).
- 393 18. Withrock, I. C. *et al.* Genetic diseases conferring resistance to infectious diseases. *Genes*
394 *& Diseases* **2**, 247–254 (2015).
- 395 19. Takadate, Y. *et al.* Niemann-Pick C1 Heterogeneity of Bat Cells Controls Filovirus
396 Tropism. *Cell Reports* **30**, 308–319.e5 (2020).
- 397 20. Haines, K. M., Vande Burgt, N. H., Francica, J. R., Kaletsky, R. L. & Bates, P. Chinese
398 hamster ovary cell lines selected for resistance to ebolavirus glycoprotein mediated
399 infection are defective for NPC1 expression. *Virology* **432**, 20–28 (2012).
- 400 21. Greninger, A. L. *et al.* Rapid metagenomic next-generation sequencing during an
401 investigation of hospital-acquired human parainfluenza virus 3 infections. *Journal of*
402 *Clinical Microbiology* **55**, 177–182 (2017).
- 403 22. Greninger, A. L. *et al.* Ultrasensitive Capture of Human Herpes Simplex Virus Genomes
404 Directly from Clinical Samples Reveals Extraordinarily Limited Evolution in Cell
405 Culture. *mSphere* **3**, 1–12 (2018).
- 406 23. Greninger, A. L. *et al.* Proteomic Reannotation of Human Herpesvirus 6. 1–17 (2018).
- 407 24. Bankevich, A. *et al.* SPAdes: A new genome assembly algorithm and its applications to
408 single-cell sequencing. *Journal of Computational Biology* **19**, 455–477 (2012).
- 409 25. Polymeropoulos, M. H. *et al.* Common effect of antipsychotics on the biosynthesis and
410 regulation of fatty acids and cholesterol supports a key role of lipid homeostasis in
411 schizophrenia. *Schizophrenia Research* **108**, 134–142 (2009).

412 **Figure Legend**

413 **Figure 1.** Figure 1A depicts the coding variants detected in the region association with severity
 414 of COVID-19 infection in FYCO1 gene. Figure 1B is a locus zoom of the entire significant
 415 chromosome 3 region associated with severe COVID-19 based on a whole genome sequencing
 416 analysis.
 417



418

419

420

421

422