## Supplemental Methods:

### Columbia University COVID-19 Biobank Cohort:

The study includes a multiethnic cohort of 1,153 COVID-19 patients treated at the Columbia University Irving Medical Center. All cases had PCR-confirmed SARS-CoV-2 infection. The patients were recruited to the Columbia University COVID-19 Biobank between March and May 2020, at the peak of the New York pandemic. A total of 480 patients were classified as having severe COVID-19. All 480 severe COVID-19 cases were defined by either death from COVID-19 (N=317), or acute respiratory failure due to COVID-19 requiring endotracheal intubation (N=163). This cohort was composed of 288 males and 192 females; the average age was 67 (range 2-101). For each patient, we performed ancestry classification into one of the six major ancestry groups (European, African, Hispanic, East Asian, South Asian and Middle Eastern) using a neural network trained on a set of samples with known ancestry labels. We used a 50% probability cut-off to assign an ancestry label to each sample and labeled samples that did not reach 50% for any of the ancestral groups as "Admixed". 45% of the participants were predicted to be Hispanic/Latinx, 40% African, 6% Middle Eastern, 6% European, 2% East Asian, 0.6% South Asian, and 0.2 % Admixed (see below for details about ancestry determination).

For comparisons of variant frequencies, we used an internal dataset of 9,589 local population controls with exome and genome sequence data generated by Columbia University's Institute for Genomic Medicine (IGM). The controls are derived from the same general patient population as the cases, served by the Columbia University Irving Medical Center and sequenced as controls or healthy family members for other studies. Additional tests were performed against a cohort of 673 patients with mild COVID-19 who were recruited to Columbia COVID-19 Biobank during the same time period, but who have recovered without the need for endotracheal intubation. This group was composed of 357 males and 316 females; the average age was 59 (range 3-97); 43% of the participants were predicted to be Hispanic/Latino, 40% African, 8% Middle Eastern, 6% European, 2% East Asian, and 0.6% South Asian.

All case exomes were captured with the IDT xGen Exome Research Panel V1.0 (Integrated DNA Technologies, Coralville, IA, USA) and sequenced on Illumina's NovaSeq 6000 (Illumina, San Diego, CA, USA) platform with 150 bp paired-end reads according to standard protocols. Exome sequencing of controls were performed on Illumina's HiSeq 2000, HiSeq 2500, and NovaSeq 6000 using various exome capture kits. Genome controls were all sequenced on Illumina's NovaSeq 6000.

All cases and controls were processed with the same bioinformatic pipeline for variant calling. In brief, reads were aligned to human reference GRCh37 using DRAGEN (Edico Genome, San Diego, CA, USA)(1) and duplicates were marked with Picard (Broad Institute, Boston, MA, USA). Variants were called according to the Genome Analysis Toolkit (GATK - Broad Institute, Boston, MA, USA) Best Practices recommendations v3.6(2, 3). Finally, variants were annotated with ClinEff(4) and the IGM's in-house ATAV(5) (https://github.com/igm-team/atav) was used to add custom annotations including gnomAD v2.1 frequencies(6) and clinical annotations provided by the Human Gene Mutation Database (HGMD)(7), ClinVar(8, 9), and Online Mendelian Inheritance in Man (OMIM). A centralized database was used to store variant and per site coverage data for all samples enabling well controlled analyses without the need of generating jointly called VCF files (see Ren et al. 2021(5) for details).

All included samples had at least 90% of the consensus coding sequence (CCDS release 20)(10) covered at 10x or more. Samples had ≤ 3% contamination levels according to VerifyBamID(11). Additionally, we removed samples with a discordance between self-declared and sequence-derived gender. We used KING(12) for the detection of related individuals and removed one of each pair that had an inferred relationship of second-degree or closer while favoring the inclusion of COVID-19 cases over controls.

For ancestry adjustment, we performed Principal Component Analysis on a set of predefined variants as previously described in Cameron-Christie et al.(13) To identify clusters that reflect ancestry we applied the Louvain method of community detection(14) to the first 6 principal components (PCs) resulting in 10 clusters reflecting not only continental populations, but also more detailed subdivision of Europeans. We performed coverage harmonization (see below) and collapsing within the clusters so that population specific variants would be easier to filter out.(15)

Coverage differences between cases and controls caused by different capture kits or sequencing depth in general can potentially introduce a bias in the analysis because without sufficient coverage no variants can be called(16). This is especially an issue if exomes and genomes are analyzed together as was the case in the analysis described by Zhang et al.(17) Therefore, we used the site-based pruning approach described in Petrovski et al. 2015(18) and removed sites with an absolute difference in percentages of cases compared to controls with at least 10x coverage of greater than 7.0%.

Using the genes tested by Zhang *et al*. (*TLR3*, *UNC93B1*, *TICAM1*, *TRAF3*, *TBK1*, *IRF3*, *IRF7*, *IFNAR1*, *IFNAR2*, *STAT1*, *STAT2*, *IRF9*, and *IKBKG*), we performed a gene set-based collapsing analysis to test for a significant difference in the proportion of cases carrying at least one qualifying variant (QV) in the gene set compared

to controls. A QV was defined as a variant that passes certain filter criteria. For the pLOF model we used basic quality control filters (e.g. GATK's Variant Quality Score Recalibration, QUAL, GQ, QD, …), an internal and external minor allele frequency (MAF) of 0.001 and restricted the variant effect to frameshift, stop gained, splice acceptor, and splice donor labeled high confidence by LOFTEE(6). For the internal MAF filter we used a leave-one-out filter that ensures that singletons are retained even if the sample size is small. For the external MAF filtering we used the continental MAFs provided by gnomAD(6) and ExAC(19). For the functional model, we also included missense variants in addition to the pLOF variants while keeping all other filters the same. From the results of the individual clusters, we extracted the number of cases/controls with and without a QV in the gene-set and used the exact two-sided Cochran-Mantel-Haenszel (CMH) test(20, 21) to test for an association between disease status and QV status while controlling for cluster membership.

For the purpose of the meta-analysis, we restricted the Biobanque Québec cohort (see below) to individuals of European ancestry and included population controls. For the Saudi cohort, we used mildly affected individuals as controls. The final analysis included a total of 696 severe COVID-19 cases and 14,918 controls. We used the CMH test as described above and treated the other cohorts as additional clusters, since each of the extra cohorts was composed of a single well-defined ancestry. The CMH odds ratios and 95% confidence intervals were reported.

**Biobanque Québec COVID-19 Cohort:**

The Biobanque Québec COVID-19 (www.BQC19.ca) is a provincial biobank that prospectively enrolls patients with suspected COVID-19, or COVID-19 confirmed through SARS-CoV-2 PCR testing. For this study, we used results from patients with available WGS data and who were recruited at the Jewish General Hospital (JGH) in Montreal. The JGH is a university affiliated hospital serving a large multi-ethnic adult population and was designated as the primary COVID-19 reference center by the Québec government early in the pandemic. In total, there were 533 participants with WGS, including 62 cases of COVID-19 who required ventilatory support (BiPAP, high flow oxygen, or endotracheal intubation) or died, 128 COVID-19 patients who were hospitalized but did not require invasive ventilatory support, 30 individuals with COVID-19 did not require hospitalization, and 313 SARS-CoV-2 PCR-negative participants. Using genetic principal component analysis derived from genome-wide genotyping, we determined that 76% of participants were of European ancestry, 9% were of African ancestry, 7% were of east Asian ancestry, and 5% were of south Asian ancestry.

We performed WGS at a mean depth of 30x on all included individuals using Illumina's Novaseq 6000 platform (Illumina, San Diego, CA, USA). Sequencing results were analyzed using the McGill Genome Center bioinformatics pipelines(22), in accordance with Genome Analysis Toolkit Best Practices (GATK) recommendations(3). Reads were aligned to the GRCh38 reference genome. Variant quality control was performed using the variantRecalibrator and applyVQSR functions from GATK. Participants had a median of 4434 (IQR: 4260-5194) autosomal single nucleotide polymorphisms of minor allele frequency less than 1% across their exomes (defined using the GENCODE(23) reference for protein coding exonic sites). Variant filtering was followed by sample filtering on missing rate (minimum 97% of sites called) and minimum average read depth (≥ 20x). Genotypes were further filtered by genotype quality (≥20x) and read depth (≥10x).

Predicted LOF variants were annotated using LOFTEE(6) and the analysis was restricted to the genes tested in Zhang *et al*. Missense variants were annotated using VEP. Qualifying variants with minor allele frequency < 0.1% annotated as high confidence for LOF or as missense variants and which passed quality control filters were extracted for the main analysis. Using these variants, we used Fisher's exact test to perform a burden test restricted to individuals of European ancestry comparing 45 of the 62 severe cases of COVID-19 to 361 controls.

**Saudi Arabian COVID-19 Cohort:**

The Saudi Human Genome Program (SHGP) aims to sequence the genomes of the Saudi patients with COVID-19 confirmed through SARS-CoV-2 PCR as part of COVID19 KACST response (https://covid19.kacst.edu.sa/en.html). The study was conducted in accordance with the ethical principles of the National Bioethical committee at KACST and approved by the Institutional Review Board Committee at King Abdullah International Medical Research Centre, Ministry of National Guard-Health Affairs, Riyadh, Ministry of Health, and King Fahad Medical City. The Institutional Review Boards of all participating hospitals also approved the study and all patients provided written informed consent. In total, we performed WES on 237 individuals with COVID-19, which included 148 (62%) cases of severe disease that required mechanical ventilation or admission to an intensive care unit for septic shock or organ failure and 89 (38%) patients with mild or asymptomatic disease. The 148 severe cases included 32% females and 68% males, and the 89 mild/asymptomatic patients were 55% females and 45% males. The age of the 237 patients included in this study ranged from 15 years to 92 years old. All samples were of Arab ancestry.

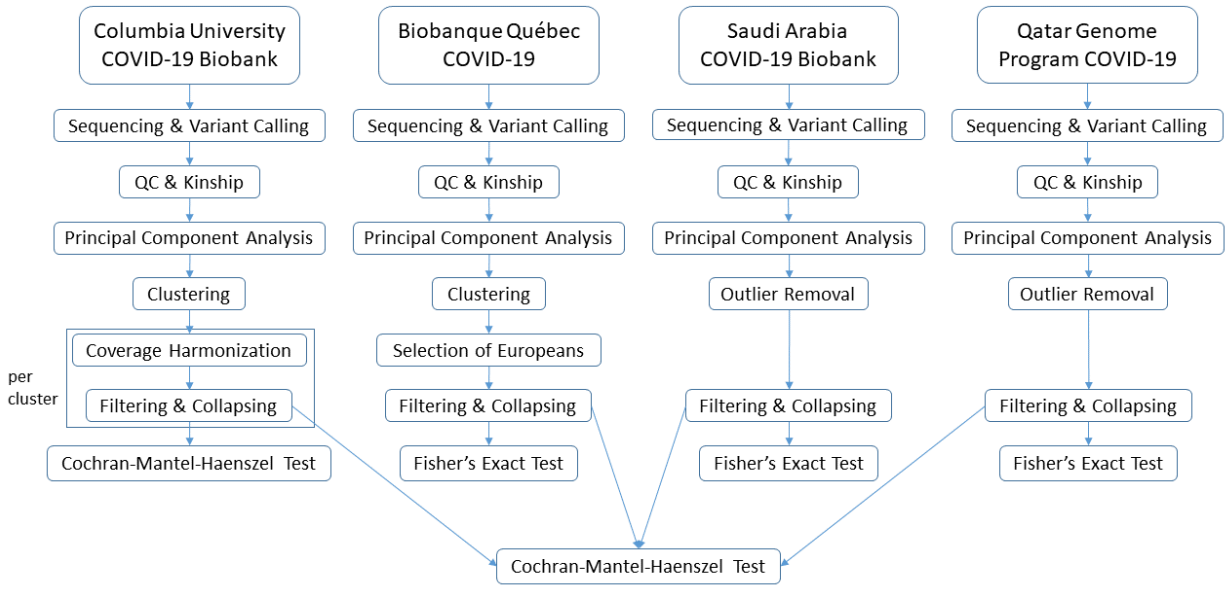
Exomes of patients with mild and severe disease were sequenced at a mean depth of 60x using Illumina's TruSeq Exome kit and the Novaseq 6000 platform (Illumina, San Diego, CA, USA) with 150 bp paired-end reads according to standard protocols. All samples were processed with the same pipeline for variant calling. Reads were aligned to GRCh38/hg38 reference genome and variant were called according to GATK's Best Practices recommendations v4 (2, 3). Quality control included removal of related samples ($2^{nd}$ degree or closer), samples with a low genotyping call rate (less than 95%), gender discordance, excess heterozygosity and PCA-based population outliers. The final multi-sample vcf file was annotated using VEP with the LOFTEE plugin. The analysis was restricted to the 13 genes tested in Zhang et al. Rare variants (minor allele frequency < 0.1%) that were annotated as high confidence for LOF by LOFTEE and passed quality control filters were extracted for the main analysis. For additional analyses we also extracted missense variants and increased the MAF threshold to 1%. We then performed a gene set-based collapsing analysis, using the same criteria as described for the Columbia University COVID-19 Biobank Cohort and used Fisher's Exact test to test for an association.

**Qatar Genome Program COVID-19 Cohort:**

The Qatar Genome Program (QGP) is a population-based project launched by the Qatar Foundation to generate a large whole genome sequence (WGS) dataset, in combination with comprehensive phenotypic information collected by the Qatar Biobank. In this retrospective cohort consisting of 14,060 WGS samples, we extracted information from electronic medical records (EMR) on the participants that were tested positive. The extraction covered the period from early March till early September. There were a total of 700 COVID-19 positive cases, including 60 severe cases with hypoxia that required ventilatory support (BiPAP, high flow oxygen, or endotracheal intubation) and 640 classified as asymptomatic or mild cases without any evidence of pneumonia or hypoxia. The severe COVID-19 cohort consisted of 30 male and 30 female patients that were hospitalized for respiratory failure. Of the mild COVID-19 cases, 290 were male and 350 were female. The median age was 38 (range 18-89) for the severe and 37 (range 18-89) for the mild group. All samples used for the analysis were of Qatari Middle Eastern Arabian ancestry. After removal of related individuals, 23 severe and 231 mild cases remained. The severe COVID-19 cohort consisted of 13 males and 10 females with a median age of 55 (range 28-68). 116 of the mild COVID-19 cohort were male, 115 female with a median age of 37 range (18-79).

All samples were sequenced on Illumina Hiseq X instrument (Illumina, San Diego, CA, USA) to a minimum average coverage of 30x. Quality control of Fastq files was performed using FastQC (v0.11.2). The secondary analyses (read mapping and joint variant calling) were performed using Sentieon's DNASeq

pipeline(24) v201808.03 (Sentieon, San Jose, CA, USA), following the GATK 3.8 best practices(2, 3). Reads were aligned to GRCh38/hg38 reference genome. Quality control analyses included removal of samples with a low genotyping call rate (less than 95%), gender discordance, excess heterozygosity and PCA-based population outliers. We then used KING(12) for the detection of related individuals and removed one of each pair that had an inferred relationship of second-degree or closer. Consequently, 5,385 unrelated samples were retained for our analysis. The remaining variants were annotated using VEP (release 99), with LOFTEE and dbNSFP plugins. We then performed a gene set-based collapsing analysis, using the same criteria as described for the Columbia University COVID-19 Biobank Cohort (pLOF with internal & external MAF < 0.1%).




***Supplementary Figure 1:*** *Flowchart of the analysis approach. Except for the last step, which represents the meta-analysis, all steps were performed independently by the respective site.*

***Supplementary Table 1:*** *Differences in study design between the analysis in Zhang et al. and the current analysis.*

| | **Zhang et al.** | **Povysil et al.** |
|---|---|---|
| **Cases/controls** | 659 severe vs. 534 mild | 713 severe vs. 1,151 mild or 15,033 population controls |
| **Avg age** | 51.8 years old | 65.9 years old |
| **Ancestry correction** | Control for 3 PCs | Clustering based on 6 PCs & stratified analysis or extraction of single homogenous ancestry |
| **WES/WGS correction** | None | Coverage harmonization or same assay for entire cohort |
| **Included variants** | 1. pLoFs for cases and controls<br>2. Functional characterization in cases only | 1. pLoFs for cases and controls<br>2. pLoFs + missense in cases and controls, no functional characterization |

References:

1. Miller NA et al. A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. *Genome Med.* 2015;7(1):100.

2. McKenna A et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.. *Genome Res.* 2010;20(9):1297–303.

3. Van der Auwera GA et al. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinforma.* 2013;43(1):11.10.1-11.10.33.

4. Cingolani P et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin).* 2012;6(2):80–92.

5. Ren Z et al. ATAV: a comprehensive platform for population-scale genomic analyses. *BMC Bioinformatics* 2021;22(1):149.

6. Karczewski KJ et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 2020;581(7809):434–443.

7. Stenson PD et al. Human Gene Mutation Database (HGMD ® ): 2003 update. *Hum. Mutat.* 2003;21(6):577–581.

8. Landrum MJ et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018;46(D1):D1062–D1067.

9. Landrum MJ et al. ClinVar: public archive of relationships among sequence variation and human phenotype.. *Nucleic Acids Res.* 2014;42(Database issue):D980-5.

10. Pruitt KD et al. The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes.. *Genome Res.* 2009;19(7):1316–23.

11. Jun G et al. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data.. *Am. J. Hum. Genet.* 2012;91(5):839–48.

12. Manichaikul A et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* 2010;26(22):2867–2873.

13. Cameron-Christie S et al. Exome-Based Rare-Variant Analyses in CKD.. *J. Am. Soc. Nephrol.* 2019;30(6):1109–1122.

14. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks[published online ahead of print: March 4, 2008]; doi:10.1088/1742-5468/2008/10/P10008

15. Povysil G et al. Assessing the Role of Rare Genetic Variation in Patients with Heart Failure. *JAMA Cardiol.* 2020;6(4):379–386.

16. Povysil G et al. Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nat. Rev. Genet.* 2019;1–13.

17. Zhang Q et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* 2020;370(6515):eabd4570.

18. Petrovski S et al. An Exome Sequencing Study to Assess the Role of Rare Genetic Variation in Pulmonary Fibrosis. *Am. J. Respir. Crit. Care Med.* 2017;196(1):82–93.

19. Lek M et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536(7616):285–291.

20. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease.. *J. Natl. Cancer Inst.* 1959;22(4):719–48.

21. Cochran WG. Some Methods for Strengthening the Common χ 2 Tests. *Biometrics* 1954;10(4):417.

22. Bourgey M et al. GenPipes: an open-source framework for distributed and scalable genomic analyses. *Gigascience* 2019;8(6). doi:10.1093/gigascience/giz037

23. Frankish A et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 2019;47(D1):D766–D773.

24. Freed D, Aldana R, Weber JA, Edwards JS. The Sentieon Genomics Tools - A fast and accurate solution to variant calling from next-generation sequence data. *bioRxiv* 2017;115717.