**Supplementary appendix**

**Supplementary results**

**Table S1.** Definition of the phenotypes.

**Table S2.** Functional enrichment analysis of the identified genes in blood.

**Table S3.** Functional enrichment analysis of the identified genes in lung.

**Figure S1.** Prioritizing gene around *IFNAR2* in association with severe COVID-19 in blood.

**Figure S2.** Prioritizing genes around *IFNAR2* in association with severe COVID-19 NEJM in blood.

**Figure S3.** Prioritizing genes around *PTPN1* in association with COVID-19 in blood.

**Figure S4.** Prioritizing genes around *FCER1G* in association with COVID-19 negative in blood.

**Figure S5.** Prioritizing genes around *BET1L* in association with COVID-19 without hospitalization in blood.

**Figure S6.** Prioritizing genes around *C7orf25*in association with severe COVID-19 in lung.

**Figure S7.** Prioritizing genes around *AP006621.1* in association with Severe COVID-19 NEJM in lung.

**Figure S8.** Prioritizing genes around *ATP5O* in association with COVID-19 in lung.

**Figure S9.** Prioritizing genes around *ZNF589* in association with COVID-19 negative in lung.

**Figure 10.** Prioritizing genes around *RP11-544L8\_\_B.4* in association with hospitalized COVID-19 in lung.

**Figure S11.** Prioritizing genes around *PSMD13* in association with COVID-19 without hospitalization in lung.

**Table S1. Definition of the phenotypes.**

|  |  |  |
| --- | --- | --- |
| **Phenotype** | **Case** | **Control** |
| **Severe COVID-19** | Hospitalized laboratory confirmed SARS-CoV-2 infection (RNA and/or serology based), AND (death OR respiratory support (intubation, CPAP, BiPAP, CNP (continue external negative pressure), Optiflow/very high flow Positive End Expiratory Pressure Oxygen\* - AND hospitalization with COVID19 as primary reason for admission | Everybody who is not a case, e.g. population |
| **Severe COVID-19 NEJM#** | Hospitalization with respiratory failure and a confirmed SARS-CoV-2 viral RNA polymerase-chain-reaction (PCR) test from  nasopharyngeal swabs or other relevant biologic  fluids | Everybody who is not a case, e.g. population |
| **COVID-19** | Individuals with laboratory confirmation of SARS-CoV-2 infection (RNA and/or serology based) OR EHR/ICD coding/ Physician Confirmed COVID-19 OR self-reported COVID-19 positive (e.g. by questionnaire) | Everybody who is not a case, e.g. population |
| **COVID-19 negative¶** | Individuals with laboratory confirmation of SARS-CoV-2 infection (RNA and/or serology based) OR EHR/ICD coding/ Physician Confirmed COVID-19 OR self-reported COVID-19 positive (e.g. by questionnaire) | (Laboratory tested for SARS-CoV-2 infection (RNA and/or serology based) AND all tests (if multiple tests) negative\*) OR self-reported tested negative for SARS-CoV-2 infection (e.g. by questionnaire) |
| **Hospitalized COVID-19** | Hospitalized laboratory confirmed SARS-CoV-2 infection (RNA and/or serology based), hospitalization due to corona-related symptoms | Everybody who is not a case, e.g. population |
| **COVID-19 without hospitalization§** | Hospitalized laboratory confirmed SARS-CoV-2 infection (RNA and/or serology based), hospitalization due to corona-related symptoms | Laboratory confirmed SARS-CoV-2 infection (RNA and/or serology based) AND not hospitalised 21 days after the test |

\*The GWAS summarized data were provided by the COVID-19 host genetics initiative and the control is the population unless otherwise noted. The eQTL data for blood and lung can be downloaded at https://cnsgenomics.com/data/SMR/#eQTLsummarydata.

#Summary data were provided by the Severe Covid-19 GWAS Group.

¶The control is the non-predicted COVID-19 and non-self-reported COVID-19.

§The control is COVID-19 without hospitalization, including laboratory confirmed or self-reported COVID-19.

**Table S2. Functional enrichment analysis of the genes tagged by the top probes in blood.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Input ID** | **Gene ID** | **Description** | **GO:0060338 regulation of type I interferon** | **GO:0060627 regulation of vesicle-mediated transport** | **GO:0034976 response to endoplasmic reticulum stress** | **GO:0006979 response to oxidative stress** |
| ***IFNAR2*** | 3455 | interferon alpha and beta receptor subunit 2 | 1 | 0 | 0 | 0 |
| ***TRIM5*** | 85363 | tripartite motif containing 5 | 1 | 0 | 0 | 0 |
| ***NLRC5*** | 84166 | NLR family CARD domain containing 5 | 1 | 0 | 0 | 0 |
| ***C7orf25*** | 79020 | chromosome 7 open reading frame 25 | 0 | 0 | 0 | 0 |
| ***MCTP1*** | 79772 | multiple C2 and transmembrane domain containing 1 | 0 | 1 | 0 | 1 |
| ***FAN1*** | 22909 | FANCD2 and FANCI associated nuclease 1 | 0 | 0 | 0 | 0 |
| ***MRPL32*** | 64983 | mitochondrial ribosomal protein L32 | 0 | 0 | 0 | 0 |
| ***PTPN1*** | 5770 | protein tyrosine phosphatase non-receptor type 1 | 1 | 1 | 1 | 0 |
| ***MARCH6*** | 10299 | membrane associated ring-CH-type finger 6 | 0 | 0 | 1 | 0 |
| ***ITGAD*** | 3681 | integrin subunit alpha D | 0 | 0 | 0 | 0 |
| ***FCER1G*** | 2207 | Fc fragment of IgE receptor Ig | 1 | 1 | 0 | 0 |
| ***ATP8B4*** | 79895 | ATPase phospholipid transporting 8B4 (putative) | 0 | 0 | 0 | 0 |
| ***ZNF589*** | 51385 | zinc finger protein 589 | 0 | 0 | 0 | 0 |
| ***RAB8A*** | 4218 | RAB8A, member RAS oncogene family | 0 | 1 | 0 | 0 |
| ***ATF4*** | 468 | activating transcription factor 4 | 0 | 0 | 1 | 1 |
| ***C7orf43*** | 55262 | microtubule associated protein 11 | 0 | 0 | 0 | 0 |
| ***SLC24A4*** | 123041 | solute carrier family 24 member 4 | 0 | 0 | 0 | 0 |
| ***BET1L*** | 51272 | Bet1 golgi vesicular membrane trafficking protein like | 0 | 1 | 0 | 0 |
| ***C18orf10*** | 25941 | tubulin polyglutamylase complex subunit 2 | 0 | 0 | 0 | 0 |
| ***CDC7*** | 8317 | cell division cycle 7 | 0 | 0 | 0 | 0 |
| ***PON2*** | 5445 | paraoxonase 2 | 0 | 0 | 0 | 1 |

**Table S3. Functional enrichment analysis of the genes tagged by the top probes in lung.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Input ID** | **Gene ID** | **Description** | **GO:0002223 stimulatory C-type lectin receptor** | **GO:0002366 leukocyte activation involved in immune response** |
| ***ARSA*** | 410 | arylsulfatase A | 0.0 | 1.0 |
| ***C7orf25*** | 79020 | chromosome 7 open reading frame 25 | 0.0 | 0.0 |
| ***HPS5*** | 11234 | HPS5 biogenesis of lysosomal organelles complex 2 subunit 2 | 0.0 | 0.0 |
| ***GABPB1-AS1*** | 100129387 | GABPB1 antisense RNA 1 | 0.0 | 0.0 |
| ***HERC2P10*** | 390561 | hect domain and RLD 2 pseudogene 10 | 0.0 | 0.0 |
| ***ATP5O*** | 539 | ATP synthase peripheral stalk subunit OSCP | 0.0 | 0.0 |
| ***AS3MT*** | 57412 | arsenite methyltransferase | 0.0 | 0.0 |
| ***C10orf32*** | 119032 | BLOC-1 related complex subunit 7 | 0.0 | 0.0 |
| ***ITGAD*** | 3681 | integrin subunit alpha D | 0.0 | 0.0 |
| ***EXOSC6*** | 118460 | exosome component 6 | 0.0 | 1.0 |
| ***ZNF589*** | 51385 | zinc finger protein 589 | 0.0 | 0.0 |
| ***FCER1G*** | 2207 | Fc fragment of IgE receptor Ig | 1.0 | 1.0 |
| ***MRPS18AP1*** | 359761 | mitochondrial ribosomal protein S18A pseudogene 1 | 0.0 | 0.0 |
| ***CNTF*** | 1270 | ciliary neurotrophic factor | 0.0 | 0.0 |
| ***NOMO3*** | 408050 | NODAL modulator 3 | 0.0 | 0.0 |
| ***MUC5B*** | 727897 | mucin 5B, oligomeric mucus/gel-forming | 1.0 | 0.0 |
| ***OR52K3P*** | 390035 | olfactory receptor family 52 subfamily K member 3 pseudogene | 0.0 | 0.0 |
| ***DDX60L*** | 91351 | DExD/H-box 60 like | 0.0 | 0.0 |
| ***PSMD13*** | 5719 | proteasome 26S subunit, non-ATPase 13 | 1.0 | 1.0 |
| ***CCHCR1*** | 54535 | coiled-coil alpha-helical rod protein 1 | 0.0 | 0.0 |
| ***SIRT3*** | 23410 | sirtuin 3 | 0.0 | 0.0 |
| ***TPGS2*** | 25941 | tubulin polyglutamylase complex subunit 2 | 0.0 | 0.0 |
| ***TRIOBP*** | 11078 | TRIO and F-actin binding protein | 0.0 | 0.0 |



**Figure S1. Prioritizing gene around *IFNAR2* in association with severe COVID-19 in blood.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of severe COVID-19, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ILMN\_1751051 probe, tagging *IFNAR2*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci

**Figure S2. Prioritizing genes around *IFNAR2* in association with severe COVID-19 NEJM in blood.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of severe COVID-19 NEJM, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ILMN\_1765146 probe, tagging *IFNAR2*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; *IFNAR2*, interferon alpha and beta receptor subunit 2; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S3. Prioritizing genes around *PTPN1* in association with COVID-19 in blood.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of COVID-19, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ILMN\_1681591 probe, tagging *PTPN1*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S4. Prioritizing genes around *FCER1G* in association with COVID-19 negative in blood.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of COVID-19 negative, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ILMN\_2123743 probe, tagging *FCER1G*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S5. Prioritizing genes around *BET1L* in association with COVID-19 without hospitalization in blood.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of COVID-19 without hospitalization, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ILMN\_2060652 probe, tagging *BET1L*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S6. Prioritizing genes around *C7orf25*in association with severe COVID-19 in lung.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of severe COVID-19, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ENSG00000136197.8 probe, tagging *C7orf25*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S7. Prioritizing genes around** ***AP006621.1* in association with severe COVID-19 NEJM in lung.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of hospitalized COVID-19, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ENSG00000177236.3 probe, tagging *AP006621.1*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S8. Prioritizing genes around *ATP5O* in association with COVID-19 in lung.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of COVID-19, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ENSG00000241837.2 probe, tagging *ATP5O*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S9. Prioritizing genes around *ZNF589* in association with COVID-19 negative in lung.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of COVID-19 negative, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ENSG00000164048.9 probe, tagging *ZNF589*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure 10. Prioritizing genes around *RP11-544L8\_\_B.4* in association with hospitalized COVID-19 in lung.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of hospitalized COVID-19, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL results in blood for top one probe, ENSG00000175967.3 probe, tagging *RP11-544L8\_\_B.4*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci



**Figure S11. Prioritizing genes around *PSMD13* in association with COVID-19 without hospitalization in lung.**

Top plot, grey dots represent the -log10(*P* values) for SNPs from the GWAS of COVID-19 without hospitalization, and rhombuses represent the -log10(*P* values) for probes from the SMR test with solid rhombuses indicating that the probes pass HEIDI test and hollow rhombuses indicating that the probes do not pass the HEIDI test. Middle plot, eQTL result in blood for the top one probe, ENSG00000185627.13 probe, tagging *PSMD13*. Bottom plot, location of genes tagged by the probes.

Highlighted in maroon indicates probes that pass the default SMR threshold (8.4×10-6).

COVID-19, coronavirus disease 2019; GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quantitative trait loci