

# Single-cell atlas of COVID-19 based on multiple sample integration

Pei Tang, MD<sup>a</sup>, Jinghan Wang, MD<sup>b</sup>, Hairui Li, MD<sup>c</sup>, Yong Qing, MD<sup>d</sup>, Zhixing Chen, MD<sup>d</sup>, Junliang Wu, MD<sup>d</sup>, Xiaolin Ran, MD<sup>a</sup>, Gang Zhong, MD<sup>e,\*</sup>

## Abstract

This study addressed the significant challenges in understanding the specific immune landscape and predicting patient outcomes of COVID-19, particularly in severe cases. To support clinical practice, we aimed to build a comprehensive single-cell atlas of the disease and develop a predictive model for severe patient outcomes. We collected single-cell RNA sequencing data from COVID-19 patients, influenza patients, and healthy controls. Our analysis encompassed alterations in cell abundance, transcriptomic profiles, and pathway activities across these groups. We then constructed a logistic regression model to predict the prognosis of severe COVID-19 patients, evaluating its performance using sensitivity, specificity, and area under the curve (AUC) metrics. Our integrated dataset comprised 434,703 cells from 233 samples across 8 datasets, categorized into major cell types including T/natural killer cells, B cells, myeloid cells, epithelial cells, and platelets. T/NK cells and myeloid cells constituted the largest proportions and exhibited the most significant changes in abundance among the different groups. We identified substantial alterations in inflammatory genes and metabolic pathways across various T cell and myeloid cell subtypes, with differential expression of these genes also correlating with COVID-19 severity and patient outcomes. The developed logistic regression model, incorporating the genes IL1R2, PI3, IGHG3, and CTTN, demonstrated a sensitivity of 80.0% and an AUC of 70.5% in the training set. These results were consistent in the test set, with a sensitivity of 81.1% and an AUC of 70.6%. By depicting a comprehensive single-cell atlas of COVID-19, this work unravels the complex cellular dynamics underlying the disease, thereby advancing our understanding of its mechanisms. The developed model shows promise for predicting outcomes in severe COVID-19 patients, which could facilitate clinical decision-making and ultimately benefit patient care.

**Abbreviations:** AUC = area under the curve, COVID-19 = novel coronavirus disease, DEGs = differentially expressed genes, HC = healthy controls, NK cells = natural killer cells, PB = peripheral blood, SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2, scRNA-seq = single-cell RNA sequencing.

**Keywords:** COVID-19, severity, single-cell RNA sequencing, survival outcome

## 1. Introduction

The outbreak of the novel coronavirus disease (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to a global health crisis with significant social and economic impacts.<sup>[1]</sup> In a clinical setting, COVID-19 patients exhibit varying symptoms, severity and disease prognosis, indicating the inherent high heterogeneity of this disease. Researchers worldwide have been tirelessly working to understand the virus's biology, host response, and to develop effective diagnostic and therapeutic strategies. Multiple mechanisms,

including abnormal expression of ACE2 (an important host receptor of SARS-CoV-2), inhibited interferon signaling, an overaggressive immune response featured by cytokine storms, age, gender, body-mass index and some comorbidities, are responsible for the COVID-19 onset and development.<sup>[2–4]</sup> For instance, SARS-CoV-2-induced ACE2 downregulation is responsible for decreased Ang II clearance and thus promote tissue damage. However, such downregulation can also limit viral cell entry and their spread, demonstrating a beneficial effect.<sup>[3]</sup> The multiple roles of ACE2 in COVID-19 highlight

PT and JW contributed equally to this work.

This work is funded by Project of Chengdu Medical Association (2023131).

This study used publicly available datasets that are anonymized and do not contain personally identifiable information. As the data are publicly accessible and anonymized, no informed consent for publication is required.

The authors have no conflicts of interest to disclose.

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

The need for research ethics approval was waived because the data were available with open access.

<sup>a</sup> Department of Medical Cosmetology, The First People's Hospital in Shuangliu District/West China Airport Hospital, Sichuan University, Chengdu, China, <sup>b</sup> West China School of Medicine, Sichuan University, Chengdu, China, <sup>c</sup> Department of Plastic Reconstructive and Aesthetic Surgery, West China Tianfu Hospital, Sichuan University, Chengdu, China, <sup>d</sup> Department of Burn and Plastic Surgery, West China Hospital, Sichuan University, Chengdu, China, <sup>e</sup> Department of

Orthopedic Surgery, West China Hospital, Sichuan University/Trauma Center, West China Hospital, Sichuan University, Chengdu, China.

\* Correspondence: Gang Zhong, No. 37 Guoxue Lane, Wuhou District, Chengdu 610041, Sichuan Province, China (e-mail: zg202510@163.com).

Copyright © 2025 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Tang P, Wang J, Li H, Qing Y, Chen Z, Wu J, Ran X, Zhong G. Single-cell atlas of COVID-19 based on multiple sample integration. *Medicine* 2025;104:52(e46038).

Received: 6 January 2025 / Received in final form: 1 October 2025 / Accepted: 16 October 2025

<http://dx.doi.org/10.1097/MD.00000000000046038>

the pathogenesis complexity of this disease. The diversity and complexity of phenotypes and pathogenesis pose significant challenges in the development of effective diagnostic and treatment strategies for COVID-19. Investigating the intricate biological behavior of individual molecule or cell at a high research resolution and further integrating it with the disease phenotype is crucial for understanding such diversity and complexity.

Single-cell RNA sequencing (scRNA-seq) is a cutting-edge technique that enables the comprehensive profiling of gene expression patterns at the single-cell level. It has emerged as powerful tools for unraveling the complexities of viral infections and host immune responses at unprecedented resolution. The application of scRNA-seq in the context of COVID-19 has already yielded remarkable discoveries. For instance, Liao M et al<sup>[5]</sup> depict the single-cell immune profile associated with COVID-19 severity, and further characterize severe and moderate COVID-19 patients with proinflammatory monocyte-derived macrophages and highly clonally expanded CD8 + T cells, respectively. Ren X et al<sup>[6]</sup> point out the important roles of megakaryocytes and monocyte subsets in contributing COVID-19-related cytokine storms at the single-cell level. Evidently, by capturing the transcriptomic profiles of individual cells, researchers have gained valuable insights into the molecular mechanisms underlying SARS-CoV-2 infection, the host immune response, and the identification of potential therapeutic targets. Moreover, scRNA-seq has shed light on the heterogeneity in immune cell populations, providing a deeper understanding of immune dysregulation and the factors influencing disease severity and outcomes. However, there are discrepancies among different research results,<sup>[5,7]</sup> which may be attributed to factors such as limited sample sizes, variations in sample types (e.g. peripheral blood (PB) or bronchoalveolar lavage fluid), or differences in population characteristics. Therefore, there is a need for higher levels of evidence with greater rigor and consistency.

Herein, this study integrated all relevant datasets (233 samples from 5 different sample types) to systematically delineate the COVID-19-specific single-cell immune landscape across different sample sources and patient characteristics. Additionally, genes associated with COVID-19 prognosis were analyzed, and based on this, prediction models were further developed and validated, aiming to provide promising tools for clinical practice and further improving patient prognosis effectively.

## 2. Materials and methods

### 2.1. Dataset retrieval and data collection

As of November 10, 2021, we conducted searches in the National Center for Biotechnology Information Gene Expression Omnibus, National Center for Biotechnology Information Sequence Read Archive, and European Bioinformatics Institute databases using keywords “novel coronavirus disease,” “severe acute respiratory syndrome coronavirus 2,” and their abbreviations. The data should be generated by scRNA-seq. Then, the retrieved datasets were filtered based on the following criteria. Samples needed to originate from clinical patients, with no restriction on sample type. If a dataset was generated based on specific cell populations, it would be excluded to avoid bias in subsequent cell abundance analysis. To prevent significant variations in cell abundance among different datasets during data integration, datasets generated based on multiplexed sequencing were excluded. Datasets without raw sequencing data were also excluded.

After the above filtering, we collected relevant information of datasets and samples. Dataset-related information included dataset ID, sequencing platform, sample types, sample sizes, and

groups. Sample-related information included age, sex, race, disease, severity, treatment, outcome, sampling timepoints, and so on. All this information was extracted from annotated files of the dataset or corresponding articles.

### 2.2. Raw data processing

Raw data files of each sample (sra format) were download and md5 validation was performed. Both Faster-dump (version 2.11.0) and gzip (version 1.10.0) were used to obtain sequencing files in fq.gz format. Quality control was performed by fastqc (version 1.4.0). For datasets generated by 10x Genomics, Cellranger (version 3.0.0) was applied to alignment and quantification. For datasets generated by other platforms such as BD Rhapsody and Seq-Well, zUMI (version 2.8.2) was used in this step. The human reference genome hg38 version was obtained from the Ensembl database.

For datasets generated by 10x Genomics, 3 files (barcodes.tsv.gz, features.tsv.gz and matrix.mtx.gz) were used to generate Seurat objects by the “CreateSeuratObject” command. For other datasets, the “dgcountrds” file was used to get the matrix and then generate the Seurat object.

### 2.3. Downstream analysis

Cells with feature counts between 500 to 4000 or a mitochondrial count  $\leq 15\%$  are included. The matrices of these cells were normalized based on 2000 variable genes. Then, harmony was used to integrate different samples and clinical information was added to metadata. Both principal component analysis and uniform manifold approximation and projection were performed for dimensionality reduction and cell clustering. Manual annotation and SingleR (version 1.8.0) were applied to annotate cells based on their marker genes. Dimensional reduction plots, bar plots feature plots, or violin plots were used to visualize cells, cell abundance and marker genes of each cell population.

Genes (DEGs) that were differentially expressed among different groups or different cell populations were captured by Wilcoxon Rank Sum test. The adjusted  $P$  value  $< .05$  were set as the filtering criteria. Volcano plots were used to visualize the differentially expressed genes (DEGs). Functional enrichment was conducted based on single-cell rank-based gene set enrichment analysis (version 1.1.2), which analyzed significantly up- or down-regulated pathways of various cell clusters among different groups. Bubble plots were drawn to show the enrichment results.

### 2.4. Predictive models

Effectively predicting disease outcomes is the key to adjust the therapeutic plan timely and benefit patients. Based on DEGs of each cluster, univariate and multifactorial logistic regression were conducted to select potential prognostic variables. Then, these candidates were integrated by logistic regression to construct predictive models for COVID-19 prognosis. According to the ratio of 7:3, a training cohort and test cohort were divided. Sensitivity, specificity, and area under curve (AUC) were calculated to quantitatively assess the predictive performance. Nomogram and receiver operating characteristic curve were plotted to exhibit the predictive model and its performance.

### 2.5. Statistical analysis

R (version 4.1.2; Vienna, Austria) and GraphPad Prism (version 7.0a; La Jolla) were applied to conduct downstream statistical analysis.  $P < .05$  was considered as significance.

**Table 1****The clinical information of included samples.**

| Sample ID*  | Group    | Age (yr) | Gender | Race  | Sample type       | Sampling post confirm (d) | Day post symptom | Symptom duration (d) | Ecmo use | Severity degree | Survival outcome |
|-------------|----------|----------|--------|-------|-------------------|---------------------------|------------------|----------------------|----------|-----------------|------------------|
| SRR14466492 | COVID-19 | <60      | Male   | White | Whole blood       | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466490 |          | <60      | Male   | White | Whole blood       | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466435 |          | <60      | Male   | NA    | Whole blood       | NA                        | NA               | NA                   | NA       | Severe          | Death            |
| SRR14466429 |          | <60      | Male   | Other | Whole blood       | NA                        | NA               | NA                   | NA       | Severe          | Death            |
| SRR14466427 |          | >60      | Male   | White | Whole blood       | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466431 |          | >60      | Female | Other | Whole blood       | NA                        | NA               | NA                   | NA       | Severe          | Death            |
| SRR14854203 |          | <60      | Male   | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14854202 |          | <60      | Male   | NA    | PBMC              | NA                        | >14              | NA                   | NA       | Severe          | Death            |
| SRR14854201 |          | >60      | Male   | NA    | PBMC              | NA                        | >14              | NA                   | NA       | Severe          | Death            |
| SRR14854200 |          | >60      | Male   | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14854199 |          | <60      | Male   | NA    | PBMC              | NA                        | >14              | NA                   | NA       | Severe          | Alive            |
| SRR14854198 |          | >60      | Male   | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14466494 |          | >60      | Male   | White | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466493 |          | <60      | Male   | White | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466491 |          | <60      | Male   | White | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466436 |          | <60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | Death            |
| SRR14466430 |          | <60      | Male   | Other | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | Death            |
| SRR14466428 |          | >60      | Male   | White | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR11804722 |          | <60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | Alive            |
| SRR11804721 |          | <60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | Alive            |
| SRR11804719 |          | >60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | Alive            |
| SRR11680223 |          | >60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11680207 |          | >60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | Yes      | Severe          | NA               |
| SRR14854209 |          | <60      | Female | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14854208 |          | <60      | Female | NA    | PBMC              | NA                        | >14              | NA                   | NA       | Severe          | Alive            |
| SRR14854207 |          | >60      | Female | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14854206 |          | <60      | Female | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14854205 |          | <60      | Female | NA    | PBMC              | NA                        | ≤7               | NA                   | NA       | Severe          | Death            |
| SRR14854204 |          | >60      | Female | NA    | PBMC              | NA                        | 8–14             | NA                   | NA       | Severe          | Death            |
| SRR14466489 |          | <60      | Female | White | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466484 |          | <60      | Female | White | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| SRR14466432 |          | >60      | Female | Other | PBMC              | NA                        | NA               | NA                   | NA       | Severe          | Death            |
| SRR11680222 |          | >60      | Female | NA    | PBMC              | NA                        | NA               | NA                   | Yes      | Severe          | NA               |
| SRR11680221 |          | >60      | Female | NA    | PBMC              | NA                        | NA               | NA                   | Yes      | Severe          | NA               |
| SRR11680216 |          | >60      | Female | NA    | PBMC              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11680215 |          | >60      | Female | NA    | PBMC              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| GSM5361200  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | NA                   | NA       | Severe          | NA               |
| GSM5361199  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361198  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361197  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361196  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | 8–14                 | NA       | Severe          | NA               |
| GSM5361195  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | 8–14                 | NA       | Severe          | NA               |
| GSM5361194  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | 8–14                 | NA       | Severe          | NA               |
| GSM5361193  |          | >60      | Male   | Asian | Nasal wash        | NA                        | NA               | 8–14                 | NA       | Severe          | NA               |
| GSM5361184  |          | >60      | Female | other | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361183  |          | >60      | Female | other | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361182  |          | >60      | Female | other | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361181  |          | >60      | Female | other | Nasal wash        | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| SRR11537951 |          | NA       | NA     | NA    | BALF <sup>c</sup> | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11537950 |          | NA       | NA     | NA    | BALF              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11537949 |          | NA       | NA     | NA    | BALF              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11181959 |          | >60      | Male   | NA    | BALF              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11181958 |          | >60      | Male   | NA    | BALF              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR11181956 |          | >60      | Male   | NA    | BALF              | NA                        | NA               | NA                   | No       | Severe          | NA               |
| SRR14466455 |          | >60      | Male   | White | Whole blood       | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466445 |          | >60      | Female | Other | Whole blood       | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466434 | COVID-19 | <60      | Female | Other | Whole blood       | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466487 |          | <60      | Male   | Other | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466486 |          | >60      | Male   | White | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466485 |          | <60      | Male   | Asian | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466456 |          | >60      | Male   | White | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR11804724 |          | <60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | Alive            |
| SRR11804723 |          | <60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | Alive            |
| SRR11804720 |          | <60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | Alive            |
| SRR11804718 |          | >60      | Male   | NA    | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | Alive            |
| SRR14466488 |          | <60      | Female | White | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466446 |          | >60      | Female | Other | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |
| SRR14466433 |          | <60      | Female | White | PBMC              | NA                        | NA               | NA                   | NA       | Moderate        | NA               |

(Continued)

**Table 1**  
**(Continued)**

| Sample ID*  | Group | Age (yr) | Gender | Race  | Sample type | Sampling post confirm (d) | Day post symptom | Symptom duration (d) | Ecmo use | Severity degree | Survival outcome |
|-------------|-------|----------|--------|-------|-------------|---------------------------|------------------|----------------------|----------|-----------------|------------------|
| GSM5361192  |       | >60      | Male   | other | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361191  |       | >60      | Male   | other | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361190  |       | >60      | Male   | other | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361189  |       | >60      | Male   | other | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361188  |       | >60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361187  |       | >60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361186  |       | >60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361185  |       | >60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361180  |       | <60      | Female | Black | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361179  |       | <60      | Female | Black | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361178  |       | <60      | Female | Black | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361177  |       | <60      | Female | Black | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361176  |       | >60      | Female | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361175  |       | >60      | Female | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361174  |       | >60      | Female | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| GSM5361173  |       | >60      | Female | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Moderate        | NA               |
| SRR14466447 |       | <60      | Female | White | Whole blood | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466443 |       | <60      | Female | Asian | Whole blood | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466451 |       | >60      | Male   | White | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR11680224 |       | >60      | Male   | NA    | PBMC        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR11680217 |       | <60      | Male   | NA    | PBMC        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR14466454 |       | <60      | Female | White | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466453 |       | <60      | Female | Other | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466452 |       | >60      | Female | Other | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466450 |       | <60      | Female | White | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466449 |       | <60      | Female | Asian | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466448 |       | <60      | Female | White | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR14466444 |       | <60      | Female | Asian | PBMC        | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| SRR11680226 |       | >60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR11680218 |       | <60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR11680208 |       | >60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| GSM5361172  |       | <60      | Male   | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361171  |       | <60      | Male   | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361170  |       | <60      | Male   | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361169  |       | <60      | Male   | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| SRR11181957 |       | <60      | Male   | NA    | BALF        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR11181954 |       | <60      | Male   | NA    | BALF        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR11181955 |       | <60      | Female | NA    | BALF        | NA                        | NA               | NA                   | No       | Mild            | NA               |
| SRR13401519 |       | >60      | Male   | NA    | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR13401514 |       | <60      | Male   | NA    | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR13401518 |       | >60      | Female | NA    | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR13401517 |       | <60      | Female | NA    | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR13401516 |       | <60      | Female | NA    | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR13401515 |       | <60      | Female | NA    | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14854192 | Flu   | >60      | Male   | NA    | PBMC        | NA                        | 8–14             | NA                   | Yes      | Severe          | Alive            |
| SRR14854191 |       | <60      | Male   | NA    | PBMC        | NA                        | 8–14             | NA                   | Yes      | Severe          | Alive            |
| SRR14854190 |       | <60      | Male   | NA    | PBMC        | NA                        | 8–14             | NA                   | Yes      | Severe          | Alive            |
| SRR14854189 |       | <60      | Male   | NA    | PBMC        | NA                        | 8–14             | NA                   | No       | Severe          | Alive            |
| SRR14854188 |       | <60      | Male   | NA    | PBMC        | NA                        | ≤7               | NA                   | Yes      | Severe          | Alive            |
| SRR14854187 |       | <60      | Male   | NA    | PBMC        | NA                        | >14              | NA                   | No       | Severe          | Alive            |
| SRR14854197 |       | <60      | Female | NA    | PBMC        | NA                        | ≤7               | NA                   | No       | Severe          | Alive            |
| SRR14854196 |       | <60      | Female | NA    | PBMC        | NA                        | 8–14             | NA                   | No       | Severe          | Alive            |
| SRR14854195 |       | <60      | Female | NA    | PBMC        | NA                        | 8–14             | NA                   | Yes      | Severe          | Alive            |
| SRR14854194 |       | <60      | Female | NA    | PBMC        | NA                        | >14              | NA                   | No       | Severe          | Alive            |
| SRR14854193 |       | >60      | Female | NA    | PBMC        | NA                        | 8–14             | NA                   | No       | Severe          | Alive            |
| SRR14854186 |       | <60      | Female | NA    | PBMC        | NA                        | 8–14             | NA                   | Yes      | Severe          | Alive            |
| GSM5361256  |       | <60      | Male   | Asian | Nasal wash  | NA                        | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361255  |       | <60      | Male   | Asian | Nasal wash  | 6                         | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361254  |       | <60      | Male   | Asian | Nasal wash  | 3                         | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361253  |       | <60      | Male   | Asian | Nasal wash  | 1                         | NA               | >14                  | NA       | Severe          | NA               |
| GSM5361244  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361243  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361242  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361241  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361240  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361239  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361238  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361237  |       | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |

(Continued)

**Table 1**  
**(Continued)**

| Sample ID*  | Group            | Age (yr) | Gender | Race  | Sample type | Sampling post confirm (d) | Day post symptom | Symptom duration (d) | Ecmo use | Severity degree | Survival outcome |
|-------------|------------------|----------|--------|-------|-------------|---------------------------|------------------|----------------------|----------|-----------------|------------------|
| GSM5361236  | Flu              | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361235  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361234  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361233  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361224  |                  | >60      | Male   | Black | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361223  |                  | >60      | Male   | Black | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361222  |                  | >60      | Male   | Black | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361221  |                  | >60      | Male   | Black | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361212  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361211  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361210  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361209  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361208  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361207  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361206  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361205  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361204  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361203  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361202  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361201  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361228  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361227  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361226  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361225  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Moderate        | NA               |
| GSM5361252  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Mild            | NA               |
| GSM5361251  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Mild            | NA               |
| GSM5361250  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Mild            | NA               |
| GSM5361249  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | 8–14                 | NA       | Mild            | NA               |
| GSM5361248  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361247  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361246  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361245  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361220  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361219  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361218  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361217  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361216  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361215  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361214  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361213  |                  | <60      | Male   | White | Nasal wash  | NA                        | NA               | NA                   | NA       | Mild            | NA               |
| GSM5361232  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361231  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361230  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| GSM5361229  |                  | <60      | Female | other | Nasal wash  | NA                        | NA               | ≤7                   | NA       | Mild            | NA               |
| SRR11680214 | Healthy controls | >60      | Male   | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR11680212 |                  | >60      | Male   | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR11680213 |                  | <60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR11680210 |                  | >60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR11680209 |                  | >60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR14466439 |                  | <60      | Male   | NA    | Whole blood | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14466495 |                  | <60      | Female | NA    | Whole blood | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14466442 |                  | <60      | Female | NA    | Whole blood | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14466441 |                  | <60      | Female | NA    | Whole blood | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14466437 |                  | <60      | Female | NA    | Whole blood | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14466440 |                  | <60      | Male   | NA    | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11804730 |                  | <60      | Male   | White | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11804729 |                  | <60      | Male   | Asian | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11804728 |                  | <60      | Male   | White | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11804726 |                  | <60      | Male   | White | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11680225 |                  | >60      | Male   | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR14466457 |                  | <60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14466438 |                  | <60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11804727 |                  | <60      | Female | White | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11804725 |                  | <60      | Female | White | PBMC        | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR11680220 |                  | >60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR11680219 |                  | <60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| SRR11680211 |                  | >60      | Female | NA    | PBMC        | NA                        | NA               | NA                   | No       | NA              | NA               |
| GSM5394400  |                  | <60      | Male   | NA    | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |

(Continued)

**Table 1**  
(Continued)

| Sample ID*  | Group            | Age (yr) | Gender | Race | Sample type | Sampling post confirm (d) | Day post symptom | Symptom duration (d) | Ecmo use | Severity degree | Survival outcome |
|-------------|------------------|----------|--------|------|-------------|---------------------------|------------------|----------------------|----------|-----------------|------------------|
| GSM5394399  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394398  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394397  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394396  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394395  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394394  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394393  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5394392  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361149  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361148  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361147  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361146  |                  | <60      | Male   | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361164  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361163  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361162  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361161  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361160  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361159  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361158  | Healthy controls | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361157  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361156  | Healthy controls | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361155  |                  | <60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361154  |                  | >60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361153  |                  | >60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361152  |                  | >60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361151  |                  | >60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| GSM5361150  |                  | >60      | Female | NA   | Nasal wash  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14144058 |                  | <60      | Male   | NA   | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |
| SRR14144057 |                  | <60      | Male   | NA   | Nasal swab  | NA                        | NA               | NA                   | NA       | NA              | NA               |

BALF = bronchial alveolar lavage fluid, NA = non-available, PBMC = peripheral blood mononuclear cell.

\*Samples would be named after their Sequence Read Archive (SRA) ID. If a sample had several SRA IDs, it would be named Gene Expression Omnibus (GEO) ID.

### 3. Results

#### 3.1. COVID-19-specific immune atlas at the single-cell level

A total of 233 samples covering 8 datasets, 5 sample types (bronchoalveolar lavage fluid, peripheral blood mononuclear cells, whole blood, nasal washing, and nasal swab), 3 groups (COVID-19, flu, and healthy controls [HCs]) were included. In detail, 112, 73, and 48 samples were in COVID-19, flu, and HCs groups, respectively. The clinical information of these samples is shown in Table 1. After quality control, 434,703 cells were included into the subsequent analysis and further divided into 7 clusters including T/natural killer (NK) cells (marker genes: CD3 and GNLY), B cells (marker genes: MS4A1 and IGKC), myeloid cells (marker genes: VCAN and FCN1), epithelial cells (marker genes: KRT19 and SCGB1A1), platelets (marker genes: PF4 and PPBP), proliferating cells (marker genes: MKI67), and undefined cells (Fig. 1A and B). In terms of cell abundance, T/NK cells and myeloid cells accounted for the largest proportion. T/NK cells were reduced in both the COVID-19 and flu groups compared to the HC group, while myeloid cells and B cells were remarkably expanded in the COVID-19 and flu groups (Fig. 1C). Considering the important roles of T/NK cells and myeloid cells in COVID-19 and the significant changes in their abundance, further attention was paid to these 2 types of cells in subsequent analyses.

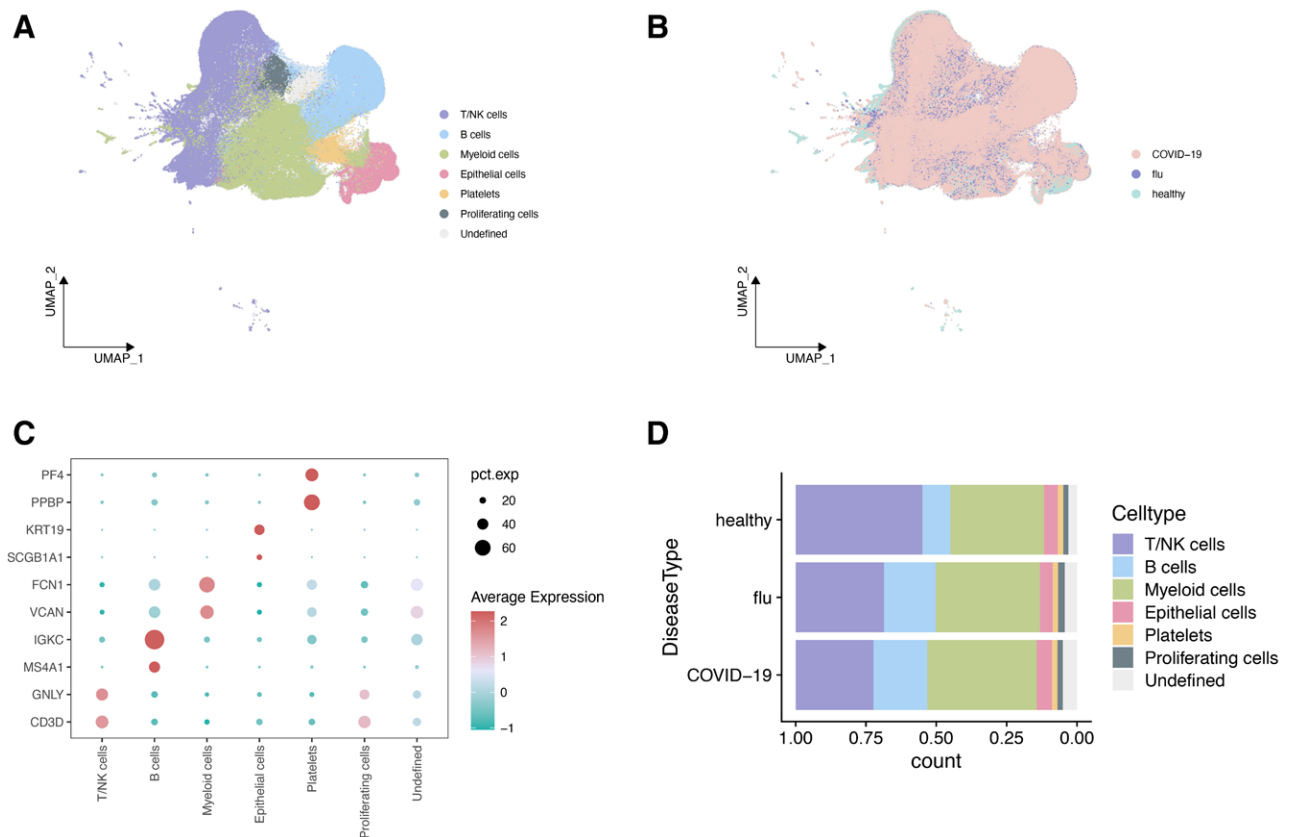
#### 3.2. COVID-19-specific T/NK cell profile

Altogether 139,592 T/NK cells were re-clustered and divided into 6 cell subtypes including CD4 + naive T cells, CD8 + central memory T cells, CD8 + effector memory T cells, CD8 + naive T cells, CD8 + terminally differentiated effector memory T cells and NK cells (Fig. 2A and B). COVID-19 patients had more

CD4 + naive T cells than HCs, while less than flu patients. Both CD8 + terminally differentiated effector memory T cells and NK cells were expanded in the COVID-19 group compared to the flu group (Fig. 2C). In terms of transcriptome, no matter compared with that of HCs or flu patients, NK cells of COVID-19 patients had the highest number of differential genes (COVID-19 vs flu: 541 DEGs; COVID-19 vs HCs: 658 DEGs) including TNFSF10, CCL3, CCL4, and others. This finding emphasizes the importance of NK cells in the onset and COVID-19. However, the transcriptome of CD8 + central memory T cells did not alter dramatically among these 3 groups. When comparing the COVID-19 group with the flu group or the HC group, there were only 16 (e.g. FOS) and 20 (e.g. HLA-C) DEGs in CD8 + central memory T cells, respectively (Fig. 3D). To further investigate the roles of these DEGs, functional analysis was carried out. Many metabolism- and inflammatory-related pathways such as fatty acid metabolism, TGF- $\beta$  signaling and IL2-STAT5 signaling were altered (Fig. 3E). For instance, apoptosis in the CD8 + effector memory T cells from COVID-19 patients were significantly enhanced, which was also reported by Kundura L et al<sup>[8]</sup> and André S et al.<sup>[9]</sup> Apoptosis is known as a kind of programmed cell death ways and it is a defense mechanism against viral infections by promoting the death of infected cells and limiting viral replication. However, excessive or dysregulated apoptosis may contribute to tissue damage and organ dysfunction seen in severe cases of COVID-19.

#### 3.3. COVID-19-specific myeloid cell profile

A total of 162,332 myeloid cells entered the re-clustering step and according to their marker genes, they were grouped as 6 subpopulations and further annotated as classical monocytes,



**Figure 1.** The single-cell atlas of COVID-19. (A) Umap plot exhibiting different cell types. (B) Umap plot exhibiting different groups. (C) Dot plot exhibiting the marker genes of each cell type. (D) Bar plot exhibiting the proportion of each cell type. COVID-19 = novel coronavirus disease, NK = natural killer.

dendritic cells, macrophages, mast cells, megakaryocytes, and neutrophils (Fig. 3A and B). Both classical monocytes and neutrophils were remarkably reduced in the COVID-19 group compared the other 2 groups, while macrophages were evidently expanded in this group (the COVID-19 group) (Fig. 3C). Notably, when comparing the COVID-19 group to the flu group or to the HC group, macrophages were the cell subtype that contained the most DEGs (Fig. 3D). In detail, there were 2378 up-regulated genes (such as ITGA4 and FLI1) and 823 down-regulated genes (such as CCL5 and COMT) in macrophages of COVID-19 patients compared with flu patients. While compared to HCs, they were 1990 (e.g., CTSS and DDX5) and 1068 (e.g., CCL4 and CCL4L2) DEGs in macrophages derived from COVID-19 patients. Functional analysis suggested that androgen response, adipogenesis, angiogenesis, apoptosis, and others were remarkably altered in these myeloid subpopulations among these 3 groups (Fig. 3E). Specifically, angiogenesis was significantly upregulated in some crucial myeloid cells including classical monocytes and macrophages from COVID-19 patients and flu patients, while this process was downregulated in these myeloid cell subpopulations from HCs. The relationship between angiogenesis and COVID-19 pathogenesis has been widely reported and SARS-CoV-2-induced angiogenesis is considered to be a major factor in the development of COVID-19.<sup>[10]</sup>

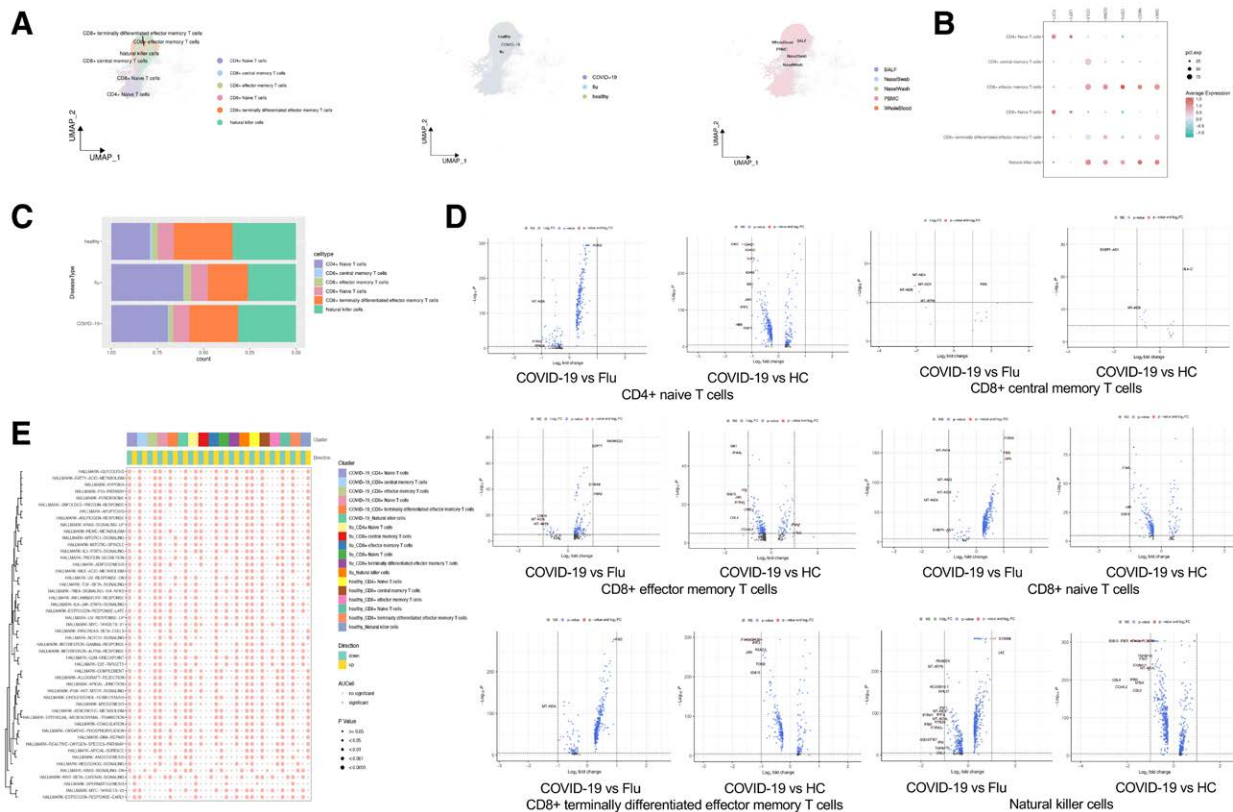
### 3.4. Gene expression patterns of COVID-19 severity

An important aspect of COVID-19 heterogeneity is disease severity. Identifying the characteristics of different severity degree is essential for early prediction and timely adjustment of treatment strategies. Therefore, this work further investigated the gene expression features of COVID-19 severity. For T cell subtypes, the transcriptome of CD8 + effector memory

T cells, CD8 + terminally differentiated effector memory T cells, and NK cells dramatically changed among COVID-19 patients with different severities and altered genes mainly related to inflammatory such as S100A8 and S100A9 and cell cytotoxicity such as GZMK. However, CD8 + central memory T cells did not undergo significantly transcriptomic reprogramming (Fig. 4A–F). The above results emphasize the value of CD8 + effector memory T cells, CD8 + terminally differentiated effector memory T cells, and NK cells in the COVID-19 development. For myeloid subclusters, except for mast cells and megakaryocytes, the transcriptome of the remaining myeloid cell subpopulations showed significant differences among COVID-19 patients with different severities (Fig. 4G–L). In these myeloid subtypes (classical monocytes, dendritic cells, macrophages, and neutrophils), DGEs among COVID-19 patients with different severities, most genes were responsible for generating inflammatory factors (e.g., CCL5, CXCL8, and IL1B). In COVID-19, an excessive immune response to viral infection can lead to the occurrence of a cytokine storm, which can result in a sustained enhancement of the inflammatory cascade, leading to tissue organ damage and severe pathological processes. The cytokine storm is considered one of the important factors contributing to the deterioration of the condition and increased risk of complications in COVID-19. Our results also underscore the importance of inflammatory factors in the development or progression of COVID-19.

### 3.5. Predictive model of COVID-19-related outcomes

The ultimate purpose of analyzing the disease severity of COVID-19 patients is to improve the prognosis of these patients, making the prediction of patient outcomes crucial. Currently,



**Figure 2.** COVID-19-specific T cell atlas. (A) Umap plot exhibiting different T cell subtypes, groups and sample types, respectively. (B) Dot plot exhibiting the marker genes of each T cell subtype. (C) Dot plot exhibiting the marker genes of each T cell subtype, as well as between COVID-19 patients and flu patients, as well as between COVID-19 patients and healthy controls, in each T cell subtype. (D) Volcano plot exhibiting the differentially expressed genes between COVID-19 patients and flu patients, as well as between COVID-19 patients and healthy controls, in each T cell subtype. (E) Functional enrichment exhibiting significantly altered pathways in each T cell subtypes for different groups. COVID-19 = novel coronavirus disease, HC = healthy control.

COVID-19-related deaths are predominantly observed in severe cases. Therefore, this study aims to construct a predictive model for the outcomes of severe COVID-19 patients.

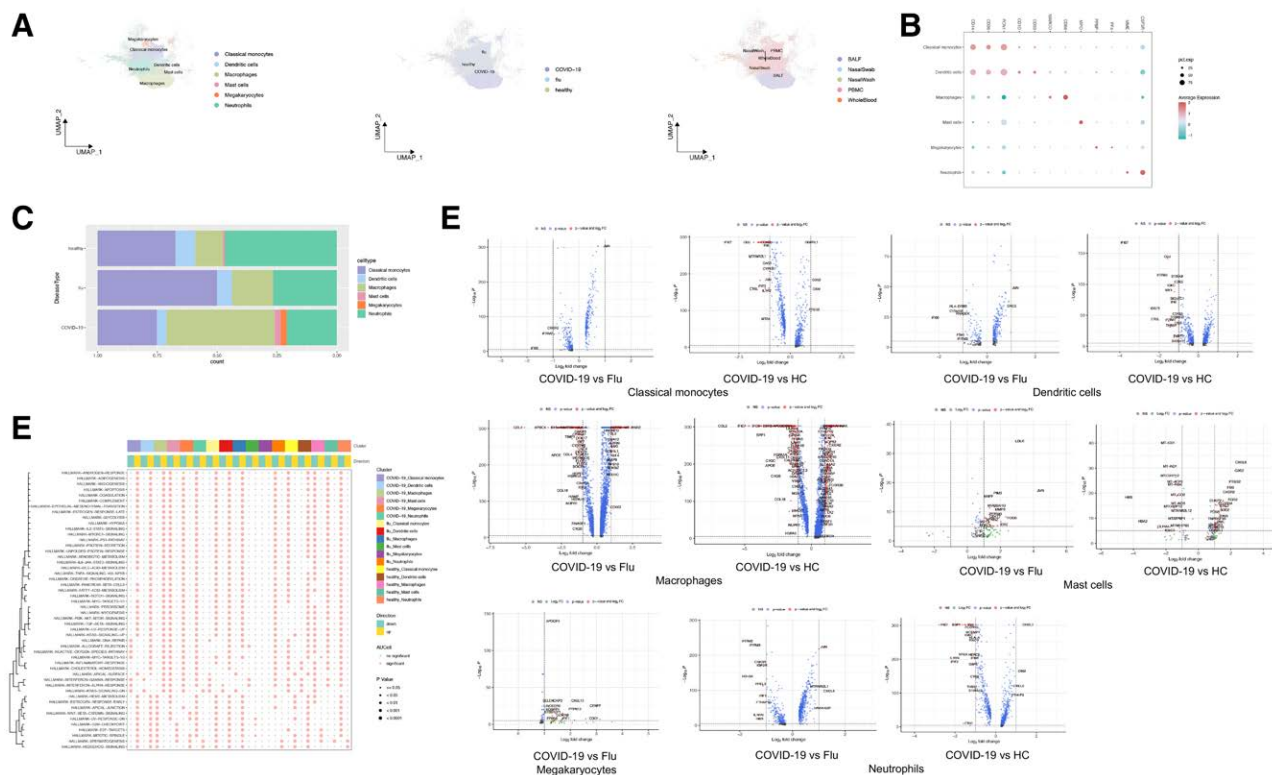
The transcriptomic changes between severe COVID-19 with different outcomes were analyzed and there were 5677 DEGs (upregulation: 3705 genes; downregulation: 1792 genes) between severe COVID-19 with different outcomes (death or alive). Based on these DEGs, univariate and multifactor logistics regression were used to select variables. A total of 4 genes, *IL1R2*, *PI3*, *IGHG3*, and *CTTN*, were considered as inputted variables. The expression of these 4 genes is shown in Figure 5A. Then, a predictive model was constructed by integrating these 4 genes. In the training set, the sensitivity and AUC of this model were 80.0% and 70.5%, respectively, while the specificity was only 56.3% (Fig. 5B). In the test set, the sensitivity, specificity and AUC of this model were 81.1%, 55.5%, and 70.6%, respectively (Fig. 5C). The above results indicate that the performance of the model is stable. In the case of predicting the outcome of severe COVID-19 patients, sensitivity bit specificity should be paid more attention and therefore, this model can well predict the survival outcomes of severe cases.

#### 4. Discussion

This work provides a systemic single-cell landscape of COVID-19 by integrating diverse patient samples and data types. Our analysis not only deepens the understanding of disease mechanisms but also identifies gene expression patterns correlated with severity and survival, culminating in a predictive model with strong clinical potential.

Our findings corroborate and extend previous observations on the critical role of specific immune cell populations. For

instance, we observed marked alterations in cytotoxic cell types, including CD8 + effector memory T cells and NK cells, as well as monocytes and macrophages. These cells are instrumental in the immune response against SARS-CoV-2, primarily through the elimination of infected cells via perforin and granzyme pathways.<sup>[11,12]</sup> Our transcriptomic data provides single-cell resolution evidence supporting the concept<sup>[13]</sup> that the inflammatory microenvironment in severe COVID-19, potentially driven by complement factors like C3a, promotes the differentiation of highly cytotoxic CD16 + T cells. These cells, through interactions with the endothelium, can exacerbate microvascular injury and multiorgan damage, thus increasing mortality risk.<sup>[14]</sup> Similarly, our analysis reinforces the significance of macrophages in COVID-19 pathogenesis. While traditionally not considered primary targets, evidence suggests SARS-CoV-2 can infect monocytes and macrophages, possibly through ACE2-independent mechanisms.<sup>[15]</sup> The subsequent pyroptosis of these infected cells is a potent trigger for the hyperinflammation central to severe disease.<sup>[15,16]</sup> The dysfunction of these myeloid cells can therefore fuel the cytokine storm and induce fibrotic complications, leading to acute respiratory distress syndrome (ARDS).<sup>[17]</sup> A key contribution of our predictive model is the high weight assigned to *IL1R2*. As a decoy receptor for the potent proinflammatory cytokine IL-1 $\beta$ , *IL1R2* acts as a crucial negative regulator of IL-1 signaling.<sup>[18–20]</sup> While IL-1 is a well-established driver in various inflammatory diseases<sup>[21]</sup> and *IL1R2* is a known therapeutic target in cancer and acute respiratory distress syndrome,<sup>[22]</sup> its direct role in COVID-19 has been unclear. Our study is among the first to quantitatively link *IL1R2* expression to COVID-19 survival outcomes, suggesting it may be a pivotal checkpoint in controlling virus-induced inflammation and a promising therapeutic target.



**Figure 3.** COVID-19-specific myeloid cell atlas. (A) Umap plot exhibiting different myeloid cell subtypes, groups and sample types, respectively. (B) Dot plot exhibiting the marker genes of each myeloid cell subtype. (C) Dot plot exhibiting the marker genes of each myeloid cell subtype. (D) Volcano plot exhibiting the differentially expressed genes between COVID-19 patients and flu patients, as well as between COVID-19 patients and healthy controls, in each myeloid cell subtype. (E) Functional enrichment exhibiting significantly altered pathways in each myeloid cell subtypes for different groups. COVID-19 = novel coronavirus disease, HC = healthy control.

Despite the comprehensive nature of our analysis, we acknowledge several limitations that also frame important avenues for future research. First, while integrating multiple datasets enhanced the robustness of our findings, the inherent technical variations across different single-cell RNA sequencing platforms and protocols could introduce batch effects, despite our efforts to mitigate them during data integration. Second, our large-scale integration, while providing a broad overview, may overlook the dynamics of rare but potentially critical cell subpopulations. The averaging effect across a massive number of cells might dilute the signatures of low-abundance cell types that could play an outsized role in disease progression. Future studies could employ targeted enrichment strategies or deeper sequencing of specific cell compartments to unmask the roles of these rare populations.

In conclusion, this study offers a valuable single-cell reference for COVID-19 and provides a robust prognostic model. By acknowledging its limitations, we pave the way for more targeted and mechanistically insightful research that can translate these findings into effective clinical strategies against current and future viral pandemics.

## 5. Conclusion

This study offered a comprehensive single-cell atlas of COVID-19 that integrates various sample types and patients with different characteristics (severity, prognosis, etc). The developed

prognostic prediction model can assist clinicians in making decisions and further improve the patient prognosis, providing evidence-based strategies to combat the ongoing pandemic and future viral threats.

## Author contributions

**Conceptualization:** Pei Tang, Yong Qing, Junliang Wu.

**Data curation:** Jinghan Wang, Junliang Wu.

**Formal analysis:** Pei Tang, Jinghan Wang, Hairui Li, Yong Qing, Gang Zhong.

**Investigation:** Pei Tang, Hairui Li, Yong Qing.

**Methodology:** Pei Tang, Hairui Li.

**Project administration:** Zhixing Chen.

**Resources:** Yong Qing.

**Software:** Yong Qing.

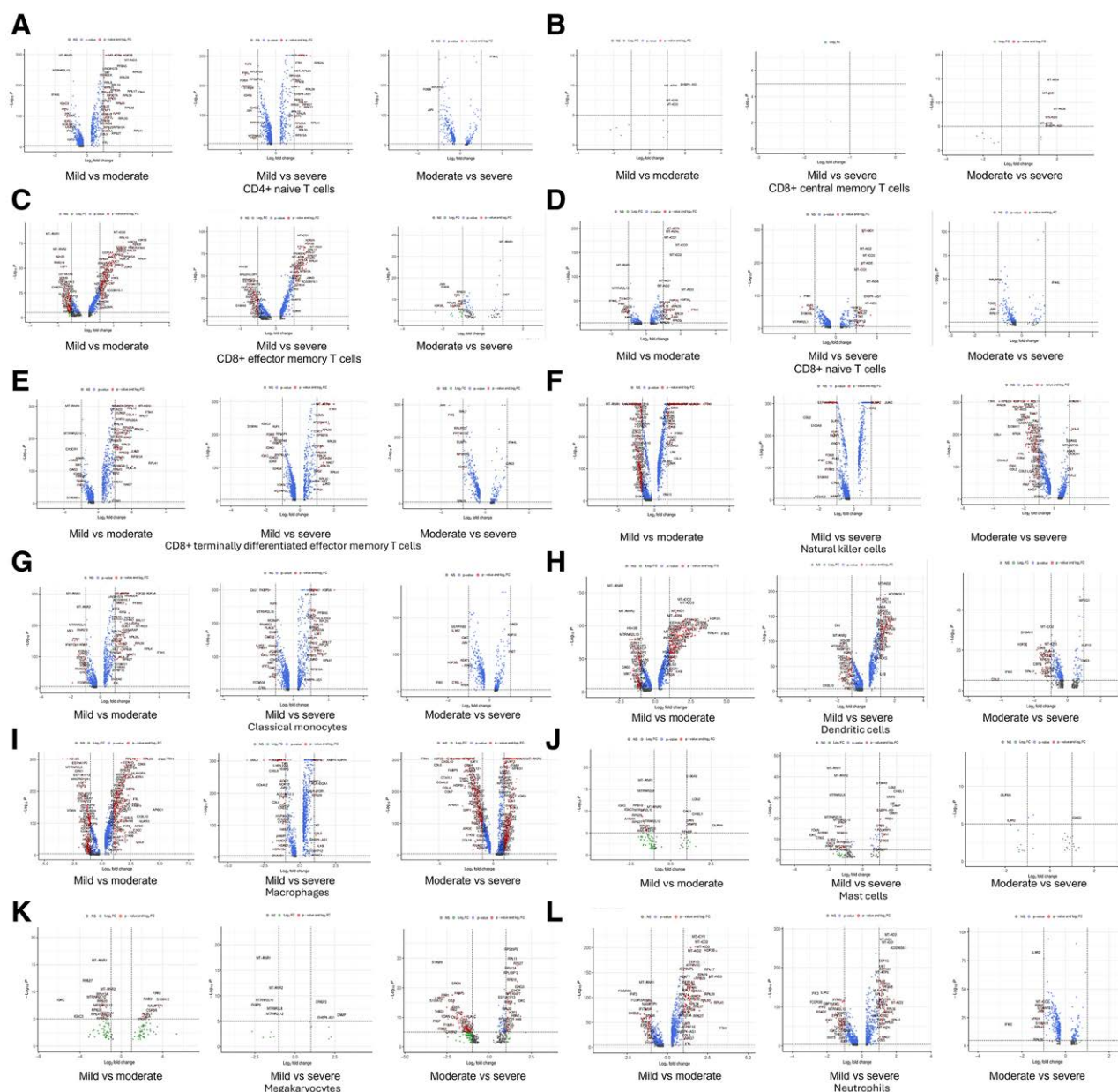
**Supervision:** Yong Qing, Zhixing Chen.

**Validation:** Jinghan Wang, Zhixing Chen, Xiaolin Ran, Gang Zhong.

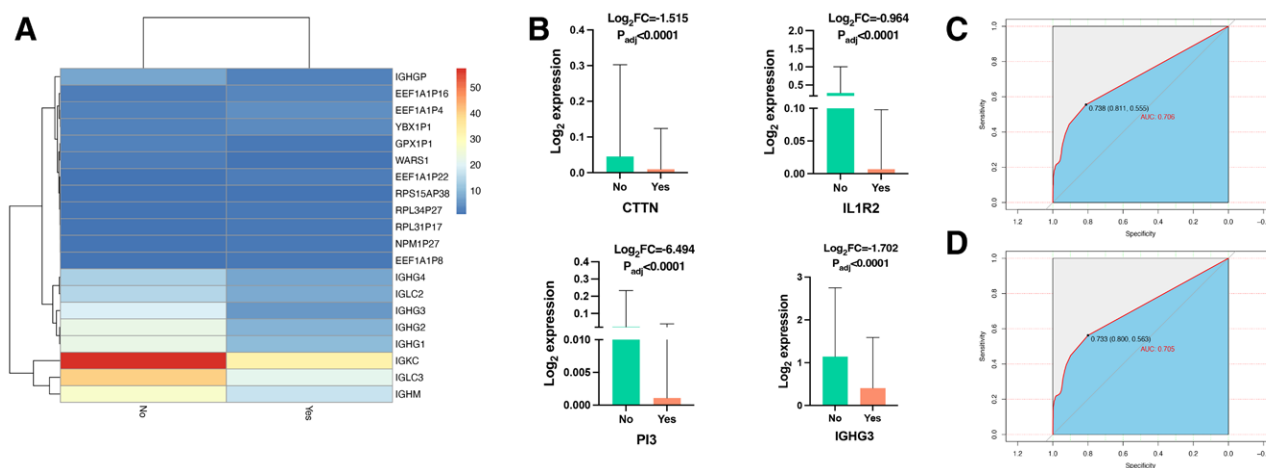
**Visualization:** Jinghan Wang, Hairui Li, Yong Qing, Xiaolin Ran, Gang Zhong.

**Writing – original draft:** Jinghan Wang, Hairui Li, Zhixing Chen, Junliang Wu, Xiaolin Ran.

**Writing – review & editing:** Zhixing Chen, Junliang Wu, Xiaolin Ran, Gang Zhong.



**Figure 4.** Gene expression pattern of COVID-19 patients with different severities. (A–F) Volcano plots showing the differentially expressed genes between mild and moderate COVID-19 patients, between mild and severe COVID-19 patients, and between moderate and severe COVID-19 patients in CD4 + naive T cells, CD8 + central memory T cells, CD8 + effector memory T cells, CD8 + naive T cells, CD8 + terminally differentiated effector memory T cells, and natural killer cells, respectively. (G–L) Volcano plots showing the differentially expressed genes between mild and moderate COVID-19 patients, between mild and severe COVID-19 patients, and between moderate and severe COVID-19 patients in classical monocytes, dendritic cells, macrophages, mast cells, megakaryocytes, and neutrophils, respectively.



**Figure 5.** The expression of inputted genes and model performance. (A) Heatmap showing the top 20 differentially expressed genes between severe patients with death or alive outcomes. (B) Bar plots showing the expression of inputted genes in patients with different outcomes. (C and D) Receiver operating characteristic curve showing the predictive performance of the constructed model in the training set and the test set, respectively. FC = fold change.

## References

- Aslam S, Goldstein DR, Vos R, et al. COVID-19 vaccination in our transplant recipients: the time is now. *J Heart Lung Transplant*. 2021;40:169–71.
- Harrison AG, Lin T, Wang P. Mechanisms of SARS-CoV-2 transmission and pathogenesis. *Trends Immunol*. 2020;41:1100–15.
- Bourgonje AR, Abdulle AE, Timens W, et al. Angiotensin-converting enzyme 2 (ACE2), SARS-CoV-2 and the pathophysiology of coronavirus disease 2019 (COVID-19). *J Pathol*. 2020;251:228–48.
- Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395:1054–62.
- Liao M, Liu Y, Yuan J, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med*. 2020;26:842–4.
- Ren X, Wen W, Fan X, et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell*. 2021;184:1895–913.e19.
- Wilk AJ, Rustagi A, Zhao NQ, et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat Med*. 2020;26:1070–6.
- Kundura L, Gimenez S, Cezar R, et al. Angiotensin II induces reactive oxygen species, DNA damage, and T-cell apoptosis in severe COVID-19. *J Allergy Clin Immunol*. 2022;150:594–603.e2.
- André S, Picard M, Cezar R, et al. T cell apoptosis characterizes severe Covid-19 disease. *Cell Death Differ*. 2022;29:1486–99.
- Ackermann M, Verleden SE, Kuehnel M, et al. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. *N Engl J Med*. 2020;383:120–8.
- Jung JH, Rha MS, Sa M, et al. SARS-CoV-2-specific T cell memory is sustained in COVID-19 convalescent patients for 10 months with successful development of stem cell-like memory T cells. *Nat Commun*. 2021;12:4043.
- Rha MS, Jeong HW, Ko JH, et al. PD-1-expressing SARS-CoV-2-specific CD8(+) T cells are not exhausted, but functional in patients with COVID-19. *Immunity*. 2021;54:44–52.e3.
- Huang P, Zhang J, Duan W, Jiao J, Leng A, Qu J. Plant polysaccharides with anti-lung injury effects as a potential therapeutic strategy for COVID-19. *Front Pharmacol*. 2022;13:982893.
- Georg P, Astaburuaga-García R, Bonaguro L, et al. Complement activation induces excessive T cell cytotoxicity in severe COVID-19. *Cell*. 2022;185:493–512.e25.
- Junqueira C, Crespo A, Ranjbar S, et al. FcγR-mediated SARS-CoV-2 infection of monocytes activates inflammation. *Nature*. 2022;606:576–84.
- Del Valle DM, Kim-Schulze S, Huang HH, et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat Med*. 2020;26:1636–43.
- Meidaninikjeh S, Sabouni N, Marzouni HZ, Bengar S, Khalili A, Jafari R. Monocytes and macrophages in COVID-19: friends and foes. *Life Sci*. 2021;269:119010.
- Garlanda C, Dinarello CA, Mantovani A. The interleukin-1 family: back to the future. *Immunity*. 2013;39:1003–18.
- Boraschi D, Italiani P, Weil S, Martin MU. The family of the interleukin-1 receptors. *Immunol Rev*. 2018;281:197–232.
- Liew F, Efstathiou C, Fontanella S, et al. Large-scale phenotyping of patients with long COVID post-hospitalization reveals mechanistic subtypes of disease. *Nat Immunol*. 2024;25:607–21.
- Peters VA, Joesting JJ, Freund GG. IL-1 receptor 2 (IL-1R2) and its role in immune regulation. *Brain Behav Immun*. 2013;32:1–8.
- Kovach MA, Stringer KA, Bunting R, et al. Microarray analysis identifies IL-1 receptor type 2 as a novel candidate biomarker in patients with acute respiratory distress syndrome. *Respir Res*. 2015;16:29.