



OPEN ACCESS

EDITED BY
Roberta Rizzo,
University of Ferrara, Italy

REVIEWED BY
Ke Lin,
Fudan University, China
Sudarson Sundarajan,
Cancyte Technologies, India

*CORRESPONDENCE

Ji Wang
✉ doctorwang2009@126.com
Qi Wang
✉ wangqi710@126.com

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 23 July 2025
ACCEPTED 15 September 2025
PUBLISHED 02 October 2025

CITATION

Zhou Y, Fan P, Zhang H, Han S, Bai M, Wang J
and Wang Q (2025) Multi-omics and machine
learning identify novel biomarkers and
therapeutic targets of COVID-19.
Front. Immunol. 16:1671936.
doi: 10.3389/fimmu.2025.1671936

COPYRIGHT

© 2025 Zhou, Fan, Zhang, Han, Bai, Wang and
Wang. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums
is permitted, provided the original author(s)
and the copyright owner(s) are credited and
that the original publication in this journal is
cited, in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Multi-omics and machine learning identify novel biomarkers and therapeutic targets of COVID-19

Yumei Zhou^{1†}, Pengbei Fan^{2†}, Haiyun Zhang^{1,3}, Shuai Han⁴,
Minghua Bai¹, Ji Wang^{1*} and Qi Wang^{1,5*}

¹National Institute of Traditional Chinese Medicine (TCM) Constitution and Preventive Treatment of Disease, Wangqi Academy of Beijing University of Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China, ²School of Traditional Chinese Medicine, Southern Medical University, Guangzhou, China, ³Medical Laboratory Center, Dalian Municipal Women and Children's Medical Center (Group), Dalian, China, ⁴Laboratory Animal Center of Inner Mongolia Medical University, Hohhot, China, ⁵Hubei Shizhen Laboratory, Hubei University of Chinese Medicine, Wuhan, China

Introduction: COVID-19 has caused over 7 million deaths worldwide since its onset in 2019, and the virus remains a significant health threat. Identifying sensitive and specific biomarkers, along with elucidating immune-mediated mechanisms, is essential for improving the diagnosis, treatment, and prevention of COVID-19. To predict key molecular markers of COVID-19 using an established multi-omics framework combined with machine learning models.

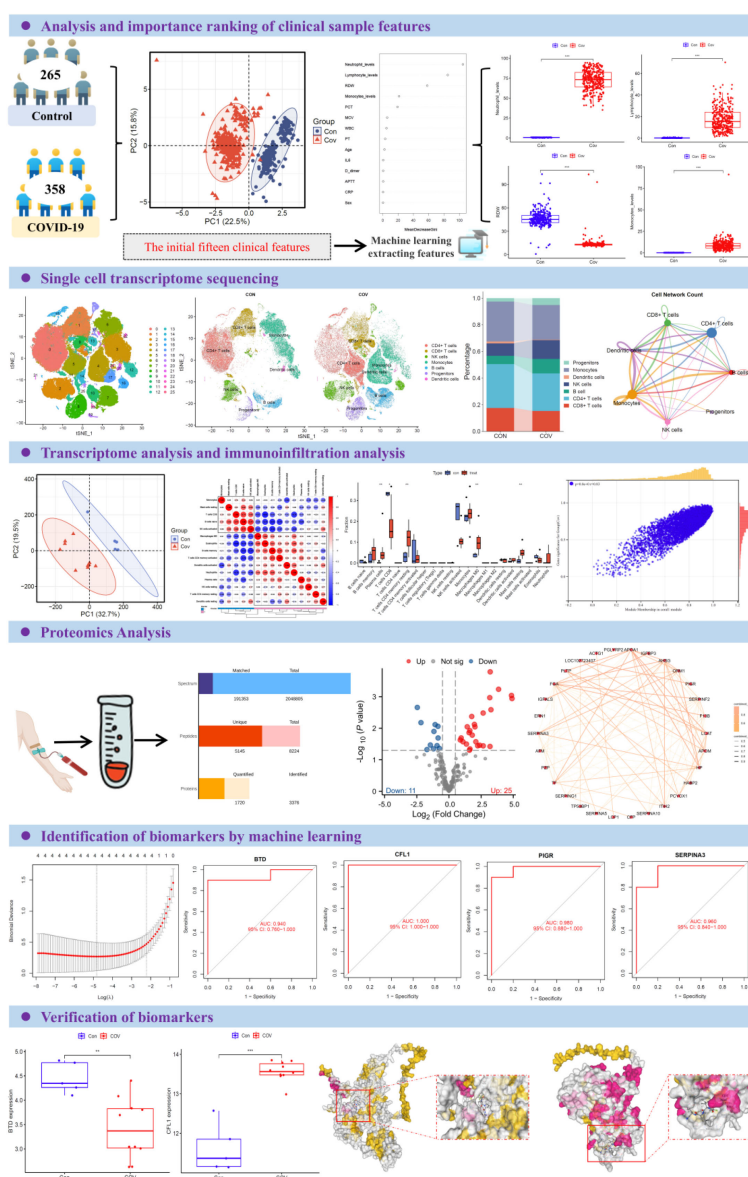
Methods: We conducted an integrated analysis of single-cell RNA sequencing (scRNA-seq), bulk RNA sequencing, and proteomics data to identify critical biomarkers associated with COVID-19. The multi-omics approach enabled the characterization of gene expression dynamics and alterations in immune cell subsets in COVID-19 patients. Machine learning techniques and molecular docking analyses were employed to identify biomarkers and therapeutic targets within the disease's pathophysiological network.

Results: Principal component analysis effectively grouped samples based on clinical characteristics. Using random forest and SVM-RFE models, we identified clinical indicators capable of accurately distinguishing COVID-19 patients. Transcriptomic analysis, including scRNA-seq, highlighted the pivotal role of CD8⁺ T cells, and WGCNA identified related module genes. Proteomic analysis, integrated with machine learning, revealed 36 DEPs. Further investigation identified several genes associated with monocyte proportions. Correlation analysis showed that BTG, CFL1, PIGR, and SERPINA3 were strongly linked to CD8⁺ T cell abundance in COVID-19 patients. ROC curve analysis demonstrated that these genes could effectively distinguish between COVID-19 patients and healthy individuals. Concordant findings from both transcriptomic and proteomic levels support BTG, CFL1, PIGR, and SERPINA3 as potential auxiliary diagnostic markers. Finally, AlphaFold-based molecular docking analysis suggested these biomarkers may also serve as candidate therapeutic targets.

Discussion: Preliminary findings indicate that *BTD*, *CFL1*, *PIGR*, and *SERPINA3* are vital molecular biomarkers related of CD8+ T cell, providing new insights into the molecular mechanisms and long-term prevention of COVID-19.

KEYWORDS

multi-omics, scRNA-seq, RNA-seq, biomarkers, COVID-19



GRAPHICAL ABSTRACT

Overview of the study design. The participants were divided into two groups: control group (n=265) and COVID-19 group (n=358). Principal component analysis was used to group samples according to clinical characteristics. Selected features were analyzed based on random forest model and SVM-REF model. By integrating scRNA-seq and RNA-seq data with an analysis of the peripheral plasma proteome, we applied machine learning models to successfully identify and predict potential biomarkers associated with CD8+ T cell responses in COVID-19 infection. ROC curve analysis was used to analyze the clinical diagnostic efficacy of inflammatory factors among different groups. The box diagram was used to show the levels of different inflammatory factors in plasma of different populations.

Background

Despite a decline in the overall burden of COVID-19, the threat posed by the virus remains significant and persistent (1, 2). COVID-19, caused by the zoonotic SARS-CoV-2, continues to spread through respiratory droplets and direct contact, posing ongoing challenges to public health (3). Despite the flood of insights into the behavior of the virus and how to prevent it from causing harm (4, 5), many questions remain regarding the duration and quality of immunity after reinfection or vaccination, the impact of coinfections, and optimal treatment protocols for different populations. Severe COVID-19 is characterized by pneumonia, lymphopenia, exhausted lymphocytes and a cytokine storm (6). However, whether this is protective or pathogenic remains to be determined. Critical COVID-19-related illnesses continue to occur, and the virus's impact on immune function remains insufficiently understood (7). Therefore, identifying sensitive and specific molecular markers, uncovering key immune mechanisms, and discovering new therapeutic targets are not only important but also critically necessary. Continued research is essential to manage COVID-19 effectively, mitigate its long-term effects, and prepare for future threats.

The integrated analysis of single cell transcriptome (scRNA-seq), bulk RNA-seq and proteomics has become a frontier means to analyze the mechanism of complex diseases and find accurate biomarkers (8). Single cell transcriptome technology can analyze the heterogeneity of cells with high resolution, and reveal the gene expression dynamics of specific cell subsets in microenvironment of COVID-19 patients (9, 10). However, the sparsity and high noise of single cell data need to be optimized by preprocessing and dimensionality reduction clustering, while the bulk transcriptome data provide the global expression characteristics of transcription level, and the two can be complemented by algorithms to infer the changes of cell subsets during the disease process (11). Proteomics data directly reflect the activity of functional molecules and make up for the lack of information in post-transcriptional regulation (12). Multimodality integration improves accuracy by combining single-cell data, bulk transcriptome, and serum proteomics, advancing early disease screening, target discovery, and personalized treatment in precision medicine.

COVID-19 infection can trigger an inappropriate immune response, leading to excessive activation of immune cells, tissue inflammation, and even multi-organ dysfunction (13). Current studies indicate that T cell-mediated immunity plays a crucial role in the effective antiviral response to COVID-19 (14). In this

context, CD8+ T cells form an initial line of defense by rapidly recognizing viral antigens (15). COVID-19 infection induces CD8+ T cells to initiate cytotoxic responses, with cytotoxic T lymphocytes being responsible for clearing infected cells, thus playing a key role in controlling the virus (16, 17). Regulatory T cells also help prevent severe COVID-19 outcomes by mitigating excessive host inflammatory responses (18). In summary, cellular immunity, particularly CD8+ T cells, is vital for monitoring and predicting COVID-19 progression, aiding in early detection of severe cases and guiding clinical treatment. While patients with COVID-19 often present with lymphopenia, the disease has also been linked to immune hyperresponsiveness (19). Notably, the mechanisms behind immune cell dysfunction are not yet fully understood.

In this study, we performed integrated analyses using scRNA-seq, RNA-seq, proteomics, and AlphaFold-based molecular docking to investigate the alterations in key biomarkers associated with COVID-19. Functional enrichment analysis and machine learning were subsequently employed to identify pathogenic genes and elucidate their role within the pathophysiological network of COVID-19. By combining genetic evidence with clinical trial data, we developed a multi-omics framework to prioritize immune-mediated drug targets for COVID-19. Preliminary findings indicate that biotinidase (BTD), Cofilin 1 (CFL1), Polymeric Immunoglobulin Receptor (PIGR), and Alpha-1-antichymotrypsin (SERPINA3) are critical molecular markers associated with CD8+ T cells in the context of COVID-19, suggesting their potential as biomarkers for diagnosis and therapeutic intervention.

Methods

Clinical data collection

In this study, we retrospectively analyzed the test data of 358 COVID-19 patients and 265 healthy people who were treated in the Second Affiliated Hospital of Mudanjiang Medical College from January to June 2023. This study was conducted in strict accordance with the Declaration of Helsinki and was approved by the Medical Ethics Review Committee of the Second Affiliated Hospital of Mudanjiang Medical College (No. 202328). All participants gave their consent to participate in the study.

Clinical features extract

Principal component analysis (PCA) was performed using the “prcomp” function from the “stats” R package based on the 14 clinical features. We utilized the randomForest v4.7 package to construct a random forest model for ranking the importance of these clinical features and evaluating their performance as indicators. Specifically, the number of decision trees (such as $n \text{ tree} \geq 500$), the maximum depth and other parameters are set, and the Gini index decline or replacement importance is used to evaluate the feature contribution.

Abbreviations: BTD, Biotinidase; CIBERSORT, Cell-type identification by estimating relative subsets of protein; CFL1, Cofilin 1; DDA, Data-dependent acquisition; GO, Gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; LASSO, Least absolute shrinkage and selection operator; Paxlovid, Nirmatrelvir/Ritonavir; PCA, Principal component analysis; PIGR, Polymeric immunoglobulin receptor; RDW, Red cell distribution width; RF, Random forest; scRNA-seq, Single cell transcriptome; SERPINA3, Alpha-1-antichymotrypsin; TOM, Topological overlap measure; WGCNA, Weighted gene co-expression network analysis.

Single-cell RNA-seq

The scRNA-seq data used in this study were screened from the public database GEO database platform. The scRNA-seq dataset (GSE192391) used is based on the Illumina NovaSeq assay platform and contains 12 patients and 6 controls (10). Subsequent analysis was performed using the R Seurat v5.2.1 package. The following quality control steps were performed to filter the count matrices: Genes expressed in <3 cells and cells expressing fewer than 200 genes were removed; Cells expressing >5000 genes were discarded as these could be potential multiplet events where more than a single cell was encapsulated within the same barcoded GEM; Cells with >10% mitochondrial content was filtered out as these were deemed to be of low-quality. The data set consists of 18 samples. In order to remove the batch effect except the processing factors between different samples, data integration is carried out based on rPCA method. After data integration, the cells were clustered by principal component analysis and dimensionality reduction. According to the markers of human cell subsets determined in previous studies, cell annotation was carried out and the proportion of each cell subset was calculated. In this study, the R CellChat v1.6.1 package was used to infer the interaction between cells based on the expression of receptors and ligands on the cell surface. The algorithm simulates the probability of cell-to-cell communication by inputting the gene expression data of cells and combining the gene expression with the prior data of the interaction between signal ligands, receptors and their co-factors.

DEG analysis, functional enrichment and immune infiltration analysis

The RNA-Seq data used in this study were GSE164805 included 10 COVID patients and 5 controls (20). Using R limma v3.62.2 package, according to the sample grouping, the differences between standard groups were analyzed. Cluster analysis was used to calculate PCA and FactoMine package drawing. Transcriptomics data were processed by R DESeq2 v1.46.0 package to analyze the difference of the original Counts matrix, and follow the standard process. In this study, $|\text{Log}_2\text{FC}| > 0.5$ and $P < 0.05$ were set for differential expression genes filtration. Gene ontology (GO) enrichment analysis of DEGs was implemented in the R GO.db v3.20.0 packages based on Wallenius non-central hypergeometric distribution (<https://geneontology.org/>). The DEGs were analyzed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<https://www.genome.jp/kegg/pathway.html>). Protein-protein interaction network was conducted in STRING (<https://string-db.org/>). Transcriptome data was transformed into the total abundance of immune cells by utilizing the Cell-type Identification by Estimating Relative Subsets of protein (CIBERSORT) analysis with the R CIBERSORT v0.1.0 package. Using the Wilcoxon test, immune cells were compared among COVID samples and control samples.

Construction of co-expression network

Weighted gene co-expression network analysis (WGCNA) was utilized to identify gene modules with similar expression patterns and analyze the correlation between monocyte proportion and gene modules. The scale independence and average connectivity of the networks were tested with different power values (from 1 to 20). The appropriate power value was determined when the independent scale was greater than 0.85 and the connectivity was high. Then, the similarity matrix was transformed into a topological matrix with the topological overlap measure (TOM) describing the correlation between genes. The genes were clustered by using 1-TOM as the distance. A dynamic hybrid cutting method was used to establish a hierarchical clustering tree to identify co-expressed gene modules. Each leaf of the tree represents a gene, and genes with similar expression data aggregate to form a branch of the tree and each branch represents a gene module. A weighted co-expression network model was established, and the gene expression matrix was divided into several related modules. Finally, the modules related to KOA and immune cell infiltration were selected for further analysis.

Label-free protein profiling detection

Protein quantification was performed using the Bradford protein quantification kit (purchased from Shanghai Biotechnology Company). Trypsin (purchased from Solebac (Beijing)) was added to each protein sample and incubated at 37°C for 4 h. CaCl_2 was added to each sample and digested overnight. Formic acid was added, centrifuged at 12,000 g for 5 min, and the resulting supernatant was loaded onto a C18 desalting column. LC-MS/MS analysis was performed using an EASY-nLC TM 1200 UHPLC and Q Exactive TM HF-X mass spectrometry system, both purchased from ThermoFisher (USA), operated in data-dependent acquisition (DDA) mode.

DEP analysis, functional enrichment and PPI analysis

Proteomics data were processed by R DEP version 1.26.0 package. Proteomic data were then normalized using “normalize_vsn” function. Cluster analysis was used to calculate PCA and R package FactoMine drawing. “add_rejections” function was applied to calculate the fold-change values of proteins. In this study, $|\text{Log}_2\text{FC}| > 0.5$ and $P < 0.05$ were set for differential expression proteins filtration. GO enrichment analysis of DEPs was implemented in the Goseq R packages based on Wallenius non-central hypergeometric distribution. The DEPs were analyzed using the KEGG database. Protein-protein interaction network was conducted in STRING.

Machine learning techniques

Machine learning was employed as a tool to improve biomarker screening. First, feature selection was performed using the Random Forest (RF) algorithm, which calculates the mean decrease in Gini (MDG) for each gene or protein, ranks them accordingly, and determines the optimal number of features by sequentially adding differential genes/proteins from highest to lowest MDG until classification accuracy is maximized. Next, the least absolute shrinkage and selection operator (LASSO), implemented via the “glmnet” R package, was applied as a regularized regression approach to further refine the selected features. Finally, an artificial neural network (ANN) analysis was conducted with the R packages “neuralnet” for model construction and training, and “NeuralNetTools” for visualization and interpretation of the results. This integrated workflow enabled robust identification and evaluation of candidate biomarkers.

ROC analysis and validation of the biomarkers

In addition, the expression level and diagnostic value of biomarkers was verified in both proteomics and transcriptomics data. We examined the diagnostic effectiveness of the biomarkers with the ROC using the R pROC v1.18.5 package. The expression levels of the biomarkers were also compared between COVID and control samples using an independent t-test, with $P < 0.05$ considered statistically significant.

AlphaFold-based molecular docking

To assess the binding affinity between the COVID-19 antiviral drug Nirmatrelvir/Ritonavir (Paxlovid) (21) and key biomarkers, the 3D structures of core target proteins were predicted using AlphaFold DB (<https://alphafold.ebi.ac.uk/>), while the 3D structure of Paxlovid (PubChem CID, 155903259) was retrieved from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>). Binding affinity was evaluated using CB-Dock2, with a docking score ≤ -5.0 kcal/mol considered indicative of strong interaction (22, 23).

Statistical analysis

The normal distribution of the data was assessed by the Shapiro-Wilk normality test using or SPSS Statistics v27.0 (IBM, USA), and the mean is expressed as mean $\pm$ standard error. If the data were normally distributed, the differences between the means were assessed by one-way analysis of variance and Tukey’s multiple comparison test. If the data were not normally distributed, scores were compared using the nonparametric Kruskal-Wallis test. When $p < 0.05$, the difference between groups was considered statistically significant. A professional biostatistician reviewed the study and confirmed that the sample size is appropriate and statistically justified.

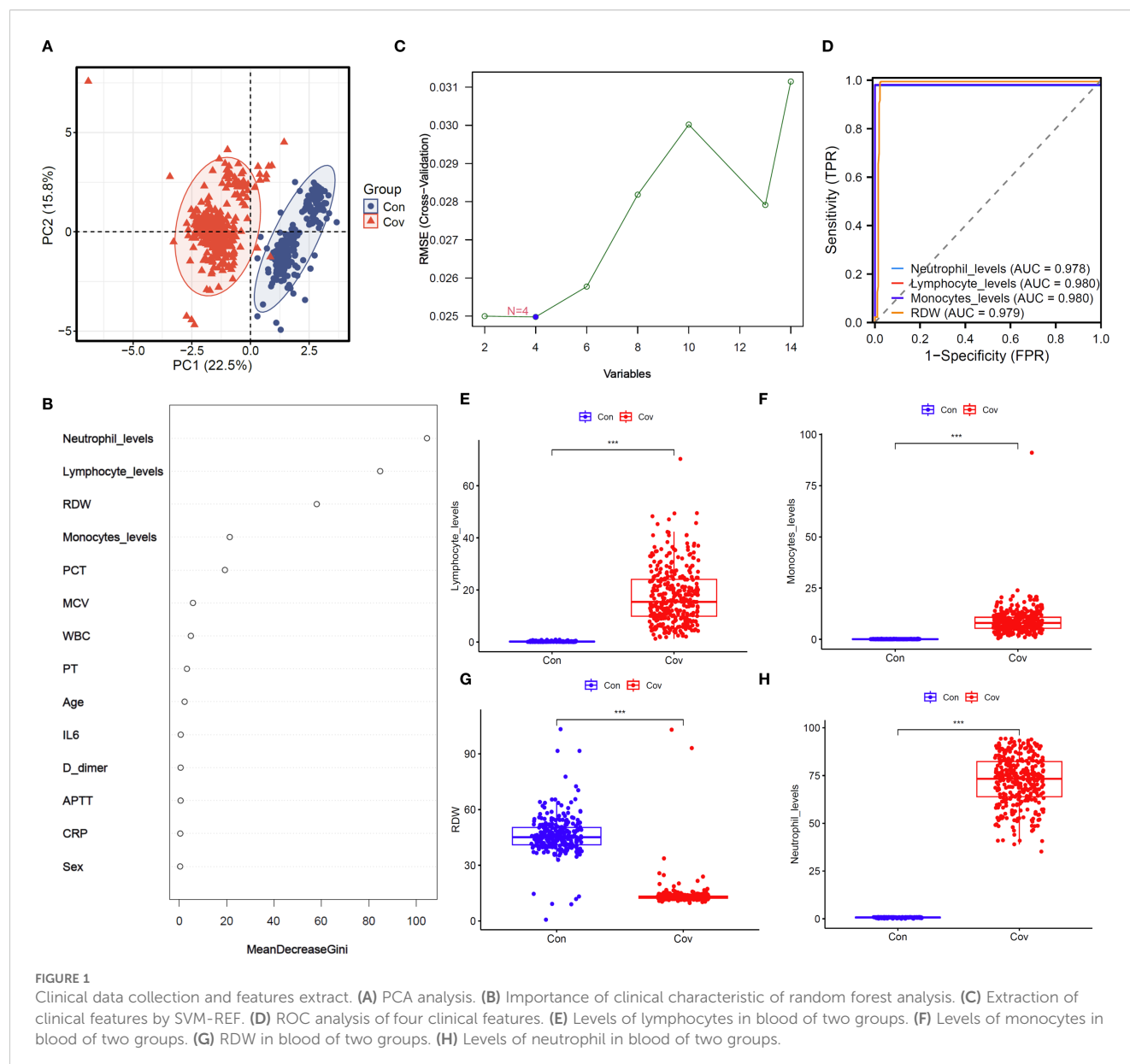
Result

Analysis of clinical sample features

A total of 623 samples were collected and divided into two groups: Control ($n=265$) and COVID-19 ($n=358$). PCA shows that the horizontal and vertical coordinates (PC 1 and PC 2) explain 22.5% and 15.8% of the total variation respectively, and the cumulative contribution rate is 38.3%, which indicates that it can effectively capture the main variation trend of samples data (Figure 1A, Table 1). The random forest results suggest that the top five clinical features of the five importance scores are lymphocyte level, monocyte level, red cell distribution width (RDW), neutrophil level and PCT (Figure 1B). The results of SVM-REF indicated that four clinical features changed obviously when the disease occurred, namely lymphocyte level, monocyte level, RDW and neutrophil level (Figure 1C). Moreover, the model exhibited favorable performance with ROC curve, lymphocyte level (AUC = 0.98), monocyte level (AUC = 0.98), RDW (AUC = 0.979) and neutrophil level (AUC = 0.978) (Figure 1D). To further screen the clinical sample features for machine learning model construction, we analyzed the changing trends of significant clinical sample features among different groups (Figures 1E–H).

scRNA-seq reveals the key immune cell components of COVID-19 patients

We compared the distribution characteristics of single cells between control group and COVID-19 group by ISNE dimensionality reduction technique. Figure 2A shows the distribution pattern of 26 cell subsets in a two-dimensional space composed of ISNE_1 (horizontal axis) and ISNE_2 (vertical axis) in the form of digital clustering, while the spatial location of CD8+ T cells (orange), monocytes (blue-green), dendritic cells (pink) and other major immune cells is clearly marked on the right (Figure 2B). Among them, the COVID-19 group showed remarkable characteristics: the density of CD8+ T cells in ISNE_1 axis 10–20 and ISNE_2 axis -10–0 regions increased obviously, showing antiviral immune activation; The distribution range of monocytes expanded to ISNE_1 axis 15–25 and ISNE_2 axis -5–5, and partially overlapped with T cell region, suggesting cell interaction in inflammatory microenvironment. Dendritic cells distributed discretely between 0–10 of ISNE_1 axis and 5–15 of ISNE_2 axis, which may be related to the disorder of antigen presentation. These differences in spatial distribution directly reveal the migration of immune cells, the reconstruction of subsets and the change of functional status caused by COVID-19 infection, which provides visual evidence for the analysis of virus-specific immune response. By comparing the immune cell composition and gene expression characteristics between the control group and COVID-19 group, the immune remodeling law of COVID-19 infection was revealed. The histogram of Figure 2C shows that the proportion of CD8+ T cells (red stripes) in COVID-19 group decreased significantly, while the proportion of monocytes (lavender stripes) increased relatively,



suggesting that the proportion of immune cells was unbalanced. **Figure 2D** volcanic map further analyzed the functional changes, the cytotoxicity-related molecules (such as GZMB and IFNG) in CD8+ T cells gene expression profile were significantly increased (red dots were dense), indicating that the remaining cells were highly activated; Monocytes are accompanied by the strong expression of pro-inflammatory genes such as IL1 β and TNF (the red dots in the fifth column on the right gather at a high level), which proves that they are polarized to the inflammatory phenotype. The up-regulation of CXCL9/10 chemokine in dendritic cells (red dot in the fourth column) and the fluctuation of NK cell activation gene (sixth column) jointly reveal the abnormal recruitment and regulation network disorder of immune cells in inflammatory microenvironment. It is worth noting that although the proportion of progenitor cells (light blue-green strips) is stable, their stem cell maintenance genes

(such as SOX4) are down-regulated (the last green dot), suggesting that the immune regeneration potential is impaired. Then, we analyzed the cell interaction network formed a dynamic regulatory system through radial connection. Among them, CD8+ T cells (green center) establish multi-directional connections with CD4+ T cells (blue), dendritic cells (uncolored), B cells (red), monocytes (orange) and NK cells (pink) through green lines, highlighting its position as a hub of cytotoxic response (**Figure 2E**); CD4+ T cells (blue center) are connected in series with B cells (red thick junction), monocytes and progenitor cells through dense blue radiation, revealing their dual functions of coordinating humoral immunity and inflammatory regulation (**Figure 2F**). The strong connection between B cells (red center) and CD4+ T cells (red thick line) confirms the core pathway of antibody production in cooperation with T-B (**Figure 2G**), NK cells and T/monocytes form rapid killing defense through pink

TABLE 1 Clinical data analysis.

n	Control (265)	Covid (358)	P value
Age (mean (SD))	60.45 (19.18)	68.77 (14.39)	<0.001
WBC.Count (mean (SD))	7.71 (4.47)	6.98 (3.39)	<0.001
Neutrophil.Proportion (mean (SD))	66.07 (17.38)	70.83 (16.05)	<0.001
Lymphocyte.Proportion (mean (SD))	21.66 (15.42)	17.10 (10.45)	<0.001
Monocyte.Proportion (mean (SD))	8.53 (4.49)	8.54 (6.21)	0.054
RBC.Mean.Volume (mean (SD))	91.30 (8.20)	88.79 (14.70)	0.016
RBC.Distribution.Width (mean (SD))	14.19 (2.58)	12.70 (2.43)	<0.001
PT (mean (SD))	8.38 (6.48)	10.29 (5.07)	<0.001
APTT (mean (SD))	18.36 (13.79)	23.26 (11.43)	<0.001
D.Dimer (mean (SD))	0.99 (1.86)	0.96 (2.30)	<0.001
IL6 (mean (SD))	1.14 (5.63)	11.25 (31.03)	<0.001
PCT (mean (SD))	0.56 (3.97)	0.22 (1.13)	<0.001
C.Reactive.Protein (mean (SD))	42.41 (52.10)	33.68 (53.20)	<0.001
Sex = 2 (%)	119 (44.9)	151 (42.2)	<0.001

connection (Figure 2H), while dendritic cells are simultaneously connected with T/B/NK cells (Figure 2I) to mediate antigen presentation and chemotaxis recruitment. Monocytes (orange center) are widely connected with T/NK/dendritic cells through orange lines, suggesting that they play a bridging role in inflammatory signal amplification (Figure 2J). From the perspective of systems biology, we reveal the precise framework of cell cooperation in immune response.

RNA-seq reveals the key immune components of COVID-19

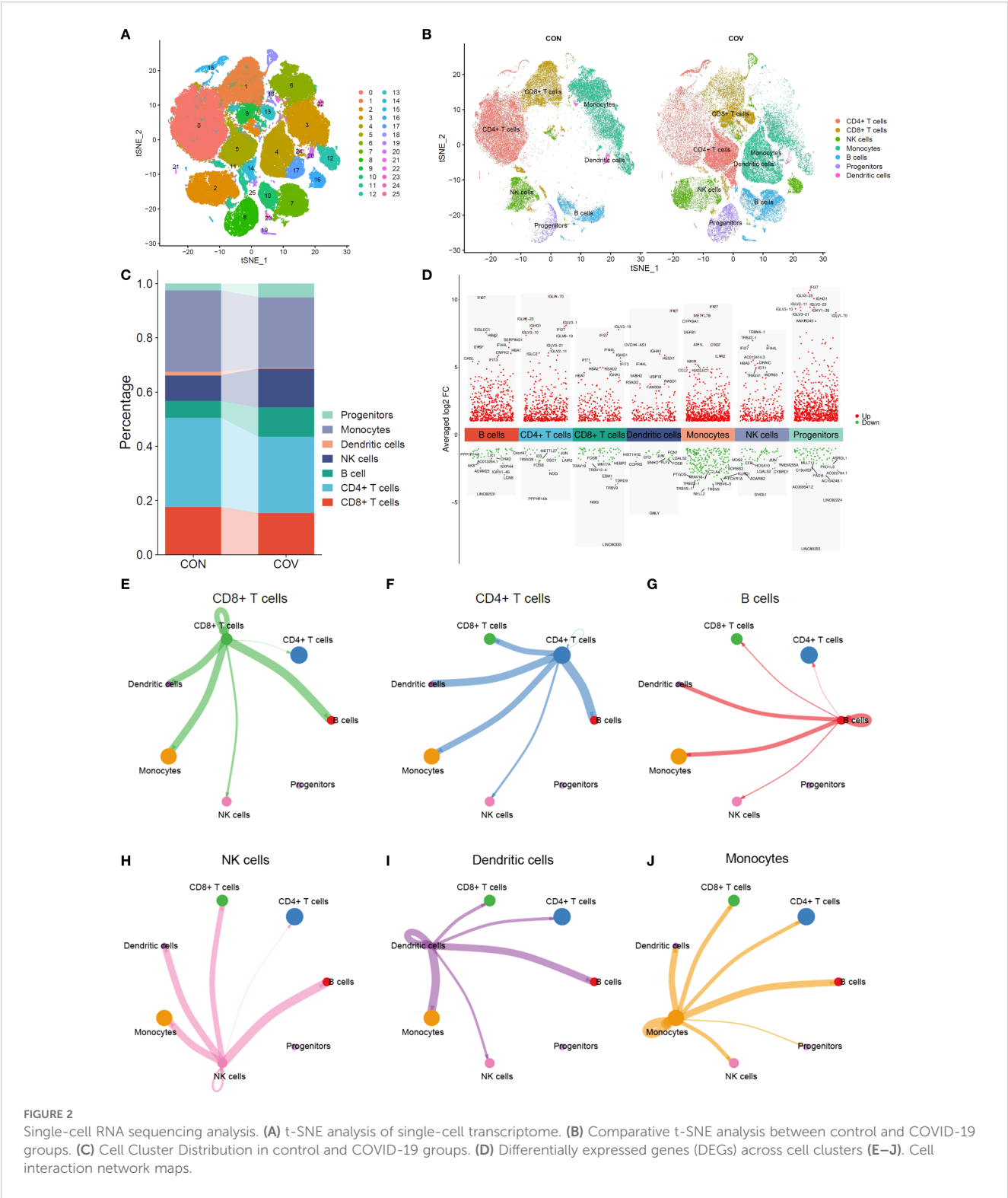
Through principal component analysis, the sequencing data of transcription group showed that the control group (Con, blue dot) and COVID-19 group (Cov, red triangle) were significantly separated along the PC1(32.7% variance) and PC2(19.5% variance) axes, among which the samples in the control group were relatively concentrated, suggesting that the transcription characteristics in the group were highly consistent (Figure 3A). Cluster thermogram shows the expression gradient of genes (rows) in different samples (columns), and blue to red respectively correspond to low expression to high expression level (Figure 3B). It can be seen that “Con” (blue label) and “Cov” (red label) samples form obvious clustering partitions on the gene expression spectrum. Further difference analysis showed that 3024 genes were up-regulated and 3812 genes were down-regulated (Figure 3C). GO analysis showed that the differential genes were significantly enriched in physiological process-related pathways such as MAPK cascade positive regulation, immune response regulation, apoptosis signal pathway and T cell differentiation, and involved ribosome structures, suggesting that cell function and structure were widely

affected during the disease process (Figure 3D). KEGG enrichment results suggest innate immune pathways such as chemokines, Toll-like receptors and T cell receptors, as well as key regulatory hubs such as FoxO, TNF and NF- κ B (Figure 3E).

The results of immune infiltration analysis showed that the correlation network among immune cell subsets was revealed by red and blue gradient (Figure 3F), in which CD8+ T cells were positively correlated with naive B Cell ($r=0.74$) and negatively correlated with neutrophils ($r=-0.79$), suggesting the cooperative or antagonistic relationship between different immune cells in the infiltration process. The difference in the proportion of immune cells between “con” (blue) and “cov” (red) groups was compared by box chart (Figure 3G). Among them, Plasma cells, CD4+ T cells memory resting, M0 macrophage, resting mast cells and CD8+ T cells showed significant distribution deviation between the two groups, and the difference in box span suggested that the treatment might specifically reshape the microenvironment of specific immune subgroups.

WGCNA analysis reveals the modular genes of CD8+ T cells

Figure 4A shows the relationship between soft threshold (power) and goodness of fit (R^2) of scale-free topological model. When power=9 (marked by red five-pointed star), R^2 reaches 0.72 and the curve tends to be flat, indicating that this threshold can effectively balance the biological characteristics and topological attributes of the network. Figure 4B shows the decreasing trend of Mean Connectivity with the increase of power, and the connectivity drops to 671.79 when power=9, suggesting that the network retains the core co-expression structure after filtering weakly related genes.



The thermogram of Figure 4C reveals the correlation pattern between modules. The dark red block in diagonal area (such as brown4 module) reflects the high degree of gene cooperation within modules, while the blue-red difference between firebrick4 and light-colored modules (such as $r=0.32$ to -0.45) implies the independence or antagonistic relationship of different functional modules. Figure 4D quantifies the similarity between samples/gene clusters

by vertical height (vertical axis), and the black branching structure shows that low-level merged clusters (such as height < 5) have high homogeneity, while high-level branches (such as height > 15) suggest significant cross-cluster heterogeneity; At the bottom, “Dynamic Tree Cut” (color strip) marks the independent clusters (such as purple and orange clusters) generated by dynamic cutting, and “Merge Dynamic” strip shows the merged large-scale

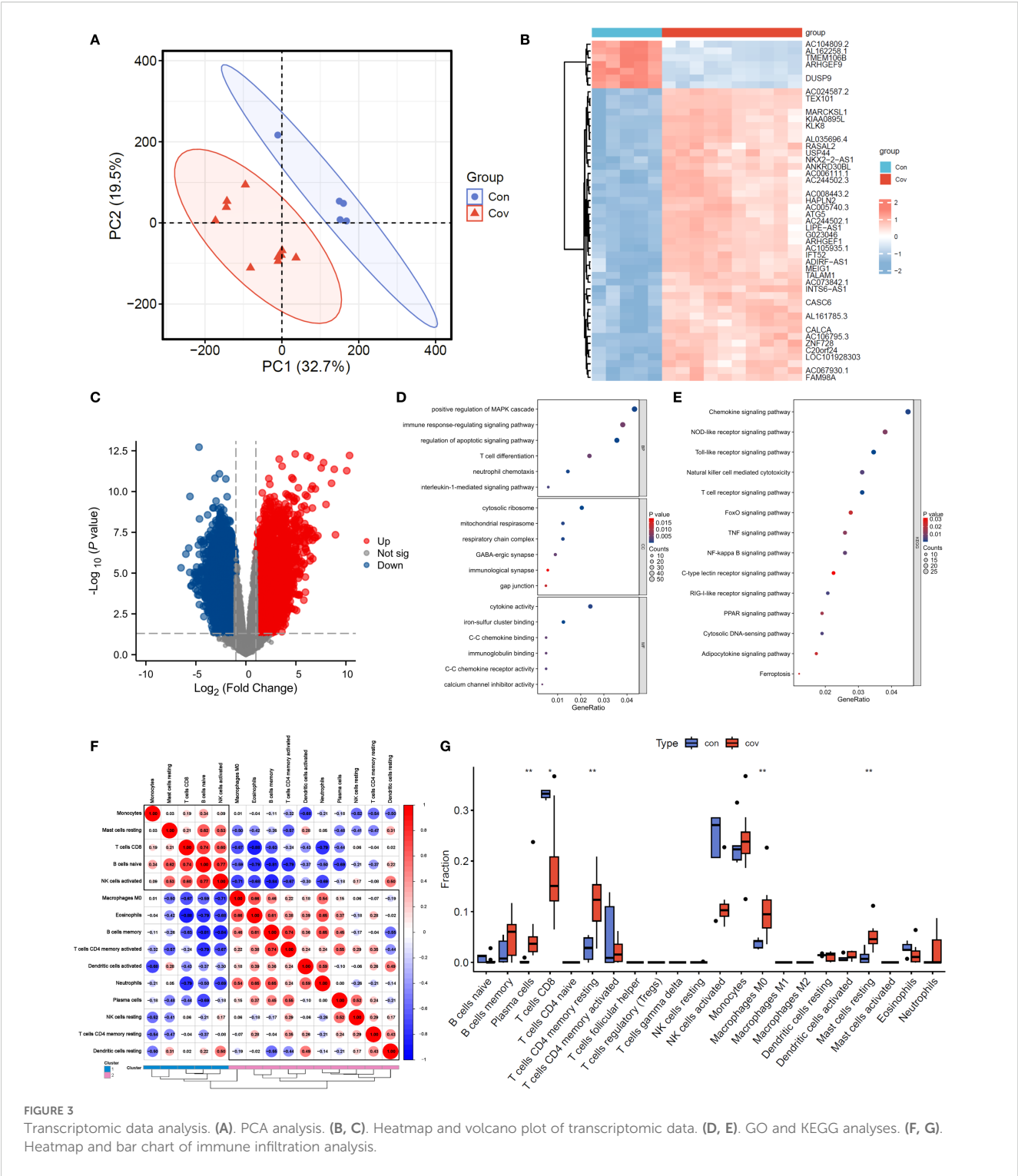


FIGURE 3
Transcriptomic data analysis. (A). PCA analysis. (B, C). Heatmap and volcano plot of transcriptomic data. (D, E). GO and KEGG analyses. (F, G). Heatmap and bar chart of immune infiltration analysis.

distribution (such as blue-green cross-sample continuous area), indicating that the data has multi-level clustering characteristics. Each module gene was correlated with the group (Figure 4E), and the results showed that coral1 module gene was significantly correlated with Cov group ($r=0.91$, $p < 0.05$). Correlation analysis of co-expression network showed that there was a strong positive correlation between the membership degree of coral1 module gene and the significance of weighted genes in Cov group (Figure 4F). By

analyzing the interaction network between gene module and immune cell-related subsets, it was found that coral1 module gene was negatively correlated with CD8+ T cells ($r=-0.81$, $p < 0.05$), which emphasized the negative regulatory effect of Coral1 module gene on CD8+T cells (Figure 4G). Then, we analyzed the correlation analysis between coral1 module gene and CD8+T cells (Figure 4H, $r=0.88$), CD4+T cells resting memory (Figure 4I, $r=0.41$) and Plasma cells (Figure 4J, $r=0.10$).

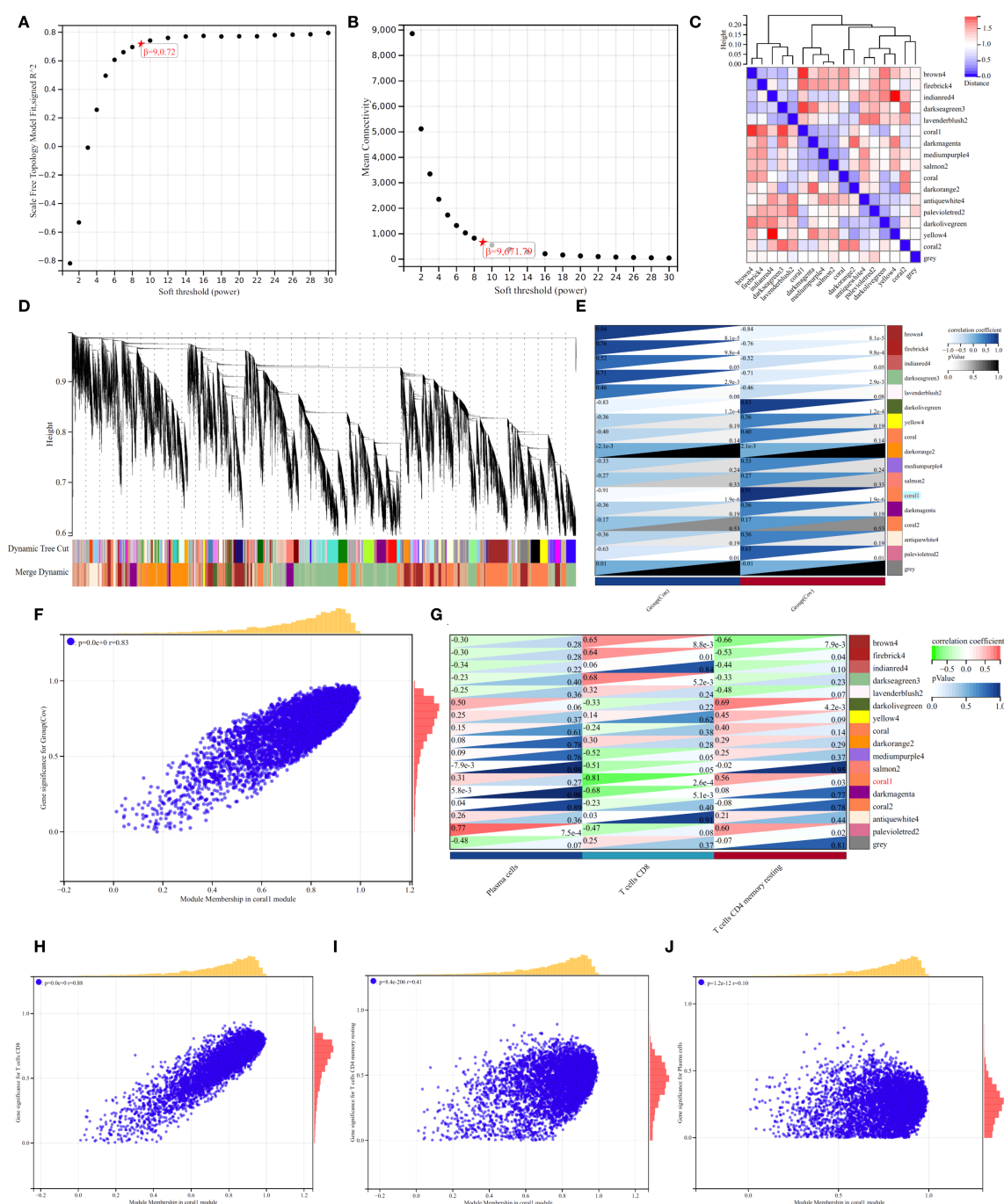


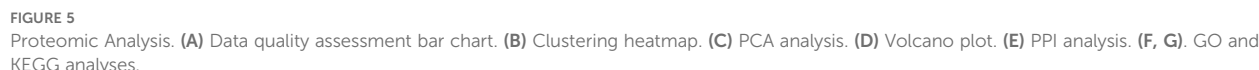
FIGURE 4

WGCNA Analysis. (A, B). Selection of optimal soft thresholds for network construction. (C). Heatmap of module-trait correlations. (D). Hierarchical clustering dendrogram for module detection. (E). Heatmap showing the correlation between modules and immune cells. (F, G, H-J) The correlation analysis between coral1 module gene and CD8+ T cells (H), CD4+ T cells resting memory (I), Plasma cells (J).

Proteomics analysis reveals the key proteins related with CD8+ T cells

The efficiency of proteomics analysis from original map to protein quantification (Figure 5A). In order to identify biomarkers of COVID-19, we conducted proteomic sequencing.

The similarity between samples is shown in Figure 5B, where there is a significant difference between the control and COVID-19. PCA results demonstrate that COVID-19 has a difference compared to Control (Figure 5C). We found a total of 36 DEPs with 11 downregulated proteins and 25 upregulated proteins (Figure 5D). To further identify potential biomarkers, we conducted PPI



Biomarker identification and molecular docking verification

frontiersin.org

Figure 6D), CFL1 (AUC = 1, Figure 6E), PIGR (AUC = 0.98, Figure 6F), SERPINA3 (AUC = 0.96, Figure 6G) had high diagnostic value. ROC analysis of proteomics results showed that BTD (AUC = 1, Figure 6H), CFL1 (AUC = 1, Figure 6I), PIGR (AUC = 0.92, Figure 6J), SERPINA3 (AUC = 0.84, Figure 6K) had high diagnostic value. ANN model was constructed using four candidate biomarkers (CFL1, PIGR,

SERPINA3, and BTD) as input variables (Figure 6L). The network comprised an input layer with four nodes, one hidden layer, and an output layer with two nodes corresponding to the Control and COVID-19 groups. The model converged with a final error of 0.000992 after 360 training steps. The diagnostic performance of the ANN was assessed by receiver operating characteristic (ROC) curve

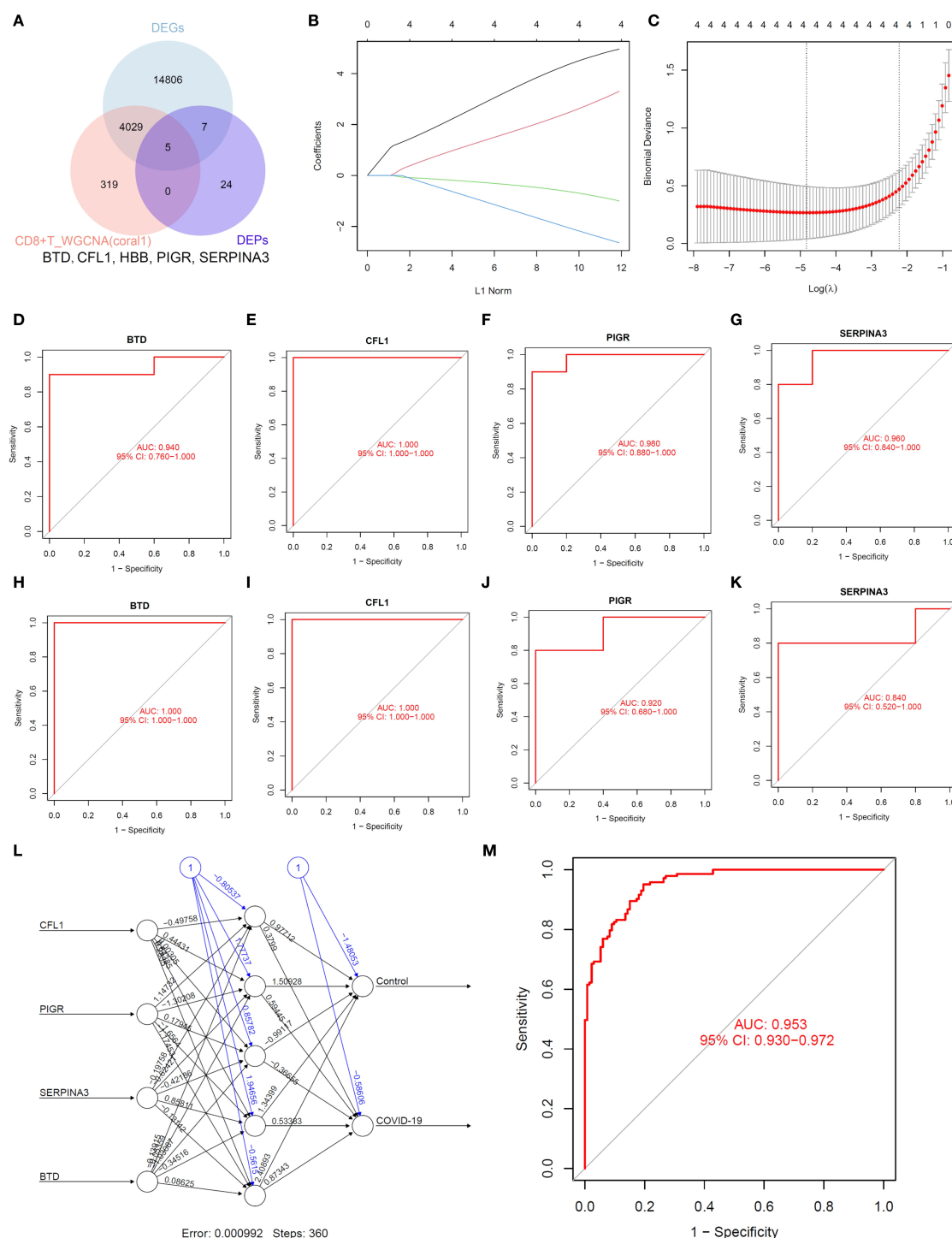


FIGURE 6

Identification of biomarkers. (A). Venn diagram. (B, C). LASSO regression analysis. (D–G). ROC analysis of transcriptomics. (H–K). ROC analysis of proteomics. (L, M). ANN analysis.

analysis. As shown in [Figure 6M](#), the model achieved an area under the curve (AUC = 0.953), indicating excellent discriminative power in distinguishing COVID-19 patients from controls. These results suggest that the identified biomarkers, when integrated into an ANN framework, hold strong potential for accurate disease classification.

To determine whether BTD, CFL1, PIGR and SERPINA3 can serve as biomarkers for COVID-19, we verified with the related biomarkers expression of transcriptomics data ([Figure 7A](#)) and proteomics data ([Figure 7B](#)). In order to determine whether Paxlovid plays a therapeutic role through biomarkers, we conducted molecular docking verification based on AlphaFold. The results indicated that Paxlovid may bind to BTD ([Figure 7C](#)), CFL1 ([Figure 7D](#)), PIGR ([Figure 7E](#)), and SERPINA3 ([Figure 7F](#)) with favorable binding energies ([Table 2](#)), potentially modulating their protein functions and exerting therapeutic effects.

Discussion

The COVID-19 pandemic, caused by the SARS-CoV-2 virus, has profoundly affected the world, with millions of confirmed cases and numerous fatalities ([24, 25](#)). The disease presents a wide spectrum of clinical manifestations, ranging from mild symptoms to severe respiratory failure, posing significant challenges for both diagnosis and treatment ([26](#)). Recent advances in multi-omics molecular profiling have greatly improved our understanding of the transmission dynamics of respiratory viruses on a global scale ([27](#)). Additionally, the substantial overlap in clinical symptoms among various respiratory illnesses continues to hinder accurate diagnosis. To address this challenge, we conducted a comprehensive framework using healthy individuals as a control group, collecting clinical diagnostic data from the peripheral blood of both COVID-19 patients and healthy controls. The omics data (GSE192391, GSE164805) and our clinical cohort come from different sources and time points, and key variables such as age, disease severity and gender can be seen in [Supplementary Table S1](#). By integrating scRNA-seq and RNA-seq data with an analysis of the peripheral plasma proteome, we applied machine learning models to successfully identify and predict potential biomarkers associated with CD8+ T cell responses in COVID-19 infection. Our findings suggest that BTD, CFL1, PIGR, and SERPINA3 may serve as promising auxiliary diagnostic and therapeutic biomarkers for COVID-19, offering significant clinical potential.

Several limitations of this study should be acknowledged. First, the cohorts used in our analysis were not fully matched in terms of demographic and clinical characteristics, which may introduce confounding factors and limit the comparability across groups. Second, while we integrated both publicly available external datasets (GSE192391, GSE164805) and our own internal proteomic data to enhance robustness, heterogeneity in sample collection, processing protocols, and sequencing platforms could have affected the consistency of the results. Third, the sample size of our internal cohort remains relatively modest, which may constrain the statistical power and generalizability of the findings. Finally, functional validation of the identified biomarkers was not

performed in the present work, and further studies in larger, well-matched, and longitudinal cohorts are needed to verify the diagnostic and preventive potential of the proposed multi-omics framework.

The findings are both clinically relevant and biologically plausible. Recent studies suggest critical roles of T cells in the clearance of SARS-CoV-2 and protection from developing severe COVID-19 ([28, 29](#)). In a study, Bilich et al. specifically explored the kinetics of SARS-CoV-2-specific T cell responses in two cohorts of patients up to 6 months after infection ([30](#)). The authors found that, whereas antibody responses wane, T cell responses to SARS-CoV-2 antigens remain consistent or increase over time. T cell responses against SARS-CoV-2 likely provide protection against severe COVID-19, but how reinfection affects T cell functionality remains unclear ([31](#)). The coronavirus typically induces an excessive immune response, with the overconsume and dysfunction of CD8+ T cells being one of the core mechanisms of immunopathological damage ([15](#)). SARS-CoV-2-specific CD8+ T cells in pre-pandemic samples from children, adults, and elderly individuals predominantly displayed a naive phenotype, indicating a lack of previous cross-reactive exposures ([32](#)). A subset of CD8+ T cells regulate chronic inflammation in COVID-19 patients by killing pathogenic CD4+ T cells ([33](#)). Nevertheless, despite extensive research, the exact role of CD8+ T cells in COVID-19 remains to be determined.

Biotinidase (BTD) plays a role in supporting immune function, and since COVID-19 affects immune responses, there may be an indirect relationship ([34](#)). BTD deficiency can impair immune function, and individuals with biotinidase deficiency or biotin deficiency may exhibit altered responses to infections, including viral ones like SARS-CoV-2 ([35](#)). While BTD has not been studied specifically in the context of COVID-19, its role in cellular processes and immunity suggests that biotinidase could potentially support immune function during viral infections, though there is no evidence yet to suggest it has a direct effect on COVID-19. Cofilin 1 (CFL1) is involved in cytoskeletal regulation and plays a significant role in immune cell migration, particularly may influence T cells migrate to sites of infection and become activated in inflammatory response ([36](#)). Therefore, CFL1 could serve as a potential marker for disease severity in COVID-19, with elevated levels or altered actin dynamics possibly correlating with more severe disease or prolonged inflammation. PIGR is essential for mucosal immunity by facilitating the secretion of dimeric IgA and IgM into mucosal surfaces like the respiratory and gastrointestinal tracts ([37, 38](#)). The dysregulation of PIGR activity could impair mucosal immune responses, particularly IgA ability to neutralize SARS-CoV-2, leading to higher viral loads in the upper respiratory tract and potentially more severe disease. SERPINA3 plays an important role in regulating inflammation and protecting tissues from protease-mediated damage ([39](#)). The decrease of SERPINA3 level may be related to poor outcome, but it may also be a protective factor to limit excessive tissue damage. Their potential role in acute COVID-19 and long-term COVID makes them potential biomarkers of disease severity and a candidate for therapeutic intervention.

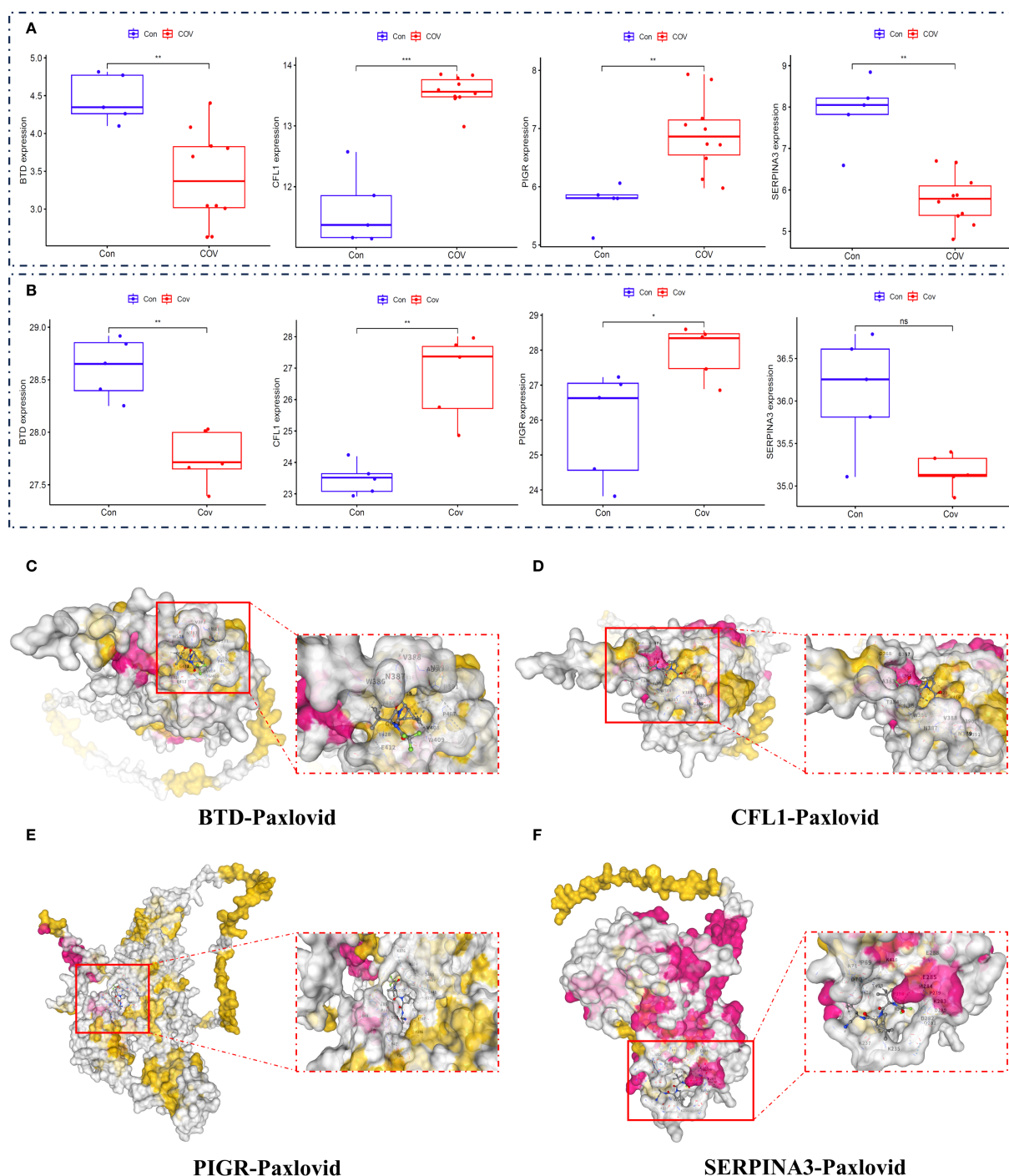


FIGURE 7

Expression level of biomarkers and Molecular Docking Visualization. (A) the gene expression levels of essential genes (BTDL, CFL1, PIGR, SERPINA3), (B) the protein expression levels of essential proteins (BTDL, CFL1, PIGR, SERPINA3). (C) The binding between BTDL and Paxlovid. (D) The binding between CFL1 and Paxlovid. (E) The binding between PIGR and Paxlovid. (F) The binding between SERPINA3 and Paxlovid.

Healthcare has shifted from a responsive model to a proactive, personalized, and preventative approach. The development of multi-omics offers a powerful framework to accurately predict an individual's disease risk and uncover complex biological interactions that may otherwise remain hidden (40). Swinnerton et al. have developed and

prospectively validated a tool to predict the absolute risk of severe COVID-19, incorporating dynamic parameters at both the patient and population levels, which could inform clinical care (41). However, it remains possible that this model may not generalize to individuals who acquired immunity via natural infection or those without exposure to

TABLE 2 Molecular docking based on AlphaFold.

Protein	AlphaFold structure ID	Vina score	Center (x, y, z)
Biotinidase	AF-P43251-F1-v4	-7.3	13, 2, -33
Cofilin-1	AF-P23528-F1-v4	-7.5	13, 2, -33
Polymeric immunoglobulin receptor	AF-P01833-F1-v4	-8.4	-15, 6, -10
Alpha-1-antichymotrypsin	AF-P01011-F1-v4	-7.7	1, 5, -19

the virus or vaccine. The SARS-CoV-2 pandemic spread rapidly worldwide, resulting in high mortality. Developing enhanced vaccination strategies that effectively protect against both disease and viral transmission is crucial for preparing for future respiratory virus pandemics. Leveraging a multi-omics approach in our research allows us to comprehensively assess the multifactorial immune response, providing deeper insights into how blood biomarkers of COVID-19 modulate immunity. Further clinical and translational studies are essential to refine these findings and bridge the gaps in our understanding.

Conclusions

Our findings suggest that BTD, CFL1, PIGR, and SERPINA3 may serve as promising auxiliary diagnostic and therapeutic biomarkers for COVID-19, offering significant clinical potential. By combining machine learning with multi-omics framework, we offer a novel approach to precision medicine, especially in early diagnosis and personalized treatment.

Data availability statement

The scRNA-seq dataset can be obtained from GEO with the primary accession code GSE192391, and RNA-Seq data can be obtained from GEO with the primary accession code GSE164805. The proteomics datasets used in this study are available from the corresponding author on reasonable request.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the Second Affiliated Hospital of Mudanjiang Medical College (Approval No. 202328). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YZ: Conceptualization, Data curation, Writing – review & editing. PF: Writing – original draft, Formal Analysis, Data curation. HZ:

Writing – review & editing, Data curation, Formal Analysis. SH: Data curation, Writing – review & editing, Formal Analysis. MB: Writing – review & editing, Data curation, Formal Analysis. JW: Conceptualization, Writing – review & editing. QW: Writing – review & editing, Conceptualization.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. High level Key Discipline of National Administration of Traditional Chinese Medicine -Traditional Chinese constitutional medicine (No. zyyzdxk-2023251), Major Science and Technology Special Projects in Hubei Province (No. 2023BCA005), the Chief Scientist Research Project of Hubei Shizhen Laboratory (HSL2024SX0002).

Acknowledgments

We thank the participants for their contributions to this research, the GEO for providing extensive data, R for offering their software free of charge, and the researchers who shared the R package.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ying H, Wu X, Jia X, Yang Q, Liu H, Zhao H, et al. Single-cell transcriptome-wide Mendelian randomization and colocalization reveals immune-mediated regulatory mechanisms and drug targets for COVID-19. *EBioMedicine*. (2025) 113:105596. doi: 10.1016/j.ebiom.2025.105596
- Chen S, Huang W, Zhang M, Song Y, Zhao C, Sun H, et al. Dynamic changes and future trend predictions of the global burden of anxiety disorders: analysis of 204 countries and regions from 1990 to 2021 and the impact of the COVID-19 pandemic. *eClinicalMedicine*. (2025) 79:103014. doi: 10.1016/j.eclinm.2024.103014
- Choe H, Farzan M. How SARS-CoV-2 first adapted in humans. *Science*. (2021) 372:466–7. doi: 10.1126/science.abi4711
- Tan L, Wang Q, Zhang D, Ding J, Huang Q, Tang Y-Q, et al. Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduction Targeted Ther*. (2020) 5:33. doi: 10.1038/s41392-020-0148-4
- Li G, Hilgenfeld R, Whitley R, De Clercq E. Therapeutic strategies for COVID-19: progress and lessons learned. *Nat Rev Drug Discovery*. (2023) 22:449–75. doi: 10.1038/s41573-023-00672-y
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
- Zhang X, Han X, Li C, Cui J, Yuan X, Meng J, et al. Clinical outcomes of hospitalized immunocompromised patients with COVID-19 and the impact of hyperinflammation: A retrospective cohort study. *J Inflammation Res*. (2025) 18:3385–97. doi: 10.2147/JIR.S482940
- Athaya T, Ripan RC, Li X, Hu H. Multimodal deep learning approaches for single-cell multi-omics data integration. *Briefings Bioinf*. (2023) 24:bbad313. doi: 10.1093/bib/bbad313
- Tang F, Lao K, Surani MA. Development and applications of single-cell transcriptome analysis. *Nat Methods*. (2011) 8:S6–S11. doi: 10.1038/nmeth.1557
- Amrute JM, Perry AM, Anand G, Cruchaga C, Hock KG, Farnsworth CW, et al. Cell specific peripheral immune responses predict survival in critical COVID-19 patients. *Nat Commun*. (2022) 13:882. doi: 10.1038/s41467-022-28505-3
- Jiang P, Zhang Y, Ru B, Yang Y, Vu T, Paul R, et al. Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. *Nat Methods*. (2021) 18:1181–91. doi: 10.1038/s41592-021-01274-5
- Kwon S, Safer J, Nguyen DT, Hoksza D, May P, Arbesfeld JA, et al. Genomics 2 Proteins portal: a resource and discovery tool for linking genetic screening outputs to protein sequences and structures. *Nat Methods*. (2024) 21:1947–57. doi: 10.1038/s41592-024-02409-0
- Rabaan AA, Al-Ahmed SH, Muhammad J, Khan A, Sule AA, Tirupathi R, et al. Role of inflammatory cytokines in COVID-19 patients: a review on molecular mechanisms, immune functions, immunopathology and immunomodulatory drugs to counter cytokine storm. *Vaccines*. (2021) 9:436. doi: 10.3390/vaccines9050436
- Bo X, Fan C-Y, Wang A-L, Zou Y-L, Yu Y-H, Cong H, et al. Suppressed T cell-mediated immunity in patients with COVID-19: a clinical retrospective study in Wuhan, China. *J Infection*. (2020) 81:e51–60. doi: 10.1016/j.jinf.2020.04.012
- Karl V, Hofmann M, Thimme R. Role of antiviral CD8+ T cell immunity to SARS-CoV-2 infection and vaccination. *J Virol*. (2025) 99:e01350–24. doi: 10.1128/jvi.01350-24
- Gozzi-Silva SC, Oliveira LDM, Alberca RW, Pereira NZ, Yoshikawa FS, Pietrobon AJ, et al. Generation of cytotoxic T cells and dysfunctional CD8 T cells in severe COVID-19 patients. *Cells*. (2022) 11:3359. doi: 10.3390/cells11213359
- Rha M-S, Shin E-C. Activation or exhaustion of CD8+ T cells in patients with COVID-19. *Cell Mol Immunol*. (2021) 18:2325–33. doi: 10.1038/s41423-021-00750-4
- Wang H, Wang Z, Cao W, Wu Q, Yuan Y, Zhang X. Regulatory T cells in COVID-19. *Aging disease*. (2021) 12:1545. doi: 10.14336/AD.2021.0709
- Weiskopf D, Schmitz KS, Raadsen MP, Grifoni A, Okba NMA, Endeman H, et al. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci Immunol*. (2020) 5:eabd2071. doi: 10.1126/sciimmunol.abd2071
- Zhang Q, Meng Y, Wang K, Zhang X, Chen W, Sheng J, et al. Inflammation and antiviral immune response associated with severe progression of COVID-19. *Front Immunol*. (2021) 12:631226. doi: 10.3389/fimmu.2021.631226

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2025.1671936/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Cohort Alignment and Characterization.

- Najjar-Debbiny R, Gronich N, Weber G, Khoury J, Amar M, Stein N, et al. Effectiveness of paxlovid in reducing severe coronavirus disease 2019 and mortality in high-risk patients. *Clin Infect Diseases*. (2022) 76:e342–e9. doi: 10.1093/cid/ciac443
- Liu Y, Yang X, Gan J, Chen S, Xiao Z-X, Cao Y. CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res*. (2022) 50:W159–W64. doi: 10.1093/nar/gkac394
- Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. (2010) 31:455–61. doi: 10.1002/jcc.21334
- Santos S, Humbard M, Lambrou AS, Lin G, Padilla Y, Chaitram J, et al. The SARS-CoV-2 test scale-up in the USA: an analysis of the number of tests produced and used over time and their modelled impact on the COVID-19 pandemic. *Lancet Public Health*. (2025) 10:e47–57. doi: 10.1016/S2468-2667(24)00279-2
- Parkins MD, Lee BE, Acosta N, Bautista M, Hubert CR, Hrudey SE, et al. Wastewater-based surveillance as a tool for public health action: SARS-CoV-2 and beyond. *Clin Microbiol Rev*. (2024) 37:e00103–22. doi: 10.1128/cmr.00103-22
- Steiner S, Kratzel A, Barut GT, Lang RM, Aguiar Moreira E, Thomann L, et al. SARS-CoV-2 biology and host interactions. *Nat Rev Microbiol*. (2024) 22:206–25. doi: 10.1038/s41579-023-01003-z
- Gao J, Zhang C, Wheelock AM, Xin S, Cai H, Xu L, et al. Immunomics in one health: understanding the human, animal, and environmental aspects of COVID-19. *Front Immunol*. (2024) 15:1450380. doi: 10.3389/fimmu.2024.1450380
- Wörzner K, Schmidt ST, Zimmermann J, Tami A, Polacek C, Fernandez-Antunez C, et al. Intranasal recombinant protein subunit vaccine targeting TLR3 induces respiratory tract IgA and CD8 T cell responses and protects against respiratory virus infection. *EBioMedicine*. (2025) 113:105615. doi: 10.1016/j.ebiom.2025.105615
- Nguema L, Picard F, El Hajj M, Dupaty L, Fenwick C, Cardinaud S, et al. Subunit protein CD40.SARS-CoV2 vaccine induces SARS-CoV-2-specific stem cell-like memory CD8+ T cells. *EBioMedicine*. (2025) 111:105479. doi: 10.1016/j.ebiom.2024.105479
- Bilich T, Nelde A, Heitmann JS, Maringer Y, Roerden M, Bauer J, et al. T cell and antibody kinetics delineate SARS-CoV-2 peptides mediating long-term immune responses in COVID-19 convalescent individuals. *Sci Transl Med*. (2021) 13:eabf7517. doi: 10.1126/scitranslmed.abf7517
- Cai C, Gao Y, Adamo S, Rivera-Ballesteros O, Hansson L, Österborg A, et al. SARS-CoV-2 vaccination enhances the effector qualities of spike-specific T cells induced by COVID-19. *Sci Immunol*. (2023) 8:eadh0687. doi: 10.1126/sciimmunol.adh0687
- Nguyen THO, Rowntree LC, Petersen J, Chua BY, Hensen L, Kedzierski L, et al. CD8+ T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope display high naive precursor frequency and TCR promiscuity. *Immunity*. (2021) 54:1066–82.e5. doi: 10.1016/j.immuni.2021.04.009
- Li J, Zaslavsky M, Su Y, Guo J, Sikora MJ, van Unen V, et al. KIR⁺CD8⁺ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19. *Science*. (2022) 376:eabi9591. doi: 10.1126/science.abi9591
- Kovarik JJ, Bileck A, Hagn G, Meier-Menches SM, Frey T, Kaempf A, et al. A multi-omics based anti-inflammatory immune signature characterizes long COVID-19 syndrome. *iScience*. (2023) 26:105717. doi: 10.1016/j.isci.2022.105717
- Kahraman AB, Yıldız Y, Çıka K, Erdal I, Akar HT, Dursun A, et al. COVID-19 in inherited metabolic disorders: Clinical features and risk factors for disease severity. *Mol Genet Metab*. (2023) 139:107607. doi: 10.1016/j.ymgme.2023.107607
- Xing J, Wang Y, Peng A, Li J, Niu X, Zhang K. The role of actin cytoskeleton CFL1 and ADF/cofilin superfamily in inflammatory response. *Front Mol Biosci*. (2024) 11:1408287. doi: 10.3389/fmolb.2024.1408287
- Zhou X, Wu Y, Zhu Z, Lu C, Zhang C, Zeng L, et al. Mucosal immune response in biology, disease prevention and treatment. *Signal Transduction Targeted Ther*. (2025) 10:7. doi: 10.1038/s41392-024-02043-4
- Laprise F, Arduini A, Duguay M, Pan Q, Liang C. SARS-CoV-2 accessory protein ORF8 targets the dimeric IgA receptor pIgR. *Viruses*. (2024) 16:1008. doi: 10.3390/v16071008
- Yang W, Huang M, Jiang L. Advancements in serine protease inhibitors: from mechanistic insights to clinical applications. *Catalysts*. (2024) 14:787. doi: 10.3390/catal14110787

40. Mohr AE, Ortega-Santos CP, Whisner CM, Klein-Seetharaman J, Jasbi P. Navigating challenges and opportunities in multi-omics integration for personalized healthcare. *Biomedicines*. (2024) 12:1496. doi: 10.3390/biomedicines12071496
41. Swinnerton K, Fillmore NR, Vo A, La J, Elbers D, Brophy M, et al. Leveraging near-real-time patient and population data to incorporate fluctuating risk of severe COVID-19: development and prospective validation of a personalised risk prediction tool. *eClinicalMedicine*. (2025) 81:103114. doi: 10.1016/j.eclinm.2025.103114