RESEARCH



Longitudinal multi-omics analysis of convalescent individuals with respiratory sequelae 6–36 months after COVID-19



Huqin Yang^{1†}, Lujia Guan^{1†}, Yi Xue¹, Xuyan Li¹, Leyi Gao¹, Zhijin Zhang¹, Haifan Zhang¹, Haomiao Ma¹, Fengjiao Liu², Xuan Huang^{1,2}, Zhaohui Tong^{1,3*} and Jieqiong Li^{1,2*}

Abstract

Background Approximately 10–30% of individuals continue to experience symptoms classified as post-acute sequelae of coronavirus disease 2019 (COVID-19 (PASC)). PASC is a multisystem condition primarily characterized by respiratory symptoms, such as reduced diffusing capacity for carbon monoxide (DLco). Although many studies have investigated the pathogenesis of acute COVID-19, the long-term molecular changes in COVID-19 convalescents with PASC remain poorly understood.

Methods We prospectively recruited 70 individuals who had been diagnosed with COVID-19 from 7 January 2020 to 29 May 2020 (i.e., COVID-19 convalescents); we performed follow-up visits at 6 months, 1 year, 2 years, and 3 years after hospital discharge. Thirty-five healthy controls (CONs), recruited from a physical examination center before the COVID-19 pandemic, served as a comparison group. We explored the proteomic and metabolomic profiles of 174 plasma samples from the 70 COVID-19 convalescents and 35 CONs.

Results We performed a comprehensive molecular analysis of COVID-19 convalescents to investigate host changes up to 3 years after hospital discharge. Our multi-omics analysis revealed activation of cytoskeletal organization and glycolysis/gluconeogenesis, as well as suppression of gas transport and adaptive immune responses, in COVID-19 convalescents. Additionally, metabolites involved in glutathione metabolism; alanine, aspartate, and glutamate metabolism; and ascorbate and aldarate metabolism were significantly upregulated in COVID-19 convalescents. Pulmonary and molecular abnormalities persisted for 3 years in COVID-19 convalescents; impaired diffusing capacity for carbon monoxide (DLco) was the most prominent feature. We used this multi-omics profile to develop a model involving one protein (heterogeneous nuclear ribonucleoprotein K (HNRNPK)) and two metabolites (arachidonoyl-EA and 1-O-(2r-hydroxy-pentadecyl)-sn-glycerol)) for identification of COVID-19 convalescents with abnormal DLco.

Conclusions These data provide insights concerning molecular sequelae among COVID-19 convalescents up to 3 years after hospital discharge, clarify mechanisms driving respiratory sequelae, and support the development of a novel model to predict reduced DLco. This longitudinal multi-omics analysis may illuminate the trajectory of altered lung function in COVID-19 convalescents.

Keywords Long COVID, Convalescence, Respiratory sequelae, Proteomic, Metabolic

[†]Huqin Yang and Lujia Guan contributed equally to this work.

*Correspondence: Zhaohui Tong tongzhaohuicy@sina.com Jieqiong Li jieqiongli2010@163.com Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Background

Advances in understanding the pathogenesis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the implementation of effective treatments have substantially reduced mortality and hospitalization rates in recent years [1]. Although most coronavirus disease 2019 (COVID-19) patients fully recover, approximately 10–30% continue to experience symptoms classified as post-acute sequelae of COVID-19 (PASC) [2, 3]. Despite extensive research concerning the pathogenesis of acute COVID-19, the long-term changes in COVID-19 convalescents and the mechanisms underlying PASC remain poorly understood.

PASC is a multisystem condition, in which respiratory symptoms constitute the most important cluster [4]. Convalescent individuals experience various degrees of pulmonary function test (PFT) impairment; reduced diffusing capacity for carbon monoxide (DLco) is the main feature [5]. Our previous investigation, a follow-up study of hospitalized COVID-19 patients, revealed that more than 40% of the participants exhibited reduced DLco at 2 and 3 years after COVID-19 discharge [6]. Multiple factors may contribute to these chronic respiratory sequelae, including lung tissue damage and pulmonary vascular abnormalities. Data remain scarce concerning the underlying etiology of persistent respiratory morbidity and impaired lung function after COVID-19.

Although PFTs are valuable tools for patients who continue to experience dyspnea after COVID-19, some clinical difficulties are evident. The equipment for DLco detection is not widely available; moreover, many patients (e.g., older patients and those with hearing impairment) cannot easily cooperate with clinical staff during detection. Thus, simple and feasible approaches are needed for the early assessment of gas transfer abnormalities and restrictive impairment. In previous studies, multi-omics investigations were conducted to elucidate the pathological landscape of PASC [7-10]. These quantitative omics technologies can facilitate the detection of novel PASC biomarkers and the development of appropriate treatment strategies. Concerning respiratory sequelae after COVID-19, Xu et al. analyzed the plasma metabolomic profile of COVID-19 convalescents 3 months after hospital discharge; their work revealed various metabolic pathway alterations among individuals with abnormal PFT results [11]. Gu et al. performed a proteomic analvsis of COVID-19 convalescents at 6 months, 1 year, and 2 years, identifying several proteins associated with reduced DLco [12]. However, it is unclear whether these abnormalities (especially for molecules involved in respiratory sequelae) persist 3 years after the initial SARS-CoV-2 infection.

This study aimed to systematically elucidate the longitudinal features of COVID-19 convalescents using multiomics analysis of plasma samples collected at 6 months, 1 year, 2 years, and 3 years after hospital discharge. Participants were divided into normal DLco (nDLco) and abnormal DLco (aDLco) groups to investigate the mechanisms contributing to aDLco. Finally, a machine learning model was developed to identify pulmonary sequelae in COVID-19 convalescents at 6 months after hospital discharge. This longitudinal multi-omics analysis may illuminate the trajectory of altered lung function in COVID-19 convalescents.

Methods

Patient enrollment

Study participants were recruited from a cohort of COVID-19 convalescents who had been discharged from a single center during the period from 7 January to 29 May 2020 [6]. According to the study protocol, patients attended the follow-up visits at the outpatient clinic at 6 months, 1 year, 2 years, and 3 years after hospital discharge. These visits included symptom assessments, laboratory tests, the 6-min walking test (6MWT), PFTs, and high-resolution chest computed tomography (HRCT) scans. Exclusion criteria were chronic lung disease, chronic infectious disease, autoimmune disease, and inability to complete the initial follow-up (e.g., due to post-discharge mortality, underlying medical conditions, or psychotic disorders). COVID-19 severity during hospitalization was stratified using a modified six-point ordinal scale [13]. PFTs and HRCT were conducted in accordance with our previously published standard protocols [6].

Age- and sex-matched healthy controls (CONs) were recruited from the physical examination center of Beijing Chao-Yang Hospital before the COVID-19 pandemic. Plasma samples were collected once from each CON and stored at -80° C.

The study was approved by the Research Ethics Commission of our hospital (approval no. KY-2021–15). Written informed consent to take part in the study was obtained from all participants during their clinic visits. To develop a biomarker panel for identifying pulmonary sequelae in COVID-19 convalescents, participants were randomly divided into a training cohort and a test cohort (using a 7:3 ratio) to confirm the diagnostic value of the panel.

Clinical data collection

Data regarding demographics, comorbidities, clinical presentation, supplemental oxygen during hospitalization, laboratory tests, and the results of PFTs and HRCT were extracted from patient medical records. PFTs were performed as previously described [6].

PFTs consisted of spirometry and DLco. Spirometric parameters included residual volume (RV), forced vital capacity (FVC), forced expiratory volume in 1 s (FEV1), FEV1/FVC ratio, total lung capacity (TLC), and maximum mid-expiratory flow (MMEF) 75/25. All PFT parameters were expressed as percentages of predicted normal values (%). Diffusion impairment was defined as DLco < 80% of the predicted value.

CT features were evaluated by an experienced radiologist and an experienced pulmonologist. CT features indicative of interstitial lung abnormalities (ILAs) included ground-glass opacities (GGOs), reticulation or consolidation, lung distortion, traction bronchiectasis and/or bronchiectasis, honeycombing, and non-emphysematous lung cysts, as defined by the Fleischner Society [14]. Semiquantitative CT scores were determined as follows, according to lung involvement in each of the five lobes: 0 (no involvement), 1 (less than 5% involvement), 2 (5–25% involvement), 3 (26–49% involvement), 4 (50–75% involvement), and 5 (more than 75% involvement). The total severity score was calculated by summing the individual five lobar scores, with possible scores ranging from 0 to 25 [15].

Sample collection

Blood samples were collected through venipuncture, in which 4 mL of whole blood were drawn into ethylenediaminetetraacetic acid (EDTA)-coated tubes (BD Vacutainer[®] EDTA K2). Samples were centrifuged at $700 \times g$ for 20 min at 4°C to separate plasma, which was then aliquoted into sterile tubes and stored at -80° C to maintain sample integrity and minimize degradation prior to analysis.

Proteomic analysis

Proteomic analysis of plasma samples was performed in accordance with previously described methods [16]. Plasma samples were prepared by adding 50 nM tetraethyl ammonium bromide to 200 µL of the sample, followed by trypsin (in 50 mM ammonium bicarbonate) at an enzyme-to-protein ratio of 1:50. Samples were digested at 37°C for 16 h. The resulting peptides were desalted, dried, and resolubilized in 0.1% formic acid solution. All data were acquired using a Thermo Scientific[™] Vanquish[™] Neo Ultra-High-Performance Liquid Chromatography system coupled to an Orbitrap Astral mass spectrometer with a Thermo Scientific[™] Easy-spray[™] source. Data were acquired in data-independent acquisition mode with a normalized collision energy of 25% and a default charge state of 2. A 110cm Thermo Scientific[™] µPac[™] Neo High-Performance Liquid Chromatography column was used for quantitative experiments with a 7-min gradient of 4% to 99% B and a flow rate of 2.5 μ L/min. For Orbitrap Astral mass spectrometry experiments, MS1 spectra were acquired at a resolving power of 240,000 every 0.6 s; tandem mass spectrometry spectra were acquired in the astral analyzer with injection times varying among experiments. For comparison searches, raw data were analyzed using DIA-NN v1.8.1 in library-free mode and searched against the UniProt Human database (UP000005640_9606_one_ gene_one_protein_2023_09_28.fasta).

Metabolomic analysis

Metabolomic analysis of plasma samples was conducted as previously described [17]. Samples stored at - 80°C were thawed at room temperature. For each plasma sample, 100 µL were added to a 1.5-mL Eppendorf tube containing 400 µL of an ice-cold mixture of methanol and acetonitrile (1/1, vol/vol); all mixtures were vortexed for 1 min, and whole samples were extracted by ultrasonication for 10 min in an ice-water bath, then stored at -20° C for 1 h. Extracts were centrifuged at 4°C (13,000 rpm) for 10 min. For each extract, 100 µL of supernatant in a glass vial was dried using a freeze concentration centrifugal dryer. Quality control samples were injected at regular intervals (every 20 samples) throughout the analytical run to obtain data for repeatability assessment. An ACQUITY Ultra-Performance Liquid Chromatography I-Class system (Waters Corporation, Milford, MA, USA) coupled with a Q-Exactive plus quadrupole-Orbitrap mass spectrometer, equipped with a heated electrospray ionization (ESI) source (Thermo Fisher Scientific, Waltham, MA, USA), was used to analyze metabolic profiles in ESI-positive and ESI-negative ion modes. Raw liquid chromatography-mass spectrometry data were processed using Progenesis QI v2.3 software (Nonlinear Dynamics, Newcastle, UK) for baseline filtering, peak identification, integration, retention time correction, peak alignment, and normalization. The main parameters were 5 ppm precursor tolerance, 10 ppm product tolerance, and 5% product ion threshold. Compound identification based on precise mass-to-charge ratio (m/z), secondary fragments, and isotopic distribution was performed with the Human Metabolome Database (HMDB), Lipid Maps (v2.3), and Metlin databases for qualitative analysis.

Statistical analysis

Categorical data were presented as counts (percentages) and analyzed using the chi-square test. Normally distributed continuous variables were presented as mean±standard deviation and analyzed using *t*-tests or one-way analysis. Non-normally distributed continuous variables were presented as median (interquartile range) and analyzed using the Wilcoxon rank sum test or the Kruskal–Wallis test. The threshold for statistical significance was set at P-value < 0.05.

Fold changes (FCs) for group comparisons were calculated using mean values. Wilcoxon rank sum tests were used to identify statistically significant differences between groups. Differentially expressed proteins (DEPs) and differentially expressed metabolites (DEMs) were selected based on the following thresholds: FC \geq 1.5 or < 0.67 and *P*-value < 0.05. All *P*-values were subjected to false discovery rate (FDR) correction using the Benjamini and Hochberg method to maximize reliability.

Partial least squares discriminant analysis (PLS-DA) was performed using MetaboAnalyst 5.0 (https://www. metaboanalyst.ca/). Volcano plots were used for data visualization. The Mfuzz package (v2.46.0) in R software was used to identify expression patterns in proteomic and metabolic profiles. To explore biological functions and pathways, Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg/) and Gene Ontology (GO, http://geneontology.org/) enrichment analyses were performed. Connected networks and interactions between proteins and metabolites were visualized using the ggplot2 package in R software and Cytoscape v3.5.1.

Machine learning and predictive model

Machine learning classifiers were constructed using logistic regression, random forest, linear support vector machine, *k*-nearest neighbor, and neural network algorithms. Receiver operating characteristic (ROC) analysis was performed to assess the discriminative performances of the models, quantified by the area under the ROC curve (AUC) values.

Results

Study design

We prospectively recruited 99 patients who had been diagnosed with COVID-19 from 7 January 2020 to 29 May 2020 and subsequently discharged from the hospital. Of these patients, 29 were excluded; 70 COVID-19 convalescents were enrolled at the 6-month follow-up. These participants were divided into two groups based on diffusing capacity: aDLco (n=27) and nDLco (n=43) (Fig. 1A and Additional file 1: Fig. S1). At the 3-year follow-up, 23 COVID-19 convalescents remained in the study (10 in the aDLco group and 13 in the nDLco group). In total, 174 plasma samples were collected from COVID-19 convalescents across four follow-up time points. Additionally, 35 CONs were recruited from a physical examination center before the COVID-19

pandemic (Fig. 1A); plasma samples were collected from these CONs.

A comprehensive molecular analysis, incorporating proteomic and metabolomic methods, was performed on the plasma samples to evaluate host changes up to 3 years after hospital discharge (Fig. 1B). Subsequent analyses focused on patients with aDLco to identify most prominent signatures (Fig. 1C). Based on these data, we used machine learning algorithms to develop a novel biomarker panel for identifying pulmonary sequelae in COVID-19 convalescents. This plasma biomarker panel was then validated in a test cohort (Fig. 1D).

Participant demographic and clinical characteristics

Patient demographic and clinical characteristics are presented in Additional file 2: Table S1. The mean ages of COVID-19 convalescents at 6 months after hospital discharge and CONs were 56.3 (12.4) and 52.9 (14.5) years, respectively. Age and sex were balanced between the two groups. Among the 70 COVID-19 convalescents, severity classifications were 23 (32.9%), score 2; 18 (25.7%), score 3; 28 (40%), score 4; and one (1.4%), score 5. At 6 months after hospital discharge, 10 (14.3%) participants experienced cough and sputum, whereas 12 (17.1%) had shortness of breath. In terms of PFT parameters, six (8.6%) patients had obstructive ventilatory impairment (FEV1/ FVC <70%), 14 (20.0%) patients had restrictive ventilatory impairment (TLC <80% predicted), and 27 (38.6%) patients had reduced DLco (<80%).

Given the substantial diffusion impairment, we divided the COVID-19 convalescents into nDLco and aDLco groups at 6 months, and then explored potential distinguishing characteristics (Additional file 2: Table S2). Pulmonary function parameters including RV, TLC, MMEF 75/25, FEV1, FEV1/FVC, and 6MWT all were significantly decreased in the aDLco group (Fig. 2A). Concerning radiological abnormalities, the aDLco group exhibited higher scores for total abnormalities, total reticulation or consolidation, traction bronchiectasis and/or bronchiectasis, and total GGO (Fig. 2B). DLco values were inversely correlated with the extent of total abnormalities, reticulation or consolidation, GGO, and traction bronchiectasis and/or bronchiectasis (Fig. 2C and D). Overall, our detailed analysis of lung imaging and functional parameters revealed that reduced DLco was the most prominent feature after SARS-CoV-2 infection.

Proteomic signature dynamics over 3 years after SARS-CoV-2 infection

Next, we examined the proteomic characteristics of hospitalized COVID-19 convalescents over 3 years of postdischarge follow-up, focusing on dynamic changes in host dysfunction. PLS-DA (Additional file 1: Fig. S2A)



Fig. 1 Overview of the study design. **A** Study design and participant timeline: This longitudinal cohort consisted of COVID-19 convalescents followed up at 6 months (n=70), 1 year (n=23), 2 years (n=23), and 3 years (n=23), along with a healthy control group (n=35). COVID-19 convalescents were divided into nDLco and aDLco groups based on their diffusing capacity at the 6-month follow-up. Clinical evaluations included pulmonary function tests, laboratory analyses, computed tomography, and face-to-face interviews. Plasma samples were collected for proteomic and metabolomic analyses. **B** Dynamic proteomic (top) and metabolomic (down) signatures of patients over 3 years of follow-up. **C** Proteomic and metabolomic profiles of patients with aDLco at 6 months after hospital discharge. **D** Identification of COVID-19 convalescents with respiratory sequelae at 6 months after hospital discharge



Fig. 2 Correlations between pulmonary function parameters and CT imaging features in COVID-19 convalescents at the 6-month follow-up. **A** Comparison of pulmonary function parameters between nDLco and aDLco groups at the 6-month follow-up. Parameters included residual volume (RV), functional residual capacity (FRC), total lung capacity (TLC), maximum mid-expiratory flow (MMEF) 75/25, forced expiratory volume in 1 s (FEV1), forced vital capacity (FRC), total lung capacity (TLC), maximum mid-expiratory flow (MMEF) 75/25, forced expiratory volume in 1 s (FEV1), forced vital capacity (FVC), FEV1/FVC, and 6-min walking test (6MWT). *P*-values were calculated using the Wilcoxon rank sum test, and significant *P*-values are indicated on the boxplots. **P* < 0.05, ***P* < 0.01. **B** Boxplots showing total radiological scores from CT imaging in nDLco and aDLco groups. Scores include total score, reticulation score, traction bronchiectasis and/or bronchiolectasis, ground-glass opacity (GGO) score, non-emphysematous lung cyst score, and honeycombing score. *P*-values were calculated using the Wilcoxon rank sum test, and significant *P*-values are indicated on the boxplots. **P* < 0.05, ***P* < 0.01. **C** Correlation matrix illustrating relationships between pulmonary function parameters and CT imaging scores. Red and blue colors represent positive and negative correlations, respectively. **P* < 0.05. ***P* < 0.01. **D** Analysis of correlations between DLco and each of the following: total score, reticulation score, traction bronchiectasis and/or bronchiolectasis, and GGO score

and volcano plots (Additional file 1: Fig. S2B–E) illustrated clear separation between groups. In total, 319 proteins were differentially expressed in samples from COVID-19 convalescents (6 months, 1 year, 2 years, and 3 years) compared with CONs (Fig. 3A and Additional file 2: Table S3.1).

To systematically evaluate the dynamic changes in DEPs among COVID-19 convalescents over the 3-year follow-up period, we performed clustering analysis on the 319 proteins. Eight expression patterns were identified, including five increasing clusters (1, 2, 4, 5, and 7), two decreasing clusters (3 and 6), and one inverted "V" cluster (8) across convalescent individuals with different timepoints (Fig. 3B). GO biological process analysis indicated that DEPs in the increasing clusters were mainly enriched for supramolecular fiber organization, cytoskeletal organization, actin cytoskeletal organization, actin filament-based processes, and hemostasis (Fig. 3C). Notably, KEGG analysis of the increasing clusters revealed that the glycolysis, carbon metabolism, and pentose phosphate pathways were upregulated in COVID-19 convalescents (Additional file 1: Fig. S3). Specifically, levels of cytoskeletal organizationrelated proteins (e.g., actin-related protein 2/3 complex (ARPC) and actin-related protein (ACTR)) and glycolysis-related proteins (e.g., L-lactate dehydrogenase A chain (LDHA), phosphoglycerate kinase 1 (PGK1), and fructose-bisphosphate aldolase A (ALDOA)) were elevated in COVID-19 convalescents (Fig. 3D). These results suggest that tissue repair processes persist in COVID-19 convalescents for up to 3 years.

Conversely, DEPs in the decreasing clusters were enriched for adaptive immune responses and gas transport (Fig. 3C). Levels of immunoglobulins (e.g., immunoglobulin heavy constant alpha 2 (IGHA2) and immunoglobulin kappa variable 2–30 (IGKV2–30)) and gas transport-related proteins (e.g., bisphosphoglycerate mutase (BPGM) and hemoglobin subunit delta (HBD)) remained low for up to 3 years after hospital discharge (Fig. 3D). These findings imply that immunosuppression and diffusion impairment persist in COVID-19 convalescents for up to 3 years.



Fig. 3 Temporal dynamics and functional insights into proteomic alterations among COVID-19 convalescents over 3 years of follow-up. A Venn diagram illustrating differentially expressed proteins (DEPs) in COVID-19 convalescents at the 6-month (6m), 1-year (1y), 2-year (2y), and 3-year (3y) follow-up visits compared with healthy controls (CONs). B Cluster analysis of temporal expression patterns for all DEPs. C GO pathway analysis of DEPs across different expression patterns, showing the top 15 GO terms. Red boxes highlight important terms. D Heatmap showing expression levels of DEPs in CONs and COVID-19 convalescents at various time points. DEPs associated with key GO pathways (supramolecular fiber organization, cytoskeletal organization, adaptive immune response, and gas transport) are highlighted

Metabolomic signature dynamics over 3 years after SARS-CoV-2 infection

Metabolomic analyses were conducted to evaluate dynamic changes in the metabolic characteristics of COVID-19 patients over the 3-year follow-up period. In total, 966 DEMs were significantly altered across the four groups; of these DEMs, 291 were overlapping (Fig. 4A and Additional file 2: Table S3.2). PLS-DA (Additional file 1: Fig. S4A) and volcano plots (Additional file 1: Fig. S4B-D) were used to illustrate group separation. Clustering analysis of the 966 DEMs identified eight clusters, including five increasing clusters (1, 2, 5, 6, and 7), two decreasing clusters (3 and 8), and one inverted "V" cluster (4) (Fig. 4B). DEMs in the increasing clusters were enriched for metabolites involved in glutathione metabolism; alanine, aspartate, and glutamate metabolism; and ascorbate and aldarate metabolism. In contrast, DEMs in the decreasing clusters were enriched for glycine, serine, and threonine metabolism; DEMs in the inverted "V" cluster were enriched for histidine metabolism (Fig. 4C).

Intriguingly, we found that lipid disturbance was a key characteristic of COVID-19 convalescents. Compared with CONs, COVID-19 convalescents exhibited increased levels of glycerophospholipids, prenol lipids, and polyketides, whereas decreased levels of fatty acyls were evident across the four follow-up time points (Fig. 4D). These pronounced metabolic alterations might be associated with repair processes and the dysregulation of transport mechanisms. Both clinical and molecular signatures indicated that reduced DLco was the most prominent feature of impaired lung function in COVID-19 convalescents.

Proteomic and metabolomic profiling of COVID-19 convalescents with aDLco at 6 months

Reduced DLco is a primary feature of COVID-19. To determine whether the molecular changes discussed



Fig. 4 Temporal dynamics and functional insights into metabolic alterations among COVID-19 convalescents over 3 years of follow-up. **A** Venn diagram illustrating DEMs in COVID-19 convalescents at the 6-month (6m), 1-year (1y), 2-year (2y), and 3-year (3y) follow-up visits compared with healthy controls (CONs). **B** Cluster analysis of temporal expression patterns for all DEMs. **C** Kyoto Encyclopedia of Genes and Genome (KEGG) pathway analysis of DEMs across different expression patterns, showing the top 15 KEGG terms. **D** Differential lipid metabolite analysis over time in COVID-19 convalescents. The y-axis represents the sum of the peak areas for all lipid metabolites (e.g., fatty acyls, glycerophospholipids, prenol lipids, polyketides, and glycerolipids). *P*-values were calculated using the Wilcoxon rank sum test and Dunn's test. The significant *P*-values are indicated on the boxplots. **P* < 0.05, ***P* < 0.01

above are related to reduced DLco, we examined DEPs and DEMs in patients with nDLco and patients with aDLco. Proteomic analysis showed that 182 proteins differed between the aDLco group and CONs, whereas 199 proteins differed between the nDLco group and CONs (Additional file 2: Table S3.3). Among these sets of DEPs, only six overlapped between the two comparisons (Fig. 5A). Four expression patterns were determined based on clustering analysis of 245 DEPs (Fig. 5B). The stratification and DEPs for each group are shown in Additional file 1: Fig. S5. DEPs in cluster 1, which exhibited a decreasing trend in COVID-19



Fig. 5 Proteomic alterations in COVID-19 convalescents with abnormal DLco at 6 months. **A** Venn diagram illustrating DEPs among COVID-19 convalescents with aDLco at 6 months (nDLco-6m), and healthy controls (CONs). **B** Differential expression and hierarchical clustering of proteins among CON, nDLco-6m, and aDLco-6m groups. Heatmap on the left displays protein expression levels; proteins are grouped into clusters based on similar expression patterns, and the number of proteins in each cluster is indicated. **C** Gene Ontology biological process (GO-BP) enrichment analysis for DEPs in cluster 1. Key biological processes include oxygen transport, immune system processes, and adaptive immune responses. **D** KEGG pathway enrichment analysis for DEPs in clusters 3 and 4. Key enriched pathways include glycolysis/ gluconeogenesis, regulation of the actin cytoskeleton, and tight junctions

convalescents, were enriched in gas transport-related pathways and immune-related pathways (Fig. 5C). Additionally, KEGG pathway analysis of DEPs in clusters 3 and 4 revealed upregulation of glycolysis/gluconeogenesis in COVID-19 convalescents compared with CONs (Fig. 5D). These results implied that patients with aDLco had higher levels of immunosuppression, abnormal gas transport, and pulmonary function repair.

Metabolomic analysis identified 565 metabolites that differed between the aDLco group and CONs, as well as 620 metabolites that differed between the nDLco group and CONs (Additional file 2: Table S3.4). Only seven DEMs overlapped among these three groups (Fig. 6A). The stratification and DEMs for each group are shown in Additional file 1: Fig. S6. Clustering analysis of these significantly altered metabolites revealed four distinct clustering patterns (Fig. 6B). DEMs in cluster 1, which showed an increasing trend in aDLco convalescents, were enriched in nitrogen metabolism, phenylalanine metabolism, and taurine and hypotaurine metabolism (Fig. 6C). DEMs in clusters 2 and 3, which exhibited



Fig. 6 Metabolic alterations in COVID-19 convalescents with abnormal DLco at 6 months. A Venn diagram illustrating DEMs among COVID-19 convalescents with aDLco at 6 months (aDLco-6m), nDLco at 6 months (nDLco-6m), and healthy controls (CONs). B Differential expression and hierarchical clustering of metabolites among CON, nDLco-6m, and aDLco-6m groups. Heatmap on the left displays metabolite expression levels; metabolites are grouped into clusters based on similar expression patterns, and the number of metabolites in each cluster is indicated. C A bubble chart illustrating KEGG pathway enrichment for metabolite clusters. Highlighted pathways include glycolysis/gluconeogenesis, pyruvate metabolism, and the tricarboxylic acid (TCA) cycle

decreasing trends in aDLco convalescents, were enriched in glycerolipid metabolism, vitamin B6 metabolism, and fructose and mannose metabolism (Fig. 6C). Notably, consistent with the proteomic analysis, metabolites associated with the glycolysis/gluconeogenesis pathway were also upregulated in COVID-19 convalescents, particularly those with abnormal DLco (Fig. 6C). Comparison of KEGG pathways between the nDLco and aDLco groups at 6 months after hospital discharge indicated that most pathways were shared between groups, but individuals with reduced DLco also exhibited unique metabolic alterations after SARS-CoV-2 infection. The specific dysregulated metabolites and associated pathways are presented in Fig. 6C. Intriguingly, despite significant phenotypic differences between the two groups, the plasma proteomic and metabolomic characteristics of individuals with reduced DLco were not highly distinct. This finding suggests that a combined diagnostic panel is needed to distinguish COVID-19 convalescents with respiratory sequelae from CONs.

Distinguishing COVID-19 convalescents with respiratory sequelae at 6 months from controls

Due to the prevalence of respiratory sequelae among individuals with PASC, there is an urgent need to identify COVID-19 convalescents with a high risk of these sequelae. We developed a biomarker panel to predict diffusion impairment using machine learning models. To facilitate model development and confirm the diagnostic value of the selected molecules, samples were randomly divided into a training cohort (Fig. 7A) and a test cohort (Fig. 7B). Considering that multiple proteins and metabolites were associated with diffusion impairment in COVID-19 convalescents, we selected six overlapping DEPs and seven overlapping DEMs for machine learning analyses with the following algorithms: logistic regression, random forest, linear support vector machine, k-nearest neighbor, and neural network. We used tenfold cross-validation to assess each model's accuracy and error rate (Additional file 2: Table S4). Ultimately, the decision tree algorithm was identified as the best diagnostic model (Fig. 7A) and validated using the test cohort (Fig. 7B). As shown in Fig. 7C, the decision tree model included four splitters and three distinct molecules-heterogeneous nuclear ribonucleoprotein K (HNRNPK), arachidonoyl-EA, and 1-O-(2r-hydroxy-pentadecyl)-sn-glycerol-to classify four terminal nodes.

As depicted in Fig. 7D, the optimal marker set included one protein (HNRNPK) and two metabolites (arachidonoyl-EA and 1-O-(2r-hydroxy-pentadecyl)-sn-glycerol). This marker set exhibited better predictive ability compared with other molecules. The model distinguished the aDLco and CON groups with 100% sensitivity and 100% specificity. Moreover, it was able to differentiate aDLco and nDLco groups with an AUC value of 0.863, suggesting that this marker set can identify respiratory sequelae in COVID-19 convalescents at 6 months after hospital discharge.

Sensitivity and specificity of ROC curves generated from the test set are presented in Fig. 7E and Additional file 1: Fig. S7. In the testing cohort, the sensitivity and specificity of this panel for distinguishing the aDLco and CON groups were both 1.00; for distinguishing the aDLco and nDLco groups, the sensitivity and specificity were 0.708 and 0.75. These results indicate that the plasma biomarker panel developed in this study can facilitate clinical identification of respiratory sequelae in COVID-19 convalescents after hospital discharge.

Discussion

Due to the high heterogeneity of PASC, there remains disagreement concerning its definition, and the underlying mechanisms are poorly understood. Previous studies have shown that symptoms and clinical parameters alone cannot distinguish PASC subtypes; thus, a combined proteomic and metabolomic analysis offers insights concerning the molecular sequelae of PASC, which may allow clinicians to predict its occurrence and develop targeted therapeutic approaches [8, 10, 18]. Respiratory sequelae are the main cluster of PASC symptoms. After the resolution of acute disease, some patients exhibit persistent respiratory symptoms, as well as residual radiologic abnormalities and impaired lung function, which seriously affect their health-related quality of life [19, 20]. To explore the contributing factors, we conducted a longitudinal assessment of the plasma proteomes and metabolomes from 6 months to 3 years after acute disease. Additionally, we developed a minimal panel of molecular markers using machine learning models to predict diffusion impairment.

We systematically characterized pulmonary sequelae over 3 years of follow-up. Consistent with the results of a previous study [21], we found that pulmonary abnormalities persisted in COVID-19 convalescents at 3 years; reduced DLco constituted the most prominent feature and was inversely correlated with the extent of lung abnormalities (e.g., GGO, reticulation or consolidation, and traction bronchiectasis and/or bronchiectasis). Considering the clinical characteristics of PASC in our cohort, we performed a longitudinal multi-omics analysis to identify the dynamic changes in the proteomes and metabolomes of COVID-19 convalescents over 3 years of follow-up. Participants were divided into nDLco and aDLco groups to explore differentially altered pathways, with the goal of elucidating molecular mechanisms that underlie diffusion impairment.



Fig. 7 Identification and validation of predictive biomarkers for diffusion impairment in COVID-19 convalescents. **A** Biomarker discovery process. Six proteins and seven metabolites were identified and then refined using a random forest model to select three markers. **B** Validation of selected markers. The model used four splitters and three distinct molecules—heterogeneous nuclear ribonucleoprotein K (HNRNPK), arachidonoyl-EA, and 1-O-(2r-hydroxy-pentadecyl)-sn-glycerol—to classify four terminal nodes. **C** Optimal marker set. The optimal marker set comprised one protein (HNRNPK) and two metabolites (arachidonoyl-EA and 1-O-(2r-hydroxy-pentadecyl)-sn-glycerol). **D** Receiver operating characteristic (ROC) curves for marker validation. ROC curves assessed the predictive performance of the identified markers. The left graph compares the aDLco at 6 months (aDLco-6m) and healthy control (CON) groups, whereas the right graph compares the aDLco-6m and nDLco at 6 months (nDLco-6m) groups. Area under the curve (AUC) values for individual markers and the logistic panel demonstrate their diagnostic accuracy. **E** Confusion matrix analysis of combined panel performance among plasma samples from the test cohort

After identifying respiratory sequelae that persisted in COVID-19 convalescents at 3 years after hospital discharge, we sought to determine whether plasma proteomic signaling components also exhibited persistent abnormalities. The present study revealed an activation signature for cytoskeletal organization in COVID-19 convalescents. The cytoskeletal system is suspected to participate in SARS-CoV-2 infection [22–24] through the upregulation of proteins such as actin beta (ACTB) [25] and tubulin alpha-1C chain (TUBA1C) [26]. SARS-CoV-2 can interact with and control the host actin cytoskeleton, leading to increases in viral load, cell-tocell spread, and immune dysfunction [23, 24]. Thus far, research concerning cytoskeletal organization has mainly focused on its role in acute COVID-19; little is known about its role in PASC. We observed the upregulation of cytoskeletal organization-related proteins in COVID-19 convalescents; these proteins were associated with PFT parameters. Grote et al. also found that cytoskeletal proteins, including talin 1 (TLN1) and parvin beta (PARVB), were increased in the high-density lipoprotein proteome of patients with post-COVID-19 symptoms [27]. Furthermore, Sfera et al. revealed that SARS-CoV-2 hijacks the actin cytoskeleton, causing host cell fusion and eliciting pathological syncytia that may contribute to cognitive impairment in COVID-19 patients [28].

Proteomic analysis identified three predominant pathways that remained abnormal at 3 years after hospital discharge and displayed associations with diffusion impairment, including the suppression of gas transport and adaptive immune responses and the activation of glycolysis/gluconeogenesis. We observed persistent suppression of gas transport pathways in COVID-19 convalescents, even at 3 years after hospital discharge. Our proteomic findings were consistent with the clinical PFT results, suggesting that SARS-CoV-2 infection caused severe and prolonged damage to gas exchange mechanisms. Most DEPs involved in gas transport pathways were hemoglobin subunits, including HBD, hemoglobin subunit beta (HBB), hemoglobin subunit gamma-2 (HBG2), and hemoglobin subunit alpha-2 (HBA2). The expression levels of these hemoglobin subunits reportedly are correlated with COVID-19 severity. During SARS-CoV-2 infection, the dysregulation of iron metabolism and anemia can cause multiorgan damage. Dysregulated expression of hemoglobin subunits leads to imbalances in globin chain synthesis, ultimately disrupting erythropoiesis [29]. Additionally, we observed that adaptive immune responses were depressed; most DEPs involved in these pathways were immunoglobulins. Our longitudinal data showed that immunoglobulin expression gradually declined over the four follow-up visits; levels were lower in the aDLco group than in the nDLco group. Consistent with our results, Yin et al. observed the downregulation of immunoglobin kappa, lambda, and heavy chain genes in individuals with long COVID (i.e., PASC) via bulk RNA-sequencing [30]. There is also evidence that immunoglobulin signatures can be used for early identification of patients with high PASC risk [31]. Our results indicate immune dysregulation in COVID-19 convalescents, especially individuals with reduced DLco. Glucose metabolism contributes to the regulation of multiple physiological processes, including viral infection, immunity, tumorigenesis, and homeostasis [32]. During acute COVID-19, enhanced glycolysis is a key metabolic feature required for the replication of SARS-CoV-2 [33, 34]. Glucose metabolism also exhibits longterm alteration after SARS-CoV-2 infection [35, 36]. In the present study, glycolysis/gluconeogenesis activation was observed in COVID-19 convalescents. Compared with the nDLco group, the extent of glycolysis/gluconeogenesis activation was higher in the aDLco group. Previous works have extensively investigated dynamic alterations in gas transport, immune response, and glycolysis/gluconeogenesis mechanisms during acute COVID-19, but the long-term effects have remained unclear. Our results provide comprehensive insights concerning proteomic and metabolic alterations, as well as their associations with long-term COVID-19 outcomes, potentially enhancing the understanding of PASC and enabling the identification of therapeutic targets.

Using the multi-omics profile, we developed a panel consisting of one protein (HNRNPK) and two metabolites (arachidonoyl-EA and 1-O-(2r-hydroxy-pentadecyl)-snglycerol), which demonstrated excellent predictive performance. The panel achieved AUC values of 1.00 and 0.863 for distinguishing the aDLco group from CONs and the nDLco group from CONs, respectively. HNRNPK is an important RNA-binding protein involved in regulating gene expression, which plays a crucial role in the replication of various viruses [37, 38]. Notably, HNRNPK may contribute to the replication cycle of SARS-CoV-2. A network analysis by Qi et al. showed that HNRNPK and its interacting host genes were significantly upregulated in patients with severe COVID-19 [39]. However, the role of HNRNPK in aDLco remains unclear. The two metabolites in our diagnostic panel constitute an organonitrogen compound and a fatty acyl, respectively. Metabolic alterations have also been observed in a cohort of COVID-19 convalescents at 3 months after hospital discharge, linking pathways such as lysine degradation and glycerophospholipid metabolism to impaired DLco [11]. These findings suggest that persistently altered proteins and metabolites comprise an effective biomarker panel for predicting diffusion impairment.

There were some limitations in this study. First, because the patients were infected with SARS-CoV-2 during the first wave of COVID-19, data were unavailable regarding vaccination status and treatments; this lack of information should be considered when interpreting the results. Second, patients in this study were infected with the original pandemic strain of SARS-CoV-2. Considering the differences in pathogenicity among SARS-CoV-2 variants, our results might not fully explain PASC caused by other SARS-CoV-2 variants. Third, although the samples were age-matched, potential confounding factors such as diet, vaccination, or environmental confounding factors (e.g., pathogen type) may influence the generalizability of our findings. Fourth, the proposed potential signatures identified by multi-omics require further validation using a more rigorous methodology. Finally, although the results were verified using a test cohort, the identified biomarkers and pathways associated with PASC (particularly respiratory sequelae) require confirmation through external validation in large prospective cohort studies.

Conclusions

In conclusion, this study provides insights concerning molecular sequelae among COVID-19 convalescents over 3 years after hospital discharge. We comprehensively characterized respiratory sequelae from 6 months to 3 years post-discharge, revealing key proteomic and metabolomic signatures associated with these complications. Additionally, we developed a biomarker panel to predict diffusion impairment among COVID-19 convalescents.

Abbreviations

SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
COVID-19	Coronavirus disease 2019
PASC	Post-acute sequelae of COVID-19
PFT	Pulmonary function test
DLco	Diffusing capacity for carbon monoxide
nDLco	Normal DLco
aDLco	Abnormal DLco
6MWT	6-Minute walking test
HRCT	High-resolution chest computed tomography
CONs	Healthy controls
RV	Residual volume
FVC	Forced vital capacity
FEV1	Forced expiratory volume in one second
TLC	Total lung capacity
MMEF	Maximum mid-expiratory flow
GGOs	Ground-glass opacities
ESI	Electrospray ionization
HMDB	Human metabolome database
FC	Fold change
DEPs	Differentially expressed proteins
DEMs	Differentially expressed metabolites
FDR	False discovery rate
PLS-DA	Partial least squares discriminant analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
ROC	Receiver operating characteristic
AUC	Area under the curve

ARPC	Actin-related protein 2/3 complex
ACTR	Actin-related protein
LDHA	L-lactate dehydrogenase A chain
PGK1	Phosphoglycerate kinase 1
ALDOA	Fructose-bisphosphate aldolase A
GHA2	Immunoglobulin heavy constant alpha 2
GKV2-30	Immunoglobulin kappa variable 2–30
BPGM	Bisphosphoglycerate mutase
HBD	Hemoglobin subunit delta
HNRNPK	Heterogeneous nuclear ribonucleoprotein K
ACTB	Actin beta
TUBA1C	Tubulin alpha-1C chain
TLN1	Talin 1
PARVB	Parvin beta
HBB	Hemoglobin subunit beta
HBG2	Hemoglobin subunit gamma-2
HBA2	Hemoglobin subunit alpha-2

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-03971-w.

Additional file 1: Fig. S1. Flowchart of participant inclusion. Related to Figure 1. Fig. S2. Analysis of differential protein expression in COVID-19 convalescents over time. Related to Figure 3. Fig. S3. Enriched KEGG pathways in COVID-19 convalescents. Related to Figure 3. Fig. S4. Analysis of differential metabolite expression in COVID-19 convalescents over time. Related to Figure 4. Fig. S5. Comparative analysis of differential protein expression based on lung diffusion capacity in COVID-19 convalescents. Related to Figure 5. Fig. S6. Comparative analysis of differential metabolite expression based on lung diffusion capacity in COVID-19 convalescents. Related to Figure 6. Fig. S7. Sensitivities and specificities of ROC curves generated from the test set. Related to Figure 7.

Additional file 2: Table S1 Comparison of demographic and clinical characteristics between COVID-19 convalescents and healthy controls. Table S2 Baseline characteristics of nDLco and aDLco groups. Table S3.1 Differentially expressed proteins (DEPs) among COVID-19-6m, COVID-19-1y, COVID-19-2y, COVID-19-3y, and CON groups. Table S3.2 Differentially expressed metabolites (DEMs) among COVID-19-6m, COVID-19-1y, COVID-19-2y, COVID-19-3y, and CON groups. Table S3.3 Differentially expressed proteins (DEPs) among aDLco-6m, nDLco-6m, and CON groups. Table S3.4 Differentially expressed metabolites (DEMs) among aDLco-6m, nDLco-6m, and CON groups. Table S4 AUC, sensitivity, and specificity of five machine learning classifiers for distinguishing COVID-19 convalescents with aDLco from COVID-19 convalescents with nDLco and CONs.

Acknowledgements

We thank all authors of this manuscript.

HQY, LJG, YX, XYL, LYG, ZJZ, HFZ, HMM, FJL, and XH collected the data. JQL, HQY, and LJG analyzed and interpreted the data. JQL, HQY, and LJG prepared the figures. HQY and LJG drafted the manuscript. JQL and ZHT revised the manuscript.

Funding

This work was supported by grants from the Ministry of Science and Technology of the People's Republic of China (2023YFC0872500), Beijing Research Center for Respiratory Infectious Diseases Project (BJRID2024-008 and BJRID2024-011), Capital's Funds for Health Improvement and Research (2022–1-1061), Reform and Development Program of the Beijing Institute of Respiratory Medicine (Ggyfz202401), and The National Natural Science Foundation of China (82200007, 82070005, and 82270009).

Data availability

The raw data have been deposited in the the ProteomeXchange Consortium (https://proteomecentral.proteomexchange.org) via the iProX partner repository with the dataset identifier PXD061113.

Declarations

Ethics approval and consent to participate

The study was approved by the Research Ethics Commission of Chao-Yang Hospital (approval no. KY-2021–15). Written informed consent to participate was obtained from all participants during follow-up visits.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Respiratory and Critical Care Medicine, Beijing Institute of Respiratory Medicine and Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China. ²Medical Research Center, Beijing Institute of Respiratory Medicine and Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China. ³Beijing Research Center for Respiratory Infectious Diseases, Beijing, China.

Received: 18 October 2024 Accepted: 26 February 2025 Published online: 05 March 2025

References

- 1. WHO. Coronavirus disease (COVID-19) dashboard. 2024. Available from: https://www.covid19whoint. Accessed 21 Apr 2024.
- Mizrahi B, Sudry T, Flaks-Manov N, Yehezkelli Y, Kalkstein N, Akiva P, Ekka-Zohar A, Ben David SS, Lerner U, Bivas-Benita M, et al. Long covid outcomes at one year after mild SARS-CoV-2 infection: nationwide cohort study. BMJ. 2023;380: e072529.
- Hastie CE, Lowe DJ, McAuley A, Mills NL, Winter AJ, Black C, Scott JT, O'Donnell CA, Blane DN, Browne S, et al. Natural history of long-COVID in a nationwide, population cohort study. Nat Commun. 2023;14(1):3504.
- Peter RS, Nieters A, Krausslich HG, Brockmann SO, Gopel S, Kindle G, Merle U, Steinacker JM, Rothenbacher D, Kern WV, et al. Post-acute sequelae of covid-19 six to 12 months after infection: population based study. BMJ. 2022;379: e071050.
- Han X, Chen L, Fan Y, Alwalid O, Jia X, Zheng Y, Liu J, Li Y, Cao Y, Gu J, et al. Longitudinal assessment of chest CT findings and pulmonary function after COVID-19 infection. Radiology. 2023;307(2): e222888.
- Yang H, Guan L, Yang T, Ma H, Liu X, Li X, Li J, Tong Z. Two- and 3-year outcomes in convalescent individuals with COVID-19: a prospective cohort study. J Med Virol. 2024;96(4): e29566.
- Liew F, Efstathiou C, Fontanella S, Richardson M, Saunders R, Swieboda D, Sidhu JK, Ascough S, Moore SC, Mohamed N, et al. Large-scale phenotyping of patients with long COVID post-hospitalization reveals mechanistic subtypes of disease. Nat Immunol. 2024;25(4):607–21.
- Wang K, Khoramjoo M, Srinivasan K, Gordon PMK, Mandal R, Jackson D, Sligl W, Grant MB, Penninger JM, Borchers CH, et al. Sequential multiomics analysis identifies clinical phenotypes and predictive biomarkers for long COVID. Cell Rep Med. 2023;4(11): 101254.
- Su Y, Yuan D, Chen DG, Ng RH, Wang K, Choi J, Li S, Hong S, Zhang R, Xie J, et al. Multiple early factors anticipate post-acute COVID-19 sequelae. Cell. 2022;185(5):881-895 e820.
- Talla A, Vasaikar SV, Szeto GL, Lemos MP, Czartoski JL, MacMillan H, Moodie Z, Cohen KW, Fleming LB, Thomson Z, et al. Persistent serum protein signatures define an inflammatory subcategory of long COVID. Nat Commun. 2023;14(1):3417.
- Xu J, Zhou M, Luo P, Yin Z, Wang S, Liao T, Yang F, Wang Z, Yang D, Peng Y, et al. Plasma metabolomic profiling of patients recovered from coronavirus disease 2019 (COVID-19) with pulmonary sequelae 3 months after discharge. Clin Infect Dis. 2021;73(12):2228–39.
- Gu X, Wang S, Zhang W, Li C, Guo L, Wang Z, Li H, Zhang H, Zhou Y, Liang W, et al. Probing long COVID through a proteomic lens: a comprehensive two-year longitudinal cohort study of hospitalised survivors. EBioMedicine. 2023;98: 104851.

- Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, Hohmann E, Chu HY, Luetkemeyer A, Kline S, et al. Remdesivir for the treatment of covid-19 - final report. N Engl J Med. 2020;383(19):1813–26.
- Bankier AA, MacMahon H, Colby T, Gevenois PA, Goo JM, Leung ANC, Lynch DA, Schaefer-Prokop CM, Tomiyama N, Travis WD, et al. Fleischner society: glossary of terms for thoracic imaging. Radiology. 2024;310(2): e232558.
- Chang YC, Yu CJ, Chang SC, Galvin JR, Liu HM, Hsiao CH, Kuo PH, Chen KY, Franks TJ, Huang KM, et al. Pulmonary sequelae in convalescent patients after severe acute respiratory syndrome: evaluation with thin-section CT. Radiology. 2005;236(3):1067–75.
- Liu F, Yang H, Yang T, Zhang Z, Guan L, Gao L, Ma H, Zhang H, Song N, Tong Z, et al. Dysregulated proteasome activity and steroid hormone biosynthesis are associated with mortality among patients with acute COVID-19. J Transl Med. 2024;22(1):626.
- Li J, Wang Y, Zhao W, Yang T, Zhang Q, Yang H, Li X, Tong Z. Multi-omics analysis reveals overactive inflammation and dysregulated metabolism in severe community-acquired pneumonia patients. Respir Res. 2024;25(1):45.
- Woodruff MC, Bonham KS, Anam FA, Walker TA, Faliti CE, Ishii Y, Kaminski CY, Ruunstrom MC, Cooper KR, Truong AD, et al. Chronic inflammation, neutrophil activity, and autoreactivity splits long COVID. Nat Commun. 2023;14(1):4201.
- Han X, Chen L, Guo L, Wu L, Alwalid O, Liu J, Zheng Y, Chen L, Wu W, Li H, et al. Long-term radiological and pulmonary function abnormalities at 3-year post COVID-19 hospitalization: a longitudinal cohort study. Eur Respir J. 2024;64(1):2301612.
- Schlemmer F, Valentin S, Boyer L, Guillaumot A, Chabot F, Dupin C, Le Guen P, Lorillon G, Bergeron A, Basille D, et al. Respiratory recovery trajectories after severe-to-critical COVID-19: a 1-year prospective multicentre study. Eur Respir J. 2023;61(4):2201532.
- Cheon IS, Li C, Son YM, Goplen NP, Wu Y, Cassmann T, Wang Z, Wei X, Tang J, Li Y, et al. Immune signatures underlying post-acute COVID-19 lung sequelae. Sci Immunol. 2021;6(65):eabk1741.
- Hekman RM, Hume AJ, Goel RK, Abo KM, Huang J, Blum BC, Werder RB, Suder EL, Paul I, Phanse S, et al. Actionable cytopathogenic host responses of human alveolar type 2 cells to SARS-CoV-2. Mol Cell. 2020;80(6):1104-1122 e1109.
- 23. Kloc M, Uosef A, Wosik J, Kubiak JZ, Ghobrial RM. Virus interactions with the actin cytoskeleton-what we know and do not know about SARS-CoV-2. Arch Virol. 2022;167(3):737–49.
- 24. Aminpour M, Hameroff S, Tuszynski JA. How COVID-19 hijacks the cytoskeleton: therapeutic implications. Life (Basel). 2022;12(6):814.
- Yang Q, Lin Z, Xue M, Jiang Y, Chen L, Chen J, Liao Y, Lv J, Guo B, Zheng P, et al. Deciphering the omicron variant: integrated omics analysis reveals critical biomarkers and pathophysiological pathways. J Transl Med. 2024;22(1):219.
- Gao T, Gao Y, Liu X, Nie Z, Sun H, Lin K, Peng H, Wang S. Identification and functional analysis of the SARS-COV-2 nucleocapsid protein. BMC Microbiol. 2021;21(1):58.
- Grote K, Schaefer AC, Soufi M, Ruppert V, Linne U, Mukund Bhagwat A, Szymanski W, Graumann J, Gercke Y, Aldudak S, et al. Targeting the highdensity lipoprotein proteome for the treatment of post-acute sequelae of SARS-CoV-2. Int J Mol Sci. 2024;25(8):4522.
- Sfera A, Rahman L, Zapata-Martin Del Campo CM, Kozlakidis Z. Long COVID as a tauopathy: of "brain fog" and "fusogen storms." Int J Mol Sci. 2023;24(16):12648.
- Zhang Z, Lin F, Liu F, Li Q, Li Y, Zhu Z, Guo H, Liu L, Liu X, Liu W, et al. Proteomic profiling reveals a distinctive molecular signature for critically ill COVID-19 patients compared with asthma and chronic obstructive pulmonary disease. Int J Infect Dis. 2022;116:258–67.
- Yin K, Peluso MJ, Luo X, Thomas R, Shin MG, Neidleman J, Andrew A, Young KC, Ma T, Hoh R, et al. Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2. Nat Immunol. 2024;25(2):218–25.
- Cervia C, Zurbuchen Y, Taeschler P, Ballouz T, Menges D, Hasler S, Adamo S, Raeber ME, Bachli E, Rudiger A, et al. Immunoglobulin signature predicts risk of post-acute COVID-19 syndrome. Nat Commun. 2022;13(1):446.

- Chen P, Wu M, He Y, Jiang B, He ML. Metabolic alterations upon SARS-CoV-2 infection and potential therapeutic targets against coronavirus infection. Signal Transduct Target Ther. 2023;8(1):237.
- Codo AC, Davanzo GG, Monteiro LB, de Souza GF, Muraro SP, Virgilioda-Silva JV, Prodonoff JS, Carregari VC, de Biagi Junior CAO, Crunfli F, et al. Elevated glucose levels favor SARS-CoV-2 infection and monocyte response through a HIF-1alpha/glycolysis-dependent axis. Cell Metab. 2020;32(3):437-446 e435.
- de Oliveira LG, de Souza AY, Yamamoto P, Carregari VC, Crunfli F, Reisde-Oliveira G, Costa L, Vendramini PH, Duque EA, Dos Santos NB, et al. SARS-CoV-2 infection impacts carbon metabolism and depends on glutamine for replication in Syrian hamster astrocytes. J Neurochem. 2022;163(2):113–32.
- Li Y, Qin S, Dong L, Qiao S, Wang X, Yu D, Gao P, Hou Y, Quan S, Li Y, et al. Long-term effects of Omicron BA.2 breakthrough infection on immunitymetabolism balance: a 6-month prospective study. Nat Commun. 2024;15(1):2444.
- George MS, Sanchez J, Rollings C, Fear D, Irving P, Sinclair LV, Schurich A. Extracellular vesicles in covid-19 convalescence can regulate t cell metabolism and function. iScience. 2023;26(8):107280.
- Poenisch M, Metz P, Blankenburg H, Ruggieri A, Lee JY, Rupp D, Rebhan I, Diederich K, Kaderali L, Domingues FS, et al. Identification of HNRNPK as regulator of hepatitis C virus particle production. PLoS Pathog. 2015;11(1): e1004573.
- Liu W, Yang D, Sun C, Wang H, Zhao B, Zhou G, Yu L. hnRNP K is a novel internal ribosomal entry site-transacting factor that negatively regulates foot-and-mouth disease virus translation and replication and is antagonized by viral 3C protease. J Virol. 2020;94(17):e00803.
- Qi HX, Shen QD, Zhao HY, Qi GZ, Gao L. Network-based analysis revealed significant interactions between risk genes of severe COVID-19 and host genes interacted with SARS-CoV-2 proteins. Brief Bioinform. 2022;23(1):bbab372.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.