Respiratory Research

RESEARCH

Open Access

Metabolomic characterization of COVID-19 survivors in Jilin province

Panyang Xu[1], Lei Zeng[2,3], Chunyu Wang[4], Jiatong Chai[5], Junguo Yin[6] and Jiancheng Xu[1*]

Abstract

Background The COVID-19 pandemic has escalated into a severe global public health crisis, with persistent sequelae observed in some patients post-discharge. However, metabolomic characterization of the reconvalescent remains unclear.

Methods In this study, serum and urine samples from COVID-19 survivors (n = 16) and healthy subjects (n = 16) underwent testing via the non-targeted metabolomics approach using UPLC-MS/MS. Univariate and multivariate statistical analyses were conducted to delineate the separation between the two sample groups and identify differentially expressed metabolites. By integrating random forest and cluster analysis, potential biomarkers were screened, and the differential metabolites were subsequently subjected to KEGG pathway enrichment analysis.

Results Significant differences were observed in the serum and urine metabolic profiles between the two groups. In serum samples, 1187 metabolites were detected, with 874 identified as significant (457 up-regulated, 417 down-regulated); in urine samples, 960 metabolites were detected, with 39 deemed significant (12 up-regulated, 27 down-regulated). Eight potential biomarkers were identified, with KEGG analysis revealing significant enrichment in several metabolic pathways, including arginine biosynthesis.

Conclusions This study offers an overview of the metabolic profiles in serum and urine of COVID-19 survivors, providing a reference for post-discharge monitoring and the prognosis of COVID-19 patients.

Keywords COVID-19, SARS-CoV-2, UPLC-MS/MS, Metabolomics, Biomarkers

Introduction

Coronavirus disease 2019 (COVID-19), characterized as a severe acute respiratory syndrome caused by SARS-CoV-2, has rapidly spread, posing a significant global public health challenge [1]. In the first 3 months of 2020 alone, over 2 million individuals were infected globally, resulting in 150,000 fatalities [2]. While the majority of research has concentrated on the epidemiology and clinical diagnostics of COVID-19, there are reports of SARS-CoV-2 PCR relapse in patients following two consecutive negative PCR tests [3]. Concurrently, concerns regarding the sequelae following acute COVID-19 recovery have intensified. A prospective follow-up study revealed that nearly half of the patients recovering from SARS-CoV-2 infection continued to exhibit persistent symptoms and decreased lung function 2 months post-infection [4].

*Correspondence:
Jiancheng Xu
xjc@jlu.edu.cn
[1] Department of Laboratory Medicine, First Hospital of Jilin University, Changchun, China
[2] Bethune Institute of Epigenetic Medicine, First Hospital of Jilin University, Changchun, Jilin, China
[3] International Center of Future Science, Jilin University, Changchun, China
[4] State Key Laboratory of Supramolecular Structure and Material Jilin University, Changchun, China
[5] Department of Laboratory Medicine, Zhongnan Hospital of Wuhan University, Wuhan, China
[6] Department of Clinical Laboratory, Changchun Hospital of Traditional Chinese Medicine, Changchun, China

Furthermore, a single-center longitudinal study indicated that clinical sequelae, encompassing cardiovascular, respiratory, and systemic symptoms, are prevalent among COVID-19 survivors [5]. Hence, research to determine the rehabilitation status of COVID-19 patients and to identify biomarkers is crucial.

This study utilized metabolomics to analyze patients recovering from COVID-19. Metabolomics is a powerful tool for qualitative and quantitative studies of small molecular metabolites in biological samples to understand cell physiological and biochemical reactions after exogenous stimulation. Various research fields, including life science, disease diagnosis, drug research and development, employ metabolomics [6].

Mass spectrometry detection enables the analysis of qualitative and quantitative information on thousands of molecules with high sensitivity, resolution, selectivity, specificity, and accuracy [7]. Recent studies applied metabolomics to identify COVID-19 biomarkers and search for therapeutic drug targets [8–10]. A cross-sectional study of serum metabolomics using UPLC-MS/MS showed differences in amino acids, carbohydrates, fatty acids, and glycerophospholipids among COVID-19 patients with different severity levels [11]. Bruzzone et al. observed abnormally elevated levels of ketone bodies (acetylacetate, 3-hydroxybutyrate, and acetone) and 2-hydroxybutyrate acid in response to SARS-CoV-2 infection [12]. Previous studies have shown that, despite post-recovery from COVID-19, a considerable proportionof survivors exhibit diffuse lung abnormalities and 13% of patients displaydecreased eGFR during follow-up after discharge [4, 13, 14]. Adittional studies have also suggested that survivors may be at risk of developing fibrosis [13, 15, 16]. Therefore, identifying differential metabolites between COVID-19 convalescent patients and healthy individuals is crucial for early intervention and accurate rehabilitation prognosis.

For the aim, this study applied non-targeted metabolomics technology, specifically ultra-performance liquid chromatography–tandem mass spectrometry (UPLC-MS/MS), to characterize the metabolic profiles of convalescent serum and urine in COVID-19 patients. Additionally, the study explored altered metabolic pathways to elucidate the underlying pathophysiology.

Methods

Study participants

A total of 32 participants were included in this prospective study. Specifically, serum samples from 16 COVID-19 recovery patients were collected within 1 month post-discharge from the Changchun Infectious Disease Hospital, along with samples from 16 healthy controls at the First Hospital of Jilin University's physical examination center. After statistical analysis, there was no statistically significant difference between the two sample groups. Urine samples were simultaneously obtained for these subjects. Upon recruitment, all participants tested negative for SARS-CoV-2 nucleic acid via real-time polymerase chain reaction (RT-PCR). COVID-19 recovery patients (Case) were diagnosed and stratified at admission according to the New Coronavirus Pneumonia Prevention and Control Program (7th edition) issued by the National Health Commission of China. Participants with underlying lung diseases were excluded. Serum and urine samples, along with laboratory findings from COVID-19 recovery patients, were collected from the Changchun Infectious Disease Hospital. Patients met the mandatory discharge criteria: normal body temperature for over 3 days, significantly improved respiratory symptoms, and negative results from two consecutive SARS-CoV-2 RNA tests at least 24 h apart. Metabolomic profiling of all 64 samples (serum and urine) was conducted using ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) to quantify identifiable metabolites. The study was reviewed and approved by the Ethics Committee of the First Hospital of Jilin University (AF-IRB-032-05). Written informed consent was wavied from the subject(s).

Non-targeted UPLC–MS/MS analysis

Non-targeted metabolomic analysis was conducted by Calibra Lab at DIAN Diagnostics (Hangzhou, Zhejiang, China) on their CalOmics metabolomics platform. Samples were extracted using methanol in a ratio of 1:4. The mixtures were shaken for 3 min and precipitated by centrifugation at 4000 × g, 10 min at 20 °C. Four aliquots of 100 µL supernatant were transferred to sample plates and dried under blowing nitrogen, then re-dissolved in reconstitution solutions for sample injection into UPLC-MS/MS systems. The instruments for the four UPLC-MS/MS methods are ACQUITY 2D UPLC (Waters, Milford, MA, USA) plus Q Exactive (QE) hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, USA). QE mass spectrometer was operated at a mass resolution of 35,000, the scan range was 70–1000 m/z. In the first UPLC-MS/MS method, QE was operated in positive ESI mode and the UPLC column was C18 reverse-phase (UPLC BEH C18, 2.1 × 100 mm, 1.7 um; Waters); the mobile solutions used in the gradient elution were water (A) and methanol (B) containing 0.05% PFPA and 0.1% FA. In the second UPLC-MS/MS method, QE was operated in negative ESI mode, and the UPLC column was C18 reverse-phase (UPLC BEH C18, 2.1 × 100 mm, 1.7 um; Waters), the mobile solutions used in the gradient elution were water (A) and methanol (B) containing 6.5 mM ammonium bicarbonate at pH 8. The

third UPLC-MS/MS method had the QE operated in ESI positive mode and the UPLC column was C18 reverse-phase (UPLC BEH C18, 2.1 × 100 mm, 1.7 um; Waters), the mobile solutions were water (A) and methanol/acetonitrile/water (B) contain 0.05% PFPA and 0.01% FA. In the fourth method, QE was operated in negative ESI mode, the UPLC column was HILIC (UPLC BEH Amide, 2.1 × 150 mm, 1.7 um; Waters), and the mobile solutions were water (A) and acetonitrile (B) with 10 mM ammonium formate.

Compound identification and quantification

After pre-processing of raw data and data quality control inspection, ion peaks were extracted using proprietary in-house IT hardware and software. Metabolites were identified by searching an in-house library generated from running reference standards commercially purchased or obtained from other sources. Identification of metabolites in samples requires strict matching of three criteria between experimental data and library entry: narrow window retention index (RI), accurate mass with variation less than 10 ppm and MS/MS spectra with high forward and reverse searching scores. For the identified metabolite, we used a single asterisk symbol (*) to indicate that the identification of this metabolite has not been validated by library data entries generated from running purified compound standards through our experimental platforms. But the identification was obtained through literature reports and searching other databases, which is also a very reliable identification. A double asterisk symbol (**) indicates that the identification of this metabolite has not been validated by corresponding standard samples, and the identification were obtained through literature reports and searching other databases, which is a relatively reliable identification. Peak area for each metabolite was calculated using area-under-the-curve.

Data normalization

Before statistical analysis, raw peak areas were normalized to adjust for system fluctuation among different run days. The normalized peak areas were then log-transformed (log_2) to reduce data distribution skewness and be in approximate normal distribution (Gaussian distribution). Missing values in peak area matrix were imputed by using the minimal detection value of a metabolite among all samples. All these analyses were conducted using MetaboAnalyst (version 5.0) [17].

Quality control of metabolome analysis

A blend of internal standards was added to each sample in order to assist with chromatographic peak alignment and monitor instrument stability. The variability of the instrument was assessed by calculating the median relative standard deviation (RSD) of all internal standards in each sample. The median RSD for this study is ≤ 5%, meeting our quality control criteria. Additionally, extracted water samples were used as blanks, and extracted commercial plasma samples were employed to monitor instrument variation.

Pathway analysis

The pathway enrichment analysis was conducted using MetPA [18] based on KEGG database and Pathview [19]. Only significantly different metabolites with associated KEGG ID were included in this analysis. Significance analysis of pathway enrichment was completed by hypergeometric test.

Statistical analysis

All statistical analyses were performed with R software (version 3.4.1). Significantly changed metabolites between case and control groups were found by parametric (student's *t*-Test, ANOVA) or non-parametric (Wilcox's rank test, Kruskal–Wallis, etc.) statistical methods. Multivariate analysis approach orthogonal partial least square discriminant analysis (OPLS-DA) and principal component analysis (PCA) were conducted using mixOmics (version 6.10.9) [20]. The random forest (RF) method was implemented in randomForest (version 4.6-14) [21].

Results visualization were provided for the performed statistical analyses, including volcano plot in differential metabolite test, scatter plot with confidence ellipse in PCA, scatter plot with confidence ellipse and variable importance dot plot in OPLS-DA, and variable mean decrease accuracy dot plot in the model construction.

Results

Non-targeted metabolomic analysis of serum and urine samples using UPLC-MS/MS

Non-targeted analyses of metabolites in serum and urine samples from two patient groups (COVID-19 survivors and healthy controls) were conducted using a UPLC-MS/MS system to identify metabolites that change in COVID-19 survivors.

Variables were selected based on the median RSD of internal standard signal fluctuations in QC samples, and metabolites with a median RSD < 5% underwent subsequent multivariate statistical analysis. In the UPLC-MS/MS dataset, all identified metabolites, both in positive and negative ion modes, were combined and classified based on their chemical taxonomic features, as illustrated in Fig. 1a (serum) and Fig. 1b (urine). A total of 1187 metabolites were detected in serum samples, and 960 metabolites in urine samples, with the three most abundant classes of metabolites in both types of samples being

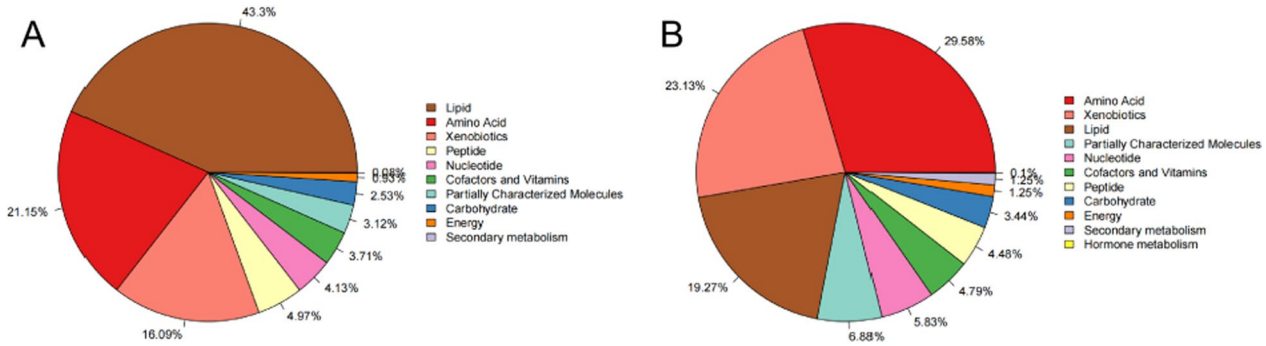

Fig. 1 Proportion of identified metabolites in each chemical class. **a** Serum **b** Urine

lipids (43.3%, 19.27%), amino acids (21.15%, 29.58%), and xenobiotics (16.09%, 23.13%), respectively.

Prior to detailed analysis of specific metabolic changes, PCA and OPLS-DA models were employed to ascertain whether there were differences in metabolic profiles between COVID-19 survivors (case) and the healthy individuals (control). In both PCA and OPLS-DA models, the two groups of serum samples did not exhibit a clear separation trend (Fig. 2a and b). However, compared to the PCA model (Fig. 2c), the OPLS-DA model demonstrated significant differences in the urinary metabolomic profiles between Case and Control, with good reproducibility within each group (Fig. 2d). Furthermore, the Q^2 and R^2 values from the

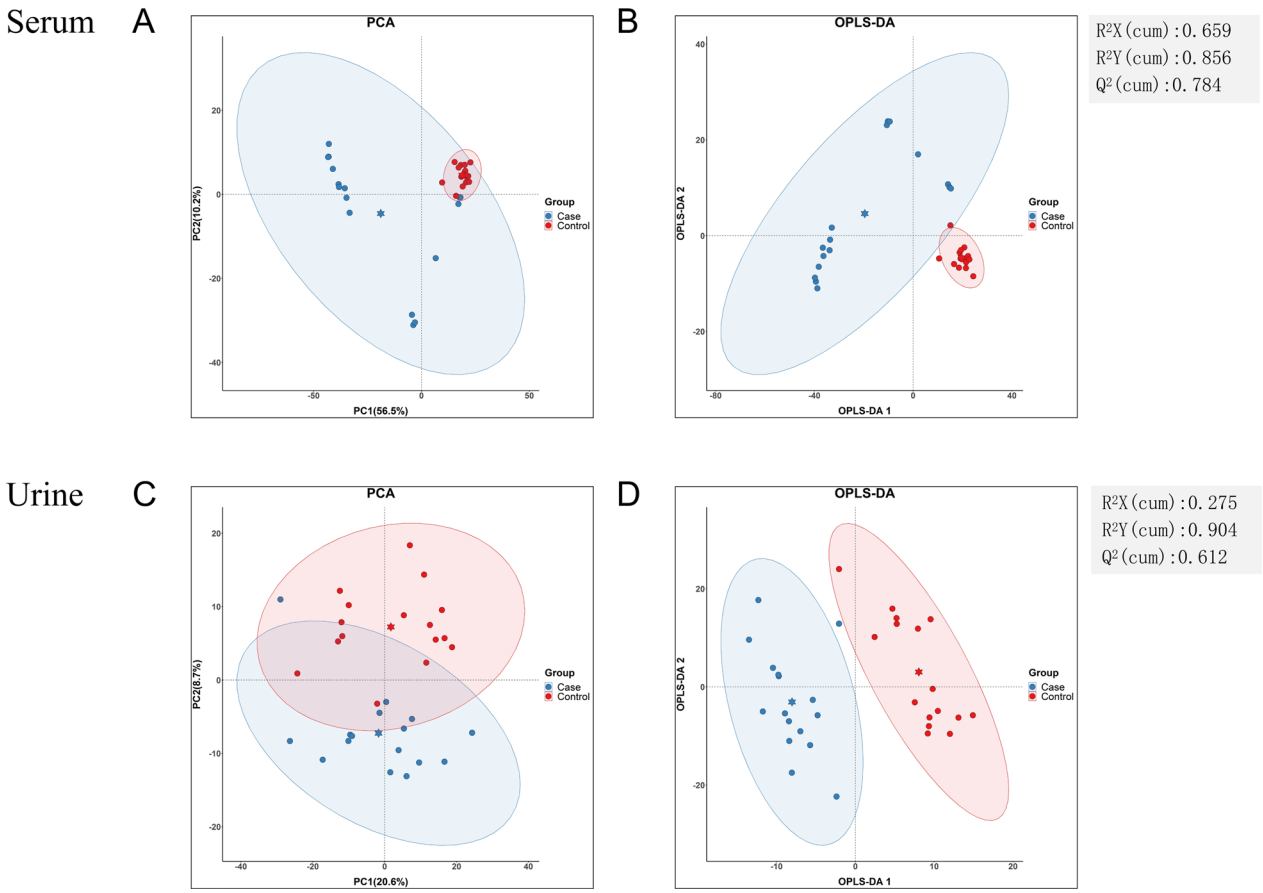

Fig. 2 Plot of the PCA scores **a** Serum **c** Urine. Plot of the OPLS-DA scores **b** Serum **d** Urine

OPLS-DA permutation test exceeded 0.5, indicating high explanatory and predictive power for categorical variables. These results indicated differences in the urinary metabolomic profiles, although they didn't preclude differences in the serum metabolomic profiles between the two groups. However, differences in urine samples were more pronounced in PCA and OPLS-DA compared to serum samples.

Following initial observations of differences, both univariate and multivariate statistical methods were employed to identify distinct metabolites in the samples of the two groups. Metabolites were deemed significantly different in this study if they had an adjusted *P*-value < 0.05, with $\log_2$FC > 1 (red) or $\log_2$FC < 1 (blue), resulting in 874 serum metabolites (457 upregulated, 417 downregulated) and 39 urine metabolites (12 upregulated, 27 downregulated), as shown in the volcano plots (Fig. 3a and b). Varian importance inprojection (VIP) scores were calculated for serum and urine metabolites using the OPLS-DA model, ranking the top 30 metabolites. The top five metabolites in blood (Fig. 3c) were identified as phenylacetylglycine, cis-4-decenoate (10:1n6), methylsuccinate, branched-chain, straight-chain, or cyclopropyl 12:1 fatty acid**, and allantoic acid; in urine (Fig. 3d), they were cis-urocanate, carnitine of C10H1402(4)**, acetylhydroquinone sulfate, pseudoephedrine, and resveratrol sulfate(1).

The Random Forest model analyzed the top 50 metabolites by importance in blood (Out-of-Bag, OOB error rate of 3.12%) and urine (OOB error rate of 6.25%) samples, identifying it as the strongest driver of overall metabolic differences between the healthy individuals and COVID-19 survivors. Based on the literature and KEGG/HMDB databases, metabolites were annotated to one of super pathways corresponding to their general metabolic processes. The most distinctive metabolites primarily originated from pathways. Including: Amino acids, Carbohydrates, Energy, Lipids, Nucleotides, Partially characterized molecules, Peptides, Secondary metabolism, and Xenobiotics (Fig. 4a and b). Based on VIP scores greater than 2 and adjusted *P*-values less than 0.05, serum and urine metabolites were analyzed together, further identifying 16 metabolites with significant differences (Table 1). A heatmap was used to display these significantly different metabolites, showing that in the Case compared to the Control, 11 metabolites were upregulated and 5 were downregulated in urine samples (Fig. 4c). Combining Random Forest and cluster analysis results, eight metabolites, including 1-ribosyl-imidazoleacetate*, carboxyethyl-GABA, cis-urocanate, glucuronide of C10H18O2

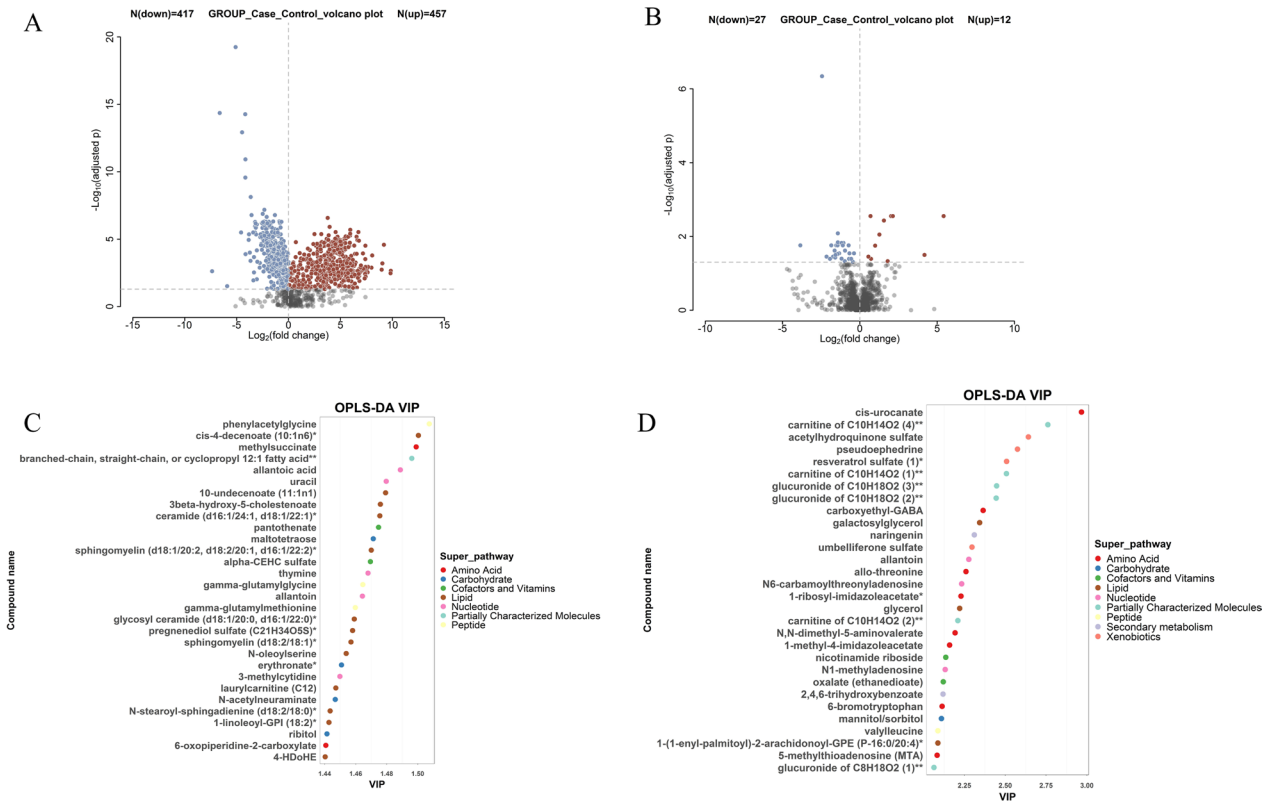

Fig. 3 Fold-change plot showing of metabolism data between Case and Control **a** Serum **b** Urine; OPLS-DA VIP score charts. **c** Serum **d** Urine

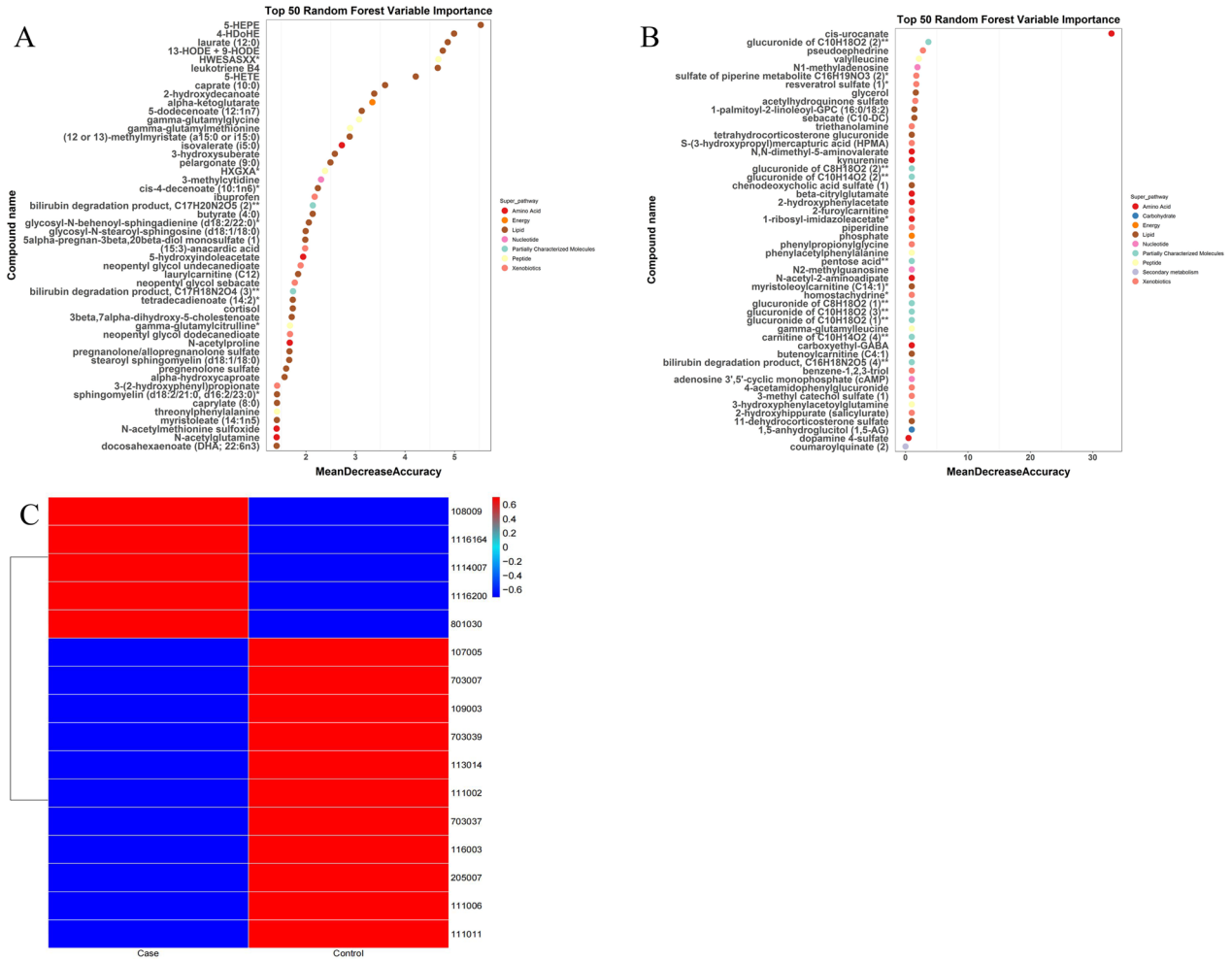

Fig. 4 Random forest model **a** Serum **b** Urine. Clustering heatmap of significant metabolism **c**

(2)**, N,N-dimethyl-5-aminovalerate, N1-methyladenosine, pseudoephedrine, and resveratrol sulfate (1)*, were found to perfectly distinguish between the healthy individuals and COVID-19 survivors, considered potential biomarkers.

Metabolic pathway analysis

To explore metabolic pathways potentially implicated in COVID-19 survivors, metabolites with significant differences between the two groups were enriched, showcasing the top 25 metabolic pathways in blood (Fig. 5a) and urine (Fig. 5b). Results indicated (Table 2) that 11 metabolic pathways exhibited significant changes (FDR < 0.05) between the two groups, namely Alanine, aspartate and glutamate metabolism; Arginine and proline metabolism; Arginine biosynthesis; beta-Alanine metabolism; Biosynthesis of unsaturated fatty acids; Butanoate metabolism; Glycine, serine and threonine metabolism; Histidine metabolism; Nicotinate and nicotinamide metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Valine, leucine and isoleucine biosynthesis.

Discussion

Metabolomics research methodologies are straightforward, with UPLC-MS being the most commonly utilized technique in metabolomics, widely applied in the screening for diagnostic biomarkers of various diseases. This study combines UPLC-MS detection methods with multivariate statistical analysis to investigate the metabolomics of serum and urine in COVID-19 survivors and healthy individuals. The findings demonstrate differences in the serum and urine metabolomic profiles between the two groups, with 874 differential metabolites identified in serum and 39 in urine. Subsequently, a combination analysis of the top-ranked important serum and urine metabolites was conducted using a random forest model and cluster analysis to control confounding factors and enhance the reliability of the results. This results

Table 1 The differential metabolites among Case and Control

Name	Case	Control	*p* (adjusted) < 0.05	VIP
Cysteine-glutathione disulfide	345,686	209,462	0.040476277	2.017095615
5-methylthioadenosine (MTA)	7,355,775	16,706,627	0.014929178	2.082195603
Mannitol/sorbitol	86,251,103	267,772,714	0.024231984	2.108868739
N1-methyladenosine	36,156,360	77,859,547	0.017427998	2.131435068
1-methyl-4-imidazoleacetate	15,800,951	57,023,752	0.017427998	2.158871982
N,N-dimethyl-5-aminovalerate	2,869,240	7,713,153	0.014463537	2.192458239
1-ribosyl-imidazoleacetate*	13,291,227	35,726,159	0.008257586	2.228461812
N6-carbamoylthreonyladenosine	3,906,783	6,482,062	0.017427998	2.23271862
Allo-threonine	877,853	2,225,573	0.028716338	2.259045228
Allantoin	10,748,845	22,428,794	0.014929178	2.27642181
Umbelliferone sulfate	19,644,798	1,085,485	0.031785244	2.295593395
Carboxyethyl-GABA	494,850	1,344,856	0.017427998	2.364157769
Glucuronide of C10H18O2 (2)**	1,352,217	308,990	0.00280968	2.443971607
Resveratrol sulfate (1)*	211,419	88,105	0.008895312	2.507556229
Pseudoephedrine	253,817,785	5,932,328	0.00280968	2.574332372
Cis-urocanate	723,855	3,961,659	4.52485E-07	2.96630827

Differences were considered statistically significant at *p* (adjusted) < 0.05 and VIP > 2. Case:COVID-19 survivors; Control:Healthy individuals

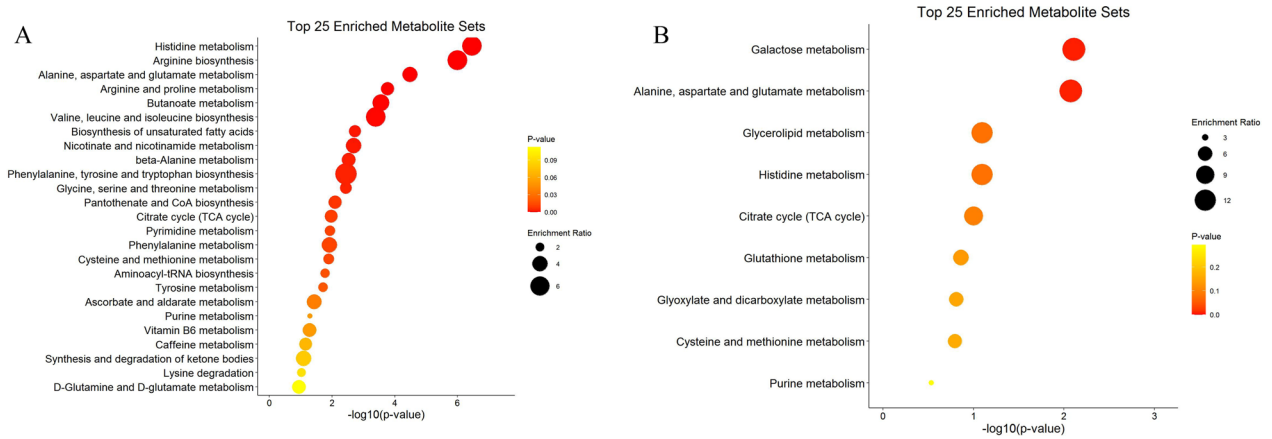

Fig. 5 Enriched KEGG iterms **a** Serum **b** Urine

indicates that, despite recovery and discharge, COVID-19 survivors still exhibit differences in endogenous substances compared to healthy individuals, aligning with the majority of research findings [22, 23].

Among the metabolites that can clearly distinguish COVID-19 survivors from healthy indivivuals in this study, 1-ribosyl-imidazoleacetate* is an intermediate in the synthesis of zoledronic acid, a drug for treating malignant hypercalcemia. In one study, the results confirmed that 1-ribosyl-imidazoleacetate* is positively correlated with ischemic stroke [24]. However, studies specifically targeting 1-ribosyl-imidazoleacetate* in relation to COVID-19 are limited. Similarly, research on the glucuronide of C10H18O2 (2)** is also limited. Carboxyethyl-GABA, although lacking genetic or cytotoxic effects, was found in one study to induce time-dependent proliferation and migration of mouse fibroblasts [25]. Fibroblasts can maintain the structural integrity of connective tissue and secrete a large amount of collagen fibers, thereby playing a role in wound healing. With the passage of time, the increase in carboxyethyl-GABA concentration leads us to hypothesize that carboxyethyl-GABA may be a potential marker for interstitial lung fibrosis, which is related to lung injury. One of the complications following COVID-19 infection is the development of fibrosis. It has been reported that lung fibrosis can be detected early in the infection, regardless of pre-existing lung conditions and disease severity [26].

Table 2 Metabolic pathways significantly altered by Case and Control

Pathway	Pathway ID	Total	Expected	Hits	Raw.*p*	Holm.*p*	FDR	Enrichment Ratio	Log.*P*.value	Metabolite.hits
Alanine, aspartate and glutamate metabolism	hsa00250	28	2.79	11	0.0000324	0.00266	0.000908	3.94265233	4.48945499	N-Acetyl-L-aspartic acid, L-Aspartic acid, Argininosuccinic acid, L-Glutamic acid, 2-Keto-glutaramic acid, Citric acid, Fumaric acid, Pyruvic acid, Ureidosuccinic acid, Succinic acid, Oxoglutaric acid
Arginine and proline metabolism	hsa00330	38	3.79	12	0.00017	0.0138	0.00357	3.166226913	3.769551079	L-Arginine, Guanidoacetic acid, Creatine, S-Adenosylmethionine, Spermidine, N-Acetylputrescine, 4-Hydroxy-L-glutamic acid, L-Proline, L-Glutamic acid, 4-Guanidinobutanoic acid, 4-Acetamidobutanoic acid, Pyruvic acid
Arginine biosynthesis	hsa00220	14	1.39	9	0.000001	0.000083	0.000042	6.474820144	6	L-Glutamic acid, L-Arginine, N-Acetylornithine, Argininosuccinic acid, L-Aspartic acid, Oxoglutaric acid, N-Acetylglutamic acid, Urea, Fumaric acid
Beta-Alanine metabolism	hsa00410	21	2.09	7	0.00297	0.225	0.0276	3.349282297	2.527243551	L-Aspartic acid, Ureidopropionic acid, Dihydrouracil, Carnosine, Uracil, L-Histidine, Spermidine
Biosynthesis of unsaturated fatty acids	hsa01040	36	3.59	10	0.00186	0.145	0.0218	2.78551532	2.730487056	Palmitic acid, Stearic acid, Arachidic acid, Oleic acid, Linoleic acid, Arachidonic acid, 8,11,14-Eicosatrienoic acid, Gamma-Linolenic acid, Docosahexaenoic acid, Eicosapentaenoic acid
Butanoate metabolism	hsa00650	15	1.49	7	0.000275	0.022	0.00463	4.697986577	3.560667306	(R)-3-Hydroxybutyric acid, Acetoacetic acid, L-Glutamic acid, Butyric acid, Oxoglutaric acid, Succinic acid, 2-Hydroxyglutarate
Glycine, serine and threonine metabolism	hsa00260	33	3.29	9	0.00361	0.27	0.0276	2.73556231	2.442492798	Guanidoacetic acid, 3-Phosphoglyceric acid, Dimethylglycine, L-Cystathionine, Glycine, Sarcosine, Creatine, 2-Ketobutyric acid, Pyruvic acid
Histidine metabolism	hsa00340	16	1.59	10	0.000000344	0.0000289	0.0000289	6.289308176	6.463441557	L-Glutamic acid, Urocanic acid, L-Histidine, Carnosine, 1-Methylhistamine, 3-Methylhistidine, Formiminoglutamic acid, Imidazoleacetic acid, Methylimidazoleacetic acid, L-Aspartic acid
Nicotinate and nicotinamide metabolism	hsa00760	15	1.49	6	0.00207	0.16	0.0218	4.026845638	2.684029655	L-Aspartic acid, Quinolinic acid, Nicotinamide riboside, Niacinamide, 1-Methylnicotinamide, N1-Methyl-2-pyridone-5-carboxamide
Phenylalanine, tyrosine and tryptophan biosynthesis	hsa00400	4	0.398	3	0.0036	0.27	0.0276	7.537688442	2.443697499	Phenylpyruvic acid, L-Phenylalanine, 4-Hydroxyphenylpyruvic acid

Table 2 (continued)

Pathway	Pathway ID	Total	Expected	Hits	Raw.*p*	Holm.*p*	FDR	Enrichment Ratio	Log.*P*.value	Metabolite.hits
Valine, leucine and isoleucine biosynthesis	hsa00290	8	0.797	5	0.000402	0.0318	0.00563	6.273525721	3.395773947	3-Methyl-2-oxovaleric acid, 2-Ketobutyric acid, L-Isoleucine, Ketoleucine, L-Valine

The decline in lung function of COVID-19 survivors can last up to 12 months and may even become permanent, especially in the case of fibrosis [27, 28].

N,N-dimethyl-5-aminovalerate is related to the catabolism of microbial corpse alkaloids [29]. A study showed that the plasma metabolic profile of N,N-dimethyl-5-aminovalerate differs significantly before and after long-term antiretroviral therapy, and its metabolite levels can clearly distinguish HIV-infected patients from healthy controls [30]. Therefore, this study speculates that N,N-dimethyl-5-aminovalerate may also be a potential marker for distinguishing between COVID-19 and healthy controls, but further confirmation is needed in future research. Many studies have proven that N1-methyladenosine is closely related to tumor response [31–33]. However, research on N1-methyladenosine in the context of COVID-19 is limited. Pseudoephedrine can be used to treat symptoms of the common cold and flu, sinusitis, asthma, and bronchitis, and is a long-standing drug. Since this study did not completely exclude drug variables, the significant metabolic profile differences in COVID-19 survivors might be due to drug residues. Resveratrol sulfate (1)* is a polyphenolic chemical, and it has been proven that resveratrol can improve inflammatory diseases involving the intestinal mucosa [34, 35]. About half of acute COVID-19 patients experience gastrointestinal symptoms, continuing inapproximately 10%–25% of COVID-19 patients continuing for up to 6 months [36, 37]. Due to the potential interaction between the immune response associated with SARS-CoV-2 infection and the immune dysregulation associated with inflammatory bowel diseases (IBD), resveratrol might offer a new therapeutic approach for COVID-19 survivors. Although research on these substances in the context of the COVID-19 pandemic remains limited, the results of this study can provide new research directions.

Enrichment analysis revealed significant enrichment of the arginine biosynthesis metabolic pathway in the serum of COVID-19 survivors. Arginine not only serves as a crucial substrate for protein synthesis but also as a precursor for the synthesis of substances like creatine, polyamines, and nitric oxide (NO) in the body, playing a significant role in human nutritional metabolism and regulation [38]. The physiologically active form of arginine in the body is L-arginine. Recent research on COVID-19 has found that serum levels of L-arginine in adults and children affected by COVID-19 are significantly lower compared to control groups [39]. Another study demonstrated that serum levels of L-arginine are inversely correlated with the severity of COVID-19 [40]. In vitro assays have shown that *T* cell proliferative capacity is significantly reduced in COVID-19 patients, which can be restored by supplementing with arginine [41]. Recent metabolomics data indicate changes in the L-arginine pathway in COVID-19 patients [42], and an increase in arginase mRNA expression was also found in peripheral blood mononuclear cells (PBMCs) of COVID-19 patients [43]. Reports suggest a close relationship between the expression of arginase or nitric oxide synthase (enzymes essential for arginine catabolism) and airway remodeling in chronic obstructive pulmonary disease (COPD) patients [44]. Data indicates that levels of arginine are reduced in the serum of COVID-19 survivors with pulmonary function abnormalities. The results of this study show that L-arginine levels in the serum of COVID-19 survivors are lower than in healthy individuals, thus suggesting that pulmonary function changes may still persist in COVID-19 survivors, necessitating timely re-examination and monitoring. Furthermore, we speculate that monitoring changes in L-arginine could also be beneficial in managing long COVID-19, as the persistence of chronic inflammation and endothelial dysfunction has been demonstrated to underlie COVID-19 sequelae [45, 46].

Despite these findings, the study has limitations: the sample size is small, and due to the unbiased nature of non-targeted metabolomics, the identified metabolites may have certain biases. Future research should aim to increase the sample size for targeted metabolomics validation.

Conclusions

In this study, UPLC-MS/MS metabolomics was applied to select for differential metabolites in COVID-19 survivors. Co-analysis of the top-ranked importance metabolites in serum and urine identified 16 metabolites with significant differences. Among themwere 1-ribosyl-imidazoleacetate *, carboxyethyl-GABA, cis-urocanate, glucuronide of C10H18O2 (2) * *, N, N-dimethyl-5-aminovalerate, N1-methyladenosine, pseudoephedrine, and resveratrol sulfate (1). * These 8 metabolites are considered as potential biomarkers in COVID-19 survivors. Our research provides new insights into the metabolomics of the COVID-19 recovery phase and may offer potential new therapeutic targets for preventing COVID-19 relapse. Future research is needed to confirm our preliminary data and identify effective diagnostic biomarkers for the COVID-19 recovery phase.

Author contributions

P. Y. X designed the study and wrote the first draft of the manuscript. L. Z, and Y. G. Y collected the clinical data and finished the metabolic experiment. C. Y. W and J. T. C performed the statistical analyses. P. Y. X finished the figures and tables. J. C. X and L. Z provided data analyses, critical revision and final approval. All authors have read and agreed to the published version of the manuscript.

Funding
This work was supported by grants from the Jilin Science and Technology Development Program (no. 20190304110YY to Dr. Jiancheng Xu), and the First Hospital Translational Funding for Scientific and Technological Achievements (no. CGZHYD202012-005 to Dr. Jiancheng Xu).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethical approval
This study was approved by the Ethics Committee of the First Hospital of Jilin University (AF-IRB-032-05). All studies were conducted under the Guidelines for Good Clinical Practice and the 1964 Declaration of Helsinki.

Consent for publication
Informed consent was waived for all subjects. Not applicable.

Competing interests
The authors declare no competing interests.

Received: 5 March 2024 Accepted: 9 September 2024
Published online: 19 September 2024

References

1. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579(7798):270–3.
2. Costa Dos Santos Junior G, Pereira CM, Kelly da Silva Fidalgo T, Valente AP. Saliva NMR-Based Metabolomics in the War Against COVID-19. Anal Chem. 2020;92(24):15688–92.
3. Hoang VT, Dao TL, Gautret P. Recurrence of positive SARS-CoV-2 in patients recovered from COVID-19. J Med Virol. 2020;92(11):2366–7.
4. Trinkmann F, Muller M, Reif A, Kahn N, Kreuter M, Trudzinski F, et al. Residual symptoms and lower lung function in patients recovering from SARS-CoV-2 infection. Eur Respir J. 2021;57(2):1.
5. Xiong Q, Xu M, Li J, Liu Y, Zhang J, Xu Y, et al. Clinical sequelae of COVID-19 survivors in Wuhan, China: a single-centre longitudinal study. Clin Microbiol Infect. 2021;27(1):89–95.
6. Bar N, Korem T, Weissbrod O, Zeevi D, Rothschild D, Leviatan S, et al. A reference map of potential determinants for the human serum metabolome. Nature. 2020;588(7836):135–40.
7. Marshall DD, Powers R. Beyond the paradigm: combining mass spectrometry and nuclear magnetic resonance for metabolomics. Prog Nucl Magn Reson Spectrosc. 2017;100:1–16.
8. Luporini RL, Pott-Junior H, Di Medeiros Leal MCB, Castro A, Ferreira AG, Cominetti MR, et al. Phenylalanine and COVID-19: tracking disease severity markers. Int Immunopharmacol. 2021;101(Pt A):108313.
9. Battaglini D, Lopes-Pacheco M, Castro-Faria-Neto HC, Pelosi P, Rocco PRM. Laboratory biomarkers for diagnosis and prognosis in COVID-19. Front Immunol. 2022;13:857573.
10. Della Corte V, Riolo R, Scaglione S, Pecoraro R, Tuttolomondo A. The role of biomarkers, metabolomics, and COVID-19 in venous thromboembolism—a review of literature. Int J Mol Sci. 2023;24(17):13411.
11. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. Cell. 2020;182(1):59–72 e15.
12. Bruzzone C, Bizkarguenaga M, Gil-Redondo R, Diercks T, Arana E, Garcia de Vicuna A, et al. SARS-CoV-2 infection dysregulates the metabolomic and lipidomic profiles of serum. iScience. 2020;23(10):101645.
13. Esendagli D, Yilmaz A, Akcay S, Ozlu T. Post-COVID syndrome: pulmonary complications. Turk J Med Sci. 2021;51(SI-1):3359–71.
14. Lai CC, Hsu CK, Yen MY, Lee PI, Ko WC, Hsueh PR. Long COVID: an inevitable sequela of SARS-CoV-2 infection. J Microbiol Immunol Infect. 2023;56(1):1–9.
15. McDonald LT. Healing after COVID-19: Are survivors at risk for pulmonary fibrosis? Am J Physiol Lung Cell Mol Physiol. 2021;320(2):L257–65.
16. Vianello A, Guarnieri G, Braccioni F, Lococo S, Molena B, Cecchetto A, et al. The pathogenesis, epidemiology and biomarkers of susceptibility of pulmonary fibrosis in COVID-19 survivors. Clin Chem Lab Med. 2022;60(3):307–16.
17. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. 2021;49(W1):W388–96.
18. Xia J, Wishart DS. MetPA: a web-based metabolomics tool for pathway analysis and visualization. Bioinformatics. 2010;26(18):2342–4.
19. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics. 2013;29(14):1830–1.
20. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput Biol. 2017;13(11):e1005752.
21. A. L, M. W. "Classification and Regression by randomForest.". R News. 2002;2(3):18–22. https://CRAN.R-project.org/doc/Rnews/.
22. Li H, Li X, Wu Q, Wang X, Qin Z, Wang Y, et al. Plasma proteomic and metabolomic characterization of COVID-19 survivors 6 months after discharge. Cell Death Dis. 2022;13(3):235.
23. Li K, Wu Q, Li H, Sun H, Xing Z, Li L, et al. Multiomic characterisation of the long-term sequelae of SARS survivors: a clinical observational study. EClinicalMedicine. 2023;58:101884.
24. He M, Xu C, Yang R, Liu L, Zhou D, Yan S. Causal relationship between human blood metabolites and risk of ischemic stroke: a Mendelian randomization study. Front Genet. 2024;15:1333454.
25. Dos Santos LV, da Silva Brum LF, de Freitas LBR, Miri JM, Pinhatti VR, Fachini J, et al. Carboxyethyl aminobutyric acid (CEGABA) lacks cytotoxicity and genotoxicity and stimulates cell proliferation and migration in vitro. Arch Dermatol Res. 2019;311(6):491–7.
26. Schwensen HF, Borreschmidt LK, Storgaard M, Redsted S, Christensen S, Madsen LB. Fatal pulmonary fibrosis: a post-COVID-19 autopsy case. J Clin Pathol. 2020;74:400–2.
27. Torres-Castro R, Vasconcello-Castillo L, Alsina-Restoy X, Solis-Navarro L, Burgos F, Puppo H, et al. Respiratory function in patients post-infection by COVID-19: a systematic review and meta-analysis. Pulmonology. 2021;27(4):328–37.
28. Wu X, Liu X, Zhou Y, Yu H, Li R, Zhan Q, et al. 3-month, 6-month, 9-month, and 12-month respiratory outcomes in patients following COVID-19-related hospitalisation: a prospective study. Lancet Respir Med. 2021;9(7):747–54.
29. Andorfer L, Holtfreter B, Weiss S, Matthes R, Pitchika V, Schmidt CO, et al. Salivary metabolites associated with a 5-year tooth loss identified in a population-based setting. BMC Med. 2021;19(1):161.
30. Lu L, Yang Y, Yang Z, Wu Y, Liu X, Li X, et al. Altered plasma metabolites and inflammatory networks in HIV-1 infected patients with different immunological responses after long-term antiretroviral therapy. Front Immunol. 2023;14:1254155.
31. Liu F, Clark W, Luo G, Wang X, Fu Y, Wei J, et al. ALKBH1-Mediated tRNA Demethylation Regulates Translation. Cell. 2016;167(3):816–28 e16.
32. Han X, Wang M, Zhao YL, Yang Y, Yang YG. RNA methylations in human cancers. Semin Cancer Biol. 2021;75:97–115.
33. Li L, Yang W, Jia D, Zheng S, Gao Y, Wang G. Establishment of a N1-methyladenosine-related risk signature for breast carcinoma by bioinformatics analysis and experimental validation. Breast Cancer. 2023;30(4):666–84.
34. Nunes S, Danesi F, Del Rio D, Silva P. Resveratrol and inflammatory bowel disease: the evidence so far. Nutr Res Rev. 2018;31(1):85–97.
35. Zhang L, Xue H, Zhao G, Qiao C, Sun X, Pang C, et al. Curcumin and resveratrol suppress dextran sulfate sodium-induced colitis in mice. Mol Med Rep. 2019;19(4):3053–60.
36. Freedberg DE, Chang L. Gastrointestinal symptoms in COVID-19: the long and the short of it. Curr Opin Gastroenterol. 2022;38(6):555–61.
37. Ghoshal UC, Ghoshal U. Gastrointestinal involvement in post-acute coronavirus disease (COVID)-19 syndrome. Curr Opin Infect Dis. 2023;36(5):366–70.
38. Wu G, Meininger CJ, McNeal CJ, Bazer FW, Rhoads JM. Role of L-arginine in nitric oxide synthesis and health in humans. Adv Exp Med Biol. 2021;1332:167–87.

39. Rees CA, Rostad CA, Mantus G, Anderson EJ, Chahroudi A, Jaggi P, et al. Altered amino acid profile in patients with SARS-CoV-2 infection. Proc Natl Acad Sci USA. 2021;118(25):e2101708118.
40. Sacchi A, Grassi G, Notari S, Gili S, Bordoni V, Tartaglia E, et al. Expansion of myeloid derived suppressor cells contributes to platelet activation by L-Arginine deprivation during SARS-CoV-2 infection. Cells. 2021;10(8):2111.
41. Reizine F, Lesouhaitier M, Gregoire M, Pinceaux K, Gacouin A, Maamar A, et al. SARS-CoV-2-induced ARDS associates with MDSC expansion, lymphocyte dysfunction, and arginine shortage. J Clin Immunol. 2021;41(3):515–25.
42. D'Alessandro A, Thomas T, Akpan IJ, Reisz JA, Cendali FI, Gamboni F, et al. Biological and clinical factors contributing to the metabolic heterogeneity of hospitalized patients with and without COVID-19. Cells. 2021;10(9):2293.
43. Derakhshani A, Hemmat N, Asadzadeh Z, Ghaseminia M, Shadbad MA, Jadideslam G, et al. Arginase 1 (Arg1) as an Up-Regulated Gene in COVID-19 patients: a promising marker in COVID-19 immunopathy. J Clin Med. 2021;10(5):1051.
44. Pera T, Zuidhof AB, Smit M, Menzen MH, Klein T, Flik G, et al. Arginase inhibition prevents inflammation and remodeling in a guinea pig model of chronic obstructive pulmonary disease. J Pharmacol Exp Ther. 2014;349(2):229–38.
45. Paneroni M, Pasini E, Vitacca M, Scalvini S, Comini L, Pedrinolla A, et al. Altered vascular endothelium-dependent responsiveness in frail elderly patients recovering from COVID-19 pneumonia: preliminary evidence. J Clin Med. 2021;10(12):2558.
46. Yan Z, Yang M, Lai CL. Long COVID-19 syndrome: a comprehensive review of its effect on various organ systems and recommendation on rehabilitation plans. Biomedicines. 2021;9(8):966.

Publisher's Note