SUPPLEMENTARY MATERIALS AND METHODS

Sample collection

Using a syringe, blood samples (10 mL) were collected from the index fingers of acute liver failure (ALF) patients after enrollment in the study. Patients with portal vein thrombosis were excluded from the study. An INR>1.5, encephalopathy, moderate elevation of aspartate aminotransferase over alanine aminotransferase, and elevated total serum bilirubin were the inclusion criteria.

Plasma depletion

Blood samples (10 mL) were collected at baseline and centrifuged for 20 minutes at 2,000 rpm and 4°C for plasma isolation. The plasma samples were subjected to plasma depletion, followed by preparation for lipidomics and fungal-peptide analyses using mass spectrometry.[1-4]

Separation of the low abundant protein (LAP) and albuminous fraction

Albumin was removed from the plasma samples of the study cohort of ALF patients (ALF-NS=32, ALF-S=8) and healthy controls (HC) (n=5) using a dye-binding affinity chromatographic technique with 1 mL HiTrap Blue HP (Cat no. 17-041201; GE Healthcare Life Sciences, Uppsala, Sweden). The analysis was performed on an AKTA Pure purification system prepacked with blue Sepharose beads immobilized by Cibacron Blue F3G-A dye, which binds specifically to albumin.[5-8] For this test, 100 µL of the baseline plasma was diluted 1:4 with binding buffer (20 mM Na_2HPO_4) and then loaded onto a 1 mL HiTrap Blue column for 2 minutes, after which we washed the column with 20 mM Na_2HPO_4 for 5 minutes. Analysis produced a sharp peak of unbound LAPs. An elution phase of 7 minutes followed, during which we increased the percentage (5–98%) of 2 M NaCl+20 mM Na_2HPO_4 to elute the dye-bound albumin, as mentioned in a previous paper about SAH and ALF.[9]

Fungal peptide analysis (fungome)

Sample preparation

After albumin was depleted from the baseline plasma, protein quantification of the LAP samples for all study groups (ALF-NS=32, ALF-S=8, HC=5) was conducted using the Bradford assay.[10]

In-solution protein digestion

1. 50 µg of equivalent protein was vacuum dried and reconstituted in 200 µL of ammonium bicarbonate buffer.
2. To reduce the disulfide bonding between the proteins, 20 µL of 10 mM dithiothreitol was added, followed by incubation at 60°C for 1 hour in a water bath.
3. Post-reduction, 15 µL of 10 mM iodoacetamide was added and incubated for 30 minutes in the dark to allow alkylation of proteins.
4. For digestion, 10 µL of modified sequencing-grade trypsin (0.1 µg/µL) was added and incubated at 37°C for 20–24 hours in a water bath. To inhibit trypsin activity in the reaction mixture, 5 µL of concentrated formic acid was added.[1,11]

C18 column–mediated desalting of the peptides

The digested LAP samples were subjected to peptide purification using a Thermo Scientific Pierce C18 Spin (Catalog number: 89873), which specifically purifies peptides. Peptide purification was performed using the following steps.[1]

1. A new Pierce C18 spin column was placed in a 1.5 mL Eppendorf tube.
2. The column was incubated with 200 µL of 100% ACN for 5 minutes to remove unbound proteins.
3. The tube was centrifuged for 2 minutes on a tabletop centrifuge; the flowthrough was discarded.
4. To activate the column, 200 µL of wash buffer, i.e., 50% ACN+50% water, were added.
5. The mixture was incubated for 5 minutes and then centrifuged for 2 minutes on a tabletop centrifuge; the flowthrough was discarded.
6. 200 µL of column equilibration buffer were added and incubated for 2 minutes, and the flowthrough was discarded.
7. 50 µg of digested protein were added and incubated for 5 minutes.
8. The flowthrough was collected, and the digested sample was loaded again.
9. Step 8 was repeated 4 times to ensure proper bind-

 https://doi.org/10.3350/cmh.2024.0554

ing of the peptides in the column.
10. The column was centrifuged for 2 minutes on a tabletop centrifuge, and the flowthrough was discarded.
11. 200 µL of equilibration buffer (5% ACN+95% water) were added and incubated for 2 minutes.
12. The column was centrifuged for 2 minutes on a tabletop centrifuge, and the flowthrough was discarded.
13. For elution, 30 µL of elution buffer (90% ACN+10% water) were added and incubated for 2 minutes.
14. An additional 30 µL of elution buffer (90% ACN+10% water) was added and incubated for 2 minutes.
15. 40 µL of elution buffer were added, and the column was centrifuged for 2 minutes on a tabletop centrifuge.
16. The flowthrough was collected, and the used column was discarded.
17. The samples were lyophilized at 4°C and 60 millibar pressure until the sample was dry.
18. The samples were reconstituted in 40 µL of 0.1% (v/v) formic acid.
19. The samples were centrifuged at 15,000 g, and 20 µL of the supernatant was loaded in an HPLC vial.

Mass spectrometry analysis of fungal-peptides (fungome)

In brief, proteins were isolated from the plasma samples of the study groups. The isolated proteins were reduced, alkylated, and digested using trypsin, followed by mass spectrometry analysis similar to that described in the proteome analysis. The MS/MS data were acquired and analyzed using Proteome Discoverer (version 2.0; Thermo Fisher Scientific, Waltham, MA, USA), and the fungal sequences were reported in FungiProteomeDB, an online database (https://vision4research.com/fungidb/). We analyzed the proteomes of 685 fungal species, which contain 7,127,141 protein sequences. We cross-validated our findings using the Mascot algorithm (Mascot 2.4, Matrix Science; http://www.matrixscience.com) to find all possible microbial species. In brief, significant peptide groups were identified at $P<0.05$, q values were identified at $P<0.05$, and the false discovery rate was set at 0.01. Only rank-1 peptides with sequence match scores >3 were subjected to biodiversity and functional analyses using Unipept. Peptides that mapped to the eukaryotic, bacterial, or viral databases were rejected. Only peptides associated with fungal species were segregated and subjected to statistical, functional, and biodiversity analyses, as detailed in our recent publication.[1,2,12]

Lipidomics

Sample preparation for lipidomics

Total lipids were extracted from 100 µL of the low-abundance fraction (ALF-NS=32, ALF-S=8, HC=5) samples using the Folch method.[13]

1. Chloroform/methanol was added at 2:1 (v/v) and held for 24 hours at –20°C.
2. The samples were centrifuged at 13,000 rpm for 10 minutes.
3. The upper phase (aqueous phase), containing different lipid species, was transferred into a new centrifuge, the interphase which consists of a protein disk was discarded and the lower rich-lipid phase (organic phase) was pooled with the upper phase and vacuum dried.
4. The dried samples were reconstituted in 100 µL of 65:30:5:5 standard solution (65% acetonitrile: 30% isopropanol: 5% water: 5% internal and external standards).
5. 80 µL of the sample solution were pipetted into an HPLC tube.
6. The remaining 25 µL were placed in a different tube to prepare a pooled quality control sample.

[Internal standards (Dinoseb: 1 mg/mL; MCPA: 1 mg/mL; Dimetrazole: 1 mg/mL) and External standards (Cholesterol: 0.1 mg/mL; Colchicine: 4 mg/ml; Impramine: 2 mg/mL; Roxithromycin: 2 mg/mL; Amiloride: 1 mg/mL; Atropine: 2 mg/mL)].

Mass spectrometry analysis for lipidomics

We subjected 15-µL samples to reverse-phase chromatography in a C18 column (Thermo Scientific?25003102130: 3 µm, 2.1 mm, 100 mm) and tandem mass spectrometry (MS/MS). Mobile phase A was 0.1% formic acid, and mobile phase B was 100% acetonitrile.[2] In brief, the samples were passed through the column into the heated electrospray source of a Q-Exactive mass spectrometer (Thermo Scientific, San Jose, CA, USA), and they were analyzed in both positive and negative ionization modes. The spray voltage was 3.7 kV in positive mode and –3.1 kV in nega-

tive ionization mode. The heated capillary temperature was maintained at 360°C, and the sheath and auxiliary gas flows were set to 15 and 10 arbitrary units, respectively. For the MS/MS mode, the isolation width was set to m/z 1.5, the normalized collision energy was 32%, and the mass resolution was 17,500 FWHM at m/z 200.

Lipid features were identified using lipidSearch 4.0 software (Thermo Fisher Scientific, Waltham, MA, USA). The following feature identification and quantitation parameters were used: Search type, Product ion search; Exp type, LC-MS; Precursor tolerance, ~5–7 ppm; Product tolerance, 7–10 ppm; Intensity threshold, 1.00%; Height-score threshold, 8; Execute quantitation, On; M/z tolerance (quantitation), –/+ 5.0 ppm; RT range (min.) (quantitation), –/+ 0.5; Top rank filter, On; Main node filter, Main isomer peak; m-Score threshold, 5; c-Score threshold, 2; FA priority, On; ID quality filter, Check A, B, and C or A–D for lower quality IDs; Target class, Select ALL lipid classes; Ion adducts (pos), +H, $+NH_4$, and +2 H; Ion adducts (neg), –H, +HCOO, and –2H.[2]

Calculation of Shannon and Simpson (α-diversity) indices for lipids and fungal peptides

In this study, the Shannon and Simpson indices (α-diversity) for ALF patients were assessed using an R script with the vegan package.

The WLCNA and WFCNA were performed for identified lipids using the WGCNA R package in Perseus software. In these analyses, each correlation matrix between the lipids or fungal-peptides in the dataset describes a fully connected, weighted network in which the weight on each edge denotes the correlations among the quantitative profiles of the lipids or fungal-proteins.[14] As settings, a soft threshold of 9, signed network, and bi-weight mid-correlation (a robust alternative to Pearson correlation) were chosen to calculate the correlations among all lipid signatures. To obtain a scale-free co-expression network, a power parameter of 12 was selected and led to an approximately scale-free network with a scale-free topology of 0.85. Hierarchical clustering of the co-expression network identified 19 and 2 modules for the identified lipids and fungal-proteins, respectively.

Correlation network analysis

MetaboAnalyst 5.0 (www.metaboanalyst.ca) was used for correlation clustering between the fungal clusters and lipid clusters, and Cytoscape-MetaScape 3.9.1. was used to establish the correlation network map among the members of each identified cluster. Correlation analysis between each module and the clinical parameters was performed using R.

Probability of death (POD) calculation

In the current study, POD among ALF patients was assessed to find indicators of non-survival. The POD was calculated using the top fungal-peptides and lipids, as identified based on their P-values, AUCs, and MDAs in each comparison (ALF vs. HC and ALF-NS vs. ALF-S) using SPSS (version 20; IBM Co., Armonk, NY, USA).

Statistical analysis

Results are shown as mean and standard deviation unless indicated otherwise. Statistical analyses were performed using GraphPad Prism (version 6; GraphPad Software Inc., La Jolla, CA, USA) and SPSS, version 20. *P*-values <0.05 using Benjamini-Hochberg correction were considered statistically significant. The unpaired (two-tail) Student's *t*-test and Mann-Whitney U-test were performed to compare two groups. For comparisons of more than two groups, a one-way analysis of variance or the Kruskal-Wallis test was performed. All correlations were calculated using Spearman correlation analyses, with $r^2>0.5$, $P<0.05$ considered to be statistically significant. Annotated features were subjected to statistical software platforms. First, missing value imputation was applied to the data, with half the minimum positive value estimated for fungal peptides that were undetected in the samples. Subsequently, data were filtered based on the non-parametric relative standard deviation (MAD/median) and subjected to log normalization and Pareto-scaling using MetaboAnalyst 5.0 (http://metaboanalyst.ca.). Inter-batch normalization was not performed. PCA and PLS-DA, heatmaps, RF, and other statistical analyses were performed. To determine the best putative indicator among the top 5 lipids and fungal-peptides, univariate and multivariate AUROC analyses were performed in SPSS, and hazard ratios were calculated. Addi-

 https://doi.org/10.3350/cmh.2024.0554

tionally, an ANN was used, and a survival analysis was performed using the Kaplan-Meier curve in SPSS on the basis of a panel of 5 lipids.

Machine learning

Approach[15] was used on the results of the discovery and validation cohort analyses. In brief, all ALF patients (270) were randomized into a 70% training cohort and 30% test cohort, followed by analysis using five ML algorithms (LDA, RF, SVM, kNN, CART) to validate the plausible biomarker lipid species. In total, we implemented 30 ML models within the 5 ML algorithms: for the 5 lipid species individually and one combined model. Fourfold (outer) nested repeated (five times) tenfold (inner) cross-validation (with randomized stratified splitting) was conducted with both the training and test cohorts in R with the Caret package. In this way, repeated tenfold cross-validation was performed 20 times, which allowed the models to obtain the best results. The overall cross-validation prediction performance was summarized using the accuracy, sensitivity, specificity, and -log10p performance measures. The equations used to quantify those performance measures are presented below, in which TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives.

Accuracy=TP+TN/ TP+TN+FP+FN

Sensitivity=TP/TP+FN

Specificity=TN/TN+FP

Artificial neural network

The ANN model described previously[16,17] is a multilayer perceptron. It was used here to develop a predictive model based on the top 5 lipid species, fungal peptides, and clinical parameters within the neural network module of IBM SPSS Statistics.

In brief, all ALF patients (270) were randomized into a 70% training cohort and a 30% test cohort and then analyzed using the ANN model. The architecture for the network model follows a feedforward method with three layers: input layer, hidden layer, and output layer. The network structure included three inputs (top 5 lipid species, fungal peptides, and clinical parameters), and the hidden layer comprised 3 units. All 20 covariates were normalized. Each node is a neuron that uses a hyperbolic tangent activation function and cross entropy error function. The default parameters in IBM SPSS were applied for the remaining settings. Training iterations were repeated until the synaptic weights yielded predictions with minimal deviation from the actual values. Following training, the ANN model validity was assessed by applying the generated .xml file to a separate test set consisting of 230 samples.[16,17] The last output layer classifies mortality status from our data.

REFERENCES

1. Tripathi G, Sharma N, Bindal V, Yadav M, Mathew B, Sharma S, et al. Protocol for global proteome, virome, and metaproteome profiling of respiratory specimen (VTM) in COVID-19 patient by LC-MS/MS-based analysis. STAR Protoc 2022;3:101045.
2. Sharma N, Yadav M, Tripathi G, Mathew B, Bindal V, Falari S, et al. Bile multi-omics analysis classifies lipid species and microbial peptides predictive of carcinoma of gallbladder. Hepatology 2022;76:920-935.
3. Sharma N, Bhat SH, Tripathi G, Yadav M, Mathew B, Bindal V, et al. Global metabolome profiling of COVID-19 respiratory specimen using high-resolution mass spectrometry (HRMS). STAR Protoc 2022;3:101051.
4. Sharma N, Pandey S, Yadav M, Mathew B, Bindal V, Sharma N, et al. Biomolecular map of albumin identifies signatures of severity and early mortality in acute liver failure. J Hepatol 2023;79:677-691.
5. Subramanian S. Dye-ligand affinity chromatography: the interaction of Cibacron Blue F3GA with proteins and enzymes. CRC Crit Rev Biochem 1984;16:169-205.
6. Naval J, Calvo M, Lampreave F, Pineiro A. Interactions of different albumins and animal sera with insolubilized Cibacron Blue, Evaluation of apparent affinity constants. Comp Biochem Physiol B 1982;71:403-407.
7. Leatherbarrow RJ, Dean PD. Studies on the mechanism of binding of serum albumins to immobilized cibacron blue F3G A. Biochem J 1980;189:27-34.
8. Andac CA, Andac M, Denizli A. Predicting the binding properties of cibacron blue F3GA in affinity separation systems. Int J Biol Macromol 2007;41:430-438.
9. Bhat A, Das S, Yadav G, Chaudhary S, Vyas A, Islam M, et al. Hyperoxidized albumin modulates platelets and promotes inflammation through CD36 receptor in severe alcoholic hepa-

titis. Hepatol Commun 2019;4:50-65.
10. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976;72:248-254.
11. Gundry RL, White MY, Murray CI, Kane LA, Fu Q, Stanley BA, et al. Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow. Curr Protoc Mol Biol 2009;Chapter 10:Unit10.25.
12. Maras JS, Sharma S, Bhat A, Rooge S, Aggrawal R, Gupta E, et al. Multi-omics analysis of respiratory specimen characterizes baseline molecular determinants associated with SARS-CoV-2 outcome. iScience 2021;24:102823.
13. Folch J, Lees M, Sloane Stanley GH. A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem 1957;226:497-509.
14. Rudolph JD, Cox J. A network module for the Perseus software for computational proteomics facilitates proteome interaction graph analysis. J Proteome Res 2019;18:2052-2064.
15. Nachtigall FM, Pereira A, Trofymchuk OS, Santos LS. Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. Nat Biotechnol 2020;38:1168-1173.
16. Thong Z, Tan JYY, Loo ES , Phua YW, Chan XLS, Syn CKC. Artificial neural network, predictor variables and sensitivity threshold for DNA methylation-based age prediction using blood samples. Sci Rep 2021;11:1744.
17. Hong WD, Chen XR, Jin SQ, Huang QK, Zhu QH, Pan JY. Use of an artificial neural network to predict persistent organ failure in patients with acute pancreatitis. Clinics (Sao Paulo) 2013;68:27-31.