nature portfolio

Corresponding author(s): Dr. Jaswinder Singh Maras

Last updated by author(s): Apr 6, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The commercial software Compound Discoverer3.0 for plasma metabolomics ; Proteome Discoverer (version 2.0) and unipept 5.0.8 for plasma metaproteomics data was used to perform protein identification and quantification.
Data analysis	Statistical analysis was performed using Graph Pad Prism v6, SPSS V20, and P-values of < 0.05 using Benjamini & Hochberg correction were considered statistically significant. Unpaired (two-tail) Student's t-test, and Mann-Whitney U test were performed for comparison of two groups. For comparison among more than two groups, a one-way analysis of variance, the Kruskal-Walli's test was performed. Annotated features were subjected to different statistical software platforms. First, missing value imputation was applied to data in which half the minimum positive value was estimated for meta-proteome that were undetected in the samples. Subsequently, data were filtered based on non-parametric relative standard deviation (MAD/median) and were subjected to log normalization and Pareto-scaling using Metaboanalyst 5.0 (http://metaboanalyst.ca.) server. PCA and PLS-DA, Heat map, Random-forest analysis, and other statistical analyses were performed for metabolomics and MicrobiomeAnalyst used for metaproteomics. MetaboAnalyst 5.0 was used for correlation clustering between meta-protein clusters bacterial phyla and metabolite pathways and Cytoscape-MetaScape 3.9.1. was used to establish the correlation network map between members of identified Cluster. Correlation analysis between each module and clinical parameters was performed using R or iMAP (https://imap.metaboprofile.cloud/ (License MPL 2.0)). We also performed WMpCNA and WMCNA on the identified Metabolites or meta-proteins using 'WGCNA' R package in Perseus software (version 1.6.14.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

For Meta-proteome analysis, MS/MS data were acquired using the plasma samples and analyzed by Proteome Discoverer (version 2.0, Thermo Fisher Scientific, Waltham, MA, United States) using the bacterial/fungal sequence (UniprotSwP_20170609 with sequences 467231 and MG_BG_UPSP with sequences 2019194). This was cross-validated using the Mascot algorithm (Mascot 2.4, Matrix Science). The identified tryptic peptide was crossed searched on the online unipept software https://unipept.ugent.be/. specifically for determination of the microbial taxa linked to the tryptic peptides . Finally the identified tryptic peptides along with the microbial taxa details was subjected to statical analysis between the study groups involved in the study. All results from statistical and bioinformatics analysis is provided in Supplementary Tables and Figures and the associated RAW data (run files) are available on request to the corresponding author : jassi2param@gmail.com
For metabolomics estimation we used Compound Discoverer 3.0 to identify metabolite features (ThermoFisher Scientific, Waltham, United States). The features were annotated using mass list, ChemSpider, mzVault, Metabolika, and mzCloud™ for metabolomics. All results from statistical and bioinformatics analysis utilizing the metabolomics data is provided in Supplementary Tables and Figures and the Raw Data for the metabolomics runs are available on request to the corresponding author : jassi2param@gmail.com
The Study protocol and Standard operating procedure and patient information is also available upon request.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	For training Cohort, (H=5;ALF=40), baseline plasma samples were collected as follows: Acute Liver Faliure-survivor Characterstics (n=8) : Male=07; Female=01, Median Age -23.5 (19-42), median MELD Score-25.5(14-29); Acute Liver Faliure-NON-survivor Characterstics (n=32) : Male=15; Female=17, median Age -32 (20-62), and median MELD Score- 40(36-41); Healthy Control Characterstics (n=5)- Male=03; Female=02, median Age 26 (22-50). For validation Cohort, (H=20;ALF=160), baseline plasma samples were collected as follows: Acute Liver Faliure-survivor Characterstics (n=53) : Male=30; Female=23, Median Age -29 (19-58), median MELD Score-26(12-54); Acute Liver Faliure-NON-survivor Characterstics (n=107) : Male=75; Female=32, median Age -30 (19-62), and median MELD Score- 40(37-49); Healthy Control Characterstics (n=20)- Male=11; Female=09, median Age 23 (19-50). The demographic profile of both the training and test cohorts were comparable with significant increase in liver function parameters (AST, INR, Bilirubin, ALP and ALT) in ALF more so in the Non-survivors (ALF-NS) respectively. Severity indices such as MELD was high in ALF-NS as compared to ALF-S ($p<0.05$). Infection at baseline found to be higher in ALF-NS (>40%) as compare ALF-S. During ICU stay ALF patients develop different complications, such as hepatic encephalopathy, acute kidney injury, sepsis, necrosis and multi organ failure are also considered as death due to acute liver failure.
Population characteristics	In this cross-sectional study, a total of 200-ALF patients were prospectively enrolled (December 2019 to December 2022) at the Institute of Liver and Biliary Sciences, New Delhi, India. Twenty-five healthy controls-(HC) were also taken. The diagnosis of acute liver failure was based on the presence of jaundice with hepatic encephalopathy within 4 weeks with laboratory evidence of increased INR>1.5. Patients with underlying features of chronic liver disease, such as splenomegaly, clinical ascites or known hepatitis B or C infection or regular alcohol intake, were excluded. ALF patients were clinically followed up till death or up to a period of 90 days. Death due to liver failure was considered if a patient died due to progressive liver failure or related complications, such as cerebral edema, renal failure, infection, sepsis or multi-organ failure during the ICU stay or within 30 days. Many patients require tracheostomy and have a rather slow recovery when shifted to HDU. Hence, follow-up of ALF patients was done to record survival outcomes. These patients were categorized as ALF non-survivors-(ALF-NS). The patients who improved and survived were shifted out of the ICU and a follow-up of 90 days was maintained and were categorized as ALF survivors-(ALF-S).
Recruitment	In this cross-sectional study, a total of 200 ALF patients along with 25 healthy controls were prospectively enrolled (December 2019 to December 2022) at the Institute of Liver and Biliary Sciences, New Delhi, India. Acute liver failure was diagnosed as; presence of jaundice with hepatic encephalopathy within 4 weeks with laboratory evidence of increased INR (>1.5); presence of infection(6). Patients with underlying features of chronic liver disease, such as splenomegaly, clinical ascites or known hepatitis B or C infection or regular alcohol intake, were excluded from the study.
Ethics oversight	All procedures involved in the study were conducted as per the Helsinki declaration. All procedures involved in the study were conducted in accordance with the institute ethical committees (IRB approval no-IEC/2019/70/NA06), and written informed consent was obtained from all subjects enrolled in this study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	While designing the study protocol (prior to recruiting patients), we planned to work with multiomics data sets using system biology. consequently, the sample size was based on experience from prior omics studies in humans. In this cross-sectional study, a total of 200 ALF patients along with 25 healthy controls were prospectively enrolled (December 2019 to December 2022) at the Institute of Liver and Biliary Sciences, New Delhi, India.
Data exclusions	Patients with underlying features of chronic liver disease, such as splenomegaly, clinical ascites or known hepatitis B or C infection or regular alcohol intake, were excluded from the study.
Replication	Metabolomics and Meta-proteomics analysis was performed on 45 biological replicates (training cohort; 40 ALF, 5 HC) these were further validated on 180 biological replicates (test cohort; 160 ALF and 20 HC). In addition each metabolomic run was spiked with internal and external standards (detailed in the supplementary section of the manuscript) in order to assess the analytical stability of the runs. We also performed a MS/ MS analysis of the biological pool samples in variable dilution to (1:1,1:2,1:4,1:8) which was used to estimate linear curve for individual metabolites.
Randomization	In this cross-sectional study, the first 40 ALF samples was characterized as training cohort and the rest 160 samples (collected prospectively) were characterized as the test cohort. The study aims to identify metabolites and bacterial peptides that corresponds to poor outcome in ALF. In the downstream statistical analysis patients were grouped on the basis of survival and non-survival. Acquisition of metabolomics amd meta-proteome data was randomized to avoid systemic bias during this measurement.
Blinding	The author collected blood samples from patients were blinded on the parameter of ALF-survivor and ALF-non-survivor at baseline. The investigators who performed metabolomics and metaproteomics sample and data acquisition were not blinded to clinical patients parameters as randomization in metabolomics and meta-proteomics data generation was anyway performed to avoid systemic bias during measurement and blinding is thus not relevant.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NA
Study protocol	NA
Data collection	200-ALF patients were prospectively enrolled (December 2019 to December 2022) at the Institute of Liver and Biliary Sciences, New Delhi, India.
Outcomes	Primary outcome of this study include metabolome and meta-proteome whose abundance level in plasma correlates with poor

Outcomes

outome in ALF. On the basis of statistical and differential analysis, we identified top 5 metabolite increased in ALF-NS such as L-Tyrosine, 4-(2-Amino phenyl)-2,4-dioxobutanoate, Chenodeoxycholic acid (linked to cell death and inflammation), Carnosine and alanyl-tyrosine (linked to negate oxidative stress) was directly correlated with clinical parameters (R2>0.85).
Secondary outcome of this study is validation of these metabolites on the basis of POD of top 5 metabolites, which showed a diagnostic efficiency of 98% (AUC=0.98(0.92-1.0)) for mortality and validated in the validation cohort using five machine learning algorithms showed >98% accuracy/sensitivity/specificity for prediction of early mortality. This can enhance our understanding of microbiome and metabolome changes in ALF-NS and also help in designing diagnostic/therapeutic approach for the better clinical treatment.
In this pilot study we for the first time found out heterogenous changes in baseline metabolite and meta-proteome in Acute liver failure patients specially who are predisposed to early mortality.