



Metaproteomics reveals age-specific alterations of gut microbiome in hamsters with SARS-CoV-2 infection

Marybeth Creskey, Fabiola Silva Angulo, Qing Wu, Levi Tamming, Emily E. F. Fekete, Kai Cheng, Zhibin Ning, Angela Wang, Patrícia Brito Rodrigues, Vinícius de Rezende Rodovalho, Marco Aurélio Ramirez Vinolo, Daniel Figeys, Xuguang Li, Francois Trottein & Xu Zhang

To cite this article: Marybeth Creskey, Fabiola Silva Angulo, Qing Wu, Levi Tamming, Emily E. F. Fekete, Kai Cheng, Zhibin Ning, Angela Wang, Patrícia Brito Rodrigues, Vinícius de Rezende Rodovalho, Marco Aurélio Ramirez Vinolo, Daniel Figeys, Xuguang Li, Francois Trottein & Xu Zhang (2025) Metaproteomics reveals age-specific alterations of gut microbiome in hamsters with SARS-CoV-2 infection, Gut Microbes, 17:1, 2505117, DOI: [10.1080/19490976.2025.2505117](https://doi.org/10.1080/19490976.2025.2505117)

To link to this article: <https://doi.org/10.1080/19490976.2025.2505117>



© 2025 Crown Copyright. Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 23 May 2025.



[Submit your article to this journal](#)



Article views: 874



[View related articles](#)



[View Crossmark data](#)

RESEARCH PAPER



Metaproteomics reveals age-specific alterations of gut microbiome in hamsters with SARS-CoV-2 infection

Marybeth Creskey^{a*}, Fabiola Silva Angulo^{b*}, Qing Wu^{c,d*}, Levi Tamming^{a,d}, Emily E. F. Fekete^a, Kai Cheng^{id c}, Zhibin Ning^c, Angela Wang^{a,d}, Patrícia Brito Rodrigues^{b,e}, Vinícius de Rezende Rodvalho^{id e S}, Marco Aurélio Ramirez Vinolo^e, Daniel Figeys^{id c}, Xuguang Li^{a,d}, Francois Trottein^{id b}, and Xu Zhang^{id a,c}

^aRegulatory Research Division, Biologic and Radiopharmaceutical Drugs Directorate, Health Products and Food Branch, Health Canada, Ottawa, Canada; ^bUniversity of Lille, CNRS, INSERM, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 9017 - CIIL - Center for Infection and Immunity of Lille, Lille, France; ^cSchool of Pharmaceutical Sciences, Faculty of Medicine, University of Ottawa, Ottawa, Canada; ^dDepartment of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Canada; ^eLaboratory of Immunoinflammation, Institute of Biology, University of Campinas (UNICAMP), Campinas, Brazil

ABSTRACT

The gut microbiome's pivotal role in health and disease is well established. SARS-CoV-2 infection often causes gastrointestinal symptoms and is associated with changes of the microbiome in both human and animal studies. While hamsters serve as important animal models for coronavirus research, there exists a notable void in the functional characterization of their microbiomes with metaproteomics. In this study, we present a workflow for analyzing the hamster gut microbiome, including a metagenomics-derived hamster gut microbial protein database and a data-independent acquisition metaproteomics method. Using this workflow, we identified 32,419 protein groups from the fecal microbiomes of young and old hamsters infected with SARS-CoV-2. We showed age-specific changes in the expressions of microbiome functions and host proteins associated with microbiomes, providing further functional insight into the interactions between the microbiome and host in SARS-CoV-2 infection. Altogether, this study established and demonstrated the capability of metaproteomics for the study of hamster microbiomes.

ARTICLE HISTORY

Received 2 December 2024
Revised 28 March 2025
Accepted 7 May 2025

KEYWORDS

Metaproteomics; gut microbiome; data-independent acquisition; COVID-19

Introduction

The coronavirus disease 2019 (COVID-19) pandemic resulted in significant morbidity and mortality, causing severe social and economic disruptions around the world.¹ Infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is correlated with observable changes in the gut microbiome, including a reduction in microbial diversity and alterations in the relative abundance of pathogenic bacterial species.^{2–6} These microbiome alterations have been associated with disease severity and may play a role in immune dysregulation and systemic inflammation observed in COVID-19 patients.^{3,7} Functional

characterization of these taxonomic changes in the gut microbiome during SARS-CoV-2 infections could provide insights into disease mechanisms and potential microbiome-directed therapeutic strategies for disease management.

Hamsters have emerged as a prominent animal model for infectious diseases, including COVID-19, and is often preferred over mice due to several advantages, including (1) hamsters are outbred animals conferring more genetic diversity over mice, and (2) the infectious disease progression seen in hamsters is more comparable to that of humans.⁸ For coronavirus disease, the hamster angiotensin-converting enzyme 2 (ACE2), the

CONTACT Xu Zhang ✉ xu.zhang@hc-sc.gc.ca; Xuguang Li ✉ sean.li@hc-sc.gc.ca Biologic and Radiopharmaceutical Drugs Directorate, Health Products and Food Branch, Health Canada, 251 Sir Frederick Banting Driveway Ottawa, ON K1A 0K9, Canada; Francois Trottein ✉ francois.trottein@pasteur-lille.fr U1019 - UMR 9017 - CIIL - Center for Infection and Immunity of Lille, University of Lille, CNRS, INSERM, CHU Lille, Institut Pasteur de Lille, 1, rue du Professeur Calmette, Lille F-59000, France; Daniel Figeys ✉ dfigeys@uottawa.ca School of Pharmaceutical Sciences, Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H8M5, Canada

*These authors contributed equally.

^SPresent address for Dr. Vinícius de Rezende Rodvalho is Bioinformatics Laboratory (LABINFO), National Laboratory for Scientific Computing (LNCC), Petrópolis, Brazil.

[†]Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19490976.2025.2505117>

© 2025 Crown Copyright. Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

main cellular receptor mediating viral entry, binds more strongly with the spike protein than the mouse ACE2, which is consistent with observations that the hamster experiences mild to severe disease with quantifiable clinical signs, weight loss, viral shedding, and lung pathology.^{9,10} Notably, recent studies have shown that the SARS-CoV-2 infection in hamsters is associated with alteration of gut microbial composition, similar to observations in human studies. Sencio et al. reported alteration of hamster gut microbiota composition along with SARS-CoV-2 infection in young-adult animals.¹¹ The infection and disease severity were associated with an increase in opportunistic pathogens such as *Enterobacteriaceae* and *Desulfovibrionaceae*, and a decrease in short-chain fatty acid (SCFA) producing bacteria such as *Ruminococcaceae* and *Lachnospiraceae*. On the other hand, Seibert et al. showed that the middle-aged hamster microbiome infected with SARS-CoV-2 shares some similarities with that of critically ill COVID-19 patients.¹² A further study in high fat/high cholesterol diet induced obese nonalcoholic steatohepatitis (NASH) hamsters demonstrated that more severe disease activity was developed following SARS-CoV-2 infection in NASH hamsters with dysbiosis.¹³ In the meantime, metabolomic analyses have also demonstrated the critical role of microbiota metabolites such as SCFA and deoxycholic acid in host resistance to SARS-CoV-2 infection and further host immune responses to infection, impacting disease severity.^{14,15}

However, despite its widespread use in hamster animal models, there remains a notable gap in the study of hamster gut microbiome with functional meta-omics methods, including metagenomics and metaproteomics. The majority of the current hamster gut microbiome studies were performed using the 16S rDNA amplicon sequencing approach. Currently, there are no published studies utilizing metaproteomics techniques in hamsters, and there are no hamster gut microbial gene/protein databases that are needed for metaproteomic identifications. To fill this knowledge gap, we constructed, evaluated, and made publicly available hamster gut microbiome protein databases by combining both an in-house and a previously published shotgun metagenomic sequencing dataset. A remarkable advance of quantitative proteomics, including

metaproteomics, in recent years is the wide application of data-independent acquisition mass spectrometry (DIA-MS).^{16–19} Therefore, this study utilized the DIA method, incorporating parallel accumulation serial fragmentation (PASEF),²⁰ to develop a DIA-PASEF metaproteomics workflow for studying the hamster microbiome. We established a two-stage PASEF metaproteomic workflow, including a first stage data-dependent acquisition (DDA) PASEF analysis of pre-fractionated pooled samples for generating a tailored spectral library/database, and a second stage of DIA-PASEF analysis for efficiently identifying and quantifying both the gut microbial and host proteins in feces of hamsters.

As a proof of concept for the methodology, we analyzed stools from young (2-month old) and old (22-month old) male Syrian golden hamsters infected with SARS-CoV-2 followed by sampling at day 0, 7, 15, 30, and 45. The established workflow enables in-depth profiling of the fecal microbial proteins, taxonomic compositions, and functions of the hamster fecal microbiomes, and reveals distinct clustering patterns at all feature levels for old hamsters at day 7 after SARS-CoV-2 infection. The hamster gut microbial protein and spectral library databases, the DIA-PASEF metaproteomic workflow, and the deep metaproteomic dataset of SARS-CoV-2 infection provided in this study altogether are valuable resources for the microbiome study in hamster disease models.

Results

Constructing hamster gut microbial protein databases for metaproteomics

A protein sequence database is needed for shotgun proteomic or metaproteomic study to efficiently interpret mass spectra through peptide-spectrum matching (PSM). For the study of gut microbiomes, extensive metagenomic sequencing has been used to generate gut microbial gene catalogs or genome databases in hosts, such as humans and mice.^{21–23} Previous study has shown limited overlap between the human and mouse gut microbial gene catalogs,²² suggesting a need for host-specific gut microbial gene/protein databases for metaproteomic identification. Since there is currently no

available hamster gut microbial gene catalog or genome database, we generated an in-house shotgun metagenomic sequencing dataset derived from three young (2-month old) and six old hamsters (22-month old), consisting of 434 million high-quality sequencing reads. This independent cohort of animals and the metagenomic sequencing data are from our companion study²⁴ and used for the establishment of the reference protein database only. We used a previously established SqueezeMeta (v1.6.4) workflow²⁵ for processing the metagenomic sequencing data. Briefly, high-quality sequencing reads were first assembled using Megahit,²⁶ and open reading frame (ORF) or genes were predicted using Prodigal²⁷ to generate a gene catalog database MGDB-V1, consisting of 1,730,340 genes or translated protein sequences (Figure 1(a); details in method section). Taxonomic annotation and quantitation of the shotgun metagenomic sequencing data showed that, similar to human and mouse gut microbiota,

Bacillota and Bacteroidota were the most abundant phyla in hamster gut microbiota. Additionally, *Thermodesulfobacteriota*, *Pseudomonadota*, *Deferribacterota*, and *Actinomycetota* were among the top abundant phyla. The most abundant families include *Oscillospiraceae*, *Lachnospiraceae*, *Muribaculaceae*, *Rikenellaceae*, *Bacteroidaceae*, *Prevotellaceae*, and *Desulfovibrionaceae* (Supplementary Table S1).

Usually, the gene catalog database is sufficient for metaproteomics identification, however due to the limited sequencing depth of a single metagenomic dataset, the coverage of the gene catalog database may be suboptimal. To overcome this limitation, we applied two strategies to improve the established hamster gut microbial protein database. In the first method, we augmented the MGDB-V1 with downloaded UniprotKB proteome databases for the abundant microbial species based on metagenomics data (MGtax1000 for species abundance >0.1%, or MGtax10000 for abundance

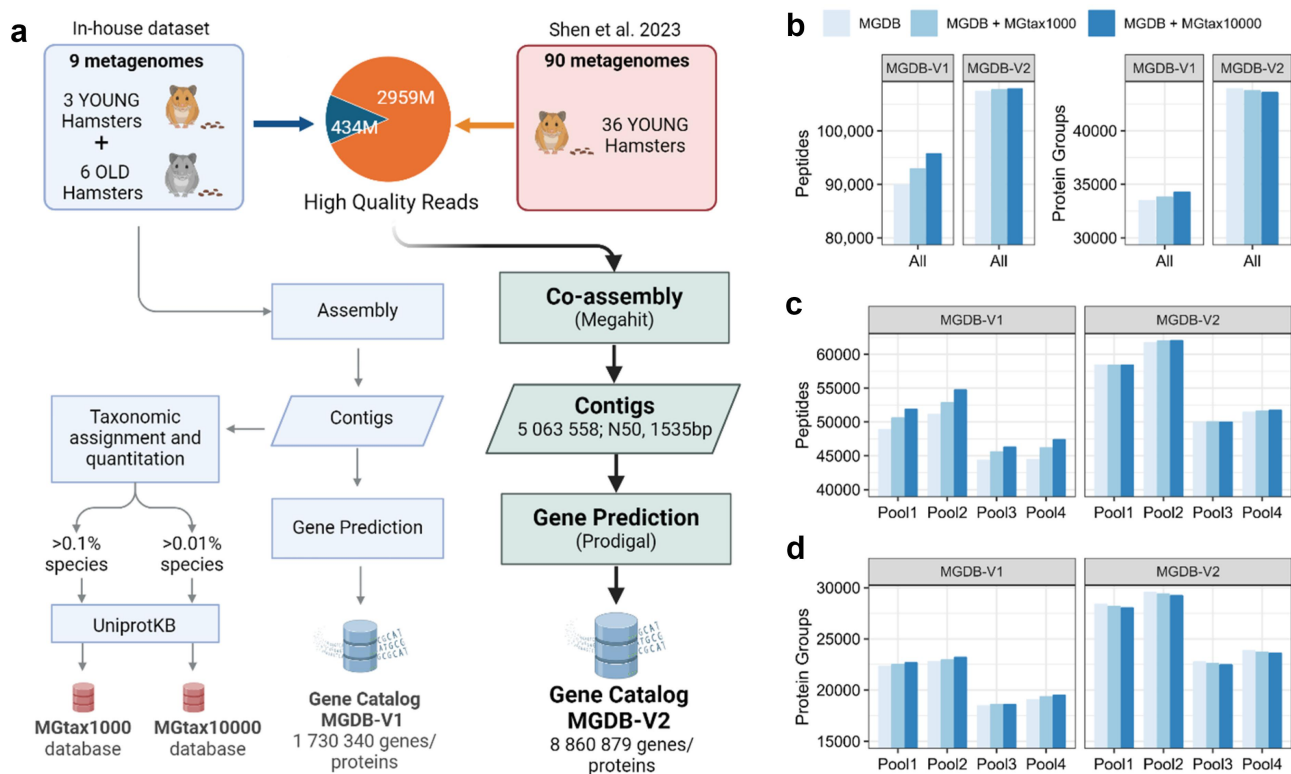


Figure 1. Hamster gut microbial gene catalog database construction and metaproteomic evaluation. (a) Gene catalog construction workflow using either in-house dataset only (MGDB-V1) or co-assembly (MGDB-V2) with a previous dataset by Shen et al. 2023.²⁸ metagenomics based taxonomy derived proteome databases (MGtax) were established based on high abundant species and UniprotKB reference protein databases. (b-d) Different databases were used for search for four pooled sample metaproteomic data generated in this study. The numbers of identified peptides and protein groups for either all samples (b) or individual samples (c-d) are shown.

>0.01% according to contig taxonomic annotation and sum abundances for all samples in the in-house young/old hamster cohort; Figure 1(a) and Supplementary Table S2). In the second method, we downloaded a previously published dataset consisting of 90 metagenomes with 2959 million high-quality sequencing reads from 30 young hamsters. To the best of our knowledge, this dataset is the only shotgun metagenomics dataset for hamster gut microbiome at the time of this project. We therefore performed a co-assembly with both datasets to generate contigs and predict genes (Figure 1(a)). This co-assembly generated an updated version of the hamster gut microbial gene catalog database (MGDB-V2) with 8,860,879 proteins, over five times that of MGDB-V1.

We then used a fractionated DDA-PASEF dataset of hamster microbiome (i.e., four pooled samples with eight fractionations per sample) to evaluate the performance of the different databases (Figure 1(b)). Notable increases in protein identifications were observed when MGtax1000 or MGtax10000 were added to the MGDB-V1 database (Figure 1(c–d)). Even higher numbers of peptides and protein groups were identified using MGDB-V2 than those with MGDB-V1 with or without MGtax augmentation. No obvious benefit was observed for MGDB-V2 when augmenting with MGtax reference proteomes, indicating sufficient coverage of the database for MGDB-V2. We therefore selected MGDB-V2 without augmentation as the database for further DIA-MS metaproteomic workflow development.

DIA-PASEF metaproteomics workflow to study hamster microbiome with SARS-CoV-2 infection

In this study, we aimed to demonstrate the capability of metaproteomics by studying the impacts of SARS-CoV-2 infection over time on hamster microbiomes. To this end, hamsters were infected with the ancestral strain of SARS-CoV-2 (BetaCoV/hCoV-19_IPL_France strain; NCBI MW575140). A total of six young and six old male Syrian hamsters were included, with one old hamster dying 24 days post-infection. Both groups experienced significant weight loss, with the lowest body weight recorded on Day 8 for the Old group and between Days 6 and 8 for the Young group

(Supplementary Figure S1). While young hamsters began recovering quickly, nearly regaining their initial weight by Day 45, the older hamsters showed slower recovery and did not fully regain their initial body weight. In total, 58 fecal samples were collected spanning five time points (day 0, 7, 15, 30, 45) post infection for both groups (Figure 2). A standard fecal metaproteomic sample processing procedure was used, including a differential centrifugation to enrich microbial cells, protein extraction with sodium dodecyl sulfate (SDS) followed by SDS removal with acetone precipitation, and an insoluble trypsin digestion to obtain peptide samples for LC-MSMS analysis (Figure 2). To generate a reduced database and spectral library for DIA-PASEF data analysis, four pooled samples representing young, old, non-infected, and infected animals, respectively, were generated, fractionated, and analyzed with a DDA-PASEF mode on a timsTOF Pro2 MS. Each individual sample was analyzed with DIA-PASEF mode.

We have previously reported that database reduction with DDA data using MSFragger followed by library-free search with DIA-NN performed the best for mouse metaproteomics data.¹⁹ This finding was further validated here with hamster metaproteomics data (Supplementary Figure S2), and thereby we chose this workflow for the analysis of hamster metaproteomics data in this study. The host protein database was appended to the reduced.fasta database. In total, DIA metaproteomics identified 32,419 proteins and 128,216 peptides (Figure 2). Of these identifications, 651 were hamster proteins, and 31,768 were from the microbiome. Up to 57,852 precursors were identified per sample. Eight samples with precursor identification less than 30,000 (~50% of the maximum identification; Supplementary Figure S3) were excluded from quantification and subsequent statistical analysis. Of all the identified protein groups, 11,627 were quantified in $\geq 70\%$ samples (present in at least 35 out of the 50 samples). Non-supervised principal component analysis (PCA) shows a time-based change after SARS-CoV-2 infection in both young and old hamsters (Figure 3(a) and Supplementary Figure S4). At Day 7, a perturbation of protein expression in both young and old animals is observed, with return to baseline at Day 15 and afterward, representing an

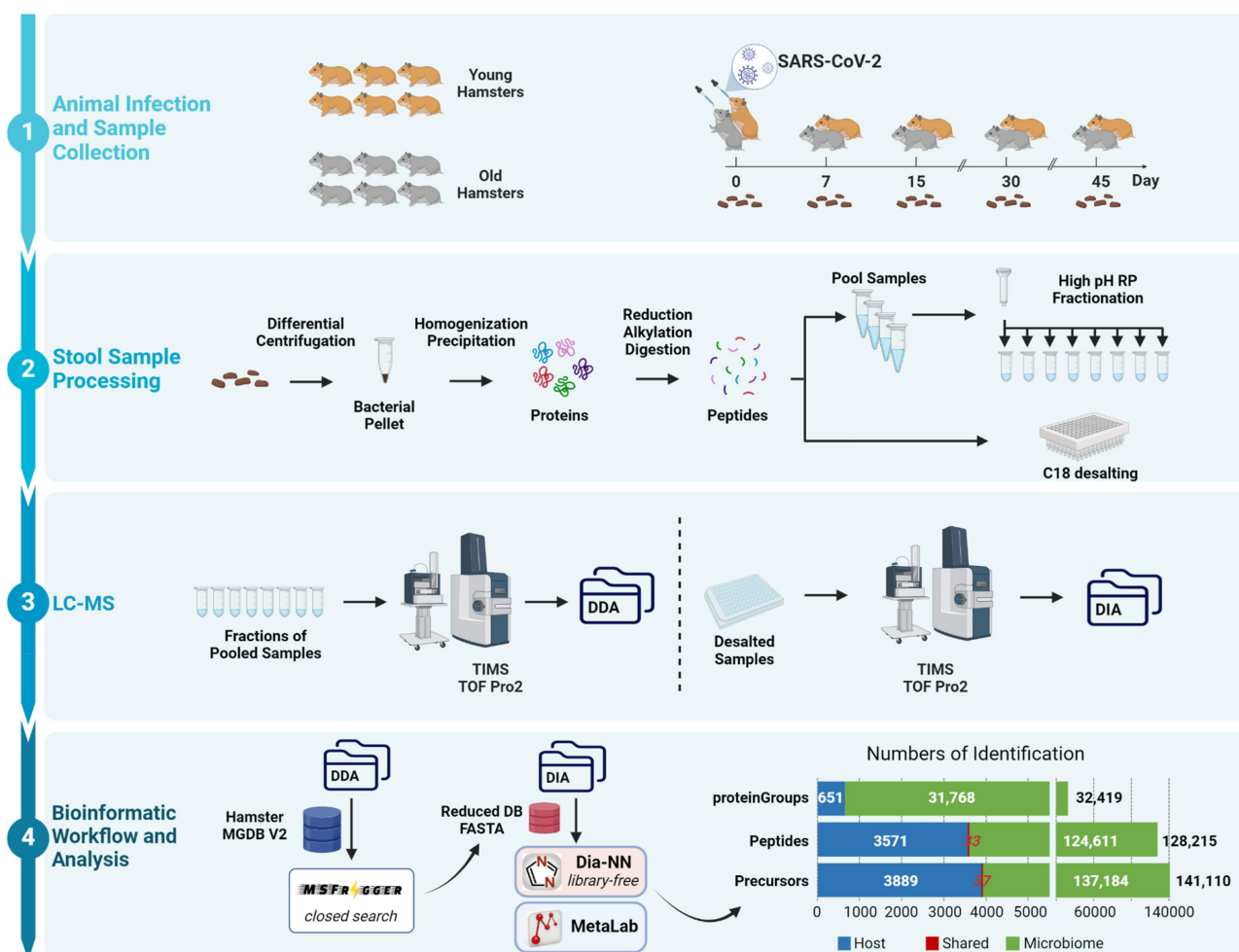


Figure 2. Experimental design and metaproteomics workflow for analyzing hamster microbiomes following SARS-CoV-2 infection. Animal experiment design, sample collection and processing steps are shown in the top panels (1 and 2). LC-MS/MS analysis and bioinformatic workflow used in this study are shown in panels 3 and 4. Numbers of precursor, peptides and corresponding protein groups identified in the whole dataset are shown as well in panel 4.

acute phase at Day 7 and recovery phase after Day 15 (Figure 3(a)). When analyzing host and microbiome proteins separately, obvious changes in host proteins were observed only in old hamsters at Day 7 following SARS-CoV-2 infection (Figure 3(b)). In contrast, marked impacts on microbiome proteins were observed in both young and old hamsters (Figure 3(c)). Interestingly, the old hamsters presented more extensive changes toward the same direction (2nd principal component in PCA score plot) with young hamsters for the microbiome protein expressions. These metaproteomic observations align well with the increased disease severity and more extensive gut microbial changes often observed in old hamsters compared to their young counterparts.^{24,29,30}

SARS-CoV-2 infection induced alteration of fecal host proteins in old hamsters

Metaproteomics has the advantage of quantifying both host and microbiome proteins at the same time. In this study, although a differential centrifugation process was included in the sample preparation to remove debris, large host cells, and proteins in the supernatant, we still identified 651 host proteins. These proteins may represent secreted host proteins closely associated with the microbial cell surface. To further examine the change of host proteins with SARS-CoV-2 infection, the relative abundance of host proteins for each sample was examined, by summing up the intensities of all hamster-derived proteins as a percentage of the total protein intensity. For young hamsters,

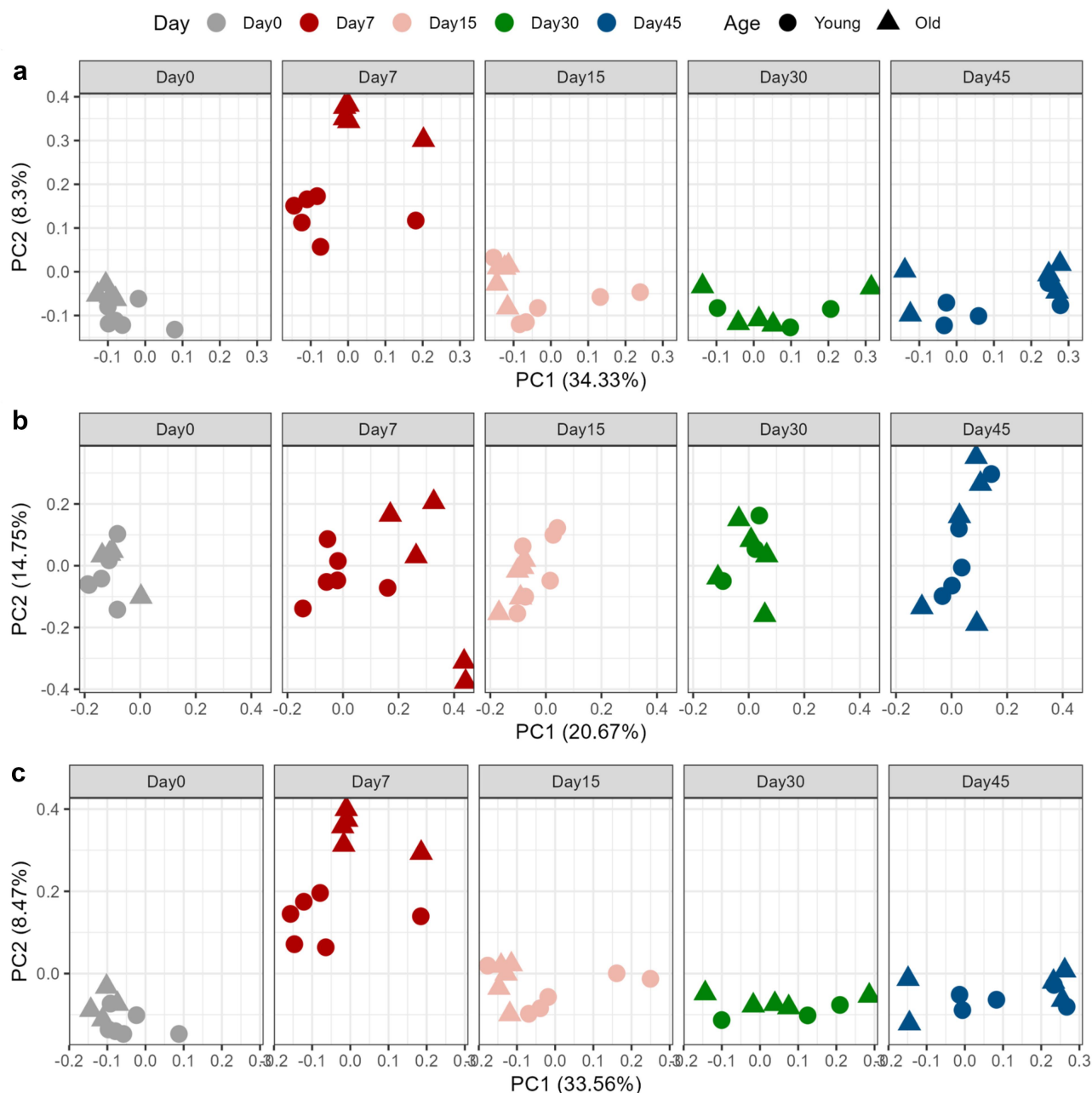


Figure 3. PCA score plots of proteins quantified in hamster feces following SARS-CoV-2 infection. PCA was performed using the normalized and log2-transformed protein intensities for all proteins (a), host proteins only (b) and microbiome proteins only (c), respectively. Only proteins that were quantified in >70% of the samples were used for analysis and, to maintain stringent quality control, 8 samples with precursor identification less than 50% of the maximum identification number were filtered out. The numbers of samples included for plotting are as follows: day 0 (young: $n=6$; old: $n=4$), day 7 (young: $n=6$; old: $n=5$), day 15 (young: $n=6$; old: $n=5$), day 30 (young: $n=3$; old: $n=5$), and day 45 (young: $n=5$; old: $n=5$). PCA score plots were generated with R ggplot2 with facet according to time points.

a progressive increase in host protein amount is seen over the course of the study timeline, achieving significant differences at Day 30 and 45. A significant increase in host protein amount is observed for old hamsters on Day 7, but returns to baseline (Day 0) by Day 15 (Figure 4). An

increasing trend is observed in old hamsters from Day 15 to Day 45, resembling the pattern observed in young hamsters.

We then sought to identify which host proteins were differentially regulated following SARS-CoV-2 infection. Since unsupervised PCA analysis

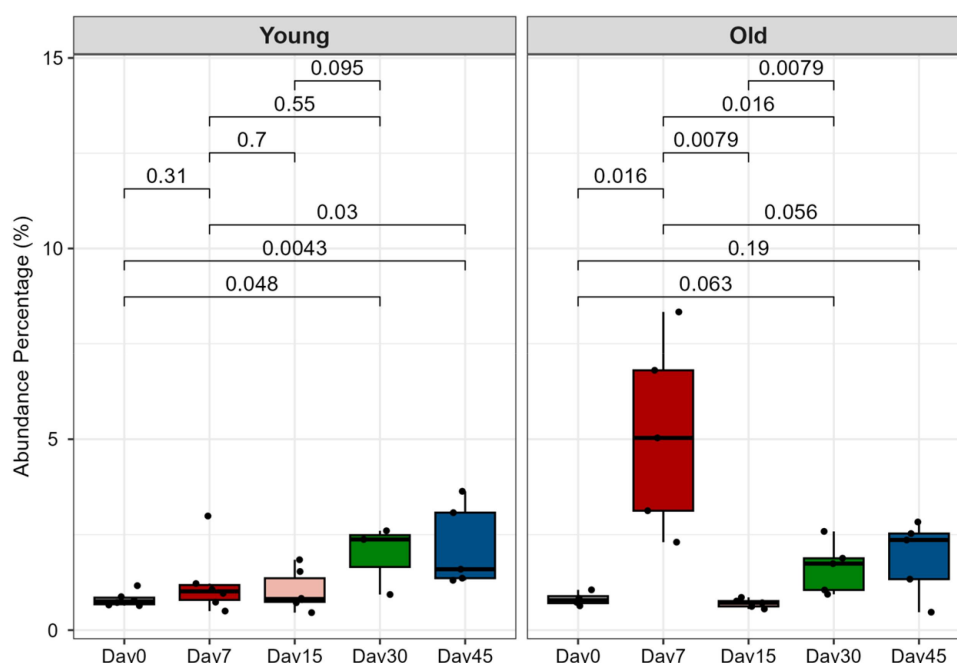


Figure 4. Relative abundances of host proteins in hamster fecal metaproteome. The percentage of sum intensities of all quantified host proteins in relation to the total intensity was calculated for each sample. Statistical significance (Wilcoxon test) and plotting were performed with ggpubr in R. *p* values are indicated for comparisons in both young and old groups for those with significance ($p < 0.05$) in either one group of hamsters.

showed that the most significant changes were at acute phase (Day 7) for both young and old groups (Figure 2 and Supplementary Figure S4), we next focused on identifying key proteins, microbial functions, or taxa that drive the changes at acute phase for Young-Day7 or Old-Day7 groups compared to other samples. We performed partial least squares-discriminant analysis (PLS-DA) analysis for normalized host protein abundances, which achieved a model goodness of prediction (Q^2) of 0.89 and the three clusters can be sufficiently separated with the first two PLS components (Supplementary Figure S5). Accordingly, 40 host proteins with variable importance projection (VIP) ≥ 1 in either PLS component 1 or 2 were selected (Figure 5 and Supplementary Table S3), among which 13 were upregulated and 27 were down-regulated in the Old-Day7 group.

The down-regulated host proteins in feces of SARS-CoV-2 infected hamsters include a diverse array of catabolic enzymes involved in carbohydrate metabolism (i.e., alpha-amylases, sucrase-isomaltase, alpha-1,4 glucan phosphorylase, and UDP-glucose 6-dehydrogenase), protein degradation (i.e., chymotrypsin-C, trypsin, elastase, chymotrypsin-like elastase family protein,

aminopeptidase, and carboxypeptidase B), lipid metabolism and transport (i.e., triacylglycerol lipase and ganglioside GM2 activator isoform X1), as well as for metabolisms of nucleotide (guanine deaminase) and vitamins (biotinidase). Additionally, there are also decreases in proteins related with intestinal barrier function or immune responses to bacteria, such as mucin-5AC, intelectin-1a-like, prolactin-induced protein, and phenazine biosynthesis-like domain-containing protein. On the contrary, the up-regulated host proteins include those involved in recognizing extracellular components of Gram-positive and negative bacteria, such as peptidoglycan recognition protein, pancreatic secretory granule membrane major glycoprotein GP2-like protein, and Ly6/plaur domain containing protein 8. In addition, we also found the up-regulation of proteins related to mitochondrial activity, including the voltage-dependent anion-selective channels (VDAC) proteins and cytochrome c oxidase. Pregnancy zone protein-like is among the most significantly upregulated proteins in old hamsters with SARS-CoV-2 infection at Day 7, and has been reported to be potential biomarker for airway or mucosal infections in humans.³¹ These alterations of host proteins associated with

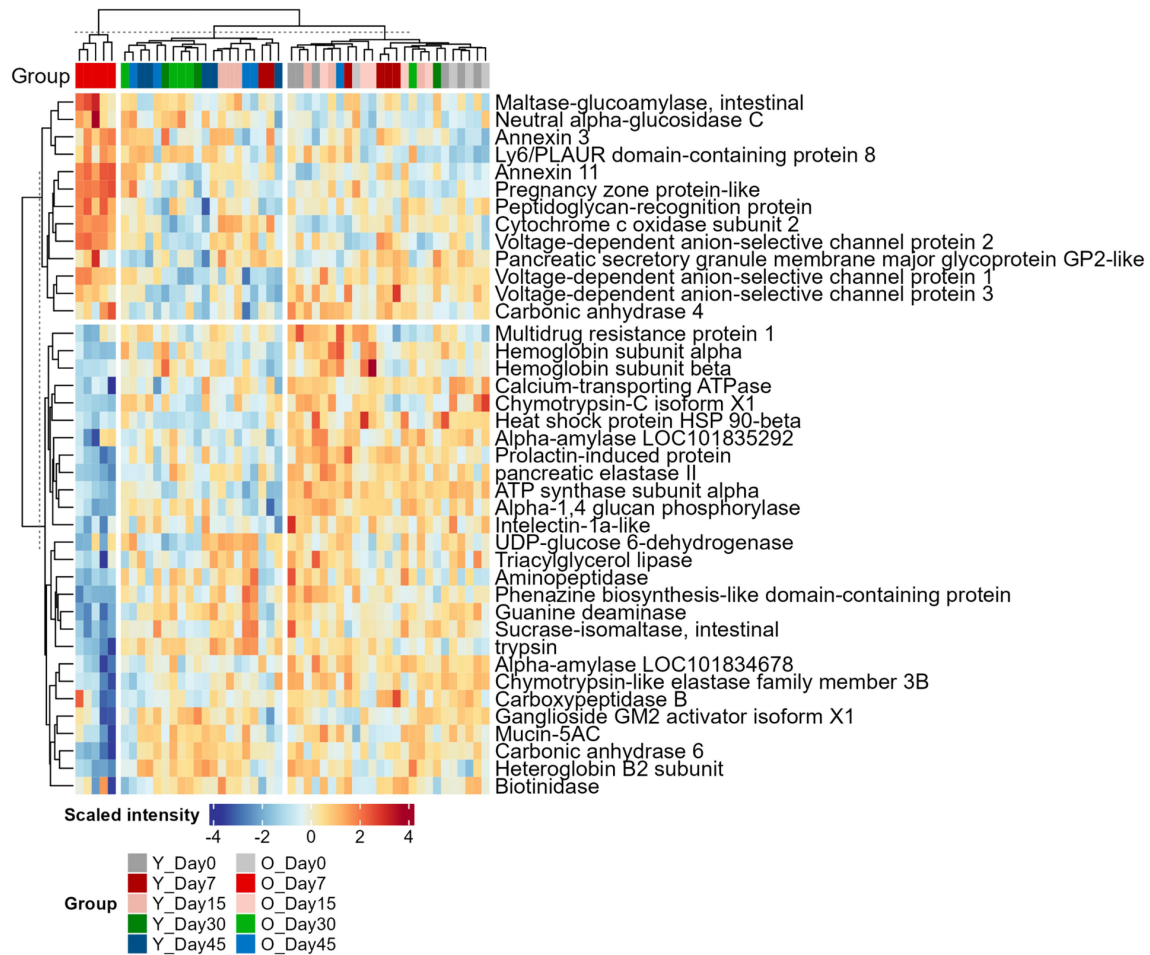


Figure 5. Heatmap of differentially abundant host proteins in young-Day7 or old-Day7 groups. Protein intensities were log2 transformed, scaled and are displayed as colours ranging from blue(low) to red(high) as shown in the key. Heatmap and clustering for both rows and columns are performed using the R ComplexHeatmap package. To maintain stringent quality control, 8 samples with precursor identification less than 50% of the maximum identification number were filtered out. The numbers of samples included for plotting are as follows: day 0 (young: $n=6$; old: $n=4$), day 7 (young: $n=6$; old: $n=5$), day 15 (young: $n=6$; old: $n=5$), day 30 (young: $n=3$; old: $n=5$), and day 45 (young: $n=5$; old: $n=5$).

microbiomes may indicate an aberrant host-microbiome crosstalk during SARS-CoV-2 infection in hamsters.

The PLS-DA analysis for host protein data excluding Day 7 samples achieved a Q^2 of 0.54 and demonstrated a gradual shift by time in the recovery phase in PLS-DA score plot (Supplementary Figure S6). This observation aligns with the increasing total amount of host proteins in the recovery phase (Day 15–45) in both young and old hamsters (Figure 4). A total of 48 proteins were identified with VIP threshold of 1 (Supplementary Figure S7 and Table S4). Among the 48 differentially abundant host proteins, eight presented increasing trend, while 40 showed decreasing trends, particularly in Day 30 and 45. We found

that the abundances of several differentially expressed host proteins in the acute phase were reversed in the recovery phases, such as peptidoglycan recognition protein, VDACs, heteroglobin B2, and GP2-like proteins. However, some others such as Ly6/plaur domain containing 8 and Annexin 3 remain at high levels in the recovery phase compared to Day 0 (Supplementary Figure S7), suggesting potential persistent impacts of viral infection on the intestinal epithelium function.

SARS-CoV-2 infection induced alterations of microbiome functions in hamsters

We next assessed the impacts of SARS-CoV-2 infection on microbiome proteins and functions

by analyzing the normalized microbial protein abundance data. To gain functional information, we annotated all the identified gut microbial proteins with GhostKOALA³² and calculated the abundance of each KEGG Orthology (KO) according to the annotation and protein abundances. In total, 22,158 (69.7%) out of the 31,768 microbial protein groups were annotated into 1571 KOs. Non-supervised PCA of the KO abundance data shows similar sample clustering to those with protein group abundances, namely an obvious shift only at Day 7 and with an age-dependent manner (Supplementary Figure S8).

We then utilized PLS-DA to identify differentially abundant KOs for acute phase in Young-Day7 and Old-Day7 groups compared to others. A Q2 of 0.93 and distinct separation of samples from Old-Day7, Young-Day7, and remainder clusters was achieved at the PLS component 1 (Supplementary Figure S9). Using a VIP threshold of 1, a total of 307 KOs were identified (Supplementary Table S5). Of the 307 differentially abundant KOs, 183 were enzymes and 52 were transporters, indicating significant alteration of microbial metabolism pathways induced by SARS-CoV-2 infections. Similarly, among the 58 KOs with VIP value >2, 31 were enzymes and 11 were transporters. There were 26 obviously up-regulated KOs and 32 down-regulated in the Old-Day7 group with the VIP threshold of 2. As can be seen in the heatmap (Figure 6), there was also a trend of shifting of Young-Day7 toward the Old-Day7 group direction, which agrees with the sample clustering in the PCA score plot. Among these KOs with VIP > 2, six were directly associated with sulfide production, and all of them were significantly upregulated in Old-Day7 and Young-Day7 groups. These KOs include dissimilatory sulfite reductases *Dsr* (K11180, K11181), sulfonate transport system permease protein (K15554), and three taurine/hypotaurine metabolism proteins to generate sulfite (K03851, K03852, K00259).

We also found an increase in peptidoglycan-associated lipoprotein (K03640), a key component of Gram-positive cell wall, in both Young-Day7 and Old-Day7 groups. This observation is in agreement with the upregulation of host peptidoglycan recognition protein in the Old-Day7 group observed in this study. It has been widely

recognized that the gut microbiota plays a pivotal role in neurological disorders through the gut-brain axis.^{33,34} We found that neurotransmitter: Na⁺ symporter-1 (NSS family, K03308) and putrescine aminotransferase (K09251) were upregulated in both Young-Day7 and Old-Day7 groups. Putrescine is one of the sources of inhibitory neurotransmitter gamma-aminobutyric acid (GABA),³⁵ and putrescine aminotransferase is the first step of putrescine degradation. The upregulation of these proteins indicates that the gut microbiota may be involved in the alteration of neurotransmitter homeostasis in the gut of hamsters with SARS-CoV-2 infection. Additionally, we also found a significant change in vitamin degradation-related functions, namely L-ascorbate 6-phosphate lactonase (K03476), being upregulated in COVID-19 hamsters, suggesting potential impacts of SARS-CoV-2 infection in shaping host intestinal homeostasis from various aspects.

A total of 32 KOs were downregulated (with VIP > 2) at both Young-Day7 and Old-Day7 groups. In addition to the enzymes involved in amino acid and polysaccharide metabolisms, these downregulated KOs include two type-IV pilus assembly proteins, PilA and PilC; three flagellar-related proteins, Flif, FliM, and FliG; as well as chemotaxis protein CheD that is closely associated with flagellar motility. These observations suggest potential significant impacts of SARS-CoV-2 infection on the gut microbial cellular motility function in hamsters.

We next sought to identify microbial functions altered during the recovery phase of SARS-CoV-2 infection (Day 15–45). PLS-DA analysis, excluding Day 7 samples, yielded a Q2 of 0.66 and identified 45 KOs with VIP scores >2 (Supplementary Table S6 and Supplementary Figure S10). Heatmap of the differentially abundant microbial KOs revealed the most pronounced changes on Day 30 and Day 45 (Supplementary Figure S11). The majority of these significantly changed functions were associated with carbohydrate (sugars) transport and metabolism, alongside changes in amino acid metabolism (involving aspartate, glycine and glutamate) and ribosomal proteins. Additionally, a significant decrease in anaerobic sulfite reductases (K16950 and K16951) was observed at Day 30 and 45,

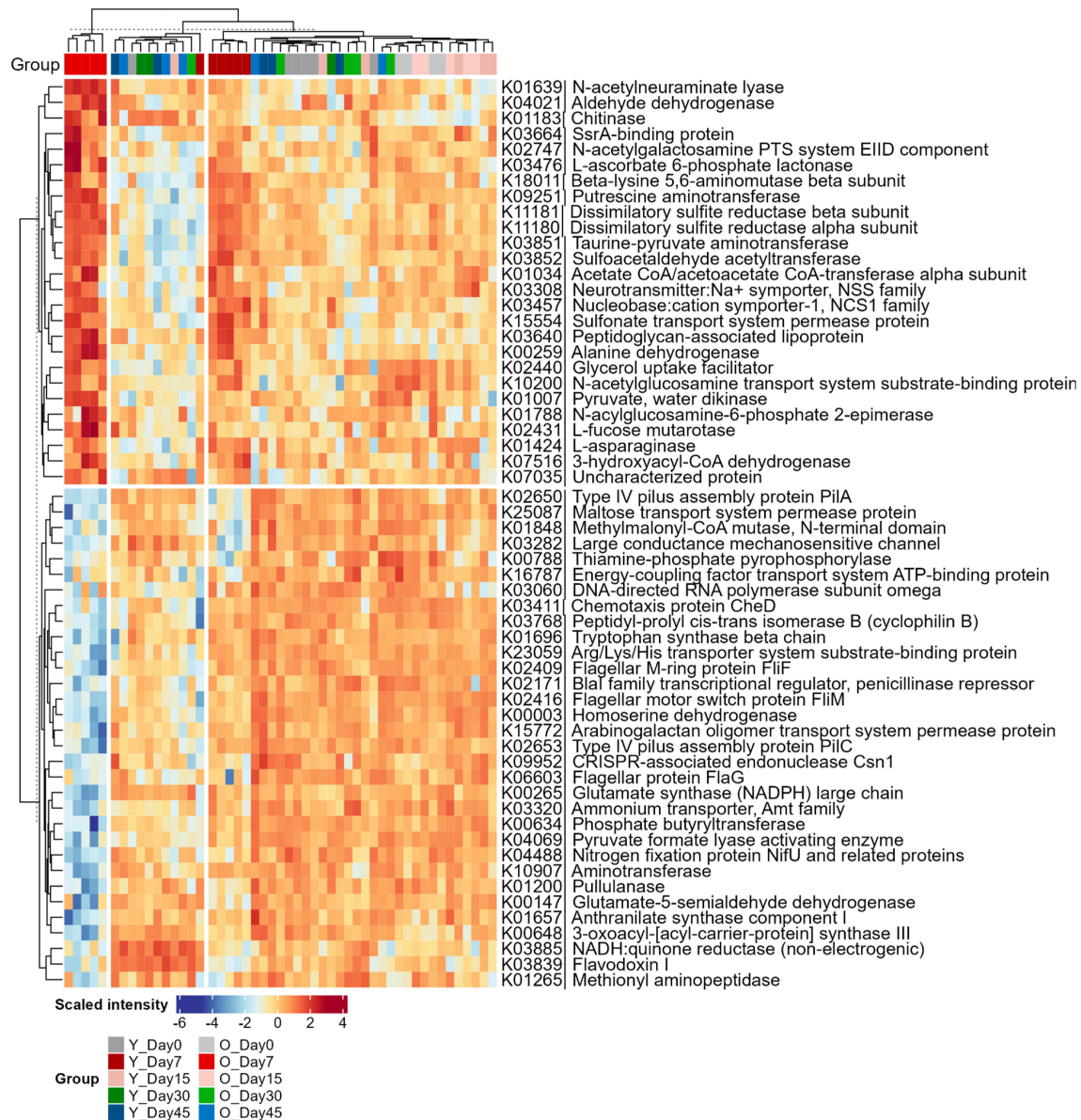


Figure 6. Heatmap of differentially abundant microbial functions. KEGG orthology (KO) intensities were calculated by summing all protein intensities annotated to that KO. KO intensities are then log2 transformed, scaled and displayed as colours ranging from blue(low) to red(high) as shown in the key. Heatmap and clustering for both rows and columns are performed using the R ComplexHeatmap. To maintain stringent quality control, 8 samples with precursor identification less than 50% of the maximum identification number were filtered out. The numbers of samples included for plotting are as follows: day 0 (young: $n=6$; old: $n=4$), day 7 (young: $n=6$; old: $n=5$), day 15 (young: $n=6$; old: $n=5$), day 30 (young: $n=3$; old: $n=5$), and day 45 (young: $n=5$; old: $n=5$).

despite no changes at Day 7 and 15 timepoints (Supplementary Figure S11).

Metaproteomics revealed extensive taxonomic alterations in COVID-19 hamsters

This study identified 128,217 peptides, with 66,997 of them being mapped to unique taxa, which were then used for taxonomic analysis using a peptide-centric workflow in MetaLab.³⁶ By using

a threshold of a minimum of three distinct peptides, all four superkingdoms can be identified with 62 phyla, 109 classes, 183 orders, 264 families, 401 genera, and 419 species (Figure 7(a) and Supplementary Table S7). Here we focus on micro-organisms, so we removed all sequences assigned to Eukaryota (except for Fungi) and calculated relative abundances at each taxonomic rank level. Similar to gut microbiomes of other mammals, Bacilota and Bacteroidota are the main bacterial

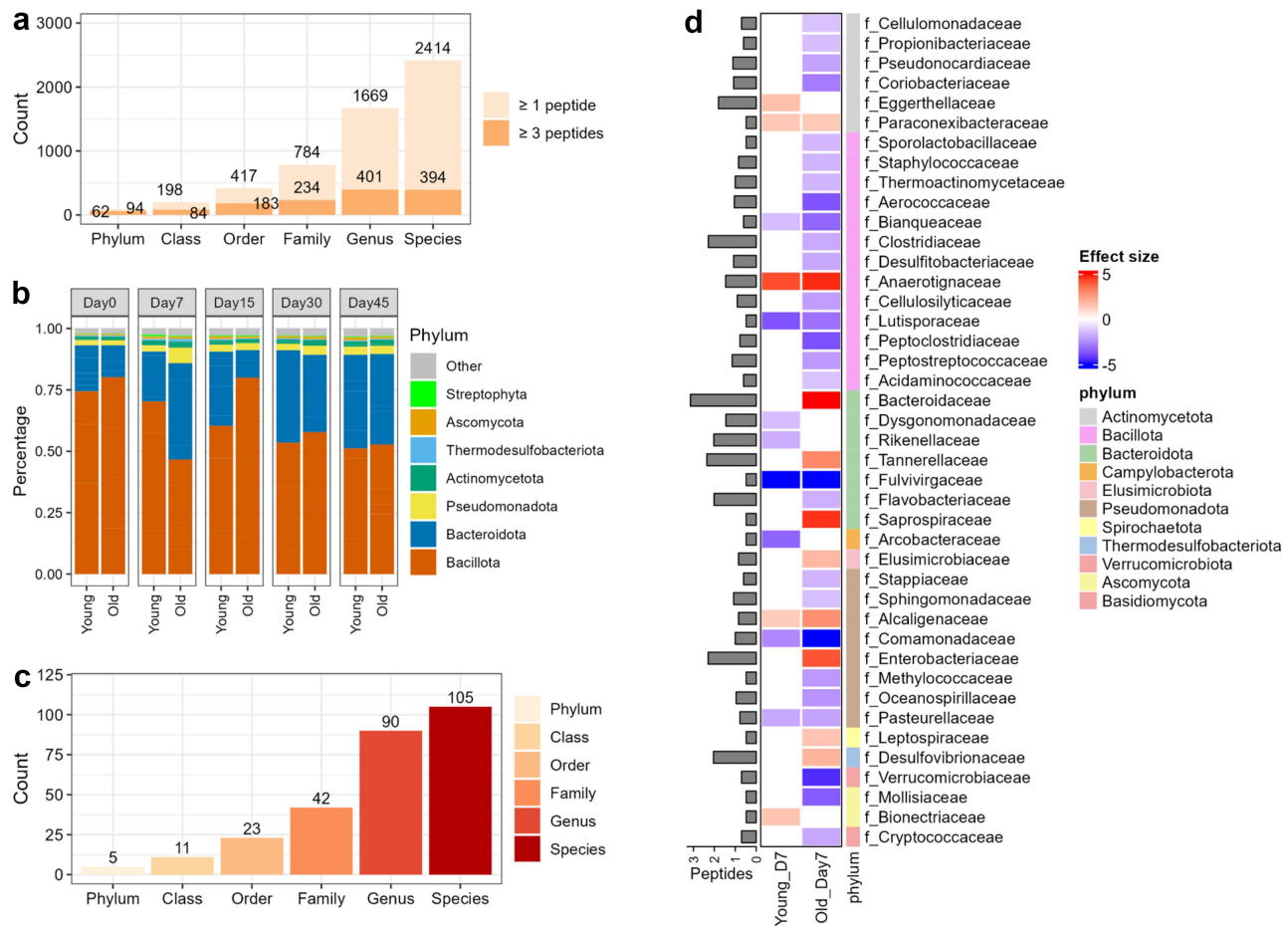


Figure 7. Metaproteome-based taxonomic composition and alterations in hamsters with SARS-CoV-2 infection. (a) Number of taxa identified at each rank level with a minimum of one or three distinctive peptides. Taxa with a minimum of three distinctive peptides were kept for further analysis in this study. (b) Phylum level composition for each group; (c) Numbers of significantly changed taxa at each rank level identified using MaAsLin2; (d) Significantly changed families identified with MaAsLin2. The heatmap shows the effect size with colors ranging from blue (down-regulated) to red (up-regulated) as shown in the key. Left side annotation bar plot shows the number (log10-transformed) of distinctive peptides for each family, while the right-side annotation shows the phylum assignment for each family as shown in the phylum legend.

phyla in the hamster fecal microbiome (Figure 7(b)). An obvious increase of Bacteroidota and Pseudomonadota and decrease of Bacillota can be observed in the Old-Day7 group. These shifts persisted in the recovery phase, particularly on Days 30 and 45, in both young and old groups. These findings are in agreement with observations with shotgun metagenomic sequencing analysis of microbiomes in an independent hamster cohort study.²⁴ Family-level analysis also showed that *Bacteroidaceae* and *Tannerellaceae*, the two abundant Bacteroidota families, are the most obviously increased in the Old-Day7 group compared to others (Supplementary Figure 12). We also observed a significant decrease in species diversity, richness, as well as evenness in the Old-Day7 group

compared to Day 0, while no significant difference was observed for young hamsters (Supplementary Figure 13).

To identify statistically significant differences of taxa at acute phase, we performed MaAsLin2 analysis³⁷ based on a general linear model comparing Young-Day7 and Old-Day7 with the other groups. In total, 277 taxa were identified as significantly changed in either Young-Day7 or Old-Day7 group, including five phyla (Elusimicrobiota, Basidiomycota, Pseudomonadota, Thermodesulfobacteriota, and Thermoproteota), 11 classes, 23 orders, 42 families, 90 genera, and 105 species (Figure 7(c) and Supplementary Table S8). As shown in Figure 7(d), the majority of the significantly different families were observed for the Old-Day7 group. Eight families

showed significant changes in both groups, and all of them exhibited the same changing direction of changes in both young and old groups. *Fulvivirgaceae* was the most significantly down-regulated in both infection groups, while *Anaerotignaceae* (including *Anaerotignum faecicola* and *A. lactatifermentans*) was the most significantly up-regulated family in both groups. *Bacteroidaceae* is among the most significantly changed family and has the highest number of distinctive peptides ($n = 1396$). Eleven species from *Bacteroidaceae* were identified as significantly changed in the Old-Day7 group and only two were decreased. *Tannerellaceae* (including *Parabacteroides distasonis* and *P. goldsteinii*), *Enterobacteriaceae* (mainly *Escherichia coli*, belong to Pseudomonadota) and *Desulfovibrionaceae* (mainly *Candidatus Bilophila faecipullorum*, *Bilophila wadsworthia*, *Mailhella massiliensis*, and *Desulfovibrio legallii*) are also significantly increased in the Old-Day7 group and have high numbers of distinctive peptides (228, 199, and 110, respectively). Interestingly, all four *Desulfovibrionaceae* species were significantly up-regulated in the Old-Day7 group, and three of them (except for *D. legallii*) were up-regulated in the Young-Day7 group as well. *Desulfovibrionaceae* species are known as sulfate-reducing bacteria.³⁸ These observations are in agreement with the findings that the KEGG functions related to sulfite reduction were up-regulated in both Young-Day7 and Old-Day7 groups.

Additionally, we also identified three fungal families that are significantly changed in either the Young-Day7 or Old-Day7 group, including *Mollisiaceae* and *Cryptococcaceae* that were down-regulated, and *Bionectriaceae* that are up-regulated (Figure 7(d)).

MaAsLin2 analysis comparing each individual group in the recovery phase to their corresponding baseline group (Day 0) identified 64 significantly changed taxa in young hamsters and 47 in old hamsters across Day 15, 30 or 45, using an adjusted P-value threshold of 0.05 (Supplementary Table S9). In the young group, the most significantly changed taxa included a decrease in *Ligilactobacillus* (Day 15), *Lachnospira* and *Roseburia inulinivorans* (Day 45), and an increase in *Bacteroides fragilis* (Day 45). In the old group, the most significant changes included an increase in *Ustilaginaceae* (belong to fungal phylum Basidiomycota, Day 30), and a decrease in

Coprococcus catus and *Yeguiia hominis* (Day 30). Additionally, in young hamsters, significant increases were observed in the kingdom Fungi as well as fungal phyla Ascomycota and Basidiomycota on Days 15 and 45, along with a notable increase in the bacterial phylum Pseudomonadota and an archaeal phylum Euryarchaeota on Day 45.

Co-occurrence analysis reveals taxon-specific functional alterations induced by SARS-CoV-2 infection

To examine the relationship between the host, microbial taxa, and function alterations at acute phase, we performed a correlation analysis which identified 115 correlations with an absolute Spearman's correlation coefficient ($|r|$) greater than 0.7 and an adjusted p-value threshold of 0.05. The most positive correlation is between *Tannerellaceae* and *Bacteroidaceae* ($r = 0.91$, $p = 0$), and the most negative correlation is between *Bacteroidaceae* and ATP5A1 ($r = 0.86$, $p = 2.66E-15$). With these significant correlations, three sub-networks were established (Figure 8(a)). The largest network is Network A, where *Tannerellaceae*, *Bacteroidaceae*, *Enterobacteriaceae*, and *Elusimicrobiaceae* were negatively correlated with a cluster of co-occurring down-regulated microbial functions, taxa, and host proteins. In particular, the microbial cellular motility function-related KOs (K03411, K02409, K02416) are positively correlated with each other and negatively correlated with *Tannerellaceae* and *Bacteroidaceae*.

Network B is characterized by a close co-occurrence of the sulfate reducing bacteria *Desulfovibrionaceae* and KOs related to sulfite reduction (K11180 and K11181), taurine metabolism (K03851 and K03852), and putrescine aminotransferase (K09251). As mentioned above, putrescine is an important source of GABA, in both bacteria and mammalian cells.³⁵ The correlation between sulfate-reducing bacteria and function with putrescine aminotransferase may indicate the role of sulfate-reducing microbiota species in aberrant gut-brain axis homeostasis in SARS-CoV-2 infection. The hub of Network C is *Saprospiraceae*, which is among the most significantly up-regulated families in Old-Day7 hamsters and is worth further investigation.

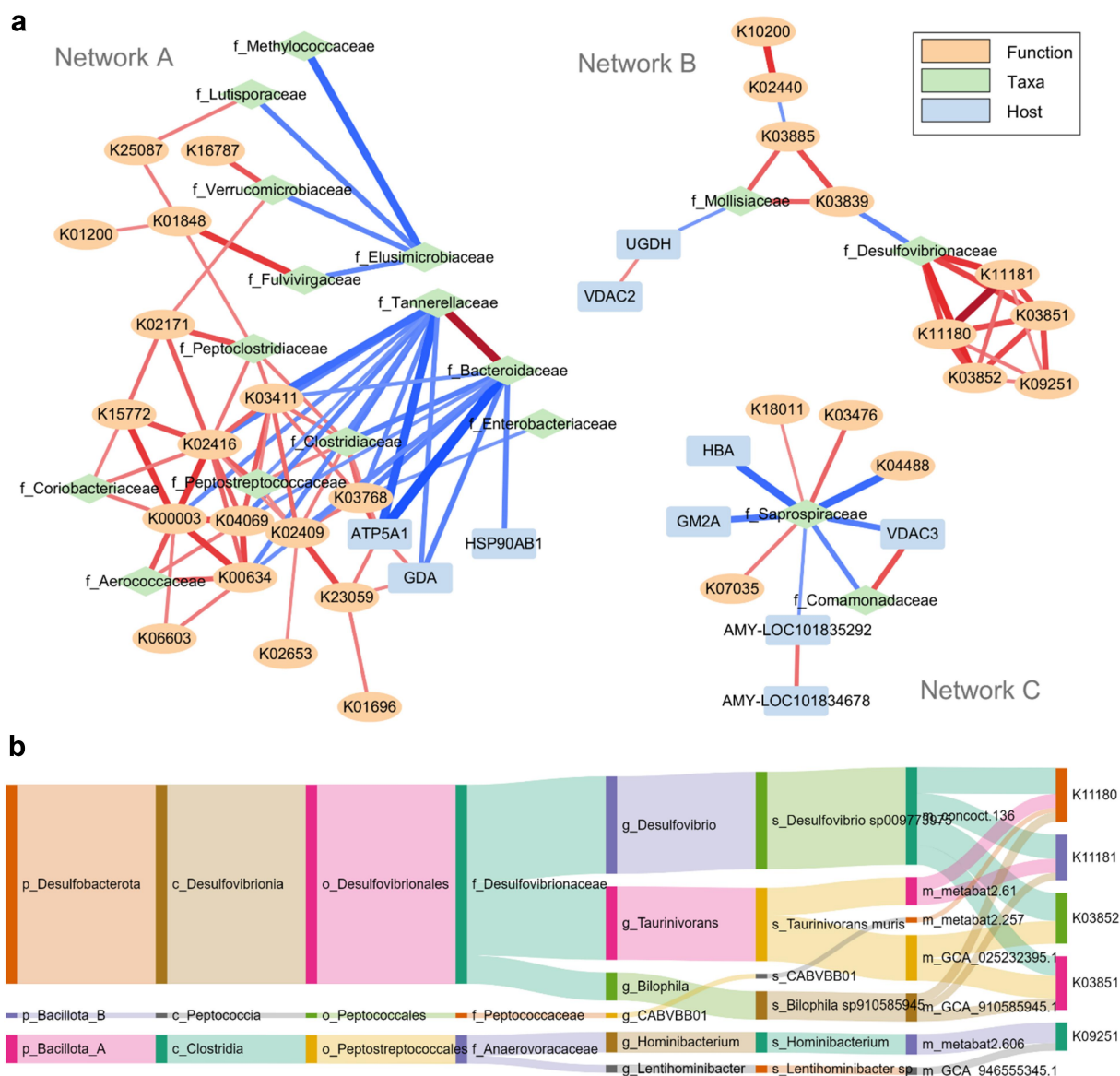


Figure 8. Correlation network of identified differentially abundant host proteins, microbial functions and taxa. (a) Pairwise Spearman correlations are calculated, and only correlations with a correlation $|r| > 0.7$ and an adjusted $p < 0.05$ were used for network visualization using Cytoscape. Three sub-networks with >4 nodes were shown. Color and shape of the nodes show host protein, microbial taxa or function type. The color and thickness of edges shows the direction (red as positive correlation, while blue as negative correlation) and coefficient ($|r|$) of correlation, respectively. (b) Sankey plot showing the OTF links between genomes, their taxonomic lineage and KEGG KOs positively associated with *Desulfovibrionaceae* in Network B.

To further explore the relationship between *Desulfovibrionaceae* and KOs related to sulfate metabolism, we then performed a taxon-function association analysis using MetaX at genome level.³⁹ The latter establishes associations between taxa and functions in metaproteomics based on the operational taxon-function unit (OTF) that is derived from linked peptides. By searching all OTFs with the five *Desulfovibrionaceae*-associated KOs in Network B, we showed that the four KOs related

to sulfite reduction and taurine metabolism are derived from five genomes, four of which belong to *Desulfovibrionaceae* (Figure 8(b)). K09251 is uniquely associated with genome metabat2.606, a strain from genus *Hominibacterium*. The genome concoct.136 (*Desulfovibrio* sp.) expresses all four KOs related to sulfur metabolism, and the genome GCA_910585945.1 (*Bilophila* sp.) expresses three KOs except for K03852. Interestingly, the two genomes belonging to *Taurinivorans muris* express

different functions, with GCA025232395.1 only associated with taurine metabolism KOs, while metabat2.61 only with sulfite reduction KOs. By examining the relative abundance of each OTF, we observed a consistent increase for all the selected OTFs in the Old-Day7 group (Supplementary Figure S14). These observations indicate potential inter- and intra-species strain interactions in sulfur metabolism that may be involved in SARS-CoV-2-induced dysbiosis of hamster microbiome functions.

Discussion

Although the COVID-19 pandemic has subsided, the disease is expected to persist as an endemic virus with seasonal surges, similar to influenza. Ongoing research into the disease's mechanisms, vaccine efficacy, and long-COVID will continue to rely on hamsters as key model organisms. In addition to coronavirus diseases, hamsters have also been considered as a preferred small animal model for pathological and vaccine research for many emerging and reemerging infectious diseases due to its ability to better meet regulatory guidelines for an appropriate animal model.⁸ Previous hamster microbiome analysis primarily with 16S rDNA sequencing has demonstrated compositional changes that are associated with SARS-CoV-2 infection.^{11–13} However, to the best of our knowledge, there has been no hamster fecal metaproteomic study to date. The lack of a hamster microbiome reference protein database is among the first challenges hampering the application of metaproteomics in hamster studies. In the current study, we have established a gene catalog as well as genome databases for the hamster gut microbiome using an in-house as well as a published shotgun metagenomic sequencing dataset of both young and old hamsters. We also reported a comprehensive metaproteomic workflow with advanced DIA-MS methods, a multi-step database search strategy, and comprehensive downstream data analysis, enabling an in-depth functional characterization of hamster microbiomes. The application to a time course study of microbiomes in young and old hamsters with SARS-CoV-2 infection offers valuable insights into the dysbiosis and aberrant crosstalk between the microbiome and

host, implicating elevated host antibacterial functions, opportunistic pathogen colonization, metabolic functions, and their potential roles in impaired gut-brain homeostasis.

A major challenge for studying microbiomes using proteomics is the accurate and efficient identification of microbiome proteins.³⁹ This is particularly challenging for model organisms that are not well characterized for their microbiomes. Most of the current bioinformatic workflows for metaproteomic analysis are based on gene and/or genome catalogs that are derived from massive shotgun metagenomic sequencing. Increasing numbers of gene catalog or genome databases have been made available along with the community efforts. For example, EMBL-EBI MGnify has made available 12 comprehensive gut microbial genome catalog databases for common model and non-model organisms.²¹ These are invaluable reference resources enabling post-genomic applications, including metaproteomics. Unfortunately, so far, no hamster microbiome gene catalog or genome databases are available, despite the critical importance of hamster animal model and the gut microbiomes in COVID-19 research. Here, we address this knowledge gap by establishing a hamster gut microbial reference database, including 8.86 million genes/proteins and 926 genomes, through co-assembly of both young and old hamsters, in-house and published shotgun metagenomic data. The size of the gene catalog database established here is close to the well-established integrated human gut microbial gene catalog (IGC, with 9.9 million non-redundant genes).²³ We demonstrated that the established gene catalog database provided sufficient coverage for metaproteomics protein identification, achieving >30,000 protein groups and >120,000 peptides with up to 57,000 peptides per sample. This depth of identification is also equivalent to the deepest metaproteomics application studies of human¹⁸ or mouse⁴⁰ microbiomes so far.

We then performed the very first hamster metaproteomic study and examined the impacts of SARS-CoV-2 infection on the microbiome functionality in young and old hamsters over 45 days post infection (covering acute and recovery phase of diseases). Interestingly, we found that the hamster metaproteome significantly changed at Day 7

but by Day 15 mostly returned to baseline. This observation aligns well with the fact that the phenotype recovery and viral clearance are completed 2 ~ 3 weeks after infection in both hamsters and humans.^{3,24} More interestingly, the current metaproteomics study also demonstrated that the microbiome responses in young and old hamsters were not the same, with a notably higher extent of changes in old hamsters than those of young animals. These observations also agree with the shotgun metagenomics as well as the plasma metabolomics analyses for samples collected in an independent cohort of hamsters.²⁴ These age-specific alterations include a significant decrease in microbial diversity only for Old-Day7 group (Supplementary Figure S13), more taxa that are significantly changed in the Old-Day7 group than the Young-Day7 group (Figure 7), as well as a greater extent of abundance alterations of key microbial functions (Figure 6). The observations in this hamster metaproteomic study align well with the more severe microbiome dysbiosis in elderly as well.⁴¹ In humans, the elderly population is more severely impacted by COVID-19 due to age-related changes in metabolism and immune function. This demographic typically experiences high levels of inflammation, stress, catabolism, and increased energy and protein needs in their gastrointestinal tract,⁴² thereby leading to more vulnerable microbiomes.

One advantage of metaproteomics over genomic approaches is that it permits simultaneous identification and quantification of both host and microbiome proteins. This advantage enables the exploration of the crosstalk between the microbiome and the host immune system as a result of SARS-CoV-2 infection. We showed that while there is a general increasing trend of microbiome cell-associated host proteins in both young and old hamsters after SARS-CoV-2 infection, a remarkable elevation of host mucosal protein secretion into the gut was observed in the Old-Day7 group (Figure 4). SARS-CoV-2 infects and replicates within enterocytes in the small intestine, specifically targeting the intestinal mucosa.⁴³ A healthy mucus layer is essential for maintaining intestinal homeostasis by supporting the symbiotic relationship between the host and gut microbiota. This mucus layer not only provides spatial

separation between microbes and the intestinal epithelium but also acts as a selective filter, facilitating crucial host-microbe interactions. Further looking at the significantly altered host proteins, this study demonstrated a wide spectrum of down-regulation of catabolic enzymes and barrier function-related proteins, while up-regulating antibacterial and mitochondria activity-related proteins. For example, we found the up-regulation of Ly6/plaur domain-containing protein 8 and peptidoglycan recognition protein in the Old-Day7 group, both of which are bacterial extracellular component-binding proteins and play key roles in maintaining the mucosal barrier by preventing the invasion of bacteria into the inner mucus layer of the colon epithelium.⁴⁴ These observations of altered host protein secretion into the gut indicate significant disruption of intestinal homeostasis in old hamsters with SARS-CoV-2 infection.

Metaproteomics can provide biomass-based taxonomic compositions as well as taxon-specific protein or functional expressions. In this study, with a metaproteome-based taxonomic analysis, we see a decrease in microbial diversity and an increase in opportunistic pathogens *Enterobacteriaceae* (mainly *Escherichia coli*), families in Bacteroidota, and sulfate-reducing *Desulfovibrionaceae* in Old-Day7 hamsters with SARS-CoV-2 infection. These taxonomic alterations also align with those of human patients with COVID-19 which showed significantly decreased bacterial diversity with enrichment of opportunistic pathogens, such as *Streptococcus*, *Veillonella*, *Fusobacterium*, and *Escherichia*.¹ In addition, the elevation of sulfate-reducing bacteria has also been reported to be implicated in many diseases, including inflammatory bowel disease. Mottawea et al. demonstrated that a bacterium *Atopobium parvulum* in the gut of pediatric IBD patients can produce H₂S, which leads to the onset of colitis, while administering H₂S scavenger can mitigate *A. parvulum*-induced colitis in animal models.⁴⁵ Depletion of sulfate-reducing *Desulfovibrionaceae* is also commonly associated with beneficial effects of dietary intervention for alleviating metabolic disorders in humans.⁴⁶ By using a co-occurrence analysis as well as taxon-function linking with MetaX, the current study demonstrated the up-regulation of *Desulfovibrionaceae* specific

dissimilatory sulfite reductase and taurine metabolism pathways in SARS-CoV-2 infected hamsters. More interestingly, these sulfur metabolism-related functions were also found to be significantly correlated with *Anaerotignaceae*-specific putrescine aminotransferase, which can lead to putrescine degradation and thereby reduce GABA synthesis from the microbiome. Significant positive correlations between sulfate-reducing bacteria and putrescine aminotransferase may indicate a potential role of these bacteria in the aberrant gut-brain axis and the development of neurological symptoms in COVID-19 patients.

In summary, this study established a reference database as well as a comprehensive workflow based on advanced DIA-MS for hamster metaproteomic study. The application to study the microbiomes in hamsters with SARS-CoV-2 infection demonstrates age- and time-specific alterations of host proteins, microbial taxonomy, and functions, as well as their cross-talks. The multi-layered functional information provided by metaproteomics enhances the preclinical evaluation of therapeutics, such as vaccines and microbiome-directed therapies, using hamsters as a model. This study also underscores the potential benefits of applying metaproteomics to clinical microbiome samples to better understand the viral-host-microbiome interactions in COVID-19 patients. The limitations of the current infection study include the small sample size and insufficient sampling density, as our primary focus was to establish and validate the metaproteomics workflow as a proof of concept. Applying this workflow to a more comprehensive experimental design with denser time points, a longer monitoring period, and a larger sample size will lead to more conclusive biological findings. Given the prominent role of the microbiome in diseases and the well-recognized suitability of the hamster model for research on emerging and reemerging high-consequence infectious diseases, the methodology developed in this study offers a valuable framework for investigating microbiome-related disease or therapeutic mechanisms. Altogether, the hamster microbiome protein databases and the tailored metaproteomic workflow are significant contributions to the disease and drug research with hamsters as animal models, enabling the opportunity to examine the associations of gut

microbiota composition and functions with the host responses.

Methods

Animal experiment and sample collection

This animal study was conducted in the biosafety level 3 laboratory (BSL3) at the Institut Pasteur de Lille. Experimental procedures adhered to national and institutional guidelines and received approval from the local biological risk evaluation committee (Institut Pasteur de Lille/B59-350009). Ethical approval was obtained from the “Comité d’Ethique en Experimentation Animale (CEEA) 75, Nord Pas-de-Calais” and authorized by the “Education, Research and Innovation Ministry” (APAFIS#25041-2020040917227851v3). Young (2-month-old, $N = 6$) and old (22-month-old, $N = 6$) male Syrian golden hamsters (*Mesocricetus auratus*) were purchased from the Janvier Laboratory (Le Genest-Saint-Isle, France). Animals were housed three per cage and were fed a chow diet (5.1% fat, 19.3% protein, 55.5% carbohydrates, minerals 4.6%, fiber 4%, and humidity 11.5%, SAFE Diets, Augy, France) with access to regular laboratory animal drinking water. Infection was carried out via intranasal administration of 100 μ l of DMEM containing 2×10^4 TCID₅₀ (50% of the tissue culture infectious dose) of BetaCoV/hCoV-19_IPL_France strain of SARS-CoV-2 (NCBI MW575140) as described previously.^{11,29} This strain contains the D614G spike mutation that appeared early in the pandemic and increases viral entry. One old hamster died at 24-day post-infection (missing Days 30 and 45). Therefore, a total of 58 fresh fecal samples were collected directly from the animal on day 0, 7, 15, 30 and 45, and stored in -20°C until processing.

Protein extraction, trypsin digestion, and desalting

Stool pre-processing

Enrichment of bacterial cells was performed by differential centrifugation as previously described.⁴⁷ Briefly, fecal samples were combined with 0.5 mL of glass beads and cold phosphate-buffered saline (PBS) per gram. After thorough vortexing, the mixture was centrifuged at $300 \times g$ for 5 min at 4°C to

isolate the supernatant. The remaining pellets were subjected to two additional extractions with cold PBS. The collected supernatant was then clarified through three sequential centrifugations at $300 \times g$ for 5 min each, all at 4°C. Subsequently, the bacterial cells were pelleted by centrifugation at $14,000 \times g$ for 20 min at 4°C. The resulting pellet was washed twice with cold PBS, with resuspension and centrifugation steps at $14,000 \times g$ for 20 min at 4°C, and finally frozen for future analysis.

Lysis and trypsin digestion

Frozen microbial pellets from differential centrifugation were lysed by resuspending them in 200 µL of lysis buffer containing 4% (w/v) SDS, 8 M urea, and 100 mM TEAB. The lysates were subjected to sonication using a Bioruptor® Plus (Diagenode, Cat# B01020001) at 50% amplitude with a cycle of 10-s pulses on/off for 20 min at 8°C. Following sonication, the samples were centrifuged at $16,000 \times g$ for 10 min at 8°C to remove debris. To precipitate the proteins, six volumes of ice-cold acetone were added to the supernatant, mixed by inversion, and incubated overnight at −20°C. After centrifuging at $16,000 \times g$ for 25 min at 4°C, the supernatant was discarded, and the resulting pellet was washed twice with 1 mL of ice-cold acetone. After the final wash, the protein pellets were briefly air-dried and stored for further processing. The protein pellets were dissolved in 100 µL of buffer containing 0.5% SDS, 8 M urea, and 100 mM TEAB. Protein concentrations were measured using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Cat# 23225), with BSA standards for calibration, according to the manufacturer's instructions. For each sample, 100 µg of protein was reduced with 10 mM DTT at 56°C for 90 min at 500 rpm. Following reduction, proteins were alkylated with 20 mM IAA for 30 min at room temperature in the dark, and the reaction was quenched by adding another 20 mM DTT. To remove SDS, proteins underwent acetone precipitation as previously described. The resulting protein pellets were briefly air-dried and reconstituted in 100 µL of 0.6 M urea in 100 mM TEAB. Trypsin digestion was performed overnight at 37°C with 4 µg of trypsin (Promega, Cat# V511B) at a 1:25 enzyme-to-protein ratio, with continuous shaking at 500 rpm. The reaction was terminated by adding 10 µL of 10% formic acid, adjusting the sample pH to 2–3.

Desalting of tryptic digest

Desalting was carried out using C18 columns (Waters Sep-Pak C18, Cat# 186002318) on a vacuum manifold, maintaining a constant pressure of 5 psi. The C18 plate was first conditioned twice with 200 µL of 50% acetonitrile (ACN) and then equilibrated three times with 200 µL of 5% ACN containing 0.5% trifluoroacetic acid (TFA). Samples were loaded onto the plate, followed by two washes with 5% ACN/0.5% TFA and three additional washes with 0.1% formic acid (FA). Desalted peptides were eluted twice with 100 µL of 75% ACN/0.1% FA. The resulting eluates containing tryptic peptides were dried using a centrivap (Labconco, Cat# 7810010). The dried peptides were resuspended in 30 µL of 0.1% FA, and their concentration was determined using the Pierce Colorimetric Peptide Assay (Cat# 23275) according to the manufacturer's instructions. Finally, the samples were diluted to a concentration of 2 µg/µL in 0.1% FA.

Pooling samples and fractionation

Four pooled samples representing four biological categories (Young Pre-infection, Old Pre-infection, Young Post-infection, Old Post-infection) were created by combining samples as follows: Pool 1–10 µl each of Young Day 0; Pool 2–10 µl each of Old Day 0; Pool 3–2 µl each of Young 7, 15, 30, 45, Pool 4–2 µl each of Old Day 7, 15, 30, 45. These pools were fractionated into eight fractions each using High pH Reversed Phase fractionation kit (Pierce cat # 84868) according to the manufacturer's directions. Fractions were dried on a centrivap and resuspended in 15 µl 0.1% FA for LC-MSMS analysis.

LC-MSMS analysis

LC-MS analysis was performed using a Bruker timsTOF Pro 2 mass spectrometer connected to a Bruker nanoElute 2 UPLC system (Bruker Daltonik, Bremen, Germany). Prior to the analysis, the instrument was calibrated using the Chip Cube High Mass Reference Standard (Agilent, G1982–85001). An HPLC setup with two columns was employed, comprising a C8 trap column followed by a PepSep Twenty-five analytical column (25 cm \times 75 µm, packed with 1.9 µm C18 particles)

from Bruker Daltonik. Peptides were separated chromatographically at a flow rate of 0.5 $\mu\text{L}/\text{min}$ over a 48-min gradient. The gradient program (solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile) was as follows: 2% B at the start, increasing to 35% B by 40 min, followed by a rapid increase to 95% B at 40.5 min, and maintained at 95% B until 48 min. The peptides eluted from the column were analyzed on the timsTOF Pro 2 mass spectrometer, operated in either DDA-PASEF or DIA-PASEF mode. In DDA-PASEF mode, the timsTOF mass spectrometer conducted an initial scan across a mass-to-charge (m/z) range of 100–1700 and an ion mobility window of 0.85–1.30 Vs/cm^2 . The MS/MS data acquisition included four PASEF ramps with an intensity threshold set at 2,500, a target intensity of 20,000, and a maximum precursor charge of 5. The trapped ion mobility spectrometry (TIMS) analyzer operated continuously with a duty cycle featuring 100 ms for both ion accumulation and ramp time, resulting in a total cycle duration of 0.53 s. Collision energy was applied linearly, decreasing from 59 eV at $1/K_0 = 1.6$ Vs/cm^2 to 20 eV at $1/K_0 = 0.6$ Vs/cm^2 . In DIA-PASEF mode, a scan covered the m/z range from 100 to 1700 with an ion mobility range of 0.6–1.60 Vs/cm^2 . The TIMS analyzer also maintained a 100% duty cycle, with 100 ms for accumulation and ramp times, and an estimated total cycle time of 1.8 s. For MS/MS acquisition, precursors within the m/z range of 400–1200 were segmented into 16 scans, each comprising 32 ion mobility steps. Each step had an isolation window of 26 Da, with a 1 Da overlap between adjacent windows. The collision energy ranged from 59 eV at $1/K_0 = 1.3$ Vs/cm^2 to 20 eV at $1/K_0 = 0.85$ Vs/cm^2 , following a linear gradient based on ion mobility.

Metagenomic database generation

This study utilized an in-house shotgun metagenomic sequencing dataset from our companion study with an independent cohort of three young (2 months old) and six old hamsters (22 months old) for the establishment of a reference protein database for metaproteomics. Details on fecal sample collections, DNA extraction, library generation, and sequencing were described in the study by Rodrigues et al.²⁴

Low-quality sequences were first removed with fastp (version 0.23.4)⁴⁸ with default parameters, and host sequences were removed with Kneaddata workflow (<https://huttenhower.sph.harvard.edu/kneaddata/>) with Bowtie 2⁴⁹ and the hamster genome as reference (GCF_017639785.1_BCM_Maur_2.0). The high-quality sequences were then subjected to assembly, gene prediction, taxonomic annotation and quantification, and binning with a previously published SqueezeMeta (v1.6.4) workflow.²⁵ Briefly, the sequence assembly was performed using Megahit, retaining contigs with >200 bpd. PRINSEQ (0.20.4 lite) was used to quality control for the contigs.⁵⁰ tRNA and ribosomal RNA sequences were removed with Aragorn⁵¹ and Barrnap,⁵² respectively, prior to gene prediction with Prodigal.²⁷ Contig abundances were used for quantitative taxonomic analysis. Taxonomic annotation of contigs was performed by DIAMOND⁵³ against the GenBank nr database. Top abundant microbial species were selected based on the calculated taxonomic abundances (MGtax1000 for species with abundance >0.1%, and MGtax10000 for abundance >0.01%). To improve the coverage of the gene catalog database, representative proteome sets were downloaded from UniprotKB Proteomes (45 proteomes for MGtax1000, and 111 proteomes for MGtax10000; Supplementary Table S2) and each was concatenated with gene catalog databases for further evaluations.

To generate a more comprehensive gene catalog database, we downloaded a hamster shotgun metagenomic dataset consisting of 90 metagenomes previously published by Shen et al.²⁸ and performed a co-assembly for gene prediction with the same workflow as described above. High-quality contigs were further used to construct draft genomes using Metabat2.⁵⁴ A total of 926 bins were obtained, and GTDB-Tk (v2.4.0)⁵⁵ was used for taxonomic annotation. Of these, 706 bins were annotated as bacteria, with 251 classified at the species level and 423 at the genus level. Whole genomes for the 251 species were downloaded from GenBank and combined with the assembled bins. The 706 annotated genomes, along with the 251 downloaded genomes, were then digested in silico using Rapid Peptides Generator (RPG) v2.0.5 with trypsin.⁵⁶ After filtering out peptides with lengths outside the range of 7–30 amino acids, a total of

82,350,794 peptides were retained for database construction for MetaX analysis. The workflow was illustrated in Supplementary Figure S15.

Reduced protein database generation with DDA dataset

To generate a reduced protein database, DDA-PASEF data of pooled fractions were searched against the MGDB-V2 database with MSFragger (version 4.0)⁵⁷ implemented in FragPipe (v21.1). The default workflow was used with the following changes: a database split factor of 20, no mass calibration and parameter optimization for MSFragger search, and spectral library generation with ciRT for RT calibration. An FDR threshold of 1% was used for both peptide and protein identifications. The sequences of all identified proteins from the MSFragger search, including indistinguishable proteins, were extracted from original MGDB-V2 using an in-house Perl script to generate a reduced FASTA database.

DIA-NN search for identification and quantitation of DIA-MS dataset

DIA-PASEF data were processed with DIA-NN (v1.9.1)⁵⁸ in library-free mode using the reduced FASTA database generated through MSFragger searches as described above. To identify and quantify host proteins, we appended the Golden hamster proteome (*Mesocricetus auratus* with 20,389 entries; downloaded from UniProtKB on July 24, 2024) to the reduced database. Default settings were used for the DIA-NN search with a precursor and protein level FDR threshold of 1%. MaxLFQ intensities of the identified proteins or peptides were used for all the downstream data analysis.

Downstream taxonomic and functional analysis

Taxonomic annotation and analysis were performed using MetaLab (v2.3.2)⁵⁹ with the peptide data matrix (report.pr_matrix.tsv) as input and UniPept for taxonomic assignment.⁶⁰ A minimum of three distinctive peptides was required for confident identification of taxa. The taxon intensity was calculated with sum intensity of

all distinctive peptides assigned to it and then normalized at each taxonomic rank level to obtain relative abundances for further statistical analysis.

For microbiome functional analysis, all host proteins were first excluded from the quantitative data matrix (report.pg_matrix.tsv) and the remaining microbiome protein abundances were then normalized for further analysis. Functional annotation of microbial proteins was performed using GhostKOALA with default parameters.³² The KEGG GENES database file was set as “genus_prokaryotes + family_eukaryotes + viruses”. The normalized intensities of microbial proteins annotated with the same KO were summed up to obtain KO abundance for further statistical analysis.

Statistical data analysis and visualization

In this study, we kept proteins or microbial functions with valid non-zero values in at least 70% samples for statistical analysis, including principal component analysis (PCA) and partial least-squares-discriminant analysis (PLS-DA). The intensities were log2-transformed and the missing values were imputed with k-nearest neighbor (KNN) algorithm when needed. PCA was performed in R with the princomp function and visualized with ggplot2. PLS-DA was performed in MetaboAnalyst 5.0⁶¹ to calculate the variable importance in projection (VIP) of each protein. VIP is a normalized value, and a VIP > 1 indicates a significant contribution to the PLS-DA model. Hierarchical clustering and heatmap of identified VIP proteins or functions were performed using the R package ComplexHeatmap.⁶²

Statistical analysis of taxonomic composition was performed using MaAsLin2 (Galaxy Version 1.8.0).³⁷ Young-Day7 and Old-Day7 groups were set as fixed effects and the others were set as reference group. The abundance data were log transformed prior to analysis using the LM (linear model) method. Statistical *p* values were corrected with Benjamini–Hochberg method and a corrected *p*-value < 0.05 was deemed as significant in this study.

Diagrams in **Figures 1 and 2** were created in BioRender (Zhang, X. (2023) BioRender.com/f82w697). Co-occurrence analysis was performed

using iMetaShiny⁶³ (<https://shiny.imetalab.ca/>) and visualized using Cytoscape (v3.10.2).⁶⁴ The Sankey plot was generated using SankeyMATIC (<https://sankeymatic.com/>).

Acknowledgments

We gratefully acknowledge Drs Simon Sauvé, Michael Rosu-Myles, Michael Johnston, and Julie Shay from Health Canada for their critical comments and edits to this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by funding from the Government of Canada through Health Canada. Q.W. is supported by a stipend from the TECHNOMISE NSERC-CREATE program. This work was also supported by the React-EU COVID2I (programme opérationnel FEDER/FSE/IEJ Nord-Pas de Calais) and the French National Research Agency (Agence Nationale de la Recherche, ANR): AAP générique 2022, ANR-23-CE15-0014-01, GUTSY).

ORCID

Kai Cheng  <http://orcid.org/0000-0002-6045-0244>
 Vinicius de Rezende Rodvalho  <http://orcid.org/0000-0003-4216-7143>
 Daniel Figeys  <http://orcid.org/0000-0002-5373-7546>
 Francois Trottein  <http://orcid.org/0000-0003-3373-1814>
 Xu Zhang  <http://orcid.org/0000-0003-2406-9478>

Author contributions

X.Z., F.T., X.L., and D.F. designed the study. M.C., F.A. and P. B. collected samples, performed the experiments, and collected the data. All authors participated in the data analysis and interpretation. X.Z., M.C., F.T., X.L., and D.F. wrote the manuscript. All authors read and approved the final manuscript.

Data availability statement

All MS proteomics data that support the findings of this study have been deposited to the ProteomeXchange Consortium (<http://www.proteomexchange.org>) with the dataset identifier PXD057158). Raw metagenomic sequencing data was submitted to NCBI sequence read archive (BioProject: PRJNA1121382). The metagenomics derived

hamster gut microbial gene catalog databases along with all FASTA files of binned genomes, annotations, and the MetaX database, have been uploaded to Zenodo (<https://zenodo.org/records/13909289>).

References

- Alexovic M, Ulicna C, Tabani H, Sabo J. In search of candidate protein biomarkers related to COVID-19 in solid tissues and non-blood fluids: an update. *Proteomics Clin Appl*. 2025;19(2):e202400117. doi: [10.1002/prca.202400117](https://doi.org/10.1002/prca.202400117).
- Zhang F, Lau RI, Liu Q, Su Q, Chan FKL, Ng SC. Gut microbiota in COVID-19: key microbial changes, potential mechanisms and clinical applications. *Nat Rev Gastroenterol Hepatol*. 2023;20(5):323–337. doi: [10.1038/s41575-022-00698-4](https://doi.org/10.1038/s41575-022-00698-4).
- Li J, Ghosh TS, McCann R, Mallon P, Hill C, Draper L, Schult D, Fanning LJ, Shannon R, Sadlier C, et al. Robust cross-cohort gut microbiome associations with COVID-19 severity. *Gut Microbes*. 2023;15(1):2242615. doi: [10.1080/19490976.2023.2242615](https://doi.org/10.1080/19490976.2023.2242615).
- Righi E, Lambertenghi L, Gorska A, Sciammarella C, Ivaldi F, Mirandola M, Sartor A, Tacconelli E. Impact of COVID-19 and antibiotic treatments on gut microbiome: a role for enterococcus spp. *Biomedicines*. 2022;10(11):10. doi: [10.3390/biomedicines10112786](https://doi.org/10.3390/biomedicines10112786).
- Zuo T, Zhang F, Lui GCY, Yeoh YK, Li AYL, Zhan H, Wan Y, Chung ACK, Cheung CP, Chen N, et al. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. *Gastroenterology*. 2020;159(3):944–955.e8. doi: [10.1053/j.gastro.2020.05.048](https://doi.org/10.1053/j.gastro.2020.05.048).
- Xiao Z, Pan M, Li X, Zhao C. Impact of SARS-CoV2 infection on gut microbiota dysbiosis. *Microbiome Res Rep*. 2024;3(1):7. doi: [10.20517/mrr.2023.48](https://doi.org/10.20517/mrr.2023.48).
- Yeoh YK, Zuo T, Lui GC, Zhang F, Liu Q, Li AY, Chung AC, Cheung CP, Tso EY, Fung KS, et al. Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut* 2021; *Gut*. 2021;70(4):698–706. doi: [10.1136/gutjnl-2020-323020](https://doi.org/10.1136/gutjnl-2020-323020).
- Warner BM, Safronetz D, Kobinger GP. Syrian hamsters as a small animal model for emerging infectious diseases: advances in immunologic methods. *Adv Exp Med Biol*. 2017;972:87–101.
- Liu Y, Hu G, Wang Y, Ren W, Zhao X, Ji F, Zhu Y, Feng F, Gong M, Ju X, et al. Functional and genetic analysis of viral receptor ACE2 orthologs reveals a broad potential host range of SARS-CoV-2. *Proc Natl Acad Sci USA*. 2021;118(12):118. doi: [10.1073/pnas.2025373118](https://doi.org/10.1073/pnas.2025373118).
- Wan Y, Shang J, Graham R, Baric RS, Li F, Gallagher T. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural

- studies of SARS coronavirus. *J Virol.* 2020;94(7). doi: [10.1128/JVI.00127-20](https://doi.org/10.1128/JVI.00127-20).
11. Sencio V, Machelart A, Robil C, Benech N, Hoffmann E, Galbert C, Deryuter L, Heumel S, Hantute-Ghesquier A, Flourens A, et al. Alteration of the gut microbiota following SARS-CoV-2 infection correlates with disease severity in hamsters. *Gut Microbes.* 2022;14(1):2018900. doi: [10.1080/19490976.2021.2018900](https://doi.org/10.1080/19490976.2021.2018900).
 12. Seibert B, Caceres CJ, Carnaccini S, Cardenas-Garcia S, Gay LC, Ortiz L, Geiger G, Rajao DS, Ottesen E, Perez DR, et al. Pathobiology and dysbiosis of the respiratory and intestinal microbiota in 14 months old golden Syrian hamsters infected with SARS-CoV-2. *PLOS Pathogens.* 2022;18(10):e1010734. doi: [10.1371/journal.ppat.1010734](https://doi.org/10.1371/journal.ppat.1010734).
 13. Sencio V, Benech N, Robil C, Deruyter L, Heumel S, Machelart A, Sulpice T, Lamazière A, Grangette C, Briand F, et al. Alteration of the gut microbiota's composition and metabolic output correlates with COVID-19-like severity in obese NASH hamsters. *Gut Microbes.* 2022;14(1):2100200. doi: [10.1080/19490976.2022.2100200](https://doi.org/10.1080/19490976.2022.2100200).
 14. Brown JA, Sanidad KZ, Lucotti S, Lieber CM, Cox RM, Ananthanarayanan A, Basu S, Chen J, Shan M, Amir M, et al. Gut microbiota-derived metabolites confer protection against SARS-CoV-2 infection. *Gut Microbes.* 2022;14(1):2105609. doi: [10.1080/19490976.2022.2105609](https://doi.org/10.1080/19490976.2022.2105609).
 15. Nagai M, Moriyama M, Ishii C, Mori H, Watanabe H, Nakahara T, Yamada T, Ishikawa D, Ishikawa T, Hirayama A, et al. High body temperature increases gut microbiota-dependent host resistance to influenza a virus and SARS-CoV-2 infection. *Nat Commun.* 2023;14(1):3863. doi: [10.1038/s41467-023-39569-0](https://doi.org/10.1038/s41467-023-39569-0).
 16. Zhang F, Ge W, Ruan G, Cai X, Guo T. Data-independent acquisition mass spectrometry-based proteomics and software tools: a glimpse in 2020. *Proteomics.* 2020;20(17–18):e1900276. doi: [10.1002/pmic.201900276](https://doi.org/10.1002/pmic.201900276).
 17. Dumas T, Martinez Pinna R, Lozano C, Radau S, Pible O, Grenga L, Armengaud J. The astounding exhaustiveness and speed of the astral mass analyzer for highly complex samples is a quantum leap in the functional analysis of microbiomes. *Microbiome.* 2024;12(1):46. doi: [10.1186/s40168-024-01766-4](https://doi.org/10.1186/s40168-024-01766-4).
 18. Sun Y, Xing Z, Liang S, Miao Z, Zhuo LB, Jiang W, Zhao H, Gao H, Xie Y, Zhou Y, et al. metaExpertpro: a computational workflow for metaproteomics spectral library construction and data-independent acquisition mass spectrometry data analysis. *Mol & Cellular Proteomics: MCP.* 2024;23(10):100840. doi: [10.1016/j.mcpro.2024.100840](https://doi.org/10.1016/j.mcpro.2024.100840).
 19. Wang A, Fekete EE, Creskey M, Cheng K, Ning Z, Pfeifle A, Li X, Figeys D, Zhang X. Assessing fecal metaproteomics workflow and small protein recovery using DDA and DIA PASEF mass spectrometry. *Microbiome Res Rep.* 2024;3(3):39. doi: [10.20517/mrr.2024.21](https://doi.org/10.20517/mrr.2024.21).
 20. Meier F, Beck S, Grassl N, Lubeck M, Park MA, Raether O, Mann M. Parallel accumulation–serial fragmentation (PASEF): multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device. *J Proteome Res.* 2015;14(12):5378–5387. doi: [10.1021/acs.jproteome.5b00932](https://doi.org/10.1021/acs.jproteome.5b00932).
 21. Gurbich TA, Almeida A, Beracochea M, Burdett T, Burgin J, Cochrane G, Raj S, Richardson L, Rogers AB, Sakharova E, et al. Mgnify genomes: a resource for biome-specific microbial genome catalogues. *J Mol Biol.* 2023;435(14):168016. doi: [10.1016/j.jmb.2023.168016](https://doi.org/10.1016/j.jmb.2023.168016).
 22. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, et al. A catalog of the mouse gut metagenome. *Nat Biotechnol.* 2015;33(10):1103–1108. doi: [10.1038/nbt.3353](https://doi.org/10.1038/nbt.3353).
 23. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, Arumugam M, Kultima JR, Prifti E, Nielsen T, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol.* 2014;32(8):834–841. doi: [10.1038/nbt.2942](https://doi.org/10.1038/nbt.2942).
 24. Rodrigues PB, de Rezende Rodovalho V, Sencio V, Benech N, Creskey M, Silva Angulo F. Integrative metagenomics and metabolomics reveal age-associated gut microbiota and metabolite markers in experimental COVID-19. *bioRxiv.* 2024;11(5):622058.
 25. Tamames J, Puente-Sanchez F. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. *Front Microbiol.* 2018;9:3349. doi: [10.3389/fmicb.2018.03349](https://doi.org/10.3389/fmicb.2018.03349).
 26. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 2015;31(10):1674–1676. doi: [10.1093/bioinformatics/btv033](https://doi.org/10.1093/bioinformatics/btv033).
 27. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf.* 2010;11(1):119. doi: [10.1186/1471-2105-11-119](https://doi.org/10.1186/1471-2105-11-119).
 28. Shen S, Wang J, Ma C, Chen Y, Ding H, Zhang J. Understanding the “individual drug reaction” from the perspective of the interaction between probiotics and lovastatin in vitro and in vivo. *Microbiome.* 2023;11(1):209. doi: [10.1186/s40168-023-01658-z](https://doi.org/10.1186/s40168-023-01658-z).
 29. Bogard G, Barthelemy J, Hantute-Ghesquier A, Sencio V, Brito-Rodrigues P, Seron K, Robil C, Flourens A, Pinet F, Eberlé D, et al. SARS-CoV-2 infection induces persistent adipose tissue damage in aged golden Syrian hamsters. *Cell Death Dis.* 2023;14(2):75. doi: [10.1038/s41419-023-05574-w](https://doi.org/10.1038/s41419-023-05574-w).
 30. Delval L, Hantute-Ghesquier A, Sencio V, Flaman JM, Robil C, Angulo FS, Lipskaia L, Çobanoğlu O, Lacoste A-S, Machelart A, et al. Removal of senescent

- cells reduces the viral load and attenuates pulmonary and systemic inflammation in SARS-CoV-2-infected, aged hamsters. *Nat Aging*. 2023;3(7):829–845. doi: [10.1038/s43587-023-00442-w](https://doi.org/10.1038/s43587-023-00442-w).
31. Finch S, Shoemark A, Dicker AJ, Keir HR, Smith A, Ong S, Tan B, Choi J-Y, Fardon TC, Cassidy D, et al. Pregnancy zone protein is associated with airway infection, neutrophil extracellular trap formation, and disease severity in Bronchiectasis. *Am J Respir Crit Care Med*. 2019;200(8):992–1001. doi: [10.1164/rccm.201812-2351OC](https://doi.org/10.1164/rccm.201812-2351OC).
 32. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428(4):726–731. doi: [10.1016/j.jmb.2015.11.006](https://doi.org/10.1016/j.jmb.2015.11.006).
 33. Cryan JF, O’Riordan KJ, Sandhu K, Peterson V, Dinan TG. The gut microbiome in neurological disorders. *Lancet Neurology*. 2020;19(2):179–194. doi: [10.1016/S1474-4422\(19\)30356-4](https://doi.org/10.1016/S1474-4422(19)30356-4).
 34. Hochuli N, Kadyan S, Park G, Patoine C, Nagpal R. Pathways linking microbiota-gut-brain axis with neuroinflammatory mechanisms in Alzheimer’s pathophysiology. *Microbiome Res Rep*. 2024;3(1):9. doi: [10.20517/mrr.2023.39](https://doi.org/10.20517/mrr.2023.39).
 35. Sequerra EB, Gardino P, Hedin-Pereira C, de Mello FG. Putrescine as an important source of GABA in the postnatal rat subventricular zone. *Neuroscience*. 2007;146(2):489–493. doi: [10.1016/j.neuroscience.2007.01.062](https://doi.org/10.1016/j.neuroscience.2007.01.062).
 36. Cheng K, Ning Z, Zhang X, Li L, Liao B, Mayne J, Stintzi A, Figeys D. MetaLab: an automated pipeline for metaproteomic data analysis. *Microbiome*. 2017;5(1):157. doi: [10.1186/s40168-017-0375-2](https://doi.org/10.1186/s40168-017-0375-2).
 37. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, Tickle TL, Weingart G, Ren B, Schwager EH, et al. Multivariable association discovery in population-scale meta-omics studies. *PLOS Comput Biol*. 2021;17(11):e1009442. doi: [10.1371/journal.pcbi.1009442](https://doi.org/10.1371/journal.pcbi.1009442).
 38. Heidelberg JF, Seshadri R, Haveman SA, Hemme CL, Paulsen IT, Kolonay JF, Eisen JA, Ward N, Methe B, Brinkac LM, et al. The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nat Biotechnol*. 2004;22(5):554–559. doi: [10.1038/nbt959](https://doi.org/10.1038/nbt959).
 39. Wu Q, Ning Z, Zhang A, Zhang X, Sun Z, Figeys D. Operational taxon-function framework in MetaX: Unveiling taxonomic and functional associations in metaproteomics. *Anal Chem*. 2025, 97(18): 9739–9747. doi: [10.1021/acs.analchem.4c06645](https://doi.org/10.1021/acs.analchem.4c06645).
 40. Gomez-Varela D, Xian F, Grundtner S, Sondermann JR, Carta G, Schmidt M. Increasing taxonomic and functional characterization of host-microbiome interactions by DIA-PASEF metaproteomics. *Front Microbiol*. 2023;14:1258703. doi: [10.3389/fmicb.2023.1258703](https://doi.org/10.3389/fmicb.2023.1258703).
 41. Magalhaes NS, Savino W, Silva PMR, Martins MA, Carvalho VF. Gut microbiota Dysbiosis is a crucial player for the poor outcomes for COVID-19 in elderly, diabetic and hypertensive patients. *Front Med (Lausanne)*. 2021;8:644751. doi: [10.3389/fmed.2021.644751](https://doi.org/10.3389/fmed.2021.644751).
 42. Rea IM, Alexander HD. Triple jeopardy in ageing: COVID-19, co-morbidities and inflamm-ageing. *Ageing Research Rev*. 2022;73:101494. doi: [10.1016/j.arr.2021.101494](https://doi.org/10.1016/j.arr.2021.101494).
 43. Lamers MM, Beumer J, van der Vaart J, Knoop K, Puschhof J, Breugem TI, Ravelli RBG, Paul van Schayck J, Mykityn AZ, Duimel HQ, et al. SARS-CoV-2 productively infects human gut enterocytes. *Science*. 2020;369(6499):50–54. doi: [10.1126/science.abc1669](https://doi.org/10.1126/science.abc1669).
 44. Okumura R, Kurakawa T, Nakano T, Kayama H, Kinoshita M, Motooka D, Gotoh K, Kimura T, Kamiyama N, Kusu T, et al. Lypd8 promotes the segregation of flagellated microbiota and colonic epithelia. *Nature*. 2016;532(7597):117–121. doi: [10.1038/nature17406](https://doi.org/10.1038/nature17406).
 45. Mottawea W, Chiang CK, Muhlbauer M, Starr AE, Butcher J, Abujamel T, Deeke SA, Brandel A, Zhou H, Shokralla S, et al. Altered intestinal microbiota–host mitochondria crosstalk in new onset Crohn’s disease. *Nat Commun*. 2016;7(1):13419. doi: [10.1038/ncomms13419](https://doi.org/10.1038/ncomms13419).
 46. Xiao S, Fei N, Pang X, Shen J, Wang L, Zhang B, Zhang M, Zhang X, Zhang C, Li M, et al. A gut microbiota-targeted dietary intervention for amelioration of chronic inflammation underlying metabolic syndrome. *FEMS Microbiol Ecol*. 2014;87(2):357–367. doi: [10.1111/1574-6941.12228](https://doi.org/10.1111/1574-6941.12228).
 47. Zhang X, Li L, Mayne J, Ning Z, Stintzi A, Figeys D. Assessing the impact of protein extraction methods for human gut metaproteomics. *J Proteomics*. 2018;180:120–127. doi: [10.1016/j.jprot.2017.07.001](https://doi.org/10.1016/j.jprot.2017.07.001).
 48. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta*. 2023;2(2):e107. doi: [10.1002/imt2.107](https://doi.org/10.1002/imt2.107).
 49. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods*. 2012;9(4):357–359. doi: [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923).
 50. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27(6):863–864. doi: [10.1093/bioinformatics/btr026](https://doi.org/10.1093/bioinformatics/btr026).
 51. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res*. 2004;32(1):11–16. doi: [10.1093/nar/gkh152](https://doi.org/10.1093/nar/gkh152).
 52. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–2069. doi: [10.1093/bioinformatics/btu153](https://doi.org/10.1093/bioinformatics/btu153).
 53. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12(1):59–60. doi: [10.1038/nmeth.3176](https://doi.org/10.1038/nmeth.3176).

54. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359. doi: [10.7717/peerj.7359](https://doi.org/10.7717/peerj.7359).
55. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH, Hancock J. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2019;36(6):1925–1927. doi: [10.1093/bioinformatics/btz848](https://doi.org/10.1093/bioinformatics/btz848).
56. Maillet N. Rapid peptides generator: fast and efficient in silico protein digestion. *NAR Genomics Bioinf*. 2020;2(1):lqz004. doi: [10.1093/nargab/lqz004](https://doi.org/10.1093/nargab/lqz004).
57. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. Msfragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods*. 2017;14(5):513–520. doi: [10.1038/nmeth.4256](https://doi.org/10.1038/nmeth.4256).
58. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17(1):41–44. doi: [10.1038/s41592-019-0638-x](https://doi.org/10.1038/s41592-019-0638-x).
59. Cheng K, Ning Z, Zhang X, Li L, Liao B, Mayne J, Figeys D. MetaLab 2.0 enables accurate post-translational modifications profiling in Metaproteomics. *J Am Soc Mass Spectrom*. 2020;31(7):1473–1482. doi: [10.1021/jasms.0c00083](https://doi.org/10.1021/jasms.0c00083).
60. Mesuere B, Debyser G, Aerts M, Devreese B, Vandamme P, Dawyndt P. The unipept metaproteomics analysis pipeline. *Proteomics*. 2015;15(8):1437–1442. doi: [10.1002/pmic.201400361](https://doi.org/10.1002/pmic.201400361).
61. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, Gauthier C, Jacques P-É, Li S, Xia J, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021;49(W1):W388–W396. doi: [10.1093/nar/gkab382](https://doi.org/10.1093/nar/gkab382).
62. Gu Z. Complex heatmap visualization. *Imeta*. 2022;1(3):e43. doi: [10.1002/imt2.43](https://doi.org/10.1002/imt2.43).
63. Li L, Ning Z, Cheng K, Zhang X, Simopoulos CMA, Figeys D. iMetalab suite: a one-stop toolset for metaproteomics. *Imeta*. 2022;1(2):e25. doi: [10.1002/imt2.25](https://doi.org/10.1002/imt2.25).
64. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504. doi: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303).