

Metaproteomic Profiling of Fungal Gut Colonization in Gnotobiotic Mice

Veronika Kuchařová Pettersen

University of Calgary

Antoine Dufour

University of Calgary

Marie-Claire Arrieta (✉ marie.arrieta@ucalgary.ca)

University of Calgary

Research Article

Keywords: Early Life Gut Microbiome, Quantitative Proteomics, Metaproteomics, Gut Fungi, Gnotobiotic Murine Model, Antimicrobials

Posted Date: December 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1193339/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Animal Microbiome on February 22nd, 2022. See the published version at <https://doi.org/10.1186/s42523-022-00163-2>.

Abstract

Background: Eukaryotic microbes can modulate mammalian host health and disease states, yet the molecular contribution of gut fungi remains nascent. We previously showed that mice exclusively colonised with fungi displayed increased sensitivity to allergic airway inflammation and had fecal metabolite profiles similar to germ-free mice. This marginal effect on the host metabolome suggested that fungi do not primarily use metabolites to modulate the host immune system.

Methods: To describe functional changes attributed to fungal colonisation, we performed mass spectrometry-based analyses of feces (Label-Free Quantitative; LFQ) and the small intestine (labeling with Tandem Mass Tag; TMT) of gnotobiotic mice colonised with defined consortia of twelve bacterial species, five fungal species, or both. We also evaluated the effect of microbiome perturbances on the metaproteome by analysing feces from mouse pups treated with an antibiotic or antifungal.

Results: We detected 6,675 proteins in the mice feces, of which 3,845 had determined LFQ levels. Analysis of variance showed changes in the different gnotobiotic mouse groups; specifically, 46% of 2,860 bacterial, 15% of 580 fungal, and 76% of 405 mouse quantified proteins displayed differential levels. The antimicrobial treatments resulted in lasting changes in the bacterial and fungal proteomes, suggesting that the antimicrobials impacted the entire community. Fungal colonisation resulted in changes in host proteins functional in innate immunity as well as metabolism, predicting specific roles of gut fungi on host systems during early developmental stages. Several of the detected fungal proteins (3% of 1,492) have been previously reported as part of extracellular vesicles and having immunomodulating properties. Using an isobaric labelling TMT approach for profiling low abundant proteins of the jejunal tissue, we confirmed that the five fungal species differentially impacted the host intestinal proteome compared to the bacterial consortium. The detected changes in mouse jejunal proteins (4% of 1,514) were mainly driven by metabolic proteins.

Conclusions: We used quantitative proteomic profiling of gnotobiotic conditions to show how colonisation with selected fungal species impacts the host gut proteome. Our results suggest that an increased abundance of certain gut fungal species in early life may affect the developing intracellular attributes of epithelial and immune cells.

Introduction

The fungal portion of the mammalian gut microbiome, also referred to as the mycobiome, is estimated to constitute less than 0.1% of the gut ecosystem (1, 2). Although it has been significantly less characterised than the bacterial microbiome, the mycobiome has an important and often overlooked role in host health and disease (3-5). Cell numbers alone do not provide a comparable measure for the microbiome communities as fungal cells are up to 100 times bigger in volume and with genomes from 4 to 200 times larger than most bacteria (6), representing large biomass with potent production capacities for proteins and metabolites. Community ecology provides many examples of low abundance species with sizable

impacts on community structure and function, showing that species' importance cannot be predicted based upon their abundance in an ecosystem (7-9).

Fungi are an essential part of terrestrial and aquatic ecosystems, engaging in a wide variety of relationships with other members of microbial communities (*i.e.*, bacteria, archaea, viruses) and their plant and animal hosts (10-12). Bacteria modulate fungi's ability to colonise mammalian hosts, as evidenced by clinical (13) and murine (14, 15) studies describing the effects of antibiotic treatment on the gut mycobiome. In general, antibiotic use promotes fungal overgrowth, and its immunological consequences can be detected at distant sites such as the lung (13, 15). Similarly, antifungals affect bacterial community structure and can have adverse effects on the host health (16). Controlled release of inbred laboratory mice into an outdoor enclosure further demonstrated the impact of fungi on the host immune functions (5). Although the mice rewilding enhanced the differentiation of immune cell populations previously associated with pathogen exposure, the resulting fungi-induced immune system changes occurred in the absence of an infection.

The inter- and intra-kingdom relations between microbes of the gut microbiome have been mostly described using the relative species abundance derived by DNA sequencing methods (17). However, this approach does not characterise molecular interactions between microbes that occur directly through secreted molecules and by physical contact, or indirectly by modulating the host immune response. Using a gnotobiotic mice model, we previously demonstrated the impact of gut fungal colonisation on a defined bacterial consortium and the host (18). Gut fungi promoted shifts in bacterial microbiome ecology, and mice colonised exclusively with fungi displayed immunological features related to atopy in early life. Intriguingly, gut fungi had a marginal effect on the host fecal metabolome (18), suggesting that other functional mechanisms are at play. The results led us to investigate the host-microbiome interactions at the protein level, taking advantage of the defined experimental conditions, the availability of the bacterial species' genomes, and reference protein databases for the fungal species and mouse host.

Here, we show proteomic analyses of feces and small intestine of gnotobiotic mice, which were colonised with either 12 bacterial species (Bacteria, -B), a group of 5 fungal/yeast species (Yeast, -Y), or a combination of the 17 fungi and bacteria (Bacteria and Yeast, -BY) (Figure 1A). We utilised the Oligo-MM¹² bacterial consortium of mouse-derived strains that are persistent, inheritable and elicit an immune response in mice similar to a complex microbiota (19, 20). For fungi, we selected six fungal strains from taxa that commonly colonise the human gut (21, 22) and have been previously linked to atopy and asthma risk (23, 24). We also evaluated the effect of microbiome perturbances to the metaproteome by analysing feces from mouse pups treated with antibiotic Augmentin (BY_ABX) or antifungal Fluconazole (BY_AFX). We selected Augmentin (Amoxicillin / Clavulanic acid) a broad-spectrum antibiotic commonly used in the paediatric population (25), as the animal model was designed to mimic adverse factors affecting the early life gut microbiome.

We observed dynamic relationships between the bacteria and fungi that were reflected by changes in the individual species proteome profiles derived from different colonisation and treatment conditions.

Metaproteomic analyses based on label-free quantification (LFQ) of fecal proteins documented an extensive impact of microbial colonisation on the host and revealed new features of the host protein response to fungal colonisation. Proteomics of the jejunal tissue based on labeling with Tandem Mass Tag (TMT) provided further insights on the host cellular pathways impacted by the microbial colonisation and confirmed that gut fungi elicit distinct effects compared to bacteria.

Materials & Methods

Gnotobiotic Mice

Germ-free (GF) C57Bl/6J mice were obtained from and housed at the gnotobiotic mouse facility of the International Microbiome Centre (IMC) at the University of Calgary. Details of the animal experiments have been previously described (18), and all the animal work was conducted following animal protocols approved by the Institutional Animal Care and Use Committee. Briefly, female adult GF mice were orally gavaged twice, three days apart, with 100 µl of a consortium of microorganisms or kept under germ-free conditions. Colonisation consisted of consortia of (i) 12 mouse-derived bacteria (19) (B - bacteria), (ii) five yeast species previously linked to atopy and asthma risk in infants (26, 27) (Y - Yeast), or (iii) a combination of all 17 bacterial and yeast species (B Y) (**Figure 1A, Table S1**). We used previously described method for mice colonisation with the Oligo-MM¹² consortium (19), in which inocula were prepared under anaerobic conditions by mixing 100 µl of 2-day-old microbial cultures of each species. Bacteria were grown in fastidious anaerobe broth (LabM, Heywood, England, United Kingdom), and yeasts were grown in yeast-mold broth (YM; BD, Sparks, Maryland, USA). After the second gavage, mice were paired for mating on a 2:1 female:male ratio per cage. Two breeding pairs were used for each group, and they produced on average 9 offspring (±2) per colonisation condition, which constituted individual biological replicates in the gnotobiotic experiment. To ensure microbial colonisation with the desired consortia in the offspring, the corresponding inocula were further spread on the dams abdominal and nipple regions on days 3 and 5 after birth (28). Microbial engraftment was confirmed by sequencing (18, 29), with an intestinal abundance of bacterial communities resembling previously described composition of the Oligo-MM¹² consortium (19). Two groups of mice colonised with both bacteria and yeasts were treated with the antibiotic Augmentin (0.2 mg/ml; MilliporeSigma, St. Louis, Missouri, USA) or antifungal Fluconazole (0.5 mg/ml; MilliporeSigma) in sterile drinking water from day 7 to 14 after birth (groups BY_ABX and BY_AFX, respectively). Treatment solutions were prepared by dissolving the antimicrobials in distilled water, followed by filter sterilisation. Mice were kept at maximum five animals per isolator cage and housed inside gnotobiotic flexible-film isolators, under a 12-h light/12-h dark cycle, 40% relative humidity, 22–25 °C, and ad libitum access to sterile food and water.

Enrichment of Microbial Cells from Fecal Sample

Fecal samples were collected from gnotobiotic mice at the end of their 3rd week of life and immediately stored at -80°C until use. After thawing at 4°C, pooled samples of ~ 300 mg, originating from co-housed mice of the same treatment group, were subjected to differential centrifugation to enrich for microbial

cells, according to previously described methodology (30). Briefly, each sample was resuspended in 4 ml of Phosphate-Buffered Saline (PBS), homogenised by using GentleMACS C tubes (Miltenyi Biotec, Bergisch Gladbach, Germany), and subjected to low-speed centrifugation at $20 \times g$ for 5 min (Centrifuge 5810 R, Eppendorf, Hamburg, Germany) to eliminate gross particulate material. The supernatant was transferred to 50 ml conical centrifuge tube and kept at 4°C, whereas the pellet was resuspended in PBS. The washing step was repeated until the supernatant appeared translucent (5-7 times). The collected supernatant was centrifuged at $3,214 \times g$ for 1 h, and the resulting pellet was subjected to cell lysis and protein digestion described below. A quality control step comprising of microscopic examination of Gram-stained fractions of the pellet was included to confirm bacterial and fungal cell extraction.

Murine Intestinal Tissue Sample

Jejunal tissue dissected from the small intestine of 4-weeks old mice was cleaned of luminal debris with PBS and snap-frozen in liquid nitrogen and stored at -80°C until further processing.

Protein Extraction

Enriched fecal microbiota and jejunal tissue samples were resuspended in lysis buffer [2% sodium dodecyl sulfate, 100 mM Triethylammonium bicarbonate (TEAB) buffer, 10 mM Ethylenediaminetetraacetic acid, and 1X Complete Mini EDTA free protease inhibitors (F. Hoffmann-La Roche AG, Basel, Switzerland), pH 8.0] in 1:4 w/v ratio and transfer into a 2 ml screw-cap tube with FastPrep Lysing matrix A (MP Biomedicals, Irvine, California, USA). Cells were mechanically disrupted by bead-beating for 2 x 3 min at 30 Hz (TissueLyser II, Qiagen, Hilden, Germany). The samples were incubated at -80°C for 10 min and at 95°C for 10 min, followed by centrifugation for 30 min at $18,000 \times g$, 4°C. To disrupt released cellular DNA that would interfere with downstream protein quantification, the supernatants were sonicated 3 x for 10s with 20s resting intervals on ice. Sonicated samples were centrifuged at $18,000 \times g$, 4°C, for 10 min, the supernatants collected, and protein concentration was measured by using DeNovix Spectrophotometer (DeNovix Inc., Wilmington, DE, USA).

Sample Preparation for Label-free Quantitative (LFQ) Proteomics

The cell lysates of fecal microbiota-enriched samples were processed according to Filter-Aided Sample Preparation protocol (30, 31). Briefly, cell lysates containing 500 µg of total protein were incubated with 10 mM DTT in 100 mM Ammonium Bicarbonate (AmBic) at the solution to total protein ratio (v/w) 1:10 for 45 min at 56°C. The samples were then mixed with 8 M Urea in 10 mM HEPES pH 8.0 (UA) in YM-30 Microcon filter units (MilliporeSigma) and centrifuged at $10,000 \times g$ for 15 min. After discarding the eluate, the filtration units were washed once with the UA buffer ($10,000 \times g$, 15 min). Next, 50 mM iodoacetamide in UA was added to the filter, and samples were incubated in the dark for 20 min. The filter was washed twice with UA, followed by two washes with 50 mM AmBic. Proteins were digested with trypsin (Promega, Madison, Wisconsin, USA) in 40 mM AmBic at 37°C for 18 h (enzyme to protein ratio (v/w) of 1:100). The resulting peptide mixtures were desalted by using SepPak C18 solid-phase extraction cartridges (Waters, Mississauga, Ontario, Canada), lyophilised at 30°C in a vacuum concentrator (Speed

Vac Plus SC110, Savant Instruments Inc., Holbrook, New York, USA) and stored at -80°C until further analysis. Prior to the LC-MS/MS analysis, the peptide mixtures were resuspended in 1% formic acid (FA). An aliquot of the tryptic digests was used to determine the concentration of the peptide mixtures by using Pierce™ Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

Tandem Mass Tag (TMT) Labelling

TMT samples were prepared according to the manufacturer's instructions (TMTsixplex™ Isobaric Label Reagent Set, Thermo Fisher Scientific). Briefly, 150 μg of jejunal tissue cell lysates were reduced by incubating with 200 mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) at 55°C for 1 h and alkylated with 375 mM iodoacetamide in the dark for 30 min. The proteins were precipitated by acetone, resuspended in 50 mM TEAB, and digested with trypsin at 37°C for 18 h. Resulting peptide mixtures originating from 4 treatment groups (GF, B, BY, Y) and 5 biological replicates ($n=5$) were labelled with TMT reagents distributed into 4 experimental groups (**Figure S1**). Each experiment had one TMT tag (131) that contained pooled samples, created by combining equal amounts of peptides (20 μg) of each sample from one experimental group. The pooled samples served as internal standards for normalising the data across the experimental setups. For labelling, the peptides were incubated with TMT reagents for 1 h at room temperature. The reaction was quenched by adding 5% hydroxylamine, followed by incubation for 15 min. The resulting four multiplexed samples were desalted and quantified as described above.

Mass Spectrometry Data Acquisition

Liquid Chromatography: Tryptic peptides were analysed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific) operated with Xcalibur (version 4.0.21.10) and coupled to a Thermo Scientific Easy-nLC (nanoflow Liquid Chromatography) 1200 system. Tryptic peptides (2 μg) were loaded onto a C18 trap (75 μm x 2 cm; Acclaim PepMap 100, P/N 164946; Thermo Fisher Scientific) at a flow rate of 2 $\mu\text{L}/\text{min}$ of solvent A (0.1% FA and 3% acetonitrile in LC-MS grade water). Peptides were eluted using a 120 min gradient from 5 to 40% (5% to 28% in 105 min followed by an increase to 40% B in 15 min) of solvent B (0.1% FA in 80% LC-MS grade acetonitrile) at a flow rate of 0.3 $\mu\text{L}/\text{min}$ and separated on a C18 analytical column (75 μm x 50 cm; PepMap RSLC C18; P/N ES803; Thermo Fisher Scientific). The Orbitrap Lumos was calibrated before each batch and 100 fmol of Pierce BSA protein digest (PI88341) was injected to control the performance of the LC and the mass spectrometer before the samples were acquired. The dates of sample runs are shown in **Table S1**.

LFQ: Peptides were electrosprayed using 2.1 kV voltage into the ion transfer tube (300°C) of the Orbitrap Lumos operating in positive mode. The Orbitrap first performed a full MS scan at a resolution of 120,000 FWHM to detect the precursor ion having m/z between 375 and 1,575 and a +2 to +7 charge. The Orbitrap automatic gain control (AGC) and the maximum injection time were set at 4×10^5 and 50 ms, respectively. The Orbitrap was operated using the top speed mode with a 3 sec cycle time for precursor

selection. The most intense precursor ions presenting a peptidic isotopic profile and having an intensity threshold of at least 5,000 were isolated using the quadrupole and fragmented by higher-energy collisional dissociation (HCD, 30% collision energy) in the ion routing multipole. The fragment ions (MS2) were analysed in the ion trap at a rapid scan rate. AGC and the maximum injection time were set at 1×10^4 and 35 ms, respectively, for the ion trap. Dynamic exclusion was enabled for 45 sec to avoid the acquisition of the same precursor ion having a similar m/z (plus or minus 10 ppm).

MS3-TMT: Analysis of TMT labelled peptide mixtures was carried out on the Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific) (control software Xcalibur™, version 4.0.21.10) using a data-dependent method with multi-notch synchronous precursor selection MS3 scanning for TMT tags as described previously¹. The Orbitrap was operated with a positive ion spray voltage of 2.1kV and a transfer tube temperature of 300C. The scan sequence began with an MS1 spectrum (Orbitrap analysis, resolution 120,000; 375–1,575 m/z , AGC target 1×10^4 , maximum injection time 50 ms). The maximum number of precursors within a 3 sec cycle time were fragmented using collision-induced dissociation (normalised collision energy 35%) and the fragmented ions were analysed in the ion trap (turbo mode; maximum injection time 50 ms). To quantify the TMT reporter ion, we collected MS3 spectra using a method in which the top 10 MS2 ions were fragmented using HCD (collision energy 65%). The MS3 were analysed in the Orbitrap (AGC 1×10^5 ; maximum injection time 105 ms; isolation window 2 m/z ; resolution 50,000; scan range m/z 100-500). All acquisition methods are available in a summary format, as Supplementary **Figures S2** (LFQ) and **S3** (TMT MS3).

Protein Identification and Quantitation

The raw spectral data were processed using MaxQuant (32) (version 1.6.5.0). For the LFQ data, the Andromeda search engine (33) integrated into the MaxQuant framework performed the spectral data search against a matched protein database (**Table S1**) composed of 12 genome-derived proteomes for the bacterial species (downloaded 26th November 2018 from NCBI) and 6 UniProtKB-derived protein databases for 5 fungal species and mouse (downloaded 30th December 2018). The MS3 spectra were searched against the mouse database.

For the MS data search, enzyme specificity was set to trypsin, allowing N-terminal cleavage to proline, and for ≤ 2 missed tryptic cleavages. Default settings were used for the MaxQuant searches, except that lysine acetylation and glutamate/glutamine conversion to pyroglutamic acid were set as variable modifications. N-terminal acetylation, methionine oxidation, and carbamidomethylation of cysteines was set as fixed modifications. The initial allowed mass deviation of the precursor ion was set to ≤ 20 ppm, and the allowed value for the fragment mass was set to ≤ 0.5 Da. Match between runs was used, with match time window 0.7 min and alignment window time 10 min. The maximum false discovery rates (FDR) at peptide and protein levels were set to 1%. Proteins LFQ intensities were determined by using the MaxLFQ algorithms (34), where the normalisation is applied on the whole dataset. This LFQ approach is based on accurate determination of spectrometric signal intensities (extracted ion chromatograms) of peptides and relies on measurements of the three-dimensional space of peptide ion intensity, m/z , and

chromatographic elution time. However, for proteins at low abundance, XICs are often contaminated by nearby signals, and although a protein can still be identified, it might not be quantified because of low-quality data. TMT-MS3 data were processed with MaxQuant using Reporter ion MS3 and TMT6plex-Lys126-131 internal labels, with the search setting described above.

Proteins Filtering and Functional Analyses

The MaxQuant output data were analysed by using Perseus (version 1.5.6.0) (38). Annotations for the identified proteins were downloaded from the UniProtKB database (35). Protein functional analyses, including metabolic pathway analysis, were performed using the DAVID (36) and STRING-db (37) tools. Information on the significant differentially produced proteins derived from jejunal tissue was used to query the DAVID knowledge database for pathway enrichment analysis, using the mouse genome as a reference list.

Filtering of protein identifications was performed as follows: first, we applied a default filtering of MaxQuant search results on proteins marked as “reverse”, “only identified by site”, and “potential contaminant”. Next, only proteins identified in at least two biological replicates were considered. To identify false-positive identifications of microbial proteins in mice groups not colonised with bacteria and/or fungi, and previously ascertained by microbiological assays and quantitative PCR (29), we checked the protein identification type and filtered out those identified by the “match between runs” algorithm of MaxQuant and not directly identified by MS. Also, proteins identified only by a single peptide and not identified by unique peptides were removed. Any other remaining proteins, whose origins were not consistent with the type of microbial colonisation of a specific mice group were checked at the peptide level, to confirm whether these are valid identifications or not. For protein quantification, we considered only proteins with LFQ intensities in at least two biological replicates and identified by two unique peptides.

Statistical Analyses

We performed LC-MS/MS analyses from 3-4 replicates of pooled fecal samples collected from mice belonging to each treatment group that underwent the same microbial colonisation and was housed in the same cage and gnotobiotic isolator. Jejunal tissue samples were derived from 20 animals (5 GF, 5 B, 5 BY, and 5 Y mice), and the TMT experimental design is described in **Figure S1**. To assess the biological variability of each experimental group, we calculated the Pearson correlation coefficients based on the protein intensities of each sample (**Table S2**). LFQ intensities were derived by the MaxLFQ algorithm, onto which we further applied a median-centring normalization strategy. Before the normalisation step and statistical analyses, the proteins LFQ intensities were \log_2 -transformed. To correct for differences in the sample amounts injected into LC-MS/MS, we normalised the relative protein amounts by dividing each protein LFQ intensity by the median intensity for all proteins in a given replicate (38) (**Figure S4**). To account for potential peptide loading differences in fractions of the 4 different 6-plex TMT batches, we applied a correction factor based on the pooled samples containing TMT⁶ label 131 (**Figure S1**). The

TMT spectra intensities were first median normalised as described above and then each relative protein amount was divided by the correction factor.

To identify proteins with levels that differ substantially among the strains, we performed analysis of variance (ANOVA) to compare the global mean level of each protein against its corresponding amount in each condition. For the ANOVA test, FDR calculations were performed in Perseus by a permutation-based procedure with 250 randomisations and a cut-off of 5%. To determine the exact pairwise differences in protein levels, Tukey's honestly significant difference (THSD) was performed on ANOVA-defined significant hits. Each species was analysed separately, and only relevant mice groups were used in the statistical analyses (e.g., only those colonised with bacteria were used to analyse bacterial proteins). Principal component analysis (PCA) was performed by using MetaboanalystR 3.0 (39), using log-transformed data normalised data by median. Hierarchical clustering analysis based on the protein levels was performed in Perseus, by using Euclidean distance, average linkage, no constraints, and pre-processing with k-means. Figures 1, and 5 were created using diagrams from BioRender.com. Data are available via ProteomeXchange with identifier PXD019355.

Results

Fecal metaproteome of defined gnotobiotic model.

Using a shotgun LFQ approach, we profiled the fecal metaproteomes of germ-free mice (GF), and mice colonised with either 12 bacterial species (B - bacteria), five fungal species (Y - yeast), both bacteria and fungi (BY), and the latter group treated with an antibiotic (BY_ABX) or antifungal (BY_AFX) (**Figure 1A**). We enriched microbial cells from feces of four-week-old mice by differential centrifugation, digested cell lysates by trypsin, and analysed the resulting peptide mixtures using LC-MS/MS. We queried the acquired mass spectra (5.3×10^6 MS/MS) against a combined protein database of the 17 microbial species and mouse (**Table S1**). The proteomic search identified 70,190 unique peptide sequences (**Table S3**), mapped to 6,675 proteins (**Table S4**). Of these, 68% were bacterial, 22% fungal, and 10% were mouse proteins (**Figure 1B**). Up to 58% of all detected proteins were characterised by LFQ intensities using the MaxLFQ method (34). Principal component analysis based on the protein relative levels identified three main clusters (**Figure 1C**), suggesting similar protein profiles of Y and GF, B and BY, and BY mice groups treated with antimicrobials (BY_ABX and BY_AFX).

From the bacterial proteomes, 4.4% of detected proteins shared peptides with other protein groups (**Table S5**). However, most of the identified peptides belonged to a single protein group (leading protein group); that is, the sequence coverage of peptides unique for the leading protein group was considerably higher than the sequence coverage of peptides belonging also to alternative protein groups. For the fungal proteomes, a high percentage of identified peptides belonged to multiple homologous proteins from different strains of the same species (e.g., *Candida albicans* strains SC5314, WO-1, CD36) (**Tables S5**). This was due to the use of UniProtKB reference protein databases for specific fungal species (**Table S1**), which included several strains of the same species. The use of reference databases was necessary

because the fungal strains used for the gnotobiotic mice colonisation were clinical isolates without sequenced genomes, contrary to the bacterial strains, which have their genome sequence determined (20). The shortcomings of using not perfectly matched protein databases were most apparent for two fungal strains (*Pichia kudriavzevii* and *C. krusei*), which were recently re-classified as strains of the same species (40) and did not have available individual databases. For these fungi, we used a common database based on reference *P. kudriavzevii* strains available in UniProtKB, resulting in decreased specificity of the MS data searches. Also, horizontal gene transfer might explain some of the homologous proteins; however, these events have been speculated to be relatively infrequent among microbial eukaryotes (41).

Using genome-specific protein databases for the bacterial strains also had the advantage of accurate assignment of proteins for the 12 different bacterial strains, for which we did not observe any species cross identifications. Nevertheless, the general UniProtKB databases did not compromise the accuracy of fungal and mouse protein identification, and we observed minimal cross-kingdom identifications (*e.g.*, same peptides mapped to a fungal and mouse protein), with these proteins being manually filtered out.

Inter-kingdom interactions influence the proteome response of gut bacterial species to antimicrobials.

From 12 bacterial strains used for the gnotobiotic mice colonisation, *Clostridium clostridioforme* YL32 had the most detected proteins (**Figure 1B**), while *Akkermansia muciniphila* YL44 had the highest predicted proteome coverage (33%) (**Table S5**). The latter result correlated with 16S rRNA amplicon sequencing data (18), which determined *A. muciniphila* YL44 as the most abundant bacterial species. The number of proteins detected per condition was relatively similar for each species (**Figure S5**).

From 2,860 quantified bacterial proteins, we identified 1,317 proteins whose levels significantly varied between the mice groups (ANOVA followed by THSD post-hoc analysis, FDR 5%, see **Table S6** for all proteins used for statistical analyses). For four bacterial species with the largest number of detected proteins (*C. clostridioforme* YL32, *A. muciniphila* YL44, *Blautia coccooides* YL58, *Muribaculum intestinale* YL27), about 50% of the quantified proteins showed differential levels (**Table S5**), prompting in-depth evaluation of the strains' protein profiles.

The proteomes of these bacterial species displayed an array of responses to antimicrobial-induced ecosystem perturbances. **Figure 2A** shows 50 of the most significantly changed proteins for four bacterial species with the largest number of detected proteins (see list of proteins and their functional annotation in **Table S7**). Antibiotic or antifungal treatment appeared to have a similar effect on the proteome of *A. muciniphila* YL44 (comparison of BY_ABX vs. BY_AFX groups), while for *M. intestinale* YL27, *B. coccooides* YL58, and *C. clostridioforme* YL32, the antibiotic treatment led to more proteins with increased levels compared to the antifungal groups (**Table S8**). For the *A. muciniphila* YL44, *M. intestinale* YL27, and *B. coccooides* YL58, many proteins had significantly increased levels in response to treatment with either antifungal or antibiotic (B/BY compared to ABX/AFX). For *C. clostridioforme* YL32, the detected proteome showed a contrasting response to the antimicrobial treatments. For instance, compared to the fecal proteome of the B only condition, there were 2.6 and 1.3 times less proteins with elevated levels

than in the BY_AFX and BY_ABX groups, respectively (**Table S8**). But when fungal species were present in the mouse gut (BY group), the effect of the antibiotic on *C. clostridioforme* YL32 appeared similar as for the other three strains (BY compared to BY_ABX), while treatment with antifungal led to a 1.3-fold increase in proteins with elevated levels in the BY group (BY compared to BY_AFX).

It should be noted that the effect of increased protein levels in the ABX/AFX groups could be further amplified by the MS detection method itself, which is less sensitive to low abundant proteins and will preferentially identify those with higher abundances (see **Table S1** - Summary of MS data search). In any case, these comparisons indicate that antimicrobial perturbation targeting bacterial or fungi cause differential responses among bacterial species that depend on species identity and presence or absence of fungi.

For the rest of the bacterial strains, low numbers of quantified proteins did not allow for in-depth investigations. Nonetheless, seven of the strains showed considerable differences in the protein profiles between the mice groups (**Figure S6**). A general picture has emerged from this comparison, in which antimicrobial treatments targeting bacteria or fungi during the second week of the mice's life resulted in lasting changes in the proteomes of the gut bacterial species.

Gut fungi differentially modulate bacterial proteomes.

Next, we statistically compared the bacterial protein levels between the B and BY groups to investigate the effect of fungi on the proteomes of the four strains with the highest numbers of detected proteins (**Figure 2B**). Similarly to what we observed for antimicrobials, fungal presence diversely influenced the proteomes of individual bacterial species. *A. muciniphila* YL44 and *M. intestinale* YL27 displayed increased amounts of proteins from various functional classes in the presence of fungi (**Figure S7**). We observed the opposite for *C. clostridioforme* YL32 and *B. coccoides* YL58, where the levels of 100 out of 107 and 51 out of 52 differentially produced proteins, respectively (t-test, FDR 5%), decreased when fungi were present (**Table S8**). The proteomic observations for *A. muciniphila* YL44 appeared to be inversely correlated with a decrease in *A. muciniphila* abundance based on 16S rRNA sequencing in the presence of fungi (18) (**Figure S7**). For *C. clostridioforme* YL32, the trend of increased protein levels in the B group compared to the BY group also appeared to inversely correlate with the 16S-based abundances; however, these trends did not reach statistical significance in the 16S sequencing data. Similarly, we could not derive a correlation with the sequencing data for *M. intestinale* YL27 and *B. coccoides* YL58, as the strains' relative abundance appeared similar across the mice groups. Metaproteomic analyses thus provided complementary information to the sequencing data and yielded a higher resolution of the ecological interactions between the gut microbial species.

Gut bacteria and antimicrobial treatments modulate fungal proteomes.

We detected 1,492 proteins for the five fungal species colonising the gnotobiotic mice, and 39% of these proteins were assigned LFQ levels. The lower numbers of proteins detected and quantified, as compared to bacteria, were a result of a lower abundance of fungi in the gut that we previously confirmed by

quantitative PCR and microbiological assays (18). Also, the LFQ approach used (34), is less sensitive to low abundant proteins, which are less likely to be quantified. Most of the identified fungal proteins were present in the Y group representing the mice colonised exclusively with fungi. We previously showed that the Y group harboured higher fungal concentration than co-colonised mice (18), confirming that a bacterial suppression of the fungal colonisation occurs in the host gut (42). Of the five fungal species, *C. glabrata* had the highest number of identified proteins, which also dominated the Y group (**Figure 3A**). However, antibiotic treatment (BY_ABX) led to higher detection of proteins derived from *Pichia kudriavzevii*/*C. krusei*. The total number of differentially abundant proteins between the mice groups and their distribution among the fungal species were rather similar for the following pairs of mice groups: Y - BY_ABX and BY - BY_AFX (**Figure 3B**). When we expressed the number of differentially produced fungal proteins as a percentage of each strain's detected proteins, we noticed that *C. glabrata* proteome showed a stronger response to the antibiotic treatment (BY_ABX), as compared to Y, BY and BY_AFX.

The presence of bacteria and early-life antimicrobials had significant effects on the fungal species protein profiles. Arrangement into categories using functional annotations based on UniProtKB (35) and other databases (36, 37), showed an increased number of fungal proteins related to stress responses in the antibiotic condition (**Figure 3C**), suggesting that the fungal species are either negatively affected by the changes in the bacterial microbiome or by the antibiotic treatment itself. Interestingly, many of the identified fungal proteins were previously reported to be excreted via extracellular vesicles (43, 44) and to have immunogenic properties (45, 46) (**Table S9**). Among these proteins were enzymes of the glycolytic pathway (*e.g.*, enolase, glyceraldehyde-3-phosphate dehydrogenase, fructose-bisphosphate aldolase, and phosphoglycerate mutase) and molecular chaperones linked to stress response (heat shock proteins). These proteins were significantly elevated in the fungal group suggesting that they are critical cytosolic proteins abundantly produced by fungal cells living in the mouse gut.

Fungal and bacterial colonisation induces distinct and persistent changes in the host fecal proteome.

We identified 674 mouse proteins as part of the mice fecal metaproteome. From proteins quantified by LFQ (405), 71% displayed differential abundance between the groups (ANOVA followed by THSD post-hoc analysis, FDR 5%, **Table S10**). PCA score plot based on the quantified proteins showed four clusters (**Figure 4A**): the first contained co-colonised mouse groups treated with antimicrobials (BY_ABX/AFX), and the second consisted of the B and BY groups. The third and fourth clusters included the Y and GF groups, and these were considerably more dissimilar from the first two clusters. Host proteins in the Y group were more clearly separated from the GF mice, in contrast to clustering when also microbial proteins were considered (**Figure 1C**), revealing that the proteome response to fungal colonisation may be more pronounced in the host proteome than in microbial cells. The GF group displayed the highest number of proteins associated with lipid metabolism, regulation, and molecular processing, while the BY and BY_AFX groups had the highest numbers of immune proteins (**Figure 4B**). In the BY_ABX group, we identified an increased number of stress response proteins. Moreover, the Y and GF groups displayed an increased number of proteins associated with energy metabolism (*e.g.*, mitochondrial proteins).

Proteins whose levels significantly varied between the mouse groups belong to the following functional classes: processing proteins such as proteases, hydrolases, and chaperones (14%), proteins linked to cell cycle regulation, motility, and cytoskeleton, including components of the tight junctions (14%), transport and membrane proteins including epithelial cell surface receptors (13%), proteins of energy and lipid metabolism (12% and 9%, respectively), and gut immune and barrier factors (9%). To tease apart the mouse proteome response to different microbes, we statistically compared pairs of treatment groups containing either fungi or bacteria (**Figure S8**). The introduction of the 12-species bacterial consortium led to a striking decrease in the mouse protein levels, and this was valid both for germ-free mice (GF vs. B) and mice colonised with fungi (Y vs. BY). Although gut fungi did not have the same significant effect on the mouse proteome as bacteria, we detected a 5-fold decrease in the number of proteins whose levels increased in the Y group compared to GF (GF vs. Y). The presence of fungi and bacteria resulted in a 2-fold increase compared to bacteria alone (BY vs. B), particularly of proteins associated with the immune system, energy metabolism, and cellular processing.

Because the mouse protein levels were strongly dependent on the gut microbial consortium, we gave special attention to the individual functional classes of differentially produced proteins. **Figure 4C** shows 45 differentially produced mouse proteins with the largest variation between the mice groups. Levels of 27 immune proteins were the highest in GF mice, followed by those colonised only with fungi. Two immune proteins were elevated in the presence of fungi: acid mammalian chitinase (Chia), implicated in the defense response against fungi, and regenerating islet-derived protein 3-gamma (Reg3g), a bactericidal C-type lectin reported to have bactericidal activity (47, 48). Thus, our results suggest that Reg3g might also be involved in the intestinal cells' response to fungal colonisation. However, more data will be needed to support the involvement of Reg3g in the epithelial cell response to fungi.

Bacterial presence resulted in a reduction in the mucosal Pentaxin (Mtx2) levels, a secreted protein involved in complement activation, Intelectin 1 (Itln1), a receptor that binds microbial glycans, and Dipeptidyl peptidase 4 (Dpp4), cell surface receptor and dipeptidyl protease involved in T-cell activation. In addition, levels of other immune proteins, such as Mucin 2 (Muc2), Poly-Ig receptor (Pigr), and immunoglobulin heavy chain (IgH), were altered by the shifts in microbial composition caused by antimicrobial treatments.

The absence of microbes resulted in elevated levels of mouse proteins linked to carbohydrate metabolism both at the cellular [*e.g.*, Pyruvate Kinase (Pkr)] and host level [glycosidases such as Pancreatic Alpha-Amylase (Amy2), Lactase (Lct), and Trehalase (Treh)]. Amy2, a glycosidase that hydrolyses alpha-linked polysaccharides such as starch and glycogen, had the highest levels in the Y group, suggestive of increased Amy2 production response to either fungal polysaccharides or fungal metabolism on dietary sugars. Fungi also elicited a strong effect on host proteins associated with lipid metabolism. The highest levels were identified in the Y group for the host enzyme Colipase (Clps), a cofactor of pancreatic lipase facilitating lipids digestion, and Phospholipase A2 (Pla2g1b), a secreted protein involved in lipid degradation and innate immune mucosal response. Closely connected to the immune system are phospho- and sphingolipid metabolism enzymes, which often have regulatory

properties towards immune cells. Of those, we identified Neutral Ceramidase (Asah2), a membrane protein hydrolysing sphingolipid ceramides into sphingosines and free fatty acids, whose relative levels were the lowest in BY and B groups.

Cytoskeletal and cell adhesion proteins were also significantly increased in the GF and Y mice. These included cadherin-related proteins 2 and 5 (Cdhr2 and Cdhr5), which regulate microvilli length, and Claudins 3 and 7 (Cldn3 and Cldn7), controlling tight junction-specific obliteration of the intercellular space. Notably, Keratin-33A and -81 (Krt33a and Krt81), structural proteins that form the cytoskeleton's intermediate filaments in epithelial cells, displayed the highest levels in the Y group, followed by the B group and had the lowest levels in the BY groups treated with antimicrobials. Adhesion proteins with regulatory properties and links to the immune system included Tetraspanin-8 (Tspan 8), Epithelial Glycoprotein 314 (Epcam), and Basigin (Bsg), a cell surface receptor. All three proteins showed increased levels in the GF and Y groups.

One of the most prevalent functional classes among the differentially produced proteins were proteases, which had the highest levels in the GF and Y groups, and the lowest in the BY and B groups. Among these was for example Angiotensin-converting enzyme (Ace2), a carboxypeptidase with multiple regulatory functions, including gap junction assembly. Other proteases with regulatory properties included membrane-bound metalloproteases Meprin A (Mep1a) and B (Mep1b), implicated in the inflammatory response. Alongside proteases, protease inhibitors followed the same quantitative pattern as described above. These included Serpins, serine protease inhibitors that negatively regulate endopeptidase activity in response to cytokines (Serp1b1a, Serp1a1d), innate immune response, inflammation, and cellular homeostasis (Serp1b1a, Serp1b6).

Numerous regulatory proteins were present among the proteins with increased levels in the GF and Y group, such as the Annexin family of Ca²⁺-regulated phospholipid-binding and membrane-binding proteins (Anxa4, Anxa11, Anxa13) and nuclear proteins such as Onzin (Plac8) suggested to regulate immune responses. Overall, the metaproteomic analyses documented an extensive impact of microbial colonisation in a controlled early life gut microbiome model, underlining the intertwined functional development of the host and its gut microbiome, and revealing new features of the host response to fungal colonisation.

Fungal colonisation drives alterations in the mouse jejunal tissue proteome.

Given the exciting findings from the mice fecal metaproteomes, we decided to investigate more closely the host proteome. We selected the small intestine as our site of interest because it contains the body's largest immune organ (the gut-associated lymphoid tissue) and is an important site of the host-microbe interactions (49). We used a tandem mass tag labelling approach (TMT 6-plex) to gain a higher sensitivity for low abundant proteins (**Figure 5A**). From the TMT 6-plex experiment, we identified 10,201 peptides (**Table S11**) matched to 1,514 mouse proteins (**Table S12**), and the levels of 45 were significantly altered between the mice groups (ANOVA, THSD post-hoc analysis, FDR 5%, **Table S13**). PCA score plot revealed different grouping results than for the mouse fecal proteomes (**Figure 5B**).

Functionally, the 45 significantly changed proteins were mainly associated with cellular metabolism and regulation of cellular processes (**Figure 5C**). Pathway analysis identified oxidation-reduction, glutathione metabolism and cell-cell adherence as significantly enriched processes (**Table S14**).

Compared to the fecal mouse proteome, proteome changes of the jejunal tissue related to microbial colonisation were more subtle. **Figure 5D** shows the relative levels of 45 differentially produced mouse jejunal proteins with significant variation between the mice groups. Among immune proteins whose levels increased with the presence of fungi were Alpha-defensin 2 (Def20) and Ig gamma-1 chain C region secreted form (Ighg1). From 10 proteins functionally connected to lipid metabolism, two Glutathione S-transferase enzymes (Gstm2 and Gsta4) were significantly decreased in the fungal group. We also detected two proteins of retinol metabolism (Retsat and Aldh1a1), which play a key role in mucosal immune responses. Fungal colonisation further appeared to impact the intestinal cells' energy metabolism, as several mitochondrial proteins had decreased levels in the fungi group, most notably, two subunits of Cytochrome C oxidase (Cox5b and Cox7a1).

Some of the differentially produced proteins have functional links to NF- κ B pathway, such as the apoptotic marker Poly (ADP-ribose) polymerase (Parp1). In some contexts, particularly in response to cellular stress, stimulation of the NF- κ B pathway promotes apoptosis (50). Selenium acts as a key element that controls NF- κ B activation and the half-life of its inhibitor I κ B α , and we detected two selenium-binding proteins (Selenbp1 and Selenbp2). Overall, the TMT 6-plex analysis of the jejunal tissue provided further insight on host cellular pathways impacted by defined microbial colonisation and confirmed that gut fungi elicit differential effects compared to bacteria.

Discussion

In this work, we used a controlled environment of gnotobiotic mice colonised with defined microbial consortia to precisely describe changes in the metaproteome profiles associated with alterations in the gut microbiome. By reducing the gut microbiome's complexity to 12- bacterial species, five fungal species, or the combination of the 17 microbes, the gnotobiotic model allowed us to isolate specific contributions and cross-kingdom interactions between the gut-associated bacteria and fungi, as well as the host response to the microbial colonisation.

Several normalisation methods for quantitative metaproteomic data have been suggested, including generation of a mock community, and normalisation of protein levels for each sample and species to a constant value (51). The latter method depends on a sufficient number of measurements for each species to be normalised. We described in our earlier study (29), using DNA sequencing data, that some of the bacterial and fungal species are in low abundance. For that reason, we chose not to normalise the data for each species to a constant value because only very few proteins would be considered for the low abundant species and possibly skew the data in favour of the more abundant species. However, it is possible to avoid normalisation to species if the reference protein sequences of the organisms under study are used for generating quantitative proteomic data (51). Indeed, we took advantage of the

genome-sequenced bacterial consortium (20) and used matched protein databases for MS data searches. This strategy yielded an accurate and specific identification of proteins from the 12 bacterial species. For fungal strains that did not have sequenced genomes, the use of UniProtKB reference protein databases resulted in decreased specificity of the MS data searches, yet still correct assignment of fungal proteins to specific species. In this way, we avoided the issue of protein identifications shared between two or more species, which would compromise the accuracy of the determination of the protein's relative quantitative levels and prevent its use in subsequent statistical analysis. Moreover, the used MaxLFQ normalisation method (34) have been applied in other metaproteomic studies for accurate label-free quantification of proteins (52).

The fungal species chosen for the gnotobiotic experiment are amongst the ones commonly reported from human samples (1, 27, 53). Still, we acknowledge that a limited number of species was investigated, and the effect of other important fungi (*Malassezia* sp., *Saccharomyces* sp.) remains to be tested. Our previous work determined the impact of the selected gut fungal species on the structure of gut microbiome community and host immune development (18). Here, we strengthened these results using a metaproteomic approach that provided an independent functional measure of the microbiome-host crosstalk. Similarly to other multi-omics studies (54, 55), we found discrepancies between the quantitative proteome profiles and the DNA-based relative abundance of individual species. These differences likely stem from complex regulatory networks along the gene to protein expression path, which differ between microbial kingdoms. The latter was recently documented in a study showing dynamic ratios between protein and RNA levels that depended on the type of microbial population (56).

For the fungal species, a meaningful comparison between the protein and DNA data was hindered by the absence of genome-specific protein databases that would guide the raw MS data searches. However, corresponding with our earlier results (18), the proteomic data showed that fungi grew in higher concentrations in bacteria's absence. This finding is in line with the interkingdom competition and antagonism between bacteria and fungi in other ecosystems, such as the rhizosphere and soil (57), and in mammalian hosts, where commensal bacteria limit fungal colonisation via activation of innate mucosal immunity (42) and directly by producing inhibitory metabolites such as short-chain fatty acids (58).

Despite the lack of a linear relationship between the DNA and protein data, our earlier results from the 16S rRNA sequencing aided the interpretation of the bacterial proteome responses, and the combined analyses revealed antagonistic and synergistic relationships between the bacterial and fungal species. For example, we identified *Lactobacillus reuteri* as a potent responder to the fungal presence from the sequencing data, as the bacterium abundance was significantly reduced in the BY group. Only a few *L. reuteri* proteins were quantified, and those displayed lower amounts in the absence of fungi, conversely to the DNA sequencing results (**Figure 6A**). Some of the most abundant bacterial proteins were detected in the B-only group (ribosomal protein L31, enolase, and translational factor Tu) along with three dehydrogenases, whose increased production often indicates cellular stress. For another low abundant bacterial species, *Clostridium innocuum*, we observed a similar reciprocal relation between the proteomic

and sequencing data (**Figure 6B**). Taken together with the observed variations in the proteome response of the four most abundant bacterial species (**Figure 2**), the proteomic data is suggestive that the cells are affected by changes in the gut environment related to the fungal presence (*L. reuteri*, *A. muciniphila*, *M. intestinale*) or absence (*C. innocuum*, *B. coccoides*, *C. clostridioforme*). Similar results were found in samples from mice treated with antimicrobials, which impacted levels of numerous proteins and often led to their elevated production (**Figure 2 and S6**). Such proteome response has been documented for bacterial cells treated with antibiotics (59), where the cells optimise their proteins' production to deal with the adverse environmental factor. Moreover, protein aggregation is a likely bacterial strategy to survive antibiotic treatment (60).

One of this study's most striking findings was a similar impact of the antibacterial and antifungal treatments on bacterial and fungal species' proteomes. Inhibitory effects of antifungal drugs on the growth of commensal bacteria are poorly described (61), but it may be that the antifungal Fluconazole inhibits intestinal bacteria to some extent via direct or indirect mechanisms, similarly to other non-antibiotic drugs (62). The antibiotic and antifungal treatments also had a strong effect on the quantities and types of fungal proteins being produced, revealing how these antimicrobial drugs impact the entire community, with similar functional consequences. Ecosystem perturbances caused by either antibiotics or antifungal drugs may impact the gut microbiome keystone species, causing a widespread effect on microbial networks and food webs (63).

Host proteins present in the intestine are essential for maintaining a mutualistic relationship with the microbes on the mucosal interface and serve as reporters on the host-microbe interactions. Host fecal proteome has been reported to exhibit signatures specific to colonisation states (64). However, to our knowledge, protein response to exclusive fungal colonisation has not been characterised before. For the first time, we described at the proteome level the cellular pathways and their protein components involved in the interactions between the mammalian host and gut fungal species. Fungal colonisation resulted in changes in host proteins functional in innate immunity as well as metabolism (**Figure 4**), suggesting specific roles of gut fungi on host systems during early developmental stages. Further research aimed at investigating these roles has great potential for novel discovery, given that most host-microbiome interactions described to date have been limited to bacteria.

In the host jejunal proteome, most of the differentially produced proteins had metabolic functions (**Figure 5**). The quantitative profiles of proteins with significantly changed levels showed closer clustering of the GF and B groups (**Figure 5D**). These results likely reflect the host control of the number of bacterial cells in the small intestine, where they would otherwise compete for easily digestible nutrients. Alternatively, the more subtle differences in jejunal proteomes across microbial colonisation conditions, as compared to profound changes identified in fecal metaproteomes, may also reflect host homeostatic pressures present in intestinal tissues. Nonetheless, these findings indicate that fungal colonisation influences host jejunal proteome in a distinct way from bacteria.

The identification of multiple fungal proteins hypothesised to be abundant components of the extracellular vesicles suggests that fungi influence the host through direct cell-to-cell contact. Possible mechanisms may be akin to recently described interactions between segmented filamentous bacteria and mice intestinal epithelial cells (65). These bacteria protruded into the epithelial cells and used adhesion-triggered endocytosis to transfer antigens into intestinal epithelial cells and modulate host T cell homeostasis. Similarly, epithelial internalisation of fungal hyphae, a morphology into which fungal cells transition to strengthen their adherence to epithelial cells (66), can deploy fungal extracellular vesicles with host cell-modulatory properties. Fungal cells produce a diverse array of biologically active compounds (67, 68), but the understanding of the mycobiome contribution to immunomodulatory substances present in the gut remains rudimentary.

It should be noted that the results of this experimental study need to be interpreted with caution, since only a limited number of microbial species was included. We used a selection of prevalent fungal species of the mammalian gut, yet many more remain to be explored and characterized. In addition, the complexity of interactions between the host and the gut microbiome, as well as within the different microbial kingdoms, will likely increase exponentially with an increasing number of species. Other limitations of this work include the lack of genome references for the fungal strains colonising the mice. Further efforts to carry out whole genome sequencing of fungal species and strains commonly associated with mammalian gut will reveal a better-resolved analysis of the functional role of these understudied microbes to their microbial ecosystems and their impact on the host. Also, fecal samples are a proxy for the colon microbiome. Although technically challenging, sampling along the gastrointestinal tract will reveal additional insights on microbial colonisation of the small and large intestine. Finally, label-free quantification is a common strategy adopted in metaproteomic studies but has low sensitivity to more scarce proteins as well as limitations in quantification accuracy. The application of metabolic labelling for improved peptide quantification (69) holds the potential to improve the accuracy of quantitative metaproteomics, and promote the application of proteomics for functional studies of intestinal microbiomes.

Conclusions

In this work, we report previously unknown interactions between specific gut bacteria, fungi, and a mammalian host by using quantitative proteomic analyses of gnotobiotic mice colonised with defined microbial communities. Broad changes in microbial proteomes reflected interkingdom interactions between bacteria and fungi, as well as a response to antimicrobials. Our data also described for the first time the key role of fungal colonisation and how it impacts the host intestinal proteome. Further, we characterised cellular pathways and their protein components involved in the interactions between the mammalian host and gut fungal species. Our results suggest that an increased abundance of certain fungal species in early life may impact the developing intracellular balance of epithelial and immune cells. The work sets the stage for future studies that will explore the details of molecular mechanisms by which gut fungi modulate host physiology.

Declarations

Ethics approval: This study does not involve the use of human subjects. The ethic statements regarding animal experimentation are included in the Methods section.

Consent for publication: Not applicable.

Availability of data: Data are available via ProteomeXchange with identifier PXD019355.

Competing Interests statement: The authors declare no competing interest.

Funding: This work was funded by The Research Council of Norway Grant No. 274296 (V.K.P). This work was also supported by the Cumming School of Medicine, the Alberta Children Hospital Research Institute, the Snyder Institute of Chronic Diseases, the Canadian Institutes for Health Research, the Sick Kids Foundation, and W. Garfield Weston Foundation (M.C.A). A.D. was supported by an NSERC Discovery Grant (DGEGR-2019-00112). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions: Conceptualization: all. Data curation: VKP. Formal analysis: VKP. Funding acquisition: all. Investigation: VKP, MCA. Project administration: VKP, MCA. Visualisation: VKP. Writing—original draft: VKP. Writing—review and editing: all. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Erik van Tilburg Bernardes and Isabelle Laforest-Lapointe for providing the mice fecal samples, and Ali Javed for a contribution to the fungal proteomes' analyses, all at the time at the Arrieta lab. We also thank Daniel Young from the Dufour lab for assistance with the proteomic sample preparations. We acknowledge the Southern Alberta Mass Spectrometry Facility and Laurent Brechenmacher for the service and support for the LC-MS/MS experiments. We also thank the reviewers for their constructive comments that have led to an improved manuscript

References

1. Nash AK, Auchtung TA, Wong MC, Smith DP, Gesell JR, Ross MC, et al. The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome*. 2017;5(1):153.
2. Zhai B, Ola M, Rolling T, Tosini NL, Joshowitz S, Littmann ER, et al. High-resolution mycobiota analysis reveals dynamic intestinal translocation preceding invasive candidiasis. *Nat Med*. 2020;26(1):59-64.
3. Bacher P, Hohnstein T, Beerbaum E, Röcker M, Blango MG, Kaufmann S, et al. Human Anti-fungal Th17 Immunity and Pathology Rely on Cross-Reactivity against *Candida albicans*. *Cell*. 2019;176(6):1340-55.e15.

4. Sokol H, Leducq V, Aschard H, Pham HP, Jegou S, Landman C, et al. Fungal microbiota dysbiosis in IBD. *Gut*. 2017;66(6):1039-48.
5. Yeung F, Chen Y-H, Lin J-D, Leung JM, McCauley C, Devlin JC, et al. Altered Immunity of Laboratory Mice in the Natural Environment Is Associated with Fungal Colonization. *Cell Host Microbe*. 2020;27(5):809-22.e6.
6. Mohanta TK, Bae H. The diversity of fungal genome. *Biol Proced Online*. 2015;17(1):8.
7. Carpenter SR, Kitchell JF, Hodgson JR. Cascading Trophic Interactions and Lake Productivity. *Bioscience*. 1985;35(10):634-9.
8. Paine RT. Food Web Complexity and Species Diversity. *The American Naturalist*. 1966;100(910):65-75.
9. Sih A, Crowley P, McPeck M, Petranka J, Strohmeier K. Predation, Competition, and Prey Communities: A Review of Field Experiments. *Annual Review of Ecology and Systematics*. 1985;16(1):269-311.
10. Zhang Y, Kastman EK, Guasto JS, Wolfe BE. Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. *Nature Communications*. 2018;9(1):336.
11. Deveau A, Bonito G, Uehling J, Paoletti M, Becker M, Bindschedler S, et al. Bacterial–fungal interactions: ecology, mechanisms and challenges. *FEMS Microbiol Rev*. 2018;42(3):335-52.
12. Alzarhani AK, Clark DR, Underwood GJC, Ford H, Cotton TEA, Dumbrell AJ. Are drivers of root-associated fungal community structure context specific? *The ISME Journal*. 2019;13(5):1330-44.
13. Noverr MC, Noggle RM, Toews GB, Huffnagle GB. Role of antibiotics and fungal microbiota in driving pulmonary allergic responses. *Infection and immunity*. 2004;72(9):4996-5003.
14. Dollive S, Chen YY, Grunberg S, Bittinger K, Hoffmann C, Vandivier L, et al. Fungi of the murine gut: episodic variation and proliferation during antibiotic treatment. *PLoS One*. 2013;8(8):e71806.
15. Kim Y-G, Udayanga KGS, Totsuka N, Weinberg JB, Núñez G, Shibuya A. Gut dysbiosis promotes M2 macrophage polarization and allergic airway inflammation via fungi-induced PGE₂. *Cell Host Microbe*. 2014;15(1):95-102.
16. Wheeler ML, Limon JJ, Bar AS, Leal CA, Gargus M, Tang J, et al. Immunological Consequences of Intestinal Fungal Dysbiosis. *Cell Host Microbe*. 2016;19(6):865-73.
17. Hoarau G, Mukherjee PK, Gower-Rousseau C, Hager C, Chandra J, Retuerto MA, et al. Bacteriome and Mycobiome Interactions Underscore Microbial Dysbiosis in Familial Crohn's Disease. *mBio*. 2016;7(5):e01250-16.
18. van Tilburg Bernardes E, Kuchařová Pettersen V, Gutierrez MW, Laforest-Lapointe I, Jendzjowsky NG, Cavin J-B, et al. Intestinal fungi are causally implicated in microbiome assembly and immune development in mice. *Nature Communications*. 2020;11(1):2577.
19. Brugiroux S, Beutler M, Pfann C, Garzetti D, Ruscheweyh H-J, Ring D, et al. Genome-guided design of a Defined Mouse Microbiota that Confers Colonization Resistance Against *Salmonella Enterica* Serovar Typhimurium. *Nature Microbiology*. 2016;2:16215.

20. Uchimura Y, Wyss M, Brugiroux S, Limenitakis JP, Stecher B, McCoy KD, et al. Complete Genome Sequences of 12 Species of Stable Defined Moderately Diverse Mouse Microbiota 2. Genome announcements. 2016;4(5):e00951-16.
21. Li J, Chen D, Yu B, He J, Zheng P, Mao X, et al. Fungi in Gastrointestinal Tracts of Human and Mice: from Community to Functions. *Microb Ecol.* 2018;75(4):821-9.
22. Ward TL, Dominguez-Bello MG, Heisel T, Al-Ghalith G, Knights D, Gale CA. Development of the Human Mycobiome over the First Month of Life and across Body Sites. *mSystems.* 2018;3(3).
23. Arrieta MC, Arevalo A, Stiemsma L, Dimitriu P, Chico ME, Loor S, et al. Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. *J Allergy Clin Immunol.* 2018;142(2):424-34 e10.
24. Fujimura KE, Sitarik AR, Havstad S, Lin DL, Levan S, Fadrosch D, et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med.* 2016;22(10):1187-91.
25. Huttner A, Bielicki J, Clements MN, Frimodt-Møller N, Muller AE, Paccaud JP, et al. Oral amoxicillin and amoxicillin–clavulanic acid: properties, indications and usage. *Clinical Microbiology and Infection.* 2020;26(7):871-9.
26. Arrieta M-C, Arévalo A, Stiemsma L, Dimitriu P, Chico ME, Loor S, et al. Associations Between Infant Fungal and Bacterial Dysbiosis and Childhood Atopic Wheeze in a Nonindustrialized Setting. *Journal of Allergy and Clinical Immunology.* 2018;142(2):424-34.e10.
27. Fujimura KE, Sitarik AR, Havstad S, Lin DL, Levan S, Fadrosch D, et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med.* 2016;22:1187.
28. Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med.* 2015;7(307):307ra152.
29. van Tilburg Bernardes E, Pettersen VK, Gutierrez MW, Laforest-Lapointe I, Jendzjowsky NG, Cavin J-B, et al. Intestinal fungi are causally implicated in microbiome assembly and immune development in mice. *Nature Communications.* 2020;11(1):2577.
30. Tanca A, Palomba A, Pisanu S, Deligios M, Fraumene C, Manghina V, et al. A straightforward and efficient analytical pipeline for metaproteome characterization. *Microbiome.* 2014;2(1):49.
31. Wiśniewski JR. Chapter Two - Filter-Aided Sample Preparation: The Versatile and Efficient Method for Proteomic Analysis. In: Shukla AK, editor. *Methods Enzymol.* 585: Academic Press; 2017. p. 15-27.
32. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc.* 2016;11(12):2301-19.
33. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J Proteome Res.* 2011;10(4):1794-805.
34. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol Cell Proteomics.* 2014;13(9):2513-26.

35. The UniProt C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 2019;47(D1):D506-D15.
36. McAlister GC, Nusinow DP, Jedrychowski MP, Wühr M, Huttlin EL, Erickson BK, et al. MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal Chem.* 2014;86(14):7150-8.
37. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):gky1131.
38. Välikangas T, Suomi T, Elo LL. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Brief Bioinform.* 2018;19(1):1-11.
39. Pang Z, Chong J, Li S, Xia J. MetaboAnalystR 3.0: Toward an Optimized Workflow for Global Metabolomics. *Metabolites.* 2020;10(5):186.
40. Douglass AP, Offei B, Braun-Galleani S, Coughlan AY, Martos AAR, Ortiz-Merino RA, et al. Population genomics shows no distinction between pathogenic *Candida krusei* and environmental *Pichia kudriavzevii*: One species, four names. *PLoS Pathogens.* 2018;14(7):e1007138.
41. Andersson JO. Horizontal Gene Transfer Between Microbial Eukaryotes. In: Gogarten MB, Gogarten JP, Olendzenski LC, editors. *Horizontal Gene Transfer: Genomes in Flux.* Totowa, NJ: Humana Press; 2009. p. 473-87.
42. Fan D, Coughlin LA, Neubauer MM, Kim J, Kim MS, Zhan X, et al. Activation of HIF-1 α and LL-37 by commensal bacteria inhibits *Candida albicans* colonization. *Nat Med.* 2015;21(7):808-14.
43. Vargas G, Rocha JD, Oliveira DL, Albuquerque PC, Frases S, Santos SS, et al. Compositional and immunobiological analyses of extracellular vesicles released by *Candida albicans*. *Cell Microbiol.* 2015;17(3):389-407.
44. Gil-Bona A, Llana-Palacios A, Parra CM, Vivanco F, Nombela C, Monteoliva L, et al. Proteomics unravels extracellular vesicles as carriers of classical cytoplasmic proteins in *Candida albicans*. *J Proteome Res.* 2015;14(1):142-53.
45. Bitencourt TA, Rezende CP, Quaresimin NR, Moreno P, Hatanaka O, Rossi A, et al. Extracellular Vesicles From the Dermatophyte *Trichophyton interdigitale* Modulate Macrophage and Keratinocyte Functions. *Front Immunol.* 2018;9:2343-.
46. Oliveira DL, Freire-de-Lima CG, Nosanchuk JD, Casadevall A, Rodrigues ML, Nimrichter L. Extracellular Vesicles from *Cryptococcus neoformans* Modulate Macrophage Functions. *Infection and Immunity.* 2010;78(4):1601.
47. Vaishnava S, Yamamoto M, Severson KM, Ruhn KA, Yu X, Koren O, et al. The Antibacterial Lectin RegIII γ Promotes the Spatial Segregation of Microbiota and Host in the Intestine. *Science.* 2011;334(6053):255-8.
48. Chen Z, Downing S, Tzanakakis ES. Four Decades After the Discovery of Regenerating Islet-Derived (Reg) Proteins: Current Understanding and Challenges. *Frontiers in Cell and Developmental Biology.* 2019;7(235).

49. Kastl AJ, Terry NA, Wu GD, Albenberg LG. The Structure and Function of the Human Small Intestinal Microbiota: Current Understanding and Future Directions. *Cell Mol Gastroenterol Hepatol*. 2020;9(1):33-45.
50. Khandelwal N, Simpson J, Taylor G, Rafique S, Whitehouse A, Hiscox J, et al. Nucleolar NF- κ B/RelA mediates apoptosis by causing cytoplasmic relocalization of nucleophosmin. *Cell Death Differ*. 2011;18(12):1889-903.
51. Kleiner M. Normalization of metatranscriptomic and metaproteomic data for differential gene expression analyses: The importance of accounting for organism abundance. *PeerJ Preprints*. 2017;5:e2846v1.
52. Zhang X, Deeke SA, Ning Z, Starr AE, Butcher J, Li J, et al. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nature Communications*. 2018;9(1):2873.
53. Wampach L, Heintz-Buschart A, Hogan A, Muller EEL, Narayanasamy S, Laczny CC, et al. Colonization and Succession within the Human Gut Microbiome by Archaea, Bacteria, and Microeukaryotes during the First Year of Life. *Front Microbiol*. 2017;8(738).
54. Kleiner M, Thorson E, Sharp CE, Dong X, Liu D, Li C, et al. Assessing species biomass contributions in microbial communities via metaproteomics. *Nature Communications*. 2017;8(1):1558.
55. Timmins-Schiffman E, White SJ, Thompson RE, Vadopalas B, Nunn BL, Roberts SB. Coupled Microbiome Analyses Highlights Relative Functional Roles of Bacteria in a Bivalve Hatchery. *Environmental Microbiome*. 2020:preprint under consideration
56. Delogu F, Kunath BJ, Evans PN, Arntzen MØ, Hvidsten TR, Pope PB. Integration of absolute multi-omics reveals dynamic protein-to-RNA ratios and metabolic interplay within mixed-domain microbiomes. *Nature Communications*. 2020;11(1):4708.
57. Boer W, Folman LB, Summerbell RC, Boddy L. Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiol Rev*. 2005;29(4):795-811.
58. Guinan J, Wang S, Hazbun TR, Yadav H, Thangamani S. Antibiotic-induced decreases in the levels of microbial-derived short-chain fatty acids correlate with increased gastrointestinal colonization of *Candida albicans*. *Sci Rep*. 2019;9(1):8872.
59. Lin MH, Potel CM, Tehrani K, Heck AJR, Martin NI, Lemeer S. A New Tool to Reveal Bacterial Signaling Mechanisms in Antibiotic Treatment and Resistance. *Mol Cell Proteomics*. 2018;17(12):2496-507.
60. Bollen C, Dewachter L, Michiels J. Protein Aggregation as a Bacterial Strategy to Survive Antibiotic Treatment. *Frontiers in Molecular Biosciences*. 2021;8(259).
61. Zarf S, Arman I. Effect of Fluconazole, An Anti-Fungal Drug, On Human Flora Bacteria. *Journal of Bacteriology and Mycology* 2020;7(2):1130.
62. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*. 2018;555(7698):623-8.
63. Herren CM, McMahon KD. Keystone taxa predict compositional change in microbial communities. *Environ Microbiol*. 2018;20(6):2207-17.

64. Lichtman JS, Marcobal A, Sonnenburg JL, Elias JE. Host-centric proteomics of stool: a novel strategy focused on intestinal responses to the gut microbiota. *Molecular & cellular proteomics* : MCP. 2013;12(11):3310-8.
65. Ladinsky MS, Araujo LP, Zhang X, Veltri J, Galan-Diez M, Soualhi S, et al. Endocytosis of commensal antigens by intestinal epithelial cells regulates mucosal T cell homeostasis. *Science*. 2019;363(6431):eaat4042.
66. Nikou S-A, Kichik N, Brown R, Ponde NO, Ho J, Naglik JR, et al. *Candida albicans* Interactions with Mucosal Surfaces during Health and Disease. *Pathogens (Basel, Switzerland)*. 2019;8(2):53.
67. Keller NP. Fungal secondary metabolism: regulation, function and drug discovery. *Nature Reviews Microbiology*. 2019;17(3):167-80.
68. Freitas MS, Bonato VLD, Pessoni AM, Rodrigues ML, Casadevall A, Almeida F. Fungal Extracellular Vesicles as Potential Targets for Immune Interventions. *mSphere*. 2019;4(6):e00747-19.
69. Zhang X, Ning Z, Mayne J, Deeke SA, Li J, Starr AE, et al. In Vitro Metabolic Labeling of Intestinal Microbiota for Quantitative Metaproteomics. *Anal Chem*. 2016;88(12):6120-5.

Figures

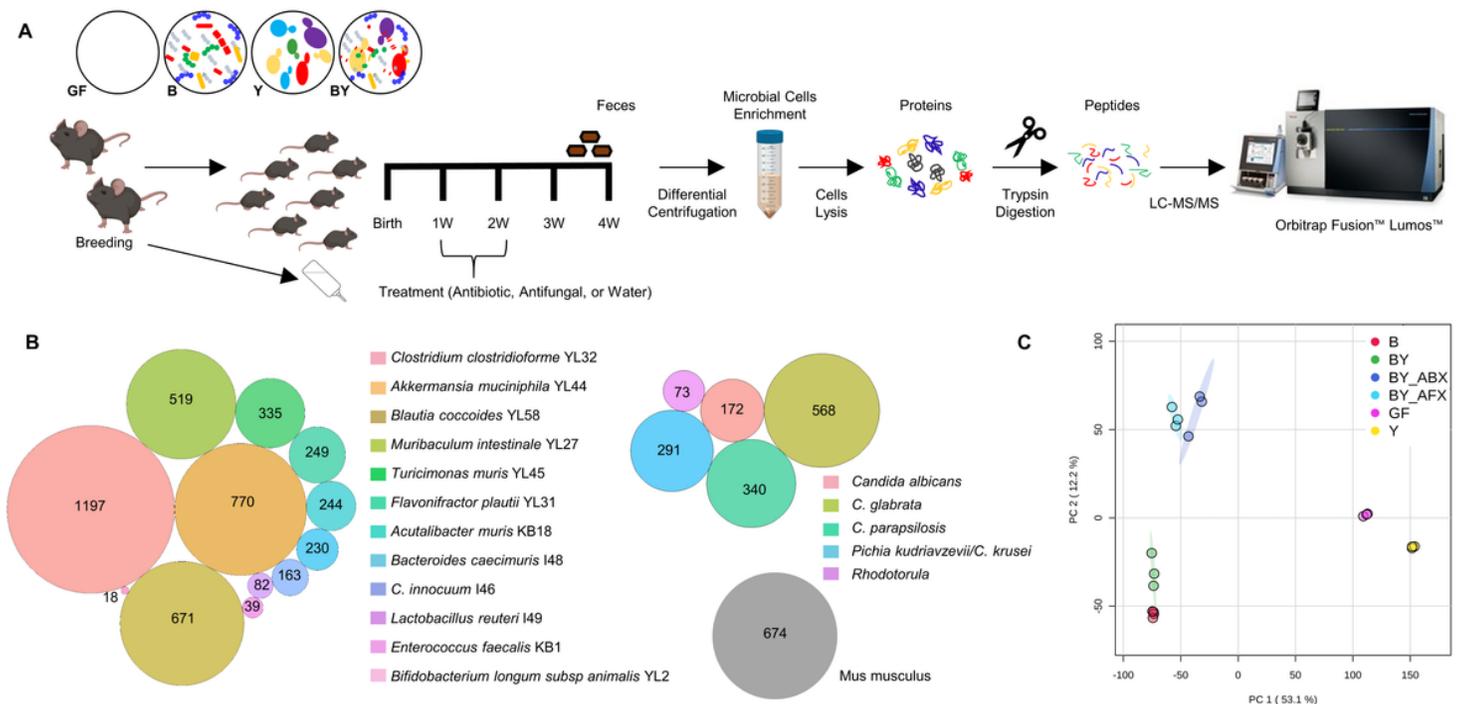


Figure 1

Description of a gnotobiotic colonisation model by label-free quantitative metaproteomics. A)

Experimental design of shotgun label-free proteomic experiment, including enrichment of fecal microbial cells by differential centrifugation. B) Detected fecal metaproteome composed of proteins originating from 12 bacterial, 5 fungal species, and mouse. C) PCA score plot based on the relative levels of 3,845

proteins quantified for mouse, bacteria, and fungi/yeast. X- and Y-axis show the first and second principal components, accounting for 53.1% and 12.2% of the total variation, respectively. Abbreviations of the mice treatment groups: B, bacteria; BY, bacteria-yeast; BY_ABX/AFX, bacteria-yeast and antibiotic or antifungal treatment; GF, germ-free; Y, fungi/yeast.

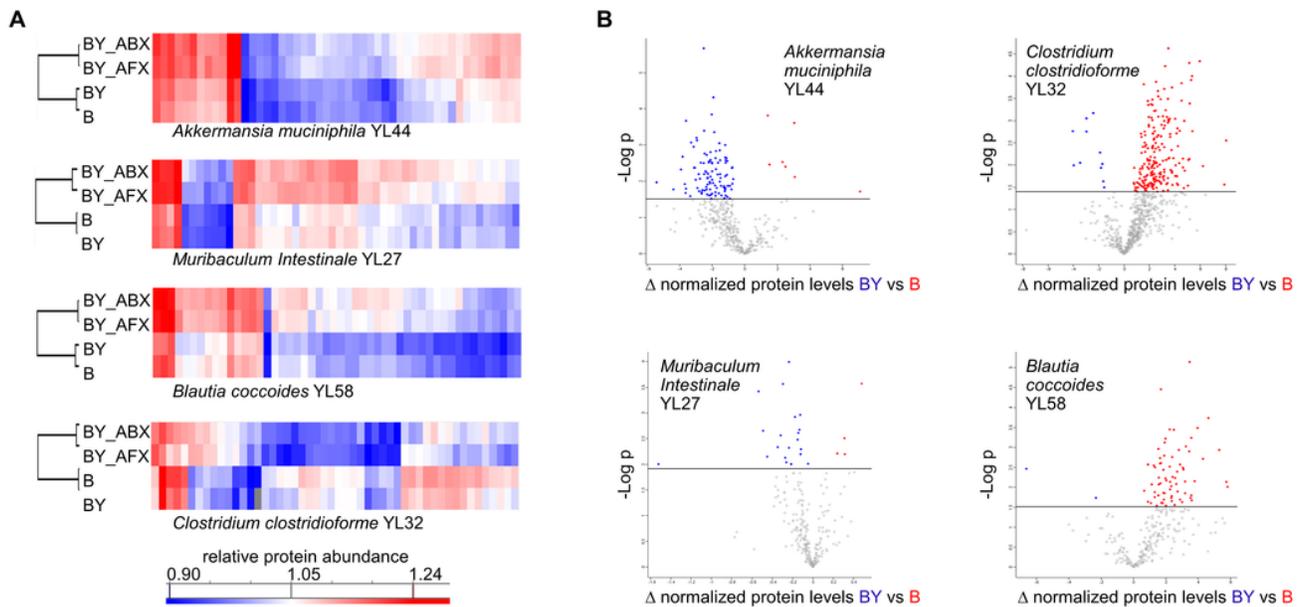


Figure 2

The response of four bacterial proteomes to treatment with antimicrobials and the presence of gut fungal species. A) A selection of 50 most significant differentially produced proteins for *A. muciniphila* YL44, *M. intestinale* YL27, *B. coccoides* YL58, and *C. clostridioforme* YL32 (ANOVA, FDR 0.05). B) Volcano plots representing the results of student t-test statistical comparison in protein levels between B and BY mice groups (FDR 0.05). Abbreviations of the mice treatment groups: B, bacteria; BY, bacteria-yeast; BY_ABX/AFX, bacteria-yeast and antibiotic or antifungal treatment.

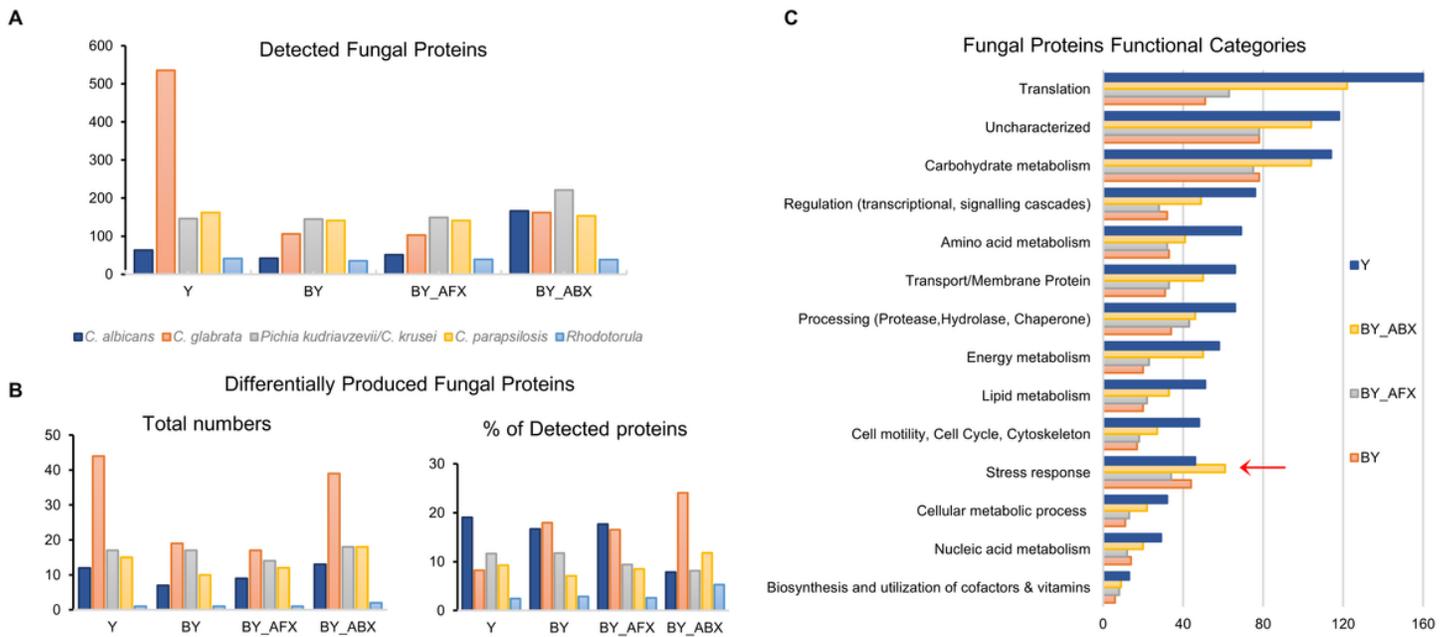


Figure 3

Response of fungal proteomes to antimicrobials and presence of bacteria. A) Distribution of detected fungal proteins across four mice group (n=1,492). B) Distribution of fungal proteins with significantly changed levels across four mice groups (n=99, ANOVA, FDR 5%). Table S6 shows the fungal proteins identified as having significantly different levels. C) Distribution of the main functional classes of detected fungal proteins (n=1,492). Proteins functional annotation was downloaded from the UniProtKB database (29) and compared to annotations obtained using the DAVID (30) and STRING-db (31) tools. Abbreviations of the mice treatment groups: BY, bacteria-yeast; BY_ABX/AFX, bacteria-yeast and antibiotic or antifungal treatment; Y, fungi/yeast.

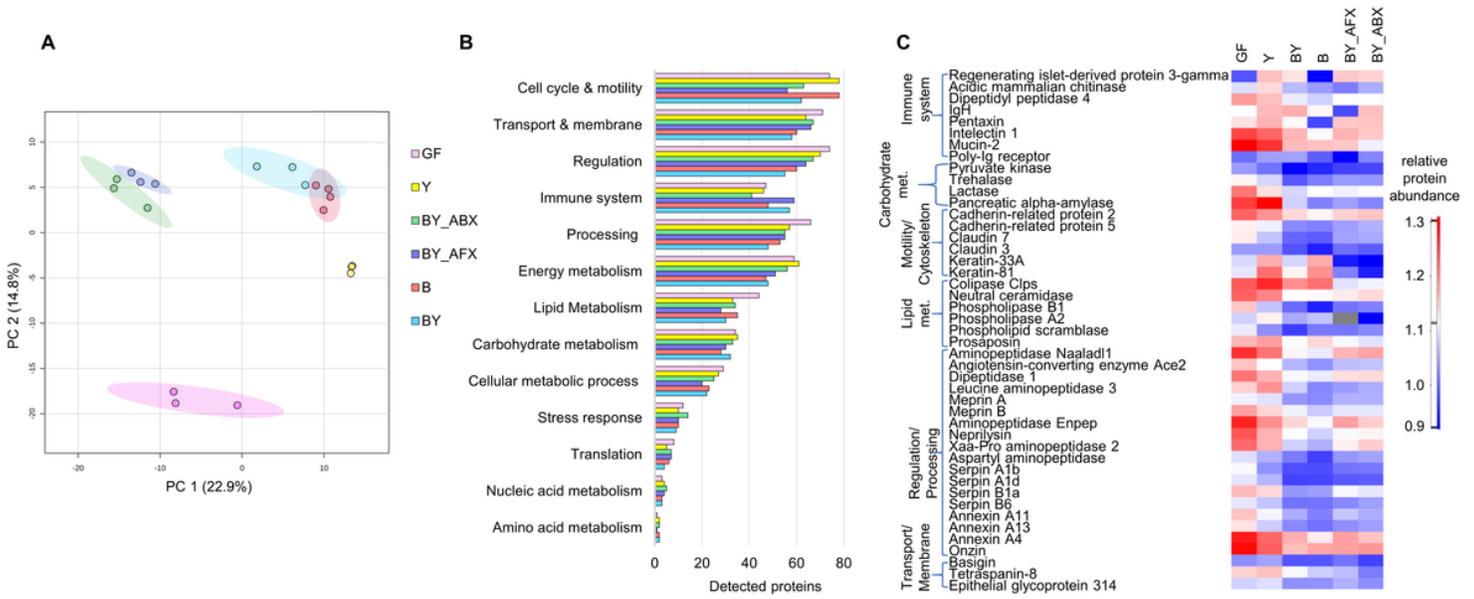


Figure 4

The response of host fecal proteome to microbial colonisation and antimicrobial treatment. A) PCA score plot based on 405 quantified mouse proteins. X- and Y-axis show the first and second principal components, accounting for 22.9% and 14.8% of the total variation, respectively. B) Functional classes of 651 detected mouse proteins, with the protein annotations derived from the UniProtKB database (29) and compared to those obtained by the DAVID (30) and STRING-db (31) tools. C) Selection of differentially produced mouse proteins extracted from feces. Abbreviations of the mice treatment groups: B, bacteria; BY, bacteria + fungi; BY_ABX/AFX, bacteria + fungi and antibiotic or antifungal treatment; GF, germ-free; Y, fungi/yeast.

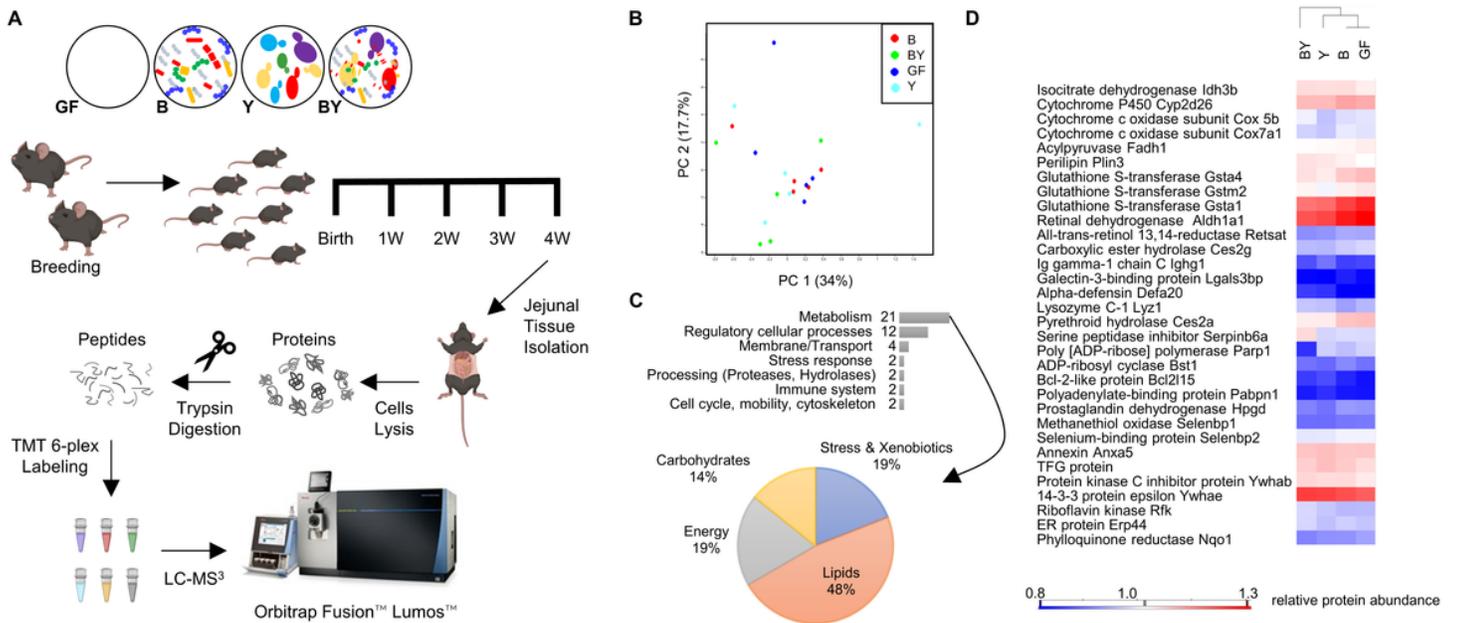


Figure 5

Response of mouse jejunal proteomes to microbial colonisation. A) Experimental design of Tandem Mass Tag (TMT) labelling experiment. B) PCA score plot based on 1,019 mouse proteins quantified in a minimum of 5 biological replicates. X- and Y-axis show the first and second principal components, accounting for 34% and 17.7% of the total variation, respectively. C) Functional classes of 45 differentially produced mouse jejunal proteins (ANOVA followed by THSD, FDR 5%). D) Selection of differentially produced mouse proteins extracted from jejunal tissue. Abbreviations of the mice treatment groups: B, bacteria; BY, bacteria + fungi; GF, germ-free; Y, fungi/yeast

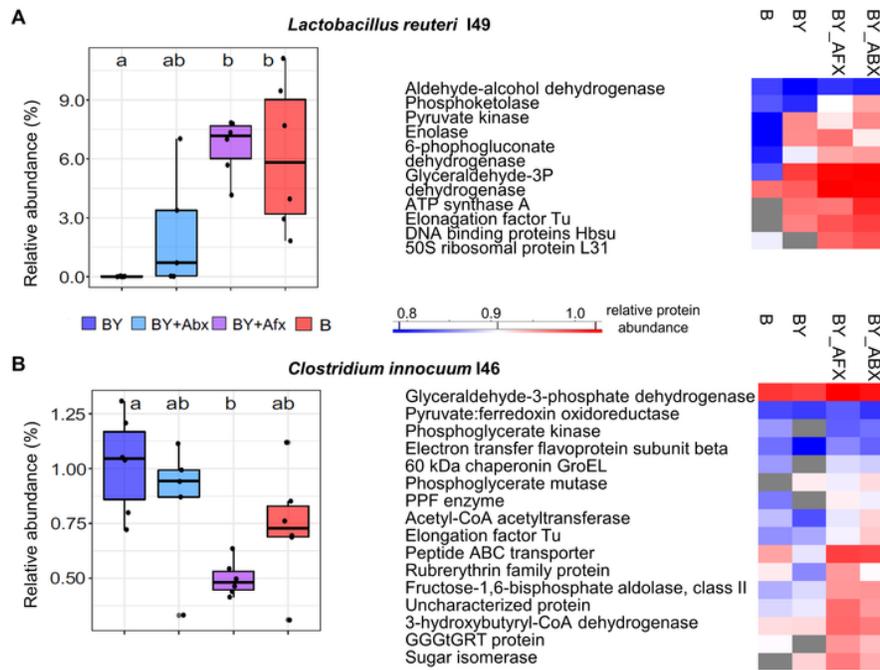


Figure 6

Detection of stress response in bacterial proteomes. The relative abundance based on 16S rRNA sequencing and differentially produce proteins of *Lactobacillus reuteri* (A) and *Clostridium innocuum* (B). Proteins quantified in a minimum of two replicates are shown; grey fields in heatmaps indicate the protein amount was below the quantification limit. Colour denotes microbial colonization (red B - bacteria, blue BY – bacteria + yeast, cyan blue BY_Abx – BY + antibiotic, purple BY_Afx - BY + antifungal)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.xlsx](#)
- [SupplementaryInformation.docx](#)