PNAS

Supporting Information for

A soil fungus confers plant resistance against a phytophagous insect by disrupting the symbiotic role of its gut microbiota

Ilaria Di Lelio[†], Giobbe Forni[†], Giulia Magoga, Matteo Brunetti, Daniele Bruno, Andrea Becchimanzi, Maria G. De Luca, Martina Sinno, Eleonora Barra, Marco Bonelli, Sarah Frusciante, Gianfranco Diretto, Maria C. Digilio, Sheridan L. Woo, Gianluca Tettamanti, Rosa Rao, Matteo Lorito, Morena Casartelli[*], Matteo Montagna[*], Francesco Pennacchio[*]

[†]Equally contributing authors.
*Corresponding authors: Morena Casartelli, Matteo Montagna, Francesco Pennacchio
Email: morena.casartelli@unimi.it; matteo.montagna@unina.it; f.pennacchio@unina.it

This PDF file includes:

Other supporting materials for this manuscript include the following:

Dataset S1A, S1B, S1C, S1D
Markdown file describing the whole pipeline of the analyses
https://figshare.com/s/b6db24859a65c391e629 (raw data and the intermediate files are available at the same link).

Supporting Results

1. Supporting Results for midgut epithelium transcriptomics of Spodoptera littoralis *larvae*
The RNA-sequencing experiment focusing on the transcriptional changes of the larval midgut epithelium associated with the feeding on T22-plants generated a total of 85.04 millions paired end reads with an average 14.17 millions reads per sample. Reads were uploaded to SRA under accessions: SRR17054439 - SRR17054444 (BioProject: PRJNA784009). FastQC showed a single GC peak at around 46% and no over-represented sequences, implying the lack of consistent leakage of rRNA/mtDNA or sources of microbial contamination. On average, the percentage of mapping reads on *Spodoptera littura* genome, obtained using STAR, was 83.3% (min: 81.9%; max: 84.2%) and FeatureCounts successfully assigned to features 47.4% of the reads (min: 46.2%; max: 49.9%). Overall, 9,791 expressed genes were identified, out of which 85 were found to be differential expressed in the DESeq2 analysis. Overall, feeding on T22-plants did not appear to have a strong impact on *S. littoralis* midgut gene expression profiles, the percentage of DE genes represented in fact < 1% of the total number of expressed features. Moreover, Principal Components Analysis (Fig. S8a) and clustering analysis (Fig. S8b) leveraging transcripts abundances did not clearly separate C-larvae from T22-larvae, supporting the evidence that feeding on T22-plants did not cause a strong perturbation of midgut epithelium transcriptional profile. Using default parameters, PANNZER2 generated 139,042 Gene Ontology terms in total, which were assigned to 17,348 genes. The GO terms enrichment analysis of upregulated genes and subsequent semantic clustering highlighted biological processes associated with immune response activation towards bacteria and oxidative processes, along with those associated with lipid and proteins metabolism (Fig. 2I). Similarly, the semantic clustering of enriched molecular function GO-terms (Fig. S3) were found to be associated with oxidative processes (GO:001649 oxidoreductase activity) and metabolism (GO:0003824 catalytic activity). Regarding this latter term, there is a clear upregulation of genes involved in the degradation of proteins (e.g., GO:0004252 serine-type endopeptidase activity; GO:0017171 serine hydrolase activity; GO:0052689 carboxylic ester hydrolase activity) and lipids (e.g., GO:0016298 lipase activity; GO:0004806 triglyceride lipase activity; GO:0008970 phospholipase A1 activity) (Fig. S3). The results of the GO-terms enrichment analysis of T22-larvae upregulated genes closely reflect what found when examining genes possibly associated with digestion (Figure 2J).

2. Supporting Results for midgut microbiota taxonomic characterization of Spodoptera littoralis *larvae*
From the sequencer a total of 136,371 paired-end 250 bp reads were obtained, corresponding to a mean of ~13,370 sequences per sample. Reads were uploaded to SRA under accessions: SRR17050410 - SRR17050419 (BioProject: PRJNA784009). After filtering, denoising, merging and chimera removal 1,569 unique biological sequences (ASVs) were identified (average length 418 bp) and a mean of 8,254 reads per sample were retained (~61 % of the original sequences). Rarefaction curves of the observed ASVs richness in 6,000 subsampled sequences showed that our sequencing effort was sufficient to capture the bacterial diversity associated with the analyzed midgut samples (Fig. S9). All the ASVs were assigned to bacterial taxa except seven ASVs assigned to Archaea (0.13%), two ASVs assigned to mitochondria (0.04%) and ten assigned to chloroplasts (0.84%), which were excluded from subsequent analyses.
The most represented bacterial classes present in the microbiota of *S. littoralis* were Bacilli (~35%), Gammaproteobacteria (~30%), Bacteroidia (~9%), Clostridia (~9%) and Alphaproteobacteria (~6%). The genus *Enterococcus* (Bacilli; Enterococcaceae), that includes well-known bacterial symbionts of *S. littoralis,* was the most abundant genus, 25% of the reads on average. The microbiota of C-larvae was characterized by a high evenness, with no dominant taxa (except for *Enterococcus*). In fact, about half of the sequences were assigned to low abundance bacterial genera, each of them representing on average less than 1% of the reads in a single sample. The effect of the treatment determined a reduced evenness and an increased richness of the microbiota, mainly driven by a taxonomic shift in its composition. In T22-larvae, besides *Enterococcus* that maintained high relative abundances, also other bacterial taxa reached abundances comparable with those of the primary symbiont of *S. littoralis* (e.g., *Erysipelatoclostridium* 24%, *Alicycliphilus* 17%). Those bacteria were present also in the microbiota of the control group but with low relative abundances (1% and 6% respectively).
The five most abundant ASVs assigned to *Enterococcus* (97.8% of the reads) showed different relative abundance in the two experimental groups. Two ASVs, identified as *Enterococcus mundtii* (confidence 0.98) were well represented in T22-larvae (~20%) and almost absent in the control (<2%). Other two

ASVs, with dubious identification (*E. gallinarum,* confidence < 0.8), were well represented in both groups but more abundant in C-larvae (control-larvae ~68%, T22-larvae ~48%). A fifth ASV (*E. casseliflavus*, confidence <0.6) had similar abundance in both groups (~29%).

3. Supporting Results for midgut microbiota transcriptomics of Spodoptera littoralis *larvae*
The sequencing run generated a total of 129,274 million paired-end reads with an average of 21,545 million reads per sample. Reads were uploaded to SRA under accessions: SRR17050383 - SRR17050387 (BioProject: PRJNA784009). Metaphlan2 taxonomic assignment found a marginal presence of Eukaryota and Viruses (< 0.11%) and the class Bacilli as the prominent contributor of microbiome activity across all samples (91.29% on average), with Actinobacteria consistently found as the second ranking in abundance (7.85% on average). At the species level, *Enterococcus casseliflavus* resulted to be the most abundant species across all samples considered (80.9% on average). When considering the core functionality of the C-larvae microbiota (mean pathway coverage > 0.8; Fig. S4), several pathways apparently involved in supporting host nutrition with essential amino acids were found; these include superpathways of L-isoleucine biosynthesis I & II, L-lysine biosynthesis II & III, L-proline biosynthesis II, superpathway of L-lysine, L-threonine and L-methionine biosynthesis II, superpathway of branched chain amino acid biosynthesis (Ile, Val, Leu). Additionally, pathways associated with ribonucleotides (purine ribonucleosides degradation) and sugars (L-rhamnose degradation I; D-galactose degradation; stachyose degradation; pentose phosphate pathway I) degradation were found along with some associated with biosynthetic processes of ribonucleotides (superpathway of 5-aminoimidazole ribonucleotide biosynthesis; adenosine ribonucleotides de novo biosynthesis), glycogen (glycogen biosynthesis I from ADP-D-Glucose) and SAM (S-adenosyl-L-methionine salvage I).
HUMAnN 2.0 quantified for each sample an average abundance of ~60,000 Uniref90 features (max: 69,570; min: 46,588) which were attributed to a total of 549 taxonomy-stratified pathways. Other than the "per-feature" analyses carried out using MaAsLin2 and described in the main text, we also tested the overall difference in midgut microbiota activity between C-larvae and T22-larvae. A Principal Component Analysis carried out on the non-taxonomy stratified pathways clearly separated the two conditions (1^{st} component: 84%; 2^{nd} component: 11%; Fig. S5a). A complementary clustering analysis, using spearman distances on the square root transformed non-stratified pathways, further corroborated a clear difference between the transcriptional activity in the two experimental conditions (Fig. S5b).

4. Supporting Results for midgut metabolomics of Spodoptera littoralis *larvae*
Among the 337 (210 down- and 127 over-) differentially accumulated metabolites (DAMs), those associated with lipid metabolism, including fatty acids, phospholipids and triglycerides from uncertain source, accounted for the highest DAMs number. They were followed by: *i.* three plant-specific classes of metabolites, *i.e*. flavonoids, saponins and alkaloids; *ii.* toxins and antibiotics, mostly from fungi; and *iii.* hormones and prostaglandins of insect origin. Among antibiotics and toxins, in T22-larvae were down-accumulated several metabolites of bacterial origin, *i.e.* 15-demethoxy-epsilon-rhodomycin-like ($FC_{T22/C\text{-}larvae}$ = 0.02), agrocin 84-like ($FC_{T22/C\text{-}larvae}$ = 0.03) and doxorubicinol-like ($FC_{T22/C\text{-}larvae}$ = 0.25). Moreover, in T22 larvae were over-accumulated compounds of fungal origin such as acetylneomycin-like ($FC_{T22/C\text{-}larvae}$ = 2.86), aphidicolin-like ($FC_{T22/C\text{-}larvae}$ = 2.51), and 5-deoxydiplosporin ($FC_{T22/C\text{-}larvae}$ = 4.43), sphingofungin A ($FC_{T22/C\text{-}larvae}$ = 2.46), brefeldin A-like ($FC_{T22/C\text{-}larvae}$ = 2.22), diplodiatoxin ($FC_{T22/C\text{-}larvae}$ = 2.13) and zearalanone-like ($FC_{T22/C\text{-}larvae}$ = 2.1).
Metabolomics analyses pointed out also many alterations in plant secondary metabolites, as flavonoids and saponins. Flavonoids and hydroxylated phenolic molecules resulted highly accumulated in T22-larvae in comparison with C-larvae. For example, (S)-naringenin 8-C-(2''-rhamnosylglucoside) ($FC_{T22/C\text{-}larvae}$ = 15.37) and quercetin 3-glucosyl-(1 → 2)-xyloside ($FC_{T22/C\text{-}larvae}$ = 7.4), as well myricetin 3-glucuronide ($FC_{T22/C\text{-}larvae}$ = 5.1) and luteolin 3',4'-dimethyl ether 7-rhamnoside ($FC_{T22/C\text{-}larvae}$ = 3.14), resulted over-accumulated in T22-larvae. Down-accumulated in T22-larvae were a series of glycosides (saponins, steroids, terpenes), such as protodegalactotigonin ($FC_{T22/C\text{-}larvae}$ = 0.24), spirosolane-15-keto-3-ol-3-O-diglycoside ($FC_{T22/C\text{-}larvae}$ = 0.3), protoprimulagenin A 3-[rhamnosyl-(1 → 4)-rhamnosyl-(1 → 4)-[rhamnosyl-(1 → 2)]-glucosyl-(1 →?)-glucuronide]-like ($FC_{T22/C\text{-}larvae}$ = 0.32) and (S)-alpha-terpinyl glucoside ($FC_{T22/C\text{-}larvae}$ = 0.62).
For lipid metabolism, several compounds taking part in plant jasmonic acid pathway (JA) were found over-accumulated in T22-larvae. Notable examples are tuberonic acid ($FC_{T22/C\text{-}larvae}$ = 5.42), 12(S),20-DiHETE ($FC_{T22/C\text{-}larvae}$ = 4.13), 12-hydroxyjasmonic acid ($FC_{T22/C\text{-}larvae}$ = 3.78), 13-OxoODE ($FC_{T22/C\text{-}larvae}$ =

3.44), and 8,13-DiHODE ($FC_{T22/C\text{-}larvae}$ = 2.64). Likewise, different members in the insect prostaglandin (PG) group showed altered amounts, either down-accumulated (13,14-dihydro-16,16-difluoro prostaglandin F2α, $FC_{T22/C\text{-}larvae}$ = 0.28; PGF2α-11-acetate methyl ester, $FC_{T22/C\text{-}larvae}$ = 0.32; 2,3-dinor-6-keto-PGF1a, $FC_{T22/C\text{-}larvae}$ = 0.65) and over-accumulated (11-deoxy-16,16-dimethyl-PGE2, $FC_{T22/C\text{-}larvae}$ = 2.89; 2,3-Dinor-11b-PGF2a, $FC_{T22/C\text{-}larvae}$ = 2.03; 2,3-dinor-6-keto-PGF1α, $FC_{T22/C\text{-}larvae}$ = 1.96) in T22 larvae. Interestingly, both JA and PG are members of oxylipins sharing the eicosanoid pathway. Moreover, a set of metabolites acting as insect hormones were identified and found over accumulated in T22-larvae: dehydrojuvabione ($FC_{T22/C\text{-}larvae}$ = 4.74) and methoprene acid ($FC_{T22/C\text{-}larvae}$ = 2.09).

Supporting Methods

1. Supporting Methods for insect bioassays on leaves of tomato plants colonized by Trichoderma afroharzianum *strain T22*

Spodoptera littoralis experimental larvae were derived from a colony established in 2009 at the University of Napoli Federico II (Department of Agricultural Sciences, Laboratory of Entomology "Ermenegildo Tremblay", Portici-NA, IT), regularly refreshed once a year with field-collected insects. Groups of newly hatched *S. littoralis* larvae were reared in plastic boxes (30×40×15 cm), bottom lined with 50 mL of 1.5% agar (w/v), on sub-apical leaves of 4 weeks-old T22-plants (T22-larvae) or of C-plants (C-larvae). Leaves were randomly selected and excised from tomato plants starting from the second whorl, then offered to larvae. To prevent any undesired effect due to mechanical damage, potentially triggering defense responses, leaves were obtained from plants that were left undisturbed (*i.e.*, not exposed to mechanical damage) for at least three days (1, 2). Leaves were replaced daily, and larvae checked to monitor their development and to select those to be used for the feeding bioassay, which was carried out by comparing two experimental groups made of 16 synchronous newly molted day one 3^{rd} instar larvae, obtained by selecting 2^{nd} instars ready to molt (*i.e.,* showing head capsule slippage) before the onset of the scotophase and collecting those molted in the morning of the following day. In detail, day one 3^{rd} instar larvae were singly weighed, and transferred into multi-well plastic rearing trays (RT32W, Frontier Agricultural Sciences), bottom-lined with 1 mL of 1.5% agar (w/v) and closed by perforated plastic lids (RTCV4, Frontier Agricultural Sciences). The larvae were singly reared on a leaf disk (4 cm^2) from a C-plant or T22-plant, which was replaced daily. Bioassays were carried out in environmental chamber at 25 ± 1°C, photoperiod 16:8 light/dark. The following parameters were recorded daily: larval survival, larval weight, pupal survival, adult emergence and adult survival. Pupal weight was recorded on day three of the pupal stage. Leaf consumption by day one of 4^{th}-6^{th} instar larvae was obtained as difference of the initial and final weight of the leaf disk. Survival curves of *S. littoralis* were compared by using Kaplan–Meier and Log-rank analysis. Differences between experimental groups (*i.e.*, C-larvae and T22-larvae) were assessed by unpaired Student's t-test or Welch's t-test. Normality of data was checked with Shapiro-Wilk test, while homoscedasticity was tested with Levene's test and Barlett's test. Statistical analyses were performed using Prism (GraphPad Software Inc. version 6.0b).

2. Supporting Methods for midgut microscopy analysis

Midgut samples, after the removal of fat body and tracheal residues, were fixed in 4% (v/v) glutaraldehyde in 0.1 M Na-cacodylate buffer, pH 7.4, for 2 hours at room temperature and then overnight at 4°C. Samples were then embedded in Epon resin as previously described (3). Semi-thin sections (600 nm-thick) were stained with crystal violet and basic fuchsin and observed with an Eclipse Ni-U microscope (Nikon), equipped with TrueChrome II S digital camera (Tucsen Photonics). Ultra-thin sections (70 nm-thick) were stained with lead citrate and uranyl acetate and observed with a JEM-1010 transmission electron microscope (Jeol) equipped with Morada digital camera (Olympus). Five midgut samples from each group (*i.e.*, C-larvae and T22-larvae) were examined.

For glycogen detection, midgut samples were fixed in 4% (w/v) paraformaldehyde in Phosphate Buffer Saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4, pH 7.4) for 2 hours at room temperature and then overnight at 4°C. Samples were then processed for paraffin embedding as previously reported (4). To detect glycogen deposits in the midgut epithelium, 7 µm-thick sections were stained with Periodic Acid-Schiff kit (PAS) (Bio-Optica) (3). To confirm the specificity of the staining, PAS reaction was performed in combination with diastase (PAS-D) which breaks down this polysaccharide. Glycogen deposits in C-larvae and T22-larvae were quantified using ImageJ (National Institutes of Health). In detail, after selecting the midgut epithelium as region of interest (ROI), the "threshold" was set

with the thresholding tool and glycogen spots were counted with the “analyze particles” command, as reported by Jensen (5). The ROI manager provided the number and the dimension of all the spots counted. For each condition (*i.e.*, C-larvae and T22-larvae), midguts obtained from 5 larvae were analyzed, measuring the glycogen deposits in 5 different midgut sections and comparing the mean values by Welch’s t-test. For the detection of lipid droplets, midgut samples were embedded in polyfreeze cryostat embedding medium (Polyscience Europe) and 7 µm-thick cryosections were stained with Oil Red O (ORO) (Bio-Optica) (6). Both stainings were performed according to the manufacturer’s instructions and sections were analyzed with Eclipse Ni-U microscope (Nikon) equipped with a digital camera (Tucsen Photonics).

3. Supporting Methods for midgut transcriptomics

Subsequently to extraction, RNA purity was evaluated by 260 nm/280 nm absorbance ratio with NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific), while the concentration and the quality was assessed with TapeStation 4200 (Agilent Technologies).

RNA samples from midgut epithelium were used as templates in libraries preparation. Each indexed library was prepared from 500 ng of purified RNA with TruSeq Stranded mRNA Sample Prep Kit (Illumina), according to the manufacturer’s instructions. Libraries were quantified using the TapeStation 4200 (Agilent Technologies) and Qubit fluorometer (Invitrogen Co.), then equimolar pooled to a final concentration of 2 nM. The pooled samples were subjected to cluster generation and sequencing using an Illumina NextSeq 500 System (Illumina) in a 2×75 paired-end format at a final concentration of 1.8 pmol.

The 75 nt paired reads generated for each sample, after the quality check performed using FastQC (7), were mapped on the reference genome of *Spodoptera litura* (GCA_002706865.1 - ASM270686v1; 8) using the bioinformatics tool STAR (v2.7.5; 9) with the standard parameters for paired reads. Transcript quantification was performed using FeatureCounts (v2.0.1; 10) and then R was used to create a matrix of all genes expressed in all samples with the corresponding read-counts. Differential gene expression analyses were performed using Deseq2 with Benjamini and Hochberg’s method for multiple tests correction (11). Treatment-biased genes were defined as those that showed an FDR < 0.05 and a LogFC either >1.5 (upregulated genes) or <1.5 (downregulated genes). GO-terms were generated with PANNZER2 using standard parameters (12). Enrichment analyses were performed separately for Biological Processes and Molecular Functions ontologies with the TopGO package in Bioconductor: a minimum node size of 3 and the “classic" algorithm were used (13). GO-terms were considered enriched when Fisher exact test resulted with *p*-value< 0.01, then enriched GO-terms were summarized and visualized as scatter plots based on their semantic similarity using REVIGO (14).

4. Supporting Methods for digestive enzyme assays

Total proteolytic activity was assayed with azocasein (Merck), as previously reported (4) with minor modifications. The buffer used to dilute the samples and resuspend the substrate was 20 mM Tris-HCl, pH 8.6. Experiments were performed in triplicate and each experiment was performed with pooled samples obtained from at least 8 larvae. One unit (U) of total proteolytic activity was defined as the amount of enzyme that causes an increase in absorbance by 0.1 unit per min per mg of proteins. α-amylase activity was measured with starch as substrate, as previously reported (4). Experiments were repeated 5 times with pooled samples (see above). One unit (U) of α-amylase activity was defined as the amount of enzyme necessary to produce 1 mg of maltose per min per mg of proteins. Lipase activity was measured in triplicate using the Lipase Activity Colorimetric Assay Kit (BioVision) according to the manufacturer’s instructions. Aminopeptidase N activity was assayed using L-leucine p-nitroanilide (Merck) as substrate, as previously reported (3), with minor modifications. After thawing, midgut tissues were homogenized in a Glass-Teflon Potter homogenizer (nine strokes at 2,000 rpm) in 100 mM mannitol, 10 mM Hepes-Tris, pH 7.2 (3 mL/100 mg tissue). Experiments were repeated 5 (for C-larvae) and 6 times (for T22-larvae) with pooled samples (see above). One unit (U) of aminopeptidase N activity was defined as the amount of enzyme that releases 1 µmol of product per min per mg of proteins. Differences in enzyme activities were accounted by using Welch’s t-test.

5. Supporting Methods for taxonomic characterization of the midgut microbiota

V3-V4 16S rRNA libraries were obtained using Bakt 341F/805R primers (15). PCRs were assembled according to Metagenomic Sequencing Library Preparation (Illumina). In the PCR step two PNA oligos

(PCR Blockers PNA Bio) were used against chloroplast and mitochondrial 16S rRNA sequences from diverse plant species. A negative control was also included in the workflow, consisting of all reagents used during sample processing (16S rRNA amplification and library preparation); the library was also prepared on a standard microbial community (Zymo Research). Libraries were quantified using Qubit fluorometer (Invitrogen Co.) and pooled to an equimolar amount of each index-tagged sample to a final concentration of 4nM, including the Phix Control Library. Pooled samples were sequenced on MiSeq platform (Illumina) in a 2×250 paired-end format.
After FastQC quality check (7), paired reads were denoised using the DADA2 pipeline (16) to obtain Amplicon Sequence Variants (ASVs). For the taxonomic assignment of the ASVs, a naïve Bayes classifier (17, 18) was trained on the 16S SILVA database release 138 (19) trimmed to the region amplified by the primers used in this study. To improve the taxonomic classification, taxonomic weights were computed with the q2-clawback plugin (20) and used to fit again a naïve Bayes classifier on the reference database to obtain the final taxonomic assignments of the ASVs (confidence>95%). The analyses were performed using the QIIME2 platform (21). Diversity analyses were performed within the statistical framework of Hill numbers (22-25) using the R packages iNEXT and iNextPD (26-28).

6. Supporting Methods for midgut microbiota transcriptomics
RNA purity was evaluated by 260 nm/280 nm absorbance ratio with NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific), while the concentration and the quality were assessed with TapeStation 4200 (Agilent Technologies). RNA samples from midgut epithelium were used as templates in libraries preparation. Each indexed library was prepared from 500 ng of purified RNA with TruSeq Stranded mRNA Sample Prep Kit (Illumina), according to the manufacturer's instructions. Libraries were quantified using the TapeStation 4200 (Agilent Technologies) and Qubit fluorometer (Invitrogen Co.), then equimolar pooled to a final concentration of 2 nM. The pooled samples were subjected to cluster generation and sequencing using an Illumina NextSeq 500 System (Illumina) in a 2×75 paired-end format at a final concentration of 1.8 pmol.
A quality check was performed on the 75 nt paired reads using FastQC (7). Functional profiling was carried out using HUMAnN 2.0 with the ChocoPhlan and UniRef100 databases (29); to allow a direct comparison among samples, abundances were normalized using copies per million (CPM). To identify relevant pathways in the microbiota of C-larvae, we selected those with a mean coverage > 0.8 across control samples only and visualized the associated metabolic map using Metacyc pathway college.
A principal component analysis leveraging CPM-normalized pathway abundances was performed with R and visualized using ggplot2 (R Core Team, 2020; 30). To identify metabolic pathways changes associated with the treatments we leveraged MaAsLin 2 using a linear model and arcsine square root transformation for proportional relative abundance data; pathway abundances were considered to be different when *p*-value<0.01 (31).

7. Supporting Methods for midgut metabolomics
Midgut samples were extracted with 0.75 mL of cold 75% (v/v) methanol and 0.1% (v/v) formic acid, spiked with 1 µg/mL of daidzein (Merck) as internal standard. After sonication (30 sec at 7Hz) samples were shaken for 30 min at 20 Hz using a Mixer Mill 300 (Retsch) and centrifuged at 20,000 × g for 30 min; the supernatant was collected, filtered with HPLC filter tubes (0.45 µm pore size) and subjected to LC-ESI-HRMS analysis using an LTQ-Orbitrap Discovery mass spectrometry system (Thermo Fisher Scientific) as previously described (32) with slight modifications in Mass Spectrometer parameters: capillary temperature was set at 250°C; sheath and aux gas flow rate at 40 and 20 units, respectively; spray voltage was set at 3.5 kV and 2.8 kV for positive and negative ion mode, respectively; S-lens RF level 70%; resolution 70,000 FWHM; ACG target 1×10^6 ions. The instrument was calibrated both in negative and positive modes with a range of 110 to 1,500 *m*/*z* using calibration solutions (Thermo Fisher Scientific). The obtained MS chromatograms were subjected to untargeted metabolomics analysis as follows. Briefly, raw files obtained from six independent replicates for each condition were converted into the mzML data format by MSConvertGUI (ProteoWizard), followed by analysis using the XCMS Online (33-35); feature detection, retention time correction, alignment, statistics, annotation and identification were defined by 3,110 parameters ID (UPLC/Q-Exactive, positive ion mode), coupled to the interrogation of the Metlin database (36). Tentative identifications were validated comparing chromatographic and spectral properties with authentic standards (when available) and reference spectra, in house database, literature data, and based on the m/z accurate masses, as reported in the Pubchem database

(http://pubchem.ncbi.nlm.nih.gov/, accessed on 1 September 2021) for monoisotopic mass identification, or on the Metabolomics Fiehn Lab Mass Spectrometry Adduct Calculator (http://fiehnlab.ucdavis.edu/staff/kind/Metabolomics/MS-Adduct-Calculator/, accessed on 1 September 2021) in the case of adduction detection, subsequently confirmed by MS/MS fragmentation. Finally, differentially accumulated metabolites (DAMs) were quantified relatively by normalizing the internal standard (formononetin) amounts. Results were then visualized using ggplot2 (30) in R (R Core Team, 2020).

8. Supporting Methods for rescue bioassay

To isolate the midgut microbiota of *S. littoralis* larvae for rescue bioassay, 4th instar C-larvae were dissected under sterile conditions (see *Methods*), the peritrophic matrix with its content was collected and placed in a 1.5 mL sterile tube containing 200 µL of sterile PBS. The suspension was gently disaggregated with a sterile pestle, vortexed for 30 seconds, and centrifuged at 500 × g for 10 min at 4°C. The supernatant, containing both midgut juice and bacteria, was recovered in a new tube and centrifuged at 5,000 × g for 10 min at 4°C. The pellet, containing the bacteria, was washed twice with 1 mL of sterile PBS and finally suspended in 200 µL of sterile PBS at a final cell concentration of 5.2×10^8 cells/mL, calculated by plating tenfold serial dilutions of bacteria suspensions on Petri dishes containing Plate Count Agar medium (PCA, Hi-Media) supplemented with Igepal and Cycloheximide (50 mg/L, Merck), then incubated in the dark at 25°C and checked daily for bacterial colony growth. The number of bacterial cells present in the midgut juice of each larva was on average 1.04×10^7 (*i.e.*, bacterial gut equivalent). To isolate and identify *E. casseliflavus*, the colonies showing different morphologies were re-isolated by picking, streaked on Petri dishes containing Luria-Bertani Agar (LBA, Hi-Media) and incubated in the dark at 25°C. After three days, two colonies per plate were picked and molecularly identified. DNA was extracted (37) and V1-V9 regions of the 16S rRNA gene amplified by PCR (Dream Taq Master Mix; Thermo Fisher Scientific) using F8 and R1492 primers pair (38, 39). The obtained PCR amplicons were sequenced by Eurofins Genomics (Ebersberg) and subjected to BLAST analysis using GenBank database (http://www.ncbi.nlm.nih.gov/), selecting sequences from type material only (40). Colonies identified as *Enterococcus* were taxonomically assigned at the species level by BLAST analysis of an amplicon of chaperonin *GroEL* gene (41). One of the colonies corresponding to *E. casseliflavus* (BLAST result: query length = 1.350 bp, best match with *E. casseliflavus* strain ATCC 25788 accession number AF245682, E-value = 0.0, percentage of sequence identity = 98.9%) was picked and cultured in 10 mL of Luria-Bertani broth (Merck), incubated overnight with shaking at 25°C. The grown bacteria were pelleted at 5,000 × g for 10 min, washed twice with 1 mL of PBS and finally suspended in PBS at a final cell concentration of 5.2×10^8 cells/mL.

For rescue bioassays with bacteria T22-larvae at day two of the 3rd instar were daily offered with T22-plant leaf disks overlaid with one of the following bacterial suspensions: I) 10 µL of PBS containing 5.2×10^6 bacterial cells of the midgut microbiota previously isolated from C-larvae (a cell number corresponding to 0.5 bacterial gut equivalent); II) 10 µL of PBS containing 5.2×10^6 cells of *E. casseliflavus* (a number of *E. casseliflavus* cells corresponding to 0.5 bacterial gut equivalent); III) 10 µL of PBS containing 1.3×10^6 cells of *E. casseliflavus* (a cell number corresponding to ½ of the estimated load of this bacterium in C-larvae based on 16S rRNA metagenomics); and IV) 10 µL of PBS containing 6.5×10^5 cells *E. casseliflavus* (a cell number corresponding to ¼ of the estimated load of this bacterium in C-larvae based on 16S rRNA metagenomics). T22-larvae and C-larvae fed on their respective leaf disks, both overlaid and not with 10 µL of PBS, were used as controls. During the bioassays, parameters on insect survival and development were recorded (see *SI Methods* 1 for insect bioassays on leaves of tomato plants colonized by *T. afroharzianum* strain T22). Differences in survival rates were assessed by using Kaplan–Meier and Log-rank analysis. Differences in larval and pupal weight and in days for adult emergence among the experimental groups were assessed by One-Way ANOVA, or Kruskal–Wallis test when ANOVA assumptions were not respected. Data normality was assessed with Shapiro-Wilk test, while homoscedasticity was tested with Levene's test and Barlett's test. Statistical analyses were performed using Prism (GraphPad Software Inc.).

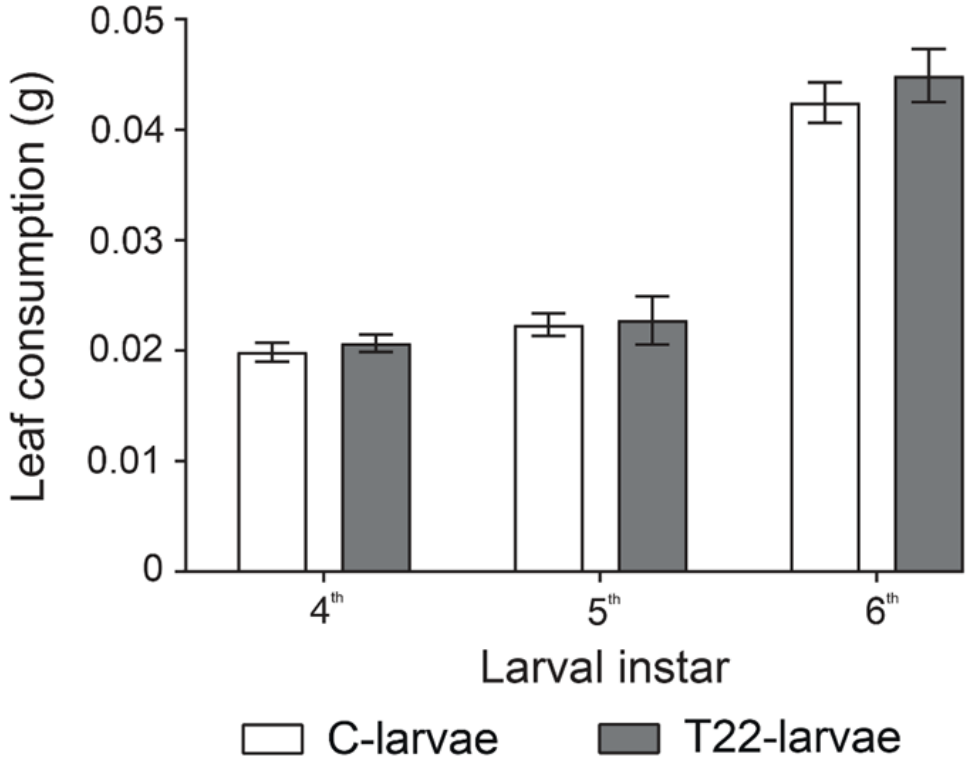

Fig. S1. Tomato leaf consumption by *Spodoptera littoralis* larvae. No differences in tomato leaf consumption were recorded in C-larvae and T22-larvae by day one of the 4th (Student's *t* test: $t(62) = 0.6795$, p-value = 0.4993), 5th (Student's *t* test: $t(52) = 0.1835$, p-value = 0.8552) and 6th instar larvae (Student's *t* test: $t(42) = 1.173$, p-value = 0.0902).

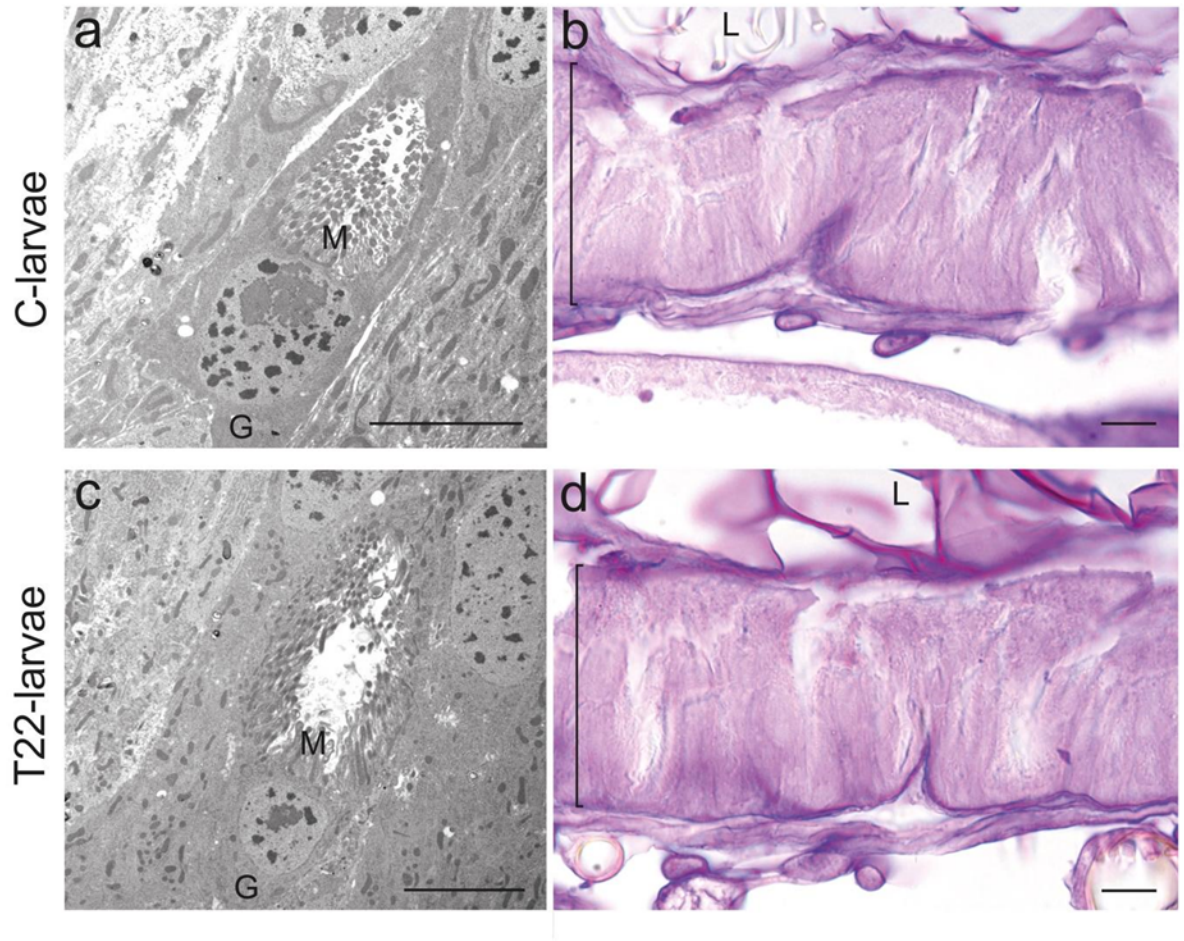

Fig. S2. Microscopy analysis of *Spodoptera littoralis* larval midgut. a, c TEM images of goblet cells in C-larvae (a) and T22-larvae (c) showing no alterations in their ultrastructural features. **b, d** Optical microscopy images of midgut in C-larvae (b) and T22-larvae (d) stained with Periodic Acid-Schiff (PAS) reaction for the detection of glycogen deposits in combination with diastase which breaks down this polysaccharide. The absence of purple/violet spots confirms the specificity of the staining; bracket: epithelium; G: goblet cell; L: lumen; M: microvilli; bars: 5µm (a, c), 10 µm (b, d).

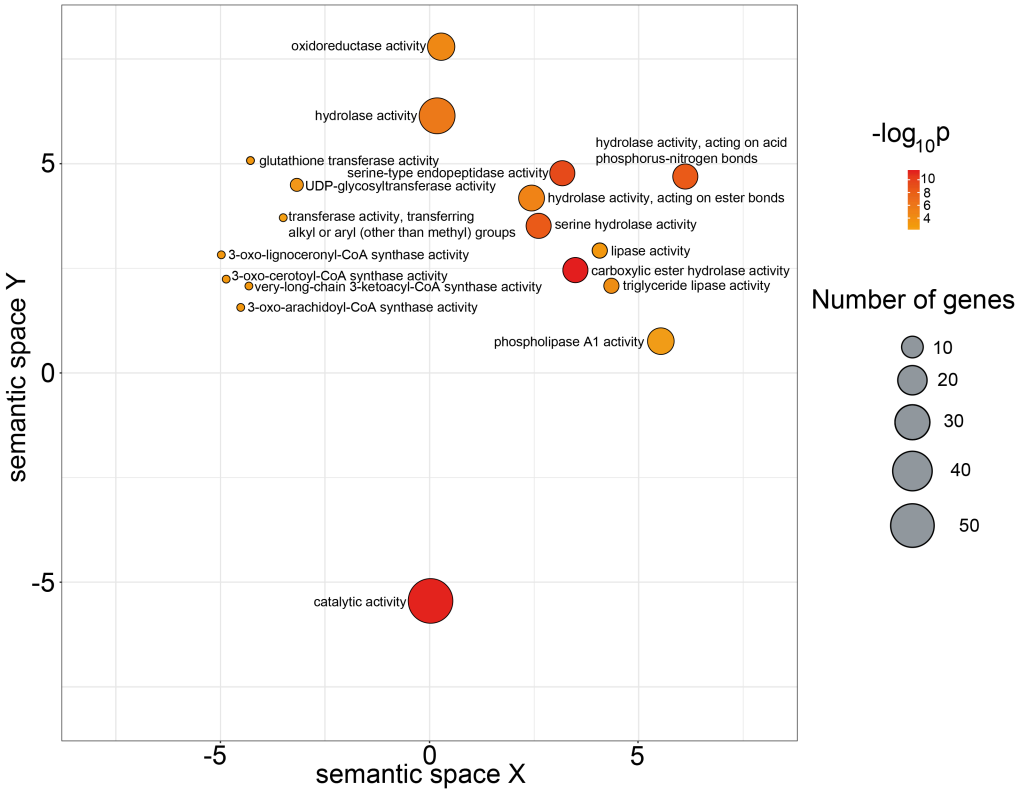

Fig. S3. Semantic clustering of enriched molecular functions GO terms associated with *Spodoptera littoralis* T22-larvae up-regulated genes resulting from midgut transcriptomic analyses. The color of circles indicates the -$\log_{10}$ *p*-value of the enriched GO terms, while the size is proportional to the number of DE genes associated with the GO term.

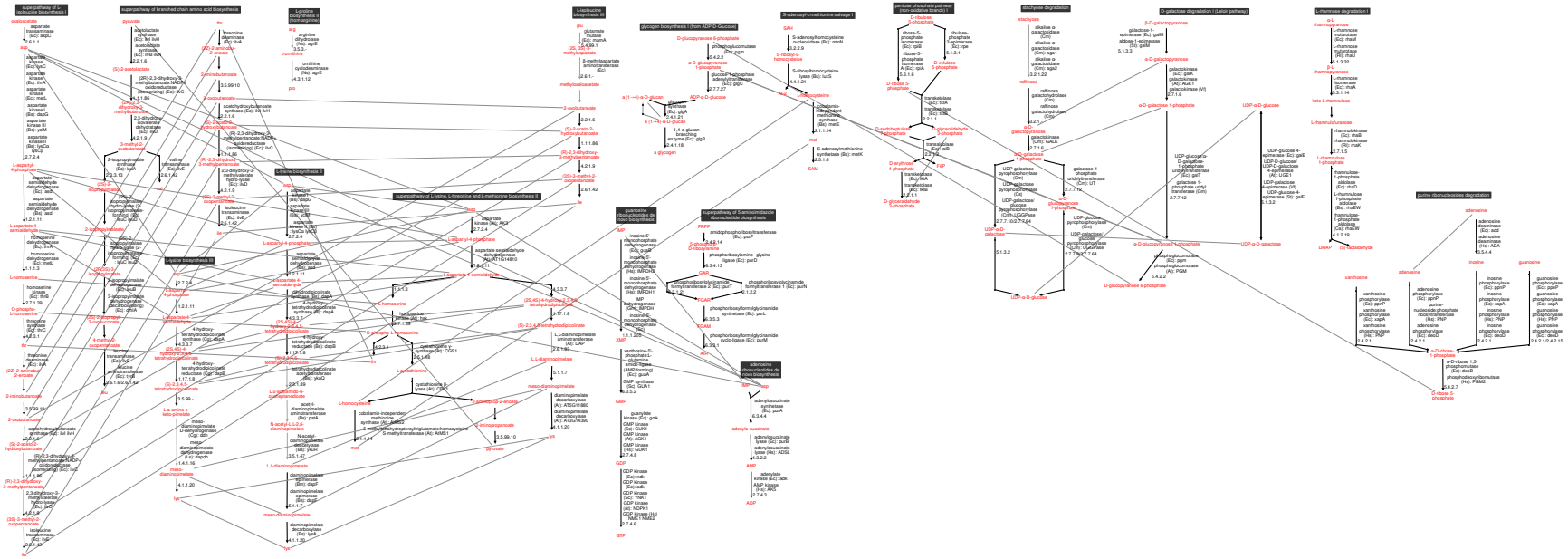

Fig. S4. Metacyc pathways of the midgut microbiota associated with *Spodoptera littoralis* C-larvae. College visualization of pathways with a mean coverage greater than 0.8.

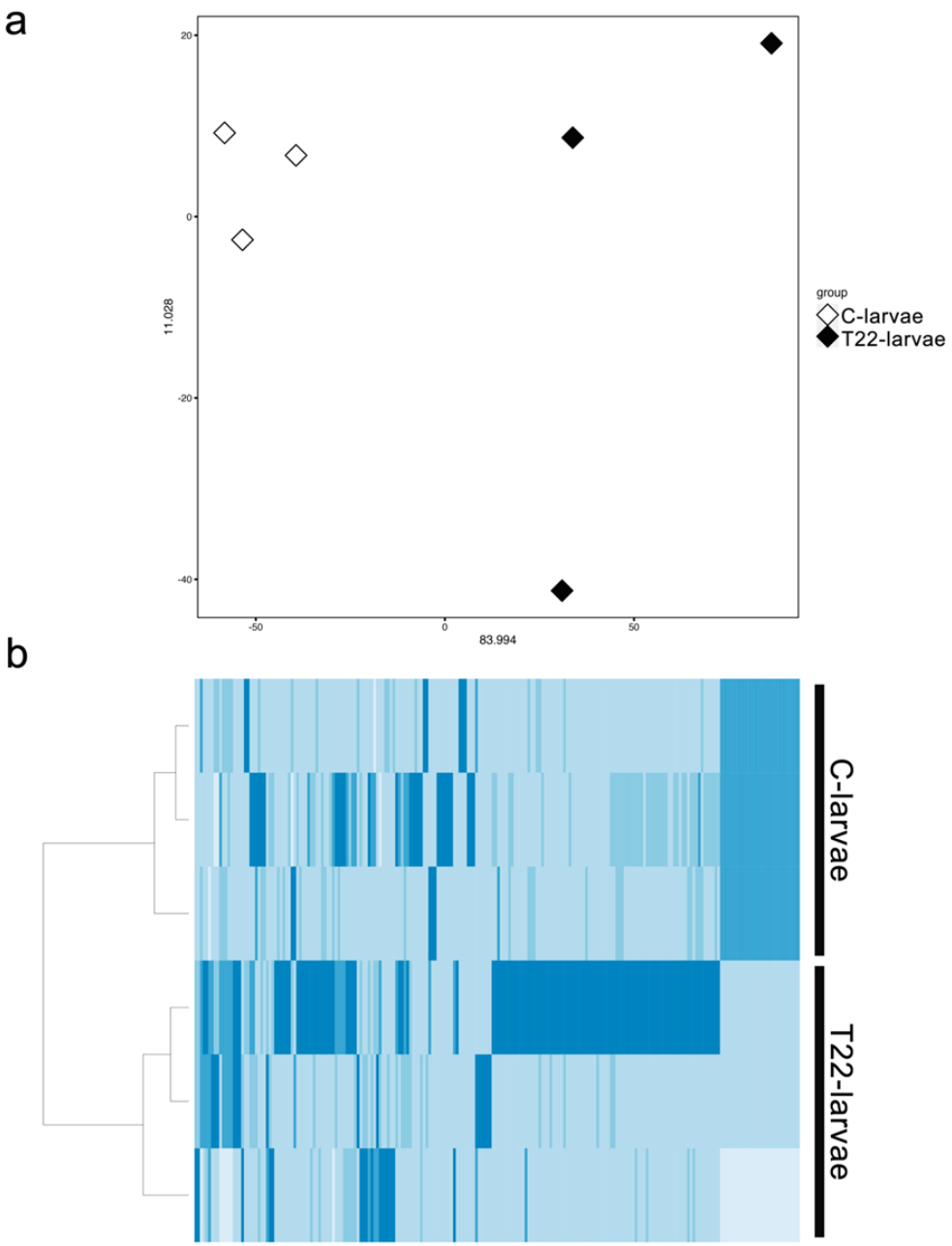

Fig. S5. Multivariate analyses on midgut microbiota transcriptomics in *Spodoptera littoralis* T22- and C-larvae. a Principal Components Analysis of the non-taxonomically stratified pathways abundances. **b** Hierarchical clustering analysis using Spearman distance on the square root transformed non-taxonomically stratified pathways.

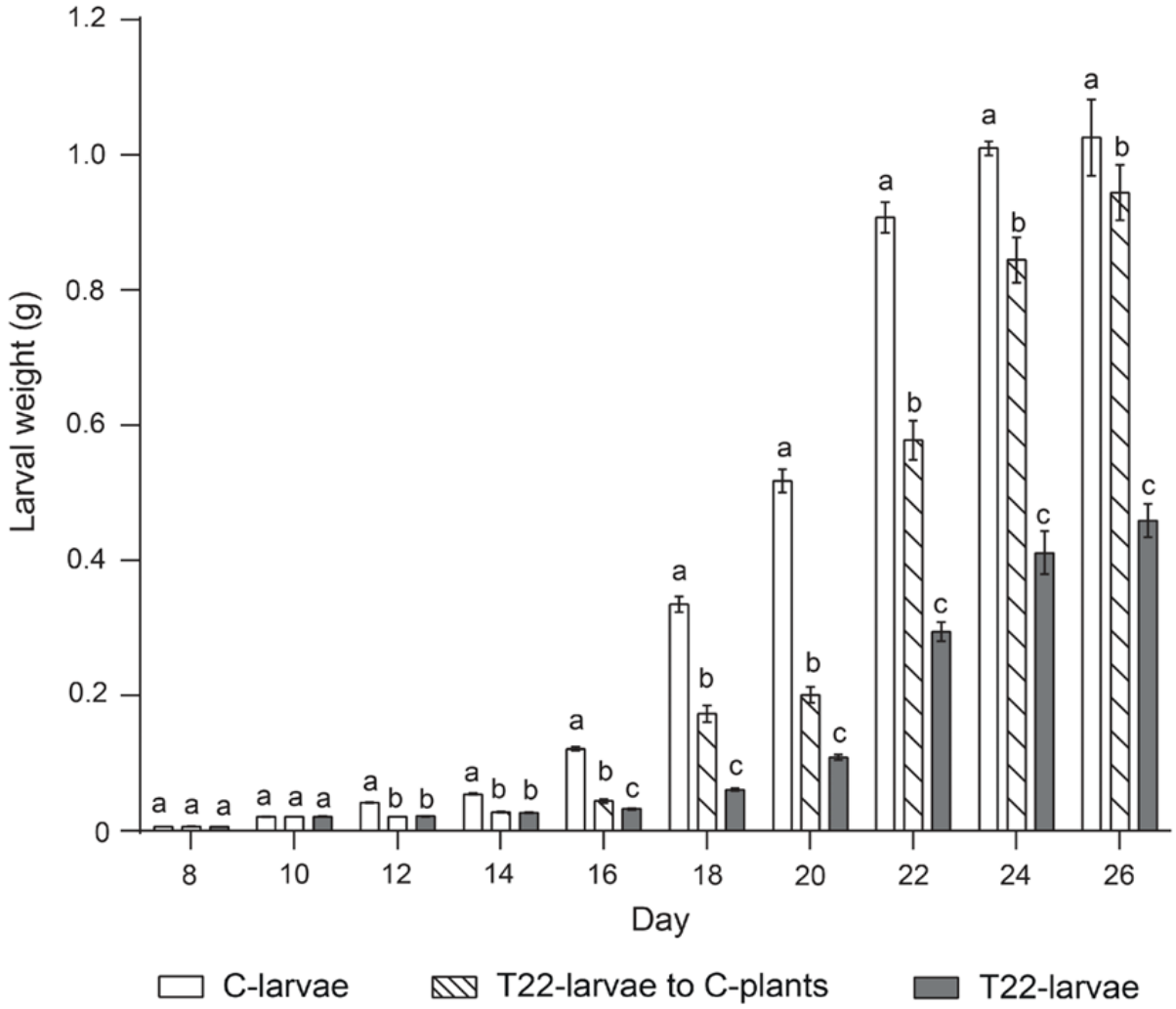

Fig. S6. Weight rescue of *Spodoptera littoralis* T22-larvae fed on C-plant leaves. A significant rescue effect on larval weight was obtained when T22-larvae were transferred to C-plant leaves from day two of the 3rd instar (day eight from hatching). Statistical analyses were performed using One Way ANOVA for the data obtained at day eight and with Kruskal-Wallis test for those obtained at all other days (day 8: $F_{(2, 93)} = 1.053$, *p*-value = 0.3529; day 10: $\chi^2 = 71.87$, *p*-value < 0.0001, dF = 97; day 12: $\chi^2 = 71.41$, *p*-value < 0.0001, dF = 97; day 14: $\chi^2 = 65.02$, *p*-value < 0.0001, dF = 88; day 16: $\chi^2 = 69{,}87$, *p*-value < 0.0001, dF = 85 ; day 18: $\chi^2 = 46.45$, *p*-value < 0.0001, dF = 77; day 20: $\chi^2 = 42.52$, *p*-value < 0.0001, dF=76 ; day 22: $\chi^2 = 47.16$, *p*-value < 0.0001, dF = 71; day 24: $\chi^2 = 46.4$, *p*-value < 0.0001, dF = 71; day 26: $\chi^2 = 46.27$, *p*-value < 0.0001, dF = 69). The values are means ± SE. Different letters indicate mean values that are statistically different.

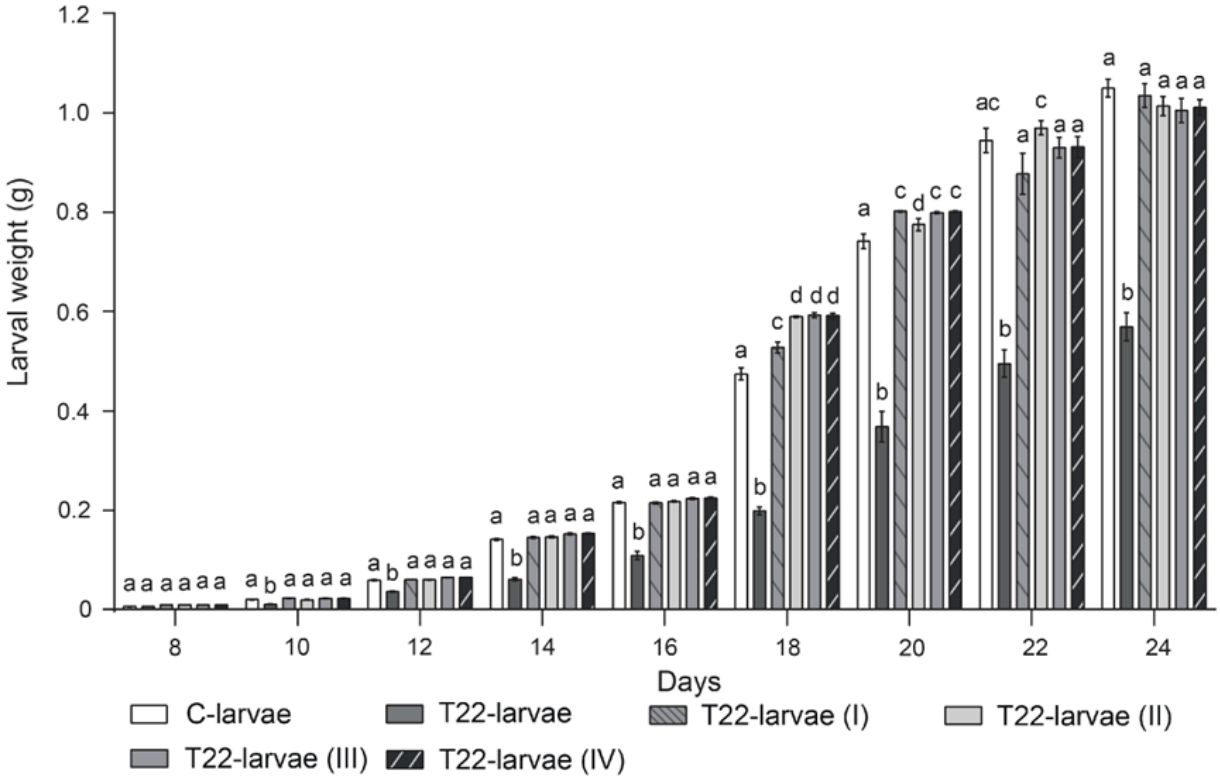

Fig. S7. Weight rescue of *Spodoptera littoralis* T22-larvae after the ingestion of midgut bacteria. A significant rescue effect on larval weight was obtained when T22-larvae were daily offered, from day two of the 3rd instar (day eight from hatching), with T22-plant leaf disks overlaid with the whole midgut microbiota obtained from 4th instar C-larvae (T22-larvae (I)) and with three different amounts of *E. casseliflavus* cells (T22-larvae (II), (III) and (IV)). Statistical analyses were performed using One Way ANOVA or Kruskal-Wallis test (punctual statistics are reported in the *SI Results*, Table S2). Different letters indicate mean values that are statistically different.

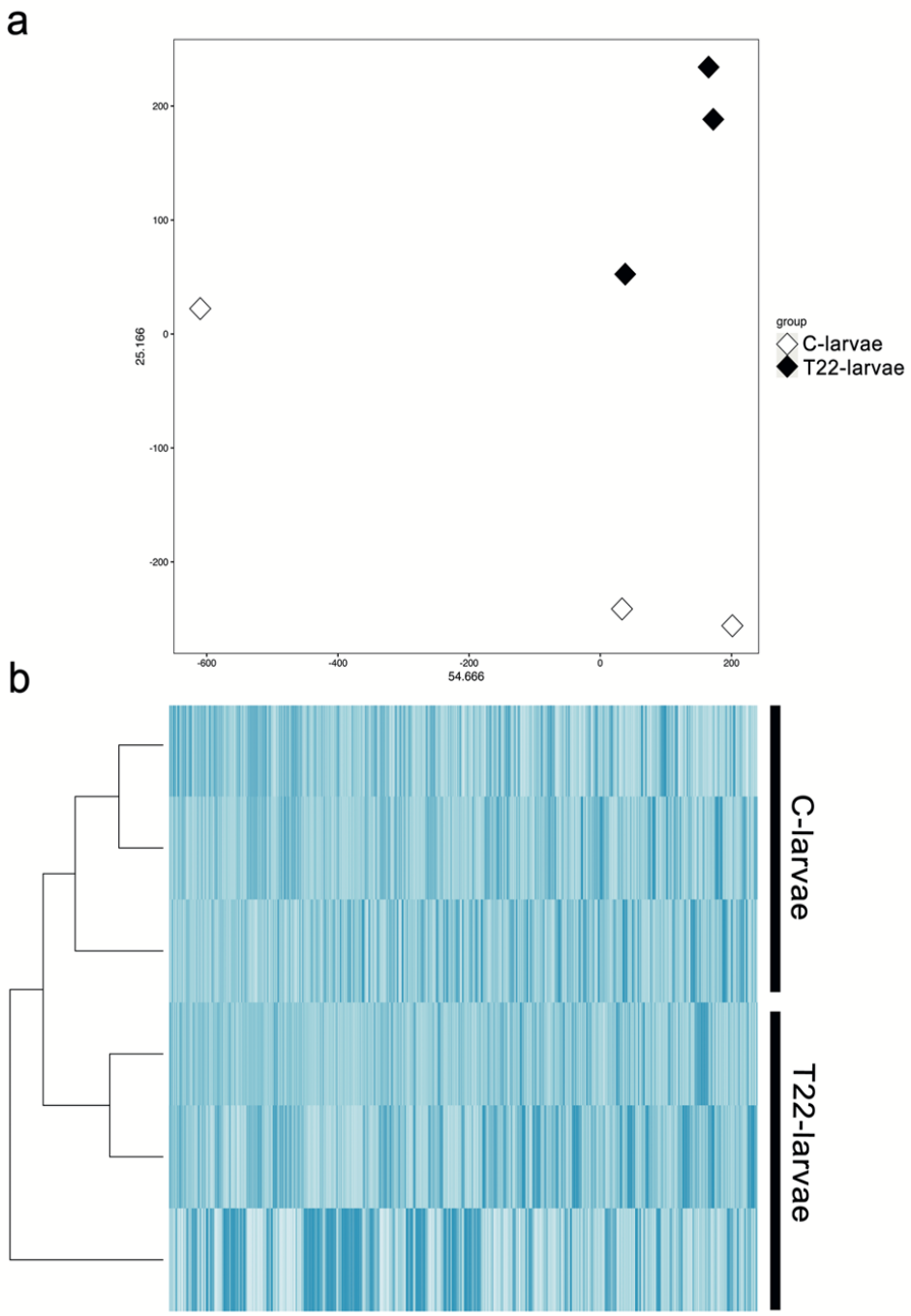

Fig. S8. Multivariate analyses on *Spodoptera littoralis* midgut epithelium gene expression in C- and T22-larvae. a Principal Components Analysis of the RNA-sequencing gene expression counts normalized by the median of ratio (MRN); **b** Hierarchical clustering analysis using Spearman distance on gene expression counts normalized by the median of ratio (MRN).

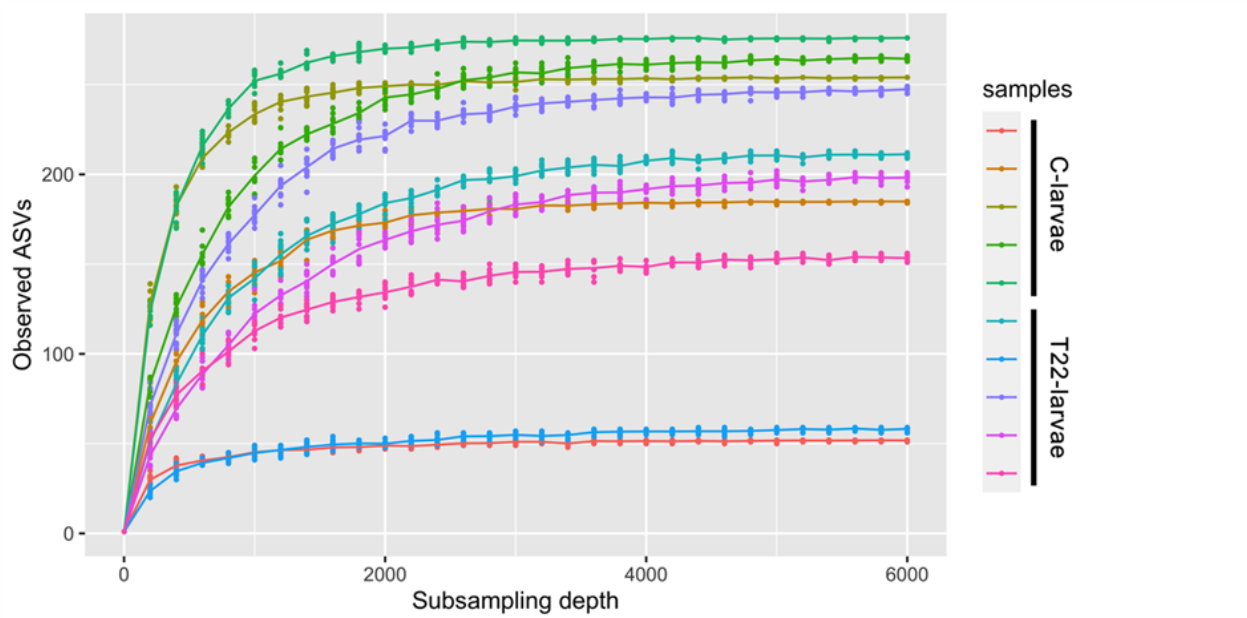

Fig. S9. Rarefaction curves of *Spodoptera littoralis* C- and T22-larvae midgut microbiota. On the x-axis is reported the sequencing effort expressed as number of 16S ribosomal RNA sequences randomly subsampled from the full dataset in each step, while on the y-axis is reported the cumulative number of observed ASVs.

SI Results, Table S1.
Larval weight, expressed in grams, from 3rd instar to pupation.

Day from eggs eclosion	Weight of C-larvae (mean ± SE)	C-larvae n	Weight of T22-larvae (mean ± SE)	T22-larvae n	*t*	dF	p-value
9	0.0035 ± 0.0000502	32	0.0036 ± 0.0004878	32	1.251	62	0.2157
10	0.0074 ± 0.00032	32	0.074 ± 0.00023	32	0.0391	62	0.9683
11	0.0102 ± 0.0004	32	0.0135 ± 0.0046	32	0.7161	62	0.4766
12	0.0261 ± 0.0011	32	0.0201 ± 0.0007	31	0.655	61	< 0.0001
13	0.0415 ± 0.0014	32	0.0312 ± 0.0001	30	5.953	60	< 0.0001
14	0.0755 ± 0.0027	32	0.047 ± 0.0023	27	7.939	57	< 0.0001
15	0.0973 ± 0.0025	32	0.0634 ± 0.0013	26	2.679	56	< 0.01
16	0.1371 ± 0.0038	32	0.0542 ± 0.0033	22	15.63	52	< 0.0001
17	0.1864 ± 0.0081	32	0.0835 ± 0.0062	20	9.296	50	< 0.0001
18	0.2638 ± 0.013	32	0.1116 ± 0.0081	18	8.016	48	< 0.0001
19	0.3717 ± 0.0194	31	0.159 ± 0.0102	17	7.720	46	< 0.0001
20	0.3958 ± 0.0166	31	0.178 ± 0.0101	17	9.176	46	< 0.0001
21	0.5086 ± 0.0195	31	0.263 ± 0.0187	16	8.074	45	< 0.0001
22	0.7223 ± 0.0264	31	0.3424 ± 0.0467	14	7.568	43	< 0.0001
23	0.8067 ± 0.0284	25	0.3749 ± 0.0225	14	10.72	37	< 0.0001
24	0.9978 ± 0.0377	17	0.4995 ± 0.0249	14	10.71	29	< 0.0001
25	1.012 ± 0.0389	5	0.5953 ± 0.03907	14	5.939	17	< 0.0001
26	1.006 ± 0.0323	5	0.7555 ± 0.0258	10	6.08	13	< 0.0001

Abbreviations: SE=standard error; n = number of observations; dF = degrees of freedom.

SI Results, Table S2.
Results of statistical analyses (One Way ANOVA or Kruskal-Wallis) performed for the experiments on the weight rescue of *Spodoptera littoralis* T22-larvae after bacterial supplementation.

Day	Statistical analysis	Statistic	p-value	dF
8	One Way ANOVA	0.195	0.9641	5,186
10	One Way ANOVA	139.3	< 0.0001	5,184
12	Kruskal-Wallis test	97.29	< 0.0001	186
14	Kruskal-Wallis test	74.1	< 0.0001	186
16	Kruskal-Wallis test	66.6	< 0.0001	182
18	Kruskal-Wallis test	98.41	< 0.0001	174
20	Kruskal-Wallis test	69.42	< 0.0001	174
22	Kruskal-Wallis test	25.41	< 0.0001	174
24	One Way ANOVA	59.02	< 0.0001	5,64

Abbreviations: dF = degrees of freedom.

Legend of Supporting Information Results Dataset S1A
Supporting Information – Results, Dataset S1A. Differentially expressed genes in *S. littoralis* midgut.

Legend of Supporting Information Results Dataset S1B
Supporting Information – Results, Dataset S1B. Expression metrics of 1,022 digestion-associated genes in *S. littoralis* midgut. Upregulated genes (LogFC >1.5 and padj < 0.05) are highlighted in red. No genes were found to be significantly downregulated.

Legend of Supporting Information Results Dataset S1C
Supporting Information – Results, Dataset S1C. ASV table reporting the number of reads per sample of each ASV. The taxonomic assignment of each ASV is reported in the taxonomy column.

Legend of Supporting Information Results Dataset S1D
Supporting Information – Results, Dataset S1D. Untargeted metabolomic LC-ESI(+)-MS analysis of the midgut of C- and T22- *S. littoralis* larvae. A t-test (p-value<0.05) was used to identify metabolites whose abundance significantly changed in T22- with respect to C-larvae. Ion retrieved and tentative identification and validation have been carried out by using the XCMS Online tool. For each metabolite, accurate m/z, retention time (RT) and, if present, chemical formula, experimental adduct, metabolic class and organism of origin have been reported. Data have been obtained using 6 biological replicates and are expressed as AVG±ST.DEV. of absolute intensities, and relative content as linear and log2 fold in the comparison T22-/C-larvae. For more details, see SI Methods 7.

SI References

1. G. Glauser, et al., Spatial and temporal dynamics of jasmonate synthesis and accumulation in Arabidopsis in response to wounding. J. Biol. Chem. 283, 16400–16407 (2008).
2. L. Yan, et al., Role of tomato lipoxygenase D in wound-induced jasmonate biosynthesis and plant immunity to insect herbivores. PLoS Genet. 9, e1003964 (2013).
3. E. Franzetti, et al., The midgut of the silkmoth *Bombyx mori* is able to recycle molecules derived from degeneration of the larval midgut epithelium. *Cell Tissue Res.* **361**, 509–528 (2015).
4. M. Bonelli, et al., Black soldier fly larvae adapt to different food substrates through morphological and functional responses of the midgut. *Int. J. Mol. Sci.* **21**, 4955 (2020).
5. E. C. Jensen, Quantitative analysis of histological staining and fluorescence using ImageJ. *Anat. Rec. (Hoboken)* **296**, 378–381 (2013).
6. A. C. Pimentel, A. Montali, D. Bruno, G. Tettamanti, Metabolic adjustment of the larval fat body in *Hermetia illucens* to dietary conditions. *J. Asia-Pac Entomol.* **20**, 1307–1313 (2017).
7. S. Andrews, FastQC: a quality control tool for high throughput sequence data. (http://www.bioinformatics.babraham.ac.uk/projects/fastqc).
8. T. Cheng, et al., Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest. *Nat. Ecol. Evol.* **1**, 1747–1756 (2017).
9. A. Dobin, et al., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
10. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
11. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
12. P. Törönen, A. Medlar, L. Holm, PANNZER2: a rapid functional annotation web server. *Nucleic Acids Res.* **46**, W84–W88 (2018).
13. A. Alexa, J. Rahnenführer, TopGO: enrichment analysis for gene ontology. R package version 2.46.0 (2021).
14. F. Supek, M. Bošnjak, N. Škunca, T. Šmuc, REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS One* **6**, e21800 (2011).

15. D. P. Herlemann, et al., Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* **5**, 1571–1579 (2011).
16. B. J. Callahan, et al., DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. methods*, **13**, 581–583 (2016).
17. F. Pedregosa, et al., Scikit-learn: machine learning in Python. *J. Mach. Learn Res.* **12**, 2825–2830 (2011).
18. N. A. Bokulich, et al., Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 1–17 (2018).
19. C. Quast, et al., The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
20. B. D. Kaehler, et al., Species abundance information improves sequence taxonomy classification accuracy. *Nat. Commun.* **10**, 1–10 (2019).
21. E. Bolyen, et al., Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
22. M. O. Hill, Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427–432 (1973).
23. A. Alberdi, M. T. P. Gilbert, A guide to the application of Hill numbers to DNA-based diversity analyses. *Mol. Ecol. Resour.* **19**, 804–817 (2019).
24. M. Brunetti, G. Magoga, F. Gionechetti, A. De Biase, M. Montagna, Does diet breadth affect the complexity of the phytophagous insect microbiota? The case study of Chrysomelidae. *Environ. Microbiol.* doi: 10.1111/1462-2920.15847 (2021).
25. M. Roswell, J. Dushoff, R. A. Winfree, conceptual guide to measuring species diversity. *Oikos* **130**, 321–338 (2021).
26. A. Chao, et al., Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monogr.* **84**, 45–67 (2014).
27. A. Chao, et al., Rarefaction and extrapolation of phylogenetic diversity. *Methods Ecol. Evol.* **6**, 380–388 (2015).
28. T. C. Hsieh, K. H. Ma, A. Chao, iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).
29. E. A. Franzosa, et al., Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).
30. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
31. H. Mallick, et al., 2021. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* 17(11), p.e1009442 (2021).
32. C. Fasano, et al., Transcriptome and metabolome of synthetic *Solanum* autotetraploids reveal key genomic stress events following polyploidization. *New Phytol*. 210, 1382–1394 (2016).
33. C. A. Smith, E. J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**, 779–787 (2006).
34. R. Tautenhahn, G. J. Patti, D. Rinehart, G. Siuzdak, XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal. Chem.* **84**, 5035–5039 (2012).
35. H. Gowda, et al., Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. *Anal. Chem.* **86**, 6931–6939 (2014).
36. C. Guijas, et al., METLIN: a technology platform for identifying knowns and unknowns. *Anal. Chem.* **90**, 3156–3164 (2018).
37. M. Bergkessel, C. Guthrie, Chapter 25 in *Methods in Enzymology vol 529*, J. Lorsch, Ed (Academic Press, San Diego, 2013), 299-309.
38. U. Edwards, T. Rogall, H. Blöcker, M. Emde, E. C. Böttger, Isolation and direct complete nucleotide determination of entire genes: characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res.* **17**, 7843–7853 (1989).
39. E. Stackebrandt, W. Liesack, Chapter 4 in *Handbook of new bacterial systematics*, M. Goodfellow, A. G. O'Donnell, Eds (Academic Press, London, 1993), 152-189.
40. D. Bulgari, S. Filisetti, M. Montagna, E. Gobbi, F. Faoro, Pathogenic potential of bacteria isolated from commercial biostimulants. *Arch Microbiol.* **204**, 162 (2022).

41. W. W. Hung, et al., Using *groEL* as the target for identification of *Enterococcus faecium* clades and 7 clinically relevant *Enterococcus* species. *J Microbiol. Immunol. Infect.* **52**, 255–264 (2019).