Title of the dataset:

"Data underlying the publication: Relative contributions of egg-associated and substrate-associated microorganisms to black soldier fly larval performance and microbiota"

Creators:

Stijn J.J. Schreven, Laboratory of Entomology, Plant Sciences Group, Wageningen University & Research, P.O. Box 16, 6700 AA, Wageningen, The Netherlands. ORCID https://orcid.org/0000-0002-5837-7025

Contributors:

Hugo de Vries, Laboratory of Microbiology, Agrotechnology & Food Sciences Group, Wageningen University & Research, P.O. Box 8033, 6700 EH, Wageningen, The Netherlands. ORCID https://orcid.org/0000-0002-8888-7202

Gerben D.A. Hermes, Laboratory of Microbiology, Agrotechnology & Food Sciences Group, Wageningen University & Research, P.O. Box 8033, 6700 EH, Wageningen, The Netherlands. ORCID https://orcid.org/0000-0003-4314-9553

Hauke Smidt, Laboratory of Microbiology, Agrotechnology & Food Sciences Group, Wageningen University & Research, P.O. Box 8033, 6700 EH, Wageningen, The Netherlands. ORCID https://orcid.org/0000-0002-6138-5026

Marcel Dicke, Laboratory of Entomology, Plant Sciences Group, Wageningen University & Research, P.O. Box 16, 6700 AA, Wageningen, The Netherlands. ORCID https://orcid.org/0000-0001-8565-8896

Joop J.A. van Loon, Laboratory of Entomology, Plant Sciences Group, Wageningen University & Research, P.O. Box 16, 6700 AA, Wageningen, The Netherlands. ORCID https://orcid.org/0000-0002-4260-0501

Related publication(s):

Schreven SJJ, De Vries H, Hermes GDA, Smidt H, Dicke M, Van Loon JJA (2021). Relative contributions of egg-associated and substrate-associated microorganisms to black soldier fly larval performance and microbiota. *FEMS Microbiology Ecology* 97 (5): fiab054. DOI: https://doi.org/10.1093/femsec/fiab054.

Schreven SJJ, De Vries H, Hermes GDA, Smidt H, Dicke M, Van Loon JJA (2021). Relative contributions of egg-associated and substrate-associated microorganisms to black soldier fly larval performance and microbiota (Chapter 4). In: Schreven SJJ, *Of maggots and microbes*. PhD thesis, Wageningen University and Research, Wageningen, The Netherlands. DOI: https://doi.org/10.18174/542262.

All sequence data have been deposited in the European Nucleotide Archive under the study accession number PRJEB40821 (https://www.ebi.ac.uk/ena/browser/view/PRJEB40821).

Description:

We quantified the relative importance of substrate-associated and black soldier fly (BSF) egg-associated microorganisms on BSF larval performance, bacterial abundance, and bacterial community composition, when larvae were fed with chicken feed or chicken manure. We found that microbes from the feed substrate have a large impact on BSF larval microbiota and performance, whereas microbes from the insect eggs only play a minor role.

Keywords:

16S rRNA gene; amplicon sequencing; qPCR; sterile; *Hermetia illucens*; chicken manure

Files:

Folder / file name	Explanation
Schreven_Ch4_data	Folder containing the datasets used for publication
Ch4_Dataset_1.txt	Data on larval performance parameters and substrate pH.
Ch4_Dataset_1_explanation.txt	Explanation of variables of Dataset 1.
Ch4_Dataset_2.txt	Metadata of 16S rRNA gene amplicon sequencing dataset as deposited in the European Nucleotide Archive.
Ch4_Dataset_2_explanation.txt	Explanation of variables of Dataset 2.
Explanation_diet_specs.txt	Explanation of variables of dataset Schreven_Ch4_diet_specs.txt in input_data subfolder.
Explanation_mocks.txt	Explanation of variables of dataset Schreven_Ch4_mocks_composition.txt in input_data subfolder.
Explanation_qPCR.txt	Explanation of varialbes of dataset Schreven_Ch4_qPCR_data.txt in input_data subfolder.
Schreven_Ch4_R_analysis	Folder containing the R project that includes all analyses of the publication. The .Rproj-file can be opened in Rstudio and will allow use of all R markdown files (.Rmd-files, containing the actual R scripts). Using these R markdown files, the user can produce the Figures and Supplements from the publication. The R scripts have also been exported as PDF-files to read the R script without use of R. R markdown file 3_subsetting.Rmd is essential to run prior to subsequent markdown files in order to create the required RDS-compressed formats of the phyloseq objects.
1. figures	Subfolder containing all figures produced in the R project
2. input_data	Subfolder containing the input files used in the R project:
Schreven_Ch4_contaminant_OTU_by_plot.txt	File with the page numbers referring to ASVs (OTUs) that are likely contaminants, based on the PDF file Schreven_Ch4_ASV_correlations_RelAbd_DNA.pdf in the "figures" subfolder.
Schreven_Ch4_data_performance.txt	File with data on larval performance and substrate pH. Contains same data as Schreven_Ch4_Dataset_1.txt

Schreven_Ch4_diet_specs.txt	File with data on amount of feed and moisture content of diets on day 0.
Schreven_Ch4_mapping_file.csv	File with metadata on sequencing data. Contains same data as Schreven_Ch4_Dataset_2.txt
Schreven_Ch4_mocks_composition.txt	File with theoretical composition of the synthetic mock communities used as positive controls in sequencing.
Schreven_Ch4_qPCR_data.txt	File with qPCR data.
Schreven_Ch4_qPCR_excl_nodoubt.txt	File with a list of sample IDs excluded from analysis of sequencing data based on samples that scored negative or had low quality in qPCR (assessed in qbase+).
Schreven_Ch4_seqdata.biom1	BIOM-file containing the sequence data per sample.
Schreven_Ch4_seqdata_tree.tre	TREE-file containing the phylogenetic relationships between ASVs of the sequences in the BIOM-file.
3. phyobjects	Subfolder containing all RDS-compressed formats of the phyloseq objects used in the R project. These files are created using R markdown file 3_subsetting.Rmd
4. renv	Subfolder required for reproducibility of the R environment (renv package).
5. tables	Subfolder containing all tables produced in the R project

Explanation of the variables:

Dataset_1 and Dataset_2 are companied with separate TXT-files with explanations of variables. Explanation of variables of the data files in the input_data subfolder is also included in the folder Schreven_Ch4_data.

Methods, materials and software:

This information can be found in the publication, see: https://doi.org/10.18174/542262 (Chapter 4 in document).