DMP title

Project Name Karen Bisschop - 12T5622N (FWO DMP) - DMP title
Grant Title 12T5622N
Principal Investigator / Researcher Karen Bisschop
Institution KU Leuven

1. General Information

Name applicant

Karen Bisschop

FWO Project Number & Title

12T5622N The role of the microbiome relative to the genome for niche width and host adaptation.

Affiliation

- KU Leuven
- Universiteit Gent

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Our overall aim is to elucidate the importance of the microbiome for the host's niche width and host adaptation. The questions are whether (i) we can find correlations between the microbiota and the host niche width using field populations, (ii) we can detect an influence of the microbial community on the adaptation of the host to novel host plants with experimental evolution, and (iii) we can discover a direct influence of the microbiota on the niche width of the host using axenic strains and transplant experiments. This will be achieved with the two-spotted spider mite, *Tetranychus urticae* Koch (Arachnida: Acari), as a model species.

The research is divided in three complementary objectives: **Objective 1:** Gaining insight into the influence of the microbiome for host niche width. Using collected spider mite populations across Flanders, we want to discover the main driver for niche width of the different populations in **WP1** (16S, ITS and rbcL amplicon sequencing, completing performance tests, whole-genome sequencing, scoring environmental variables).

Objective 2: Investigating the importance of the microbiome for host adaptation (**WP2**). We will create iso-female lines with the same genetic background that are either having a poor microbiome (due to an antibiotic treatment) or an enriched microbiome (via bacterial inoculation). The spider mite lines are selected based on their overall performance on the different host plant species. We will select spider mite lines from WP1 that are either having a broad niche width (i.e. generalist) or a narrow niche width (i.e. specialist). These spider mite lines will be exposed to challenging host plant species and their performance (fecundity, longevity, leaf depletion, and growth rate) will be measured through time together with their microbiome (16S rRNA sequencing).

Objective 3: Searching for the direct influence of the microbial community on host niche width. In **WP3**, we will perform transplant experiments of the microbial community between different spider mite iso-female lines. We plan to create axenic spider mite lines and inoculate them with their own microbiome or the microbiome of another spider mite strain.

Type of data	**Format**	**Volume**	**How created**
Scoring field observations	xlsx	1 file, max. 2 GB	Field survey of 32 spider mite populations (scoring of e.g. spatial traits) in WP1
Photographs	jpeg	300-500 GB	Daily photographs of performance tests of spider mites in WPs 1, 2, and 3
Amplicon sequencing	fastq	50 GB	16S rRNA, RbcL, and ITS amplicon sequencing in WP1 and WP2
Whole Genome Sequencing	fastq	500 GB	whole genome sequencing Miseq data from 32 populations in WP1

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference:
Short description of the kind of personal data that will be used:

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

1. An extra tab will be provided in the xlsx file including all the metadata to clarify the collected field data and the results from the performance tests.

2. An extra file will be made to clarify which picture corresponds to which specific performance test.

3. The amplicon data will be uploaded in the NCBI GenBank Sequence Read Archive (SRA) database as a project. The necessary documentation will be provided within the project.

4. All R scripts will be properly documented within the script and ReadMe files will be constructed for every publication.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

The metadata will be stored in an Excel file with detailed information for columns, tabs and abbreviations. There wille be Readme files with information about used programs and what to find in which document. This will make it possible to find and reuse all my data by colleagues in the future.

5. Data storage and backup during the FWO project

Where will the data be stored?

1. The time-stamped master copy of the data will be kept on the shared drive (J:) of our research unit. Copies can be made and kept on personal devices.

How is backup of the data provided?

The data will be stored on the university's central servers with automatic daily back-up procedures.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

There is sufficient storage & backup capacity available.

What are the expected costs for data storage and back up during the project? How will these costs be covered?
No extra costs are expected.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?
The data will be stored on the research unit shared drive (J:) which is protected with a login and password. We will not work with sensitive personal data.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).
All data used in publications will be freely available as supplementary files (this refers to xlsx files and R scripts). The amplicon and whole genome sequences will be on the NCBI GenBank Sequence Read Archive (SRA) database. The raw pictures from which the performance data are obtained will be stored on the university's central servers (K:).

Where will the data be archived (= stored for the longer term)?
1. The data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?
1. The data will be hosted on the servers of KU Leuven. In view of the expected size of the database (< 2 TB), no extra costs are expected.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?
All data will be made available after the end of the project.

Where/how will the data be made available for reuse?

- In an Open Access repository

The specific repository is depending on the request of the journal (e.g. Dryad, DataverseNL).

When will the data be made available?

- Upon publication of the research results

Who will be able to access the data and under what conditions?
1. The full dataset will be uploaded as an open access dataset under a CC-BY license. Therefore, it will be available to anyone for any purpose, provided that they give appropriate credit to the creators.

What are the expected costs for data sharing? How will the costs be covered?
No extra costs are expected for data sharing.

8. Responsibilities

Who will be responsible for data documentation & metadata?
The responsible for data documentation and metadata will be Karen Bisschop.

Who will be responsible for data storage & back up during the project?
The responsible for data storate and back up during the project will be Karen Bisschop.

Who will be responsible for ensuring data preservation and reuse ?
The responsible for ensuring data preservation and reuse will be Karen Bisschop.

Who bears the end responsibility for updating & implementing this DMP?
The PI bears the end responsibility of updating & implementing this DMP.