
The evolution of tolerance and cross-tolerance to pesticides in the water flea *Daphnia magna*

A Data Management Plan created using DMPonline.be

Creator: Marlies Van de Maele

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 1171323N

ID: 199165

Start date: 01-11-2022

End date: 31-10-2026

Project abstract:

Freshwater biodiversity is increasingly threatened by pesticide pollution. Rapid evolution of pesticide tolerance is therefore crucial for the persistence of populations and the stability of communities in polluted water bodies, especially if it confers cross-tolerance to other pesticides. Global warming also poses a threat to freshwater biodiversity and moreover magnifies the toxic effects of many pesticides. I will study the evolution of tolerance to single and combined exposure to a pesticide and warming in experimental evolution trials with the water flea *Daphnia magna*, which plays an important role in aquatic food webs, and determine (i) costs of tolerance to each single stressor in terms of increased sensitivity to the other stressor, (ii) evolution of cross-tolerance to other pesticides, and (iii) whether (cross-)tolerance to pesticides reduces the synergism between pesticides and warming. To increase understanding of the evolution of host (cross-)tolerance and costs, which has been identified as a major knowledge gap, I will focus on the causal role of the gut microbiome. I will do so by using gut microbiome transplant experiments between *Daphnia* clones that evolved or not evolved tolerance. I will also directly study the evolution of (cross-)tolerance to pesticides in bacteria using experimental evolution trials, and thereafter, by inoculating germ-free sensitive *Daphnia*, determine whether this contributes to the (cross-)tolerance of the *Daphnia* host to pesticides.

Last modified: 27-04-2023

The evolution of tolerance and cross-tolerance to pesticides in the water flea *Daphnia magna*

Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects ...) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

I will collect/generate (i) datasets (.xls files) including response variables in the experiments (survival, life history traits, bacterial growth parameters), measured pesticide concentrations, and ASV (amplicon sequence variant) tables, (ii) experimental designs and protocols (.docx files), (iii) data analysis scripts (R scripts), (iv) 16S rRNA sequencing reads (.fastq files), (v) figures (.tiff files), (vi) manuscripts and supplementary data (.docx files), and (vii) presentations (.ppt files).

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

During my PhD, I will be the main responsible for data storage, together with Prof. Robby Stoks who will be main responsible after completion of my PhD. Obtained data will be stored on my personal computer, on the LUNA server, in a folder on KU Leuven OneDrive shared with Robby Stoks, and on external hard disks as additional long-term back-up. I will provide an index file (.txt file) for each WP, linking the name, location and description of all files. Published manuscripts and appendices will be saved on LIRIAS, while published data will be shared on data repositories like Dryad. The reads of the 16S rRNA sequencing will be shared on NCBI SRA, while R scripts will be stored on Github.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

I do not wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years. All data will be preserved for the expected 5 year period.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

NA

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

NA

The evolution of tolerance and cross-tolerance to pesticides in the water flea *Daphnia magna*

FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
WP1.1_experimentalevolutiontrial_Daphnia	Data of the measured response variables during the experimental evolution trial: survival, water quality parameters, pesticide concentrations.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP1.1_experimentalevolutiontrial_clones	Microsatellite analysis data	Generate new data	Digital	Experimental	.csv .fsa	< 100 MB	NA
WP1.2_lifetable	Data of measured response variables during the life-table experiment: body size, survival, age and size of first two broods, calculation intrinsic growth rate, pesticide concentrations.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP1.3_cross-tolerance	Data of measured response variables during the experiment: survival, pesticide concentrations.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP2.1_gutmicrobiome	Sequencing data of gut microbiome samples from the life table experiment: 16S rRNA sequences of V4 region. Files obtained from analysing on supercomputer cluster (VSC).	Generate new data	Digital	Experimental	.txt .fastq .qza	< 100 GB	NA
WP2.2_transplant	Data of measured response variables following the gut microbiome transplants: body size, survival, age and size of first two broods, calculation intrinsic growth rate, pesticide concentrations.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP2.2_transplant_gutmicrobiome	Sequencing data of gut microbiome samples following the transplant: 16S rRNA sequences of V4 region. Files obtained from analysing on supercomputer cluster (VSC).	Generate new data	Digital	Experimental	.txt .fastq .qza	< 100 GB	NA
WP3.1_experimentalevolutiontrial_bacteria	Data of the measured response variables during the experimental evolution trial: bacterial growth, pesticide concentrations.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP3.1_experimentalevolutiontrail_bacterial communities	Sequencing data of bacterial community samples in the experimental evolution experiment: 16S rRNA sequences of V4 region. Files obtained from analysing on supercomputer cluster (VSC).	Generate new data	Digital	Experimental	.txt .fastq .qza	< 100 GB	NA
WP3.2_tolerancetest	Data of the measured response variables during tolerance test: bacterial growth, pesticide concentrations, presence/absence organophosphate and carbamate degrading genes.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP4_inoculationexperiment	Data of measured response variables during the life-table experiment following inoculation to test for tolerance to malathion and cross-tolerance: body size, survival, age and size of first two broods, calculation intrinsic growth rate, pesticide concentrations.	Generate new data	Digital	Experimental	.csv	< 100 MB	NA
WP4_inoculation_gutmicrobiome	Sequencing data of gut microbiome samples following the inoculation: 16S rRNA sequences of V4 region. Files obtained from analysing on supercomputer cluster (VSC).	Generate new data	Digital	Experimental	.txt .fastq .qza	< 100 GB	NA
Materials and methods	Information about protocols and experimental design	Generate new data	Digital	Experimental	.pdf	< 100 MB	NA
Data analysis scripts	Script of data analysis: both phenotype (life-history) and gut microbiome analysis	Generate new data	Digital	Experimental	.R	< 1 GB	NA
Manuscripts	Written manuscripts, revisions, revised manuscripts	Generate new data	Digital	Experimental	.pdf	< 100 MB	NA

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

NA

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

A record will be maintained of the following for each WP (if applicable):

- Experimental designs and protocols (.doc)
- Abbreviations used (.doc)
- Data structure (.doc)
- Raw data and analysed data (.xls and .csv)
- Raw sequencing data and analysed data (.fastq, .txt and .qza)
- Script of data analysis with code and extra information about code for analysing phenotypical and sequencing data (.doc and .R)
- Read me file (.txt): name, folder location (OneDrive, local server, hard disk), description of abovementioned files, named according to WP.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

The metadata standard of the research data repository (e.g. Dryad or RDR) will be used.

3. Data storage & back-up during the research project

Where will the data be stored?

The collected data will be stored on a personal computer and on OneDrive.

How will the data be backed up?

The data will be backed up on a local server of the Department of Biology, OneDrive and on a hard disk.

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

The available space will be enough to store all data, as there can be stored up to 2 TB data on OneDrive.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data on personal computer, server and OneDrive are protected by passwords. The back-ups on the hard disks are kept in the office of the PI.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

No extra costs are expected since enough storage is already available for the data.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data will be retained for 10 years (KU Leuven guidelines) after the end of the project.

Where will these data be archived (stored and curated for the long-term)?

All data will be archived on the local server of the Department of Biology and on OneDrive for at least 10 years, conform the KU Leuven RDM policy. Published data will be available online on data repositories such as Dryad, Mendeley Data or RDR, R scripts will be available on GitHub, and the microbiome sequencing data will be stored on the Sequence Read Archive (SRA).

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

There are no extra costs for most data repositories (typically one-time submission cost for Dryad which is mostly covered by the submission journal; Mendeley Data, RDR, GitHub and SRA are free to use).

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

Each dataset will be made available after publication in an online data repository.

If access is restricted, please specify who will be able to access the data and under what conditions.

NA

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Depending on the journal guidelines upon publishing: Dryad, RDR or Mendeley Data for phenotypic data, SRA for microbiome sequencing data.

When will the data be made available?

Upon publication of the results.

Which data usage licenses are you going to provide? If none, please explain why.

Data shared on Dryad will be made available under a Creative Commons Public Domain License (CC0), conform Dryad regulations.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

There are no expected costs for public data repositories, except for Dryad as there is a one-time submission cost which will be covered by the personal bench fee (if the submission journal does not cover these costs).

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Marlies Van de Maele

Who will manage data storage and backup during the research project?

Marlies Van de Maele

Who will manage data preservation and sharing?

Marlies Van de Maele

Who will update and implement this DMP?

Both Marlies Van de Maele and the PI, Robby Stoks, bear the end responsibility of updating and implementing this DMP.

*