FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information		
Name Grant Holder & ORCID	Sarah Jorissen (0000-0002-3666-2132)	
Contributor name(s) (+ ORCID) & roles	Robby Stoks: promotor (0000-0003-4130-0459)	
	Janne Swaegers: co-promotor (0000-0003-1952-3170)	
Project number 1 & title	Thermal evolution during range expansion of the pace-of-life and the mitochondrial phenotype: a	
	complex life cycle perspective (11PBB24N)	
Funder(s) GrantID ²	FWO	
Affiliation(s)	⊠ KU Leuven	
	☐ Universiteit Antwerpen	
	☐ Universiteit Gent	
	☐ Universiteit Hasselt	
	□ Vrije Universiteit Brussel	
	□ Other:	
	ROR identifier KU Leuven: 05f950310	

¹ "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

² Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short p	project description
--------------------------	---------------------

Range expansions toward warmer regions provide natural experiments to study the rapid thermal evolution of the pace-of-life (POL). The POL refers to the alignment along a fastslow continuum, whereby fast-paced animals have a faster life history and metabolic rate, but an increased sensitivity to oxidative stress and a shorter lifespan. I will study the thermal evolution and plasticity of the POL during the range expansion of the damselfly Ischnura elegans from France into Spain, and this in an integrated way across its complex life cycle by conducting a common-garden rearing experiment. I will also investigate the thermal evolution and plasticity of the mitochondrial phenotype (mitochondrial density and efficiency) as possible driver of the POL patterns during range expansion. Moreover, by studying the gene-expression levels using RNA-sequencing, I will increase the understanding of the underlying mechanisms contributing to the POL and mitochondrial phenotype. Finally, using whole-genome sequencing data, I will assess whether hybridization of I. elegans with the locally adapted Spanish I. graellsii has contributed to the evolution of I. elegans to the new warmer thermal regime, thereby focusing on the role of adaptive introgression of the POL-associated genes and the entire mitogenome. Identifying the mechanisms driving adaptive divergence is crucial for understanding, modelling and predicting the (evolutionary) trajectories of species in novel environments.

2. 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 ³.

				ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL
		1			DATA		DATA
Dataset Name	Description	New or	Digital or	Digital Data Type	Digital Data	Digital Data	Physical Volume
		Reused	Physical		Format	Volume (MB, GB,	
						TB)	
WP1.1	Data of the measured	⊠ Generate	□ Digital	☐ Audiovisual	.CSV	□ < 1 GB	NA
Phenotyping	response variables during	new data	☐ Physical	☐ Images		□ < 100 GB	
the embryonic	the experiment	☐ Reuse		☐ Sound		□ < 1 TB	
pace-of-life	(phenotyping the pace-of-	existing data		☐ Numerical		□ < 5 TB	
	life variables in damselfly			☐ Textual		□ > 5 TB	
	eggs): developmental rate,			☐ Model		Other:	
	egg size and metabolic rate			☐ Software		< 100 MB	
				⊠ Other:			
				Experimental			
WP1.2	Data of the measured	⊠ Generate	□ Digital	☐ Audiovisual	.csv	□ < 1 GB	NA
Phenotyping	response variables during	new data	☐ Physical	☐ Images		□ < 100 GB	
the larval pace-	the experiment	☐ Reuse		☐ Sound		□ < 1 TB	
of-life	(phenotyping the pace-of-	existing data		☐ Numerical		□ < 5 TB	
	life variables in damselfly			☐ Textual		□ > 5 TB	
	larvae): developmental			☐ Model		Other:	
	rate, growth rate, food			☐ Software		< 100 MB	
	intake, assimilation			⊠ Other:			
	efficiency, conversion			Experimental			
	efficiency, metabolic rate						

³ Add rows for each dataset you want to describe.

	and sensitivity to oxidative stress						
WP1.3 Phenotyping the adult pace- of-life	Data of the measured response variables during the experiment (phenotyping the pace-of-life variables in damselfly adults): growth rate, food intake, metabolic rate and sensitivity to oxidative stress	☑ Generate new data☐ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other: Experimental	.CSV	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☑ Other: < 100 MB	NA
WP2. Thermal evolution and plasticity of the mitochondrial phenotype as driver of POL	Data of the mitochondrial respiration of damselfly eggs, larvae and adults: mass of used fibers, CI-LEAK, CI-OXPHOS, Cytochrome C, CI+ProDH-OXPHO, CI+ProDH+CII-OXPHOS, CI+ProDH+CII+G3PDH-OXPHOS, CI+ProDH+CII+G3P-ETS, CIV, CytochromeC_effect, P-L ratio_(RCR ratio), Succinate_contribution, G3P_contribution, Proline_contribution, Cytochroom_increase, P-L ratio_(RCR ratio)_corrected.	⊠ Generate new data □ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other: Experimental	.csv .DLD .docx	□ < 1 GB □ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB ⊠ Other: < 2GB	NA

	Files (.DLD) obtained from the analysis in Oroboros Datlab.						
WP3. Gene expression patterns underlying thermal evolution and plasticity of the POL and the mitochondrial phenotype	RNA-sequencing data about the gene expression levels in the three different life stages and regions. Files (.fastq) obtained from analysing on supercomputer cluster (VSC). Scripts to analyse data (.slurm).	⊠ Generate new data □ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other: Experimental	.slurm files .fastq	□ < 1 GB □ < 100 GB ⊠ < 1 TB □ < 5 TB □ > 5 TB □ Other:	NA
WP4. Adaptive introgression of POL- and mitochondrial phenotype-associated genes, and mitogenome	Available dataset with whole-genome sequencing data of the studied species. Further analysis to investigate the role of adaptive introgression will be done on the supercomputer cluster (VSC), .fastq files. Script to analyse data (.slurm).	☑ Generate new data☑ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other: Experimental	.slurm files .fastq	□ < 1 GB □ < 100 GB ⊠ < 1 TB □ < 5 TB □ > 5 TB □ Other:	NA
Materials and methods	Information about protocols and experimental design	☑ Generate new data☐ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other:	.docx .pdf	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☑ > 5 TB ☑ Other:	NA

				Experimental		< 100 MB	
Data analysis scripts	Script of data analysis: both phenotype (life-history) and gut microbiome analysis	⊠ Generate new data □ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other: Experimental	.R	<pre></pre>	NA
Manuscripts	Written manuscripts, revisions, revised manuscripts	⊠ Generate new data □ Reuse existing data	⊠ Digital □ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☑ Other: Experimental	.docx .pdf	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☑ > 5 TB ☑ Other: < 100 MB	NA

GUIDANCE:

The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.

RDM Guidance on data

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 NA
Are there any ethical issues concerning the	☐ Yes, human subject data; provide SMEC or EC approval number:
creation and/or use of the data	☐ Yes, animal data; provide ECD reference number:
(e.g. experiments on humans or animals, dual	Yes, dual use; provide approval number:
use)? If so, refer to specific datasets or data	⊠No
types when appropriate and provide the	Additional information:
relevant ethical approval number.	
Will you process personal data ⁴ ? If so, please	☐ Yes (provide PRET G-number or EC S-number below)
refer to specific datasets or data types when	
appropriate and provide the KU Leuven or UZ	Additional information:
Leuven privacy register number (G or S number).	
Does your work have potential for commercial	□ Yes
valorization (e.g. tech transfer, for example spin-	⊠ No
offs, commercial exploitation,)?	If yes, please comment:
If so, please comment per dataset or data type	
where appropriate.	
Do existing 3rd party agreements restrict	☐ Yes
exploitation or dissemination of the data you	⊠ No
(re)use (e.g. Material/Data transfer agreements,	If yes, please explain:
research collaboration agreements)?	
If so, please explain to what data they relate and	
what restrictions are in place.	

⁴ See Glossary Flemish Standard Data Management Plan

Are there any other legal issues, such as	☐ Yes
intellectual property rights and ownership, to be	⊠ No
managed related to the data you (re)use?	If yes, please explain:
If so, please explain to what data they relate and	
which restrictions will be asserted.	

	3. Documentation and Metadata			
Clearly describe what approach will be followed	A record will be maintained of the following for each WP (if applicable):			
to capture the accompanying information	Experimental designs and protocols (.docx)			
necessary to keep data understandable and	Abbreviations used (.docx)			
usable , for yourself and others, now and in the	Data structure (.docx)			
future (e.g. in terms of documentation levels and	Raw data and analysed data (.xls, .csv and .DLD)			
types required, procedures used, Electronic Lab	Raw sequencing data and analysed data (.fastq, .pbs, .txt)			
Notebooks, README.txt files, Codebook.tsv etc.	Script of data analysis with code and extra information about code for analysing phenotypical and			
where this information is recorded).	sequencing data (.docx and .R)			
	Read me file (.txt): name, folder location (OneDrive, local server, hard disk), description of			
RDM guidance on documentation and metadata.	abovementioned files, named according to WP.			
Will a metadata standard be used to make it	□ Yes			
easier to find and reuse the data?	⊠ No			
	If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:			
If so, please specify which metadata standard	NA NA			
will be used. If not, please specify which				
metadata will be created to make the data	If no, please specify (where appropriate per dataset or data type) which metadata will be created:			
easier to find and reuse.	The metadata standard of the research data repository (e.g. RDR) will be used.			
REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN				
FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E.				
STANDARD LISTS WITH UNIQUE IDENTIFIERS.				

4. Data Storage & Back-up during the Research Project			
Where will the data be stored?	☐ Shared network drive (J-drive)		
	☐ Personal network drive (I-drive)		
Consult the <u>interactive KU Leuven storage guide</u> to	☐ OneDrive (KU Leuven)		
find the most suitable storage solution for your data.	☐ Sharepoint online		
	☐ Sharepoint on-premis		
	☐ Large Volume Storage		
	☐ Digital Vault		
	☐ Other: NCBI (whole genome sequencing data) & Gene Expression Omnibus (RNA-seq data)		
How will the data be backed up?	☐ Standard back-up provided by KU Leuven ICTS for my storage solution		
	□ Personal back-ups I make (specify)		
WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?	☐ Other (specify)		
	The data will be backed up on the OneDrive of the KULeuven as well as on a hard disk.		
Is there currently sufficient storage & backup	⊠ Yes		
capacity during the project? If yes, specify	□ No		
concisely. If no or insufficient storage or backup capacities are available, then explain how this	The available space will be enough to store all data, as there can be stored up to 2 TB data on OneDrive. If no, please specify: NA		
will be taken care of.	ii iiu, piease specify. NA		

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?	The data on OneDrive is protected by passwords. The back-ups on the hard disks are kept in the office of the PI.
CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE. Guidance on security for research data	
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.

5. 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). Guidance on data preservation	 ✓ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain) 		

Where will these data be archived (stored and curated for the long-term)? Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.	 ⊠ KU Leuven RDR □ Large Volume Storage (longterm for large volumes) □ Shared network drive (J-drive) □ Other (specifiy):
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 (RDR is free to use as well as GitHub for the R scripts).

6. Data Sharing and Reuse		
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available.	 ✓ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) 	
NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-AccessRights	□ Other, please specify:	
If access is restricted, please specify who will be able to access the data and under what conditions.	NA 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 per dataset or data type where appropriate.	 Yes, privacy aspects Yes, intellectual property rights Yes, ethical aspects Yes, aspects of dual use Yes, other No If yes, please specify:
Where will the data be made available?	
If already known, please provide a repository	☐ Other data repository (specify)
per dataset or data type.	☐ Other (specify)
When will the data be made available?	 ☑ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
Which data usage licenses are you going to	
provide? If none, please explain why.	Data Transfer Agreement (restricted data)
	☐ MIT licence (code)
A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS	GNU GPL-3.0 (code)
GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY	☐ Other (specify)
REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A	
LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.	
Check the RDR guidance on licences for data and	
software sources code or consult the <u>License selector</u>	
<u>tool</u> to help you choose.	
I	

Do you intend to add a PID/DOI/accession	☑ Yes, a PID will be added upon deposit in a data repository
number to your dataset(s)? If already available,	☐ My dataset already has a PID
please provide it here.	□ No
INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE	
IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.	
What are the expected costs for data sharing?	There are no expected costs for public data repositories such as RDR.
How will these costs be covered?	

7. Responsibilities		
Who will manage data documentation and	Sarah Jorissen	
metadata during the research project?		
Who will manage data storage and backup during the research project?	Sarah Jorissen	
Who will manage data preservation and sharing?	Sarah Jorissen	
Who will update and implement this DMP?	Both Sarah Jorissen and the PI, Robby Stoks, bear the end responsibility of updating and implementing this DMP.	