
On the edge of extinction? Quantifying the evolutionary and conservation implications of forest edge effects in understory plant species

A Data Management Plan created using DMPonline.be

Creator: Lore Hostens

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Lore Hostens

Grant number / URL: 11P8224N

ID: 202687

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

Habitat fragmentation jeopardizes the ability of populations to deal with climate change by compromising the adaptive genetic variation upon which natural selection can act. In forest plant species, this threat may be exacerbated by deteriorating microclimates in forest edges that arise as a consequence of severe habitat fragmentation. With over 20% of European forests being characterized by compromised microclimates, we need an understanding of how edge effects impact the evolutionary potential of understory plant species. Such species are typically long-lived, resulting in low rates of genetic and demographic change after environmental disturbances. This renders them prone to extinction debts and time-delayed evolution; understudied phenomena that are nevertheless critical to long-term forest biological integrity. I hypothesize that (i) understory plant species face range-wide time-delayed extinctions, (ii) forest edge populations show reduced potential to adapt to climate change, and (iii) the presence of core populations can provide evolutionary rescue of forest edge populations. I will address these objectives through range-wide genome analyses of the forest specialists *Paris quadrifolia* and *Circaea lutetiana*. By studying genomic variation in light of past environmental factors in addition to the current microclimate, I will be able to provide novel insights into the spatial distribution of extinction debt and climate resilience within and across European forests.

Last modified: 24-01-2024

On the edge of extinction? Quantifying the evolutionary and conservation implications of forest edge effects in understory plant species

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)

The following data will be collected/generated:

- Silica-dried leaves of populations across Europe of *Circaea lutetiana* and *Paris quadrifolia*
- Genome-wide sequence reads for both species obtained by individual and pooled RADseq or GBS
- Four-five manuscripts reporting the impact of habitat fragmentation on the climate vulnerability of the study species
- Soil samples taken from the upper layer of the soil across the extent of the Belgian *Paris quadrifolia* populations
- Soil data (phosphorus, nitrate, soil pH and organic matter) extracted from soil samples

The following existing data will be used:

- Macroclimate data from CHELSA
- Microclimate data from ForestTemp
- Historical maps from the Royal Library of Belgium (KBR) and other European researchers
- Landcover maps from ESA WorldCover
- Herbarium specimens provided by the Botanical Garden Meise
- Soil data from SoilGrids

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)

1. Designation of responsible person (If already designated, please fill in his/her name.)
 2. Storage capacity/repository
 - during the research
 - after the research
1. Responsible person for the data management during the project: Lore Hostens (KU Leuven) & Olivier Honnay (KU Leuven).
2. -Kasper Van Acker, a permanent member of the research group will overlook the sample collection after the project has finished
- OneDrive for Business will be used to store all documents related to this project
 - GenBank, an annotated collection of all publicly available DNA sequences, will be used to make the obtained DNA sequences publicly available
 - An extra backup of all genetic markers will be stored on an external hard drive
 - LIRIAS, KU Leuven's institutional repository where publications and associated files are archived
 - All data will be stored on the Teams of the Agronomic Ecology and Conservation Biology research group

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)

Not applicable.

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)

The samples collected in France are subjected to the Nagoya protocol. To align with the protocol, the needed permits are obtained through the ministère de la transition écologique et de la cohésion des territoires.

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

Not applicable.

On the edge of extinction? Quantifying the evolutionary and conservation implications of forest edge effects in understory plant species
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
		Please choose from the following options: <ul style="list-style-type: none"> Generate new data Reuse existing data 	Please choose from the following options: <ul style="list-style-type: none"> Digital Physical 	Please choose from the following options: <ul style="list-style-type: none"> Observational Experimental Compiled/aggregated data Simulation data Software Other NA 	Please choose from the following options: <ul style="list-style-type: none"> .por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ... NA 	Please choose from the following options: <ul style="list-style-type: none"> <100MB <1GB <100GB <1TB <5TB <10TB <50TB >50TB NA 	
Silica dried leaf samples	Leaf samples dried in silica gel for DNA extraction of <i>Paris quadrifolia</i> and <i>Circaea lutetiana</i>	Generate new data	Physical	NA	NA	NA	Max. 50 000 individuals
Herbaria specimens	Plants of both species with leaves dried in an herbarium voucher for DNA extraction	Reuse existing data	Physical	NA	NA	NA	Max. 150 individuals
Pooled allele frequency datasets	For both forest species	Generate new data	Digital	Observational	.txt	<100 GB	NA
Sequence datasets	For both extant and herbarium samples of both forest species	Generate new data	Digital	Observational	.fasta	<1 TB	NA
Genetic variant datasets	For both extant and herbarium samples of both forest species	Generate new data	Digital	Observational	.vcf	<100 GB	NA
Climate data	Data on 19 climate variables from BioClim database	Reuse existing data	Digital	Observational	.tif	<1 GB	NA
Adaptability and extinction debt maps	for both forest species	Generate new data	Digital	Simulation data	.tif	<100 GB	NA

Historical maps	Ferraris map of Belgium available at KBR and maps of European countries provided by researchers of that country or from Old Maps Online	Reuse existing data	Digital	Observational	.tif	<1 GB	NA
ForestTemp	Sub-canopy microclimate temperatures of European forests	Reuse existing data	Digital	Observational	.tif	<1 TB	NA
Land cover maps	Land cover maps from WorldCover esa	Reuse existing data	Digital	Observational	.tif	<1 GB	NA
Soil samples	Soil samples collected from the upper soil layer across the extent of every Belgian Paris quadrifolia population	Generate new data	Physical	NA	NA	NA	NA
Soil data	Data on phosphorus, nitrate, soil pH and organic matter of the collected soil samples	Generate new data	Digital	Observational	.csv	NA	Max 20 kg of soil
Soil data	Data on 5 soil variables from SoilGrids database	Reusing existing data	Digital	Observational	.tif	<1 GB	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:

This project will reuse herbarium specimens of Paris quadrifolia and Circaea lutetiana stored at the Herbarium of Meise Botanic Garden (stored with accession numbers in BGBase) and the corresponding passport data.

WorldClim database: [Bioclimatic variables — WorldClim 1 documentation](#)

ForestTemp: <https://doi.org/10.6084/m9.figshare.14618235>

ESA WorldCover land-use maps: <https://esa-worldcover.org/en>

Historical Ferraris maps from the Royal Library of Belgium (KBR): <https://www.kbr.be/nl/kaart-van-ferraris/>

SoilGrids database: <https://soilgrids.org/>

Old Maps online: <https://www.oldmapsonline.org/>

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.

- Yes

The use of silica dried leaf samples and the derived data from these samples are subjected to the Nagoya protocol, more specifically for the samples collected in France. In accordance with the Nagoya Protocol, the Ministry of Ecological Transition and Territorial Cohesion, signed a declaration providing the KU Leuven access to genetic resources on the national territory of France.

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).

Per work package, output data will be collected, including a text file with a clear description of what the data represent and how they were generated. A text file with a description of the type and format of all input data and scripts used to generate the output data will be kept inside the same folder. The name of the folder will specify the applied algorithm and the date the results were generated (A .txt file explaining the naming will be maintained).

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

For data collected on the field and analysed in the lab, a text file will be made containing the date that the data was collected, the coordinates and/or study area, the people that collected the data, a short explanation of each variable and its units, the file format, and user rights and acknowledgements.

For scripts, a text file will be made containing the date the script was made, the developer of the script, a short explanation of the target of the script, necessary software, necessary input data and explanation of the output, reference to data to which this script was applied, and user rights and acknowledgements.

For the generated results, a text file will be made containing the date that the results were obtained, the coordinates and/or study area, the author of the data, a short description of the results, a reference to the script used to generate the results (if available), the file format, and the user rights and acknowledgements.

For the genetic data obtained by sequencing companies, a pdf report will be provided to detail on the sequencing and bio-informatics procedure.

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

Where will the data be stored?

- OneDrive (KU Leuven)
- Other (specify below)
- Large Volume Storage

The time-stamped master copy of the data will be kept on the Teams of the project and the Teams of the Agronomic Ecology and Conservation Biology research group. Copies can be made and kept on personal OneDrive accounts.

The raw read files obtained by sequencing companies will be deposited on GenBank.

The genetic variant and allele frequency files will be stored on Onedrive, with a copy on a large-volume storage device.

Physical data (herbarium specimens) will be stored at the herbarium of Meise Botanic Garden. The silica-dried leaves and soil samples will be stored at the KU Leuven Kasteelpark Arenberg 31, 3001 Leuven.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

The data storage is backed up via Teams and OneDrive (see above)

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

Yes, personal OneDrive accounts have 2 TB and the Teams 5 TB of storage.

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

The data are on secure servers, with centralized security controls and multi-layered encryption.

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

There are no additional costs expected for data storage and back-up. Only a large volume storage device of 1 TB will be purchased. This cost and any additional costs are covered by Meise Botanic Garden, university/department or bench fee FWO.

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 preserved for 10 years according to KU Leuven RDM policy

At the end of the project, all data will be stored on the Teams of the Agronomic Ecology and Conservation Biology research group.

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

Microsoft Teams

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

The costs will be covered by the project PI and amount max to 1000 euros.

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
- Yes, in a restricted access repository (after approval, institutional access only, ...)

Data, metadata and code will be deposited to online open-access public repositories such as Dryad, Zenodo and Github. Biodiversity can, in addition to the deposition as described above, also be uploaded to biodiversity repositories such as GBIF (<http://gbif.org/>). Sequence read data deposited on GenBank will be embargoed for the duration of the project.

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

The files uploaded to figshare, Zenodo and Github will be public. Data stored on the university's central servers will be accessible after permission of the promoters.

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.

See answer above

When will the data be made available?

Data and code underlying publications in each of the WPs will be deposited at the time of publication as much as possible, such that they are fully open and can be accessed and re-used from the time of publication onwards. Only in the case that data will be (partly) reused in subsequent publications (e.g. some of the data such as the temperature data per plot) or due to other restrictions (e.g. journal policies), a reasonable delay/embargo in access to the data to publish our findings first and/or to secure intellectual property protection might be applied. All data and code will be openly accessible at the latest at the end of the project (if restrictions apply, with a maximum of 6 months embargo beyond that date).

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

CC-BY 4.0 (data)

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.

- Yes

Yes, a PID will be added upon deposit in a data repository.

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

There are no expected costs related to data sharing, apart from data handling and data transfer.

6. Responsibilities

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

Lore Hostens, Dr. Hanne de Kort and prof. Olivier Honnay will be responsible for data documentation and metadata.

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

Lore Hostens, Dr. Hanne de Kort and prof. Olivier Honnay will be responsible for data storage. Data back-up happens automatically.

Who will manage data preservation and sharing?

Lore Hostens, Dr. Hanne de Kort and prof. Olivier Honnay will be responsible for compiling a folder with all data and corresponding metadata that needs to be preserved. Our Teams' data storage team will be responsible for storing the data thereafter.

Who will update and implement this DMP?

Lore Hostens, Dr. Hanne de Kort and prof. Olivier Honnay bear the overall responsibility for updating and implementing this DMP.

On the edge of extinction? Quantifying the evolutionary and conservation implications of forest edge effects in understory plant species

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

On the edge of extinction? Quantifying the evolutionary and conservation implications of forest edge effects in understory plant species

DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable