

Study of the prothoracic gland in the desert locust

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.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		<i>Indicate: N(ew data) or E(xisting data)</i>	<i>Indicate: D(igital) or P(hysical)</i>	<i>Indicate: Audiovisual Images Sound Numerical Textual Model SOftware Other (specify)</i>		<i>Indicate: <1GB <100GB <1TB <5TB >5TB NA</i>	
1	Bulk RNA sequencing	N & E	D	N	Mainly .fastq & .xlsx	<5TB	N/A
2	Single cell RNA sequencing	N	D	N	Mainly .fastq & .xlsx	<5TB	N/A
3	Microscopy	N & E	D	I	Mainly .czi, .tiff, .jpeg & .png	<5TB	N/A
4	RT-qPCR	N & E	D	N	Mainly .xlsx	<1GB	N/A
5	Receptor deorphanization	N & E	D	N	Mainly .xlsx	<1GB	N/A
6	Lipidomics (LC-Q-TOF-MS)	N	D	N	Mainly .raw and .xlsx files	<1TB	N/A
7	Observations	N	D	I & T	Mainly .xlsx, .docx, .tiff, .jpeg & .png	<1GB	N/A
8	GC-MS	N	D	N	Mainly .raw and .xlsx files	<1TB	N/A
9	Peptidomics (MS)	N	D	N	Mainly .raw and .xlsx files	<1TB	N/A
10	Proteomics (MS)	N	D	N	Mainly .raw and .xlsx files	<1TB	N/A
11	Western blots	N	D	I	Mainly .tiff, .jpeg & .png	<1GB	N/A
12	Gel electrophoresis	N	D	I	Mainly .tiff, .jpeg & .png	<1GB	N/A
13	Manuscripts	N	D	T	Mainly .docx,	<1GB	N/A

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:

No

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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- 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 or 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

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).

For each experiment, one folder will be created and stored on the lab NAS (Network Attached Storage), which will contain: the raw data, as directly observed/counted by the researchers or automatically generated by lab instruments with accompanying settings; derived processed/analysed data; and an explanatory document. This explanatory document will include the experimental details of each dataset. If applicable, it will also refer to the software (version) and hardware used to generate it and analyze it. In parallel, all researchers in the lab are also obliged to take detailed written notes about their work (chronologically ordered) in their individual lab notebook, which is to be kept in the lab.

Will a metadata standard be used to make it easier to find and reuse the data?

If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

No general metadata-standard is available for this project (considering the diversity in experimental approaches). A text file will be created per experiment and stored in the corresponding folder, with the precise information regarding the organization and naming (including experiment identification and date) of the dataset files.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Other (specify below)

Before being published, the data will be kept on a Network Attached Storage (NAS) system available at our lab. Upon publication, data will be archived on KU Leuven RDR.

How will the data be backed up?

- Personal back-ups I make (specify below)

The data stored on the lab NAS are regularly backed up to an external device using Synology backup software (frequent incremental backups and periodic total backups). The backup location is separated from the main storage space.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

Yes. Our lab has a NAS system, dedicated solely to data storage. We therefore have storage capacity for the foreseeable future.

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

The security of the data stored on our NAS is ensured by both physical and digital barriers: the data are physically located inside a restricted access area (badge required) and locked in place. The servers are only reachable via SSH through a LAN that is restricted to the lab.

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

Since our NAS is already available and in function, we do not expect significant storage costs.

Data Preservation after the end of the Research Project

Which data will be retained for 10 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

All data will be stored for at least 10 years as previously described, and uploaded to publicly available repositories upon publication.

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

- KU Leuven RDR
- Other (specify below)

The data will be archived on our NAS system as previously described. Upon publication, data will be made publicly available via online repositories and archived on KU Leuven RDR.

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

Since our NAS is already available and in function, we do not expect significant storage costs in this case. Regarding the KU Leuven RDR, every researcher can store 50 GB per year for free, which we expect to be sufficient for this project.

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 restricted data (upon approval, or institutional access only)
- Yes, as open data

Most data will become available when published. Unpublished data can be shared upon request and approval by the PI (Jozef Vanden Broeck).

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

Third parties may access restricted (unpublished) data upon agreement with the PI (Jozef Vanden Broeck).

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.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- KU Leuven RDR (Research Data Repository)
- Other data repository (specify below)

KU Leuven RDR

Public databases (NCBI)

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)

Data that can be shared will be made available under a creative commons attribution license (CC-BY 4.0), so that users have to give credit to the original research authors.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

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

Upon publication, a DOI will be provided through KU Leuven RDR.

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

When applicable, fees for open access publications will be covered by the project. For KU Leuven RDR, every researcher can store 50 GB per year for free, which we expect to be sufficient. There are normally no costs associated with sharing data via online publicly available depositories, such as NCBI databases.

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

The PhD students Stijn Goossens, Rong Liu and Solmaz Ghanbari, under the supervision of the PI Jozef Vanden Broeck.

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

The PhD students Stijn Goossens, Rong Liu and Solmaz Ghanbari, under the supervision of the PI Jozef Vanden Broeck.

Who will manage data preservation and sharing?

The PhD students Stijn Goossens, Rong Liu and Solmaz Ghanbari, under the supervision of the PI Jozef Vanden Broeck.

Who will update and implement this DMP?

The PhD students Stijn Goossens, Rong Liu and Solmaz Ghanbari, under the supervision of the PI Jozef Vanden Broeck.