
DMP_12AKR24N_Watteyne

A Data Management Plan created using DMPonline.be

Creator: Jan Watteyne

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 12AKR24N

ID: 207324

Start date: 01-11-2023

End date: 31-10-2026

Project abstract:

Neuromodulators are potent regulators of brain function and physiology. Large efforts in mapping the expression of these molecules across brains is revealing common features of the brain's chemoarchitecture. Yet, it remains unclear how neuromodulator pathways change throughout development or to what extent non-neuronal cells participate. This project will uncover how the neuromodulator network matures at a systems level by analyzing recent transcriptional atlases of all somatic cells of the nematode *C. elegans* across larval and adult stages. First, the expression of neuromodulators and their receptors will be systematically compared to trace all putative cell-to-cell connections across life stages. Next, I will use network science to investigate the interactions between neuromodulators and synaptic connectivity between neurons, glia, and muscles, which will pinpoint cells where aminergic and/or neuropeptide signaling converge to control behavior and physiology. Finally, neuromodulator signaling will be connected to neural function by assessing to what extent signaling motifs predict functional connectivity in brain-wide neuronal imaging datasets. Doing so for adult worms will guide targeted in vivo experiments to characterize changes in neuromodulator function throughout development. Given neuromodulator pathways show evolutionary consistency in their functions and structure, this work promises to uncover principles of neuromodulator function generalizable across animals.

Last modified: 15-05-2024

DMP_12AKR24N_Watteyne
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
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Generate new data • Reuse existing data 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Digital • Physical 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • .por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ... • NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA 	
Images	Confocal fluorescence microscopy images of reporter transgenes	Generate new data + reuse from previous publications	Digital	Experimental	.dzi, .tif, .tiff, .oib	<1GB	
Processed data files	quantitative data (f.i. confocal images using fluorescent reporters), + (re-) analyses of existing datasets (transcriptome analyses, network models)	Generate new data + reuse from previous publications	Digital	Experimental	csv, .xlsx, .mat, .RData, .rda	<100GB	
data representations	visual representations of quantitative and qualitative data	Generate new data	Digital	Compiled/aggregated data	.ai, .svg, .jpg, .png, .pdf	<100GB	
notebooks	Experimental logbook	Generate new data	Physical (notebook) + Digital (Word + Excel)	Experimental	.xlsx, .doc	<100GB	

software code	code written in MATLAB or Python programming languages to analyze microscopy images and neuromodulator signaling pathways	Generate new data	Digital	Experimental	.m, .py, .ipyn	<100GB	
transgenic <i>C. elegans</i> strains; transformed <i>E. coli</i> bacteria; plasmid DNA	non-pathogenic laboratory organisms frozen for long-term storage + pure plasmid DNA	Generate new data	Physical	Experimental			approx. 6 cryogenic storage boxes (100 samples, 2 mL each) - divided across -20°C and -80°C freezers

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:

Transgenic *C. elegans* strains: public repositories like the Mitani lab; *Caenorhabditis* Genetics Center (CGC) and other labs

Plasmid DNA: other research labs + public repositories like Addgene

C. elegans RNAseq datasets:

- Taylor et al., 2021 *Cell*; Analyzed datasets are available at www.cengen.org. Analysis code is available at [github](https://github.com/cengenproject) <https://github.com/cengenproject>. Raw data available at NCBI GEO (Accession Numbers GSE136049 and GSE169137).
- Roux et al., 2023 *Cell Reports*; Raw counts after filtering, age and cell annotation publicly available at c.elegans.aging.atlas.research.calicolabs.com/. Raw single-cell RNA sequencing data deposited on NCBI GEO under the ID GSE208154.
- Ghaddar et al., 2023 *Science Advances*; web interface to query gene expression at WormSeq (wormseq.org)

Anatomical *C. elegans* datasets:

- Witvliet et al., 2021 *Nature*; Connectivity matrices for all datasets available at <http://www.nemanode.org> and as Supplementary Tables.
- Cook et al., 2019 *Nature*; Connectivity matrices for all datasets available at <http://www.wormwiring.org> and as Supplementary Tables.

Neural activity *C. elegans* datasets:

- Randi et al., 2023 *Nature*; raw data publicly accessible through OSF repository at <https://doi.org/10.17605/OSF.IO/E2SYT>. Interactive browsable versions of the same data are available online at <https://funconn.princeton.edu>
- Atanas et al., 2023 *Cell*;
-
- Original code and data related to neural activity and behavior available through Zenodo <https://doi.org/10.5281/zenodo.8150515> and www.wormwideweb.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

No. Only common and non-pathogenic lab organisms that do not require ethical oversight will be used in this project: *E. coli* and related bacteria; transgenic *C. elegans* worms

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).

All digital data (raw and analyzed data files) will be stored on Desktop File Storage and Large Volume Storage. Digital files will be organized in folders per research objective and experiment, including a txt file with a clear description of what the data represent and how they were generated.

Experimental procedures will be fully documented as word files and in hardcover notebooks. Details on samples, including plasmid maps and strain genotypes, will be archived in excel files with an overview of their location in frozen stock collections.

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.

- Yes

Biological imaging data will be stored following the OME (Open Microscopy Environment) standard to encode metadata on microscopy experiments in image files metadata (*.tif, *.tiff, *.hdf5, *.czi).

Metadata about strain and plasmid collections will be created manually, following the community guidelines as published on the Nomenclature section of the community resource Wormbase (www.wormbase.org).

For software code (MATLAB with *.m; Python with *.py or *.ipynb; R with *.r) core calculation functionality is programmed and commented in the source code. Comments and Notebooks will be used to keep all information necessary to understand and reuse the code. Reference to any additional and necessary data/functions are made directly in the code and/or README.txt files bundled together with the software code.

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

Where will the data be stored?

During the research, all digital data will be stored on Desktop File Storage. The ICT team of the Biology Department will facilitate the technical infrastructure and authentication to access stored files through KU Leuven's Active Directory (Luna). Hardcover notebooks will be kept personally by all researchers involved during the project, and by the PI after the end of the project. Samples will be stored in central stock collections of the lab.

How will the data be backed up?

Network storages (fee-based) for digital data are hosted in the KU Leuven ICTS data center, with incremental backups on at least a daily basis. All data is mirrored to a second ICTS data center.

**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

Network storages for digital data hosted by the KU Leuven ICTS data center can provide sufficient storage and back-up capacity (fee-based).

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

We will not be working with personal, confidential or sensitive data but will ensure data security by storing data at secured KU Leuven Network storages and buildings.

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

Expected costs for data storage and back-up during the project are estimated 3500 EUR, which will be covered by the allocated FWO project bench fee or research project funds.

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, including digital files and samples, will be stored at least 5 years after the end of the project.

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

After the research, digital data will be archived for at least 5 years on Large Volume Storage hosted in the KU Leuven ICTS data center with automatic back-up procedures. Samples will be kept for long-term storage in frozen stock collections (-80 freezer or liquid nitrogen).

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

Expected costs for data storage and back-up after the project are estimated at 3500 EUR, which will be covered by research grant budgets.

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

Digital data and samples resulting from this project will be made available after publication of results.

Software code will be placed on GitHub.com in an open accessible manner to the public upon completion of the project.

Publications resulting from this project will be archived in the KU Leuven Lirias 2.0 repository. Samples can be requested from the principal investigator by mail.

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.

Digital data and samples resulting from this project will be made available after publication of results.

Software code will be placed on GitHub.com in an open accessible manner to the public upon completion of the project.

Publications resulting from this project will be archived in the KU Leuven Lirias 2.0 repository. Samples can be requested from the principal investigator by mail.

When will the data be made available?

Upon publication of the research results

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

Software code:

GNU General Public License 2 or later (GPL-2.0)

You may copy, distribute and modify the software as long as you track changes/dates of in source files and keep all modifications under GPL. You can distribute your application using a GPL library commercially, but you must also disclose the source code.

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

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

Expected costs for data sharing encompass publication fees (estimated 3000 EUR), which will be covered by the research project funds.

6. Responsibilities

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

The main researcher will be primarily responsible for data documentation and metadata management during the project.

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

The main researcher will be responsible for data storage and back-ups on KU Leuven servers during the project.

Who will manage data preservation and sharing?

The principal investigator will be responsible for ensuring data preservation and reuse.

Who will update and implement this DMP?

The main researcher and PI share the end responsibility of updating & implementing this DMP.