
ALIEN NEUROGENESIS IN A DISH: NOVEL TECHNIQUES TO STUDY NEURODEVELOPMENT IN OCTOPUS VULGARIS

A Data Management Plan created using DMPonline.be

Creator: Mark Lassnig  <https://orcid.org/0000-0001-6005-2724>

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 11L1923N

ID: 199048

Start date: 01-11-2022

End date: 31-10-2026

Project abstract:

Within the phylum Mollusca the Cephalopoda arose approximately 530 million years ago. They comprise the shelled nautiloids and the soft-bodied coleoids, for example, cuttlefish and octopus. All cephalopods share a basic organisation of the nervous system: a supra- and a suboesophageal brain mass surrounding the oesophagus, a pair of optic lobes laterally and an axial nerve cord spanning into the appendages. The octopus' nervous system is the most complex one within the cephalopods, comprising more than 40 lobes in the brain. Interestingly, the octopus has independently evolved a complex nervous system, analogous to the human nervous system expansion in the vertebrate clade. The octopus thus represents an alternative, highly structured brain which enables to perform fascinating cognitive tasks. The goal of the PhD project is to investigate the embryonic development of the brain of *Octopus vulgaris*. We aim to elucidate similarities or differences in the mechanisms of neurogenesis compared to vertebrate brains. Therefore, we will deeply characterise the neurogenic zone of the developing embryo by establishing innovative in-vitro models, such as explants and organoids. These systems will allow to disentangle the intrinsic properties of the neurogenic stem cells from external influencing factors and reveal what is needed to establish a complex functional brain.

Last modified: 27-04-2023

ALIEN NEUROGENESIS IN A DISH: NOVEL TECHNIQUES TO STUDY NEURODEVELOPMENT IN OCTOPUS VULGARIS

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)

Used Omics datasets are single-cell seq, single-nuclei seq, single-cell ATAC-seq and bulk seqRNA sequences. Further genomic files include annotation files (gtf), ncbi sequence files (fasta). Collected data consists of video's (avi, mov), image files (tiff and jpg), excel sheets (xlsx), ncbi files (fasta), annotation files (gtf), phylogenetic files (tree). Protocols for experimental work and bioinformatic analysis (docx, pdf, R file) will be stored at Github and protocols.io. Omics datasets will be publically available on GEO and a user-friendly interface (https://scope.aertslab.org/#/Octopus_Brain//welcome). Manuscripts describing these results will be published in appropriate journals.

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

Eve Seuntjens, assisted by ICT staff from Department of Biology. During research, omics data (OD) are stored on the Flemish Super Computer clusters (VSC, HPC Leuven, 50TB, expandable), also on SCRATCH and ARCHIVE. Primary data (PD) in lab notebooks is also stored on LVS (Large Volume Storage) drives. DNA, RNA, protein samples are kept in designated places in laboratory. Published datasets and images will be deposited in publicly accessible databases. Archive costs are budgeted. After research, OD remains at VSC and is shared with other databases (SRA, GenBank). PD is transferred from LVS to KU Leuven Archive drive. Storage drives are in ICTS data center and a mirror in 2nd ICTS data center.

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)

NA

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

ALIEN NEUROGENESIS IN A DISH: NOVEL TECHNIQUES TO STUDY NEURODEVELOPMENT IN OCTOPUS VULGARIS

DPIA

DPIA

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

Question not answered.

ALIEN NEUROGENESIS IN A DISH: NOVEL TECHNIQUES TO STUDY NEURODEVELOPMENT IN OCTOPUS VULGARIS

GDPR

GDPR

Have you registered personal data processing activities for this project?

Question not answered.

ALIEN NEUROGENESIS IN A DISH: NOVEL TECHNIQUES TO STUDY NEURODEVELOPMENT IN OCTOPUS VULGARIS

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 	
multiomics	R scripts	Generate new data	Digital	Experimental	.R	<100MB	
multiomics	Jupyter notebooks	Generate new data	Digital	Experimental	.ipynb	<100MB	
multiomics	figure outputs	Generate new data	Digital	Experimental	.pdf	<1GB	
multiomics	pickle files for scenicplus object	Generate new data	Digital	Experimental	.pkl	<100GB	
multiomics	seurat files for scRNAseq	Generate new data	Digital	Experimental	.rds	<100GB	
ocxplants	HCR imaging	Generate new data	Digital	Observational, experimental	.czi	<1TB	
ocxplants	timelapse imaging	Generate new data	Digital	Observational, experimental	.czi	<10TB	
ocxplants	timelapse videos final	Generate new data	Digital	Observational, experimental	.mp4, .avi	<100GB	
ocxplants	figure generation	Generate new data	Digital	Software	.ai	<100GB	
ocxplants	final figures	Generate new data	Digital	NA	.tif, .jpeg	<100GB	
ocxplants	HCR probes	Generate new data	Digital	Compiled/aggregated data	.xlsx	<100MB	
presentation	powerpoint	Reuse existing data	Digital	Software	.pptx	<100GB	
HCR	probes	Generate new data	Physical	NA	NA	NA	2 drawers in freezer
HCR	slides	Generate new data	Physical	NA	NA	NA	one big fridge
literature	papers	Reuse existing data	Physical/Digital	NA	.pdf	<100GB	

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:

The work is based on the data gathered in previous publications from the Seuntjens lab.
 DOI: 10.7554/eLife.69161
 DOI: 10.1186/s12861-020-00212-6

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).

To keep data on the VSC understandable I produce readme.txt files, which specify what the files actually contain, where the files are coming from and how they were generated. To keep an organization on the personal laptop and about the daily lab routine I am keeping an electronic lab notebook, e.g. on Benchling and Obsidian. In the R scripts I provide further information about specific lines of code following the "#". For wet lab work, I keep a physical lab book, in which I document the necessary information about experiments and where I also refer to general protocols used in the lab.

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

The data will be standardized according to the experiments. The metadata applicable for this project will be created using the Dublin core (<http://www.dcc.ac.uk/resources/metadatastandards/dublin-core>).

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

Where will the data be stored?

The physical data, e.g. HCR slides and HCR probes, will be stored in the lab fridges of the Research group of developmental neurobiology. Digital data in use is temporarily stored on OneDrive, the Personal (I:) and Shared (J:) server at KU Leuven. Long term storage of finalized data will be stored on the LVS (L:) and Archive (K:) servers at KU Leuven.

How will the data be backed up?

For this, the back-up facilities of the KU Leuven ICTS will be used.

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

There is currently sufficient storage at KU Leuven ICTS.

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

I can control who I give access to the files on my personal OneDrive. To access the KU Leuven servers, access is provided and controlled by the group leader, Eve Seuntjens. The KU Leuven ICTS data center hosts the network storage, with a mirror available in the second ICTS center. This ensures additional back-up capacity, recovery of lost data and long term data availability. The access is controlled by KU Leuven security groups and it is password corrected.

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

Back-up cost per Tb (KU Leuven ICTS): 295€/year Expected amount of data (5 Tb).
Digital vault for private data: windows server (KU Leuven ICTS): 1302 €/year.
The costs of storage are covered by internal funds in the Research Group of Developmental Neurobiology, which is managed by Eve Seuntjens.

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 digital data, will be retained for 10 years. Data will be made public as soon as it is published.

Physical data, like HCR slides, cannot be preserved for more than 1 month because the fluorophore activity decreases rapidly and cannot be preserved. Octopus paraffin sections which haven't been used for HCR will be preserved at 4 degrees for 1 year.

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

Digital data will be stored at the Archive (K:) server from KU Leuven ICTS.

HCR probes will be stored in the freezers from the Research Group of Developmental Neurobiology.

Code scripts will be stored on Github.

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

The costs are expected to rise to 3000 EUR/year, which will be covered by internal funds by the Research Group of Developmental Neurobiology, managed by Eve Seuntjens.

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

Written reports will be stored internally and can be accessed upon request. Relevant findings will be made available in publications in high profile, peer-reviewed international journals within the life science field. Data will also be presented on regional and international conferences.

Code scripts will be accessible on GitHub for open access.

Gathered sc-multio-mics datasets will be made available for download after publication for open access.

Raw data concerning the experiments, like .czi, .tiff, .pdf files, will be restrictedly available.

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

Team members of the Research Group of Developmental Biology and collaborators will be able to access the data upon request. If the data will be used, reference to the original publication shall be ensured.

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.

Code scripts will be made available on GitHub.

When will the data be made available?

After publication of the results

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

So far, data usage licenses have not been discussed.

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

Planning to add DOI numbers to the publications.

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

The costs depend on the chosen data repository. Coverage is ensured by internal funds in the Seuntjens lab, managed by Eve Seuntjens.

6. Responsibilities

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

Eve Seuntjens, Adrian Ranga, Mark Lassnig

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

Eve Seuntjens, Adrian Ranga, Mark Lassnig

Who will manage data preservation and sharing?

Eve Seuntjens, Adrian Ranga, Mark Lassnig

Who will update and implement this DMP?

Mark Lassnig