Plan Overview

A Data Management Plan created using DMPonline.be

Title: Validation of host cell factors involved in the nuclear import of LINE-1 and discovery of LINE-1 inhibitors.

Creator: Kune Herrebosch

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Project abstract:

In this project LINE-1 will be studied, the only retrotransposons still active in the human genome. Due to its autonomous ability to reinsert itself into the genome and the reverse transcriptase activity that has recently been established to be present in the cytosol, LINE-1 retrotransposition has been linked to a multitude of different diseases. Therefore, I will perform a high throughput screening in this project, aiming to find a potent LINE-1 inhibitor. As well as study the role of two host factors, TRN-1 and TRN-SR2, and their role in LINE-1 nuclear entry.

ID: 213415

Start date: 01-11-2024

End date: 31-10-2028

Last modified: 01-04-2025

Validation of host cell factors involved in the nuclear import of LINE-1 and discovery of LINE-1 inhibitors.

Application DMP

Questionnaire

The questions in this section should only be answered if you are currently applying for FWO funding.
Are you preparing an application for funding?

- No

Validation of host cell factors involved in the nuclear import of LINE-1 and discovery of LINE-1 inhibitors.

DPIA

DPIA

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

- Not applicable

Validation of host cell factors involved in the nuclear import of LINE-1 and discovery of LINE-1 inhibitors.

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Validation of host cell factors involved in the nuclear import of LINE-1 and discovery of LINE-1 inhibitors.

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
sequencing (sanger and whole plasmid)	Sanger sequencing by LGC genomics. Receiving 800 bp read length of DNA. .AB1 file used to compare DNA in snapgene software. Whole plasmid sequencing via oxford nanopore technology. Receiving analysis report and sequence information	Generate new data Reuse existing *d*ata	digital	experimental	txt csv dna	<1GB	
western blot	Used to detect proteins and separate them based on size. Visualization is done using primary and secondary antibodies.	Generate new data	digital	experimental software	tif jpg	<1GB	
Imagelab	ImageLab is a software from BioRad used to visualize gels stained with a DNA- or protein-binding dye. This will typically be used for checking PCR results, molecular clonings, …	Generate new data	Digital	Experimental	jpg	<100 MB	

BCA	Bovine serum albumin assay is used to determine protein concentrations of lysates. Data is measured using the envision software.	Generate new data	Digital	Experimental	xml	<100 MB	
FACS	Fluorescence-activated cell sorting is used to detect eGFP positive cells as well as live-dead staining	Generate new data	Digital	software experimental	csv pdf	<100 MB	
Imaging	either confocal or HTS	Generate new data	Digital	experimental	czi tiff	<1 TB	
cell lines	Different stable cell lines will be generated to use in experiments	Generate new data Reuse existing *d*ata	Physical	NA	NA	NA	+- 20 cell lines
qPCR	qPCR (quantitative PCR) is a technique used to accurately amplify and quantify DNA sequences in real-time.	Generate new data	Digital	software	csv pdf dwg pcdr	<100 MB	

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:

For sanger sequencing previous data will be used to verify existing plasmids in the lab. These data will be stored on the shared J drive of KULeuven.
Cell lines that were previously generated in the lab will be used as well. These cell lines are descripted in an excel file and also modifications are tracked here.

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.

- Yes

A part of the project is drug discovery that will happen in collaboration with CD3. Active lead compounds will be patented and possibly licensed out to industry. To this purpose we collaborate with IOF manager Dr. Frauke Christ and with LRD to enable efficient valorization of lead compounds

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 in-house collected data will be accompanied by explanation and rationalization in an electronic lab notebook that is kept on the shared J drive. These notebooks will be updated regularly and have a link to the data that is discussed in them.
Data collected not in-house or digital data will be assigned to a specific folder in the FWO fellow's (Kune Herrebosch) KU Leuven OneDrive. This will be easy to find since there is a standard known by everyone working in the Laboratory for Molecular Virology and Gene Therapy.

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

No metadata standard will be used. However, data will be easy to find since a OneDrive notebook will be used in which all dates of experiments are mentioned with a description. In this way, it should be easy to find all data, which are divided into different work packages and subfolders.

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

Where will the data be stored?

Data will be stored at the FWO fellow's (Kune Herrebosch) OneDrive made available by KU Leuven with a capacity of 2TB (can be expanded upon my request). For back-up storage, the L drive (Large Volume Storage) of the Laboratory of Molecular Virology and Gene Therapy will be used.

How will the data be backed up?

The data on the fellow's lab PC is synchronized with the fellow's KU Leuven OneDrive. Data from the FWO fellow's KU Leuven OneDrive will be backed up on the large volume storage of the Laboratory of Molecular Virology and Gene Therapy.

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

J-drive, L-drive, OneDrive KULeuven with a capacity of 2 TB (can be expanded upon my request)

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

Concerning the FWO fellow's lab PC: an antiviral system is installed, no connection will be made with unknown networks and no illegal programs/software will be downloaded.
Concerning the Laboratory of Molecular Virology and Gene Therapy: access is only provided for strictly assigned persons and protected by a password.

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

The expected costs for data storage during the project will be around 1000 euro. Costs will be partially covered both by the research budget provided alongside the FWO fellowship and the laboratory of Molecular Virology and Gene Therapy.

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 obtained data will be archived for 10 years according to KU Leuven RDM policy.

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

All obtained data will be archived on a L drive (Large Volume storage) of the laboratory of Molecular Virology and Gene Therapy at the KU Leuven.

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

The expected costs for data preservation will be around 1000 euro. Costs will be covered by the Laboratory of Molecular Virology and Gene Therapy and by the research budget provided by the
FWO fellowship.

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, …)

Only data which were already published/mentioned in papers will be made available after the end of the project for public use. However, for a possible continuation of this research project, all unpublished data will be available in the laboratory of Molecular Virology and Gene Therapy. In this way, data can still be accessed on demand. In addition, data will be shared between close collaborators when necessary

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

Everyone will be able to request access to published data by e-mail. Nevertheless, prof. Zeger Debyser (PI) and Kune Herrebosch (FWO fellow) will, under consultation, evaluate if the request will be accepted. For example, ongoing collaborations will, by great change, be accepted while competitive labs will be denied access.

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.

All obtained data will be archived at the L drive (Large Volume storage) of the laboratory of Molecular Virology and Gene Therapy at the KU Leuven. Everyone will be able to request access to published data by mail.

When will the data be made available?

Upon publication of the research results or upon request via mail.

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

NA

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?

The expected costs for data sharing will be largely based on the publication of papers. The price of publication is estimated at 3000 euro. This will be covered by the allocated project budget (FWO) and the Laboratory of Molecular Virology and Gene Therapy.

6. Responsibilities

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

Kune Herrebosch

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

Kune Herrebosch

Who will manage data preservation and sharing?

Zeger Debsyer (PI) and Kune Herrebosch

Who will update and implement this DMP?

Kune Herrebosch