DMP title

Project Name Understanding lysosomal dysfunction in the frontotemporal dementia-amyotrophic lateral sclerosis spectrum of neurodegenerative disorders. (FWO DMP) - DMP title

Grant Title G073222N

Principal Investigator / Researcher Philip Van Damme

Project Data Contact Philip Van Damme

Description Frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) are two related incurable neurodegenerative disorders typically striking people in mid-life. Loss of neurons in the frontal and temporal lobes in FTD gives rise to behavioural changes and language disturbances. Loss of neurons of the motor system in ALS causes progressive muscle weakness and wasting. Due to the progressive nature of FTD and ALS it has devastating effect on patients and their families. Although the disease presentation of FTD and ALS is very different, there is substantial overlap in the disease mechanisms leading to neuronal loss. A key feature in about 50% of FTD and in more than 95% of ALS is the aggregation in of a protein called TDP- 43. How aggregation of this protein is brougth about and how it causes neuronal death is not known. Such TDP-43 aggregation is also present in two hereditary forms of FTD and ALS, caused by mutations in the progranulin gene and in the C9orf72 gene, allowing us to study the mechanisms leading to TDP-43 aggregation in stem cell models derived from patients suffering from the disease. We recently observed that the lysosomes look abnormal and dysfunctional in neurons made from such patient-derived stem cells. Lysosomes play an important role in the removal of insoluble proteins such as TDP-43. Therefore, a better understanding of how the lysosomal problems can contribute to the aggregation of TDP-43 will be crucial to devolop treatments for these disorders.

Institution KU Leuven

1. General Information

Name applicant

Philip Van Damme

FWO Project Number & Title

G073222N

Understanding lysosomal dysfunction in the frontotemporal dementia-amyotrophic lateral sclerosis spectrum of neurodegenerative disorders.

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of Data	Format	Volume	How Created
Analysis of confocal images	.xls, .pzfx	100MB-1GB	Quantification of confocal image data performed using ImageJ, in Microsoft excel and graph pad prism.
Analysis of western blot images	.xls, .pzfx	100MB-1GB	Numerical data. Quantification of western blot images, performed using imageJ, in Microsoft excel and graph pad prism.

Confocal images	.LIF , .ometiff, .nd2, .czi, .tif	100GB-4TB	Confocal microscopy images taken using a Leica SP8, Nikon-A1R, Zeiss-LSM900 confocal microscopes, and N-SIM S Super-Resolution Microscope
DNA sequencing	.abi	200-500MB	Sequencing performed by LGC genomics
Electron microscopy (TEM) images	.tif	100GB-1TB	EM images used to determine mitochondrial and lysosome morphology
FACs sort reports	Numerical and multimedia (Flowjo files, .fcs, .TIF, PDF)	200MB-1GB	Acquisitions on flow cytometers (KU Leuven/VIB core) by myself or an operator from the core (only for cell sorting)
Figures of data for publication	.ai, .ppt, .pdf, .tif, .jpg	10-100GB	Figures of amalgamated data produced during the project, created using adobe illustrator or Microsoft PowerPoint.
Immunohistochemistry images	.zen, .ometiff, .tiff	10-100GB	Images of human brain sections analyzed using immunohistochemistry, taken using Zeiss imager.M1 microscope with axio cameras
Lab books	Electronic lab notebook and paper notes	5 lab books + 2GB	Dated written notes associated with carrying out experimental procedures
Mass spectrometry (proteomics) data for analysis	.xls	100MB-50GB	Statistical analysis lists of identified protein targets in mass spectrometry experiments
Mass spectrometry (proteomics) data raw	.raw .mzID	10-1000 GB	Data readable as LC chromatograms, mass spectra or identified protein lists produced using the Max Quant program
Microscopy slides	Glass Microscopy slides	250-1000 (nos.)	Microscopy slides used during imaging, consist of formaldehyde-fixed tissue or cells immunostained or dyed using chemicals.
Plotted graphs	.pzfx	200-500MB	Graphs of data from .xls files produced using GraphPad prism

Protein samples	tubes of liquid containing protein	200-1000	Samples of denatured or undenatured proteins stored at -20°C or -80°C extracted from tissue or cells using detergents.
qRT-PCR data analysed files and statistical analysis	.xls, .pzfx	100-500MB	Analysis of qRT-PCR data performed using Quant Studio software (Thermo fisher), statistical analysis performed in graph pad prism.
qRT-PCR data raw files	.eds, .xls	100-500MB	Raw qRT-PCR data collected using Quant studio 3 thermocycler and associated quant studio software (Thermo Fisher)
Risk assessments (RAs)	.docx, .pdf	50-200MB	Written risk assessments associated with standard operating procedures for experimental procedures performed within the lab
RNA samples	tubes of liquid containing RNA	10-100	RNA samples extracted from tissue or cells
Standard operating procedures (SOPs)	.docx, .pdf	50-200MB	Written protocols for experimental procedures performed in the lab
Text manuscript for publication	.docx	1-20GB	Text files associated with submitted publications
Western blot images	.gel, .ometiff, .tiff	750MB-50GB	Tif images of chemiluminescent signal taken using ImageQuant LAS 4000 instrument.

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference:

Short description of the kind of personal data that will be used:

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, add the reference to the formal approval by the relevant ethical review committee(s)

- Yes

Regarding the work related to induced pluripotent stem cellls (iPSCs) ethical approval was granted by the Ethische Commissie Onderzoek UZ/KU Leuven.
(S50354)

Regarding the work related to mice studies ethical approval was granted by the Ethische

Commissie Proefdierenonderzoek UZ/KU Leuven
(P185/2021)

If post-mortem material will be used, ethical approval was granted to Prof. Dr. Dietmar thal who will be
responsible for the samples.

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- Yes

We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. VIB has a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed.

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- Yes

Yes, resource sharing with external partners was always accompanied by the generation of a material transfer agreement (MTA) between the labs and clearly mentions the restrains.

To my knowledge, the only restraint concerns the isogenic iPSC lines that may not be used for drug screening applications.

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

Type of Data	Associated Documentation
Analysis of confocal images	ROIs drawn in images will be saved as metadata attached to TIFs of images and stored with image analysis
Analysis of western blot images	Names of analyzed images will be stored as part of the image analysis excel files
Confocal images	Microscopy images the following information will be noted: dimensions, image type, bit-depth, pixel sizes, and microscope settings. The methodology and protocol will be described in detail in the lab book. A ReadMe file of the image collection will be written.
DNA sequencing	Sequencing files will be stored with associated plasmid sequences on the lab server. Details of dates of sequencing will be recorded in the lab book.
Electron microscopy (TEM) images	TEM images the following information will be noted: dimensions, image type, bit-depth, pixel sizes, and microscope settings. The methodology and protocol will be described in detail in the lab book. A ReadMe file of the image collection will be written.

FACs sort reports	All associated metadata will be recorded and stored with the raw data
Figures of data for publication	Names of source images and source data will be stored with the created final figures.
Immunohistochemistry images	Microscopy images the following information will be noted: dimensions, image type, bit-depth, pixel sizes, and microscope settings. The methodology and protocol will be described in detail in the lab book. A ReadMe file of the image collection will be written.
Lab books	Lab books will be stored in an agreed location in the lab. Digital scans of lab books will also be taken and stored in the lab server. E-lab notebooks will be stored on the lab server.
Mass spectrometry (proteomics) data analyzed	Analyzed data will be stored with a read me file
Mass spectrometry (proteomics) data raw	Raw data will be associated with the analyzed data described above.
Microscopy slides	The location of stored microscopy slides, the date of their creation, and sample types will be noted in the lab book.
Protein samples	Information on the date of creation, sample source, and concentration will be stored as excel files and in the lab book
qRT-PCR data raw files	The following metadata will be noted: Ct values, concentration, plate overview. The methodology and protocol will be described in detail in the lab book.
RNA samples	Information on the date of creation, sample source, and concentration will be stored as excel files and in the lab book
Western blot images	The following metadata will be noted: dimensions, image type, bit-depth, pixel sizes. The methodology and protocol will be described in detail in the lab book.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- Yes

Type of Data	Metadata standard
Confocal images	OME-TIFF
Electron microscopy (TEM) images	OME-TIFF
FACs sort reports	MIFlowCyt
Immunohistochemistry images	OME-TIFF
Mass spectrometry (proteomics) data raw	mzML
Western blot images	OME-TIFF-GEL

Where no metadata standard exists, metadata will be stored based on the Dublin core standard. The following information will be stored:

Title: free text

Creator: Last name, first name, organization

Date and time reference

Subject: Choice of keywords and classifications

Description: Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) (including version number) used to produce and read the data, the purpose of the experiment, etc.

Format: Details of the file format,

Resource Type: data set, image, audio, etc.

Identifier: DOI (when applicable)

Access rights: closed access, embargoed access, restricted access, open access.

The final dataset will be accompanied by this information in the form of a README.txt or MS Word document. This file will be located in the top-level directory of the dataset and will also list the contents of the other files and outline the file-naming convention used (see section 7 below). This will allow the data to be understood by other members of the laboratory and add contextual value to the dataset for future reuse.

5. Data storage and backup during the FWO project

Where will the data be stored?

Digital files will be stored on KU Leuven servers, with hourly on-site backup and mirroring or stored on a cloud-based service offered by KU Leuven (OneDrive). except for private data that will be stored on KU Leuven secure server (digital vault).
- Tissue samples: Tissues will be stored locally in the laboratory.
- Omics data: omics data generated during the project will either be stored on KU Leuven servers or on The Flemish Supercomputer Centre (VSC), initially in the staging area and later in the archive area.
- Cell lines: Human cell lines will be stored locally in liquid nitrogen cryostorage of the laboratory when actively used for experiments. Animal cell lines will be stored in liquid nitrogen cryostorage in the laboratory.
- Biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.
- Algorithms, scripts, and software: All the relevant algorithms, scripts, and software code driving the project will be stored in a private online git repository from the GitHub account of the department (https://github.com/vibcbd).

How is backup of the data provided?

KU Leuven drives are backed-up according to the following scheme:
- data stored on the “L-drive” is backed up daily using snapshot technology, where all incremental changes in respect of the previous version are kept online; the last 14 backups are kept.
- data stored on the “J-drive” is backed up hourly, daily (every day at midnight), and weekly (at

midnight between Saturday and Sunday); in each case, the last 6 backups are kept.
- data stored on the digital vault is backed up using snapshot technology, where all incremental changes in respect of the previous version are kept online. As standard, 10% of the requested storage is reserved for backups using the following backup regime: an hourly backup (at 8 a.m., 12 p.m., 4 p.m., and 8 p.m.), the last 6 of which are kept; a daily backup (every day) at midnight, the last 6 of which are kept; and a weekly backup (every week) at midnight between Saturday and Sunday, the last 2 of which are kept. Regarding the cell lines and plasmids, a backup vial is always stored at another location (ON4 instead of ON5)

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

Data storage is arranged at a departmental level and can be increased if needed.
There is sufficient storage and backup capacity on all KU Leuven servers:
- The “L-drive” is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp series storage systems, and a CTDB samba cluster in the front-end.
- The “J-drive” is based on a cluster of NetApp FAS8040 controllers with an Ontap 9.1P9 operating system.
- Onedrive provides 2TB of storage data which should be sufficient for (most of) the data generated in the project.

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

(The costs of digital data storage are as follows: 173,78€/TB/Year for the “L-drive” and 519EUR/TB/Year for the “J-drive”.)

The housing of mice typically costs 1 €/day per cage. The mouse colony will require at least 50 cages for 2 years (18500/year). By adding the costs of genotyping, phenotyping, and standard laboratory techniques on brain samples from these mice we will need 22500€/year to execute these experiments.

These costs will be covered by the FWO grant.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The “L-drive” and “J-drive” servers are accessible only by laboratory members and are mirrored in the second ICTS data center for business continuity and disaster recovery so that a copy of the data can be recovered within an hour.

Access to the digital vault is possible only through using a KU Leuven user-id and password, and user rights only grant access to the data in their own vault. Sensitive data transfer will be performed according to the best practices for “Copying data to the secure environment” defined by KU Leuven.
The operating system of the vault is maintained on a monthly basis, including the application of upgrades and security patches. The server in the vault is managed by ICTS, and only ICTS personnel (bound by the ICT code of conduct for staff) have administrator/root rights. A security service monitors the technical installations continuously, even outside working hours.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

The minimum preservation term of 5 years after the end of the project will be applied to all datasets.

Where will the data be archived (= stored for the longer term)?

The data will be stored on the university's central servers (with automatic backup procedures) for at least 10 years, conform the KU Leuven RDM policy.

For all datasets, long term storage will be ensured as follows:
Digital datasets: files will be stored on the “L-drive”.
Biological samples (protein, RNA, cDNA, etc.) will be stored locally in the laboratory.
Tissue samples: Tissues will be stored locally in the laboratory.
Omics data: datasets will be stored on the “L-drive”.
Other biological and chemical samples: storage at 4°C and/or frozen samples in cryovials as appropriate.

If appropriate, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The costs of digital data storage are 3000€ (based on the costs in section 5). These costs are covered by the lab and or the FWO grant.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

Participants to the present project are committed to publish research results to communicate them to peers and to a wide audience. All research outputs supporting publications will be made openly accessible. Depending on their nature, some data may be made available prior to publication, either on an individual basis to interested researchers and/or potential new collaborators, or publicly via repositories (e.g.negative data).

Where/how will the data be made available for reuse?

- In an Open Access repository
- Upon request by mail

When will the data be made available?

- Immediately after the end of the project
- After an embargo period. Specify the length of the embargo and why this is necessary
- Upon publication of the research results

As a general rule all research outputs will be made openly accessible at the latest at the time of publication. No embargo will be foreseen unless imposed e.g. by pending publications, potential IP requirements – note that patent application filing will be planned so that publications need not be delayed - or ongoing projects requiring confidential data. In those cases, datasets will be made publicly available as soon as the embargo date is reached.

Who will be able to access the data and under what conditions?

Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. As detailed above, metadata will contain sufficient information to support data interpretation and reuse and will be conform to community norms. These repositories clearly describe their conditions of use (typically under a Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication or an ODC Public Domain Dedication and Licence, with a material transfer agreement when applicable). Interested parties will thereby be allowed to access data directly, and they will give credit to the authors for the data used by citing the corresponding DOI. For data shared directly by the PI, a material transfer agreement (and a non-disclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted.

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

It is the intention to minimize data management costs by implementing standard procedures e.g. for metadata collection and file storage and organization from the start of the project, and by using free-to-use data repositories and dissemination facilities whenever possible. Data management costs will be covered by the laboratory budget.

8. Responsibilities

Who will be responsible for data documentation & metadata?

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the laboratory notebook that refer to specific datasets, and additionally compiling applicable metadata along with the data in the manner described above.

Who will be responsible for data storage & back up during the project?

The researcher and technical staff will ensure data storage and back up, with support from René Custers and Alexander Botzki for the electronic laboratory notebook and from Raf De Coster for the KU Leuven drives.

Who will be responsible for ensuring data preservation and reuse ?

The PI (Philip van Damme) is responsible for data preservation and sharing, with support from the research and technical staff involved in the project, from René Custers and Alexander Botzki for the electronic laboratory notebook and from Raf De Coster for the KU Leuven drives.

Who bears the end responsibility for updating & implementing this DMP?

The PI bears the end responsibility of updating & implementing this DMP.