Investigating the pathological mechanisms underlying KIF5A mutations in ALS

A Data Management Plan created using DMPonline.be

Creators: Nicole Hersmus, Ludo Van Den Bosch

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Ludo Van Den Bosch

Grant number / URL: G026924N

ID: 209015

Start date: 01-01-2024

End date: 31-12-2027

Project abstract:

Amyotrophic lateral sclerosis (ALS) is an adult-onset neurodegenerative disease caused by the selective degeneration of motor neurons. It is the loss of contact between motor neurons and muscles that causes the paralysis and death of the ALS patient, on average 2 to 5 years after the diagnosis. In 10% of patients, ALS is a familial disease with causative mutations identified in more than 20 genes. Recently, we identified mutations in kinesin family member 5A (KIF5A, encoding kinesin-1). These mutations induce removal of exon 27, leading to the replacement of the C-terminal tail by a completely new tail (KIF5AΔe27). The aim of this proposal is to investigate the molecular changes and mechanisms resulting from mis-splicing of KIF5A. More in particular, we will investigate whether KIF5A mutations produce defects in axonal transport, axonal reinnervation/denervation and the formation of neuromuscular junctions using our iPSC-derived models not overexpressing KIF5A or KIF5AΔe27. As the replacement of the C-terminal sequence affects the liquid-liquid phase separation behavior of the KIF5A protein, we will next investigate the role of this aberrant behavior in the observed phenotypes. As mis-splicing of different mRNAs is more and more linked to ALS, we will investigate whether the mis-splicing of KIF5A also occurs in other forms of ALS. Last but not least, our aim is to reproduce our in vitro observations in an in vivo model. For this, we will use the zebrafish.

Last modified: 19-07-2024

Investigating the pathological mechanisms underlying KIF5A mutations in ALS
DPIA

DPIA

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

- No

Investigating the pathological mechanisms underlying KIF5A mutations in ALS
GDPR

GDPR

Have you registered personal data processing activities for this project?

- No

Investigating the pathological mechanisms underlying KIF5A mutations in ALS
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)

We will generate new data and no personal data will be used. Several different data types will be generated in the different WPs. In several WPs we will systematically investigate the axonal transport of cargos in differentiated neuronal cultures (obtained from iPSCs) as well as in zebrafish. The proportions of antero- and retrogradely moving and stationary organelles, as well as their transport velocity will be quantified using kymographs obtained after time-lapse imaging. This is done using ImageJ and the files contain RGB images (both TIFF and raw) in interleaved format. Image width is the number of pixel in each row of image data and Image Height is the number of rows in the image. Offset to First Image is the number of bytes in the file before the first byte of image data.
We will determine axonal outgrowth in microfluidic chambers. This will be done using ImageJ and the files contain RGB images. Digital images will be acquired from fixed cells using a Leica SP8 confocal microscope. The fluorescent intensity of reporters in the nucleus and cytoplasm was analyzed with ImageJ. The different KIF5A isoforms will be determined by Western blot using specific antibodies. Optical densities of the different bands will be determined using the integrated density measurement tool of ImageJ (NIH).
The data will in several cases be converted to a MS Excel (.xls/.xlsx) format. Commercially obtained control iPSC lines as well as established iPSC-lines from patients with or without mutations in the gene encoding KIF5A will be stored in -80°C. In addition to human satellite cell-derived myoblasts and mesoangioblasts to differentiate myotubes. Also these cells will be stored in -80°C. Estimated size for the obtained digital data: 10 gigaBytes.

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: Nicole Hersmus/Ludo Van Den Bosch
2. Storage capacity/repository
    - during the research: All digital files will be stored on KU Leuven servers, except for private data that will be stored on KU Leuven secure server (digital vault). 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. Both the “L-drive” and “J-drive” servers are accessible only by laboratory members, and are mirrored in the second ICTS datacenter 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. Only de-identified data will be exported from the database by the PI, and stored on KU Leuven servers from where it can be accessed by the research and technical staff from the laboratory. Databases are encrypted, password protected and within KU Leuven firewalls.
      Cells: Human cell lines will be stored locally in the laboratory in liquid nitrogen storage and will be deposited in the future in the UZ Leuven-KU Leuven Biobank. 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).
    - after the research: The minimum preservation term of 5 years after the end of the project will be applied to all datasets. All datasets will be stored on the university's central servers with automatic back-up procedures for at least 5 years, conform the KU Leuven RDM policy. The costs (€156 per TB per year for “Large volume-storage”) will be covered by the budget of the lab.

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)

There are no deviations from the minimum preservation of 5 years.

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)

No

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

None

Investigating the pathological mechanisms underlying KIF5A mutations in ALS
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
Text files	Description of experimental procedures and outcome of the different experiments	New data	Digital	Experimental	Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Mark-up Language (.xml), Adobe Portable Document Format (.pdf), LaTex (.tex) format	<1GB	
Digital images	Images obtained during the different experiments (e.g. immunostainings)	New data	Digital	Experimental	uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	<100GB	
Digital video data	Videos obtained for axonal transport	New data	Digital	Experimental	MPEG-4 High Profile (.mp4), motion JPEG 2000 (.mjp2), Audio Video Interleave (.avi)	<100GB	
Survey data	Datasets obtained in the different WPs	New data	Digital	Experimental	excel (.xls), SPSS (.spss)	<1GB	
Nucleotide sequences	Determination of missplicing (WP4)	New data	Digital	Experimental	raw sequence data trace (.ab1), text-based format (.fasta/.fa) and accompanying QUAL file (.qual), Genbank format (.gb/.gbk)	<1GB	
Sequence alignment data	Determining identity of misspliced KIF5A RNA	New data	Digital	Experimental	(.sam), .bam;	<1GB	
Biological and chemical samples	Frozen samples obtained from cell cultures or from animals (zebrafish)	Samples	Physical	NA	NA	NA	500 samples
Live animals	Zebrafish	Animals	Physical	NA	NA	NA	150 aquaria
Postmortem material	Slices obtained from postmortem material (WP4)	Samples	Physical	NA	NA	NA	50 samples

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:

NA

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.

- Yes, human subject data

S50354: Blood sample and stockage of blood, skin biopsies and liquor of patients with neurodegenerative disease and controles.
S67294: Study the mechanisms of acute and chronic axonal and neuronal degeneration and regeneration, aiming to contribute to the

development of new therapeutic strategies for neurodegenerative disorder. Acronym: Umbrella project for human cell lines

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.

- Yes

Human tissue samples, specifically fibroblasts from ALS patients, along with various cell lines, are utilized in this study. Personal data is pseudonymized and managed securely using a decentralized tool provided by UZ Leuven

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).

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the electronic laboratory notebook (E-notebook) that refer to specific datasets. All datasets will be accompanied by a README.txt file containing all the associated metadata.

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

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

Where will the data be stored?

Protocols are stored on the lab's shared drive. Researchers use ELabJournal in the lab for written notes and to maintain a daily overview of research tasks. Raw data files, including images and videos, will be stored in personal folders on KU Leuven servers, accessible only to members of the Lab of Neurobiology. Samples will be stored in the -80°C freezer in the Laboratory of Neurobiology.

How will the data be backed up?

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.

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 are KUL managed storage systems and back up garanteed by ICTS

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

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

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

The total estimated cost of data storage during the project is 4,000 euro per year. Costs will be covered by our research group

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 datasets will be preserved for a minimum of 5 years after the conclusion of the project. Physical data will be stored for 25 years in accordance with the guidelines for clinical results.

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

For all datasets, long-term storage will be ensured by storing digital datasets on the “L-drive”. Physical data will be stored in the vapor phase of liquid nitrogen for up to 25 years, in accordance with the protocol.

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

The total estimated cost of data storage during the 5 years after the end of the is 20,000 euro. This estimation is based on the different costs described in section 5.

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, …)

Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support data sharing. All research outputs supporting publications will be made accessible upon request by mail.

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

Access will be granted only to researchers who submitted a request via email and received approval

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

As the patients are deceased, GDPR regulations do not apply.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

As of now, there is no available details.

When will the data be made available?

Upon publication

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

A Data Transfer Agreement (DTA) will be utilized.

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

The repository adds automatically a DOI.

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

Up to 50GB it is free of charge.

6. Responsibilities

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

Ludo Van Den Bosch

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

The researcher and the PI.

Who will manage data preservation and sharing?

PI

Who will update and implement this DMP?

The researcher and the PI