DMP title

Project Name DMP FWO RNA - DMP title

Grant Title G0C1620N

Principal Investigator / Researcher Ludo Van Den Bosch

Description Amyotrophic lateral sclerosis (ALS) is a dramatic neurodegenerative disorder characterized by the selective death of the motor neurons in the brain and the spinal cord and leads to the death of the patient on average 2 to 5 years after the detection of the first symptoms. In 1 out of 10 patients, ALS is a familial disease and in the majority of these cases a hexanucleotide repeat expansion in a non-coding region of the C9orf72 gene is responsible for the disease (C9-ALS patients). We have established a transient zebrafish model for this form of ALS and published data strongly indicating that RNA toxicity (induced by the hexanucleotide repeat containing RNA) plays an important role in this ALS model. Preliminary data also indicate that we can counteract this RNA toxicity by co-expression of different RNA-binding proteins. The aim of this proposal is to further characterize this protective effect and to investigate the underlying molecular mechanism using both a transient and a stable zebrafish model. In addition, we will use motor neurons obtained from induced pluripotent stem cells (iPSCs) derived from C9-ALS patients. In a second part, we will check whether the presence of the repeats has an effect on the expression of the RNA-binding proteins using the different models as well as post-mortem material of C9-ALS patients. Overall, we will get a better insight into the role of RNA-binding proteins in ALS, which could lead to new therapeutic targets.

Institution KU Leuven

1. General Information

Name applicant

Ludo Van Den Bosch

FWO Project Number & Title

G0C1620N

Affiliation

- KU Leuven

2. Data description

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

- Generate new data

Describe the origin, type and format of the data (per dataset) and its (estimated) volume, ideally per objective or WP of the project. You might consider using the table in the guidance.

The research will generate the following type of data

WP1: Digital images, Antibodies, synthetic compounds, recombinant compounds, Western Blots
WP2 + 4: Genetically modified organisms, digital images, Antibodies, synthetic compounds,
recombinant compounds, Western Blots
WP3 + 5: Human tissue samples (fibroblasts from ALS patients), cell lines, viral vectors, digital
images, Antibodies, Synthetic compounds, recombinant compounds, Western Blots
WP6: Human tissue samples (spinal cord and brain tissue from ALS patients), Western Blots,
sequencing data, antibodies
For all WP: Research documentation (text, spreadsheets, protocols, notes and diaries), Manuscripts

3. Legal & ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

- No

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

S-60803 is already obtained.
By September 30, 2020 a P-number will be obtained for ECD research.

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?

Research can indeed result in results that have a potential for tech transfer and valorization. No IP restrictions will be claimed

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?

- No

No third-party agreement restricts dissemination or exploitation of the data from this project. In particular, existing agreements between VIB and KU Leuven do not restrict publication of data.

4. Documentation & metadata

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

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? 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

While specific data types might require particular metadata, as a general rule the metadata will be based on a generalized metadata schema such as Dublin Core or DataCite, including the following elements:

- 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 to 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.

For specific datasets, additional metadata will be associated with the data file as appropriate.
When depositing data in a repository, the final dataset will be accompanied by this information under the form of a README.txt 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 & back up during the FWO project

Where will the data be stored?

During: Protocols are stored on lab shared drive. In the lab, researchers use the E-notebook for

written notes and daily overview of research tasks. Raw data files and images/videos will be stored in personal folder on KU Leuven servers - only members of the Lab of Neurobiology have access.
Samples will be stored in the -80°C freezer of the Laboratory of Neurobiology.
After: Data stored in the E-Notebook system will be available for at least 5 years.
Manuscripts: will be published and archived in public repositories. Intermediate analysis files will be on KU Leuven
servers for 5 years.

How is back up 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.

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 sufficient storage and back-up capacity on all KU Leuven servers:
- the “L-drive” is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp eseries storage systems, and a CTDB samba cluster in the front-end.
- the “J-drive” is based on a cluster of NetApp FAS8040 controlers with an Ontap 9.1P9 operating system.

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

The total estimated cost of data storage during the project is 4,000 euro per year. This estimation is based on the following costs:
The costs of digital data storage are as follows: 173,78€/TB/Year for the “L-drive”, 519€/TB/Year for the “J-drive”, 70,00€/TB/Year for VSC-archive and 130,00€/TB/Year for VSC-staging.

These costs are covered by the University.

Data security: 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.

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)?

For all other datasets, long term storage will be ensured as follows:
- Digital datasets: files will be stored on the “L-drive”.

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

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)?

- Yes. Specify:

The University holds the IPR of data apt for exploitation.

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

- Other (specify):

Open-access publications in peer-reviewed journals, including supplemental information

When will the data be made available?

- 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 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 electronic laboratory notebook (E-notebook) that refer to specific datasets.

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

The research 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 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.