DMP title

Project Name My plan (FWO DMP) - DMP title

Project Identifier 12P5922N

Grant Title 12P5922N

Principal Investigator / Researcher Sriram Balusu

Project Data Contact sriram.balusu@kuleuven.vib.be

Description Accumulation of misfolded, aggregated proteins and the associated neuronal cell loss are considered the most fundamental hallmarks of neurodegenerative diseases, including Alzheimer's disease (AD). Despite the prevalent neuronal cell death in AD, the modality of cell death and molecular events leading to such neuronal cell death mechanisms are seldom investigated. Lack of preclinical models that sufficiently recapitulate AD-relevant neuronal cell death has impeded progress in this area. We recently demonstrated the activation of the necroptosis pathway in post-mortem AD brains. Necroptosis markers are highly enriched inside AD lesions, correlated with TAU pathology, and inversely correlated neuronal cell density in the hippocampus & frontal cortex (Koper et al., 2019). However, the contribution of necroptosis towards neuronal cell loss and upstream activators of the pathway is unknown. Hence in the present proposal, we take full advantage of an already established iPSC-derived neuronal xenograft model, which displays classical hallmarks of AD, including TAU pathology and necroptosis activation. We will interfere with the necroptosis critical molecules of the pathway and modulate the neuroinflammation to assess the contribution of necroptosis pathway in neuronal loss. Finally, capitalizing on the cell death phenotype, we will perform in vivo full-genome CRISPR-CAS9 screen to identify the modulators of neuronal cells and validate them in clinical and preclinical models.

Institution KU Leuven

1. General Information

Name applicant

Sriram Balusu

FWO Project Number & Title

12P5922N

Necroptosis pathway activation and its role in neuronal cell loss in Alzheimer's disease

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
Microscopy images	.tif .ZEN	100-200GB	Confocal microscope, and Axioscans from the mouse brains
Analysis of western blot images	.xls, .pzfx	500MB-1GB	Numerical data. Quantification of western blot images, performed using imageJ, in microsoft excel and graph pad prism

FACs sort reports	.CSV	1-2GB	Reports describing fluorescently activated particle sorting of FUS droplets.
DNA sequencing files	.abi	1GB-5GB	Sequencing of plasmids, sequencing of PCR products performed by BGI genomics or VIB nucleomics core.
Microscopy slides	commercial glass slides	200-400	Microscopy slides used during imaging, consisting of formaldehyde-fixed tissue or cells immunostained or dyed using chemicals/antibodies
PCR gel images	.tiff .jpeg	100-200MB	TIFF images of ultraviolet irradiated DNA samples run on agarose gels containing gel green reagent. Images collected using geldoc imaging system.
Western blot images	.gel, .ometiff, .tiff	750-1000MB	750-1000MB
RNA samples	tubes of liquid containing RNA	200-400	RNA samples extracted from tissue or cells using phenol/chloroform extraction or commercial kits

qRT-PCR data analysed files and statistical analysis	.xls, .pzfx	300-500MB	Analysis of qRT-PCR data performed using QBASE, statistical analysis performed in graph pad prism.
Protein samples	tubes of liquid containing protein	200-300	Samples of denatured or undenatured proteins extracted from tissue or cells using detergents.
DNA plasmids	tubes of liquid containing DNA plasmids		DNA plasmids produced during the project, derived from existing DNA plasmids provided by commercial and non-commercial suppliers. New plasmids will be deposited in Adgene.

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

All the experiments related to the mouse are approved by the KU Leuven ethical committee.
ECD 154/2020

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 might possibly identify new targets to treat Alzheimer's disease.

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

4. Documentation and metadata

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

1. 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 Electronic lab notebook. A ReadMe file of the image collection will be written. The file name will contain all the relevant information about the samples and the antibodies used.

2. RNA and cDNA samples: Information on date of creation, sample source, and concentration will be
stored as excel files and in Electronic lab notebooks. All the samples will be named with specific codes corresponding to the mouse or cells that they are derived from.

3. DNA plasmids and Sequencing files: Sequence of plasmids will be stored digitally on the ELN using snapgene file formats. Details of construction will be stored in the ELN. 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.

4. FACS sort: All associated metadata (.FCS files) will be recorded and stored with the raw data.

5. PCR gel 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.

6. Protein samples: Information on date of creation, sample source, and concentration will be stored as excel files and in ELN.

7. RNA samples: Information on date of creation, sample source, and concentration will be stored as excel files and in lab book.

8. 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 of mouse brain tissue	OME-TIFF
Fluorescent imaging files	OME-TIFF
FACs sort reports	MIFlowCyt
Immunohistochemistry images	OME-TIFF
PCR gel images	OME-TIFF
Western blot images	OME-TIFF

Where no metadatastandard 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 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.

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. 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?

All other electronic files (text documents, images, sequences, facs files): will be stored on KU Leuven servers (L-drrive), with hourly on-site backup and mirroring or stored on a cloud-based service offered by KU Leuven (OneDrive). In addition, data will be regularly backed -up and stored on an external hard drive.
All samples will be stored as appropriate: -80°C for nucleic acids, protein samples, will be stored with appropriate backup copies of the stock in the Prof. De Stroopers lab at KU Leuven. Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer). One copy will be submitted to the lab plasmid databse system. All published vectors and the associated sequences will be sent to the non-profit plasmid repository Addgene, which will take care of vector storage and shipping upon request. Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.

How is backup of the data provided?

Data will be stored on OneDrive, which is synced approximately every 10 minutes. Backup to external hard drives will be performed once a month. KU Leuven servers have hourly on-site backup and mirroring.

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

Onedrive provides 1TB of storage data which should be sufficient for the data generated in the project.

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.

- All omics data stored on the Flemish Supercomputer Centre (VSC) will be transferred on a weekly basis to the archive area which is backed up.

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

The costs of the storage of 1-2 TB data for 5 years on KU Leuven servers are maximally €1737.80.
These costs will be met partially by the benchfee provided by the FWO and partially by existing lab grants.

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)?

As a general rule, 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.

For all other datasets, long term storage will be ensured as follows:

- Digital datasets: files will be stored on the “L-drive”.
- Tissue samples: Tissues will be stored locally in the laboratory.
- Omics data: datasets will be stored on the “L-drive” or, for larger datasets, on the Vlaams Supercomputer Centrum.
- Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer) and as a bacteria glycerol stock (-80°C).
- Cell lines: cell lines will be stored locally in the laboratory (-80°C).
- Genetically modified organisms: lines that are no actively used for experiments will be cryopreserved.
- Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.

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 project is about 4500€ (for 5 TB). This estimation is based on the different costs described before.

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

As a general rule, datasets will be made openly accessible via existing platforms that support

FAIR data sharing (www.fairsharing.org). Sharing policies for specific research outputs are detailed below:

- Omics datasets will be deposited in open access repositories such as the PRIDE Archive for proteomics data, the EMBL-EBI platform for genomics and epigenomics data, the LIPID MAPS Lipidomics Gateway for lipidomics data, the Metabolomics Workbench Data Repository for metabolomics data, or the NCBI Gene Expression Omnibus (GEO) or the EBI ArrayExpress databases for functional genomics data.
- Vectors: Upon publication, all vectors supporting a manuscript will be made publicly available via the non-profit plasmid repository Addgene, along with the corresponding DNA sequences. Addgene in turns performs quality control on the DNA, curates the plasmids online with all relevant information (maps, sequences), and for a minimal cost (typically $65) ships the vectors upon simple request and signature of a material transfer agreement. The MTA will be prepared before depositing the vectors with the help of our organization's Tech Transfer office. For transfer between nonprofit or academic institutions, Addgene typically uses the Uniform Biological Material Transfer Agreement (https://www.addgene.org/terms/1047/). All non-published vectors and the associated documentation will be shared by the PI upon request and after signature of a material transfer agreement, at no cost except the cost of shipment.
- Cell lines: All human pluripotent cell lines supporting publications will be registered in hPSCreg, the European human embryonic stem cell registry supported by the European Commission (https://hpscreg.eu/). Information about the deposited lines (including donor information, derivation method, availability and characterization) will also be made accessible. Registration of cell lines in hPSCreg will provide visibility, confirm ethical procurement and facilitate comparison with other hPSC lines. The PI will remain the distributor of the pluripotent cell lines.

All other cell lines supporting publications will be deposited in the American Type Culture Collection (ATCC) database (https://www.atcc.org/), which is a private, non-profit biological resource center. This will provide a secure back-up for this material. Investigators can purchase cell lines from the ATCC database upon signature of a material transfer agreement (https://www.lgcstandards-atcc.org/~/media/PDFs/MTA_2.ashx) and, in some cases, of a Limited Use/Label License (e.g. for CRISPR products or iPSC materials) and/or a Customer Acceptance of Responsibility (for potentially highly pathogenic materials). Information about the cell lines (including organism, cell type, tissue, biosafety level and disease if applicable) will also be made accessible.

- Genetically modified organisms: All genetically modified organisms used in publications will be made available to researchers upon request at the time of publication.
- Antibodies, synthetic and recombinant compounds: samples will be stored as appropriate in the laboratory. Within availability, they will be shared with interested researchers upon request.
- Research documentation: All protocols used to generate published data will be described in the corresponding manuscript(s), and the related documentation will be included as supplementary information. These data and all other documents (daily logs, raw data) deposited in the E-Notebook are accessible to the PI and the research staff, and will be made available upon request.
- Manuscripts: All scientific publications will be shared openly. Manuscripts submitted for publication will be deposited in a pre-print server such as bioRxiv, arXiv, Nature Precedings or ASAPbio). At the time of publication, research results will be summarized on the PI's website (add website address) and post-print pdf versions of publications will be made available there if allowed by copyright agreements, possibly after an embargo as determined by the publisher. Before the end of the embargo or in cases where sharing the post-print is not allowed due to copyright agreements, a pre-print version of the manuscript will be made available. Publications will also be automatically added to our institutional repository, Lirias 2.0, based on the authors name and ORCID ID.
- Nucleic acid and protein sequences: All nucleic acid and protein sequences generated during the project will be stored on KU Leuven servers. Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes), NCBI Gene Expression Omnibus (microarray data / RNA-seq data / CHIPseq data), the Protein Database (for protein sequences).
- Data that do not support publication will be either deposited in an open access repository or made available upon request by email.

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

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?

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.

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 (ELN) 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.