FWO DMP Template - Flemish Standard Data Management Plan

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Céline Vrancx** ORCID: 0000-0003-2594-9020
Contributor name(s) (+ ORCID) & roles	**Wim Annaert**, mentor, ORCID: 0000-0003-0150-9661
Project number & title	**Unraveling a novel role for PSEN2/γ-secretase in late endosome/lysosome-endoplasmic reticulum contact sites: impact on Alzheimer's disease pathogenesis**
Funder(s) GrantID	**12B7423N**
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR (Research Organization Registry Community) identifier when possible:
Please provide a short project description	In the context of Alzheimer's disease (AD), presenilin-2 (PSEN2) is majorly responsible for the generation of toxic intraneuronal β-amyloid owing to its restricted location in late endosomes and lysosomes (LE/Lys). Recent evidence of the lab confined PSEN2 to membrane contact sites (MCSs) of LE/Lys with the endoplasmic reticulum (ER). In-house interactomics identified an interaction of PSEN2 with an ER contact site protein and with nutrient sensing complexes, underscoring a yet unknown role for PSEN2 in inter-organellar communication between LE/Lys and ER. I will combine state-of-the-art super-resolution imaging and spatial interactomics to unravel the integral nano-environment of PSEN2 in LE/Lys-ER MCSs, with a particular attention to explore the functional connection with nutrient sensing and identify substrate involvement. These outcomes will be validated in iPSC-derived neurons and microglia, with a focus on how PSEN2 deficiency vs a familial AD causing mutation impact on endolysosomal homeostasis through modulating organellar communication, and this in both cell-autonomous and non cell-autonomous settings. These novel insights will provide a better understanding on the etiology of endolysosomal dysfunctions at early, preclinical stages of AD.

2. 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. Add rows for each dataset you want to describe.

	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
Digital images	Digital images obtained from electron, confocal and super-resolution microscopy via EM-tags or fluorescently-labelled antibodies; digital images obtained from densitometry analysis of western blots, gel scans; illustrations and figures derived from experimental data sets.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☒ .xml ☐ .tab ☒ .csv ☒ .pdf ☒ .txt ☒ .rtf ☐ .dwg ☒ .tab ☐ .gml ☒ other: .tif/.tiff, .jpg, .jpg2, .bmp, .gif, .svg, .eps, .svg, .ai, .xls/.xlsx.docx/.docx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

Video and audio files	Video recordings will be made from live imaging experiments on the different cell models using fluorescently tagged proteins and organelles.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☒ .xml ☒ .tab ☒ .csv ☒ .pdf ☒ .txt ☒ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .tif/.tiff, .jpg, .jpg2, .bmp, .gif, .svg, .eps, .svg, .avi, .xls/.xlsx.docx/.docx ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Cytometry data	Flow Cytometry and fluorescence-activated cell sorting (FACS) data will be generated for the analyses of organelles (e.g. lysosomal/endosomal pH) as well as the possible phenotypic	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☒ .txt ☐ .rtf ☐ .dwg ☒ .tab ☐ .gml	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

	characterization and isolation of specific cell types.				☒ other: .fcs, .xls/xlsx ☐ NA		
Omics data	This study includes proteomics data to identify the nano-environment of PSEN2; generating DNA, protein, and peptide sequences, as well as text files describing omics analyses and representative lists for quantifications.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☒ .xml ☐ .tab ☐ .csv ☒ .pdf ☒ .txt ☐ .rtf ☐ .dwg ☒ .tab ☐ .gml ☒ other: .fasta/.fa, .qual, .gb/.gbk, .xls/xlsx ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Vectors	Bacterial vectors, mammalian expression vectors, viral vectors and shuttling vectors will be used to generate molecular tools to alter the expression of PSEN2 and its interactors in the different cell	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

	models, in order to study their co-localization and dynamics in intracellular trafficking. The work on recombinant-DNA is covered by an environmental permit (number: D/PMVC/00A13/28156) and a biosafety authorization (number: AMV/18092017/SBB219.2017/0518).				☒ other: .fasta/.fa, .qual, .gb/.gbk, .xls/xlsx ☐ NA		
Cell lines	Bacterial strains will be used for the production of expression vectors (DNA); induced pluripotent stem cells (iPSCs) and stable human cell lines will be used to study the functions of PSEN2 at the cellular and molecular levels.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☒ .xml ☐ .tab ☐ .csv ☒ .pdf ☐ .txt ☒ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .doc/.docx, .tex	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Biological samples will be stored physically; frozen cell lines, pellets and organelle extracts, bacterial glycerol stocks, viral particles.

	These iPSC lines will be differentiated to neuronal and microglial cells.				☐ NA		
Organisms and tissue samples	This project may require the use of primary neuronal and microglial cells isolated from murine animals. The experiments planned in this project have all been approved by the institutional Ethical Committee of the KULeuven for Animal Experimentation and designated as project P173-2022.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☒ .xml ☐ .tab ☐ .csv ☒ .pdf ☐ .txt ☒ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .doc/.docx, .tex ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Biological samples will be stored physically: primary cells cultured for immunocytochemical and protein analysis.

GUIDANCE:
DATA CAN BE DIGITAL OR PHYSICAL (FOR EXAMPLE BIOBANK, BIOLOGICAL SAMPLES, …).
DATA TYPE: DATA ARE OFTEN GROUPED BY TYPE (OBSERVATIONAL, EXPERIMENTAL ETC.), FORMAT AND/OR COLLECTION/GENERATION METHOD.
EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA (E.G. TEXT & DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.
EXAMPLES OF DATA FORMATS: TABULAR DATA (.POR,. SPSS, STRUCTURED TEXT OR MARK-UP FILE XML, .TAB, .CSV), TEXTUAL DATA (.RTF, .XML, .TXT), GEOSPATIAL DATA (.DWG,. GML, ..), IMAGE DATA, AUDIO DATA, VIDEO DATA, DOCUMENTATION & COMPUTATIONAL SCRIPT.
DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.
PHYSICAL VOLUME: PLEASE ESTIMATE THE PHYSICAL VOLUME OF THE RESEARCH MATERIALS (FOR EX THE NUMBER OF RELEVANT BIOLOGICAL SAMPLES THAT NEED TO BE STORED AND PRESERVED DURING /AFTER THE PROJECT).

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.	-
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, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☒ Yes, animal data ☐ Yes, dual use ☐ No If yes, please describe: The Ethics Committee Research (EC Research) of University Hospitals Leuven (UZ Leuven) has provided ethics clearance for the work on iPSCs (number: S65730). Animal experiments are in accordance with the Belgian and European laws, guidelines and policies for animal experimentation, housing and care (Belgian Royal Decree of 29 May 2013 and European Directive 2010/63/EU on the protection of animals used for scientific purposes of 20 October 2010). The experiments planned in this project, have all been approved by the institutional Ethical Committee of the KU Leuven for Animal Experimentation and designated as project P173-2022.
Will you process personal data? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:

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 ☐ No If yes, please comment: Participants to the present project are committed to publish research results to communicate them to peers and to a wide audience. Results will be published in accordance with ethical guidelines set by the International Committee of Medical Journal Editors. Existing agreements between VIB and KU Leuven do not restrict publication of data. 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 are not delayed
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 to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. 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).	Digital data will be stored on KU Leuven servers and will be made available together with the accompanying metadata, at the latest at the time of publication. The principle of preservation of data and the minimum preservation term of 10 years after the end of the project will be applied without restriction, both to raw and processed data. No embargo will be foreseen unless imposed by e.g. 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. As detailed below, metadata will contain sufficient information to support data interpretation and reuse, and will be conform to community norms. Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. 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 License, 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. In this regard, plasmids can be submitted to Addgene (https://www.addgene.org/depositing/start-deposit/). 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. Before the start of an experiment, suitable metadata standards will be checked at FAIR sharing; using standards as the Minimal Information for Biological and Biomedical Investigations (MIBBI) and OME-TIFF for images. The latter is used by the local OMERO platform. In particular, the following datasets will be stored: **Dataset 2.1 – Research documentation** Research documentation generated or collected from online sources (e.g. Pubmed) and from collaborators, including publications, tutorials, laboratory notes, protocols, animal husbandry data. Data formats: -Text files**:** 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. Estimated yearly storage: 10 MB.

	Dataset 2.2 – Manuscripts Includes text files, illustrations and figures derived and compiled from experimental data. Data formats: - Text files: 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; - Quantitative tabular data: comma-separated value files (.csv), tab-delimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb); - Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif; - Digital images in vector formats: scalable vector graphics (.svg), encapsulated postscript (.eps), Scalable Vector Graphics (.svg), Adobe Illustrator (.ai); - Digital video data: MPEG-4 High Profile (.mp4), motion JPEG 2000 (.mjp2), Audio Video Interleave (.avi). Estimated yearly storage: 150 MB.
Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes that refer to specific datasets. While specific data types may 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 dataset 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: dataset, 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. The final dataset as deposited in the chosen data repository 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/scientific community and add contextual value to the dataset for future reuse.

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	As a rule, digital data will be stored on KU Leuven servers. All omics data generated during the project will be stored on KU Leuven servers or, for larger datasets, on The Flemish Supercomputer Centre (VSC) in the staging area, at first. Upon publication, all omics data supporting a manuscript will be made publicly available via open access repositories such as the PRIDE Archive for proteomics data, the EMBL-EBI platform for genomics and epigenomics data, and the LIPID MAPS Lipidomics Gateway for lipidomics data.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS? DESCRIBE THE LOCATIONS, STORAGE MEDIA AND PROCEDURES THAT WILL BE USED FOR STORING AND BACKING UP DIGITAL AND NON-DIGITAL DATA DURING RESEARCH.*[1] *REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.*	The operating system of the KU Leuven 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. Stored data 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 8am, 12pm, 4pm and 8pm), the last 6 of which are kept; a daily backup at midnight, the last 6 of which are kept; and a weekly backup at midnight between Saturday and Sunday, the last 2 of which are kept. A security service monitors the technical installations continuously, even outside working hours.

[1] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

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 ☐ No If yes, please specify concisely: If no, please specify: We give preference to the use of robust, managed storage with automatic backup. Options include central storage facilities of the research unit, the group or KU Leuven, or a cloud service offered by KU Leuven, all of which have sufficient storage & backup capacity during the project.
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* [1]	Animal administrative, husbandry and animal welfare data are sensitive data and are stored in the LAIS database according to security procedure of KU Leuven. For any other sensitive data, we will abide by the Belgian law on the protection of individuals with regard to the processing of personal data (30th July 2018) and the General Data Protection Regulation 2016/679. The Privacy Team of KU Leuven will be notified before the start of the project research starts and the Data Stewart will therefore: - designate the categories of persons who have access to the sensitive data, with a precise description of their capacity in relation to the processing of these data; - keep the list of the designated categories of persons at the disposal of the competent supervisory authority (Data Protection Authority); - ensure that the designated persons are obliged by a legal or statutory obligation, or by an equivalent contractual provision, to observe the confidential nature of the data concerned.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Our intention is to minimize data storage costs by sharing data, and this through the implementation of standard procedures for e.g. 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. Unless mentioned otherwise, data management costs will be covered by the laboratory budget. **All digital files** Digital files will be stored on KU Leuven servers: - the “L-drive” costs 173,78€/TB/Year. This server is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp e series storage systems, and a CTDB samba cluster in the front-end. Stored data 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.

	- The “J-drive” costs 519€/TB/Year. This server is based on a cluster of NetApp FAS8040 controllers with an Ontap 9.1P9 operating system. Stored data is backed up using snapshot technology where all incremental changes in respect of the previous version are kept online. Backups are performed hourly, daily and weekly; in each case the last 6 backups are kept. Both 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. We will use free-to-use repositories to share digital files, so that there will be no additional cost required to make the data open access. **Vectors** 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. The associated costs are thus minimal (only shipment costs). All other vectors generated during the project will be shared with researchers upon request (handling by the technical staff of the laboratory, shipping costs supported by the receiver). Management of the vector collection is under the responsibility of the PI and the lab manager. Long-term preservation of this collection is of extremely high value for the laboratory, and 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). These will be stored for the remainder of the PI’s research career. Note that any DNA sequences derived from human subjects will be de-identified. **Genetically modified organisms** Maintaining a mouse colony alive costs about 1,200 euro per year (for 6 cages), excluding the costs of genotyping. When no experiment is planned with a particular mouse strain, and in compliance with the 3R’s rule (https://www.nc3rs.org.uk), cryopreservation will thus be used to safeguard the strain, prevent genetic drift, loss of transgene and potential infections or breeding problems. Cryopreservation of sperm/embryos costs about 500 to 700 euro per genotype, plus a minimal annual storage fee (25 euro per strain for 250 to 500 embryos). Frozen specimen are kept in two separate liquid nitrogen tanks at two different sites on campus. When necessary, the costs of revitalization from cryopreserved sperm/embryos are about 1,100/600 euro respectively.

5. 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...).	The data will be stored for a minimum of 10 years, i.e. at least 5 years after the end of the project. After this period, the PI will regularly evaluate whether retention of the data is still necessary and, if applicable, delete data.
Where will these data be archived (stored and curated for the long-term)?	Images will be archived using the OMERO platform. Upon publication, all omics data supporting a manuscript will be made publicly available (and archived) via open access repositories such as the PRIDE Archive for proteomics data, the EMBL-EBI platform for genomics and epigenomics data and the LIPID MAPS Lipidomics Gateway for lipidomics data.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Similarly to the data management costs during the project, data preservation after the end of the FWO project will be covered by the laboratory budget.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATASET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☒ Yes, in an Open Access repository ☐ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☒ Other, please specify: upon request by email
If access is restricted, please specify who will be able to access the data and under what conditions.	-
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 per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify: -
Where will the data be made available? If already known, please provide a repository per dataset or data type.	Metadata will contain sufficient information to support data interpretation and reuse, and will be conform to community norms. Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing.
When will the data be made available? *THIS COULD BE A SPECIFIC DATE (DD/MM/YYYY) OR AN INDICATION SUCH AS 'UPON PUBLICATION OF RESEARCH RESULTS'.*	As previously stated, no embargo will be foreseen unless imposed by e.g. 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.

Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *EXAMPLE ANSWER: E.G. "DATA FROM THE PROJECT THAT CAN BE SHARED WILL BE MADE AVAILABLE UNDER A CREATIVE COMMONS ATTRIBUTION LICENSE (CC-BY 4.0), SO THAT USERS HAVE TO GIVE CREDIT TO THE ORIGINAL DATA CREATORS."* [2]	Datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. 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 License, 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.
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes ☐ No If yes: not available yet
What are the expected costs for data sharing? How will these costs be covered?	As stated above, it is the intention to minimize data sharing 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.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The research and technical staff will generate, collect, process, analyze and store the data listed above, as detailed in the project description. All staff members are committed to conduct high quality research. In particular, standard protocols will be followed to collect data, if needed after appropriate training. Data and methods used will be regularly discussed during team and lab meetings to ensure a high level of confidence in the data generated.

[2] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

Who will manage data storage and backup during the research project?	Regarding data security, transfer of sensitive data 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. Stored data 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 8am, 12pm, 4pm and 8pm), the last 6 of which are kept; a daily backup at midnight, the last 6 of which are kept; and a weekly backup at midnight between Saturday and Sunday, the last 2 of which are kept. A security service monitors the technical installations continuously, even outside working hours.
Who will manage data preservation and sharing?	The PI is responsible for data management. Access to the digital vault is possible only by using a KU Leuven user-ID and password, and user rights only grant access to the data in their own vault.
Who will update and implement this DMP?	The PI bears the overall responsibility for updating & implementing this DMP.