DMP title

Project Name My plan (Internal Funds DMP) - DMP title
Project Identifier u0035360
Grant Title C3/21/049
Principal Investigator / Researcher Joris Vriens
Project Data Contact Joris.Vriens@kuleuven.be
Description The TRPM3 protein is an ion channel that is functionally expressed in peripheral pain-sensing neurons of human. TRPM3 is functionally linked to the pain pathway and genetic ablation or pharmacological inhibition of TRPM3 prevent or revert hypersensitivity associated with inflammation or nerve injury. A collaboration with the centre of drug discovery and development has resulted in the identification of a class of TRPM3 antagonists, including a preclinical candidate drug with efficacy in different preclinical pain models and an excellent safety profile (two patents filed). However, our drug development project still has some important lacunas that hinder partnering and progression toward phase I studies. The aim of this project is to investigate the role of TRPM3 and the long-term efficacy, and safety of our first-in-class TRPM3 antagonist in relevant models of neuropathic pain. Moreover, we will evaluate the possibility that TRPM3 inhibitors can be use prophylactically, to prevent the development of the neuropathic pain. The successful execution of this project will provide the needed preclinical data to advance our program towards clinical studies.
Institution KU Leuven

1. General Information

Name of the project lead (PI)
Joris Vriens

Internal Funds Project number & title
C3/21/049
Validation of TRPM3 as a target for neuropathic pain

2. Data description

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

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

1. Observational data
2. Experimental data

Dataset 2.1. – Digital images

-Digital images of processed neuronal tissue; stored as uncompressed TIFF; total estimated size: 500 GB -Gel scans; stored as uncompressed TIFF; total estimated size: 20 GB

Dataset 2.2. – Video and audio files

-Video files of in vivo and ex vivo calcium imaging in sensory neurons; stored as uncompressed TIFF or AVI; total estimated size: 10 TB Dataset 2.6. – Omics data

-single-cell RNA-seq data of sensory neurons innervating healthy or diseased tissue; stored as .fastq(.gz) files; total estimated size: 10 TB

-Sequence alignment data; stored as .bam files; total estimated size: 500 GB

3. Simulation data
4. Derived and compiled data

Dataset 4.1 – Research documentation

Research documentation in the form of digital lab notebooks, ethical approval documents and protocols Dataset 4.2 – Manuscripts

Manuscripts arising from the research in MS WORD; total estimated size: 1 GB

Dataset 4.3 – Algorithms and scripts

Analysis scripts written in R or Igor; total estimated size: 10 MB

5. Canonical data

These datasets represent an important source of information for the laboratory of the PI (including future staff), for scientists, journalists and higher education teachers working in the field of neuroscience and pain, but also for non-profit organizations and industries active in the field of drug development.

3. Ethical and legal issues

3.1. 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

3.2. 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).
Animals are housed in facilities of the Laboratory Animal Center of KU Leuven, which applies Standard Operation Procedures concerning housing, feeding, health monitoring to assure consistent care in accordance with European and national regulations and guidelines. Animal administrative, husbandry and animal welfare data are sensitive data and are stored in the LAIS database according to security procedure of KU Leuven. The experiments are covered by ECD

3.3. Does your research 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?
We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. Ownership of the data generated belongs to KU Leuven. 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.

3.4. 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 regarding reuse and sharing are in place?

No

4. Documentation and metadata

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

Data will be generated following standardized protocols. 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) and/or in hard copy lab notebooks that refer to specific datasets. The data will be generated following standardized protocols. Clear and detailed descriptions of these protocols will be stored in our lab protocol database, and published along with the results.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

No

Metadata will include 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. Additionally, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards more specific to our data type. For specific datasets, additional metadata will be associated with the data file as appropriate. Give details as needed for the project. Specific examples (adjust as required): - SOPs for biological data generation are kept on a dedicated KU Leuven shared drive. A central excel file is stored on that same drive, detailing for examples: (1) sample ID; (2) SOP with which data generation was performed; (3) abnormalities or deviations from SOP in data generation; (4) experimental QC values (e.g. DNA concentrations); (5) location of the source sample in the freezer. - For bioinformatics processing, a data analysis log will be kept that details: (1) sequencing run ID; (2) the bioinformatics SOPs/scripts that were applied; (3) location of source files; (4) abnormalities or deviations. 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 project

5.1. Where will the data be stored?

Digital files will be stored on KU Leuven servers, except for private data that will be stored on KU Leuven secure server (digital vault). Add information about other data types as required (adjust and add more as needed): - Tissue samples: Tissues will be stored locally in the laboratory. All human tissue samples will be registered with a Belgian biobank, in compliance with the Belgian law on human body material (dd 19-12- 2008). - 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. - 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). 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. - Cell lines: Newly created human cell lines will be stored locally in the laboratory in liquid nitrogen storage and will be deposited in the UZ Leuven-KU Leuven Biobank. Other 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 of the laboratory. - Bacterial and yeast strains will be stored in a -80°C freezer in the lab of xxx. Costs are covered by general lab expenses. - Blood, saliva and fecal samples of study participants will be stored in -80°C freezer facilities with restricted access. - Psychosocial data gathered with the questionnaires will be kept on Qualtrics servers within KU Leuven containing sample ID and SOP with the data generation. We will work with unique identifiers that enable us to link both the biological with the psychosocial and mental health data. The Qualtrics server meets both the criteria stipulated by the Privacy Commission and the ethical committee of the KU Leuven. - Raw as well as derived microbiome data will be stored on a secured Raes lab storage server hosted by VIB (PSB-Gent). - A batch of each EO or EO fraction will be stored in duplicate in two separate refrigerators at KU LeuvenMCB lab under the best possible conditions, but small changes in the composition over time cannot be excluded. - Genetically modified organisms: Mice / Zebrafishes / Xenopus will be maintained in facilities of the Laboratory Animal Center of KU Leuven, which applies Standard Operation Procedures concerning housing, feeding, health monitoring to assure consistent care in accordance with European and national regulations and guidelines. All animals will be registered in the Leuven Animal Information System (LAIS) database, along with corresponding genotyping information, ethical approval documents and animal provider receipts. Drosophila lines will be stored in a dedicated room and managed using a specific database for storage of the corresponding information (including genotype, origin, number of vials and date of transfer, crossing schemes) and vial tracking via unique QR codes. Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate. - Algorithms, scripts and softwares: 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). - 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), the EBI European Genome-phenome Archive (EGA) for personally identifiable (epi)genome and transcriptome sequences

5.2. 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. - 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. - 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. Incremental backups are done daily from one 20 TB QNAP NAS to a second 20 TB QNAP NAS. - Microbiome data on VIB servers is secured by daily

snapshot backups, as well as weekly and monthly long-term backups on independent servers. The data storage servers have internal redundancy (RAID) to avoid data loss. All data will be also duplicated on KU Leuven servers for long term perservation. - For bacterial and yeast strains: A backup of selected strains will be stored in -80°C freezers in xxx in a more compressed form (96-well plates instead of cryotubes)

5.3. 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

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

The total estimated cost of data storage during the project is 5000€. 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” and 519€/TB/Year for the “J-drive”. -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.

Electricity costs for the -80° freezers present in the labs are included in general lab costs.Data storage and backup costs are included in general lab costs.

5.5. 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. 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. All private data will be rendered anonymous before processing outside the digital vault. Only the PI will be granted access to the server to deposit private data. The PI will be the only responsible for linking patient information, survey data and/or tissue samples, and will strictly respect confidentiality. All 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.

6. Data preservation after the end of the project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

The minimum preservation term of 10 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 10 years, conform the KU Leuven RDM policy. The costs (€156 per TB per year for “Large volume-storage”) will be covered by the PIs.

6.2. Where will these data be archived (= stored for the long 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

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The total estimated cost of data storage during 10 years after the end of the project is . 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” and 519€/TB/Year for the “J-drive”.

-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. Data storage and backup costs are included in general lab costs

7. Data sharing and re-use

7.1. 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 or because of IP potential)?

No

7.2. 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). We aim at communicating our results in top journals that require full disclosure upon publication of all included data, either in the main text, in supplementary material or in a data repository if requested by the journal and following deposit advice given by the journal. Depending on the journal, accessibility restrictions may apply. Biological material will be distributed to other parties if requested

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

- In a restricted access repository

After contacting the PI.

7.4. When will the data be made available?

- Upon publication of the research results

Once the data are published they will be available for the broad audience by sharing the link to the dataset.

7.5. 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, a Creative Commons Attribution (CC-BY) 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 nondisclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted

7.6. What are the expected costs for data sharing? How will these 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-touse data repositories and dissemination facilities whenever possible. Data management costs will be covered by the laboratory budget. A budget for publication costs has been requested in this project

8. Responsibilities

8.1. Who will be responsible for the 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.

8.2. 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 Andrei Segal Stanciu for the electronic laboratory notebook (ELN) and KU Leuven drives.

8.3. Who will be responsible for ensuring data preservation and sharing?

The PI is responsible for data preservation and sharing, with support from the research and technical staff involved in the project, fromAndrei Segal Stanciu for the electronic laboratory notebook (ELN) and KU Leuven drives.

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

The PI is ultimately responsible for all data management during and after data collection, including implementing and updating the DMP