MY PLAN (PDM DMP)

Identifying disease-critical cell types and pathways to guide new treatment approaches for Houge-Janssens syndrome type 2 (HJS2).

PDMt1/24/003

ADMIN DETAILS

Project Name: Identifying disease-critical cell types and pathways to guide new treatment approaches for Houge-Janssens syndrome type 2 (HJS2).
Principal Investigator / Researcher: Iris Verbinnen
Institution: KU Leuven

1. GENERAL INFORMATION

Name applicant
Iris Verbinnen

PDM Project Number & Title
PDMt1/24/003
Identifying disease-critical cell types and pathways to guide new treatment approaches for Houge-Janssens syndrome type 2 (HJS2).

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).

WP1: A multidisciplinary approach to characterize disease mechanisms in 3D and 2D human stem cell models of HJS2

Task1.1. To determine the cellular consequences of HJS2 in 3D and 2D stem cell models

Type of data	Format	Estimated volume	How created
PGP-1 cells (human induced pluripotent stem cells (hiPSC)): parental (wildtype), *PPP2R1A* M180T/+ heterozygous and *PPP2R1A* R182W/+ heterozygous cells	cells	20 tubes per genotype (60 in total)	Commercialy available cell line. Mutations were made in UC Davis (US).
qPCR to determine neuronal cell types	QuantStudio Design & Analysis Software single experiment document and .xls	50 MB	QuantStudio software, Excel
Immunostainings for several markers of neuronal differentiation, proper formation of cortical zones, apoptosis, proliferation, cell cycle exit	.jpeg, .xls	1 TB	Images made with microscope TissueFAXS. Analysis with Excel.
Single cell RNA-seq: 1) RawFastqFiles 2) PreProcessedFastqFiles 3) Mapping ((a) Bamfiles; (b) Reference-Genome; (c) Unmapped) 4) Summarization 5) Statistical analysis	1) .gz 2) .gz 3) (a) .bam; .bai; (b) .gtf; (c) .gz 4) .xls 5) .xls	1) 18 GB 2) 18 GB 3) (a) 12 GB; (b) 60 MB; (c) 1.2 GB 4) 25 MB 5) 70 MB	Data are provided by the Leuven Institute for Single Cell Omics (LISCO – KU Leuven).

Task1.2. To determine the molecular consequences of HJS2 in 3D and 2D stem cell models

Type of data	Format	Estimated volume	How created
PGP-1 cells (human induced pluripotent stem cells (hiPSC)): parental (wildtype), *PPP2R1A* M180T/+ heterozygous and *PPP2R1A* R182W/+ heterozygous cells	cells	20 tubes per genotype (60 in total) (same vials as described for Task 1.1.)	Commercialy available cell line. Mutations were made in UC Davis (US).
(Phospho-)proteomics: Tables with raw and processed data	.raw .xls	20 MB	Data are provided by the Proteomics Core Facility (KU Leuven)
Kinome profiling: 1) Statistical analysis 2) Summarization	.xls, .pdf	20 MB	Data are provided by PamGene International (The Netherlands)
Targeted immunoblot or immunohistochemical analyses to validate potential hits	.jpeg, .xls	50 GB	Images of immunoblots made with ImageQuant800 (Amersham). Images of stainings made with microscope TissueFAXS. Quantifications in Excel.

Task1.3. To determine the functional consequences of HJS2 in 3D and 2D stem cell models

Type of data	Format	Estimated volume	How created
PGP-1 cells (human induced pluripotent stem cells (hiPSC)): parental (wildtype), *PPP2R1A* M180T/+ heterozygous and *PPP2R1A* R182W/+ heterozygous cells	cells	20 tubes per genotype (60 in total) (same vials as described for Task 1.1. and Task 1.2.)	Commercialy available cell line. Mutations were made in UC Davis (US).
Recording at microcircuitry level (Maxwell electrodes)	HDF5 (*.h5) for raw data, .xls for processed data	50 GB	Data are analysed using Matlab and Phyton

Immunostainings for neuronal morphology and synapse formation	.jpeg, .xls	500 GB	Images made with microscope TissueFAXS. Analysis with Excel.

WP2: Characterization of cross-species conservation of PPP2R1A-regulated neurodevelopmental processes using HJS2 mouse models

Task2.1. To determine the early cellular consequences in HJS2 mouse models

Type of data	Format	Estimated volume	How created
Mouse models (M180T/+ and R182W/+ heterozygous mice)	sperm	50 straws (5 straws per male, 5 males in total)	Sperm cryopresevation is done by the Mouse Expertise Unit (VIB-KU Leuven)
Immunostainings for several markers of neuronal differentiation, proper formation of cortical zones, apoptosis, proliferation, cell cycle exit	.jpeg, .xls	1 TB	Images made with microscope TissueFAXS. Analysis with Excel.
Single cell RNA-seq: 1) RawFastqFiles 2) PreProcessedFastqFiles 3) Mapping ((a) Bamfiles; (b) Reference-Genome; (c) Unmapped) 4) Summarization 5) Statistical analysis	1) .gz 2) .gz 3) (a) .bam; .bai; (b) .gtf; (c) .gz 4) .xls 5) .xls	1) 18 GB 2) 18 GB 3) (a) 12 GB; (b) 60 MB; (c) 1.2 GB 4) 25 MB 5) 70 MB	Data are provided by the Leuven Institute for Single Cell Omics (LISCO – KU Leuven).

Task2.2. To determine the early molecular consequences in HJS2 mouse models

Type of data	Format	Estimated volume	How created
Mouse models (M180T/+ and R182W/+ heterozygous mice)	sperm	50 straws (5 straws per male, 5 males in total) (same as in Task2.1.)	Sperm cryopresevation is done by the Mouse Expertise Unit (VIB-KU Leuven)
(Phospho-)proteomics: Tables with raw and processed data	.raw .xls	20 MB	Data are provided by the Proteomics Core Facility (KU Leuven)
Kinome profiling: 1) Statistical analysis 2) Summarization	.xls, .pdf	20 MB	Data are provided by PamGene International (The Netherlands)
Targeted immunoblot or immunohistochemical analyses to validate potential hits	.jpeg, .xls	50 GB	Images of immunoblots made with ImageQuant800 (Amersham). Images of stainings made with microscope TissueFAXS. Quantifications in Excel.

Task2.3. To determine rescue strategies for the observed phenotypes of HJS2 in mice

Type of data	Format	Estimated volume	How created
Mouse models (M180T/+ and R182W/+ heterozygous mice)	sperm	50 straws (5 straws per male, 5 males in total) (same as in Task2.1. en in Task2.2.)	Sperm cryopresevation is done by the Mouse Expertise Unit (VIB-KU Leuven)
(Phospho-)proteomics: Tables with raw and processed data	.raw .xls	20 MB	Data are provided by the Proteomics Core Facility (KU Leuven)
Kinome profiling: 1) Statistical analysis 2) Summarization	.xls, .pdf	20 MB	Data are provided by PamGene International (The Netherlands)
Targeted immunoblot or immunohistochemical analyses to validate	.jpeg, .xls	50 GB	Images of immunoblots made with ImageQuant800 (Amersham). Images of

consequences of therapy			stainings made with microscope TissueFAXS. Quantifications in Excel.
Table with behavioral tests/milestones for each day to evaluate during the preweaning period	.xls	7 MB	Excel and table completed each day during testing (answers: yes/no; +/-; sec; degrees)
24h cage activity: Table of beam crossings per mouse for every 30 minutes (24h)	.xls	4 MB	Table generated by program: MouseWin
Open field: 1) Table with different parameters 2) ANY-maze backup files 3) ANY-maze data files 4) Track plots (movies and pictures)	1) .xls 2) .szk 3) .szd 4) movies: .mp4 and pictures: .png	1) 700 KB 2) 150 MB 3) 500 MB 4) movies: 200 MB and pictures: 900 KB	All data are generated by ANY-maze
Passive avoidance: Table with latencies	.xls	1.6 MB	Table is made in Excel

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: N.A.

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

ECD P129/2020: Karakterisering van muismodellen voor PP2A-gerelateerde syndromen

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?

- No

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?

I will provide detailed descriptions of data acquisition in paper lab notebooks, according to good laboratory practices. All data belonging to a certain work package (and task) of the project are properly linked to each other and referred to each other in the lab notebooks and detailed protocols are provided (.doc files + printed and stored in a folder). The complete list of sub-projects is summarized as a final list (.doc file); the detailed protocols are also digitally archived on the lab's dropbox, on an external hard drive and J-drive. In the notes (lab notebook), clear links are provided to (1) digital data (pictures, excel files, movies, other raw data), and where to find these (Dropbox, J-drive, external hard drive etc); file names, types and dates are provided; and to (2) physical data (cell lines, mice, tissue samples, antibodies) and where/how they are stored. General protocols and information are also stored on the shared J-drive and printed instructions will be stored in a folder. Lists of antibodies and primers for PCR and qPCR are stored on the J-drive of the lab.

The following documentation is available:

WP1: A multidisciplinary approach to characterize disease mechanisms in 3D and 2D human stem cell models of HJS2

- An overview of the available vials of the PGP-1 cell lines with different genotypes (WT, M180T, R182W) is kept in the Biobank file, and in an excel file of the Dropbox of the lab.
- Protocols (.docx) related to immunostainings and qPCR are stored on specific drives and are printed and stored in a folder. For each protocol, the necessary antibodies (stainings) and primers (qPCR) are listed and can be found in the shared .xls files on the J-drive. Storage of the organoid samples in -80°C freezer and overview of the boxes with gelatin/sucrose blocks are listed in an .xls file.
- An overview of the organoids (time point, batch, date) that are used for the scRNA-seq, proteomic and kinome analyses is available on the shared J-drive. Protocols for isolation of mRNA and proteins from organoids are saved on several drives, and are printed and stored in a folder. Storage of the organoid samples in -80°C freezer and overview of the boxes with gelatin/sucrose blocks are listed in an .xls file.
- Protocols (.docx) related to electrophysiology measurements are stored on the J-drive.

WP2: Characterization of cross-species conservation of PPP2R1A-regulated neurodevelopmental processes using HJS2 mouse models

- Protocols of the breeding (.docx) and genotyping (.docx) of mice are stored and printed instructions are found in a folder. Genotypes of the mice are collected in an .xls file. Administration of the mice (genotypes, room and cage of housing) are stored on the Tick@Lab website of the KU Leuven animal facility.
- Protocols (.docx) related to immunostainings and qPCR are stored on specific drives and are printed and stored in a folder. For each protocol, the necessary antibodies (stainings) and primers (qPCR) are listed and can be found in the shared .xls files on the J-drive. Storage of the brain samples in -80°C freezer and overview of the boxes with paraffin blocks are listed in an .xls file.
- An overview of the mouse embryo's that are used for the scRNA-seq, proteomic and kinome analyses is available on the shared J-drive. Protocols for isolation of mRNA and proteins from brains are saved on several drives, and are printed and stored in a folder. Storage of the brain samples in -80°C freezer and overview of the boxes with paraffin blocks are listed in an .xls file.
- Protocols of neonatal behavioral tests (.docx) and standardized tables (.xls) to complete for each pup are made and stored. Printed versions of protocols are also available. Completed and analyzed tables (written) are pasted in a Lab Notebook and stored on an external hard drive, dropbox and the K-drive.
- Standard operating procedures (SOPs) of adult behavioral tests (.docx), standardized tables for passive avoidance measurements (.xls) and tables generated by MouseWin and ANY-maze software (.xls) are kept. Completed and analyzed tables (.xls) are stored on specific drives (section 5). Printed protocols are also stored in a folder.

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.

- No

No real metadata standard will be used.
Excel files stored within each experiment folder will contain the following information:

1) Experimental design
> Definition of experimental and control groups
> Number within each group (n)
> Researcher
> Date
2) Samples
> Organoid/2D differentiation batch, start date, end date > Mouse IDs, cage number, birth date, date of sacrificing
> Genotypes
> Type of tissue
> Processing procedure (frozen, fixed)
> Storage location (freezer box, parafin box)
3) Protocol
> Reference to protocol (printed version, drives)
> Adjustments
> Remarks
4) Data analysis
> Computational programs
> Parameters
5) Data storage
> Notebook pages
> Drives (+ sub-folder)

On the K-drive, the data are structured/ordered per paper (separate 'archive': see 5.1).
Furthermore, the data of my project are structured/ordered per sub-project (work package and task) on the K-drive.

5. DATA STORAGE AND BACKUP DURING THE PDM PROJECT

Where will the data be stored?

I will provide detailed descriptions of data acquisition in paper lab notebooks. All data belonging to a certain work package (and task) of the project are properly linked to each other and referred to each other in the

lab notebooks and detailed protocols are provided (.doc files + printed and stored in a folder). General protocols and information are stored on the shared J-drive, dropbox, external hard drive and printed instructions will be stored in a folder. Lists of antibodies and primers for PCR and qPCR are stored on the J-drive of the lab. Data are stored in dropbox, external hard drive, and the Archive drive (K-drive).

For each work package, the following documentation is stored:

WP1: A multidisciplinary approach to characterize disease mechanisms in 3D and 2D human stem cell models of HJS2

- An overview of the available vials of the PGP-1 cell lines with different genotypes (WT, M180T, R182W) is kept in the Biobank file, and in an excel file of the Dropbox of the lab.
- Protocols (.docx) related to immunostainings and qPCR are stored on specific drives and are printed and stored in a folder. For each protocol, the necessary antibodies (stainings) and primers (qPCR) are listed and can be found in the shared .xls files on the J-drive. Storage of the organoid samples in -80°C freezer and overview of the boxes with gelatin/sucrose blocks are listed in an .xls file on an external hard drive, dropbox and K-drive. Printed protocols are also stored in a folder.
- An overview of the organoids (time point, batch, date) that are used for the scRNA-seq, proteomic and kinome analyses is available on the shared J-drive. Protocols for isolation of mRNA and proteins from brains are saved on several drives, and are printed and stored in a folder. Storage of the organoid samples in -80°C freezer and overview of the boxes with gelatin/sucrose blocks are listed in an .xls file on an external hard drive, dropbox and K-drive. At the end of the experiment, raw data (see section 2) and processed data (.xls) will be saved on the K-drive.
- Protocols (.docx) related to electrophysiology measurements are stored on the J-drive and final protocols on the K-drive. Printed protocols are also stored in a folder.

WP2: Characterization of cross-species conservation of PPP2R1A-regulated neurodevelopmental processes using HJS2 mouse models

- Protocols of the breeding (.docx) and genotyping (.docx) of mice are stored and printed instructions are found in a folder. Genotypes of the mice are collected in an .xls file on an external hard drive, J-drive and K-drive. Administration of the mice (genotypes, room and cage of housing) are stored on the Tick@Lab website of the KU Leuven animal facility.
- Protocols (.docx) related to immunostainings and qPCR are stored on specific drives and are printed and stored in a folder. For each protocol, the necessary antibodies (stainings) and primers (qPCR) are listed and can be found in the shared .xls files on the J-drive. Storage of the brain samples in -80°C freezer and overview of the boxes with paraffin blocks are listed in an .xls file on an external hard drive, dropbox and K-drive.
- An overview of the mouse embryo's that are used for the scRNA-seq, proteomic and kinome analyses is available on the shared J-drive. Protocols for isolation of mRNA and proteins from brains are saved on several drives, and are printed and stored in a folder. Storage of the brain samples in -80°C freezer and overview of the boxes with paraffin blocks are listed in an .xls file on an external hard drive, dropbox and K-drive.

- Protocols of neonatal behavioral tests (.docx) and standardized tables (.xls) to complete for each pup are made and stored. Printed versions of protocols are also available. Completed and analyzed tables (written) are pasted in a Lab Notebook and stored on an external hard drive, dropbox and the K-drive.
- Standard operating procedures (SOPs) of adult behavioral tests (.docx), standardized tables for passive avoidance measurements (.xls) and tables generated by MouseWin and ANY-maze software (.xls) are kept. Completed and analyzed tables (.xls) are stored on an external hard drive, dropbox and the K-drive. Raw data are stored on an external hard drive. Printed protocols are also stored in a folder.

How is backup of the data provided?

Besides regularly provided automated backups by ICTS (of J-drive, K-drive), the data stored on personal PCs and J-drive will be weekly backed up on the personal KU Leuven I-drive, and the own Dropbox of the labs concerned (password-protected). Every two months, additional backups will be made on external hard disks, as an extra security measurement. MS data, RNA-seq data and imaging data will be backed up by copying to an SSD. External SSD hard-disks (up to 5 TB storage capacity) keep the storage costs feasible since cost of an external hard drive is cheaper than ICTS storage costs for such big data volumes.

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

Extra external hard disks will be bought, as appropriate. Same for liquid nitrogen containers or freezers, but sufficient capacity is currently available.
K-drive: up to 100 GB per PI; in addition, extensions of this volume can be asked for at ICTS at all times (for an additional cost).

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

- K-Archive drive costs: 20€/year/100 GB; daily back-ups.
- Dropbox: free
- External hard disks: max. 750€ (=5 disks of 1-2 TB)

Costs are covered by my PDM/FWO bench fee.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The main data archive on the K-drive is a read-only drive for lab members only (password-protected); only the PI of the lab concerned can write on this drive (= add data), but not modify data.
J-drive and OneDrive: restricted to account of researcher.
Lab Dropbox: access password-protected.
Paper lab notebooks are kept in locked cabinets.

6. DATA PRESERVATION AFTER THE PDM 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, ...).

All generated (raw and processed) data will be preserved, for at least 5 years.
Data underpinning publications: see above (section 5, K-drive).
Unpublished data from unfinished work will be kept for longer than 5 years (ideally, until use in a publication, and then preserved or deposited as per our criteria for published data).

Where will the data be archived (= stored for the longer term)?

Raw and processed data: K-drive and external SSD hard disks.
In addition, after publication, all Next Generation Sequencing Data will be uploaded to the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/); and all proteomics data will be deposited to the ProteomeXchange Consortium via the PRIDE partner repository.
The different genotypes of the PGP-1 cell lines are stored in the cryotheek at KU Leuven (KU Leuven Biobank).
Organoid samples embedded in 7.5%gelatin/10%sucrose solution are stored in well-defined boxes in -80 freezers.
Frozen sperm of mouse strains (50 straws (5 straws per male, 5 males in total)) are stored in 2 separated LN2 tanks on 2 different sites on the KU Leuven Gasthuisberg campus (done by Mouse Expertise Unit, VIB-KU Leuven).
Tissue samples of mice are stored in -80 freezers and paraffin blocks in well-defined boxes.

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

K-Archive drive costs: 20€/year/100 GB
External hard disks: max. 1500€ (=10 disks of 5 TB)
Costs are covered by my PDM or FWO bench fee.

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:

Data sharing with peers will exclusively occur through publications, taking into account the Open Access policy of KU Leuven. The planned depositions of data in the relevant responsible repositories (as described above) will only occur **after** publication.

Which data will be made available after the end of the project?

Publications; all data deposited in (public) repositories.

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

- In an Open Access repository

Publications; all data deposited in (public) repositories.

When will the data be made available?

Upon publication during the project, or as soon as possible upon publication after the project.

Who will be able to access the data and under what conditions?

Publications and repositories mentioned: all open access.
For unpublished data: only the PI involved (or scientific collaborators who will continue and follow up on the research after the project).

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

Publication costs (Open Access) will be covered by bench fees.

8. RESPONSIBILITIES

Who will be responsible for data documentation & metadata?

Iris Verbinnen

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

Prof. Veerle Janssens
Iris Verbinnen

Who will be responsible for ensuring data preservation and reuse ?

Prof. Veerle Janssens

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

The PI bears the end responsibility of updating & implementing this DMP.