DMP title

Project Name DMP_FWO_11K9122N

Project Identifier 11K9122N

Principal Investigator / Researcher Ruth Lannoo

Description This doctoral research aims at identifying novel genes involved in the Alternative Lengthening of Telomeres (ALT) pathway to deliver potential targets for the design of an ALT-targeting tumor-agnostic therapy that will help patients suffering from ALT positive tumors (e.g.: Soft-tissue sarcoma). For this purpose, at least three genome-wide CRISPR knock-out screens will be performed in SW26 cells (relying on the ALT pathway) and SW39 cells (relying on telomerase).

Institution KU Leuven

1. Data Description

What data will you collect or create? Fill out the table below and/or describe.

	Type of data	Format	Volume	How created?
1	Primary experimental data	numerical and textual (txt, csv, docx, pdf, xlsx)	<1 TB	Lab experiments: MTS assay (absorbance data), qPCR (>Quantstudio, Ct values), ELISA (absorbance data), Nanodrop (concentrations), sanger sequencing.
2	Images	jpeg, png	<5 MB	Lab experiments: Western Blot (Protein simple - WES), Gels (FAS-V). Graphs created of experiments (GraphPad Prism). Images created for presentations and projects (Powerpoint, Biorender)
3	Genomic data	Fasta/Fastq files (.fasta/.fastq)	1 MB	NGS (Illumina sequencing) of genomic DNA extracted from SW26 and SW39 cell lines
4	Cell lines (SW26, SW39)	physical data	100*10^6 cells per cell line	Gifted by Prof. Oliver Bechter (UZ Leuven/KU Leuven) - Stored in liquid nitrogen (A3.A340 Rega Institute)
5	CRISPRCloud2 Script	R-studio file	10 KB	https://crispr.nrihub.org/
6	Lab notes	written (non-digital)	2 books/year	Tracking daily experiments

Do you intend to reuse existing data?

Not in the near future.

When we have identified potential drug targets, we might use anonymized patient-derived gene expression data from soft-tissue sarcoma and glioblastoma patients to analyze the expression of identified hit genes for the establishment of novel biomarkers. These data can be obtained, for example, from the GDC portal (TCGA-SARC database) and the Laboratory of Precision Cancer Medicine (Prof. dr. De Smet - KU Leuven).

This item will be updated as soon as we expect to reuse existing data.

Do you use personal data (i.e. all data possibly identifying an individual)?

- No

2. Documentation and Metadata

Describe the documentation that will be created for the data. This section deals with the way in which you will document how the dataset was created and subsequently processed.

1. Experimental data: Protocol and methodology of each individual experiment will be described in a detailed non-digital lab book (chronological). In addition, per larger experiment, a digital folder will be made containing excel, word, and/or PDF files with background info on the experiment, the protocol(s), raw data, (statistical) analysis of the raw data, and an optional graph to visualize the results (made in GraphPad or excel). All the files and folders will have a date and title in their name to easily find back experimental results.

2. Images: Images will be stored in the appropriate digital folder of the experiment during which the images were created (see point 1.). Images for presentations will always be stored in the same folder as the other presentation files (eg. PowerPoint file).

3. Genomic data: Genomic data (fasta/fastq files) will be stored in a (zipped) folder per experiment and will be analyzed with the CRISPRCloud2 script in R-studio. The NGS processed data will be stored in a large excel file (including a 'ReadMe' file) in the same folder. GraphPad will be used to visually represent the data. These GraphPad files and graphs (PDF or jpeg) will also be stored in the same folder.

4. Cell lines: The cells will be stored in liquid nitrogen at the Rega Institute (100*10^6 cells per cell line - room 03.A340). An excel file will be updated regularly to keep track of the number of vials per cell line.

5. CRISPRCloud2 script: This script will be kept as an R-file. The text function of R-studio will be used to provide information about the different steps in the script.

6. Lab notes: My lab books will be kept in a locked cabinet in my office.

Describe the metadata for the data. This section deals with metadata: information contained in your dataset about the research data.

The metadata generated by instruments (Simple WES, TECAN Saffire II, Quantstudio II, etc) will be present in the raw data files, which will be stored in a folder per experiment. These metadata consist of the date/time of read-out and conditions/settings of measurement (specific wavelengths, duration of measurement, temperatures, exposure time, etc).

3. Ethical, Legal and Privacy Issues

Are there any ethical issues concerning the creation and/or use of the data?

PRET questionnaire filled in and accepted (G-2021-4345): There are no ethical issues concerning the creation and/or use of the data for my project.

Did you consider all issues about copyrights and IPR?

Not applicable. This topic might be updated in the future.

Are the collected data considered to be â€œdata containing personal informationâ€ and are all the requirements about the collection of these data met?

PRET questionnaire filled in and accepted (G-2021-4345): I will not collect "data containing personal information".

4. Data storage and Backup during Research

How and where will the data be stored during research?

- Centrally on storage facilities of the research unit
- In a cloud service offered by the university

The generated data will be stored on a shared lab drive, provided by KU Leuven. In order to guarantee data protection, this drive is only accessible to members of our research unit

(password protection). This shared drive is automatically backed-up by KU Leuven. The data will also be synchronized with my OneDrive profile licensed by KU Leuven, which is only accessible to me (password protection).

Which back-up procedures are in place?

The data will be stored on the research unit servers with automatic daily back-up procedures. OneDrive by KU Leuven is also automatically backed-up.

Describe the data security procedures and who has access to the data.

The data on the KU Leuven lab drives is only accessible for members of our research unit (password protection). OneDrive by KU Leuven is only accessible by me (password protection).

5. Data selection and Preservation after Research

What is the long-term preservation plan for these dataset(s)?

During and after this research project I will be responsible for the correct storage and preservation of the generated data. In addition, the head of our research group, Prof. dr. Daelemans bears the overall responsibility for long-term data storage. The data will be stored on the research unit shared drives (with automatic backup procedures) for at least 10 years, conform the KU Leuven RDM policy. My written lab notes will also be stored in a locked cabinet for at least 10 years after the ending of the project at the Rega Institute.

Data Selection: Which data will have long time value for the research and will be preserved?

All the datasets described in section 1 will be preserved for at least 10 years, according to the KU Leuven RDM policy.

6. Data Sharing

Are there any restrictions for sharing the data?

There are no restrictions concerning data sharing. Data can be shared when there is a data-sharing agreement with me or Prof. Dirk Daelemans.

If there are no restrictions, which mechanisms will be in place to assure that the data are discoverable, accessible and intelligible?

Discoverable: All data are identifiable and searchable on the KU Leuven shared servers of our research unit (to the ones who have access to these servers). All the experiments and data are stored in folders with clear titles and dates to easily relocate them.

Accessible: All the data are accessible when there is given permission by me or Prof. Dirk Daelemans to access the lab servers.

Intelligible: All the data are intelligible due to the Read-me files and experiment information files in the folders per experiment. Moreover, the written lab notes also provide background information to understand the performed experiments and generated data.

How will you share the data?

- Publication

The generated data will be published in multiple manuscripts (articles). The datasets described in the articles (eg: dataset of genes involved in a specific pathway) will be accessible through the manuscripts. Specific data and experimental procedures/protocols can be requested via email to Prof. Daelemans, following the signing of a data-sharing agreement.

With whom will the data be shared?

- On request

Articles will be published in journals with Open Acces policy. Specific data and experimental procedures/protocols can be requested via email to Prof. Daelemans, following the signing of a data-sharing agreement.

7. Responsabilities and Resources

Who is responsible for Data Management during the project? This will be the person

who might receive questions on the data management aspects of the research project.

During and after this research project I will be responsible for the correct storage and preservation of the generated data. In addition, the head of our research group, Prof. Dirk Daelemans, bears the overall responsibility for long-term data storage and implementing this DMP.

Which additional resources are needed for the execution of the Data Management Plan?

Not applicable

Did you read the KU Leuven Data Management Policy? (find the link to the policy in the guidance).

- Yes