DMP title

Project Name DMP FWO-SB Mattijs Bulcaen - DMP title
Project Identifier 1SE8122N
Principal Investigator / Researcher Zeger Debyser
Project Data Contact Mattijs Bulcaen
Description DMP for FWO-funded Strategic Basic PhD research project entitled "A prime editing and translational delivery strategy for drug-refractory CFTR mutations causing cystic fibrosis".
Institution KU Leuven

1. General Information

Name applicant

Mattijs Bulcaen

FWO Project Number & Title

Application number: 1SE8122N
Project title: A prime editing and translational delivery strategy for drug-refractory *CFTR* mutations causing cystic fibrosis

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	100-300 GB	Photos taken from (transduced) cell cultures. Fluorescence photos stained samples or of cells expressing YFP/GFP. Confocal images for single particle tracking. Histological staining. Forskolin-induced swelling intestinal organoids.
Gel electrophoresis images	.tif, .jpg	5 GB	Imaging of agarose gel electrophoresis, PAGE Western blot
Numeric data	.xls, .txt, .fcs, .csv	1 GB	Data on BCA analysis, editing outcomes, flow cytometry measurements, p24 ELISA, transduction efficiencies, cell counts.
Sanger sequencing chromatograms	.abi	10 GB	Outcome gene editing experiments, control sequencing of new plasmid constructs
Next generation sequencing	.fastq, .zip, .fastq.gz	40 GB	GUIDE-Seq experiments, targeted amplicon deep sequencing on/off-target sites
Ussing chamber measurements	.cla	5 GB	Short-circuit recordings of transepithelial voltage, resistance and currents.
Fluorometric readout Perkin Elmer-Envision	.csv	2 GB	HS-YFP quenching assay and fluorometric readout of mucin-cyanoacetamide reaction.
Protocols, lab book, metadate description	.docx, .ppt, .txt	5 GB	All experiments

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

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)

- No

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

IP protection will be prioritised before disclosure. For this matter, I can rely on the in-house expertise of IOF manager dr. F. Christ and on the IP-cell of the LRD office at KU Leuven.

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

No third-party agreements restrict dissemination.

4. Documentation and metadata

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

For all experiments, all metadata is saved in both hardcopy notebooks as well as digital files in such a way that a non-expert researcher will be able to reconstruct the experiment and analyse the data. All raw data, together with detailed annotations and techniques are stored both as printed hard copy, as well as online via OneNote. These OneNote digital lab books contain introduction, rationale, materials&methods, execution, raw data, analysis method, results, conclusion section and rationale for future experiments. Reported metadata includes sample coding eg. "exp117-c29" and concomittant storage of experiment sheets that in detail describe how every sample was treated. For every type of experiment, standard protocols are saved and referred to in experiment descriptions. Deviations from these standard protocol versions are always mentioned. Sequencing files are always annotated according to this sample coding and include the primer that was used during sequencing. For NGS, a sample description .docx file is kept along .docx and .txt files that describe the used analysis commands. For readouts in multiwell-plates .txt files describing plate layout are saved next to the raw data files. For gel electrophoresis/Western blot/microscopy images/... .ppt files that show the location of each condition or nature of each staining on the gel/blot/image are saved in the same folder as the raw data files. For cloning/PCR procedures all details (incl. ordered DNA sequences) documented.

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

Currently, the use of metadata standards has not been implemented in the daily routine of the research group. In case this will change in the course of the project, this change and its timing will be reported at the end of the project. For all data common metadata are collected: title, author, data type, date created and data modified, file size, equipment reference (such as manufacturer and model identification). Depending on th enature of data additional metadata are collected. For instance, for microscopy; channels used, fluorescence: wavelength.

5. Data storage and backup during the FWO project

Where will the data be stored?

Time-stamped master copies of all data will be stored centrally. All lab work will be reported both in an electronic lab book (OneNote) as well as in a physical lab book with chronological reporting of all experiments and results including cross references to electronic storage of the data. Both OneNote as well as most data and metadata are stored on a 2TB harddrive and synchronised with a KUL OneDrive account. For large data the laboratory has reserved 5 TB of server storage capacity which can be extended on demand.

How is backup of the data provided?

Daily backup is secured for all data stored on central university servers. All other data stored on the main desktop is synchronised via OneDrive and weekly transferred to an extra physical backup drive of 2TB.

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

Both the OneDrive account, as well as the storage capacity of the desktop computer have a limit of 2TB. Large data sets, as well as periodic backups of the OneDrive-stored data are stored on university servers (5TB available).

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

OneDrive is free of use for KUL members. For large data storage (L-drive), the devision of

Molecular Medicine pays €1138.40 per year. These costs are financed through grant applications.

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

Access to any of the locations used for data storage (KUL servers, personal desktop, extra harddrive, physical lab books and lab archive) require explicit authorisation and are protected from outside access.

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, ...).

All data obtained during this FWO project will be retained for the expected 10 year period, thereby complying to the data preservation rules of KU Leuven.

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

Long term data preservation will be achieved via the central university servers and the laboratory physical archive. Digital versions of the lab books will be kept on OneDrive.

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

These costs directly depend on the pricing of large storage drives of KU Leuven. These costs will be covered through grant applications.

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

In respect to the data produced/collected no restriction for data sharing apply.

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

Relevant digital data will be published and made available after the end of the project. Data with valuable IP will be protected prior to publication. We will comply with open access regulations of FWO.

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

- In a restricted access repository
- Upon request by mail

Data will be available after signing a data sharing agreement which will be established with the support of KUL R&D after a request. Once KU Leuven has established a university managed and owned data repository sharing of data (or a subset of data), this repository will be evaluated for use based on policy and repository conditions. Changes to the data sharing policy will be reported in the final DMP of this project.

When will the data be made available?

- Upon publication of the research results

Data will be made available after publication in peer reviewed journals. Additional data will be made available on basis of data sharing agreements if requested by third party. Additional material (and associated data) will be made available on basis of material transfer agreements (MTAs).

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

All members of the laboratory of Molecular Virology and Gene Therapy are authorized to have access to all obtained digital and physical data after the project. Due to potential commercial value of the data, no general full open access will be provided by default before IP protection and publication. Data which will be shared with third parties will exclude commercial use and will require appropriate credit to the data owners.

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

The KU Leuven repository will not require any cost contribution for KU Leuven researchers. Data shared through journal repositories will be covered by publication costs. Bilateral agreements for data sharing will be established through the services of KU Leuven R&D. The costs expected for data sharing are thus low.

8. Responsibilities

Who will be responsible for data documentation & metadata?

The individual researcher producing the data - ie. Mattijs Bulcaen - will have the final responsibility for data documentation and metadata. Supervision is provided by prof. Zeger Debyser and prof. Marianne Carlon.

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

Prof. Zeger Debyser and prof. Marianne Carlon are responsible for data preservation and reuse.

Who will be responsible for ensuring data preservation and reuse ?

Prof. Zeger Debyser and prof. Marianne Carlon are responsible for data preservation and reuse.

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

Prof. Zeger Debyser and prof. Marianne Carlon bear the end responsibility for updating and implementing this Data Management Plan.