DMP_FWO_11L0822N

Project Name DMP_FWO_11L0822N - DMP_FWO_11L0822N

Project Identifier u0141151, r0589374

Grant Title 11L0822N

Principal Investigator / Researcher Maxime Vanmechelen

Project Data Contact maxime.vanmechelen@kuleuven.be, maxime.vanmechelen@uzleuven.be, vanmechelen.maxime@gmail.com, +32474/59.87.93, +3216/34.44.39

Description Despite intensive research investments, glioblastoma (GBM) remains an incurable condition with a dismal prognosis. While trials in the last decades have mainly focused on targeting tumor cells, recent years showed a shift towards immunomodulatory approaches, including checkpoint inhibitors. Unfortunately, immunotherapies have so far not met expectations which is partly caused by a brain-specific, immunosuppressive microenvironment, different from other peripheral cancer types. In this light, microglia, the brain-resident macrophages that constitute a highly prevalent cell type in GBM, can significantly modulate tumor progression. However, compared to T- cells, the molecular mechanisms of microglial-tumor cell interactions remain largely understudied and underappreciated as potential targets for GBM therapy. In this project, I will therefore study the interactions between tumor and microglial cells using state-of-the-art patient-derived in vitro (GBM-microglia co-cultures) and in vivo models (humanized xenografts), in addition to extended patient- sample profiling using advanced single cell analyses. Using these models, I will also engage in pathway mapping to unravel the underlying biological mechanisms, evaluate the therapeutic potential of new targets, and position these insights across the spectrum of various GBM tumor subtypes.

Institution KU Leuven

1. General Information

Name applicant

Maxime Vanmechelen

FWO Project Number & Title

- Project number: 11L0822N
- Project title: Mapping the Microglial-Tumor cell interplay in Glioblastoma using advanced in vitro and in vivo single-cell profiling

Affiliation

- KU Leuven

KU Leuven – UZ Leuven

2. Data description

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

- Generate new data
- Reuse existing 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).

I will mostly collect and generate new data, but also make use of existing material and data. A part of the patient information (clinical characteristics) and already stored material and cell lines in our biobank will be used. All the other data will be generated during the project.

WP	Type of data	Format	Volume	How created

WP1	Multiplex IHC (MILAN), spatial proteomics	.czi, .jpg, .xls	>15TB	Immunofluorescence (ZeissAxioscan) with MILAN technique, excel files with the calculations for the antibody cocktails, screenshots taken during quality control
WP1	Generation of tissue microarrays (TMAs)	Tissue blocks, .xls structured overview of included samples	+- 10 TMAs	TMA grand master, ROIs identified by an expert neuropathologist
WP1	Data processing multiplex IHC	.tiff, .ome tiff, .Rmd (R-scripts & R-studio), software tool, GUI, .html files	>25TB	Newly developed spatial single cell data analysis pipeline (MILAN pipeline), large set of intermediary and final data files
WP1	Sequencing	.vcf, .maf files (VEP and ONCOKB annotated)	25GB	Whole exome sequencing for characterization of paired GBM patient samples
WP1	MGMT promotor methylation	.xls	1GB	Quantitative methylation-specific PCR (qMSP)
WP1	Clinical database of included patients UZ Leuven	.xls, text fields in MySQL database	500MB	Both retrospective and prospective data, where the data is manually extracted from patient files (KWS). Decoded personal data, clinical parameters, pathology information and genetic information (if available) and longitudinal follow-up. KWS data is stored for at least 25 years, all patient data has been anonymised (TSSxx number, GBMxx number) = partially existing data, partially gathered by me

WP1	Clinical database of included patients from external hospitals	.xls	500MB	Comparable clinical characteristics and pathology/genetic information from GBM patients in Maastricht (MUMc) and Genk (ZOL) = existing data
WP2	Collection of biopsies and generation of patient-derived cell lines (PDCL)	.xls, text fields in MySQL database, lab notebook, electronic lab notebook (ELN) in labcollector	10GB	Collection, processing and storage of biopsies. Cell lines derived from fresh tumor biopsies will be generated to perform the described experiments. Cell lines stored in liquid nitrogen. Most of the cell lines used in the co-culture experiments where already generated beforehand in the available biobank = existing data (Leuven Living Tissue Bank)
WP2	Generation of iPSC-derived microglia	.xls, .txt, lab notebook, electronic lab notebook (ELN)	2GB	Protocols available from literature, cells freshly generated for co-culturing experiments
WP2	Genetics	.fastq files for raw data, .vcf files for analysed data	10GB	Genetic information of each cell line (chromosomal aberrations, mutations, transcriptional information)
WP2	Microscopy	.tiff, .czi	1TB	Light microscopy, confocal microscopy and fluorescent microscopy for monitoring of tumor and microglial interactions
WP2	Sequencing + barcoding (Multi-seq)	.abi, .count, .fastq, .xls, .pdf	100GB	Single cell RNA-seq for genotyping of patient biopsies + micoglia-GBM and macrophage-GBM co-cultures, gene expression profiling

WP2	Functional information of the generated co-culture models	.txt, numbers and graphs in MySQL data base, Prism files and R-codes for data analysis	15GB	General physiological parameters, drug responses using viability assays and molecular analyses, migratory patterns, chemotaxis,
WP2	qPCR	.csv	1GB	Validation of scRNAseq expression data, study changes of expression of genes after co-culturing
WP2	FACS	.fcs, Flowjo v.10 software	5GB	Sorting of cells on basis of cell surface marker expression, validation of cellular phenotypes
WP2	ELISA	.xls	250MB	Expression of inflammatory mediators in co-cultures to quantitatively measure enzyme activity
WP2	Isoplexis - Isospark	Isospeak software	10 GB	Multiplex measurement of cytokine secretion (up to 20 cytokines) after stimulation, single cell secretome analysis, functional immune landscaping, single-cell polyfunctionality,
WP2	Spatial transcriptomics	DCC files, .abi, .count, .fastq, .xls	50GB	CosMx Nanostring spatial transcriptomics for analysing the transcriptome next to spatial proteomics on consecutive slides
WP3	Crispr/cas9 editing	.abi, .fastq	10GB	Gene-edited cell lines using the crispr/cas9 technology (+- 5 targets for a strategic selection of cell lines) + genotyping of Crispr

WP3	Pathway mapping and biomarker/therapeutic target elucidation, GSEA, GO enrichment analysis	.abi, .count, .fastq, .xls	50GB	Looking deeper into pathways from literature (CSF1R, KIT, MAPK, TGFB1, CD47) and upregulated / downregulated pathways from WP2, selective pathway inhibitors
WP3	Incucyte (real-time imaging)	IncuCyte.exe files	5GB	Cell imaging and analysis platform that enables quantification of cell behavior over time
WP3	Microscopy	.tiff	25GB	Morphometric analysis of microglia and glioblastoma cells, migration analysis, expression of apoptotic, inflammatory, immunosuppressive markers
WP4	PDX mouse models (MIGRATE)	FFPE tissue blocks, frozen material	5GB	Orthotopic microglia and GBM injections for the generation of partly humanized PDX mouse models, immunodeficient mice (+- 5 mice for proof of concept, if successful validation in larger cohort (+- 15 mice)), tissues stored in liquid nitrogen
WP4	Follow-up of tumor growth	DICOM	5GB	Bioluminescence and MRI imaging
WP4	Single-cell analysis on PDX models: MILAN, single-cell RNAseq	.czi, .jpg, .xls, .abi, .count, .fastq, .pdf	5TB	Further elucidate these interactions, which will also be compared to the in vitro findings

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.

- Yes

Privacy Registry Reference:

- G-2021-3789 (S59804)
- G-2021-3895 (S61081)

Short description of the kind of personal data that will be used:

Clinical patient follow-up is collected and includes coded patient-related information, including: age, gender, prior history if cancer related, date of surgery, location of the tumor, information on Gliolan® positivity/negativity, clinical imaging parameters (e.g., from MRI), pathological diagnostics, overall survival, disease-free survival, date of tumor progression/relapse, treatment regimen and response, date of decease. All patient information is anonymized (TSSxx, GBMxx). As part of prior research, extensive genetic profiling has also been done already. Importantly, all obtained research data points and clinical information will be added in a coded manner to the LabCollector/Glioma2015 database of the Translational Cell and Tissue Research group, located on a KU Leuven hosted and secured server and is password protected. A dedicated, trained person will add all genetic and research information from this project to a local, KULeuven based cBioportal database. Only coded information will be extracted and used for the downstream research analyses. Personal data will only be kept as long as necessary for the research activities, but is available in the clinical files (KWS) for at least 25 years (legal obligation).

Biobank

Our laboratory has established a unique library of well characterized patient-derived cell lines (PDCLs) derived from fresh glioblastoma tumor samples (collaboration with Prof. S. De Vleeschouwer, UZLeuven). To date, a collection of +-40 newly diagnosed and 15 recurrent PDCLs have been generated and extensively characterized genomically and transcriptionally. These cell lines will be used for co-culture experiments as well as mouse experiments during the project.

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

Ethical approval for work packages 1-3 is already obtained from the Medical Ethical committee from UZLeuven/KULeuven, including informed consent for every patient. No ethical issues concerning research data are to be expected. The goal of the project is to collect data beyond 5 years and data collection, data sharing and long-term preservation has been defined in the informed consent.

Ethical approval for the project

- S59804
- S61081
- S62248

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

There is potential tech transfer/valorization in both the novel models we develop (genetic, proteomic and heterogeneity profiles with clinical information and drug responsivity) and the putative assays/biomarkers we measure. Depending of the translatability and strength of the generated models, they could possibly be used for broader research endeavors in the future. We are in touch with LRD regarding these issues.

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 restricting dissemination are in place, since we do the multiplex IHC and generate the in vitro and in vivo models in our laboratory.

4. Documentation and metadata

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

All collected data points (i.e., cell line information, diagnosis, clinical information, genetic information, proteomic information, protocols such as for MILAN and TMA grand master, etc) are and will be stored in a coded manner in LabCollector (an online labmanagement system with integrated Electronic Lab Notebook (ELN)) which runs on a secured and backed up server of

KULeuven (managed by ICT of the Biomedical Sciences Group). All patient information (age, sex, disease) will be registered in a pseudononymised way in the file containing all collected samples. Every patient will receive an identification number which can only be decoded by the responsible data manager. This system also provides a logging system so no data can ever be erased, making that everything will be tracable and stored longterm (way beyond the common 5-year requirement). Only coded information will be extracted and used for the downstream research analyses. Patient material, patient-derived cell lines and PDX mouse tissues are and will be stored in -80°C freezers or liquid nitrogen, where the exact location is defined in labcollector.

Images will be clearly labeled to identify experiment, sample, and date. The labeling of every individual image taken is standardized and defined in a standard operating procedures (SOP) lab protocol. For the images analysis, the algorithm used for the software analysis will be noted in an excel database per experiment with the info about the relevant parameter changed.

All files will be clearly labeled, to include experiment number and conditions (eg. sequencing experiments). For experiments such as Elisa, qPCR and Isospark, including output like dose response curves with company-delivered cytokines/growth factors, I will save the results in Excel formats on the KUL server.

All experimental data will also be noted in the lab book including all detailed experimental data. All files and folders will be labeled in a clearly structured way. The explanation of the labeling and the performed analysis will also be written down in the lab book and digital lab book (ELN, Labcollector). Whenever SOPs are available, they will be saved in the same location to facilitate reuse in the future.

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.

All data is searchable and includes various levels of metadata. The LabCollector platform is structured according to projects and topics where all relevant information is directly linked to each experiment. The integration of a lab inventory to the ELN system in LabCollector, makes that information can be retrieved both from the experiment point of view or from the sample point of view (e.g., which experiments have been performed with cell line A, or which cell lines were used in experiment B). Metadata concerning the technical specifications of the multiplex IHC images (e.g., channel/gating…) are collected with the .czi files, or else will be noted down with each experiment. The used medical terminology (e.g., diagnosis) is according to the standards in the neuro-oncology and neuropathology field. Cell biology-based terms are explained in the database.

5. Data storage and backup during the FWO project

Where will the data be stored?

The collected cell lines are stored in cryogenic tanks (Liquid Nitrogen) in our laboratory (Campus Gasthuisberg, O&N4, 5th floor). The extracted DNA/RNA/Protein/snap frozen materials are stored in -20°C/-80°C freezers. Fixed materials are stored in designated closets in the lab. Storage locations are included in the labcollector database to keep track of each sample.

The data generated during the project will be safely stored in our research unit central storage facility. Copies can be made and kept on personal devices. I will make use of a physical lab book with the chronological reporting of all related experiments and results including a cross reference to electronic storage of data. These lab books are owned by the research group and remain in the laboratory during the entire time. All obtained research data points are also kept in electronic lab note books in the LabCollector database of the Translational Cell and Tissue Research group, located on a KU Leuven hosted and secured server and is password protected.

Large data sets, such as images from microscopy, are stored on the KU Leuven K-drive and L-drive (large storage servers). In addition, external hard drives for temporary storage are bought as deemed necessary. On top of the large storage servers from KU Leuven, we regularly expand the longterm storage capacity with a user-friendly NAS (network-attached storage) coupled to the KUL storage servers.

How is backup of the data provided?

Backup is secured automatically and daily on central hosting KUL servers of the university. The vast majority of the generated data is stored in 2 different places, e.g., external hard drive and KUL server, external hard drive and NAS. In case of data loss, one can contact IT to retrieve the back-upped data.

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

Yes, we have unlimited space available – payable on yearly basis. Storage of cell lines/biological samples is limited to the size of our storage rooms. However, we are planning to move the less relevant materials to the UZLeuven Biobank from the moment it will be ready to accept materials (will be for free initially).

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

<1000 euro/year for server space and backup. Labcollector has been purchased (only updates required), where general lab budget is foreseen to keep this system up and running. Other costs are currently connected to the pricing model of KULeven for large volume and cold storage. External hard drives for temporary storage are bought as deemed necessary (20 Tb per drive, +-550 euro).

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

Secured KULeuven server with automatic back-up + password access by users + person-based decision on rights to access and modify data + every modification to data is logged + no data can be erased. In case necessary, Belnet filesender will be used for a secure transfer of files.

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 research data generated will be retained for at least 10 years after the project, thus complying to the data preservation rules of KU Leuven. Data will be stored in a safe, secure and sustainable way for purposes of reproducibility, verification and potential reuse. For optimal preservation of research data, I aim to convert data to open or standard file formats to easy future use of the data.

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

The data will be stored on the university's central servers (with automatic, daily back-up procedures) for at least 10 years, conform the KU Leuven RDM policy. Data is also stored on Labcollector, hosted on a secured KUL server with password-protected access. The server is automatically backed up using KUL services. The system provides a logging system so no data can ever be erased and everything will be tracable and stored longterm (beyond the 5-year requirement).

Cell lines and other (co-culture) models are stored in cryogenic tanks (Liquid Nitrogen) in our laboratory. Extracted materials are stored in -20°C/-80°C freezers.

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

Costs directly depend on the pricing of storage drives of KULeuven for large volume and cold storage. These costs will be covered through the general lab budget (non-related grant applications). As stated before, we except <1000 euro/year for server space and backup, where Labcollector has been purchased (only updates required). The Translational Cell and Tissue Research Unit also decided to invest in this system for our entire group.

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

There are no concrete factors restricting or preventing the sharing of produced/collected data as defined with a 3rd party or legal restriction. The informed consent includes coded data sharing with both academic and non-academic parties. The data will be published before making it partly available. Our goal is to fuel collaborations where our data and materials can be used.

Considering the use of patient materials, data will be either shared in a coded manner through publications, or, if more specific information is required, following ethical approval. Data sharing can also happen with specific access controls in place, for example only allowing reuse for research purposes.

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

Data will be made publicly available post publication depending on the journals policy (post-publication data repository). Non-published data will remain confidential until a final decision on publication of the data has been taken. All data will be available in a collaborative setting (i.e. any other internal and external research group with whom we may work in the future that could benefit from data and materials gathered in this project).

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

- In a restricted access repository

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) on this repository, we will evaluate data sharing and data reuse depending on the policy and conditions of this repository. The conditions of access will be determined by the need of the third party that wants access (e.g., only data, only materials, industrial or academic party, etc). External users will have access to our database through our publications and after setting up a new collaboration. For academic users the requirement will be to be included as co-authors in publications where our data and materials are used. For industrial users, the conditions will be determined in collaboration with LRD.

When will the data be made available?

- Upon publication of the research results

Data will be made available after publication in peer reviewed journals and/or after signing a data sharing agreement which will be established with the support of KUL R&D after a request. Additional material (and associated data) will be made available, on basis of material transfer agreements (MTAs) if the objectives fulfill the aims referred in the informed consent signed by the donors of the biopsies and after approval by the UZLeuven ethical committee, if requested by a third party. Changes to the data sharing policy will be reported in the final DMP.

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

All data will be available in a collaborative setting. Due to the potential commercial value of data, no general and full open access to data will be provided by default. Data which will be shared with third parties, will exclude commercial use and will require appropriate credit to the data owners. Detailed data sharing agreements will therefore be implemented.

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

The KU Leuven repository will not request 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 and will be reported in the final DMP at the end of the project. They will be covered through funds available in the laboratory. For co-culture sharing if legally approved, the KUL biobank where the samples will be registered will calculate the cost for sample handling and shipment and also for administrative work. Access to the lab materials (e.g., cell lines) will be covered by the receiving party.

8. Responsibilities

Who will be responsible for data documentation & metadata?

As an individual researcher producing data, I will have the final responsibility for data documentation and metadata. If necessary, I can reach out to technicians and administrative personnel for some help and guidance.

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

Data storage and back-up underlies the responsibility of the individual researcher, who will be supervised by the scientific coordinator. The responsibility for maintaining the infrastructure access for data storage lies in the hands of the IT responsible of the research team. Finally, the

maintenance of servers and integrity of data stored on these servers underlies the ITC services of the university. The maintenance of the integrity of the external drives and the data stored on them will be responsibility of the researcher and the lab manager.

Who will be responsible for ensuring data preservation and reuse ?

Both me and my PI, Prof. Frederik De Smet, will take responsibility to ensure data preservation, access and reuse.

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

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