MY PLAN (INTERNAL FUNDS DMP)

PREDICTING METASTASIS FORMATION BASED ON CANCER CELLS WITH A DISTINCT ENERGY METABOLISM

ADMIN DETAILS

Project Name: My plan (Internal Funds DMP) - Predicting metastasis formation based on cancer cells with a distinct energy metabolism
Project Identifier: CELSA/22/031
Grant Title: CELSA/22/031
Principal Investigator / Researcher: Sarah-Maria Fendt
Project Data Contact: Sarah-Maria Fendt
Description: Breast cancer is the leading cancer-associated cause of death in women1 and metastasis formation accounts for almost 90% of the associated mortality resulting in more than 685,000 deaths per year. Unfortunately, we currently cannot predict which patient will develop metastases due to our limited understanding of how cancer cells gain metastatic capacity. This greatly reduces the application of personalized follow up screening and preventative measures. We hypothesize that increased copper transporters Slc31a1 and Slc25a3 expression potentiates mitochondrial function and as an extension of this, is predict of metastasis formation and may be targeted for therapy. Thus, we will 1) identify the mitochondrial function of increased Slc31a1 and Slc25a3 expression in cancer cells; 2) Determine whether inhibition of either Slc31a1 and Slc25a3 impairs specifically breast cancer metastasis in mouse models; 3) Define whether Slc31a1 and Slc25a3 protein expression is predictive of metastasis formation in an independent cohort of breast cancer patients. To address this hypothesis, we will rely on the expertise of Prof. Chinopoulos (Semmelweis University, Hungary, SE) in assessing mitochondrial biological function combined with the cutting-edge expertise of Prof. Fendt (KU Leuven, Belgium, KUL) in assessing metastasis formation and biology and the renowned expertise of Prof. Gijsbers at the Leuven Viral Vector Core (LVVC) in genetic engineering using viral vectors (KU Leuven, Belgium, KUL).
Institution: KU Leuven

1. GENERAL INFORMATION

Name of the project lead (PI)
Prof. Sarah-Maria Fendt

Internal Funds Project number & title
CELSA/22/031- Predicting metastasis formation based on cancer cells with a distinct energy metabolism

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.

Work package 1 (Lead: SE)

The Gijsbers lab (KUL-Gijsbers) will generate miRNA-based viral vectors against mouse and human genes of interest. Knockdown breast cancer cell lines will be generated and validated in the Fendt Lab (KUL-Fendt). Mitochondrial function of control and knockdown cells will be assessed in the Chinopoulos Lab (SE).

AIM: Deliver a functional evaluation of the role of genes of interest in cancer cells with metastatic potential.

1. **Genome engineering**: Plasmids, oligonucleotides and lentiviral vectors to generate knockdown or overexpression cell lines will be stored at the appropriate temperatures (-20 or -80 degrees). They will be linked to a unique identifying number and a digital Snapgene map (.dna). Inventory excel files (.xlsx) for plasmid/oligonucleotides/lentiviral vectors will contain detailed information about the targeted genes, species, and storage location. Following sequencing of the plasmids, sequencing files (.seq files) are attached to the Snapgene map. They will be maintained on the Gijsbers lab server (KU Leuven J-drive). *Estimated volume: <10MB*
2. ***In vitro* cell culture measurements**: Mouse breast cancer cell line (4T1) and human breast cancer cell line (MCF10A-RAS^{v12}) will be used as *in vitro* models for cell culture data collection. Frozen stocks will be stored in liquid nitrogen tanks and excel files (.xlsx) with a detailed description of the frozen stocks and their location will be updated regularly. Cell counting will be performed using a cell counter and resulting data will be manually recorded in an excel file (.xlsx). Microscopy images of the cell cultures will be stored as image files (.jpg). All above mentioned files will be stored on the Fendt and Chinopoulos lab server (KU Leuven L-drive and SE kozoslo server (SE-K), respectively). *Estimated volume: <1GB*
3. **Gene expression analysis for validating knockdown**: RT-qPCR data from RNA, extracted from cells/tissues, will be exported to excel files (.xlsx). Oligonucleotide primers for qPCR will be stored at -20 degrees, and will be listed in the laboratory primer inventory excel files (.xlsx) with unique identifying numbers, base pair sequences, the target gene/species and storage location. These files will be stored in the Fendt lab server stored on the KU Leuven L-drive. *Estimated volume: <10MB*
4. **Oxygen consumption measurements**: The oxygen consumption rate of control and knockdown breast cancer cells will be assessed *in* vitro using the Seahorse XF analyzer. Raw data files will be stored as XFD files (.xfd) and directly exported as excel files (.xlsx). All data will be stored on SE-K server. *Estimated volume: <10GB*

5. **Fluorescent microscopy**: The *in situ* mitochondrial membrane potential, NAD(P)H autofluorescence, adenine nucleotide translocase (ANT) directionality and mitochondrial morphology measurements in intact cells will be assessed using an Olympus IX81 inverted microscope equipped with a UAPO 20 × 0.75 NA lens, a Bioprecision-2 xy-stage (Ludl Electronic Products Ltd., Hawthorne, NY) and a 75W xenon arc lamp (Lambda LS, Sutter Instruments, Novato, CA). Time lapses of 1342 × 1024 pixels frames (digitized at 12 bit with 4 × 4 binning, will be acquired by an ORCA-ER2 cooled digital CCD camera (Hamamatsu Photonics, Hamamatsu, Japan) under control of MetaMorph 6.0 software (Molecular Devices; Sunnyvale, CA, USA). Data output will be in .nd format. Data will be analyzed offline using Image Analyst MKII (Image Analyst Software, Novato, CA), and exported as time-lapse acquisition of TIFF files. Fluorescent intensities of regions-of-interest (ROIs) will be digitized and converted to excel files (.xlsx). All data will be stored on SE-K server. *Estimated volume: <500GB.*

Work package 2 (Lead: KUL-FENDT)

The Fendt lab will validate the metastatic potential of CD90.1$^+$ control and knockdown breast cancer cells *in vivo.*

AIM: Determine whether inhibition of the genes of interest impairs breast cancer derived metastasis formation in mice.

1. ***In vivo* mouse data**: We will use syngeneic mouse models: CD90.1$^+$ knockdown and control 4T1 breast cancer cells will be injected intravenously or in the mammary fat pad of Balb/c mice. Body weight and tumor burden will be recorded manually in Excel files (.xlsx). Images of mouse tissues/metastasis (lung) will be taken with a digital camera and saved as .jpg image files and saved on the KU Leuven L-drive. *Estimated volume: 1GB*
2. **Flow cytometry**: The lungs of the mice will be harvested and the CD90.1$^+$ cell fraction will be analyzed via flow cytometry. Raw flow cytometric data will be stored as FCS files (.fcs) and processed using FlowJo software, leading to the generation of FlowJo workspace files (.wsp). All data will be stored on the Fendt lab server (KU Leuven L-drive). *Estimated volume: 5GB*
3. **H&E stainings**: Isolated primary tumor tissue and lungs will be sectioned and assessed using H&E staining. High quality scans of the H&E stainings will be taken (.qptiff) and metastatic burden will be analyzed using the QuPath software (.xlsx). Files and images will be stored on the Fendt lab server on the KU Leuven L-drive. *Estimated volume: 100GB*

Work package 3 (Lead: SE)

The Chinopoulos Lab will assess the expression of the proteins of interest in metastasis formation in patient samples via reverse phase protein array.

AIM: Define the predictive power of the expression of the proteins of interest for metastasis formation.

1. **Human sample data**: We will use 293 healthy and 293 paired breast tumor samples available to Prof. Chinopoulos at the Semmelweis university for Reverse Phase Protein Array (RPPA) analysis. All patient data are being alphanumerically encrypted using a 128-bit encryption algorithm and stored in a remote server (db.rppa.hu). Prof. Chinopoulos is the project leader of the RPPA facility

of SE, currently hosting >4,000 human healthy and tumor samples pre-processed for RPPA (http://rppa.hu).

2. **Reverse Phase protein Array**: Because of the high-throughput format (thousands of samples can be analyzed simultaneously on the same array) and the small amount of biological sample that is required, it facilitates large-scale multiplex analysis of multiple markers from *in vitro*, preclinical and clinical studies. We anticipate to use ~50 antibodies directed against proteins of interest. Extended protocols can be found at http://rppa.hu/Protocols.html. Final data output will be in excel files (.xlsx) in a format that is recognized by a custom-made python script (.py) for data visualization. Analyzed data will be stored as .txt files on the SE-K server. *Estimated volume: 10GB*

General:

1. **Statistical analysis and graphical figures**: GraphPad Prism 9 will be used for all statistical analyses and graphical designs, leading to the generation of program-specific files (.pzfx). Final figures will be exported in the following file types: .pdf, .png, .wmf, .svg. All data will be stored on the server of the lab that generated the data (KU Leuven J/L-drive or SE-K server). *Estimated volume: 5GB*
2. **Protocols, documentation, metadata, and papers**: Experimental and computational protocols, as well as details related to collection and processing of data (both documentation and metadata) will be stored in Word (.docx), Excel (.xlsx), or text (.txt) files. Manuscripts originating from the project will be stored as Word documents (.docx), and final versions will be exported to PDF format (.pdf). All data will be stored on the server of the lab that generated the data (KU Leuven J/L-drive or SE-K server). *Estimated volume: <1GB*

Data will be shared between laboratories via One Drive.

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.

Yes, clinical data from patients will be accessed through Prof. Chinopoulos (SE) at the RPPA facility at Semmelweis University and will only contain information on type of tissue, hazard level (i.e. hepatitis C infected, …), gender, date of birth and site of origin of the sample (i.e. clinic from where the sample was obtained).

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).

Yes, animal experiments will be performed during this project. All experiments have been approved by the Ethical Committee for Animal Experimentation (ECD) at KU Leuven and are available and outlined in ECD #P048-2021.

Human samples will be used during this project. The collection and use of samples obtained through Prof. Chinopoulos has been approved by the local Ethical Committee (ETT-TUKEB), file number: 35302-5/2017/EKU. The approval is available and outlined in http://rppa.hu/ETT-TUKEB%20ethical%20permission%20for%20VEKOP-2.3.3-15-2016-00012.pdf

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?

Yes, intellectual Property arising from this work is managed as per the framework agreement between the VIB (VIB Tech Transfer), the KU Leuven and Semmelweis University, the three participating institutes in this study.

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?

Yes, as above, dissemination or exploitation of the data is managed according to the framework agreement between the VIB, the KU Leuven and Semmelweis University.

4. DOCUMENTATION AND METADATA

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

Protocols and details related to data collection and processing will be recorded in physical and electronic lab books and transcribed to Word or Excel files by lab technician Maxime Smits (GIJSBERS lab), post-doctoral scientist Dr. Konstantinos Kolliopoulos (FENDT lab) and PhD student Noemi Karnok (CHINOPOULOS lab). Data folders containing raw and processed data will be hierarchically organized and labeled based on the source of the data, the type of experiment, the date of data generation, and the different experimental conditions analyzed. Data analysis methods and particularities (including metadata) will be described in text documents and Excel files included in these folders. All files will be stored in the KU Leuven drives (J/L-Drive) and SE-K server with sharing possibilities via One Drive.

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, text documents and Excel files stored within each experiment folder in the KU Leuven J/L-Drive or SE-K server will contain guidelines describing data collection/analysis methods and all relevant metadata (including experimental conditions, sample keys, computational analysis pipelines and their parameters) to ensure the reusability of the data and the reproducibility of any further data generation.

5. DATA STORAGE AND BACKUP DURING THE PROJECT

5.1. Where will the data be stored?

Upon data collection/preprocessing, temporary copies of the data will primarily be stored in the KU Leuven/Semmelweis University-managed personal computer of the lab technician Maxime Smits (Gijsbers

lab), post-doctoral scientist Dr. Konstantinos Kolliopoulos (Fendt lab) and PhD student Noemi Karnok (Chinopoulos lab). A copy of the data will be immediately uploaded to the KU Leuven drives (J/L-Drive) and SE-K server for long-term preservation and backup.

5.2. How will the data be backed up?
Data stored on the KU Leuven J/L-Drive is managed, maintained, and backed up by KU Leuven IT services. Data stored on the SE-K server is managed, maintained, and backed up by Semmelweis University IT services.

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.
The KU Leuven J/L-drive and SE-K server have sufficient storage capacity for the outlined project.

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?
The annual cost of KU Leuven L-Drive storage is 569.2 euro/5TB/year (Fendt lab). The annual cost of the KU Leuven J-Drive storage is 51.9 euro/100Gb/year (Gijsbers lab). The SE-K server is cost-free for the Semmelweis lab (Chinopoulos lab). This cost and capacity include the performance of mirror copies of the stored data, for safety backup purposes. We expect that 5TB will be sufficient to store all data generated as part of the project. These costs will be covered by the budget of each PI independently.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?
For KU Leuven laboratories: Data stored intermediately on personal computers are protected via password access to the computers. Off-site access to J/L-drive data is available by KU Leuven data access points and is password protected. Access to modify these files are limited to lab members with access to the Gijsbers or Fendt lab J/L-Drive folders. For Semmelweis laboratory: Data stored intermediately on personal computers are protected via password access to the computers. Off-site access to SE-K server data is available through a VPN and is password protected. Access to modify these files are limited to lab members with access to the Chinopoulos lab SE-K folders.

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, ...).
All raw data will be retained for 10 years on the KU Leuven J/L-Drive and SE-K server. Publication data is further organized and catalogued by figure for future reference to raw datasets used for figure generation.

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

Long term data archival is maintained in specific archive folders on the KU Leuven J/L-Drive and SE-K server for at least 10 years, conform the KU Leuven RDM policy.

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

The annual cost of KU Leuven L-Drive storage is 569.2 euro/5TB/year (Fendt lab). The annual cost of the KU Leuven J-Drive storage is 51.9 euro/100Gb/year (Gijsbers lab). The SE-K server is cost-free for the Semmelweis lab (Chinopoulos lab). This cost and capacity include the performance of mirror copies of the stored data, for safety backup purposes. We expect that 5TB will be sufficient to store all data generated as part of the project. These costs will be covered by the budget of each PI independently.

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)?

Yes, we will use healthy and paired breast tumor samples available to Prof. Chinopoulos at the Semmelweis University at the RPPA facility. Prof Chinopoulos is project leader of the RPPA facility of SE, currently hosting >4,000 human healthy and tumor samples pre-processed for RPPA (http://rppa.hu). Therefore, he has a Material and Data Transfer Agreement (MTA) in place/filled. Third party agreements will not restrict dissemination of the data that will be generated within this project.

The results of this project will be communicated in established, peer-reviewed (non-predatory) academic journals, that require disclosure of all included data. Patient data will therefore be made available in restricted access repositories. Access to these data will be made available to any individuals making a specific request to Prof. Christos Chinopoulos, project leader of the RPPA facility at Semmelweis University. Personal data will only be published after de-identification to protect the identity of the patients. The identifiers will not be published.

In order to protect the privacy of the patients, their tumor samples will only be made available to lab members involved in the project, not to other groups or studies unless ethical approval has been obtained. Other data will be shared externally upon request of collaborating scientist, which will be reviewed and approved on case-by-case basis by Prof. Fendt, Prof. Chinopoulos and/or Prof. Gijsbers.

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

The key findings of the project and their interpretation will be made available to peers and the wide audience through publication in established, peer-reviewed (non-predatory) academic journals. Data related to patient samples will be made available in repositories with appropriate access control. Access to these data will be made available to any individuals making a specific request to Prof. Christos Chinopoulos. Any data shared will only be released prior to a Data Transfer Agreement that will have to include the necessary conditions to guarantee protection of personal data. The personal data will be de-identified and the identifiers will not be published to protect the privacy of the patients.

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

- In a restricted access repository

- Upon request by mail

Datasets will be made publicly available, if possible, through repositories to support FAIR data sharing.
All patient data will be made available in repositories with appropriate access control upon de-identification of the personal data. Access to these data will be made available to any individuals making a specific request and this request will be handled by Prof. Christos Chinopoulos. The remaining data can be approved for reuse upon request and will be assessed for approval on a case-by-case basis by the Prof. Sarah-Maria Fendt, Prof Rik Gijsbers and Prof. Christos Chinopoulos.

7.4. When will the data be made available?

- Upon publication of the research results

Summaries of key findings of the project and their interpretation will be made available through the publication of journal articles in reputable academic journals. Upon publication, the human data will be made available in a restricted access repository. Patient data will be de-identified to protect the privacy of the patients.

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

Unpublished data will only be accessible to lab members involved in the project. Upon publication of the data in reputable journals, the published data will be made publicly available through data repositories to support FAIR data sharing. Any user can place reasonable data requests for non-commercial purposes, and these requests will be assessed on a case-by-case basis by the PIs (Prof. Sarah-Maria Fendt, Prof. Christos Chinopoulos and Prof. Rik Gijsbers), and in case of patient data, assessed in consultation with Prof. Christos Chinopoulos. Commercial-based requests will be navigated in coordination between Prof. Sarah-Maria Fendt, Prof. Christos Chinopoulos and Prof. Rik Gijsbers, if applicable: UZ/KU Leuven, Semmelweis University and the VIB/VIB Tech Transfer team.

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

Costs for data sharing will be discussed with collaborators on a case-by-case basis. To minimize data management costs, free-to-use data repositories will be used when possible. Data management will be covered by own funding.

8. RESPONSIBILITIES

8.1. Who will be responsible for the data documentation & metadata?

Prof. Sarah-Maria Fendt, Prof. Rik Gijsbers and Prof. Christos Chinopoulos accept all responsibility for data documentation and metadata. Lab technician Maxime Smits (Gijsbers lab), post-doctoral scientist Dr. Konstantinos Kolliopoulos (Fendt lab) and Prof. Christos Chinopoulos will be responsible for experimental data.

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

The lead postdoctoral scientist (Dr. Konstantinos Kolliopoulos (Fendt lab)) and Noemi Karnok (Chinopoulos lab) will be primarily responsible for collecting/generating data, and for correct documentation and upload onto the L-Drive and SE-K storage space. The KU Leuven IT department will be responsible for

maintenance and back up of the L-Drive data storage space. The Semmelweis University IT department will be responsible for maintenance and back up of the SE-K server storage space.

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

Prof. Sarah-Maria Fendt, Prof. Rik Gijsbers and Prof. Christos Chinopoulos will be responsible for ensuring data preservation and reuse.

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

Prof. Sarah-Maria Fendt, Prof. Rik Gijsbers and Prof. Christos Chinopoulos bear the end responsibility of updating & implementing this DMP.