DMP OLYMPUS

ADMIN DETAILS
Project Name: OLYMPUS
Project Identifier: G098025N
Grant Title: Exploring the role of lymphatics after lung transplantation

Principal Investigator / Researcher: Bart Vanaudenaerde
Description: Lung transplantation is a critical treatment for end-stage lung diseases, yet long-term survival remains challenging due to complications like acute and chronic graft rejection, of which the mechanisms are poorly understood. Our project aims to unravel the role of the lymphatic system, a crucial component of the body's immune defense and fluid balance, in the development of rejection after lung transplantation. We hypothesize that disruptions and regenerations of the lymphatic network in transplanted lungs contribute to complications such as primary graft dysfunction, acute rejection and chronic lung allograft dysfunction. Employing innovative imaging methods, we will map the entire lymphatic network in mouse and human lung samples, providing a holistic view of its interplay with blood vessels and airways. Our unique approach, including longitudinal analysis and modulation of lymphangiogenesis, set the stage for a comprehensive understanding of the lymphovascular role in rejection. Transcriptomics analysis will identify key molecular factors driving rejection, potentially serving as biomarkers. Access to rare human samples and a well-established mouse lung transplant model positions us to make groundbreaking strides, with potential implications for improved management of lung transplant recipients and insights into various lung diseases.

Institution: KU Leuven

1. GENERAL INFORMATION

Name of the project lead (PI)
Bart Vanaudenaerde

C1-C2 Project number & title
Exploring the role of lymphatics after lung transplantation. FWO G098025N

2. DATA DESCRIPTION

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

- Generate new data
- Reuse existing data

2.2. You have set a number of objectives in your proposal. Describe the origin, type and format of the data and its (estimated) volume that will be used to obtain each objective, preferably per objective. You might consider using the table in the guidance.

WP1. Implementing a novel and cutting-edge optical imaging methodology to facilitate 3D visualization of the lymphatic network in parallel with the vasculature and lung airways in intact mouse and human lung samples. Co-registration of 3D optical imaging with micro-CT and histopathology will yield novel insights into the intricate interplay between the lymphatic network, blood vasculature, and airways.

Optical imaging and MicroCT Scans of lungs (8-14GB) and small cores (5-6GB) of lungs tissue from human and mouse exist of large datasets in tiff and mha format. Depending on the application software packages include RadiAnt, ITK-SNAP, NeuronStudio, ImageJ, Fiji, Vision4D, Arivis, Imaris, ClearMAP,Mimics, Dragonfly. For the conversion from DICOM to tiff and/or mha ImageJ is used. Histology digital images are saved as tiff format.

WP2. Assessment of lung lymphovascular development in high-resolution using *ex vivo* imaging post-transplantation and during progression to chronic rejection in mouse (isograft and allograft) lung transplant samples and human explant lung samples.

Optical imaging and MicroCT Scans of lungs (8-14GB) and small cores (5-6GB) of lungs tissue from human and mouse exist of large datasets in tiff and mha format. Depending on the application software packages include RadiAnt, ITK-SNAP, NeuronStudio, ImageJ, Fiji, Vision4D, Arivis, Imaris, ClearMAP,Mimics, Dragonfly. For the conversion from DICOM to tiff and/or mha ImageJ is used. Digital datasets include tiff/jpeg format images and avi format movies all stored on GBW-0017-LTx.

WP3. Identification of lymphangiogenic key factors as possible biomarkers of rejection using (single cell) transcriptomics on mouse (isograft and allograft) lung transplant samples and human explant lung samples, with a focus on local and circulating lymphatic biomarkers.

The physical data include human and mouse tissue, sections, cells and RNA. Histology digital images are saved as tiff format. For tissue and single cell/nucleus RNA profiling datasets are in txt files. Datasets have a size of about 50MB and sub analytical file 2-5GB.

WP4. Modulating lymphangiogenic development by targeting key lymphangiogenic factors. Using the mouse lung transplant model, we will investigate the impact of stimulation or inhibition of lymphangiogenesis on lung rejection by evaluating structural alterations and biomarker analysis.

The same methodology and digital file formats as all previous WPs in addition to xml and bak format in a zip file for raw data of lung function measurement.

For all WPs extracted analyzed data are converted to MicroSoft Excel, Graphpad prism and R-statistical software for statistical analysis and presented in Word and PowerPoint files.

3. ETHICAL AND LEGAL ISSUES

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use (and add the reference to your file in your host institution's privacy register - not relevant yet)

Yes

Personal data collected are general demographic information (age, gender, length, body weight, home address...) but also specific medical information on the lung condition (lung function, radiological evaluation,

pathology of the lung) and general health status and therapy (CRP, smoking behavior, steroid use, immunosuppression…).
We will be using intact human lungs collected within our biobank S51577 for studying airway morphometry, cellular composition, aerodynamics, and specific in vitro experiments on the airway epithelial cells.

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).
The project is dealing with explant human lungs collected at the time of transplantation and mouse experiments are considered.
Running ethical approval for WP 1 to 4 to collect the human lungs: biobank S51577.
Running ethical approval for WP 1 to 4 for microCT/OMICS: S52174.
Running ethical approval for WP3 snRNA OMICS in mice: P08/2017
A new ethical request is needed for mouse experiments for WP1, 2 and 4: this will be arranged within the first 6 months of the start of the project.

3.3. Does your work possibly result in research data with potential for tech transfer and valorization? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?
Yes, this might be possible
IP may include:
- lymphatics imaging
- RNA profiles from single nuclear RNA profiling resulting in pharmacological targets
This will be explored on time together with KULeuven LRD department in case findings demonstrate potential.

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?
no

4. DOCUMENTATION AND METADATA

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?
Standard experimental procedures (SOPs) and practices will be fully documented as word (and PDF) and saved on the GBW-0076_LTx server.
SOPs include:
WP1: lung collection and processing
Optical imagaing and microCT scanning and digital processing of lung casts
Immunostaining and histology of lymphatics
WP2 identification and understanding of lymphatics (in transplanted lungs) of human and mice
WP3 identifycation of algorithms of lung lymphatics endothelial cells in snRNAseq data set

WP4 mouse handling and generating lung emphysema and fibrosis together with analytical parameters
Raw experimental data from the analytical SOPs are collected per experimental test, and will include a MS Word file with a clear description of what the data represent and how they were generated. This description will be documented in notebooks (with page numbers), as well as in electronic format (word files).
The name of the folder always contains the date, name of the experiment, and the name of the person who performed the experiment.
Each individual file with experimental data contains information on the study design, the origin of the samples, and all necessary information for an independent analyst to use or reuse the data accurately and efficiently.

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.
Raw data are stored at the slower long-term storage GBW-0017_LTx which includes an accompanying meta datafile of the experimental settings. The format is method and software dependent and includes xml, Dicon, lif, xml, fcs, vsi,… or excel. Other metadata from experiments in a searchable database format (in an excel file) include data, initials of the investigator, type of the study (clinical, pre-clinical, in vitro, compound used). Project specific analyzed data per fellow, PI will include an excel file, statistical file (Graphpad Prism, R software, SAS), powerpoint file and word file allowing to reconstruct and reanalyze each manuscript. Temporary files per project will be stored on the faster GBW-0076_LTx server per fellow/PI during the project but once finished and published they will be stored on the larger server to back-up the study for a period of minimal 5 years.

5. DATA STORAGE AND BACKUP DURING THE C1-C2 PROJECT

5.1. Where will the data be stored?
The physical data will be stored in appropriate storage places including histology rooms, fridges, freezers and cryotanks.
Digital data will be stored and protected in accordance with the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
Sensitive personal data will be pseudonymized if they go out of our KUL Pneumology research group. Our data will be stored at the KUL university's secure environment, of which daily backup are made by the ICT to secure the data. Our policy to protect data within our group is to code files at the server and personal PCs of Laptops.
Our dataservers are:
GBW-0076_LTx is a smaller (0.5Tb) server but faster server on which each fellow stored analyzed/processed data).
GBW-0017_LTx is a larger (15Tb) server but slower on which we store all raw data for long time. The annual costs are 156,6 euro.
Copies of the raw data and unanalyzed data are made and kept on personal devices only by members of our research group. To protect data that is locally stored or cached on computer drives, all Windows OS

hard disks are encrypted with centrally managed Bitlocker technology and protected with a PIN code and a centrally backed up recovery key (in case of lost PIN). Users of Mac hardware are provided with configuration instructions and asked to enable FileVault for hard disk encryption. Researchers are encouraged to encrypt removable media such as USB sticks and external hard drives using the same techniques. To exchange pseudonymized data with collaborating centers during the project, we transfer data by BelnetSender.
Access to these code files is controlled by a research fellow working with the data and by the lab manager and PIs of the project.

5.2. How will the data be backed up?

Snapshots

Backups of project data on faculty network shares are made using “snapshot” technology, which is the online storage of incremental data changes. The standard backup regime is as follows:

• An hourly backup (at 8 AM, 12 PM, 4 PM and 8 PM) the last 6 of which are stored on our servers
• A daily backup, at midnight, the last 6 of which are stored on our servers
• A weekly backup, Saturday night at midnight, the last 12 of which are stored on our servers

The end user can use his own Windows PC to restore files to an older version using the “previous versions” function. According to the above backup scheme, it is possible to go back in time up to 12 weeks (~3 months).

Mirror

For the purpose of “business continuity” or “disaster recovery”, a mirror (exact copy) of all data is created in a second datacenter. The data are copied every hour to the second datacenter. In the event that the primary storage unit is corrupted, the ICTS team can get this copy online within the hour.

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.

Yes and this is still expandable:
Our data will be stored at the KUL university's secure environment, of which daily backups are made by the ICT to secure the data. Our dataservers are:
GBW-0076_LTx (a smaller but faster server on which each fellow stored analyzed/processed data).
GBW-0017_LTx (a larger server but slower on which we stored all raw data for long time).

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?

For GBW-0076_LTx the annual cost is 519 euro/terrabyte. We currently use 377 Gigabyte of the 500 Gigabyte reserved, resulting in an annual cost of 195 euro/year currently.
For GBW-0017_LTx the annual costs are 156,6 euro. We currently use 10,7 terrabytes of the 15 terrabytes reserved resulting is a yearly cost of 2350 euro currently.
The overall currently yearly cost for both servers is 2550 euro. The total budget needed specific for this project for the 4 years of the project is about 2500 euro. These costs are already covered for many years by the running projects and will be included in the budget of the laboratory by left budget and new projects.

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

The data are secured on both servers by the university ICT. All ICT solutions at KU Leuven are subject to the university-wide ICT information security standards. The group ICT service organizes the raw network storage, it procures from central ICT services in such a way that access permissions are limited, fixed, delegated to and audited by data managers who do not need to have an IT background.

Access to these servers are restricted to the research group members. The lab manager (Karen Maes) and the PI of the project control who has access to the servers.

6. DATA PRESERVATION AFTER THE END OF THE C1-C2 PROJECT

6.1. Which data will be retained for the expected 5-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, ...).

Certainly, raw data will be stored for a longer period.

Big dataset that impact on the need of server space are datasets of microCT scan, digital slides scans, RNA profiling datasets and immune/chemical stainings. Datasets of clinical data, analyzed data in excel or Graphpad prism files do not impact on server space.

Raw digital data of scans, slide picture and molecular profiles are stored on the large volume server. Analyzed subdatasets and clinical and end-analytical excel or Graphpad prism data stored at the smaller server by the individual researchers are always evaluated at the end of the PhD. The excess of data is removed and valuable data remain at the large volume server under the name the fellow in a folder ‘old fellows”. Currently this concept is permanent storage but this can be reconsidered in the future. A minimum of 5 years is considered which allows us to reanalyze all published work if needed.

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

Data will remain on the servers mentioned before. The large volume server (GBW-0076_LTx) is more used for the raw data and long-time storage.

6.3. What are the expected costs for data preservation during these 5 years? How will the costs be covered?

Considering the currently yearly cost we expect costs for data preservation to be about 2500 euro. The department CHROMETA reserves for each separate group per years a small budget which is enough to cover these annual (and total) cost of basic storage.

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)?

This project is restricted to the KULeuven and so no data sharing issue are involved.

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

For most part of the project, raw data will not be made open access available.
The only expected open access will be for RNA profiling (anonymized) at the time of publishing in international peer reviewed as this is international obligatory.

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

- In an Open Access repository

The full (anonymized) dataset will be uploaded in a cvs format in Zenodo

7.4. When will the data be made available?

- Upon publication of the research results

RNA profiling will be made available upon acceptance of the manuscript. For most other data, no release of raw data is requested and considered unless asked during manuscript revision. IP if applicable will request specific embargo's and will only be considered at the time needed.

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

Only RNA dataset will be made accessible upon acceptance of the manuscript and this for everybody without restrains.

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

No cost expected.

8. RESPONSIBILITIES

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

Data documentation and metadata will be organized by the PIs, postdocs, PhD fellow and lab technician of the project namely Bart Vanaudenaerde, Birgit Weynand, Max Nobis, Birger Thielemans and Celine Aelbrecht.

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

Both servers are dedicated to the PI of the project and access is managed by the PI and the lab technician. ICT (Gert Goos as contact person and PI) is handling back-up and if needed expansion of storage capacity.

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

All persons mentioned before being the PI, post-doc and lab Technician.

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

Prof Bart Vanaudenaerde (PI)