Plan Overview

A Data Management Plan created using DMPTuuli

Title: DEVELOPMENT OF FLUORESCENCE-TAGGED VEGF-C FOR IN-VIVO LIVE IMAGING TO MAKE IT A BETTER THERAPEUTIC TARGET FOR LYMPHEDEMA TREATMENT

Creator: Khushbu Rauniyar

Principal Investigator: Khushbu Rauniyar

Data Manager: Khushbu Rauniyar

Affiliation: University of Helsinki

Template: General Finnish DMP template

ORCID iD: 0000-0001-5485-7040

Project abstract:

Lymphatic vessels are an integral part of the immune system that defends us against bacteria and viruses. These vessels serve as a drainage system for returning interstitial tissue fluid and immune cells into the blood circulation. Dysfunctional lymphatic system causes swelling of tissues due to excess fluid, known as lymphedema. Primary lymphedema is a genetic disorder, whereas secondary lymphedema is the consequence of another, underlying disease. In developed countries, it is most frequently caused by cancer surgery ("breast cancer-associated lymphedema" alone accounts for 150,000 cases diagnosed annually in Europe and the USA). The primary factor that stimulates the growth of lymphatic vessels (lymphangiogenesis) is vascular endothelial growth factor (VEGF-C). Although this has been known since 1997, many mechanistic details of VEGF-C's action have only been recently uncovered. VEGF-C needs to be activated before it can generate new vessels and that this activation requires the CCBE1 helper protein and one of several proteases. Worldwide, there is no approved lymphedema drug and Lymfactin (a novel adenovirus type 5-based gene therapy expressing VEGF-C) from Finnish company Herantis Pharma is the only investigational drug to reach Phase II clinical trial. Despite the encouraging initial results from Lymfactin, its main limitations are the weak response to VEGF-C as a single agent and our incomplete understanding of the actual mechanisms of VEGF-C activation for the development of a functional lymphatic vasculature. My aim is to identify additional puzzle pieces in VEGF-C activation with the ultimate goal to generate better, different, and safer lymphedema drugs. Besides lymphedema treatment, identification of novel VEGF-C activating proteases such as KLK3 aka prostate-specific antigen and cathepsin D in my recent paper has opened ways to therapeutically target VEGF-C in reproduction and metastatic cancer. To further understand the lymphangiogenic signaling in detail, I have two different approaches: 1. Most of our knowledge about VEGF-C activation is based on in-vitro studies. While valuable, there is yet no possibility to monitor or detect VEGF-C in-vivo. In this respect, I want to develop a method that allows fluorescence-tagging of VEGF-C for in-vivo live imaging. I plan to jointly test the fluorescence-tagged VEGF-C in the zebrafish model with Kaska Koltowska (Uppsala University), an expert in zebrafish imaging. 2. Based on a mass spectrometric screen of purified VEGFR-3 activating capability (from human saliva and semen), I have identified additional proteins that are potentially important for lymphangiogenesis. My goal is to characterize the roles of additional proteases and cofactors necessary for general VEGF-C signaling.

ID: 14929

Last modified: 28-04-2021

DEVELOPMENT OF FLUORESCENCE-TAGGED VEGF-C FOR IN-VIVO LIVE IMAGING TO MAKE IT A BETTER THERAPEUTIC TARGET FOR LYMPHEDEMA TREATMENT

1. General description of the data

1.1 What kinds of data is your research based on? What data will be collected, produced or reused? What file formats will the data be in? Additionally, give a rough estimate of the size of the data produced/collected.

The data from my research will consist of nucleotide sequences for new plasmids cloned, data on a list of peptides and proteins from mass spectrometry analysis, raw data files from experiments, experimental analysis data files, microscopy images, videos, nucleotide and amino acid sequences, alignments. Data on protein purification will be reused for different proteins. Data on cloning new constructs will be recorded and submitted to Addgene. Data obtained from mass spec and N-terminal sequencing will be published in the supplementary material section of open access journals. The data will be in the following file formats:
.xls, .doc, .pdf, .ppt, .fasta,
The image files are stored in Web standard formats (.jpg, .gif, .png). Original, large images are stored as .svg or .tiff files. The data analyzed using graphpad prism will be saved as .pzfx file.
The approximate size of the data generated would be around 500GB.

1.2 How will the consistency and quality of data be controlled?

Electronic lab journal (eLabFTW) will be maintained throughout the project documenting every step of the research. Data will be version controlled and documented. Dates of data retrieval and changes will be recorded, making all data related actions traceable and repeatable. Everyone handling the data has been introduced to good laboratory practices in the research.

2. Ethical and legal compliance

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

No ethical issues are associated with data management in my project.

2.2 How will you manage the rights of the data you use, produce and share?

Standard Material Transfer Agreements (SMTA) will be provided when any plasmid or cell line is transferred between lab., which clearly states the use of materials rights. Most of the data will be shared in Addgene, GitHub, and open-access papers and will be encouraged to acknowledge my work. I will also patent my innovation if possible.

3. Documentation and metadata

3.1 How will you document your data in order to make it findable, accessible, interoperable and re-usable for you and others? What kind of metadata standards, README files or other documentation will you use to help others to understand and use your data?

All published and relevant data from the project will be openly available for other researchers either by depositing accepted manuscripts on HELDA or by request from the authors if results are not subject to patenting issues. When applicable, data generated by the project will also be deposited in various internet-based open data repositories. While sharing our data we will ensure that it is available for re-use under Creative Commons licenses. We will preserve the data by using persistent identifiers (PID) which enable us to access the data via a persistent link (e.g. DOI, URN).
For me and my lab members, we have shared folders for each published paper in google drive, onedrive and eLab notebook where every detail about the data generated is recorded.

4. Storage and backup during the research project

4.1 Where will your data be stored, and how will it be backed up?

We store our data in google drive (extra space bought) and the University group directory. Backups of the data are being created each hour by two external drives which we have in our lab and also in our PI's home. So, no data is ever lost. We also use backup and sync. CSC long-term storage facilities are utilized if server space will be an issue.

4.2 Who will be responsible for controlling access to your data, and how will secured access be controlled?

No sensitive data is associated with this project. Normal IT safety measures are followed by all team and project members to keep data from being stolen before publication.

5. Opening, publishing and archiving the data after the research project

5.1 What part of the data can be made openly available or published? Where and when will the data, or its metadata, be made available?

The data obtained from the research project will be well preserved by the lab members and will be published following the guidelines of open science in an open-access journal. In order to support

the "open access" of the University of Helsinki (HY), we will send the final draft of our research article along with the supporting information to the University library of HY. Metadata entries will be published immediately when they can be considered sufficiently complete, even if the data itself is not yet public.

5.2 Where will data with long-term value be archived, and for how long?

Long term data archival will be done using the CSC- Finnish IT Centre for Science Ltd IDA data-archival service. Data will be minimally stored for 10 years.

6. Data management responsibilities and resources

6.1 Who (for example role, position, and institution) will be responsible for data management (i.e., the data steward)?

The PI is the chief responsible person for data management during the project period. The PhD students and Postdocs will also share the responsibility to save the data.

6.2 What resources will be required for your data management procedures to ensure that the data can be opened and preserved according to FAIR principles (Findable, Accessible, Interoperable, Re-usable)?

eLabFTW to record data by all researchers in the lab. Each member in the lab is committed to using electronic lab journal and spend time each day to document data properly. External hard drives and subscription fees for elab books. Costs for publishing data in open-access journals.