Plan Overview

A Data Management Plan created using DMP Tool

DMP ID: https://doi.org/10.48321/D1DK5F

Title: HETEROTYPIC SIGNALING IN MELANOMA PROGRESSION: FROM CELL CULTURE TO SYSTEMS-WIDE ANALYSES

Creator: André Zelanis - **ORCID:** 0000-0003-3066-6527

Affiliation: Universidade Federal de São Paulo (unifesp.br)

Funder: São Paulo Research Foundation (fapesp.br)

Template: Digital Curation Centre

Project abstract:

Melanoma is aggressive skin cancer and a lethal melanocytic neoplasm with an increasing annual number of cases (faster than any other solid tumor). As changes in the secreted proteins (secretome) composition of cancer cells might often represent a portrait of the cellular status during disease progression, the profiling of cancer secretomes can lead to the discovery of cancer-related biological markers and, eventually, to new drug targets. The secretion of bioactive molecules by primary tumors enables the organotropism of metastatic cells for specific target organs and is involved in the reprogramming of nonmalignant stromal cells which, in turn, facilitate neoplastic growth and metastatic dissemination in a spatiotemporal manner. In this context, the interplay between tumoral cells and endothelial cells is of paramount importance for the transition from a localized invasion to the inner of blood and lymphatic vessels (intravasation) and, eventually, the leaving of these locations for invading the surrounding tissue (extravasation). Thus, the main objectives of this project are (i) to identify molecular patterns associated with cross-talk of endothelial cells (ECs) and melanoma cells in a co-culture system (mixed cells), including qualitative and quantitative analysis of secreted proteins, (ii) to evaluate proteolytic signaling events in the secretome of this co-culture system, (iii) to monitor the development and organotropism of melanoma cells in vivo, in animals treated or not with conditioned medium from the cells in co-culture; as well as (iv) to interrogate liquid biopsies (plasma) obtained from patients with different levels of progression of melanoma to search for markers associated with disease progression. It is hoped that the integration of these data may assist in understanding the interaction between tumor and endothelial cells for the intra/ extravasation process as well as for the development of the pre-metastatic niche and metastasis in melanoma. In addition, it is expected that, collectively, the results will provide a systemic panorama of the heterotypic signaling in

melanoma and of the biological circuits activated by proteolytic processing in melanoma, with significant translational potential.

Start date: 03-17-2022

End date: 02-17-2027

Last modified: 07-08-2024

Copyright information:

HETEROTYPIC SIGNALING IN MELANOMA PROGRESSION: FROM CELL CULTURE TO SYSTEMS-WIDE ANALYSES

Data Collection

What data will you collect or create?

Experimental approaches will generate 3 main types of data:
(1) Data regarding the profile of proteins from both melanoma cells and endothelial cells after co-culturing in metabolic labeling conditions;
(2) Proteomics data derived from metastatic tissues in mice injected with co-cultured cells;

(3) Data from targeted proteomics analysis of the plasma from patients with melanoma.

The mass spectrometric data format will be made available under the original vendor (RAW) format as well as will be converted to the universal data format 'mzML'. Assembled contig data will be made available in '.fasta' format. Proteomics and data will be submitted to public repositories such as the ProteomeXchange Consortium (http://www.proteomexchange.org/) In addition, for bioinformatic analysis purposes, data from the above repository might also be interrogated.

How will the data be collected or created?

Briefly, Mass spectrometric (RAW) data will be analyzed within the Trans Proteomics Pipeline platform (v.4.8 ; Build 201411201551-6764) with Comet search engine (version 2014.02, rev. 2) against the UniProt/SwissProt database restricted to the taxonomy 'Homo sapeins' or 'Rodentia'(latest releases). Proteolytic signaling events (N-terminomic data) will be analyzed using WebPICS, a web-based platform freely available (http://clipserve.clip.ubc.ca/pics/) and the TopFIND knowledge base, a database and analysis resource for protein termini and protease processing (http://clipserve.clip.ubc.ca/topfind). Search results will be further filtered with PeptideProphet to a > 99% confidence interval, corresponding to a False Discovery Rate (FDR) of less than 1%. For both proteomics and transcriptomics data, custom-made
scripts will be designed and the analyses will be performed in R scripting and statistical environment.

Documentation and Metadata

What documentation and metadata will accompany the data?

The data from the bioinformatics analyses (proteomics) will be presented in the form of a spreadsheet, with the respective RAW files (raw file referring to each mass spectrometry run) as well as the information
associated (generally, the output of proteomics data search software). File size as well as any other associated information will be available in the same spreadsheet.

Ethics and Legal Compliance

How will you manage any ethical issues?

The experimental protocol that will be used in this project will be submitted to (and must have the approval of) the Ethical Committee of Animal Experimentation of the Federal University of São Paulo-SP, Brazil. All animals will have free access to food and water ad libitum during the assays. The use of human plasma samples already have the approval of the institutional ethical commitee (CEP/UNIFESP n: 1477/2016)

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

There will be no restrictions on the reuse of third-party data.

Storage and Backup

How will the data be stored and backed up during the research?

In addition to the public repositories mentioned above, data will be stored locally in a server located at the Functional Proteomics Laboratory, at the Institute of Science and Technology of the Federal University of São Paulo, at São José dos Campos, Brazil. A backup will be done on a weekly basis, using external hard drives.

How will you manage access and security?

Data in public repositories are available to the general public. Ideally, our local storage will mirror the public repositories.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

All data will be retained and used in throughout the development of this and other research projects at the laboratory.

What is the long-term preservation plan for the dataset?

At this time (June 2021), there is no budget limitation for time and effort to preserve data, as we have plenty of space on our server. However, as the project starts and data will come, it will be necessary a minimum of financial support for ordinary maintenance. However, as mentioned before, all data will be made public through public repositories of 'omics' data.

Data Sharing

How will you share the data?

All data will be made public through public repositories of 'omics' data.

Are any restrictions on data sharing required?

No

Responsibilities and Resources

Who will be responsible for data management?

As of today, all data handling in our server is managed by graduate students and myself. Additionally, IT personnel from the University are available for assisting in any issue.

What resources will you require to deliver your plan?

In fact, financial support is of paramount importance for the ordinary activities of data handling and hardware updating. The IT personnel of the University is constantly assisting us in any issue that may occur with our server or with data handling.