Plan Overview

A Data Management Plan created using DMPonline

Title: Characterization of postbiotics derived from fermented dairy products: potential application for digestive health

Creator: Laura Grasa

Principal Investigator: Laura Grasa , Mª Lourdes Sánchez

Data Manager: Laura Grasa , Mª Lourdes Sánchez

Affiliation: Other

Template: DCC Template

ORCID iD: 0000-0002-5306-9365

ORCID iD: 0000-0001-5964-823X

ID: 89532

Start date: 01-09-2022

End date: 01-09-2025

Last modified: 08-02-2022

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Characterization of postbiotics derived from fermented dairy products: potential application for digestive health

Data Collection

What data will you collect or create?

Most of the experimental data will be captured in the course of instrumental analysis with laboratory equipment (liquid chromatography/mass spectrometry, western blotting, spectrofotometry, StepOne Plus Real-Time PCR System, organ bath capturing electrophysiological signals, microbiota sequencing with Illumina, etc). Observational data will also be obtained (optical and fluorescence microscopy observations, colitis assessment). Raw data will originate from the measures collected by the researchers during the execution of the working plan of the project. Processed data will be obtained from the analysis of the data with different softwares and statistical comparisons between each condition or group of study, as described in the work plan of the project.
In our project, we will obtain the following data:

Description	Type	Format	Volume
Data of samples	Metadata	-Spreedsheets in comma-separated values files (.csv), tab-delimited file (.tab) and MS Excel (.xlsx)	Small (MB)
Quantitative experimental measurements	Raw data	-Spreedsheets in comma-separated values files (.csv), tab-delimited file (.tab) and MS Excel (.xlsx) -Sequencing data in fasta files (.fasta or .fastq) -Graphics in PDF (.pdf) or Tagged Image File Format (.tiff)	Medium (GB)
Observational data	Raw data	-Figures in JPEG (.jpg) or Tagged Image File Format (.tiff)	Medium (GB)
Analysis of the data and statistical comparisons	Processed data	-Spreedsheets in comma-separated values files (.csv), tab-delimited file (.tab) and MS Excel (.xlsx) -Tables in MS Excel (.xlsx) and MS Word (.docx) -Graphics in GraphPad (.pzf), TIFF (.tiff), PDF (.pdf) or EPS (.eps). -Figures or images in JPEG (.jpg), TIFF (.tiff), PDF (.pdf) or EPS (.eps). -Text in PDF (.pdf) or MS Word (.docx)	Medium (GB)
Software tools	Sofware program code in R, bash or phyton	-Scripts in R code (.R), bash (.sh) or phyton (.py)	Small (MB)
Lab reports containing objectives, hypothesis, protocols, partial results, conclusions of the experiments	Documentation	-Text in PDF (.pdf) or MS Word (.docx)	Small (MB)
Electronic copies of conference presentations and published articles	Documentation	-Text in PDF (.pdf) or MS Word (.docx)	Small (MB)
Doctoral Thesis, Master Dissertations and Undergraduate Dissertations	Documentation	-Text in PDF (.pdf) or MS Word (.docx)	Medium (GB)

How will the data be collected or created?

Quality assurance

We will ensure the good quality of our data by the following actions:
-Researchers of the project will maintain all the devices calibrated to check the precision, bias and/or scale of measurement.
-Researchers of the project will take multiple measurements, observations or samples in our experiments.
-Researchers of the project will use standardized methods and protocols for getting the experimental measurements.
-Data will be checked by others (4 eyes principle).
-Raw data will be fully up to date and stored in the Electronic Lab Notebook LabArchives protected by a password
-We will apply a policy for handling missing data and outliers.

Organising files and folders

Well-organized file names and folder structures make it easier to find and keep track of data files. In our project, the file names will be meaningful and brief, avoiding the use of spaces and special characters.

A root folder with the name of the project will be created and additional folders will be created inside this folder. Files will be organized in folders according to the objectives of the project. For each objective of the project we will have a folder containing several folders with the tasks of the objective. In each task folder, other folders containing the sample metadata, raw data, processed data and protocols will be organized. Files related with conference papers and presentations, articles for journals, Doctoral Thesis and Undergraduate and Master dissertations will be organized in folders inside the root folder of the project.

- Proyect
  - Objective 1
    - Task 1
      - Sample_metadata
      - Raw_data
      - Processed_data
      - Protocols
    - Task 2
    - ....
  - Objective 2
  - ....
  - Conferences
  - Papers_journals
  - Doctoral_Thesis
  - Undergraduate_Dissertations
  - Master_Dissertations

Version control and authenticity

As the files generated during the project will be used by multiple users and in multiple locations (PC, laptop, cloud), we will keep a track of master versions of files.

The version of a file will be identified by a version number in file name (v1, v2, v3) and the date of the last modification. We will keep a version control table, where versions, dates, authors and details of changes to the file are recorded.

VERSION CONTROL TABLE FOR A DATA FILE			
Title:	Results of TJ gene expression by qPCR		
File Name:	TJ_PCR_v3		
Created by:	Name Surname		
Maintained by:	Name Surname		
Created:	Andrea Bellés		
Last Modified:	Laura Grasa		
Version	**Responsible**	**Notes**	**Last amended**
V3	Laura Grasa	Entries checked by LG	01/06/2021
V2	Andrea Bellés	Test results 11-20 entered	01/05/2021
V1	Andrea Bellés	Test results 1-10 entered	01/03/2021

We will keep in PC, a hard disk and Google Drive the three last versions of the file. Archive copies of master files will be performed every week. The responsibility of keeping the version control of the files will be the researchers of the project responsible of the different tasks of the project.

Documentation and Metadata

What documentation and metadata will accompany the data?

Metadata of samples, processed data, pre-print version of articles published in Subscription-based Journals and conference presentations will be deposited in the repository ZENODO. The deposit of these data in ZENODO ensures that the data will be FAIR:

To be Findable:

-In ZENODO, the data are assigned a globally unique and persistent identifier. A DOI is issued to every published record on Zenodo.

-Data are described with rich metadata. Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichements.

-Metadata clearly and explicitly include the identifier of the data it describes. The DOI is a top-level and a mandatory field in the metadata of each record.

-Metadata and data are registered or indexed in a searchable resource. Metadata of each record is indexed and searchable directly in Zenodo's search engine immediately after publishing. Metadata of each record is sent to DataCite servers during DOI registration and indexed there.

To be Accessible:

-Metadata and data are retrievable by their identifier using a standardized communications protocol. Metadata for records are harvestable using the OAI-PMH protocol by the record identifier. Metadata is also retrievable through the public REST API.OAI-PMH and REST are open, free and univesal protocols for information retrieval on the web.

Metadata are publicly accessible and licensed under public domain. No authorization is ever necessary to retrieve it.

-Metadata are accessible, even when the data are no longer available. Data and metadata will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least. Metadata are stored in high-availability database servers at CERN, which are separate to the data itself.

To be Interoperable:

-Metadata and data use a formal, accessible, shared, and broadly applicable language for knowledge representation. Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML.

-Metadata and data use vocabularies that follow FAIR principles. For certain terms we refer to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE).

-Metadata and data include qualified references to other (meta)data. Each referrenced external piece of metadata is qualified by a resolvable URL.

To be Reusable:

-Metadata and data are richly described with a plurality of accurate and relevant attributes. Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments.

-Metadata and data are released with a clear and accessible data usage license. License is one of the mandatory terms in Zenodo's metadata, and is referring to an Open Definition license. Data downloaded by the users is subject to the license specified in the metadata by the uploader.

For accessing the data, standard publicly available software will be required uniquely. Specifically, the data will be accessible using the following software: MS Office, Open Office, Adobe Reader and Image Viewer, which are widely available as commercial products or as freeware.

Ethics and Legal Compliance

How will you manage any ethical issues?

Not Applicable.

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

The principal investigators for this work are responsible for creating the data and will own the copyright.

Part of our data will be published in subscription-based, non-open access journals that require authors to transfer the copyright of their article to the publisher. However, the authors will retain the right to deposit a pre-print version of the article in the institutional repository of the University of Zaragoza, ZAGUAN, in order to be able to share it publicly.

Another part of the data will be published in open-access journals, in which the copyright of the data generally remains with the authors. This means we will be able to share always and publicly our open access publications. Creative Commons (CC) licenses are generally used for Open Access articles. The Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/), permit unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

The original datasets will be deposited in ZENODO, a public repository in which the data can be shared under Creative Commons licenses (CC-BY license).

Storage and Backup

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

The raw data and lab reports will be stored on the Electronic Lab Notebook LabArchives and in PC, and back-ups will be performed every week in a hard disk and Google Drive. Raw data will not be published nor openly accessible.

Sample metadata and processed data will be stored in PC and Google Drive and back-ups will be performed every week in a hard disk. Datasets of the experiments performed during the project used to create the figures and plots in scientific publications will be openly accessible in ZENODO repository. In addition, the sample metadata and fasta files obtained from microbiota sequencing will be deposited in SRA database at NCBI.

Documentation of conference papers and presentations, published papers, doctoral thesis, undergraduate and Master´s dissertations will be stored in PC and Google Drive and back-ups will be performed every week in a hard disk.

Doctoral thesis, undergraduate and Master´s dissertations will be deposited in ZAGUAN, the institutional repository of the University of Zaragoza, which is part of RECOLECTA, an open science platform that brings together all the national scientific repositories in Spain.

Upon completion of the research objectives of the project, the results will be published in Open Access Journals and in Subscription-based Journals. In both cases, the articles will be identified with the Digital Object Identifier System (DOI). In the cases of articles published in Subscription-based Journals, a pre-print version of the article will be deposited in the repository ZAGUAN.

We won´t need to include any charges for storing the data.

How will you manage access and security?

The access to rooms where data, computer and hard disks are held, will be controlled by locking the doors with a key, that only the researchers of the project have.

We will keep firewall protection and security-related upgrades and patches to operating systems of the computers containing the research data to avoid viruses and malicious code. Computer systems will be locked with a password.

Selection and Preservation

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

Destruction of our data will not be necessary for ethical reasons, as we do not have any data related to personal data.

We will keep the processed experimental data obtained during the development of the project and we will share them in freely accessible repositories, so that they can be used by other researchers in other studies. The data will be deposited in the repositories indefinitely.

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

The articles, conference papers, datasets, undergraduate and Master's dissertations will be kept on the servers of the Open Access journals and in the ZAGUAN (RECOLECTA) and ZENODO repositories, whose existence is guaranteed indefinitely, or at least for many years. The cost of depositing on these servers is currently free.

Data Sharing

How will you share the data?

As mentioned above, the data will be shared in ZENODO and ZAGUAN repositories with free access and subject to Creative Commons licenses, so that they can be used by other researchers in other studies, as long as they cite the reference (DOI) and authors of the data.

Are any restrictions on data sharing required?

Data of the repositories can be used by other researchers, as long as they cite the reference (DOI) and authors of the data.

Responsibilities and Resources

Who will be responsible for data management?

The principal investigators of the project will be in charge of keeping the DMP revised and updated.
Each researcher of the project responsible for each objective and its tasks will be in charge of the data capture, metadata production, data quality, storage and backup.
The principal investigator of the project will be responsible for data archiving and data sharing in the repositories mentioned above.

What resources will you require to deliver your plan?

Project researchers will be informed of the existence of the DMP, as well as its content, and will be trained to carry out all the practices described in the DMP.
To carry out the DMP, in addition to the resources that we already have: computers, hard drives, access to servers in the cloud, we will need to acquire the LabArchives Electronic Laboratory Notebook software.
The data will be deposited in public (ZENODO) or institutional (ZAGUAN) repositories that are free.