Data Sharing in Plant Space Biology: The Role of Metadata as a Form of Experimental Control

Paola Castaño and Sabina Leonelli
Egenis Centre for the Study of the Life Sciences, University of Exeter

Background

- PHIL_OS Project (Philosophy of Open Science for Diverse Research Environments): A multi-year international study of open science practices in biology.
- Inquiry into open science beyond easy slogans:
  1. It is not a one-size-fits-all endeavour.
  2. It has to adapt to different settings, methods, and questions.
  3. It is a labour-intensive process that requires infrastructures and systems of recognition.
  4. Open science policies must consider resource disparities between individuals and institutions.

Key Questions

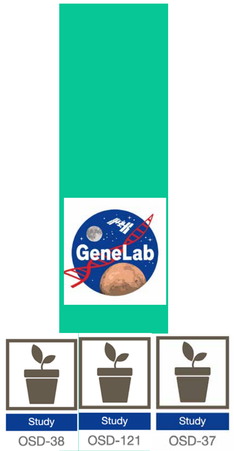
- What is the impact of open science in plant space biology?
- What is the relationship between data re-analysis and experimental practices in this field?
- How does the work of NASA's GeneLab Analysis Working Groups enable new insights from comparing datasets and probing experimental designs?

Methods

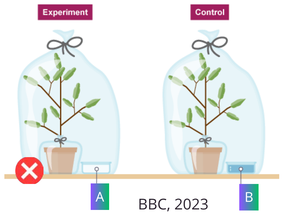
- Philosophical and sociological analysis of data sharing and interpretation.
- Participant observation in the Plants Analysis Working Group, interviews with members of the group, and textual analysis of their publications (2022-2023).
- Case study: "Meta-analysis of the space flight and microgravity response of the Arabidopsis plant transcriptome" (Barker et al. 2023).

Case Study

Studying Plant Stress in Space: An Entangled Data-Collection Enterprise



Azuma Makoto, 2014

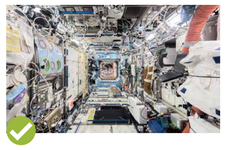


BBC, 2023

For plant studies, spaceflight is not a natural environment nor like the administration of a single treatment in a laboratory setting.

NASA, 2015

NASA, 2014



Unique and multiple environmental factors interact simultaneously with the organisms

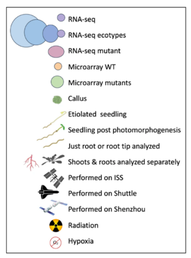
+

Combination of operational constraints, hardware, and diverse experimental designs

=

Datasets resulting from these experiments are difficult to compare directly.

Metadata Curation and Comparative Data Re-Analysis: a Disentanglement Exercise



Barker et al. 2023

1. Provides a detailed picture of the many intervening factors in experiments.
2. Identifies confounding variables "imposing effects on patterns of gene expression over and above those of the spaceflight treatment" (Barker et al. 2023, 3).
3. Narrows down the frame of comparison keeping some conditions constant (e.g. analysis methodology and hardware) for more targeted analyses that yield new results.
4. Points to future experimental disentanglements of dependent and independent variables (e.g. lighting as independent variable within a type of hardware.

Findings

- The work of metadata curation is an extension of experimental practice: it breaks down 'space' into manageable components for data analysis.
- Just like controls in an experiment, metadata curation treats some factors as targets of analysis and stabilizes others as background conditions.
- This work is providing novel ways to compare, interpret, and potentially integrate results obtained in spaceflight experiments.
- Why does this matter? Metadata practices play a key role in discovery:
  1. Retrospective reconstruction: They enable new biological insights from data already generated.
  2. Prospective building: They can guide future experimental designs.

References

Barker, Richard, et al. 2023. "Meta-Analysis of the Space Flight and Microgravity Response of the Arabidopsis Plant Transcriptome." *Npj Microgravity*, vol. 9, no. 1.
Castaño, Paola. 2023. "Plant Biologists and the International Space Station: Institutionalising a Scientific Community." In *Routledge Handbook of Social Studies of Outer Space,* 423–36, .
Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study*. The University of Chicago Press.
----------. 2023. *Philosophy of Open Science*. Cambridge University Press.
Schickore, Jutta. 2019. The Structure and Function of Experimental Control in the Life Sciences. *Philosophy of Science, 86*(2), 203-218.

For more resources and publications from this project, visit: https://opensciencestudies.eu

Acknowledgments

The PHIL_OS project is funded by the European Research Council (ERC) Horizon 2020 Research and Innovation programme [Grant Agreement 101001145].

Thanks to the NASA GeneLab management team, all members of the Plants AWG, and Richard Barker for their time and engagement with this project.