

RECONSTRUCTING EVOLUTION OF THE PLANT ENERGY SENSING SYSTEM

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		<i>Indicate: N(ew data) or E(xisting data)</i>	<i>Indicate: D(igital) or P(hysical)</i>	<i>Indicate: Audiovisual Images Sound Numerical Textual Model Software Other (specify)</i>		<i>Indicate: <1GB <100GB <1TB <5TB >5TB NA</i>	
LUC/GUS	reporter assays	N	D	I/N	.xlsx	<1GB	-
qRT-PCR	quantitative real-time PCR	N	D	I/N	.xlsx	<1GB	-
WB	Immunoblots	N	D	I	.tiff/jpg	<100GB	-
Phenotype	Images of phenotypes	N	D	I	.tiff/jpg	<100GB	-
Sequences	DNA and protein sequences	N/E	D	N/T	diverse (.fastq/csv/sam/bam/vcf/txt/py/sch)	>5TB	-
Plasmids	DNA constructs	N/E	P	-	-	-	Freezer space
Plants	Seeds, cells and spores of mutant and transgenic plants	N/E	P	-	-	-	Freezer and fridge space
Scripts	Bioinformatics pipeline	N/E	D	N/T	diverse	<100GB	-

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Sequence data from (and to) repositories, such as GenBank, ENA, TreeBASE.

Mutant plants from (and to) stock centers, such as NASC and the Chlamydomonas Resource Center.

Scripts from (and to) Github.

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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.

- No

This is very fundamental biological research into the evolution of plants.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

In addition to the physical notebooks and individual storage on OneDrive, we will create a Teams channel, where all data related to the project will be shared with all collaborators (2 promoters, 2 PhD students), in the formats listed above.

Will a metadata standard be used to make it easier to find and reuse the data?

If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Other (specify below)

. OneDrive

. Teams

. Repositories, as listed above.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution
- Personal back-ups I make (specify below)

All data are also regularly transferred to an external hard drive.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Secured by KU Leuven ICTS.

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

Difficult to estimate (the amount of sequencing data) at the moment.

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- Large Volume Storage (longterm for large volumes)
- Other (specify below)

. OneDrive

. Teams

. Repositories, as listed above.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

Difficult to estimate (the amount of sequencing data) at the moment. We do not expect excessive costs that cannot be covered by this or follow-up projects.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?

Please explain per dataset or data type which data will be made available.

- Yes, as open data

Upon publication, all data will be made available to the research community.

If access is restricted, please specify who will be able to access the data and under what conditions.

NA

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other data repository (specify below)

Repositories specified above.

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

Not clear at this point.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

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

Publications costs (depending on the journals, not clear at this point).

Responsibilities

Who will manage data documentation and metadata during the research project?

Promotors Filip Rolland and Steven Janssens.

Who will manage data storage and backup during the research project?

Promotors Filip Rolland and Steven Janssens.

Who will manage data preservation and sharing?

Promotors Filip Rolland and Steven Janssens.

Who will update and implement this DMP?

Promotors Filip Rolland and Steven Janssens.