ETHYLUTION

A Data Management Plan created using DMPonline.be

Creator: n.n. n.n.

Affiliation: KU Leuven (KUL)

Funder: European Research Council (ERC)

Template: ERC DMP +

Principal Investigator: n.n. n.n.

Grant number / URL: 10.3030/101087134

ID: 208256

Start date: 01-09-2023

End date: 31-08-2028

Project abstract:

Ethylene is a plant hormone that influences plant development and response to stress. However, it is unclear when plants started producing it. Funded by the European Research Council, the ETHYLUTION project is investigating the hypothesis that ethylene emerged when ancient plants transitioned from water to land habitats. Researchers will use model plant species representing early life on earth, and investigate the ethylene biosynthesis pathway and the diverse signalling pathways it triggers. Results will unveil important information on the role and functions of ethylene signalling during the evolutionary history of plants.

Last modified: 18-06-2024

GDPR record

Have you registered personal data processing activities for this project?

- No

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- No

Created using DMPonline.be. Last modified 18 June 2024

3 of 6

ETHYLUTION
ERC DMP +

Project information

Project Acronym

ETHYLUTION

Project Number

101087134

Data summary

Summary

Work Package	Task	Data type	File type	Data size	Methods/equipment/source
WP1	1.1	Numerical (chromatograms, mass spectra, peak areas)	mzDATA or similar; .xls	<1 GB	LC-MS and HPLC instruments
	1.2	Biological samples	/	/	Plant material (gemmae/spores)
		Numerical (chromatograms, peak areas)	.xls	<1 GB	GC instrument
	1.3	Biological samples	/	/	Plant material (gemmae/spores)
		Numerical (chromatograms, peak areas)	.xls	<1 GB	GC instrument
	1.4	Numerical (sequences, gene expression)	.fasta and .xls	<1 GB	Sequencing, qPCR
		Images (microscopy; blots; gels)	.tiff or .jpeg	1-10 GB	Images
	1.5	Numerical (chromatograms, mass spectra, peak areas)	mzDATA or similar; .xls	<1GB	LC-MS and GC-MS instrument
WP2	2.1	Numerical (sequences)	.fasta	10-100 GB	Sequencing
	2.2	Biological samples	/	/	Plant material (gemmae/spores)
		Numerical (chromatograms, peak areas, mass spectra)	.xls; mzDATA or similar	<1 GB	GC and GC-MS instrument
		Numerical (sequences, gene expression)	.fasta and .xls	<1 GB	Sequencing, qPCR
		Images (microscopy; blots; gels)	.tiff or .jpeg	1-10 GB	Images
	2.3	Numerical (gene expression)	.fasta and .R	10-50 GB	Sequencing
WP3	3.1	Numerical (phylogenetic analysis)	.xml and .meg	1-5 GB	Processed sequence data in Geneious
		Images	.tiff or .jpeg	<1 GB	Phylogenetic trees
	3.2	Biological samples	/	/	Plant material (seeds)
		Numerical (chromatograms, peak areas, mass spectra)	.xls; mzDATA or similar	<1 GB	GC and GC-MS instrument
		Numerical (sequences, gene expression)	.fasta and .xls	<1 GB	Sequencing, qPCR
		Images (microscopy; blots; gels)	.tiff or .jpeg	1-10 GB	Images
	3.3	Biological samples	/	/	Plant material (algal cultures)
		Numerical (chromatograms, peak areas, mass spectra)	.xls; mzDATA or similar	<1 GB	GC and GC-MS instrument
		Numerical (sequences, gene expression)	.fasta and .xls	<1 GB	Sequencing, qPCR
		Images (microscopy; blots; gels)	.tiff or .jpeg	1-10 GB	Images
Overarching	all	Numerical (data analysis and statistics)	.R or .xls	<1 GB	Statistical analyses
	all	Text (writing and presenting data/papers/presentations)	.doc; .pdf; .ppt	<1 GB	Text writing and presentations

FAIR data

1. Making data findable

- Folders with raw data are stored per user in folders that have dates and information on the experiments.
- Filenames and raw data-files have a consistent naming.
- Metadata for experiments and obtained data will be developed in separate txt, xls, doc or ppt file. These files will contain sample information, experimental conditions, instrument settings, and data-processing steps.
- Metadata from large-scale data-sets from omics experiments (whole genome resequencing; RNA-sequencing) will be deposited alongside the raw and processed data in online available data repositories containing a doi number.
- Plant material and biological resources (seeds, gemmae, spores) will be deposited in the lab's seed stocks which is curated by an accompanying excell that contains all the metadata information (species, source, date generated, genetic perturbations, user name, location, resistance/selection marker).

2. Making data openly accessible

Processed and polished data will be presented in publications. Raw data will be made available to third parties upon a reasonable request.

Large-scale datasets (e.g. .xls) and its metadata will be added as supplemental files in published papers.

Large-scale datasets obtained from genome and transcriptome sequencing (or other omics technologies) and its metadata will be deposited in an online repository database (e.g. NCBI Sequence Read Archive) with a doi number.

Obtained plant material (biological resources such as seeds, spores, gemmae) will be made available to the community. Seeds of Arabidopsis will be deposited in the international NASC (Nottingham Arabidopsis Stock Center). Plant material of other species will be kept in the lab (due to the absences of official stock centers) but made available upon request. Exchange of plant material is conform the Nagoya protocol and the necessary phytosanitary and biosafety regulations.

3. Making data interoperable

Most data will be processed in excell or R. Given that these are two very common software used in life-sciences, data files will be stored in these files (.xls and .R).

Large-scale mass-spec data will be stored in its original format (.mzDATA) that can be operated with specific software. But the analyzed and processed data will be stored in .xls and .R files.

Large-scale sequencing data will be stored in its original format (.fasta) that can be operated with specific software. But the analyzed and processed data will be stored in .xls and .R files.

Phylogenetic analyses will be stored in its original format (.fasta) that can be operated with specific software. But the analyzed and processed data will be stored in .xls and .R files or final output images in .tiff or .jpeg formats.

4. Increase data re-use

All data and biological resources generated in the ETHYLUTION project are available to the community without restrictions (open access). No embargo is foreseen.

Data quality is assured by the high-level academic standards of the PI and it's staff and according the the plant science community-accepted methodologies.

5. Allocation of resources and data security

Budget to publish findings (processed and analyzed data) and its accompanying original data-sets (e.g. supplementary material to papers and large-scale data-sets deposited in repositories) is foreseen in the grant and will be used to ensure all work is publicly available under the BY-CC license.

Digital storage of the data during the project is done on the labs sharepoint, curated, backupped and maintained by the central IT services of KU Leuven. We have 100 TB of data storage capacity, but this volume can be increased if needed. Extra volume can be purchased via the central IT services.

After the project is finished, all the data will be transferred to the labs digital archive, a long-term storage drive (>10 years) curated, backupped and maintained by the central IT services of KU Leuven. This data-transfer is overseen by a central IT person and the PI. Fees for the archive

drive are carried by the lab, but are substantially lower than for the working sharepoint.

Biological resources are stored in specific conditions and are curated and maintained by the lab technician under the supervision of the PI. Expenses needed for storage and maintenance are part of the daily lab budget given it's prime importance for a plant-science lab. The lab technician is 60% paid by the Department, and is hence available for maintenance tasks of the plant collections.

- Marchantia spores are stored at 4 C in the cold-room in labeled eppendorf tubes. These spores are kept dry and can be stored for 5-15 years.
- Marchantia gemmae are stored at 4C in the cold-room in labeled eppendorf tubes on Gambourg B5 medium supplemented with sucrose. This collection is transferred to new medium every 6-12 months.
- Arabidopsis seeds are stored in a dry condition in a dedicated seed cabinet at room temperature containing silica-gel (replaced every 6 months to prevent moist build up). Seeds are stored in individual eppendorf tubes. Seeds can be stored up to 15 years. Additionally, seeds of generated lines will be deposited to the International stock center (NASC).
- Tomato seeds are stored in a dry condition in a dedicated seed cabinet at 4 C in the cold-room, containing silica-gel (replaced every 6 months to prevent moist build up). Seeds are stored in individual eppendorf tubes.
- Algal cultures are stored in the lab at 4 C in the cold-room in culture medium. Cultures can be stored up to 4-6 months and will be sub-cultured regularly. Additionally, cultures are available from the International stock center (UTEX algal culture collection of the University of Texas-Austin).