

➤ Long Term Plant Phenomic Data Sharing in Generic Data Repositories (Zenodo, Dataverse, e!Dale) Using MIAPPE

Cyril Pommier, Erwan le Floch, Daniel Arend, Flora D'Anna

<https://tess.elixir-europe.org/materials/workshop-on-resources-for-plant-sciences-2023>



EMPHASIS



INRAE

www.elixir-europe.org

<https://emphasis.plant-phenotyping.eu>



- Plant Phenomics data challenge and data repositories
- Guidelines portals
- Dataverse data submission
- Plant data standard and their usage in Dataverse
- Impact in data portals

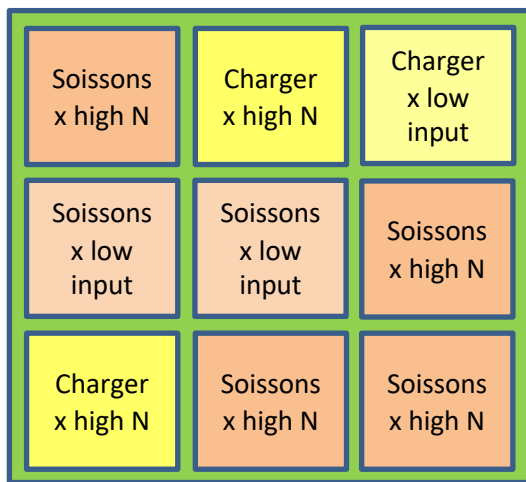


Introduction : Plant Phenomics data challenge and data repositories

Phenotype data lifecycle

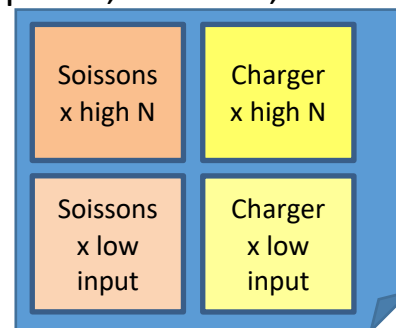


« Raw » data, pheno/env measurement, variables

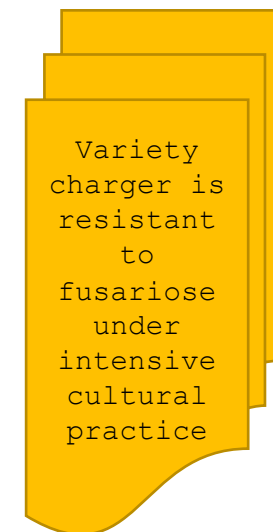


Derivation, Reduction

« Derived » data, computed, reduced, indicators



Publication



Genotype	Treatment	N input	Date	Rep	Fusariose
Soissons	low input	15,3225129	15/11/2011	1	5
Soissons	low input	15,3430556	16/11/2011	2	7

Genotype	Treatment	Fusariose
Soissons	low input	6

VARIABLES

- Raw Measures Pheno & Env
- Data cleaning
- Traceability, Reproducibility & Provenance

INDICATORS

- New computation for each scientific question
- One experimental dataset → many derived dataset

Data repositories

Thematic repositories

Opened



Access Control



Generalist repositories



Publisher repository



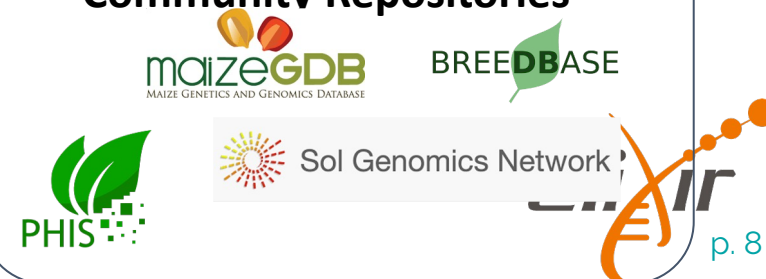
Institutes Repositories + partners



National Repositories + partners



Community Repositories



Data repositories

Thematic data repositories

- **Pros**
 - Precise dataset description and metadata
 - Good quality
- **Cons**
 - Not available for all data types (e.g. phenotyping)
 - Time-consuming publication/curation of data

Generalist repositories

- **Pros**
 - Any data type \Rightarrow flexible
 - Minimal publication metadata
 - author, title, description, DOI
 - Rapid submission/publication
- **Cons**
 - Weak description (biological material, methods, etc.)
 - Non-standard formats
 - FAIR reduced (less Findability, Interoperability, Reusability)
 - Need for Guidance and optional validators



Knowledge Hub introduction : RDMKit & FAIRCookbook

Plant Pheno publication Guidelines

Ecosystem of knowledge documentation portals

- **RDMkit**: entry point, with links to registries & recipes, provides best practices & guidelines to make your data FAIR
- **FAIR Cookbook**: detailed recipes repository, that help to make & keep data FAIR



<https://rdmkit.elixir-europe.org/>



<https://faircookbook.elixir-europe.org>

RDMkit - Best practices & guidelines to make your data FAIR

- A “wikipedia-like” knowledge base website, free and open
- Describes how to manage research outputs according to FAIR principles
- Portal to other online resources used by RDM professionals and researchers

→ it is no peer-review journal, repository, registry or manual, but it provides you information & links

- Entry to the meaningful piece of documentation for a precise task.
- Assembly of related but dispersed documentation (e.g. [readthedoc.io/isa/chapter 1](https://readthedoc.io/isa/chapter1) then readthedoc.io/Biosamples/chapter5)

The screenshot shows the RDMkit website header with the logo and navigation links: Data management, About, Contribute, and GitHub. The main content area features the title "The Research Data Management toolkit for Life Sciences" and a subtitle "Best practices and guidelines to help you make your data FAIR (Findable, Accessible, Interoperable and Reusable)". A URL is provided: <https://rdmkit.elixir-europe.org/>

Recommended in the **Horizon Europe Program Guide** as the “resource for Data Management guidelines and good practices for the Life Sciences”

RDMkit in numbers



- Contributors are experts in RDM and/or in scientific domain in ELIXIR and beyond
- Managed by an editorial board
 - 16 ELIXIR members from several Nodes
 - lead for “FAIR Data & Resources” at the Office of Data Science Strategy at NIH
- The content is curated by members of the ELIXIR RDM Community

RDMkit: content organisation



Data life cycle

Start here to get an overview of research data management based on stages in the data life cycle.



Your role

Identify your role in research data management, find data management resources relevant for you, and information to help you progress in your career path.



Your domain

Learn about data management tasks that affect your domain or research community, and the solutions adopted to address them.



Your tasks

Find guidelines and solutions for tackling common data management tasks.



Tool assembly

Find concrete combinations of tools and resources assembled into an ecosystem for research data management.



National resources

Find pointers to country specific information resources and national research data management practices.

Data life cycle

RDMkit Data management About Contribute GitHub Search RDMkit

Data management

Data life cycle

Sharing ✎ ! ↺

What is data sharing?

Sharing data means making your data known to other people.

You can share your data with collaboration partners in the context of a collaborative research project, or you can publish your data to share it with the global research community and society at large.

On this page

- What is data sharing?
- Why is data sharing important?
- What should be considered for data sharing?
- Related pages
- More information

General definitions and considerations for each stage of the data life cycle


Your role

RDMkit Data management About Contribute GitHub Search RDMkit

Your role

In this section, information is organised based on the different roles a professional can have in research data management. You will find:

- A description of the main tasks usually handled by each role.
- A collection of research data management responsibilities for each role.
- Links to RDMkit guidelines and advice on useful information for getting started with data management specific to each role.

Search Type here... 

Data Steward	Policy maker
Data management guidance for data stewards	Data management guidance for policy makers

Role-specific tasks in frame of research data management

Your domain



Considerations

- Did you collect the metadata for the identification of your plant material according to the recommendation provided in the [above section](#)?
- Have you documented your phenotyping and environment assays (i.e. measurement or computation methodology based on the trait, method, scale triplet) both for direct measures (data collection) and computed data (after data processing or analysis)?
 - Is there an existing [Crop Ontology](#) for the species you experiment and does it describe your assay? If not, have you described your data following the trait, method, scale triplet?
- Do you have your own system to collect data and is it compliant with the [MIAPPE](#) standard?
- Are you exchanging data with individual researchers?
 - In what media is data being collected?
 - Is the data described in a [MIAPPE](#) -compliant manner?
- Are you exchanging data across different data management platforms?
 - Do these platforms implement the Breeding API [BrAPI](#) specification?
 - If not, are they MIAPPE-compliant and do they enable automated data exchange?

On this page

- Introduction
- Data management planning
- Plant biological materials: (meta)data collection and sharing
- Phenotyping: (meta)data collection and publication**
- Genotyping: (meta)data collection and publication
- Related pages
- More information
- Relevant tools and resources

Solutions

Checklists and ontologies

- The metadata standard applicable to plant phenotyping experiments is [MIAPPE](#).
 - There is a section dedicated to the identification of plant biological materials that follows [Multi-Crop Passport Descriptor \(MCPD\)](#) described [above](#).
 - There is a section to describe the phenotyping assays based on the [Crop Ontology](#) recommendations.

Data management practices in a domain

- what aspects should be taken into account (considerations)
- available solutions
- links to tools and resources explained in clear context (when to use them and for what purpose)

Tool assembly

RDMkit
Data management
About
Contribute
[GitHub](#)

Data management

Data life cycle ▾

Your role ▾

Your domain ▾

Your tasks ▾

Tool assembly ▴

COVID-19 Data Portal

CSC

Galaxy

IFB

Marine Metagenomics

MOLGENIS

NeLS

OMERO

Plant Genomics

Plant Phenomics

TransMed

Tool assembly ✎ ! ↺

Tool Assemblies are examples of combining tools to cover data management tasks across several stages of the data life cycle. These can be tools that one or several communities combine to support RDM that can be picked up or accessed and used by others. The assemblies are aimed for users in a specific location and/or for users within a specific domain.

Filter by affiliation

Choose...

✕

Search

Type here...

✕

Related pages ▾

- Your tasks
- Data sensitivity
- Existing data
- Data publication
- Data analysis
- Your domain
- Human data

COVID-19 Data Portal

The COVID-19 Data Portal brings together relevant datasets for sharing and analysis to accelerate coronavirus research.

Related pages ▾

Affiliations:

CSC

The Center of Science (CSC) provides high-quality ICT expert services for researchers in Finland and their collaborators.

Related pages ▾

Affiliations:

IFB

The French Bioinformatics Institute (IFB) offers IT infrastructure and bioinformatics expertise to support researchers in Life Sciences.

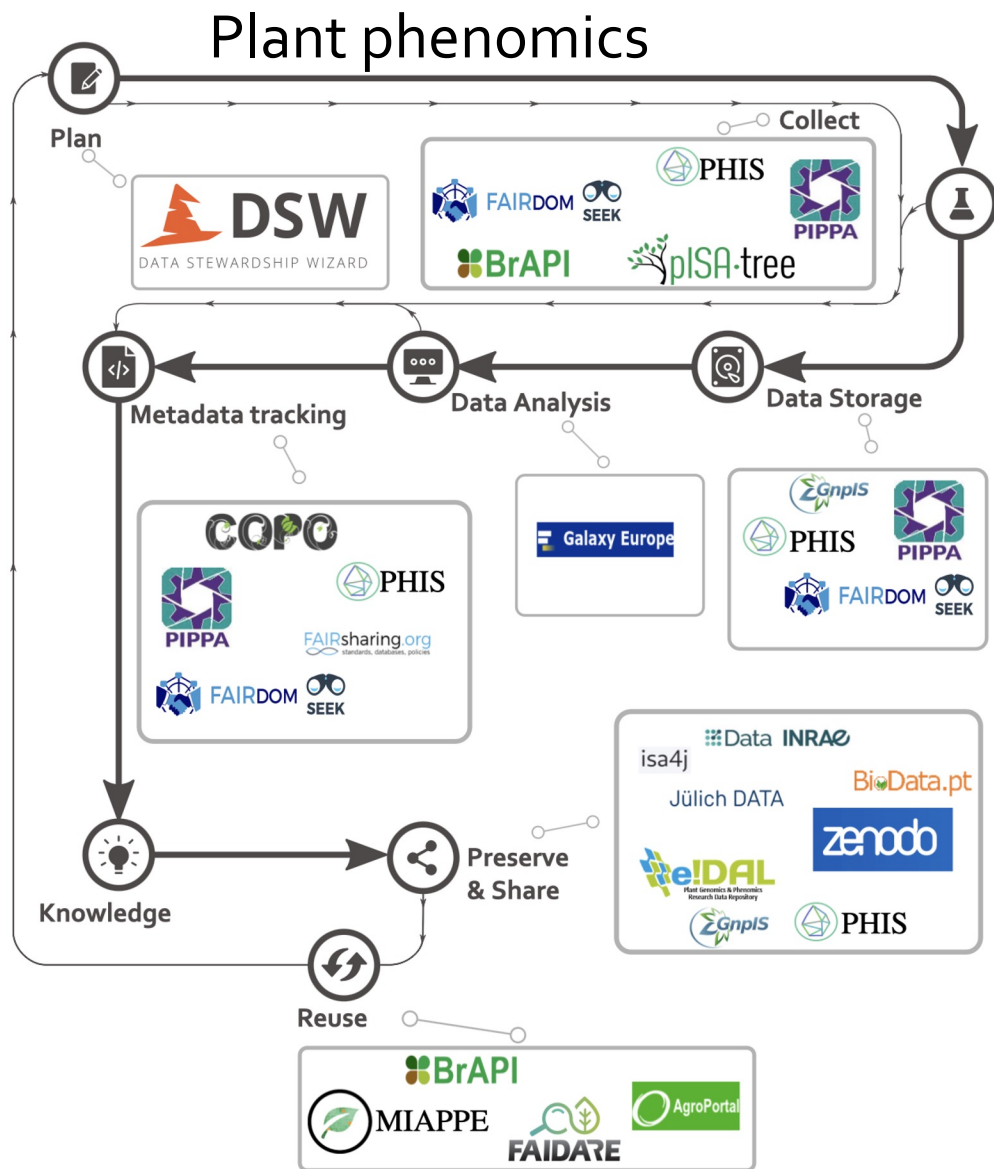
Related pages ▾

Affiliations:

2024 Jan 13 / Cyril Pommier

p. 18

Tool assemblies for plant data



do, Datavers

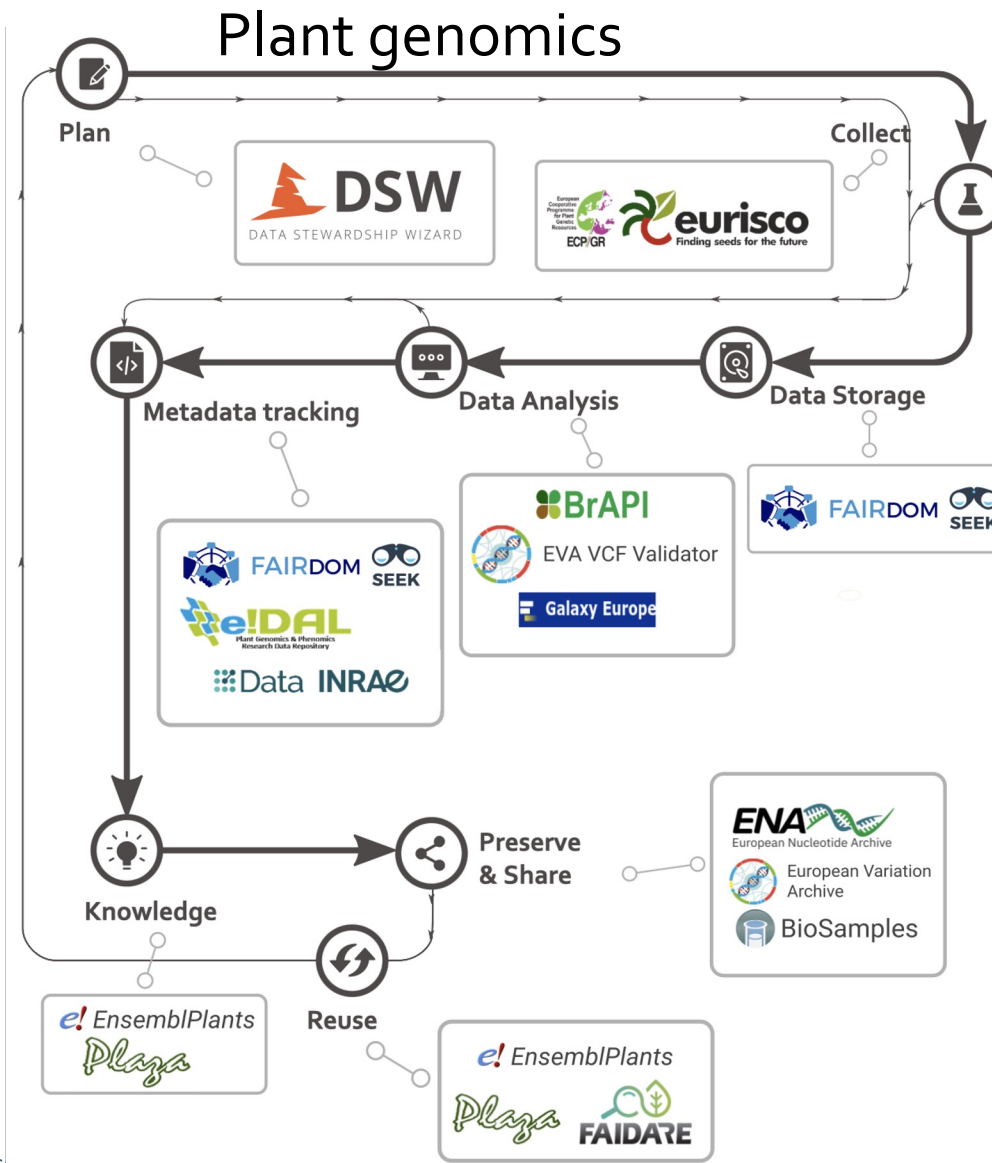
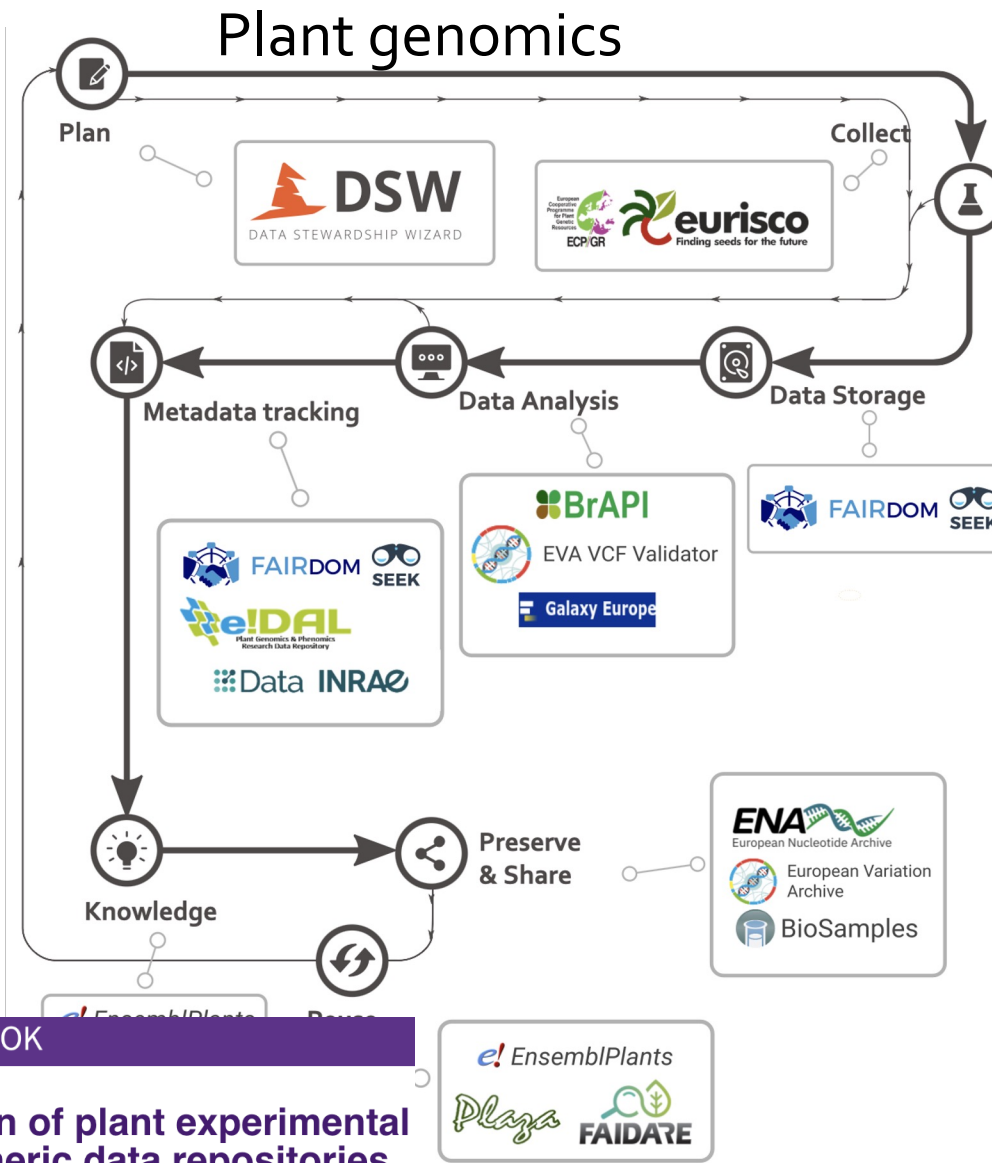
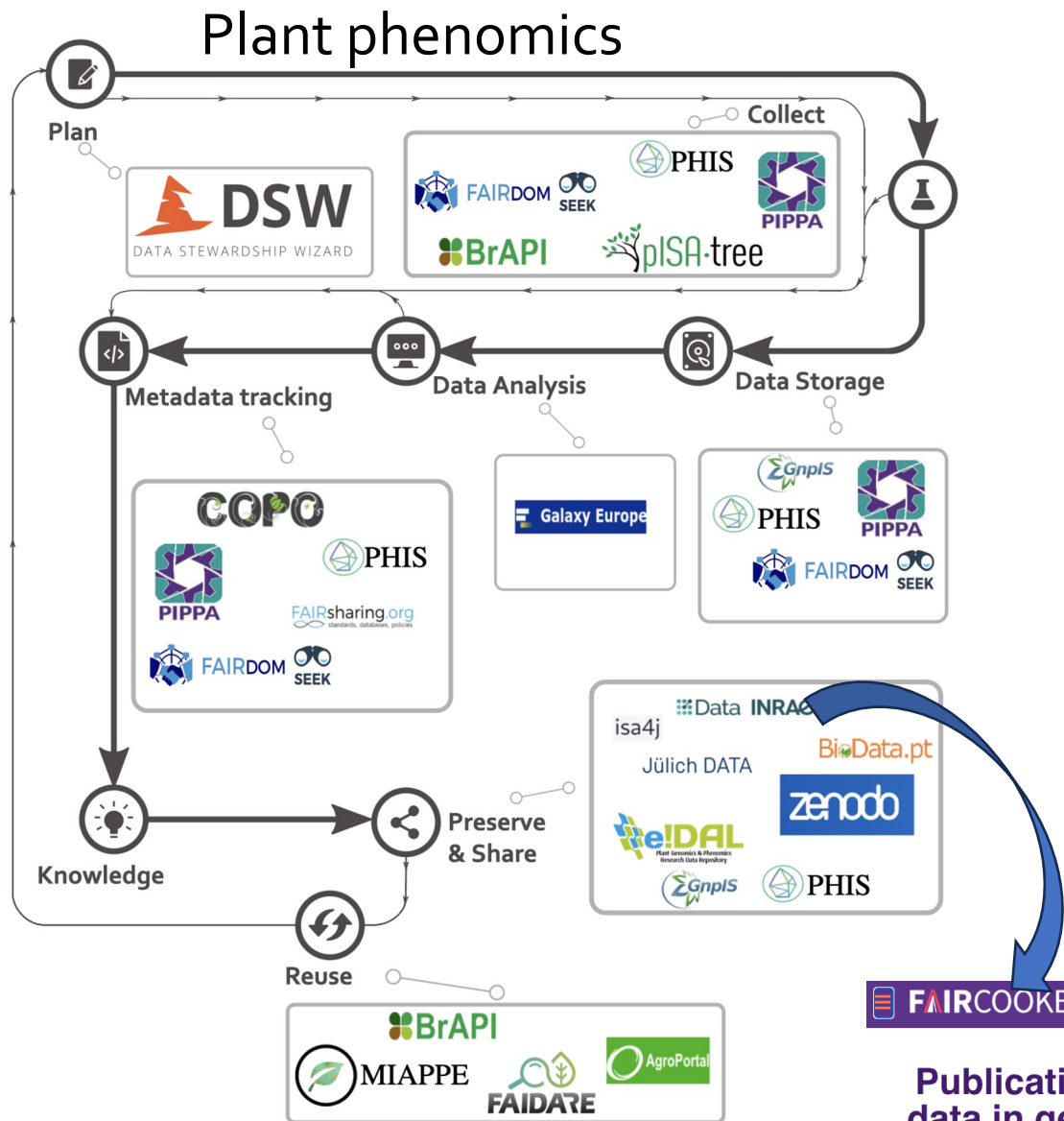


Figure 1. The plant genomics tool assembly.

Figure 1. The plant phenomics tool assembly.

Tool assemblies for plant data



FAIRCOOKBOOK

Publication of plant experimental data in generic data repositories

Figure 1. The plant phenomics tool assembly.

Figure 1. The plant genomics tool assembly.



Dataverse

<https://faircookbook.elixir-europe.org/content/recipes/reusability/plant-pheno-data-publication.html#step-by-step-process-for-data-submission-and-publication-in-dataverses>

<https://demo.dataverse.org/>

Publication of plant experimental data in generic data repositories

Recipe Overview

Reading Time 30 minutes
 Executable Code No
 Difficulty

Publishing plant phenotypic data

Recipe Type Guidance
 Audience Everyone

Maturity Level & Indicator

DSM-2-C1 DSM-2-C6

Cite me with FCB083

Contents

- Main Objective
- Summary
- Graphical overview
- FAIRification Objectives, Inputs and Outputs
- Table of Data Standards
- Table of Software and Tools
- Step-by-Step process for data submission and publication in Dataverses
- Step-by-Step process for data submission and publication in e!DAL-PGP
- Conclusion
- Authors
- Licence

Main Objective

Generic data repositories offer high level metadata, but they are not sufficient to **guarantee interoperability and the possibility to reuse your data later**. They must be completed by the Minimum Information About Plant Phenotyping Experiments (MIAPPE) international standard. The [Plant Sciences domain page](#) of the RDMkit describes how MIAPPE is needed and used to clearly document the biological material as well as the traits, environmental and phenotypical, recorded and computed. Those MIAPPE metadata can be stored using: MIAPPE full template, selected sheets from the full template (recommended since it can be stored as CSV), and ISA Tab archives. The first

Step 1: Dataset creation

Step 1: Dataset creation

Find an appropriate dataverse

You need first to select the appropriate dataverse and/or sub-dataverse for your use case depending on the constraints of your consortium or institute. The guidelines explained below are applicable to all of those dataverse instance. You will find below a non -exhaustive list of dataverse instances. Also, find the right sub-dataverse for your submission, like the one of the research group or the project you belong to: dataverses can contain dataverses.

- recherche.data.gouv.fr (FR)
 - Open to submission from any consortium involving at least one member of
 - Examples: [Data INRAE](#) and [CIRAD](#).
- [Jülich DATA](#) (NRW - DE)
 - Open to submission from any research activity done from partners at Forsch
 - Meant for data and software submissions.
 - Maintained by the central library of FZJ.
 - Examples under subject "[Agricultural Sciences](#)".
- dmpportal.biodata.pt (PT)
 - Open to submission of biological data from Portuguese research & innovati
 - Example: [Plant BioDataVerse](#).

The screenshot shows the homepage of recherche.data.gouv.fr. The main heading reads: "Recherche Data Gouv, l'écosystème au service du partage et de l'ouverture des données de recherche célèbre ses 1 an". Below this, there is a section for "Recherche Data Gouv" with a "Génération datapaper" button and a metrics bar showing "675,807 Downloads". A paragraph in French describes the platform as a sovereign solution for sharing and opening up data. Below the text is a carousel of logos for various sub-dataverses: CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement), CNRS (CNRS Research Data), INRAE (Data INRAE), and CY UNIVERSITÉ (Data repository CY Cergy Paris Université). The INRAE logo is circled in blue, and the text "sub-dataverse" is written below it.

Step 1: Dataset creation

Authenticate

Click on “Log in” in the top right corner, and choose the right authentication mechanism. Use your institutional login if it is present:

- in recherche.data
- in Jülich DATA:
- in [dportal.bio](#)

Connecting with a

Create a new dataset

Fill in the metadata

Fill in the metadata

Fig. 31 Fill in the metadata of the dataset

Save the dataset

The DOI will be automatically generated but the dataset will remain in draft mode (with the “Draft” and “Unpublish” buttons).

Manage collaborative editing and access rights

It is possible to ask partners to complete certain metadata by giving them the appropriate rights.

- Ask the partner to create an account on the Dataverse containing the dataset, and to give you its username.
- On the dataset, click on Edit > Permissions > Dataset.
- Click on the Assign Roles to Users/Groups button.
- Fill in the name of the producer and assign the role of Contributor
- Save changes

Step 3: Add generic metadata to the dataset

Step 3: Add generic metadata to the dataset

These are the metadata that are not specific to plant sciences. Select the “Metadata” tab then click on the “add + edit metadata” button

Add and edit generic metadata

Mandatory metadata:

- Language: in section “Citation Metadata” > “Language” : dataset language
- Organism: in section “life science metadata” > “Organism” : species name

Recommended metadata:

- Contributor
- Keywords: In the list of Dataverses, Datasets and Files of your Dataverse, looking at existing entries for “Keyword Term” can help you find appropriate keywords.
- Related publication
- Grant Information
- Project Information
- Time Period Covered (start and end date)
- Geospatial Metadata > Geographic Coverage

Step 4: Publish the dataset

Step 4: Publish the dataset

Private publication: temporary

Give access to an unpublished dataset (private URL): Edit dataset > Private URL

Give access to an unpublished dataset

The submitter of a dataset can generate a private URL, to give access to a dataset not yet published to a person who has no access rights on it.

Create a private URL

Unpublished dataset private URL

Publication: definitive, with unlimited updates possible

Publish the dataset:

Publish the dataset

The screenshot shows the 'Your Dataset Name' page in the INRAE data repository. At the top, there is a notification for an 'Unpublished Dataset Private URL'. Below this, the dataset title 'Your Dataset Name' is displayed with a 'Draft' and 'Unpublished' status. The author information is 'Author Name, 2023, "Your Dataset Name", https://doi.org/10.57745/FYDBT0, Recherche Data Gov, DRAFT VERSION'. There are buttons for 'Cite Dataset' and 'Learn about Data Citation Standards'. On the right side, there is a menu with options: 'Access Dataset', 'Publish Dataset', 'Edit Dataset', 'Contact Owner', and 'Share'. Below the menu, 'Dataset Metrics' are shown: 0 Views, 0 Downloads, and 0 Citations. At the bottom, there is a 'Files' section with '1 File' and an 'Upload Files' button.

Add an existing dataset to a specific dataverse

Open your dataset page and login, then choose "Link Dataset":

Add an existing dataset to a specific dataverse

Find the dataverse in which you want to add your dataset, select it and click on "Save Linked Dataset".

Save Linked dataset

Dataverse dataset example

<https://doi.org/10.15454/RGMM07>

Survival and growth of 711 forest tree taxa in eight French arboretums from three different climate regions

Version 2.2



Ducatillon, Catherine; Lamant, Thierry; Bellanger, Richard; Bouttier, Valentin; Bastien, Jean-Charles; Michotey, Celia; Gautier, Laurent; Fady, Bruno; Musch, Brigitte, 2021, "Survival and growth of 711 forest tree taxa in eight French arboretums from three different climate regions", <https://doi.org/10.15454/RGMM07>, Recherche Data Gouv, V2, UNF:6:RA5DHg7p6j/KNJJo39rydQ== [fileUNF]

Cite Dataset ▾

[Learn about Data Citation Standards.](#)

Access Dataset ▾

Link Dataset

Contact Owner

Share

Dataset Metrics ⓘ

1,434 Views ⓘ

126 Downloads ⓘ

0 Citations ⓘ

Description ⓘ

Abstract

The dataset is checklist of 711 native and exotic forest tree taxa (species, subspecies and varieties) from 177 genera worldwide, their survival and growth in eight French arboretums from three contrasted bioclimates. Planted between 1969 and 1976, 338 taxa have survived and occur as at least one individual in one plot.

Purpose

A series of eight arboretums was installed some 50 years ago for testing the ability of a wide range of mostly exotic tree species to survive in marginal, stressful habitats (high elevation, low fertility, ...)

[Read full Description \[+\]](#)

Subject ⓘ

Earth and Environmental Sciences; Medicine, Health and Life Sciences; Agricultural Sciences

Keyword ⓘ

Adaptation, Common garden, Forest tree, Ecology, Forest sciences

Related Publication ⓘ

Ducatillon, C., Lamant, T., Bellanger, R., Bouttier, V., Bastien, J.-C., Michotey, C., Gautier, L., Fady, B., & Musch, B. (2022). Survival and growth of 711 forest tree taxa in eight French arboretums from three different climate regions. *Annals of Silvicultural Research*, 47(1). <https://doi.org/10.12899/asr-2325> doi: <https://doi.org/10.12899/asr-2325>

License/Data Use Agreement

[Custom Dataset Terms](#)

More precise dataset description using MIAPPE

Generic data repositories limitations

- (Too) broad metadata
 - Year, Data type, Author
 - Organism: no controlled vocabulary
 - Keyword, Subject
- Lacking plant phenomic specific metadata
 - Biological material, from species to genetic resource accession
 - Traits
 - Experiment locations
 - ...
- Need of dedicated metadata scheme, added to generic data repositories
 - As additional metadata → Filling long list in the repository form is not an option
 - As companion files

→ MIAPPE data standard

Semantic

- ◆ Description of the data
- ◆ Controlled vocabularies: term name and definitions
- ◆ Ontologies: semantic links between terms
- ◆ *Biologist driven*



Persistent Unique Identifiers

URI, gene ID, accessions ID, Trait ID, DOI,...

Structure

- Formatting and Organizing the data
- Data Models
- Standards : CSV, VCF, GFF, MIAPPE (www.miappe.org) , etc...
- *Biologist & Computer scientist driven*



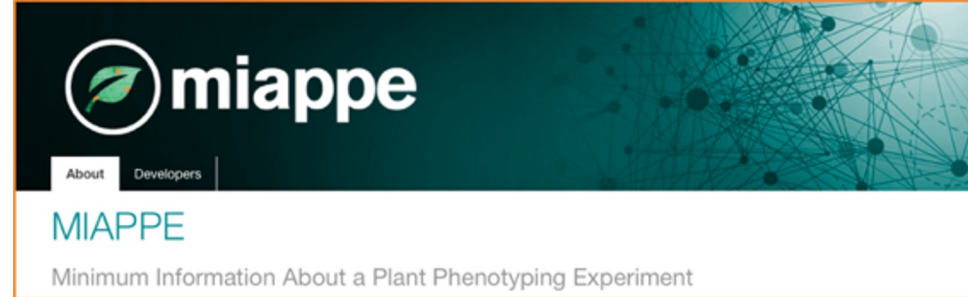
Technical

- Data integration and sharing
- Interoperability : tools and systems
 - GA4GH
 - Breeding API www.brapi.org



Computer scientist driven

• Phenotype Structure Standard



Minimal Information About Plant Phenotyping Experiment

www.miappe.org

- Many stakeholders
 - Elixir, Emphasis, Bioversity, North American PPN
- Open Community
- Crops and woody plants
- Many Implementation
 - Repositories: Generic, GnpIS, BioSample, PHIS, PIPPA

Papoutsoglou *et al.* (2020) Enabling reusability and interoperability of plant phenomic datasets with MIAPPE 1.1. *New Phytol*, 227:260-273; <https://doi.org/10.1111/nph.16544>

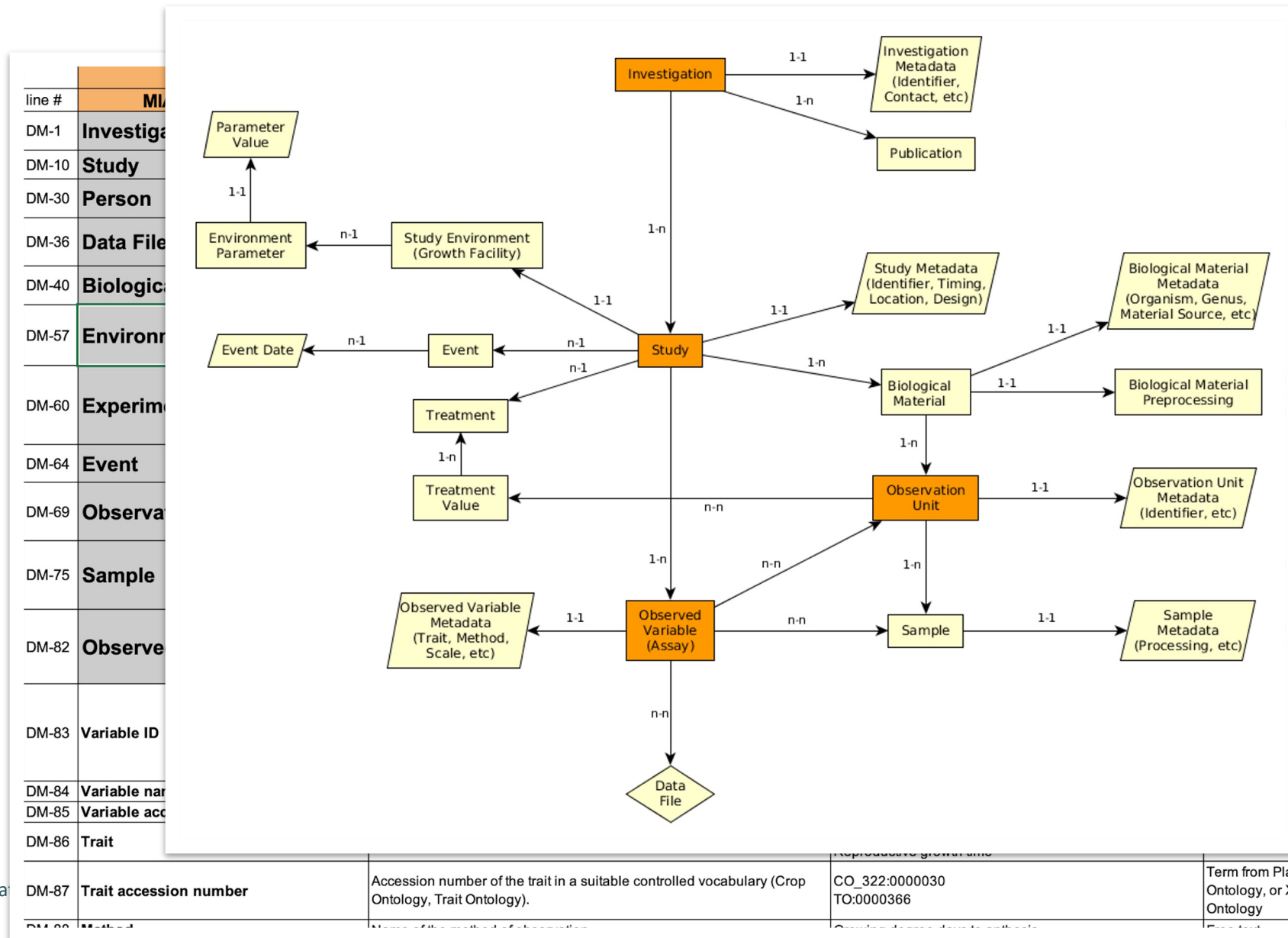
MIAPPE					
line #	MIAPPE Check list	Definition	Example	Format	Cardinality
OM-1	Investigation	Investigations are research programmes with defined aims. They can exist at various scales (for example, they could encompass a grant-funded programme of work, the various components comprising a peer-reviewed publication, or a single experiment).			1 per MIAPPE submission
OM-2	Investigation unique ID	Identifier comprising the unique name of the institution/database hosting the submission of the investigation data, and the accession number of the investigation in that institution.	EBI:12345678	Unique Identifier	0-1
	Investigation title	Human-readable string summarising the investigation.	Adaptation of Maize to Temperate Climates: Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with	Free text (short)	1
Environment					
ENV-1	Non-exhaustive list of Environment Parameters.				
line#	Environment parameters	Definition	Example	Format	
ENV-2	Growth facility				
ENV-3	Air temperature	List of hourly air temperature throughout the experiment.	22 °C	Numeric	
ENV-4	Organ temperature	List of hourly organ temperatures throughout the experiment.	18 °C	Numeric	
ENV-5	Experimental Factors				
TR-1	Non-exhaustive list of Experimental Factors that can be applied.				
line #	Factor type	Definition	Example factor values	Format	
TR-2	Seasonal environment	A plant treatment (EO:0001001) involving an exposure to a given conditions of regional seasons.	Spring season; dry season	Plant Environment Ontology:EO_0007038'	
TR-3	Air treatment regime	The treatment involving an exposure to wind/air with varying degree of temperature, which may depend on the study type or the regional environment.	28/25°C (Day/Night)	Plant Environment Ontology:EO_0007161'	
TR-4	Soil temperature regime	A physical plant treatment (EO:0007316) involving an exposure to varying degree of temperature, which may depend on regional environment.	27/25°C (Day/Night)	Plant Environment Ontology:EO_0007161'	
TR-5					

MIAPPE Specifications

- Specification table
- Sections
- Metadata Fields

line #	MIAPPE Check list	MIAPPE			Cardinality
		Definition	Example	Format	
DM-1	Investigation	Investigations are research programmes with defined aims. They can exist at various scales (for example, they could encompass a grant-funded programme of work, the various components comprising a peer-reviewed publication, or a single experiment).			1 per MIAPPE submission
DM-10	Study	A study (or experiment) comprises a series of assays (or measurements) of one or more types, undertaken to answer a particular biological question.			1+ per investigation
DM-30	Person	A human involved in the investigation or specifically any of its studies.			1+ per investigation / 0+
DM-36	Data File	A file or digital object holding observation data recorded during one or more assays of the study, typically in tabular form. Multiple data files may be provided per study, and each file can include observations for several observation units and several observed variables.			0+ per study
DM-40	Biological Material	The biological material being studied (e.g. plants grown from a certain bag or seed, or plants grown in a particular field). The original source of that material (e.g., the seeds or the original plant cloned) is called the material source, which, when held by a material repository, should have its stock identified.			1+ per study; 0+ per observation unit
DM-57	Environment	Environmental parameters that were kept constant throughout the study and did not change between observation units or assays. Environment characteristics that vary over time, i.e. environmental variables, should be recorded as Observed Variables (see below).			0-1 per study
DM-60	Experimental Factor	The object of a study is to ascertain the impact of one or more factors on the biological material. Thus, a factor is, by definition a condition that varies between observation units, which may be biotic (pest, disease interaction) or abiotic (treatment and cultural practice) in nature. Depending on the level of the data, an experimental factor can be either "what is the factor applied to the plant" (i.e. Unwatered), or the "environmental characterisation" (i.e. if no rain on unwatered plant : Drought ; if rain on unwatered plant: Irrigated)			0+ per study; 0+ per observation unit
DM-64	Event	An event is discrete occurrence at a particular time in the experiment (which can be natural, such as rain, or unnatural, such as planting, watering, etc). Events may be the realization of Factors or parts of Factors, or may be confounding to Factors. Can be applied at the whole study level or to only a subset of observation units.			0+ per study/observation unit
DM-69	Observation Unit	Observation units are objects that are subject to instances of observation and measurement. An observation unit comprises one or more plants, and/or their environment. There can be pure environment observation units with no plants. Synonym: Experimental unit.			1+ per study
DM-75	Sample	A sample is a portion of plant tissue harvested, non-harvested or extracted from an observation unit for the purpose of sub-plant observations and/or molecular studies. A sample must be used when there is a physical sample that needs to be stored and traced. Otherwise, observations made at the sub-plant level should be recorded as plant level observations using the observed variables to characterize the object of the observation (e.g. Berry sugar content, Fruit weight, Grain Protein content, Leaf 1 width, Leaf 2 width, Leaf 2 length).			0+ per observation unit
DM-82	Observed Variable	An observed variable describes how a measurement has been made. It typically takes the form of a measured characteristic of the observation unit (plant or environmental trait), associated to the method and unit of measurement. Multiple variables with the same combination of trait, method and scale can be used in association with different plant parts (leaf 1, leaf 2), when this distinction is necessary for observations referring to different parts of the same observation unit.			1+ per study
DM-83	Variable ID	Code used to identify the variable in the data file. We recommend using a variable definition from the Crop Ontology where possible. Otherwise, the Crop Ontology naming convention is recommended: <trait abbreviation>_<method abbreviation>_<scale abbreviation>. A variable ID must be unique within a given investigation.	Ant_Cmp_Cday	Unique identifier	1
DM-84	Variable name	Name of the variable.	Anthesis computed in growing degree days	Free text	0-1
DM-85	Variable accession number	Accession number of the variable in the Crop Ontology	CO_322:0000794	Crop Ontology term	0-1
DM-86	Trait	Name of the (plant or environmental) trait under observation	Anthesis time Reproductive growth time	Free text	1
DM-87	Trait accession number	Accession number of the trait in a suitable controlled vocabulary (Crop Ontology, Trait Ontology).	CO_322:0000030 TO:0000366	Term from Plant Trait Ontology, Crop Ontology, or XML Environment Ontology	0-1

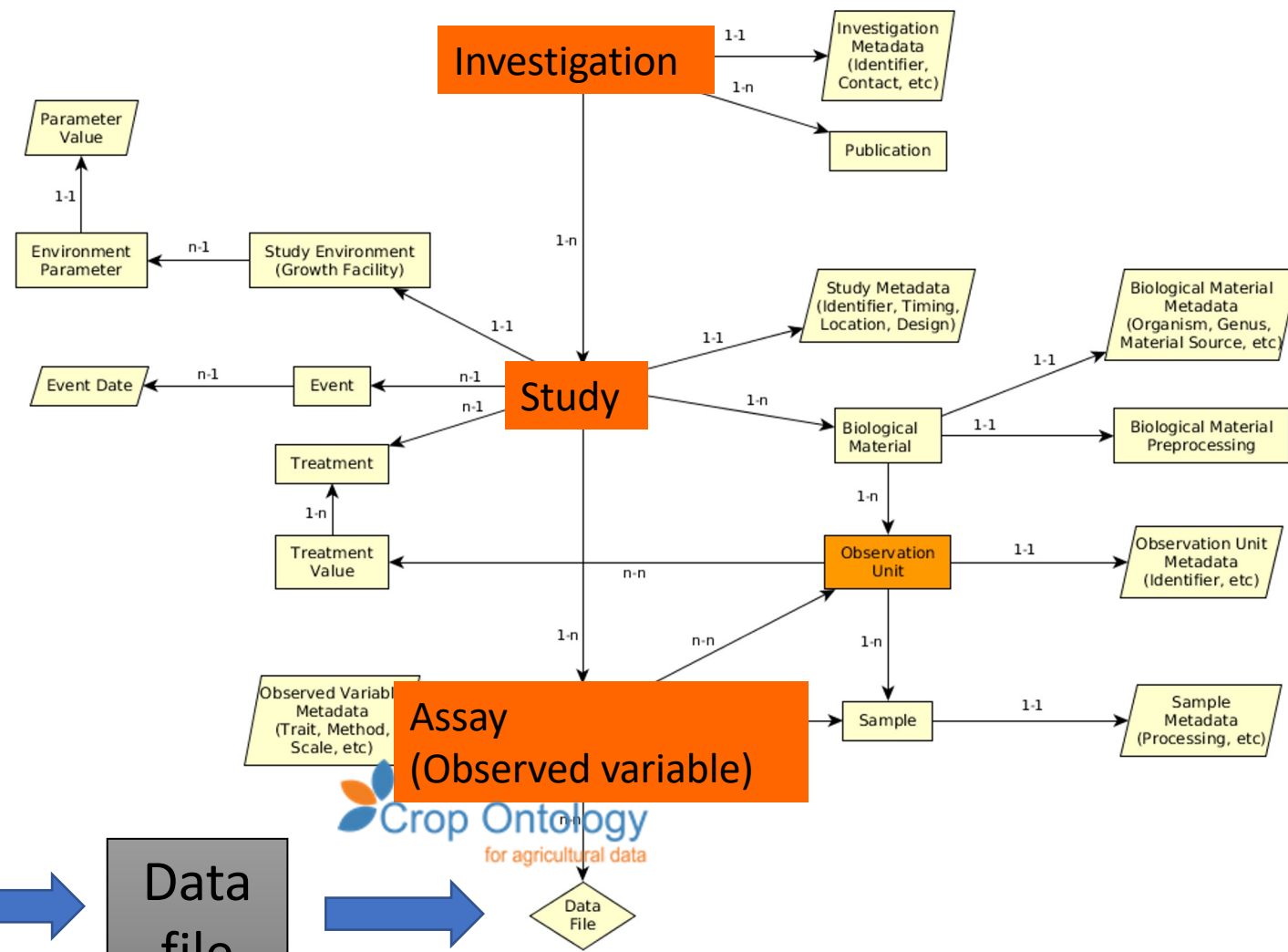
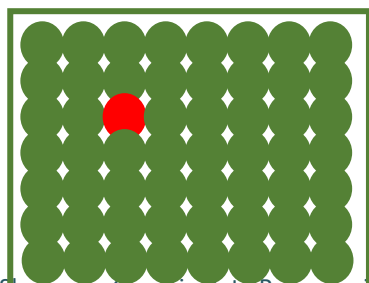
- Specification table
- Sections
- Metadata Fields
- Linked and organized



line #	Metadata Field
DM-1	Investigation
DM-10	Study
DM-30	Person
DM-36	Data File
DM-40	Biological Material
DM-57	Environment
DM-60	Experiment
DM-64	Event
DM-69	Observation Unit
DM-75	Sample
DM-82	Observed Variable
DM-83	Variable ID
DM-84	Variable name
DM-85	Variable accession number
DM-86	Trait
DM-87	Trait accession number

Format	Description	Cardinality
	programme of work, the	1 per
		1+ per
		1+ per
	vided per study, and	0+ per
	aterial (e.g., the seeds or	1+ per
		0+ per
	eristics that vary over	0-1 per
	tween observation	0+ per
	mental factor can be	0+ per
	in on unwatered plant:	0+ per
	Events may be the	0+ per
	traits in their environment.	1+ per
	olecular studies. A	0+ per
	be recorded as plant	0+ per
	ent, Leaf 1 width, Leaf 2	0+ per
	ent or environmental	1+ per
	iation with different	1+ per
		1
		0-1
	term	0-1
		1
		0-1

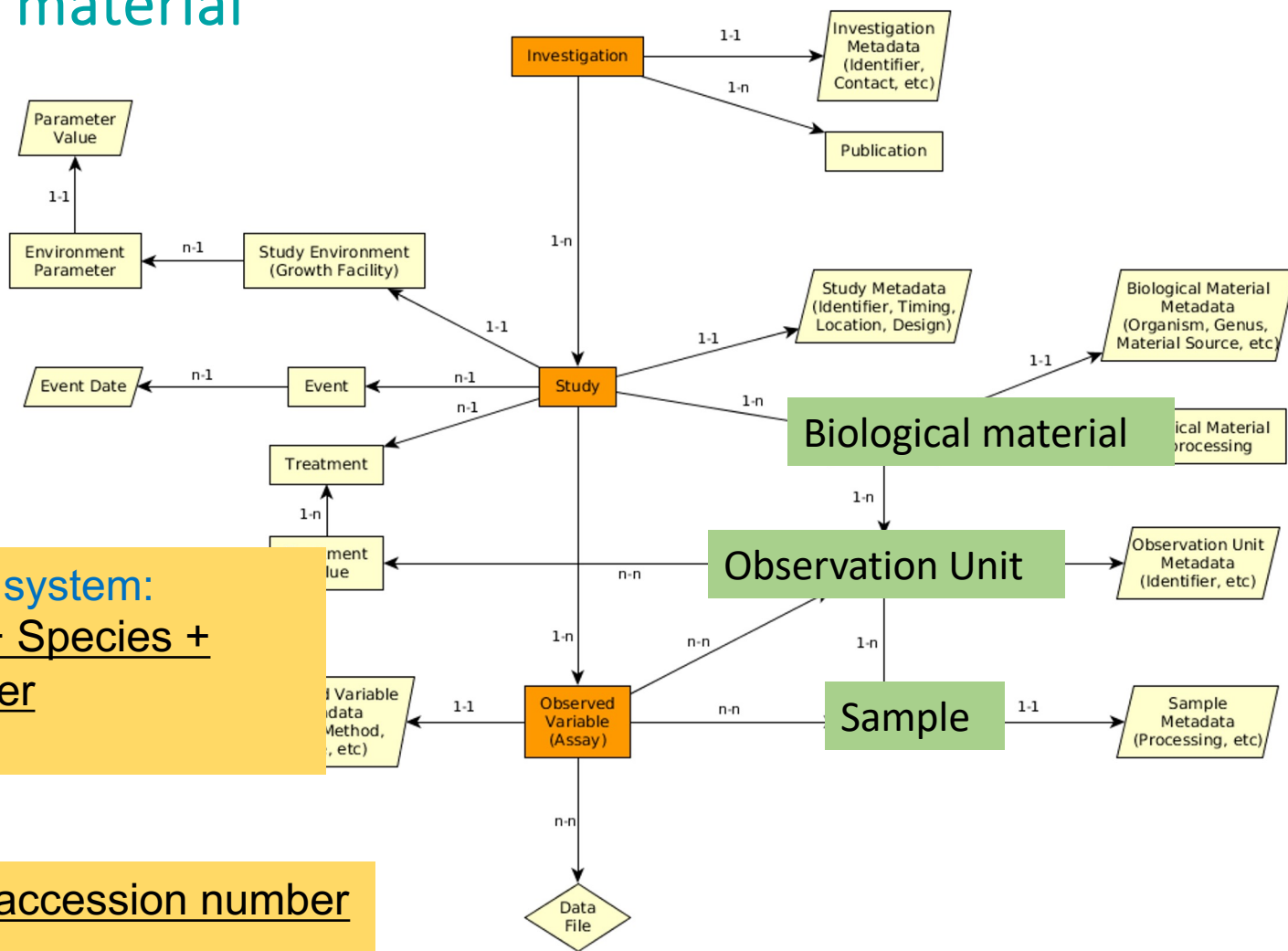
- **Investigation:** whole dataset
- **Study :** one experiment in one location for one to several year
- **Assay:**
 - trait or indice (Pheno or Env) observed
 - Level + Trait + Method + Scale/Unit
- **Level:**
 - Plant
 - Microplot
 - Block
 - Trial
 - ...



MIAPPE main sections – Biological material

- Plant Material
 - ◆ Identification
 - ◆ Description

• Multi Crop Passport Descriptor



MCPD identification system:

- Genebank/Lab + Species + accession number
- DOI

Material Source: accession, cultivar/variety, region of provenance, laboratory cross, ...



Biological material used in the study: seed lot, cuttings...



Plant Samples used in the study: detached leaves, ...

- Lab + internal accession number
- URI

- Lab + internal accession number (mandatory)
- BioSample ID

MIAPPE main sections – Observed Variable (Assay)



- Plant Phenomic Specific Ontology Model:

Phenotyping/environment variable = *Trait + Method + Unit/Scale*

Trait

+

Method

+

Unit

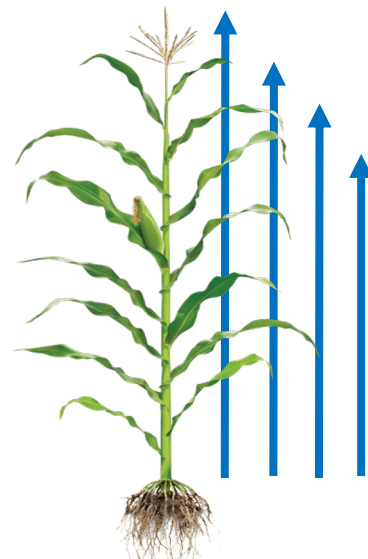
M1: Total height

M2: First tassel branch

M3: Last expanded leaf

M4: Youngest growing leaf

U1: cm
U2: mm



T1: Plant Height

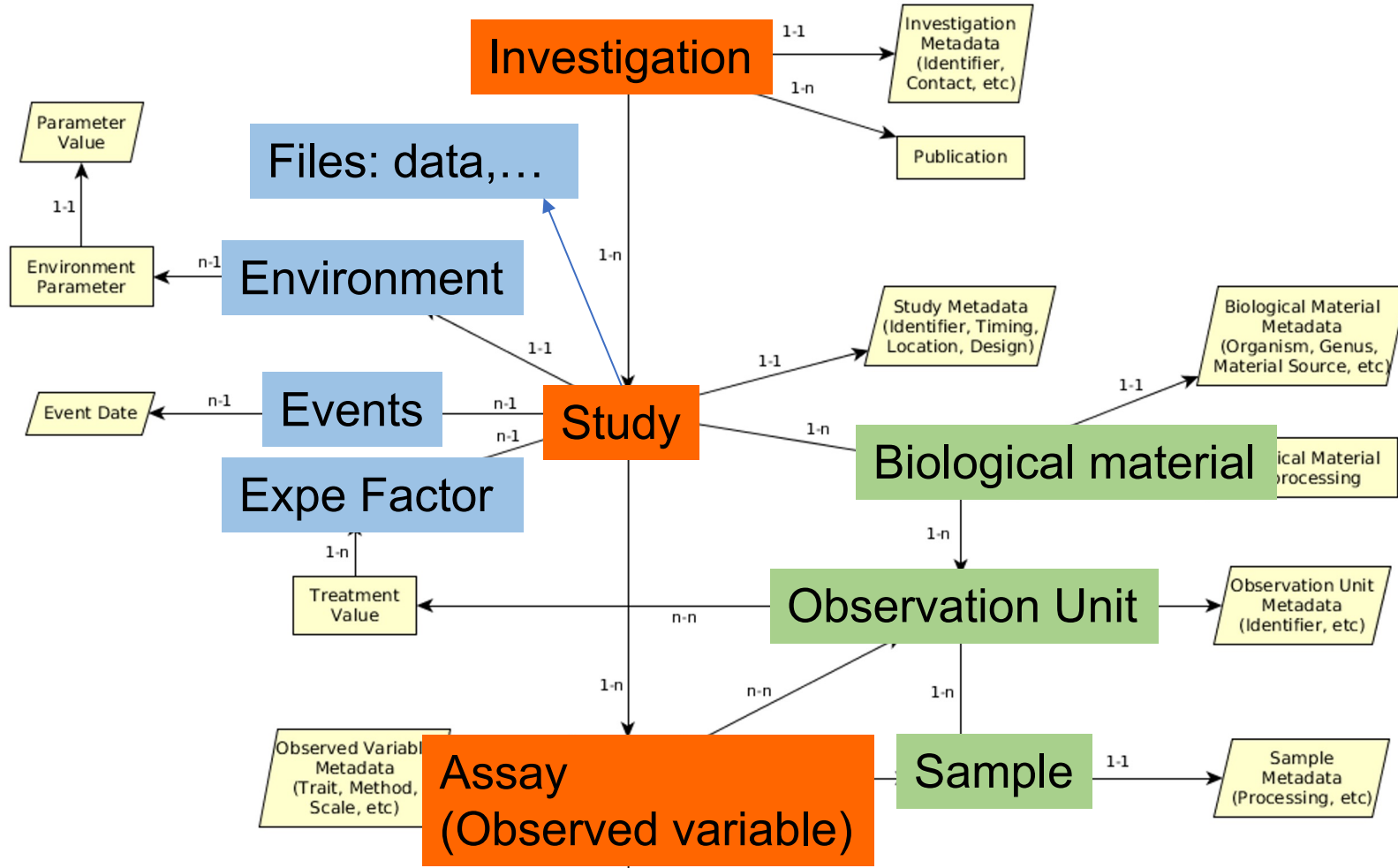


M5: Highest pixel corresponding to plant

U3: pixel

Slide:
L Cabrera-Bosquet

- MIAPPE – Other Important sections



- MIAPPE V1.1 Overview
Data file content
- Any format (Near Infra Red Spectrum, Images, Image Archives references,)
- Mostly tabular

A	B	C	D	E	F	G	H	I
Accession Number	Trial Site	Campaign	Circum1: Tree circumference at 1 year	Date [Circum1]	Height1: Tree total height at 1 year	Date [Height1]	Shoots3: Number of resprouts at 3 years	Date [Shoots3]
661300270	Ardon	2004	45.645632645603683	12/01/2004	284.3	12/01/2004		
661300270	Ardon	2005					14.630625	12/05/2005
661300444	Ardon	2004	38.96112577281653	12/01/2004	228.8	12/01/2004		
661300444	Ardon	2005					8.5030559999999991	12/05/2005
661300312	Cavallermaggiore	2004	52.4	01/01/2004	249.9	01/01/2004		
661300312	Cavallermaggiore	2005					12.981609000000001	01/05/2005
661300371	Cavallermaggiore	2004	45.74	01/01/2004	230.2	01/01/2004		
661300371	Cavallermaggiore	2005					10.3041	01/05/2005
661300487	Cavallermaggiore	2004	72.52	01/01/2004	309.8	01/01/2004		
661300487	Cavallermaggiore	2005					10.679823999999998	01/05/2005
661300585	Cavallermaggiore	2004	71.739999999999995	01/01/2004	305.7	01/01/2004		
661300585	Cavallermaggiore	2005					10.956100000000001	01/05/2005
661300468	Headley	2004	45.27	01/01/2004		247	01/01/2004	
661300468	Headley	2005					15.888196000000002	01/05/2005
661300469	Headley	2004	70.930000000000007	01/01/2004		313	01/01/2004	
661300469	Headley	2005					13.271448999999999	01/05/2005
661300533	Headley	2004	57.67	01/01/2004	258.8	01/01/2004		



MIAPPE @ Dataverse

Step 2: Add mandatory metadata for plant phenotyping data

Biological material

Use [BiologicalMaterial.xlsx](#). This spreadsheet contains the following fields: Mandatory fields:

- “Biological material ID” (ex: INRA:W95115_inra_2001): Lot number or material identifier in the data files
- “Material source ID” (ex: INRA:B73_usda) OR “Accession_number” (B73_usda) + “Holding_institute” (ex: INRA)
- Accession Number
- Genus
- Species
- Optional fields:
- “Material source DOI”: accession DOI
- Organism: NCBITAXON:4577
- “Infraspecific name”: variety names, cultivar names, etc...
- Genealogy:
 - Parent1or2_AccessionNumber
 - Parent1or2_TaxonGroup
 - Parent1or2_HoldingInstitutionName
 - Parent1or2_Type (father/mother/undefined)
- All MIAPPE Biological Material fields ([DM-40 to DM-56](#))
- Free input: synonyms, project IDs, any relevant information on the plant material.

Observed variables

Use [ObservedVariables.xlsx](#). This file is needed for the description of phenotyping experiments traits and methods.

Studies or experiments

It is recommended to list the experimentation done in this dataset, including in particular the GPS location, the site name and the environmental parameters which characterize the experimental sites. Use

[Studies.xlsx](#)

How to add metadata files in a Dataverse

- Click on “+ Upload Files” in the “Files” tab
- Add the file(s). Please note that :
 - the file size is limited to 15.0 GB
 - compressed files are automatically decompressed at the time of import
 - tabbed files must use the “,” separator and “UTF-8” encoding to avoid problems during import (see the dedicated section in the [user guide](#))
- Fill in the “Description” field for each added file
- Update the file labels by selecting “File options” > “Tags” for each file

Update the file labels

- Add a custom label, “Biological_Material” or “Observed_Variable” depending on the file type. If the label exists, it will be available in the “File labels” section, otherwise you will have to create it in the “Customize file label” section and apply it.

Add a custom label

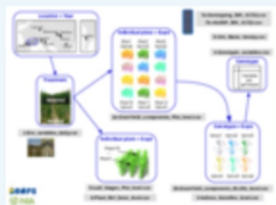
- Save your modifications

Maize Example

DROPS project dataset: <https://doi.org/10.15454/IASSTN>

A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios

Version 4.0



Millet, Emilie J.; Pommier, Cyril; Buy, Mélanie; Nagel, Axel; Kruijjer, Willem; Welz-Bolduan, Therese; Lopez, Jeremy; Richard, Cécile; Racz, Ferenc; Tanzi, Franco; Spitkot, Tamas; Canè, Maria-Angela; Negro, Sandra S.; Coupel-Ledru, Aude; Nicolas, Stéphane D.; Palaffre, Carine; Bauland, Cyril; Praud, Sébastien; Ranc, Nicolas; Presterl, Thomas; Bedo, Zoltan; Tuberosa, Roberto; Usadel, Björn; Charcosset, Alain; van Eeuwijk, Fred A.; Draye, Xavier; Tardieu, François; Welcker, Claude, 2019, "A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios", <https://doi.org/10.15454/IASSTN>, Recherche Data Gouv, V4, UNF:6:zS2/ccOQxFrKIUt+1S0Cvg== [fileUNF]

Cite Dataset ▾

Learn about [Data Citation Standards](#).

Access Dataset ▾

Contact Owner

Share

Dataset Metrics ?





15,218 Views ?

7,235 Downloads ?

3 Citations ?

Description ?

This dataset comes from the European Union project DROPS (DROught-tolerant yielding PlantS). A panel of 256 maize hybrids was grown with two water regimes (irrigated or rainfed), in seven fields in 2010 and 2012, respectively, spread along a climatic transect from western to eastern Europe, plus

<input type="checkbox"/>	 <p>10-Info-ObservedVariable.tab Tabular Data - 21.8 KB Published Nov 3, 2022 196 Downloads 17 Variables, 91 Observations UNF:6:bxOj...iaQ== </p> <p>List of phenotypic and environmental variables used in the dataset, following the MIAPPE data standard (https://www.miappe.org/)</p> <p>Observed Variable</p>
<input type="checkbox"/>	 <p>11-Info-Study.tab Tabular Data - 6.4 KB Published Nov 3, 2022 180 Downloads 11 Variables, 19 Observations UNF:6:9nO7...GEQ== </p> <p>List of studies, including locations, used in the dataset, following the MIAPPE data standard (https://www.miappe.org/)</p> <p>Study</p>

VariableID	VariableName	VariableAccessionNumber
Tnight	Night temperature	EIPO:0000001
Ri	Solar radiation	EIPO:0000002

StudyTitle	StudyUniqueID	ExperimentalSiteName	GeographicLocationLatitude	GeographicLocationLongitude
Biogemma Gaillac 2012	Gai12	Gaillac	43.9	1.89
Biogemma Gaillac 2013	Gai13	Gaillac	43.9	1.89

MIAPPE Metadata



8-Info-BiologicalMaterial.tab

Tabular Data - 49.9 KB

Published Nov 3, 2022

159 Downloads

20 Variables, 256 Observations UNF:6:xjhU...4SA==

This file contains the description of the genotypes. Briefly, all studied hybrids result from a F1 cross

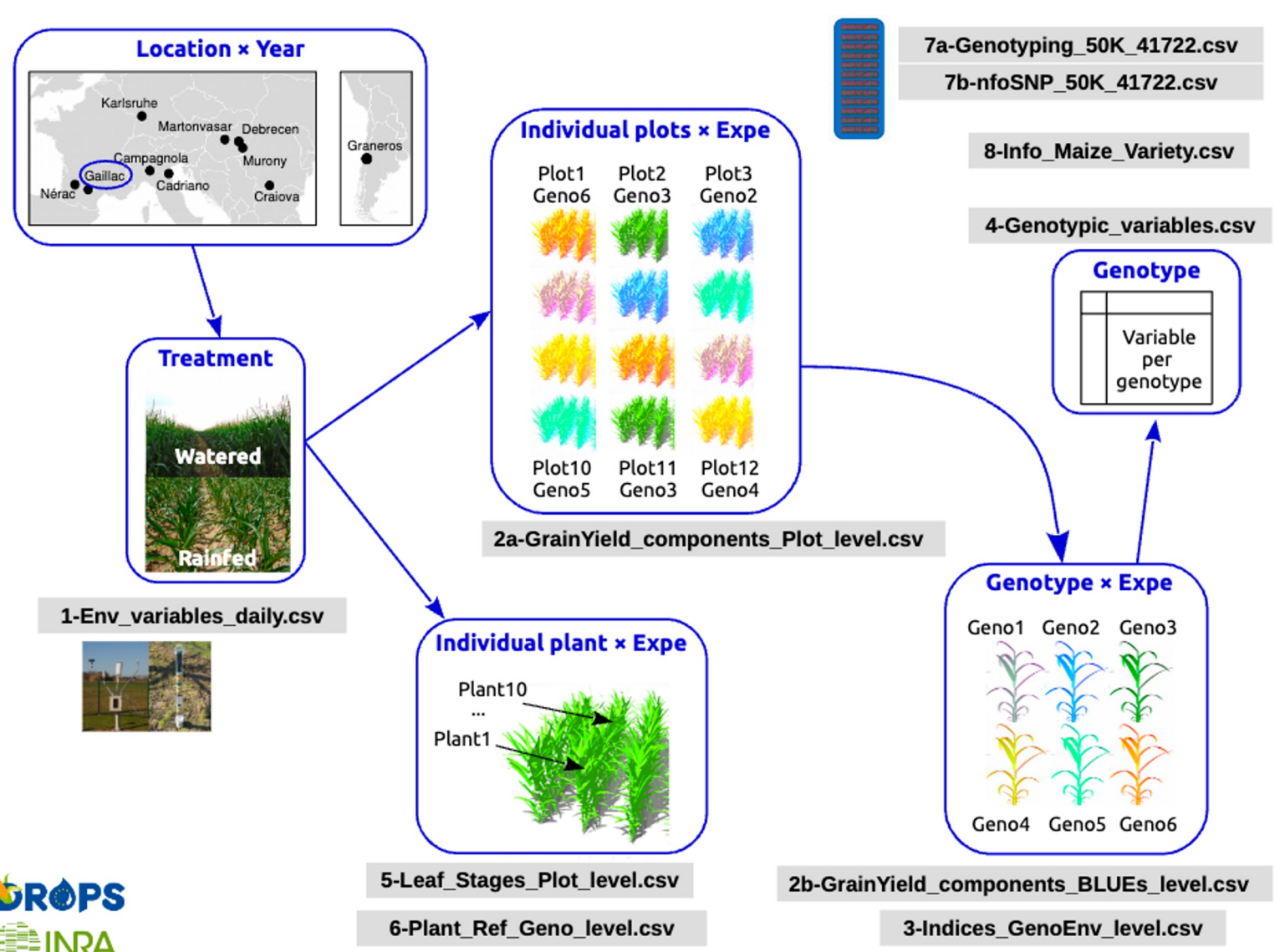
name»: Identifiers of the holding institutions. The Following columns have been added to comply to MIAPPE: "MaterialSourceDOI": DOI of the Accession identifying this variety, "MaterialSourceID": ID of the Accession identifying this variety, "BiologicalMaterialID": ID used in the data files, "Organism":NCBI Taxonomy identifier, "Genus", "Species", "InfraspecificName": variety name

Biological Material

VarietyID	AccessionID	AccessionHolding	MaterialSourceDOI	MaterialSourceID
11430	11430_H	inra	https://doi.org/10.15454/MDSCXQ	inra:11430_H
A3	A3_H	inra	https://doi.org/10.15454/GAHOFA	inra:A3_H
A310	A310_H	inra	https://doi.org/10.15454/CNJTVT	inra:A310_H

Good practice: Human friendly Provenance

Links between files



Publication of plant experimental data in generic data repositories

Step 2: Add mandatory metadata for plant phenotyping data

i Note

This section is also applicable to Zenodo.

i Note

Some dataverse instances (such as [dportal.biodata.pt](https://portal.biodata.pt)) have a MIAPPE metadata section that allows the use of dataverse web pages to store Biological material, Observation Variable and Study metadata. In that case you can use your dataverse's MIAPPE metadata instead of the MIAPPE metadata files described in this step.

The description of the plant material used in the experimentation proposed by MIAPPE is common to all types of experimentation (phenotypic, genetic, omic, etc...) and enables long-term reuse and interoperability between data sets. You can either use the full **MIAPPE template** or only selected sheets, available as individual file templates which can be found below.

Publication of plant experimental data in generic data repositories

Used by Metadata indexer in the future

<input type="checkbox"/>		<p>10-Info-ObservedV Tabular Data - 21.8 KB Published Nov 3, 2022 196 Downloads 17 Variables, 91 Observed List of phenotypic and (https://www.miappe.c Observed Variable</p>
<input type="checkbox"/>		<p>11-Info-Study.tab Tabular Data - 6.4 KB Published Nov 3, 2022 180 Downloads 11 Variables, 19 Observed List of studies, includi (https://www.miappe.c Study</p>

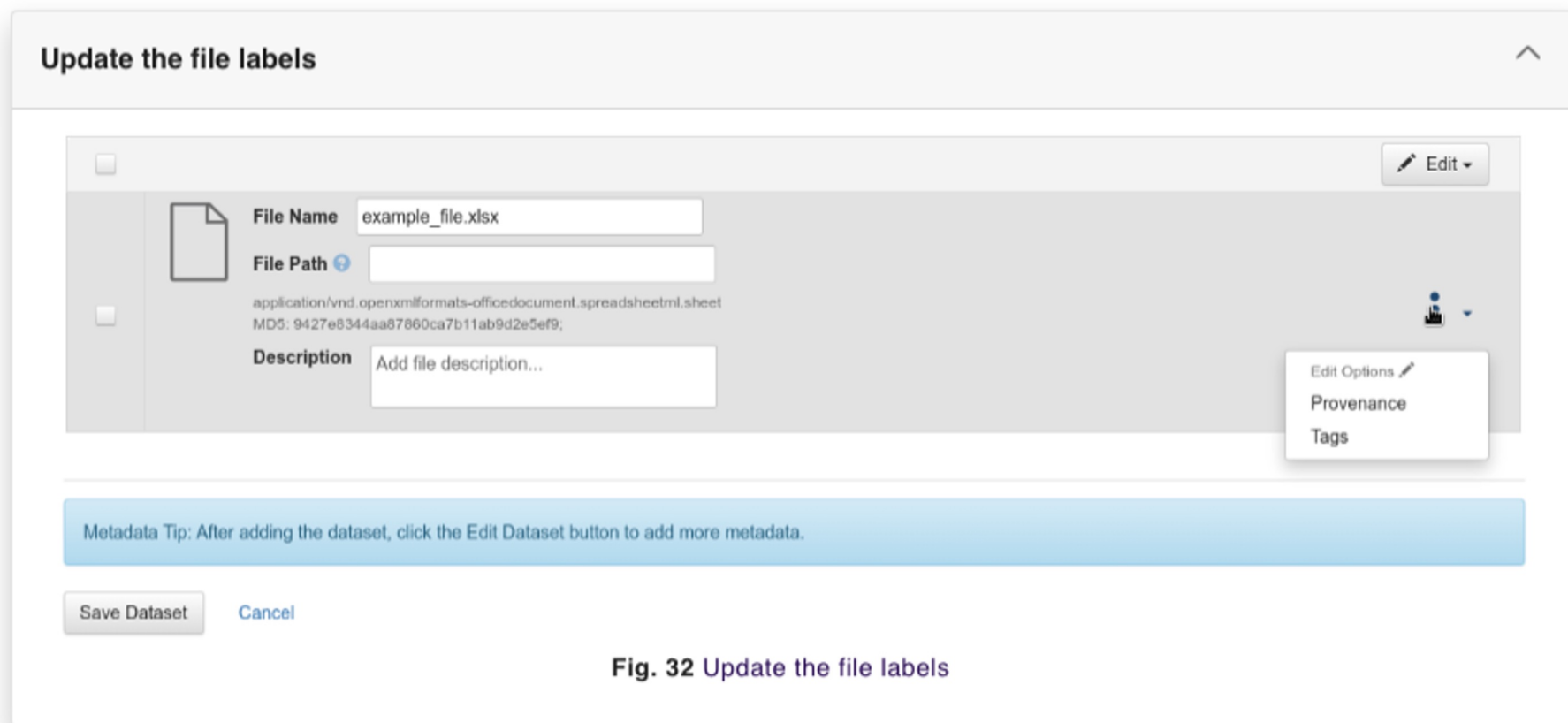


Fig. 32 Update the file labels

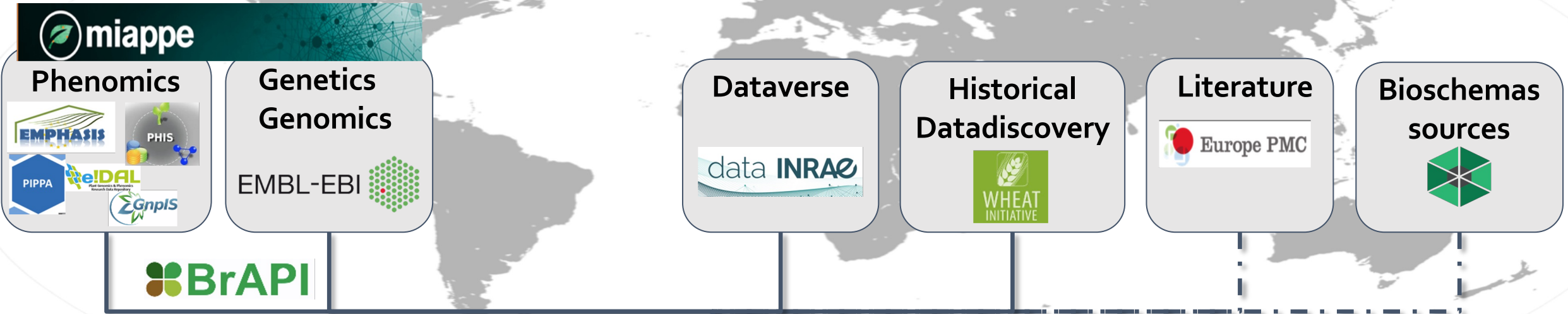
- Add a custom label, “Biological_Material” or “Observed_Variable” depending on the file type. If the label exists, it will be available in the “File labels” section, otherwise you will have to create it in the “Customize file label” section and apply it.



Benefit: Better visibility in FAIDARE Data portal

<https://urgi.versailles.inrae.fr/faidare/>

FAIDARE: Global Plant Research Data discovery portal



<https://urgi.versailles.inrae.fr/faidare/>

URGI More...

yield

Results 1 to 20 of 156

Species (21)
Filter on Species...

Data type
 Bibliography [151]
 None [5]

Ontology annotation (20)

10.3389/fpls.2018.00529 - OpenMinTeD@GnpIS
 Bibliography **Triticum Triticum aestivum**
 Global QTL Analysis Identifies Genomic Regions on Chromosomes 4A and 4B H...
 Related Traits Across Different Environments in Wheat (*Triticum aestivum* L.). 20...
 Genomic Regions on Chromosomes 4A and ... (expand)

10.1186/s12864-019-6005-6 - OpenMinTeD@GnpIS
 Bibliography **Triticum Triticum aestivum**
 Genome-wide association study reveals new loci for **yield**-related traits in Sichu...
 stripe rust stress. 2019 Genome-wide association study reveals new loci for **ye**...

trichocarpa (poplar) located between positions 7073249 and 7073695 on chr03_scaffold_3 and which

Ontology variable selection

Filter English

- Woody Plant Ontology **Ontology**
 - Biochemical **Trait class**
 - Morphological **Trait class**
 - Other **Trait class**
 - Phenological **Trait class**
 - Budflush **Trait**
 - BF_score_BI: Broadleaves budflush scoring **Variable**
 - Budget date **Trait**
 - BS_date: Budget date **Variable**

Identifier CO_357:1000009
 Name Budget date
 Description Assessment of the date when budget score will be reached for the first time
 Entity bud
 Attribute budset
 Class Phenological
 Main abbreviation BS_date
 Status Standard for INRAE
 Bud date protocol **Method**
 Identifier CO_357:2000014
 Name Bud date protocol
 Description Estimated date from polynomial regression of a time series of budflush or budget scores
 Class Computation
 Calendar day **Scale**
 Identifier CO_357:3000043
 Name Calendar day
 Data type Date
 Min 0
 Max 0
 Documentation <https://urgi.versailles.inrae.fr/>
 Context of use Research-intensive characterization
 Trial evaluation
 Breeding criterion
 Status Standard for INRAE

OK Cancel

Full text
 +
 Fine criteria
 +
Link back



- 33 databases indexed all over the world

Database (33)

Filter on Database...

Data provider

- INRAE-URGI [47,461,178]
- EBI [12,794,551]
- IPK [402,128]
- Gramene [257,561]
- T3 [223,013]
- UWA [167,167]
- PGSB [140,138]
- Rothamsted Research [137,917]
- EVA [66,510]
- GrainGenes [23,339]
- WUR [6,660]
- CIMMYT [1,788]
- TERRA-REF [1,098]
- CIRAD [722]
- NIB [694]
- VIB [692]
- IBET [67]
- DDBJ [58]
- IPGPAS [8]

France

- URGI SouthGreen
- arcad PHENOME
- MaiAGE

Germany

- HELMHOLTZ GEMEINSCHAFT
- IPK GÄTERSLEBEN
- PLANTS DB

United Kingdom

- EnsemblPlants
- EMBL-EBI
- TGAC The Genome Analysis Centre™
- ROTHAMSTED RESEARCH

United States

- USDA | NIFA | Triticace CAP
- GrainGenes
- Gramene

Mexico

- CIMMYT International Maize and Wheat Improvement Center

France

- BrAPI
- URGI cirad

Europe

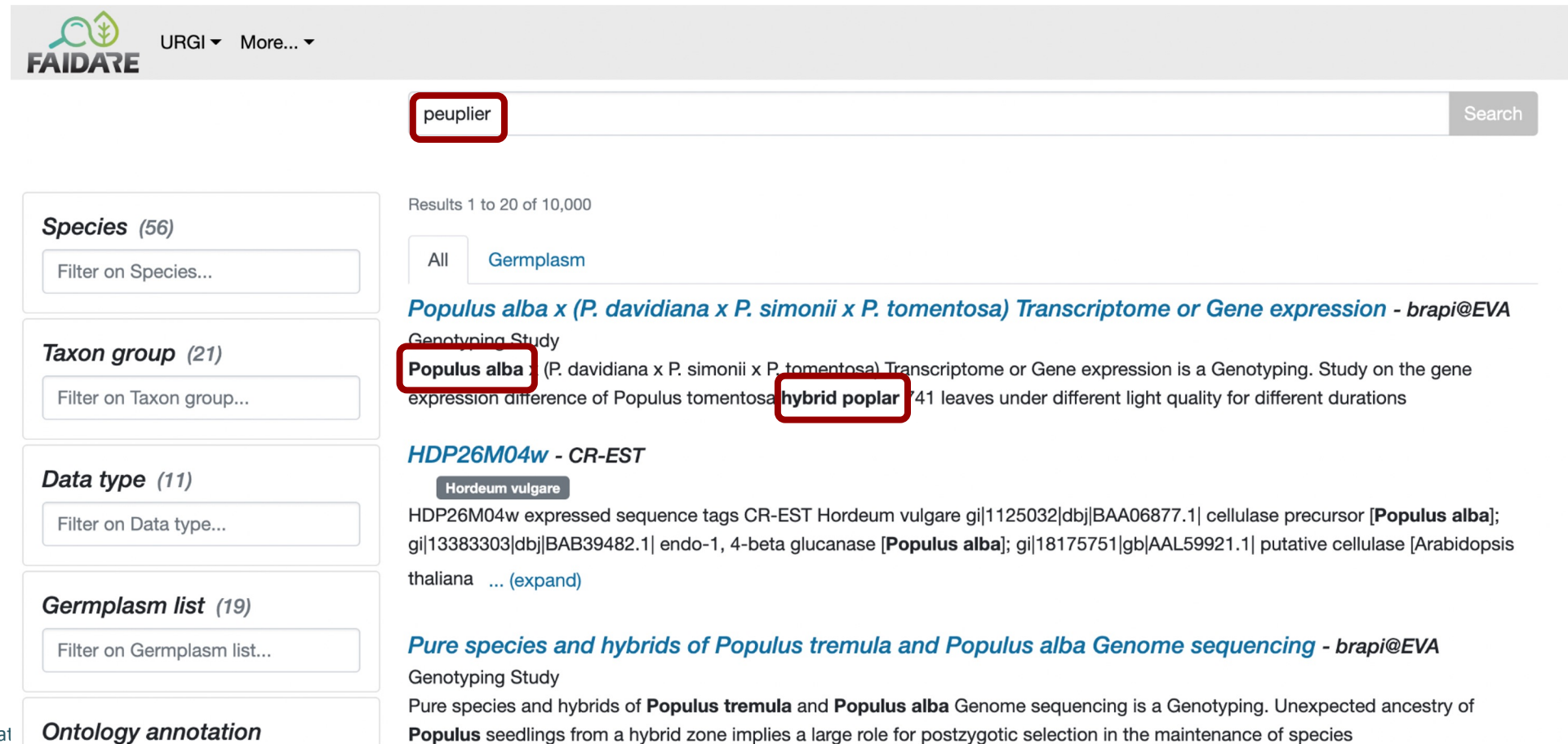
- IBET
- VIB
- WAGENINGEN UNIVERSITY & RESEARCH
- EMBL-EBI
- EVA
- NIB NACIONALNI INŠTITUT ZA BIOLOGIJO

United States

- TERRAPHENOTYPING REFERENCE PLATFORM

Minimal generic (Data Discovery)

- 33 databases indexed all over the world
- Simple full text search
 - Ontology based synonyms



The screenshot shows the FAIDARE search interface. At the top, the FAIDARE logo and navigation links 'URGI' and 'More...' are visible. A search bar contains the query 'peuplier' and a 'Search' button. Below the search bar, the results are displayed as 'Results 1 to 20 of 10,000'. On the left side, there are several filter panels: 'Species (56)', 'Taxon group (21)', 'Data type (11)', 'Germplasm list (19)', and 'Ontology annotation'. Each panel has a 'Filter on...' input field. The main results area shows a list of search results. The first result is titled '*Populus alba* x (*P. davidiana* x *P. simonii* x *P. tomentosa*) Transcriptome or Gene expression - brapi@EVA'. Below the title, it says 'Genotyping Study'. The text of the result includes '**Populus alba** (P. davidiana x P. simonii x P. tomentosa) Transcriptome or Gene expression is a Genotyping. Study on the gene expression difference of *Populus tomentosa* **hybrid poplar** 41 leaves under different light quality for different durations'. The second result is titled '*HDP26M04w* - CR-EST' and includes a 'Hordeum vulgare' tag. The text of this result includes 'HDP26M04w expressed sequence tags CR-EST Hordeum vulgare gi|1125032|dbj|BAA06877.1| cellulase precursor [**Populus alba**]; gi|13383303|dbj|BAB39482.1| endo-1, 4-beta glucanase [**Populus alba**]; gi|18175751|gb|AAL59921.1| putative cellulase [*Arabidopsis thaliana* ... (expand)']'. The third result is titled '*Pure species and hybrids of Populus tremula and Populus alba* Genome sequencing - brapi@EVA' and includes a 'Genotyping Study' tag. The text of this result includes 'Pure species and hybrids of **Populus tremula** and **Populus alba** Genome sequencing is a Genotyping. Unexpected ancestry of **Populus** seedlings from a hybrid zone implies a large role for postzygotic selection in the maintenance of species'.

- 33 databases indexed all over the world
- Simple full text search
 - Ontology based synonyms
- Dedicated filters
 - Data type
 - Taxonomy
 - Genetic resources
 - Crop Ontology
 - Gene Ontology
 - ...

Data type (37)

gen|

- Genome annotation [14,181,486]**
- Gene annotation [3,298,434]
- Gene [117,609]
- Genotyping Study [748]
- Genetic map [165]
- Genotyping Experiment [13]

Data type (37)

ph|

- Physical map feature [2,157,405]**
- Bibliography [2,988]
- Phenotyping Experiment [2,179]
- Phenotyping study [829]
- Phenotyping Study [816]
- Physical map [12]

Species (661)

Filter on Species...

Taxon group (156)

Filter on Taxon group...

- Triticum [20,150,119]**
- Hordeum [88,286]
- Brachypodium [67,367]
- Oryza [53,579]
- Sorghum [52,772]
- Setaria [48,627]
- Italica [48,625]
- Zea [46,973]

Other results are available. Refine your search.

Ontology variable selection

Filter English

- Wheat Crop Ontology **Ontology**
 - Abiotic stress **Trait class**
 - Agronomical **Trait class**
 - Grain weight **Trait**
 - Grain yield **Trait**
 - Lodging incidence **Trait**
 - LS_score: Lodging score **Variable**
 - Plant height **Trait**
 - Precocity **Trait**
 - Biotic stress **Trait class**
 - Other **Trait class**
 - Quality **Trait class**

LS_score **Variable**

Ontology name Wheat Crop Ontology

Identifier CO_321:1000099

Name LS_score

Synonyms Lodging score
Susceptibility to lodging

Institution INRA

Scientist Jacques Le Gouis

Date 15/06/2016

Crop Wheat

Cross reference WIPO:0000099

Lodging incidence **Trait**

Identifier CO_321:0000167

Name Lodging incidence

Description Indicates incidence of lodged plants.

Entity Plant

Attribute Lodging incidence

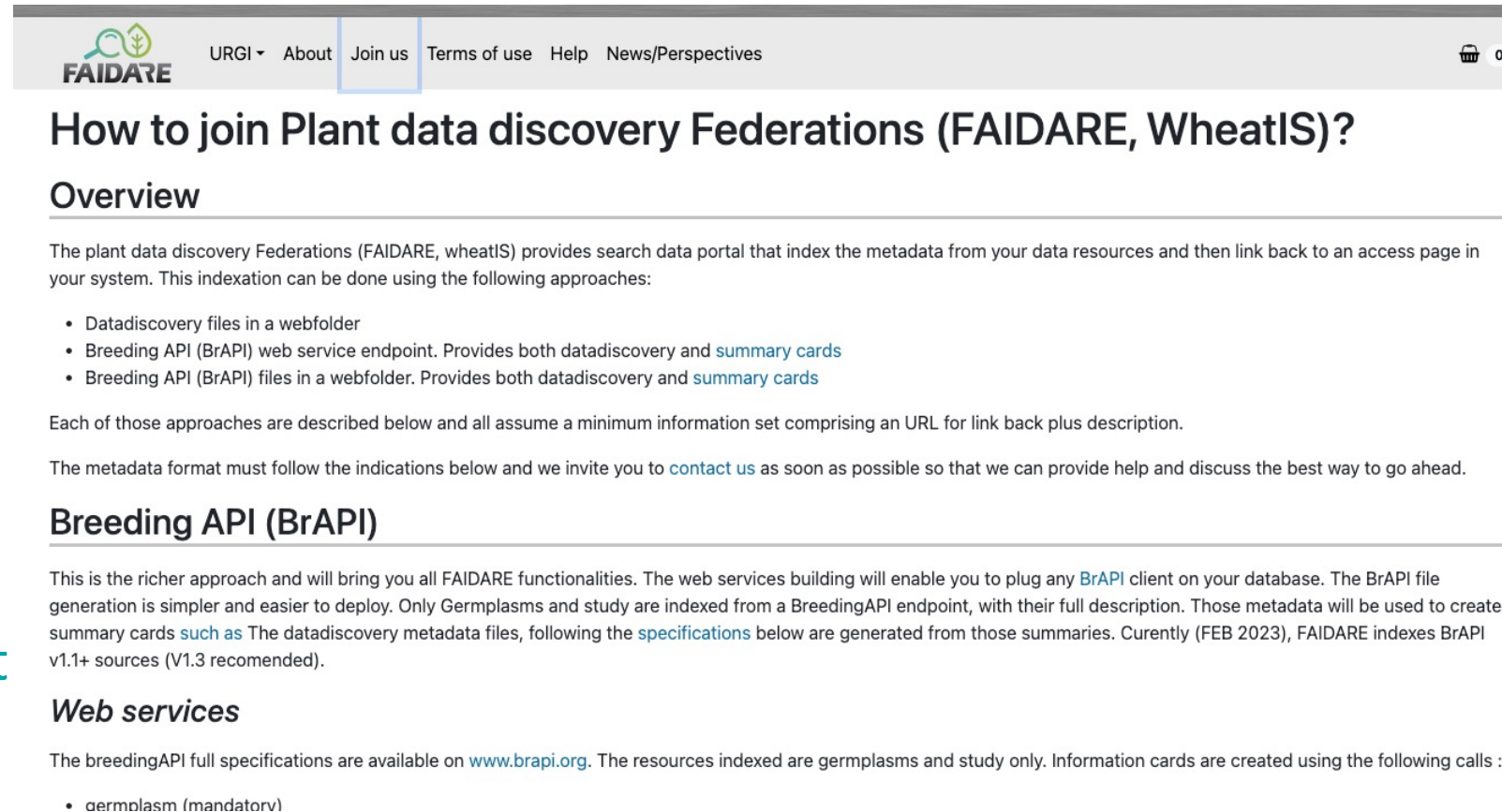
Class Agronomical

Main abbreviation Lodg

Alternative LOD

OK Cancel

- 33 databases indexed all over the world
- Simple full text search
 - Ontology based synonyms
- Dedicated filters
 - Data type
 - Taxonomy
 - Genetic resources
 - Crop Ontology
 - Gene Ontology
 - ...
- Friendly community management
 - Join the Federation support



The screenshot shows the FAIDARE website header with navigation links: URGI, About, Join us, Terms of use, Help, and News/Perspectives. The 'Join us' link is highlighted with a blue box. Below the header is the article title 'How to join Plant data discovery Federations (FAIDARE, WheatIS)?' and the section 'Overview'. The overview text describes the plant data discovery Federations and lists three approaches: Datadiscovery files in a webfolder, Breeding API (BrAPI) web service endpoint, and Breeding API (BrAPI) files in a webfolder. It also mentions that each approach is described below and that the metadata format must follow the indications below. The section 'Breeding API (BrAPI)' follows, explaining that this is the richer approach and will bring all FAIDARE functionalities. The 'Web services' section mentions that the breedingAPI full specifications are available on www.brapi.org.

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ About Join us Terms of use Help News/Perspectives



DROPS

Search

Species

- Zea mays [4]
 [2]
 None [1]

Data type

- dataset [4]

Database

- Research Data Gouv [4]

Data provider

- INRAE-URGI [4]

Results 1 to 4 of 4

[Flowering-related traits and florigen expression in a collection of maize inbred lines - Research Data Gouv](#)

dataset **Zea mays**

This dataset comes from a collection of 354 maize inbred lines corresponding to the "Public Amaizing Dent Panel". This panel includes: (i) 254 inbred lines from "Drops panel" assembled in the frame of F7P European project **DROPS** (FP7-244374, PI: F. Tardieu) ... [\(expand\)](#)

[DROPS maize genotyping elaborate dataset - Research Data Gouv](#)

dataset **Zea mays**

These three dataset correspond to three elaborate genotyping matrix after filtering and imputation published by Negro et al., (2019) in BMC Plant Biology: " [...] tion of 247 maize inbred lines corresponding to **Drops** Panel (Millet et al., 2016). These data were o ... [\(expand\)](#)

[DROPS maize GWAS analysis dataset - Research Data Gouv](#)

dataset **Zea mays**

This dataset corresponds to the most significant SNP ($-\log(p\text{-value}) > 5$) identified by genome-wide association studies between the three traits (Flowering Tim [...] red lines. These analysis were conducted in the **DROPS** DROught-yielding PlantS european project (F. T ... [\(expand\)](#)

[A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios - Research Data Gouv](#)

dataset **Zea mays**

This dataset comes from the European Union project **DROPS** (DROught-tolerant yielding PlantS). A panel of 256 maize hybrids was grown with two water regimes (irrigated or rainfed), in seven fields in 2012 and 2013, respectively, spread along a climatic trans ... [\(expand\)](#)

<https://urgi.versailles.inrae.fr/faidare/>

URGI ▾ About Join us Terms of use Help News/Perspectives  0

maize drought

Results 1 to 20 of 10,000

All **Germplasm**

Species (482)
Filter on Species...

Taxon group (121)
Filter on Taxon group...

Data type (25)
Filter on Data type...

Collection (24) ?
Filter on Collection...

Ontology annotation (2,000)
Filter on Ontology annotation
 Expand search ?

Database (26)
Filter on Database...

Data provider
 INRAE-URGI [4 129 467]

Regulatory networks of gene expression in maize (Zea mays) under drought stress and re-watering - brapi@EVA

Genotyping Study **Zea mays**
Regulatory networks of gene expression in **maize (Zea mays)** under **drought** stress and re-watering is a Genotyping. we investigated gene expression profiles in **maize** during **drought**/re-watering. These results provide new insights into the molecular mechanisms o ... (expand)

RNA-seq of maize ear primordia in response to drought - brapi@EVA

Genotyping Study
RNA-seq of **maize** ear primordia in response to **drought** is a Genotyping. RNA-seq data from a controlled **drought** stress experiment, which profiled immature **maize** ears during response to **drought**

Transcriptome of maize seminal root under drought - brapi@EVA

Genotyping Study
Transcriptome of **maize** seminal root under **drought** is a Genotyping. These data contained 48 seminal root samples of 8 genotypes (**drought**-resistant and -sensitive **maize** inbred lines) × 2 treatments (well watered and water stressed). Our goals were exploring t ... (expand)

maize drought tolerance mutant sequencing - brapi@EVA

Genotyping Study
maize drought tolerance mutant sequencing is a Genotyping.

Transcriptome of maize leaves under drought - brapi@EVA

Genotyping Study
Transcriptome of **maize** leaves under **drought** is a Genotyping.

<https://urgi.versailles.inrae.fr/faidare/>

The screenshot shows the FAIDARE web interface. At the top, there is a navigation bar with the FAIDARE logo and links for 'URGI', 'About', 'Join us', 'Terms of use', 'Help', and 'News/Perspectives'. A search bar contains the text 'drought maize' and a 'Search' button. Below the search bar, the results are displayed as 'Results 1 to 5 of 5'. There are two filter tabs: 'All' and 'Germplasm'. The first result is titled 'DROPS maize GWAS analysis dataset - Research Data Gouv' and is labeled as a 'dataset' with a 'Zea mays' tag. The description states: 'This dataset corresponds to the most significant SNP (-log(p-value)>5) identified by genome-wide association studies between the three traits (Flowering Time, Plant Height, Grain Yield) measured in 22 environments on a panel of 247 maize inbred lines. This ... (expand)'. The second result is titled 'A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios - Research Data Gouv' and is also labeled as a 'dataset' with a 'Zea mays' tag. The description states: 'This dataset comes from the European Union project DROPS (DROught-tolerant yielding PlantS). A panel of 256 maize hybrids was grown with two water regimes (irrigated or rainfed), in seven fields in 2012 and 2013, respectively, spread along a climatic trans ... (expand)'. The third result is titled 'DROPS maize genotyping elaborate dataset - Research Data Gouv' and is labeled as a 'dataset' with a 'Zea mays' tag. The description states: 'These three dataset correspond to three elaborate genotyping matrix after filtering and imputation published by Negro et al., (2019) in BMC Plant Biology: " [...] ; Glaubitz et al., 2014) on a collection of 247 maize inbred lines corresponding to Drops Panel (M...'.

On the left side of the interface, there are two filter panels. The first panel is titled 'Species (187)' and contains a button for 'Zea mays [5] x' and a text input field 'Filter on Species...'. The second panel is titled 'Data type' and contains a list of checkboxes with their respective counts: 'Genome annotation [1,702,658]', 'Protein [33,919]', 'Repeat reference [7,31]', 'Germplasm [2,993]', 'Biochemical reaction [357]', 'Genotyping Study [15]', 'Bibliography [13]', and 'dataset [5]'. The 'dataset [5]' option is checked. The third panel is titled 'Database' and contains a checkbox for 'Research Data Gouv [5]'.

<https://urgi.versailles.inrae.fr/faidare/>

- In the future
- Indexation
 - biological material from dataverse
 - Observation Variable from dataverse
- Demonstration with BrAPI indexed data

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ About Join us Terms of use Help News/Perspectives

Night temperature soil

Species (103)

Filter on Species...

Taxon group (35)

Filter on Taxon group...

Data type (21)

Filter on Data type...

Collection ?

- None [17,162]
- AKER collection [9]
- Collection AKER [9]
- Cherry Europe [1]
- Collection Européenne de cerisier [1]
- European Collection of Cherry [1]
- Plum Europe [1]

Ontology annotation

(1,781)

Results 1 to 20 of 10,000

All **Germplasm**

[Cucumis sativus Transcriptome or Gene expression - brapi@EVA](#)

Genotyping Study

Cucumis sativus Transcriptome or Gene expression is a Genotyping. The sex differentiation of cucumber dir affects cucumber fruit yield, and low **night temperature** promotes female flower differentiation. However, s on the effects of low **night temp** ... (expand)

[Oryza sativa Transcriptome or Gene expression - brapi@EVA](#)

Genotyping Study

Oryza sativa Transcriptome or Gene expression is a Genotyping. rice (Oryza sativa L.) in response to high **ni temperature** stress

[Brachypodium mexicanum Bmex347 Gene Expression Profiling - R19_010 transcripton brapi@EVA](#)

Genotyping Study

Brachypodium mexicanum Bmex347 Gene Expression Profiling - R19_010 transcriptome is a Genotyping. Brachypodium mexicanum Bmex347 Gene Expression Profiling; Sample from Roots grown under conditions day / 20C **night** and a well fertilised, well watered ... (expand)

[Brachypodium mexicanum Bmex347 Gene Expression Profiling - R19_012 transcripton brapi@EVA](#)

Genotyping Study

Brachypodium mexicanum Bmex347 Gene Expression Profiling - R19_012 transcriptome is a Genotyping. Brachypodium mexicanum Bmex347 Gene Expression Profiling; Sample from Roots grown under conditions

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ About Join us Terms of use Help News/Perspectives

Night temperature soil

Species

- Zea mays [144]
- Zea mays L. [2]

Taxon group (35)

Zea [146] ×

Filter on Taxon group...

Data type

- Protein [124]
- Phenotyping Study [19]
- Marker [2]
- Genotyping Study [1]

Ontology annotation (99)

Filter on Ontology annotatio

Expand search ?

Results 1 to 20 of 146

OptiV1S26663 - OPTIMAS-DW

Protein **Zea mays**

OptiV1S26663 Sequence OPTIMAS-DW Zea mays AT5G58070 **TEMPERATURE-INDUCED I** Encodes a **temperature**-induced lipocalin TIL1. Involved in thermotolerance. Peripherally as membrane.,GO:0005215; GO:0008289; GO:0005488; transporter ... (expand)

OptiV1S34336 - OPTIMAS-DW

Protein **Zea mays**

OptiV1S34336 Sequence OPTIMAS-DW Zea mays AT5G58070 **TEMPERATURE-INDUCED I** Encodes a **temperature**-induced lipocalin TIL1. Involved in thermotolerance. Peripherally as membrane.,GO:0005215; GO:0008289; GO:0005488; transporter ... (expand)

OptiV1C04673 - OPTIMAS-DW

Protein **Zea mays**

OptiV1C04673 Sequence OPTIMAS-DW Zea mays AT5G58070 **TEMPERATURE-INDUCED I** Encodes a **temperature**-induced lipocalin TIL1. Involved in thermotolerance. Peripherally as membrane.,GO:0005215; GO:0008289; GO:0005488; transporter ... (expand)

MTA_ATK Martonvasar 2013 - brapi@INRAE-URGI

Phenotyping Study **Zea mays**

MTA_ATK Martonvasar 2013 is a Phenotyping Study conducted from 2013-05-14 to 2012-11 (Hungary)

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ About Join us Terms of use Help News/Perspectives

Night temperature soil

Species

- Zea mays [144]
- Zea mays L. [2]

Taxon group (35)

Zea [146] ×

Filter on Taxon group...

Data type

- Protein [124]
- Phenotyping Study [19]
- Marker [2]
- Genotyping Study [1]

Ontology annotation (99)

Filter on Ontology annotatio

Expand search ?

Results 1 to 20 of 146

OptiV1S26663 - OPTIMAS-DW

Protein **Zea mays**

OptiV1S26663 Sequence OPTIMAS-DW Zea mays AT5G58070 **TEMPERATURE**-INDUCED I
Encodes a **temperature**-induced lipocalin TIL1. Involved in thermotolerance. Peripherally as
membrane.,GO:0005215; GO:0008289; GO:0005488; transporter ... (expand)

OptiV1S34336 - OPTIMAS-DW

Protein **Zea mays**

OptiV1S34336 Sequence OPTIMAS-DW Zea mays AT5G58070 **TEMPERATURE**-INDUCED I
Encodes a **temperature**-induced lipocalin TIL1. Involved in thermotolerance. Peripherally as
membrane.,GO:0005215; GO:0008289; GO:0005488; transporter ... (expand)

OptiV1C04673 - OPTIMAS-DW

Protein **Zea mays**

OptiV1C04673 Sequence OPTIMAS-DW Zea mays AT5G58070 **TEMPERATURE**-INDUCED I
Encodes a **temperature**-induced lipocalin TIL1. Involved in thermotolerance. Peripherally as
membrane.,GO:0005215; GO:0008289; GO:0005488; transporter ... (expand)

MTA_ATK Martonvasar 2013 - brapi@INRAE-URGI

Phenotyping Study **Zea mays**

MTA_ATK Martonvasar 2013 is a Phenotyping Study conducted from 2013-05-14 to 2012-11
(Hungary)

<https://urgi.v...>

Ontology variable selection ✕

English ▾

- Environmental Traits Ontology
 - Experimental practice Trait
 - Maximal temperature Trait
 - Tmax: Maximum air temperature Variable
 - Night temperature Trait
 - Soil water potential Trait
 - Solar radiation Trait
 - Agronomical Trait class
 - Morphological Trait class
 - Phenological Trait class
 - Anthesis silking interval Trait
 - Anthesis time Trait
 - Silking time Trait

Tmax Variable

Ontology name Environmental Traits
 Identifier EIPO:0000005
 Name Tmax
 Synonyms Maximum air temperature
 Scientist Melanie Buy
 Crop Maize

Maximal temperature Trait

Identifier EIPO:0000010
 Name Maximal temperature
 Description Maximum air temperature
 Entity Air
 Attribute Temperature

Maximum air temperature measurement Method

Identifier EIPO:0000015
 Name Maximum air temperature measurement
 Description Measured temperature
 Class Measurement

Celsius degrees Scale

Identifier EIPO:0000016
 Name Celsius degrees

OK
Cancel

Optiv IC05763 Sequence OPTIMAS-DW Zea mays AT3005890 RARE-COLD-INDUCIBLE ZB (RC1ZB) RARE-COLD-INDUCIBLE MTA_ATK Martonvasar 2013

Phenotyping Study Zea mays
 MTA_ATK Martonvasar 2013 is a Phenotyping Study conducted from 2013-05-14 to 2012-11-11 (Hungary)

<https://urgi.versailles.inrae.fr/faidare/>



[URGI](#) ▾ [About](#) [Join us](#) [Terms of use](#) [Help](#) [News/Perspectives](#)

Night temperature soil Se

Species

Zea mays [19]

Taxon group

Zea [19]

Data type

Phenotyping Study [19]

Database

brapi@INRAE-URGI [19]

Data provider

INRAE-URGI [19]

Ontology (17)

2 selected variables

Modify the selection

Results 1 to 19 of 19

[MTA_ATK Martonvasar 2013 - brapi@INRAE-URGI](#)

Phenotyping Study Zea mays

MTA_ATK Martonvasar 2013 is a Phenotyping Study conducted from 2013-05-14 to 2012-10-31 in Martonva (Hungary).

[KWS Murony 2013 - brapi@INRAE-URGI](#)

Phenotyping Study Zea mays

KWS Murony 2013 is a Phenotyping Study conducted from 2013-05-21 to 2013-10-05 in Murony (Hungary).

[Biogemma Nerac 2013 - brapi@INRAE-URGI](#)

Phenotyping Study Zea mays

Biogemma Nerac 2013 is a Phenotyping Study conducted from 2013-05-15 to 2013-10-11 in Nerac (France)

[Syngenta Campagnola 2011 - brapi@INRAE-URGI](#)

Phenotyping Study Zea mays

Syngenta Campagnola 2011 is a Phenotyping Study conducted from 2011-05-10 to 2011-10-04 in Campagn (Italy).

[Biogemma Nerac 2011 - brapi@INRAE-URGI](#)

Phenotyping Study Zea mays

Biogemma Nerac 2011 is a Phenotyping Study conducted from 2011-05-13 to 2011-10-08 in Nerac (France).

[Syngenta Debrecen 2011 - brapi@INRAE-URGI](#)

Phenotyping Study Zea mays

Syngenta Debrecen 2011 is a Phenotyping Study conducted from 2011-05-13 to 2011-10-11 in Debrecen

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ About Join us Terms of use Help News/Perspectives

Night temperature soil

Se

Species

Zea mays [19]

Taxon group

Zea [19]

Data type

Phenotyping Study [19]

Database

brapi@INRAE-URGI [19]

Data provider

INRAE-URGI [19]

Ontology (17)

2 selected variables

Modify the selection

Results 1 to 19 of 19

[MTA_ATK Martonvasar 2013 - brapi@INRAE-URGI](#)

Phenotyping Study **Zea mays**

MTA_ATK Martonvasar 2013 is a Phenotyping Study conducted from 2013-05-14 to 2012-10-31 in Martonva (Hungary).

[KWS Murony 2013 - brapi@INRAE-URGI](#)

Phenotyping Study **Zea mays**

KWS Murony 2013 is a Phenotyping Study conducted from 2013-05-21 to 2013-10-05 in Murony (Hungary).

[Biogemma Nerac 2013 - brapi@INRAE-URGI](#)

Phenotyping Study **Zea mays**

Biogemma Nerac 2013 is a Phenotyping Study conducted from 2013-05-15 to 2013-10-11 in Nerac (France)

[Syngenta Campagnola 2011 - brapi@INRAE-URGI](#)

Phenotyping Study **Zea mays**

Syngenta Campagnola 2011 is a Phenotyping Study conducted from 2011-05-10 to 2011-10-04 in Campagn (Italy).

[Biogemma Nerac 2011 - brapi@INRAE-URGI](#)

Phenotyping Study **Zea mays**

Biogemma Nerac 2011 is a Phenotyping Study conducted from 2011-05-13 to 2011-10-08 in Nerac (France).

[Syngenta Debrecen 2011 - brapi@INRAE-URGI](#)

Phenotyping Study **Zea mays**

Syngenta Debrecen 2011 is a Phenotyping Study conducted from 2011-05-13 to 2011-10-11 in Debrecen

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ À propos Rejoignez-nous (EN) Conditions d'utilisation Aide Nouveautés/Perspectives Web services

Study Phenotyping Study: University_of_Bologna Cadriano 2012



📍 Origin site 📍 Collecting site 📍 Evaluation site 📍 Multi-purpose site

Identification	
Name	University_of_Bologna Cadriano 2012
Identifier	dXJuOkIOUkFFLVSR0kvc3R1ZHkvQm9sMTI=
Source	
Data link	Link to this study on URGI GnplS

<https://urgi.versailles.inrae.fr/faidare/>



URGI ▾ À propos Rejoignez-nous (EN) Conditions d'utilisation Aide Nouveautés/Perspectives Web services

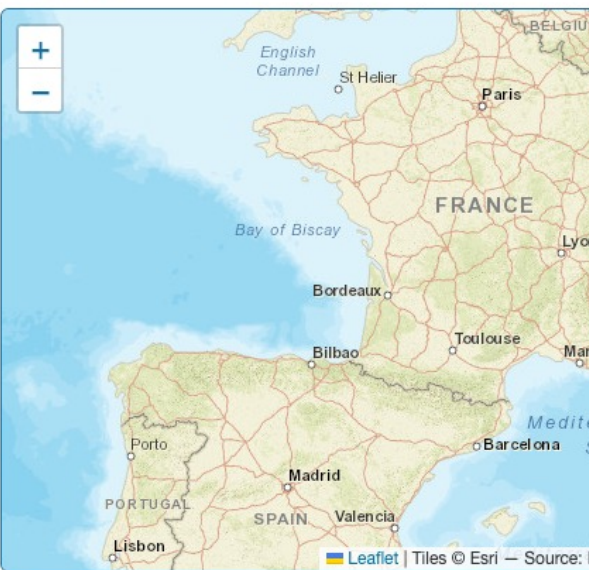
Study Phenotyping Study: University_of_Bologna Cadriano 2012

Genotype

Accession number	Name	Taxon
11430_H	11430_H	Zea may
A310_H	A310_H	Zea may
A347_H	A347_H	Zea may
A374_H	A374_H	Zea may

Variables

Variable ID	Variable short name	Variable long name	Ontology name	Trait description
EIPO:0000001	Tnight	Night temperature	Environmental Traits	Air night temperature
EIPO:0000002	Ri	Solar radiation	Environmental Traits	Solar radiation
EIPO:0000003	Psi	Soil water potential	Environmental	Water potential at th



Identification

Name	Univ
Identifier	dXJuOkIOUKFFLVVSR0I
Source	
Data link	Link to this study on UR



Erwan Le Floch



Daniel Arend



Sebastian Beier



Inês Chaves



Daniel Faria



Nicolas Francillonne



Vanita Haurheeram



Célia Michotey



Cyril Pommier



Philippe Rocca-Serra

Crop Ontology

for agricultural data

Crop Ontology

Arnaud E, Laporte MA, ...

MIAPPE community

ELIXIR Plant Community,

Krajewsky P, Cwiek H, Tardieu F, Usadel B, Arend D, Arnaud E, Junker A, King G, Laporte MA, Poorter H, Reif J, Rocca-Serra P, Sansone SA, Kersey P,
And many more!



Breeding API

Selby P, Mueller L, Robbins K, Backlund JE, ... ,
And many more!

Emphasis

Tardieu F, Usadel B, Arend D, Junker A, Poorter H, Neveu P, Pierushka R, Shur U...
And many more!



Acknowledgments & Questions

Thanks



Anne-Françoise Adam-Blondon

Bert Droesbeke

Cyril Pommier

Daniel Arend

Daniel Faria

Erwan Le Floch

Isabelle Alic

Matthias Lange

Philippe Rocca-Serra

Sebastian Beier

Elixir Plant community & platforms

Beier S., Gruden C., Pommier C., Coppens F, Scholz U., Lange M., Contreras B., Adam Blondon AF, Faria D, Chavez I, Miguel C, Droedsbek B, Finkers R, Papoutsoglou E, Olster R, Ramsak Z, ...

