

Data management in Plant Sciences: update on the BioData.pt Plant Sciences community

Pedro M Barros^{1,7}, Bruno Costa^{2,7}, Inês Chaves^{1,3,7}, Oscar Dias^{4,7}, Daniel Faria^{5,7}, Liliana Marum^{6,7}, Célia Miguel^{2,7}, M. Margarida Oliveira^{1,7}, Hugo Rodrigues^{1,7}, Filippo Bergeretti^{1,7}, Nelson Saibo^{1,7}, Marta Silva^{1,7}, Anabel Usié^{6,7}

1 - ITQB NOVA, 2 - FC.Ulisboa, 3 - iBET, 4 - U.Minho, 5 - ISTécnico, 6 - CEBAL-MED-CHANGE, 7 - BioData.pt | ELIXIR-PT

1st BioData.pt All Hands
3rd Technical Meeting

CCMAR, 28-30/09/2022



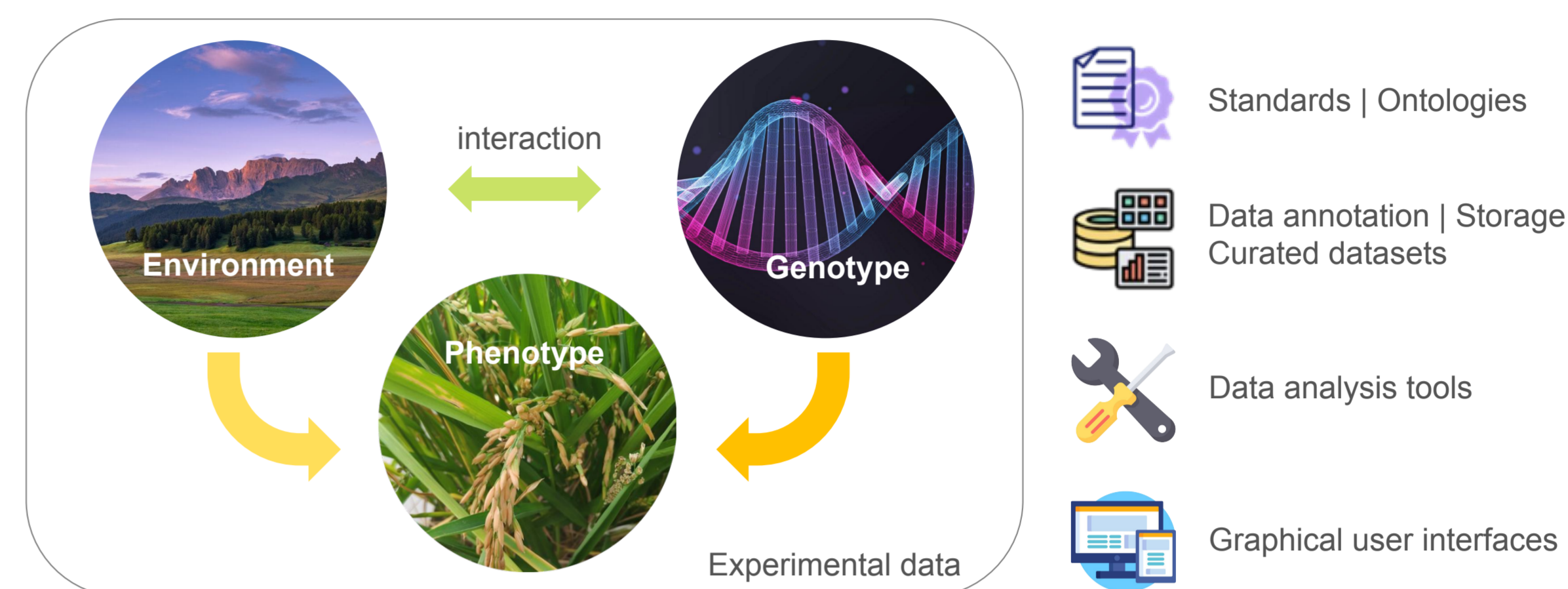
MOTIVATION

To address the challenge of sustainable production of food and non-food plant products (e.g. wood, cork, paper), particularly in the scope of climate change, many plant sciences researchers aim to study the **interplay between plant development, genomic features and environmental signals**.

The **Plant Sciences Community** aims to tackle the challenge of **managing plant genotypic and phenotypic data according to the FAIR data principles**. Aligned with ELIXIR Europe, this effort is determinant to guarantee efficient integration and reuse of plant phenotyping and genotyping datasets, so that full value can be extracted from the increasing wealth of data collected by researchers and producers worldwide. **In Portugal, this community is coordinated by ITQB NOVA and also includes several researchers from iBET, FCUL, U. Minho and CEBAL.**

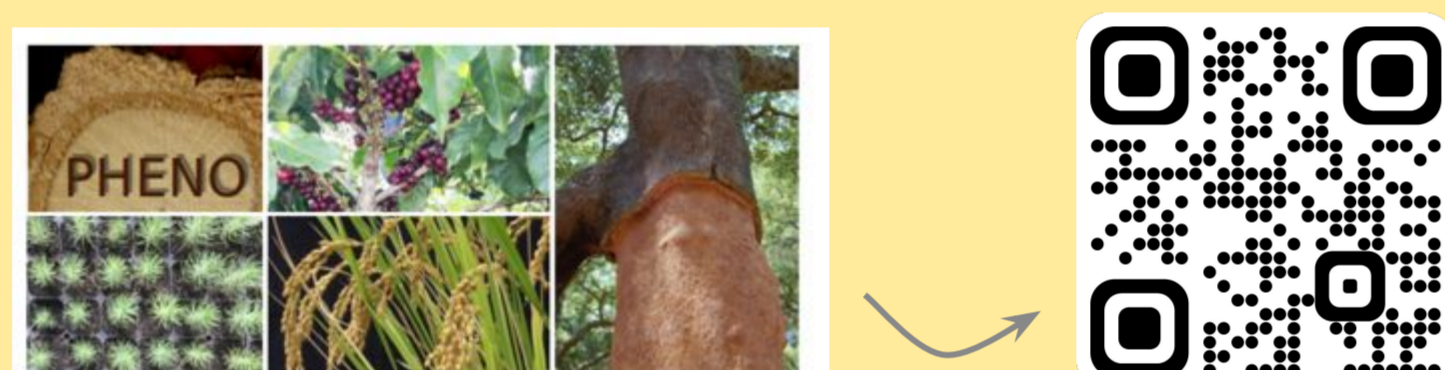
OBJECTIVES


- Development of tools and standards for efficient research data management (RDM) in plant sciences, following FAIR data principles.
- Development of node services for data analysis, visualization and retrieval



1 Tools and standards for Research Data Management

- The Plant community at ELIXIR collaborated in the development of:
 - **MIAPPE**, a standard for metadata annotation from plant phenotyping experiments [1]
 - PT **BrAPI endpoint - PHENO** - to access plant phenotypic data, fully interoperable with other European endpoints



- Ontologies to describe plant traits and related experiments - Curators of Crop Ontology and Woody Plant Ontology 
- The team is promoting the adoption of the developed resources by other plant scientists, and helping on the implementation of RDM Plans
- Development of a GUI to assist metadata annotation, following the MIAPPE standards, and integration into PHENO

2 Cork Oak Genome Portal

- The **CorkOakDB** is the **web portal** to access and browse the Cork Oak (*Quercus suber*) genomic data in a single, focused repository [2]
- Public access to the **curated genomic and transcriptomic data** on available in this species
- Allows gene sequence retrieval and data visualization (e.g. gene expression).
- Digital content was indexed using Bioschemas markup to improve findability by web search engines.

- > 12 000 page views
- Users from Portugal, Spain, Germany, USA and China, among others



4 Plant Small RNA Portal

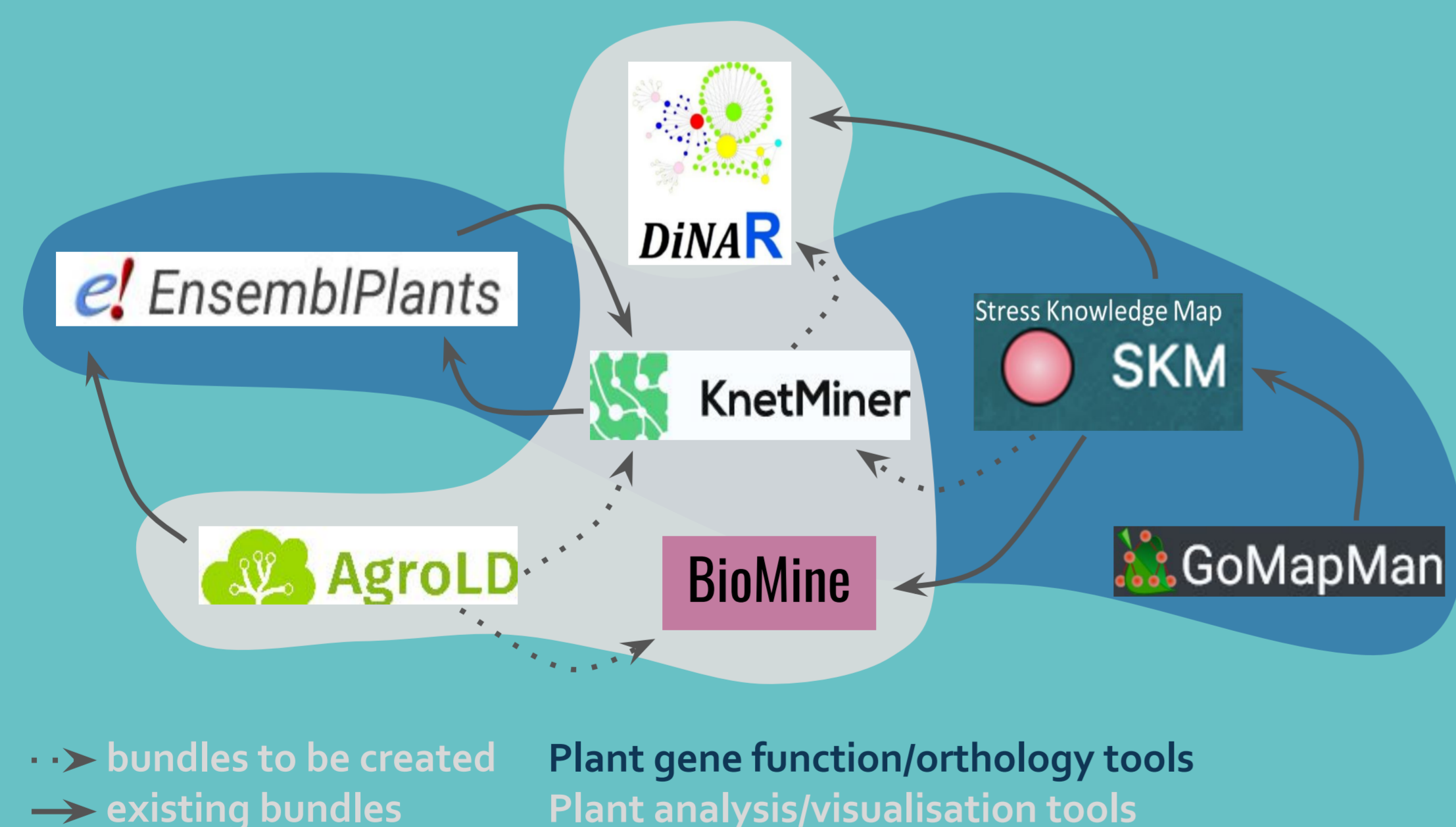
- The plant sRNA portal provides a **database of sRNAs annotated by the miRPursuit pipeline**.
- It is designed to provide a **GUI**, to interact with results and perform further research on the annotated data.
- Includes tools to integrate external information, aid the visualization of the results, and predict miRNA targets and functions.

- > 92 page views
- Users from Portugal, USA, Indonesia and Hong Kong, among others



3 INCREASING Plant data findability and reuse beyond ELIXIR

- **Implementation study** (ELIXIR Commissioned Services)
- Adapt and integrate ELIXIR plant services and tools through service bundles
- **Integrate data from non-model species:** Cork oak genomic data used as test case for the data visualization bundle for plant 'omics'



- Preparation of training materials applied to non-model species

5 Industry Knowledge exchange

- Bringing the ELIXIR Plant Data Infrastructure to the Portuguese pulp and paper industry - partnership with **The Navigator Company**
- Transfer knowledge on standards for FAIR plant data access
- Organizing data on eucalyptus breeding according to the MIAPPE and BrAPI standards; establish an access protocol for data submission in PHENO.



Final Remarks

- The Plant community at Biodata.pt is engaged in **maintaining and improving the developed node services** and continue **promoting the adoption of the developed RDM resources** by plant scientists.
 - Engagement with EMPHASIS by active participation in HORIZON projects such as PHENET (Bringing Emphasis to Operation: European Infrastructure for Multi-Scale Plant Phenomics and Simulation for Food Security in a Changing Climate)
 - Development of a RDM Plan and metadata annotation in the PRIMA R&D project TRACE-Rice
- Collectively the work being developed on Plant RDM will contribute to a major goal of establishing a National High-Throughput Phenotyping Infrastructure.

References:

- [1] Papoutsoglou et al. (2020) New Phyt. 227(1), 260-273.
[2] Arias-Baldrich et al. (2020) Database 2020, baaa114



Acknowledgements:

We are grateful to ELIXIR Implementation Study: INCREASING