# DIVERSify: Designing InnoVative plant teams for Ecosystem Resilience and agricultural Sustainability

| | |
|---|---|
| **Grant Agreement No.:** | 727284 |
| **Project Acronym:** | DIVERSify |
| **Project Title:** | Designing Innovative Plant Teams For Ecosystem Resilience And Agricultural Sustainability |
| **Project Co-ordinator:** | Dr Alison Karley, JHI |
| **Tel:** | +44 (0)1382 568820 |
| **Email:** | Alison.Karley@hutton.ac.uk |

## Plant and trait database V3 (Report, Public)

## Deliverable 2.9 (D16)

| | |
|---|---|
| Deliverable Lead: | UCPH |
| Deliverable Due Date: | 31-March-2021 |
| Actual Submission Date: | 30-March-2021 |
| Version: | 1.0 |

| | |
|---|---|
| Work Package: | 2 |
| Lead Author: | Lars Pødenphant Kiær (UCPH) |
| Contributing authors: | Christoph Scherber (ZFMK), Silvia Pappagallo (ZFMK), Jana Brandmeier (ZFMK), Martin Weih (SLU), Stefano Tavoletti (UNIVPM), Diego Rubiales (CSIC), Ángel Villegas-Fernández (CSIC), Eveline Adam (SZG), Sebastian Raubach (JHI), Paul Shaw (JHI), Adrian Newton (JHI), Alison Karley (JHI) |
| Reviewers: | M. Carlota Vaz Patto (ITQB), Joëlle Fustec (ESA) |

| History of Changes | | |
|---|---|---|
| **Version** | **Publication Date** | **Change** |
| 0.1 | 28th March 2021 | Initial version sent to authors and reviewers |
| 0.2 | 29th March 2021 | Version with author and reviewer edits and contributions incorporated |
| 1.0 | 30th March 2021 | Final version |

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

## Table of Contents





This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

## Executive Summary

A comprehensive database of crop traits in monocultures and intercrops or 'plant teams' has been assembled from data collected from experimental trials in the DIVERSify project. Within the trials, trait data have been collected for cereal, legume, herb and grassland species grown in plant teams and monocultures under different environmental and pedo-climatic conditions. The report summarises how the database was constructed by collating and aligning trial data and metadata provided by partners in standardised templates. Storage, further development and use of the database in DIVERSify outputs is described, along with details about how the database is being made available to the scientific community as an open access data resource.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

# 1. Background

In the DIVERSify project, a central trait database is maintained containing data on crop traits across different plant species and cultivars when grown together as intercrops or 'plant teams'. The database contains data collected within the consortium with a specific focus on crop traits measured in plot-scale field experiments in Work Package 2 (Ecological Approach to identify mechanisms and traits for optimised plant teams). Trait data was collected according to standardised protocols (Kiær et al. 2020) from co-ordinated field trials summarised in Scherber et al. (2019). The main aim of the database is to store all measured trait values for monocultures and plant teams under high- vs. low-intensity management. The first version of the database on crop traits - of plants mainly grown in individually or in monocultures - was assembled from existing open-access data (Scherber 2017). This report summarises how trait data collected from field trials is being used to develop a DIVERSify database of plant traits in monocultures and plant teams.

# 2. Data sources

## 2.1. Field trials

Trait data were collated from a large number of legume-cereal intercropping field trials established across seven countries and four pedoclimatic regions of Europe (**Table 1**; photos of example trials are shown in **Appendix 1**).

Coordination of field activities is crucial for evaluating trait and performance data from parallel field trials across sites and/or years, such as those operated in the DIVERSify project. For this purpose, we developed, for WP2 trials, a set of 47 standardized protocols for trait assessment distributed across six categories: Phenology, Morphology, Physiology, Reproduction, Yield and post-harvest, and Non-crop biodiversity. For each trait included in the handbook, method(s) of collection and units were described, and a standard name reported. To ensure that a minimum dataset was collected, given the high number of possible measurements, traits were classified as 'essential' (representing minimum basic data required to achieve project goals) or 'useful' (providing valuable additional data).

These protocols for field measurements of plant traits have been collated into a 110pp handbook with Introduction, References, and Appendices (Kiær et al. 2020) and made publicly available at the project legacy website (www.plant-teams.org). Besides increasing the quality of the dataset and acting as documentation for the many field assessments, it is a tool for improving future coordinated field assessments.

## 2.2. Online repositories

Using open-source software and online databases, we developed a toolset to set up and query a scalable plant trait database to be used within DIVERSify and elsewhere for the conception of



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

complementary plant teams. This was used to construct a 'virtual' database by collating traits of species (individual plants or in monoculture stands) using publicly available trait databases (Scherber, 2017). The database is implemented in the open-source software package R (R Core Team 2017) and RStudio (RStudio Team 2017), using the TR8 library (Bocci 2015) to query plant traits across publicly available databases (Scherber 2017). We successfully extracted information for 25 numeric traits across 128 European grassland and cropland species from nine online data sources. Trait data were collated for most cereal, legume, herb and grassland species currently utilised in plant teams across Europe. The database and associated R scripts are stored both locally at ZFMK/WWU on a RAID system and on a network folder with mirror copies distributed across different server locations.

## 3. Harmonisation and management of data

Data from multiple field trials must be harmonised and documented to facilitate proper merging, analysis and sharing of datasets. More and more datasets are being shared online. However, accessibility and practical use of these datasets is encumbered by the fact that many of them lack clear semantics and metadata standardization, and existing data reporting tools (e.g. from biodiversity experiments) are inadequate to fully accommodate the specificities of intercropping data. This represents a major barrier to exchange of data on crop mixtures. Therefore, a robust approach was developed to increase consistency and quality of collated field trial data.

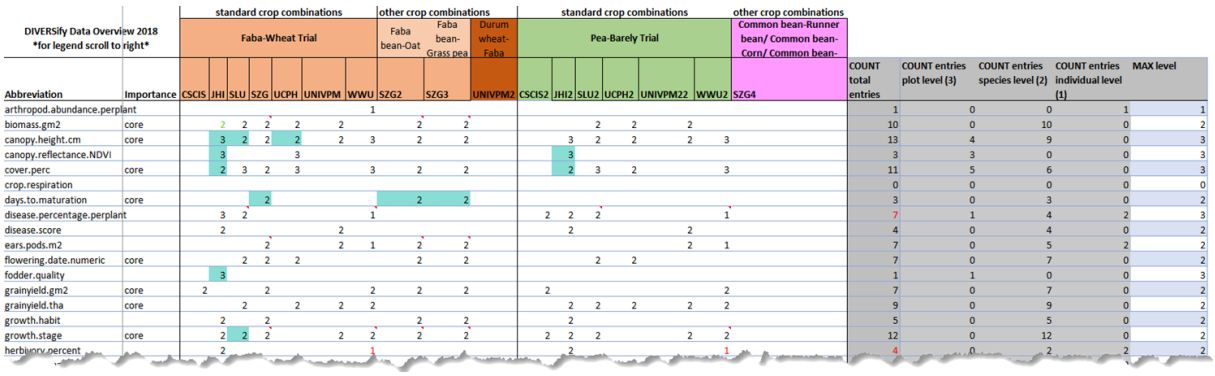
Data generated by partners was reported using a 'Data Template', a standardized Microsoft Excel file designed to accommodate intercropping data. To facilitate dataset handling, data were reported on standardised spreadsheets in a specified format (see Pappagallo et al. 2021), using standardised trait names and units of measurement. This template is not yet available as an open access resource, but it could form part of an open access publication in the future and it has been shared with partners in the ReMIX project (sister project to DIVERSify) as a resource for data collation.

Plot experiments generated a total of 24 datasets for 2017 and 2018 trials. Once these were returned by partners, datasets were first screened for errors, formatting was checked, and partner-specific differences were harmonized, described in Pappagallo et al. (2021). Formatting steps were documented using the open-source software R (version 3.6.3, R Core Team 2020) implemented in RStudio (version 1.3.959). Data formatting mainly required name and descriptors standardisation (especially among variables that were partially overlapping), and screening for evident errors and outliers. During the formatting work, two catalogues were developed. A 'data overview' summarized and kept track of correspondence between partner, trial type and level at which traits were measured in each year (**Figure 1**), whereas a 'trial master' spreadsheet was created to track metadata from all partners and years. These files have also been structured by 'TrialID', which enables linking to the data files for analysis. Datasets were ultimately merged according to sampling level (plot, species and individual level). Aggregated versions of all levels to plot or species level scale may also be created for analyses (Pappagallo et al. 2021).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

This intensive work generated a fully merged version of plot and species level data. In total, 61 traits were collected in year 1 (2017) trials by at least one partner, and 58 traits were collected by at least one partner in year 2 (2018) trials, out of a total of 118 reported traits. Of these, 13 traits were reported by most partners, allowing a good analytic framework to be set up across partner datasets.



| DIVERSify Data Overview 2018 *for legend scroll to right* | | standard crop combinations | | | | | | | other crop combinations | | | standard crop combinations | | | | | | other crop combinations | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Faba-Wheat Trial | | | | | | | Faba bean-Oat | Faba bean-Grass pea | Durum wheat-Faba | Pea-Barely Trial | | | | | | Common bean-Runner bean/ Common bean-Corn/ Common bean- | | | | |
| Abbreviation | Importance | CSCIS | JHI | SLU | SZG | UCPH | UNIVPM | WWU | SZG2 | SZG3 | UNIVPM2 | CSCIS2 | JHI2 | SLU2 | UCPH2 | UNIVPM22 | WWU2 | SZG4 | COUNT total entries | COUNT entries plot level (3) | COUNT entries species level (2) | COUNT entries individual level (1) | MAX level |
| arthropod.abundance.perplant | | | | | | | | 1 | | | | | | | | | | | 1 | 0 | 0 | 1 | 1 |
| biomass.gm2 | core | | 2 | 2 | 2 | 2 | 2 | | 2 | 2 | | | | 2 | 2 | 2 | | | 10 | 0 | 10 | 0 | 2 |
| canopy.height.cm | core | | 3 | 2 | 2 | 2 | 2 | 3 | 2 | 2 | | | 3 | 2 | 2 | 2 | 3 | | 13 | 4 | 9 | 0 | 3 |
| canopy.reflectance.NDVI | | | 3 | | | 3 | | | | | | | 3 | | | | | | 3 | 3 | 0 | 0 | 3 |
| cover.perc | core | | 2 | 3 | 2 | 3 | | 3 | 2 | 2 | | | 2 | 3 | 2 | | 3 | | 11 | 5 | 6 | 0 | 3 |
| crop.respiration | | | | | | | | | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 |
| days.to.maturation | core | | | | 2 | | | | 2 | 2 | | | | | | | | | 3 | 0 | 3 | 0 | 2 |
| disease.percentage.perplant | | | 3 | 2 | | | | 1 | | | | 2 | 2 | 2 | | | 1 | | 7 | 1 | 4 | 2 | 3 |
| disease.score | | | 2 | | | | 2 | | | | | | 2 | | | 2 | | | 4 | 0 | 4 | 0 | 2 |
| ears.pods.m2 | | | | | 2 | | 2 | 1 | 2 | 2 | | | | | | 2 | 1 | | 7 | 0 | 5 | 2 | 2 |
| flowering.date.numeric | core | | | 2 | 2 | 2 | | | 2 | 2 | | | | 2 | 2 | | | | 7 | 0 | 7 | 0 | 2 |
| fodder.quality | | | 3 | | | | | | | | | | | | | | | | 1 | 1 | 0 | 0 | 3 |
| grainyield.gm2 | core | 2 | | | 2 | | | 2 | 2 | 2 | | 2 | | | | | 2 | | 7 | 0 | 7 | 0 | 2 |
| grainyield.tha | core | | | 2 | | 2 | 2 | 2 | | | | | 2 | 2 | 2 | 2 | 2 | | 9 | 0 | 9 | 0 | 2 |
| growth.habit | | | 2 | | 2 | | | | 2 | 2 | | | 2 | | | | | | 5 | 0 | 5 | 0 | 2 |
| growth.stage | core | | 2 | 2 | 2 | | 2 | 2 | 2 | 2 | | 2 | 2 | 2 | | 2 | 2 | | 12 | 0 | 12 | 0 | 2 |
| herbivory.percent | | | 2 | | | | | 1 | | | | | 2 | | | | 1 | | 4 | 0 | 2 | 2 | 2 |

**Figure 1.** Screenshot of one of the data overview files created to provide an overview of which traits were measured by which project partner at which level.

# 4. Overview of trait data

## 4.1. Field trial database

### 4.1.1. Structure

The database contains 97,620 registrations made in a total of 30 combinations of plant teams in field trials carried out across two years in seven European countries: Austria, Denmark, Germany, Italy, Spain, Sweden, and Scotland (**Table 1**). It represents a total of 239 treatment combinations of site, year, input level, crop(s), and seeding rate(s), deriving from 3,732 field plots. The database currently includes a total of 15,173 entries at plot level, and 39,452 entries at species level.

The database consists of three files: Trial metadata, Plot data, and Species data. Each registration on plot or species level is affiliated with date, PlotID and TrialID, allowing the files to be linked.

Each field trial has been assigned a unique trial ID and each physical field plot has been assigned a unique plot ID, allowing tracking and analysis across time and registrations.

Trial metadata describe characteristics of the trial site (e.g. country, name and geographical coordinates), environmental conditions (e.g. weather sums, management/input level and seeding rate), and experiment (e.g. design, replicates and plot size).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

**Table 1**. Intercrops of six legume species: Faba bean (F), Common bean (C), Runner bean (R), Pea (P), Grass pea (G), Lentil (L), and five cereal species: Wheat (W), Barley (B), Oat (O), Maize (M), Sorghum (S).

| Zone | Country | Partner | 2017 | 2018 |
|---|---|---|---|---|
| Nemoral | Sweden | SLU | F-W, P-B | F-W, P-B |
| Atlantic | Denmark | UCPH | F-W, P-B | F-W, F-O, P-B |
| Atlantic | UK | JHI | F-W, P-B | F-W, F-O, P-B, L-B |
| Continental | Austria | SZG | F-W, F-O, F-G, C-M, R-M, C-S, R-S | F-W, F-O, F-T, F-G, C-M, R-M, C-S, R-S |
| Continental | Germany | WWU | F-W, P-B | F-W, P-B |
| Mediterranean | Spain | CSIC | F-W, P-B | F-W, P-B |
| Mediterranean | Italy | UNIVPM | F-W, F-D, P-B | F-D, P-B |

### 4.1.2. Plant teams and varieties

A total of 11 different plant teams are currently included in the database, encompassing combinations of six legume species and five cereal species (**Table 2**).

**Table 2**. Number of included varieties for each crop species.

| Crop species | No. varieties |
|---|---|
| Pea | 14 |
| Faba bean | 9 |
| Common bean | 3 |
| Lentil | 2 |
| Grass pea | 1 |
| Runner bean | 1 |
| Barley | 15 |
| Wheat | 11 |
| Oat | 5 |
| Maize | 2 |
| Durum wheat | 2 |
| Sorghum | 1 |
| Triticale | 1 |

### 4.1.3. Plant traits

The field trial database contains 169 uniquely named traits, describing standard agronomic performance parameters (e.g. harvested grain/biomass yield, seed quality measures, lodging), plant growth and development traits (e.g. tillering, branching, plant height, time to flowering, specific leaf area), and characteristics arising from biotic interactions (e.g. disease incidence, weed cover, herbivore damage) (**Figure 2**). Some of the traits are relevant to all the crop species assessed in the



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

trials, while others are crop species-specific observations, like fungal diseases or insect pests. In total, more than 40 traits or trait groups are found in the database.

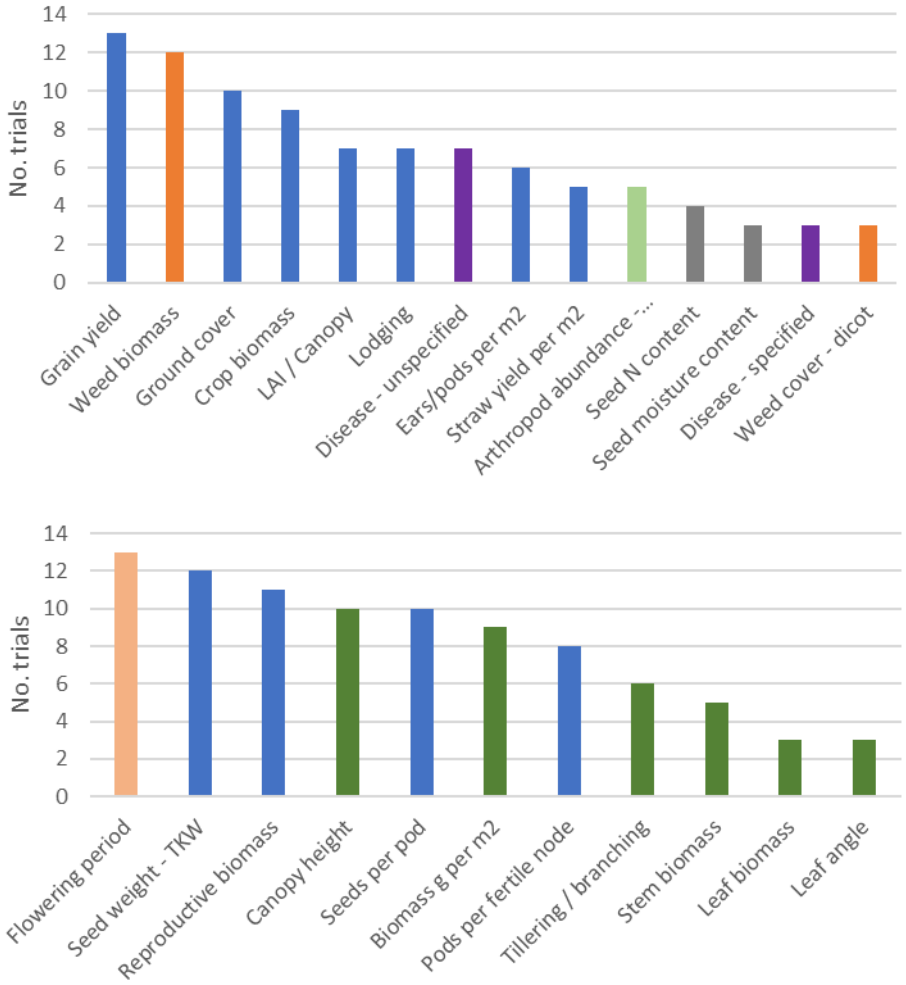



**Figure 2**. Summary of plant traits collected for wheat-faba bean plant teams in the DIVERSify trait database.

## 5. Database storage and accessibility

The trait database is stored on a OneDrive site hosted by WWU/ZFMK, with duplicate copies held at JHI (see below) and UCPH. The original datasets are stored on servers of the partners who generated the data and on a SharePoint site at JHI.

### 5.1. Online repositories

The database is being made publicly available with open access under the Creative Commons license CC BY-NC-SA, via the DIVERSify community at zenodo.org (link), following an embargo period until 31st



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

March 2023 (latest date) to allow time for project publications, although the database will be released sooner if database-associated publications are released quickly.

The database, in whole or part, will be uploaded to other repositories (e.g. dryad and/or PANGAEA) when used to support open access publications, depending on the journal requirement; the uploaded database will be registered in online registries (such as OpenAIRE and re3data.org). A scientific paper is being prepared, providing further details on the structure and querying of the dataset (Pappagallo et al., in prep).

## 5.2. Visualisation

Visualisation is a powerful approach in data exploration, helping to uncover correlations, clustering and other hidden knowledge in data. Hence, DIVERSiplotter was developed as an online tool to allow users to explore and interact with intercropping field trial data collected as part of the scientific field trials of the DIVERSify project (**Figure 3, Figure 4**). This expandable platform is freely available online (https://ics.hutton.ac.uk/diversify/#/). It currently includes integrated data for 94 traits, 11 crop species, and 48 varieties distributed among three intercropped plant teams and seven countries across Europe. Details of the tool development and use are described in Raubach et al. (2021) and Brooker et al. (2021).

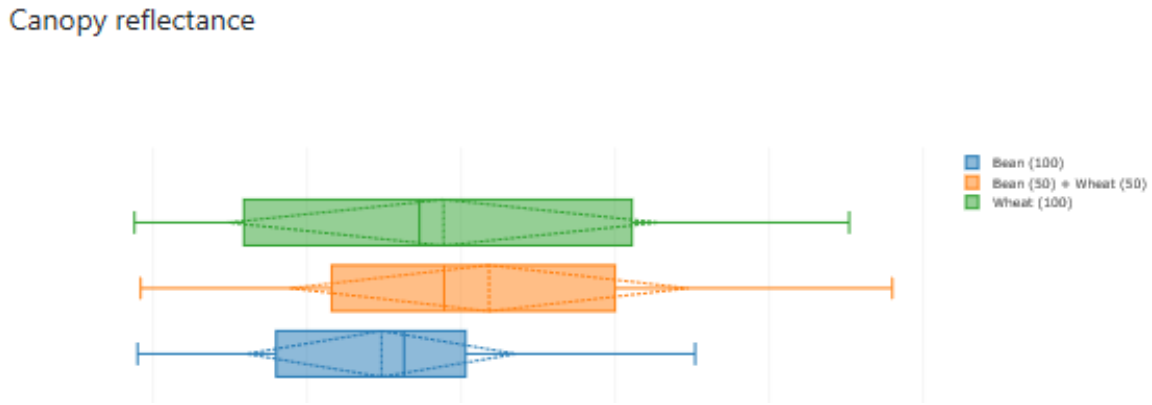



**Figure 3**. Example of species-level trait data visualization using DIVERSiplotter: Distributions of data on canopy reflectance for faba bean and spring wheat monoculture and plant teams.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

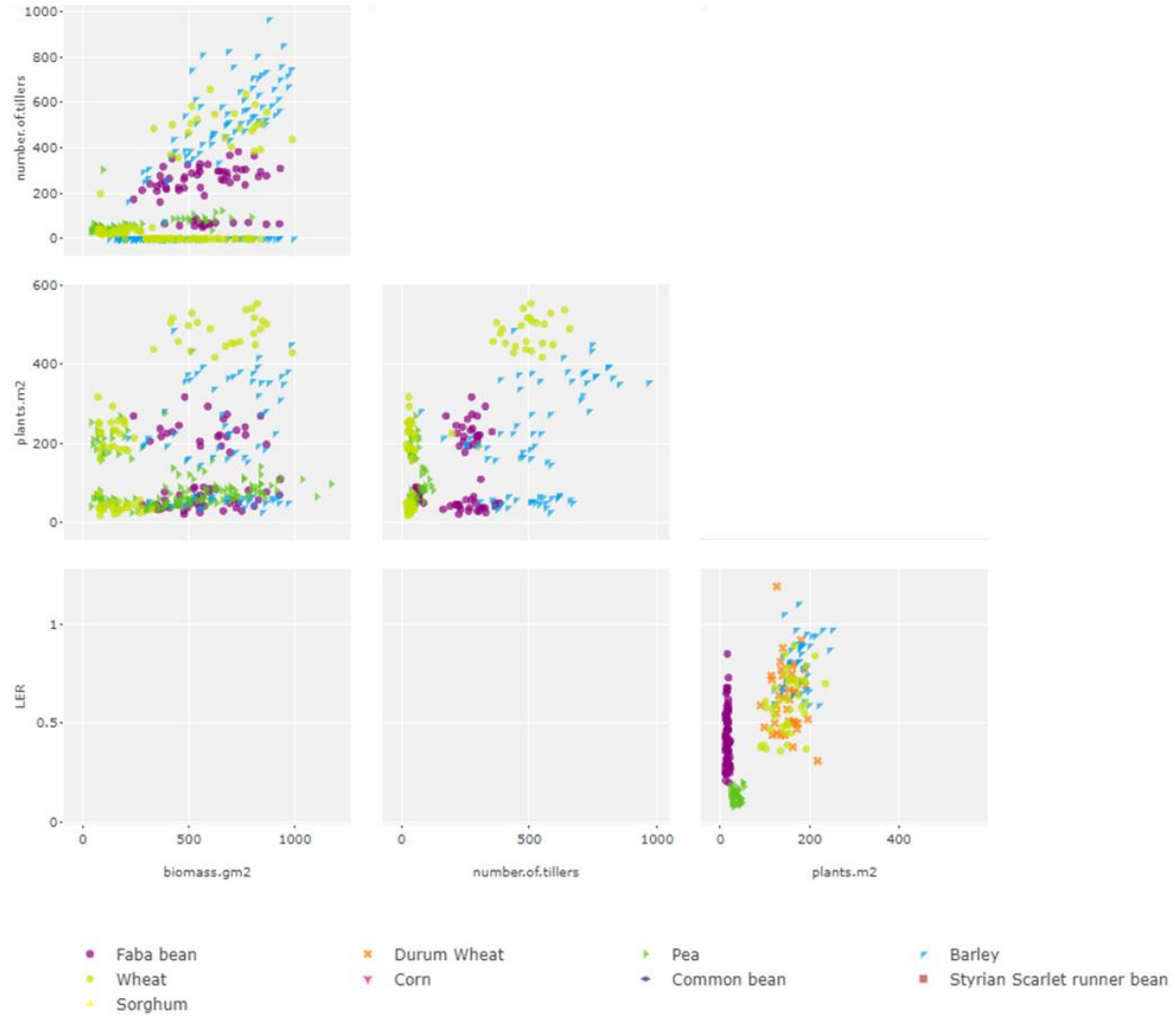



**Figure 4**. Example of individual-level data visualization using DIVERSiplotter: Scatter matrix of four trait variables (plants per $m^2$, number of tillers per plant, biomass in $gm^{-2}$, Land Equivalent Ratio - LER) plotted against each other to highlight correlations, clusters, and outliers in the data.

## 6. Outlook

The crop trait data collected in DIVERSify so far represent a unique collection at unprecedented resolution across European partners and pedo-climatic conditions. Data collected can be used for dissemination activities in scientific manuscripts and recommendations of traits suitable as breeding targets for plant team cropping. The trait database is made openly accessible via the project Zenodo site and other open access repositories. Future developments include incorporating trait data



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

collected from DIVERSify WP2 trials conducted in 2019. The database could also be expanded to include trait data from individual plants/monoculture stands collected in the 'virtual' trait database and from additional phenotyping trials of crop species in monoculture that were conducted in DIVERSify to extend the range of species-level trait data available.

## References

Bocci, G. (2015) TR8: an R package for easily retrieving plant species traits. Methods in Ecology and Evolution, 6(3): 347-350.

Brooker R.W., Pakeman R., Raubach S., Shaw P. (2021). Deliverable 5.5 (D37) Meta-analysis and visualisation tools. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability'), funded by the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement Number 727284

Kiær L., Scherber C., Weih M., Rubiales D., Tavoletti S., Adam E., Vaz Patto M.C., Leitão S.T., Schmutz A., Schöb C., Pakeman R., Newton A.C., Karley A.J. (2020). Deliverable D2.2 (D17) Handbook for trait assessment in agricultural plant teams. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability'), funded by the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement Number 727284.

Pappagallo S., Brandmeier J., Banfield-Zanin J.A., Kiær L., Shaw P., Raubach S., Karley A.J., Scherber C. (2021) Coordinating data collection in intercropping: A feasible example. Aspects of Applied Biology 146, 249-256.

Raubach S., Pappagallo S., Karley A.J., Scherber C., Kiær L., Shaw P. (2021) DIVERSiplotter – A tool for field data visualisation based on the DIVERSify intercropping trials. Aspects of Applied Biology 146, 407-415.

Scherber C. (2017) Deliverable 2.1 (D12) Trait database for plant team species. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability'), funded by the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement Number 727284.

Scherber C., Kiær L., Weih M., Rubiales D., Villegas-Fernandez A., Tavoletti S., Vaz Patto M.C., Adam E., Barradas A., Karley A.J. (2019). Deliverable 2.10 (D19) - Summary report on performance of stakeholder-driven plant teams. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability'), funded by the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement Number 727284.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

## Disclaimer

The information presented here has been thoroughly researched and is believed to be accurate and correct. However, the authors cannot be held legally responsible for any errors. There are no warranties, expressed or implied, made with respect to the information provided. The authors will not be liable for any direct, indirect, special, incidental or consequential damages arising out of the use or inability to use the content of this publication.

## Copyright



## Citation

Please cite this report as follows:

Kiær L., Scherber C., Pappagallo S., Brandmeier J., Weih M., Tavoletti S., Rubiales D., Villegas-Fernández A.M., Adam E., Raubach S., Shaw P., Newton A., Karley A.J. (2021). Deliverable 2.9 (D16) Plant and trait database V3. Developed by the EU-H2020 project DIVERSify ('Designing innovative plant teams for ecosystem resilience and agricultural sustainability'), funded by the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement Number 727284.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

## Appendix 1

Supplementary images of some of the plant teams trials conducted by DIVERSify partners.

**Figure S1.** Plots of the barley/pea intercropping experiment showing monocrops and intercropping (CSIC, April 2018).

**Figure S2.** Plots of the barley/pea intercropping experiment showing monocrops and intercropping (CSIC, April 2018).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

**Figure S3.** Images of the UNIVPM durum wheat-faba bean field trial in April 2018 (top) and in May 2018 (bottom).



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284

**Figure S4.** Barley-pea trial at UNIVPM in April (top) and June (bottom) 2018.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727284