CROPINNO

STEPPING UP SCIENTIFIC EXCELLENCE AND INNOVATION CAPACITY FOR CLIMATE-RESILIENT CROP IMPROVEMENT AND PRODUCTION

Grant Agreement No 101059784

D3.1 – Guidelines and best practices on the use of multi-omics and phenotyping technologies in crop improvement

Funded by the European Union

Lead beneficiary	Forschungszentrum Jülich GmbH (FZJ)
Autor List	Kerstin Nagel
Due date	30/06/2025
Completion date	23/06/2025

Dissemination level		
PU	Public, fully open	X
SEN	Sensitive, limited under the conditions of the Grant Agreement	
Classified R-UE/EU-R	EU RESTRICTED under the Commission Decision No2015/444	
Classified C-UE/EU-C	EU CONFIDENTIAL under the Commission Decision No2015/444	
Classified S-UE/EU-S	EU SECRET under the Commission Decision No2015/444	

Document History

Version	Issue date	Changes
1.0	18/06/2025	First draft by Kerstin Nagel (FZJ)l, Serena Varotto (UNIPD), Renate Horn (UROS), Leire Molinero-Ruiz (CISC) and Dragana Miladinović (IFVCNS)
1.1	23/06/2025	Final version approved by Project Management Committee

Disclaimer

TABLE OF CONTENTS

Contents

Funded by the European Union

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. Grant agreement No. 101059784.

EXECUTIVE SUMMARY

Guidelines and best practices on the use of multi-omics and phenotyping technologies in crop improvement is a deliverable of the CROPINNO project, funded as a HORIZON Coordination and Support Action by the European Commission under its Horizon Europe (HE) Programme. It is produced in the scope of Task 3.1 within Work Package 3: Stepping up excellence and innovation capacity. This document summarizes the guidelines and best practices for crop phenotyping and genotyping for improved resilience as developed, standardized and validated by CROPINNO partners.

Reports on phenotyping and genotyping experiments was drafted by FZJ, which is the leader of WP3, with input from all partners.

1. INTRODUCTION

The aim of task 3.2. was to develop guidelines and share best practices on how the multi-omics and phenotyping tools used in research projects at all participants can be utilized to set up future research projects and the Climate Crop Centre in Serbia. The best-practices at the established phenotyping centre of Jülich Plant Phenotyping Center (JPPC) at FZJ and UNIPD Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), with long record of application of -omics technologies and cooperation with the industry, were collected guidelines, along with insights from University of Rostock and recommendations from CSIC, who is one of the founders of The Centre for Research in Agricultural Genomics (CRAG) that has extensive experience in leading-edge research of the molecular basis of genetic characters of interest in plants and farm animals, and in the applications of molecular approaches for breeding of species important for agriculture and food production.

In the following we report the developed guidelines and summarize the shared best practices on the use of multi-omics (2.1) and phenotyping (2.2.) technologies in crop improvement.

2. DESCRIPTION

2.1 Guidelines and best practices on the use of multi-omics technologies in crop improvement

2.1.1 Best practice examples for development and optimization of SNP-based markers for drought tolerance for use in crop improvement programs

Drought tolerance associated SNP-based markers can originate from DNA or RNA sequencing, or even SNP arrays. To test the markers association with drought tolerance, either a mapping population that segregates for drought tolerance or an association panel with varieties that differ in drought tolerance is necessary. Both, F_2-population and association panel need to be tested for drought tolerance in at least three replications.

When starting the marker development with DNA, the fundament should always be plants that are precisely grouped according to their drought tolerance. With access to a segregating F_2-population, a bulked segregant analysis (BSA) can be conducted on whole genome level (BSAseq). BSA identifies genomic areas associated with a trait. The segregating plants are separated into bulks according to their drought tolerance. The more precise the drought tolerance is measured, the lower the number of individual plants in a bulk can be. To reduce noise for drought tolerance relevant variants only the most and least drought tolerant plants should be picked for the bulks and no plants that are somewhere in the middle in terms of drought tolerance. Pooled DNA samples from the bulks are whole genome sequenced. Genome coverage should be chosen regarding the ploidy of the analysed species, so that every allele is represented by at least 15 reads. The resulting reads are mapped to a reference genome e.g. HanXRQr2.0. If multiple reference genomes are available, the one that is the most related should be picked in its latest version. This again can decrease the number of variants not involved in drought tolerance. For variant calling the program GATK can be used and SNP annotation should be performed with SnpEff. Quantitative trait loci (QTL) associated with drought tolerance can now be identified with programs like e.g. QTLseqr. Variants should be filtered prior to this step. For this QTLseqr has a built-in function. Variants with more than twice the coverage should be discarded, as well as variants with a fourth of the coverage. Besides, variants should have a frequency of 10% or higher. Filter settings should be changed according to the ploidy of the species. The resulting QTL markers, that are unique for one of the bulks, can be considered candidate markers or markers in candidate genes. If available, drought tolerance association panels (size: 200-500 genotypes) should be used to verify the markers as associated with drought tolerance using programs like TASSEL.

Developing markers from RNASeq is not as straightforward as it is with BSAseq, because primarily differences in the expression levels of various genes are detected. However, the RNAseq data can be also used for SNP calling and annotation by applying the programs GATK and SnpEff. However, SNPs in the differentially expressed genes (DEGs) are not per se associated with drought tolerance, but have to be verified with an association panel for drought tolerance.

With knowledge about the location and the flanking sequences, markers for high-throughput screening like KASP markers can be developed. For a SNP to be suitable as a KASP marker, the 3'-end of the forward or reverse primer has to be set on the SNPs location. If possible, no other SNPs should be in the binding region of the primer. Primer design and screening in an association panel can be performed by contracting companies. With higher ploidy level, dosage information should also be considered. A marker on one allele might have a different influence on the trait than a marker present in two or more alleles. But a marker does not necessarily need to have an effect on the trait, as it is more likely that the marker due to linkage disequilibrium is co-expressed on the same allele.

Funded by the European Union

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. Grant agreement No. 101059784.

2.1.2 Best practice examples for development and application of epi-QTLs in breeding

Some examples of guidelines and good practices on the protocols developed and shared with CROPINNO partners are listed for the phenotyping of stressed sunflower plants.

a) Adoption of standardized and repeatable protocols to ensure reproducibility and comparability of results across experiments and institutions.

b) Application of stress treatments under controlled conditions that mimic natural field environments, increasing the ecological relevance of the findings.

c) Use of validated and widely recognized instruments for physiological measurements, coupled with a consistent sampling strategy (e.g., specific leaf positions), to minimize variability.

d) Implementation of time-course analyses to monitor dynamic physiological responses throughout stress induction and recovery phases.

e) Comprehensive documentation of the post-stress recovery phase, enabling the assessment of plant resilience and potential stress memory mechanisms.

f) Integration of physiological measurements with environmental data (e.g., greenhouse temperature records) to enhance the interpretability and robustness of results.

g) Knowledge sharing among project partners, including dissemination of detailed experimental protocols through academic outputs such as student thesis work, promoting transparency and collaborative improvement.

Regarding the protocols for RNA extraction, sequencing and data analysis from sunflower leaves after stress application and recovery the following guidelines and best practices were adopted:

a) Use of validated RNA extraction kits with clearly described protocol modifications.

b) Immediate processing and preservation of RNA integrity through liquid nitrogen grinding.

c) Multi-step quality control: NanoDrop, Bioanalyzer (RIN), and post-sequencing QC.

d) Use of established and peer-reviewed software tools (FastQC, STAR, Samtools, featureCounts, DESeq2).

e) Application of statistical thresholds and enrichment analysis to ensure biological relevance.

Funded by the European Union

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. Grant agreement No. 101059784.

f) Sharing of protocols and analysis pipelines with collaborators to support reproducibility and capacity building.

Similarly, a series of guidelines and good practises have been employed for chromatin extraction and immunoprecipitation (ChIP) protocol development:

a) Integration of multi-omics data by using the same biological material for RNA and chromatin analysis.

b) Immediate sample preservation using liquid nitrogen to maintain chromatin integrity.

c) Adaptation and optimization of standard ChIP protocols for sunflower tissue.

d) Use of validated antibodies and quality control steps throughout the process.

e) Clear focus on reproducibility and standardization of each stage of the workflow.

f) Sharing of detailed protocols within the research network to ensure consistency and foster collaborative progress.

2.2 Guidelines and best practices on the use of phenotyping technologies in crop improvement

2.2.1 Best practice examples for in vitro pre-screening of genotypes for drought tolerance under controlled conditions

For pre-screening genotypes in an in vitro system for drought tolerance there are several factors to consider.

Using an in vitro system, it is especially important to work under sterile conditions, otherwise fungi and other contaminants can lead to a bias in the experiments. To surface sterilize the seeds it is crucial to find the right protocol and to determine the right percentage of solutions to be used. For our experiments, we used the NaClO solution (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) with a few drops of Tween 20. High concentrations of NaClO will negatively affect the germination rate and the seedling growth while too low concentrations might not be enough to ensure complete surface sterilization. Also, it is essential to determine whether the seeds need to be peeled before sterilization. For our experiment, 4% of NaClO solution worked best, some genotypes needed to be peeled before surface sterilization while others had to be surface sterilized with the peel. In case of severe problems with sterility, a new charge of seeds, perhaps from a different year, should be used.

Also, it is crucial to ensure complete sterile conditions during all processes of the experiment (laying out of the seeds and transferring the seeds) under a sterile workbench. All equipment that is being used needs to be sterilized (jars, frames, nets, tweezers). All solutions used need to be autoclaved or filter sterilized. For our experiments all equipment was either autoclaved (121°C, 20 min) or sterilized (160°C, 4 hours).

Funded by the European Union

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. Grant agreement No. 101059784.

In addition, it is important to determine the right substance for drought stress simulation and the right concentration to apply the drought stress. The stress should be strong enough to induce a reaction but not too severe for the plants to completely stop growing or to initiate senescence. We used polyethylene glycol 6000 (PEG 6000) (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) as drought stress simulant. Concentrations between 10% and 20% of PEG 6000 were tested. For us, 15% PEG 6000 in liquid MS-medium (Duchefa, Haarlem, Netherlands) gave the best results to induce moderate drought stress while still allowing the plants to grow. It is necessary to be able to distinguish the differences in the response to drought stress between the different genotypes, so that differences between tolerant and sensitive genotypes are visible.

For evaluation and sampling it is important to be consistent. If sampling is done for RNA extraction and analysis it is essential to freeze the samples in liquid nitrogen as fast as possible and to store them at -80°C until further use. Samples from different rounds of experiments should be taken at the same time during the day (e.g. noon). We prepared different jars for phenotypic evaluation and sampling for RNA extraction to ensure as little manipulation of RNA samples as possible. For the phenotypic evaluation it is important to use the same measuring equipment for all experiments that need to be comparable and to have clear instructions for all staff to ensure the same technique for measurements (length of hypocotyl, leaf area).

With regard to how many genotypes and treatments can be tested in one experiment, the evaluation is the critical point as this part takes the longest time. Before the experiment it should be tested how many plants can be evaluated in a certain amount of time to determine the number of replications, treatments and genotypes that are possible in one round of experiments.

In vitro pre-screening of genotypes for drought tolerance under controlled conditions is a strong tool for a fast, cost-effective and easy method to evaluate genotypes. Several genotypes can be tested in a short amount of time with limited resources.

2.2.2 Best practice examples for high-throughput phenotyping of genotypic variation in root and shoot traits

The guidelines for high throughput plant phenotyping experiments were evaluated and published in a pre-reviewed paper from Poorter et al. (2023). In the following, we focus on the considerations, which are relevant for operating a rhizotron phenotyping system at IFVCNS, to identify genotypic variation in sunflower. The rhizotron system at IFVCNS consists of rhizotrons, which are manually photographed using an imaging box for measuring root traits by manually tracing the roots in the images. Because of the labor-intensive work, experiments must be designed to optimize the use of rhizotrons and labor costs.

Important aspects to be considered are the number of genotypes and number of replicates, the duration of the experiment to quantify root architecture during unhampered growth, i.e. prior to roots touching the sides and bottom of rhizotrons, the timing of measurements with respect to plant development and also diurnal cycle, and how image-based traits can be related to destructive measurements, i.e. which calibration curves should be established.

With respect to the number of replicates per experiment, the desired analysis and outcome should be considered when planning the experiment. For the comparison of genotypes, number of replicates must be high enough to ensure sufficient statistical power for general conclusions. Generally, higher genetic variability within genotypes and lower variation between genotypes call for a higher number of replicates. In contrast, analyses such as genome-wide association studies (GWAS) or Quantitative Trait Loci (QTL) rather benefit from including more genotypes with low replicate number.

Despite the non-invasive measurements of root development by imaging, the frequency of measurements should be carefully determined due to the time-consuming manual processes involved in imaging and image analysis. In the project CROPINNO, advantage was taken from performing sunflower experiments using the automated high throughput phenotyping platform GrowScreen-Rhizo III at FZJ. Based on daily imaging of GrowScreen-Rhizo III, two timepoints were determined as suitable to determine genotypic variation in sunflower root development when performing manual measurements of rhizotrons. First, a measurement during the early plant development should be made to assess the establishment of plants based on the development of the primary root to identify outliers. The second measurement should occur at the end of the experiment to quantify genotypic variation in root traits. Ideally, measurements are made shortly before roots reach the bottom or sides of the rhizotrons, to evaluate unhampered root growth with maximized variation between genotypes.

In order to perform comparison of many genotypes, it may be necessary to conduct a row of experiments. Despite a high degree of standardization in performing experiments, even when growing plants in climate chambers, it is advisable to include at least one reference genotype in each experiment, in order to ensure comparability of results across genotypes, and to be able to perform a normalization of results based on the reference genotype. By repeated evaluation of the reference line, one is also able to assess the repeatability of experiments.

Conducting high throughput phenotyping experiments under controlled conditions are a powerful prerequisite for genotype testing. Comparison of many genotypes under controlled conditions, allows to identify particularly well and rather bad performing genotypes, whose genetic background can be further investigated to identify suitable markers for breeding, and which can be further studied e.g. by destructive measurements and / or physiological measurements, to gain a better understanding for contribution of specific traits to plant performance and ultimately crop yield.

2.2.3 Best practice examples for phenotyping of biotic stress

In CROPINNO, phenotyping of biotic stresses focuses on broomrape or the parasite *O. cumana*. Phenotyping experiments follow protocols and methodologies that are quite developed and established at CSIC. Some of them have been used in several research publications, for instance in Ortiz-Bustos et al. 2016 (DOI: 10.3389/fpls.2016.00884) or Ortiz-Bustos et al. 2017 (DOI: 10.3389/fpls.2017.00833). In this deliverable we make emphasis on the considerations that are relevant for phenotyping of biotic stress under greenhouse conditions.

In CSIC facilities, space is the limiting factor in the design of phenotyping experiments, since the greenhouse is shared between different research groups. There is a reservation system that assigns use, in terms of space (square meters) and time slot, to those wishing to conduct experiments in the greenhouse. Therefore, experiments must be designed to optimize space available as related to number of genotypes, number of *O. cumana* populations and number of replicates (individual plants) to be included in the phenotyping. Generally, several experiments having similar design are planned subsequently in time.

The timing of measurements is a crucial aspect for non-invasive phenotyping. Since target of the sensors are plant leaves, these must be large enough to offer an area that allows effective measurements. Besides, infections by broomrape are clearly visible when broomrape stems emerge from the soil. Thus, there are only some weeks of time window to take measurements: between week 3 and week 9 after inoculation, when leaves have a sufficiently large size and first broomrape stems are visible aboveground, respectively.

Concerning number of replicates, the genetic background of plant material, i.e. source of resistance to biotic stress and how much “fixed” that trait of resistance is in the genotype, should be considered when planning the experiment. In general, the higher fixation level, the low number of biological replicates will be necessary. It is also important to note that non-inoculated control treatments should be included as reference. These allow the confirmation of the correct progress of inoculation with the parasite *O. cumana*, and provide the information that corresponds to the biological inherent behaviour of the plant material. Having the non-inoculated plants as reference, different outcomes are possible, like the identification of particularly well (resistant) or bad (susceptible) performing genotypes.

Despite the use of non-invasive methods to detect plants’ signals, these methods are still manually operated by using hand-held devices. Thus, they are time-consuming and the time needed to perform measurements should be carefully considered. In some occasions, the work of two persons can be necessary in order to simultaneously hold the device, handle plant’s leaves and record measurements. At the point of conducting measurements, it is advisable to always follow the same order of plants and recording all incidences that might happen (breaking of a leaf, confusion of replications, etc.).

Finally, data pre-processing, synchronization and identification are crucial. Both, measurements and data management, are advisable to be conducted by the same operator along the time and for each experiment.

Funded by the European Union

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them. Grant agreement No. 101059784.