**WP3: transPLANT – Milestone 7 Report**

This report synthetises the work done on WP3 in Transplant focused on Milestone 7

Contributors to this work :

Paweł Krajewski, Hanna Ćwiek (IPG PAS), Paul Kersey (EBI), Cyril Pommier (INRA), Matthias Lange (IPK), Delphine Steinbach (INRA)

Transplant project needs **the definition of a set of mandatory or optional fields associated with ontology formalism for descriptors**. T

he aim is to construct standards and data formats for phenotypic data exchange. To achieve this goal, we need ontologies and we need to work on data consistency, data scope definition, standard format for data exchange and data repository identification.

Concerning ontologies, we decided to capitalize on existing work at international level, conducted within the PhenotypeRCN coordination network. This RCN gathers reference of international plant trait ontologies and particularly the Plant Ontology Consortium, which coordinates Plant Ontology and Gramene Trait Ontology and the Crop Ontology coordinated by Bioversity within the Generation Challenge Program. Several members of the Transplant project are involved in this group, namely Paul Kersey, Pawel Krajewski and Cyril Pommier.
This ontology community is building a Reference Plant Trait Ontology. It will serve as a foundation to a collection of specific ontologies which will cross reference the reference ontology to ensure consistency. Those specific ontologies will therefore be much easier to maintain by their community, for instance for a species group.

For the repository identification aspect, we will capitalize on INRA-URGI’s Information System, GnpIS. GnpIS, thanks to the Ephesis Genotype X Environment project, is able to integrate various and numerous phenotypic data experiments. GnpIS EPHESIS is already chosen today by INRA to be a strategic and long term repository of elaborated phenotypes. It will be used in the frame of the starting French project called PHENOME, involving a lot of phenotypic facilities that will submit their data at central level in EPHESIS. The first tests done with data submitted by INRA URGI collaborators and also by TransPlant partner (Pawel) this year, were successful. GnpIS EPHESIS is indeed able to store phenotyping experiments and can also be used to store Transplant’s Phenotypic experiments, but without the raw data. Raw data represent a huge amount of data, hundreds if not thousands of Terabytes and should be dealt with outside of Transplant. That is why, it was decided in Transplant to set up a network of Phenotypic Experiment Repositories and GnpIS Ephesis will be the first node.

Standard format will serve two purposes: data submission in the repositories and data exchange between repositories and software from WP10 and WP6. The existing data exchange format used for GnpIS Ephesis will serve as a basis in the work but the aim is now to build a real standard. We will follow the “Minimal Information” (MI) approach as represented by MIBBI website, approach already applied for expression data to describe expression data that conducts to the development of the MIAME standard. Our aim in the workpackage, is to set up a Minimal Information About Plant Phenotypic Experiment (MIAPPE) that will present requirements necessary to describe any experiment with plants, in which any phenotypic (transcription levels, protein levels, metabolite levels, morphology traits, image phenotyping) traits are collected. Our aim is also to develop web services from theses formats, to ensure a distributed architecture and a distribution between data storage and data processing.

Concerning data scope definition, phenotypic experiment can commonly be used as sample source for transcriptome, proteome or metabolome experiments. Furtheremore, those raw transcriptome data can serve as a basis to produce elaborated Traits. The file format and repositories we are setting up must therefore only store elaborated transcriptome Trait, and leave raw transcriptome data to their dedicated file format, MIAME and repositories like Array Express.

To go on interaction at international level, we will participate at a meeting in Oregon in autumn. Presentations will be done, one to present TransPlant project and the WP3 workpackage (Paweł) and one on needs and solutions in the area of ontologies as seen by INRA in the context of WP3 (Cyril).