



Thaliadb : A database dedicated to association genetic in plants

Yannick de Oliveira, Guy-Ross Assoumou-Ella, Alain Charcosset, Johann Joets

► To cite this version:

Yannick de Oliveira, Guy-Ross Assoumou-Ella, Alain Charcosset, Johann Joets. Thaliadb : A database dedicated to association genetic in plants. European Conference on Computational Biology 2014, Sep 2014, Strasbourg, France. <hal-02417594>

HAL Id: hal-02417594

<https://hal.science/hal-02417594v1>

Submitted on 18 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

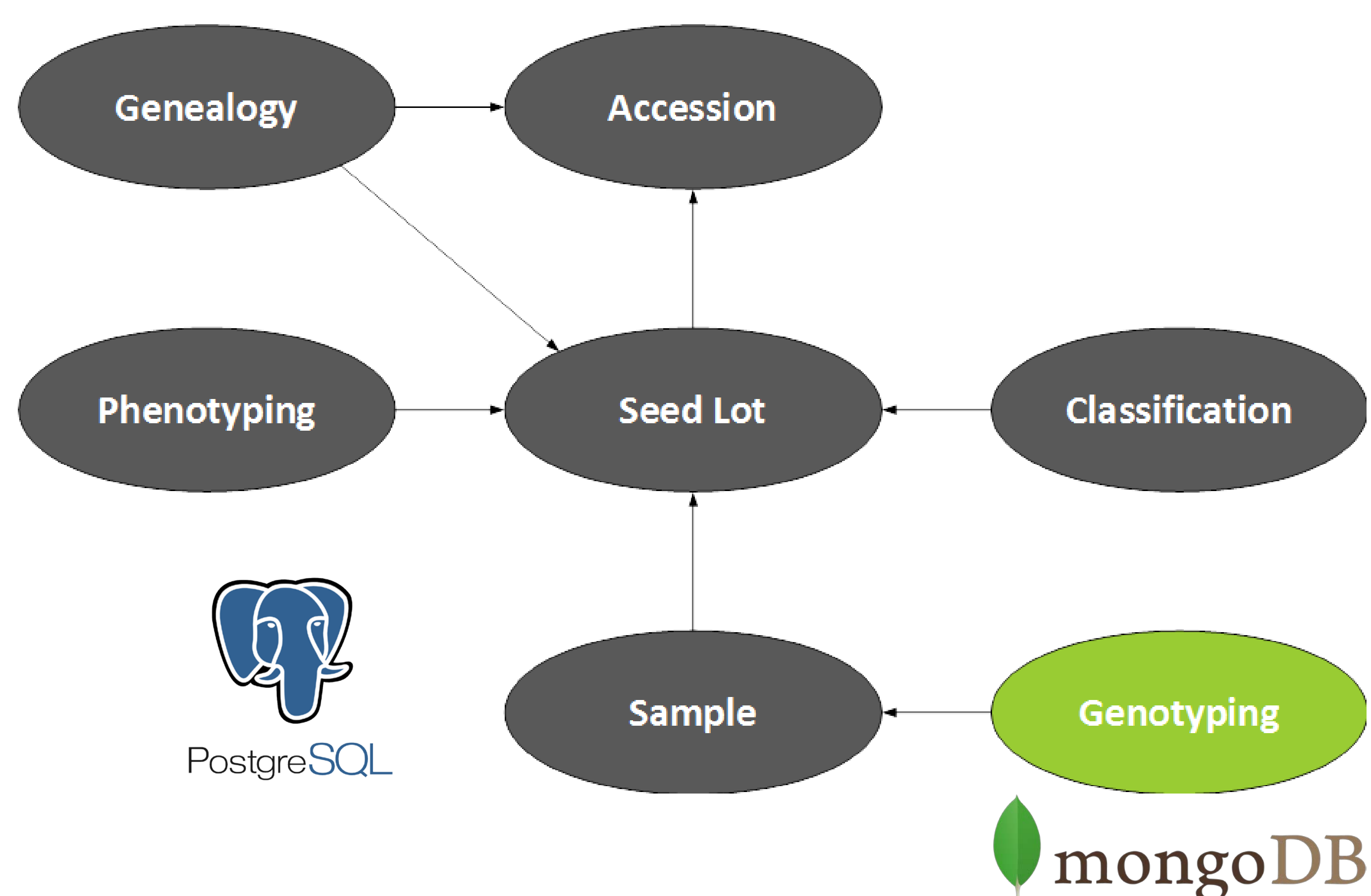
Thaliadb : A database dedicated to association genetic in plants

Yannick de Oliveira, Guy-Ross Assoumou-Ella, Alain Charcosset and Johann Joets
INRA, UMR 8120 Génétique Végétale, F-91190 Gif-sur-Yvette, France

Diversity and association genetics studies lead to manipulate a large number of individual, lines, clones and/or populations. Moreover, emergence of high-throughput technologies for both genotyping and phenotyping generates a large amount of data. These need to be stored and managed in order to perform requests and organize datasets to conduct association genetics studies.

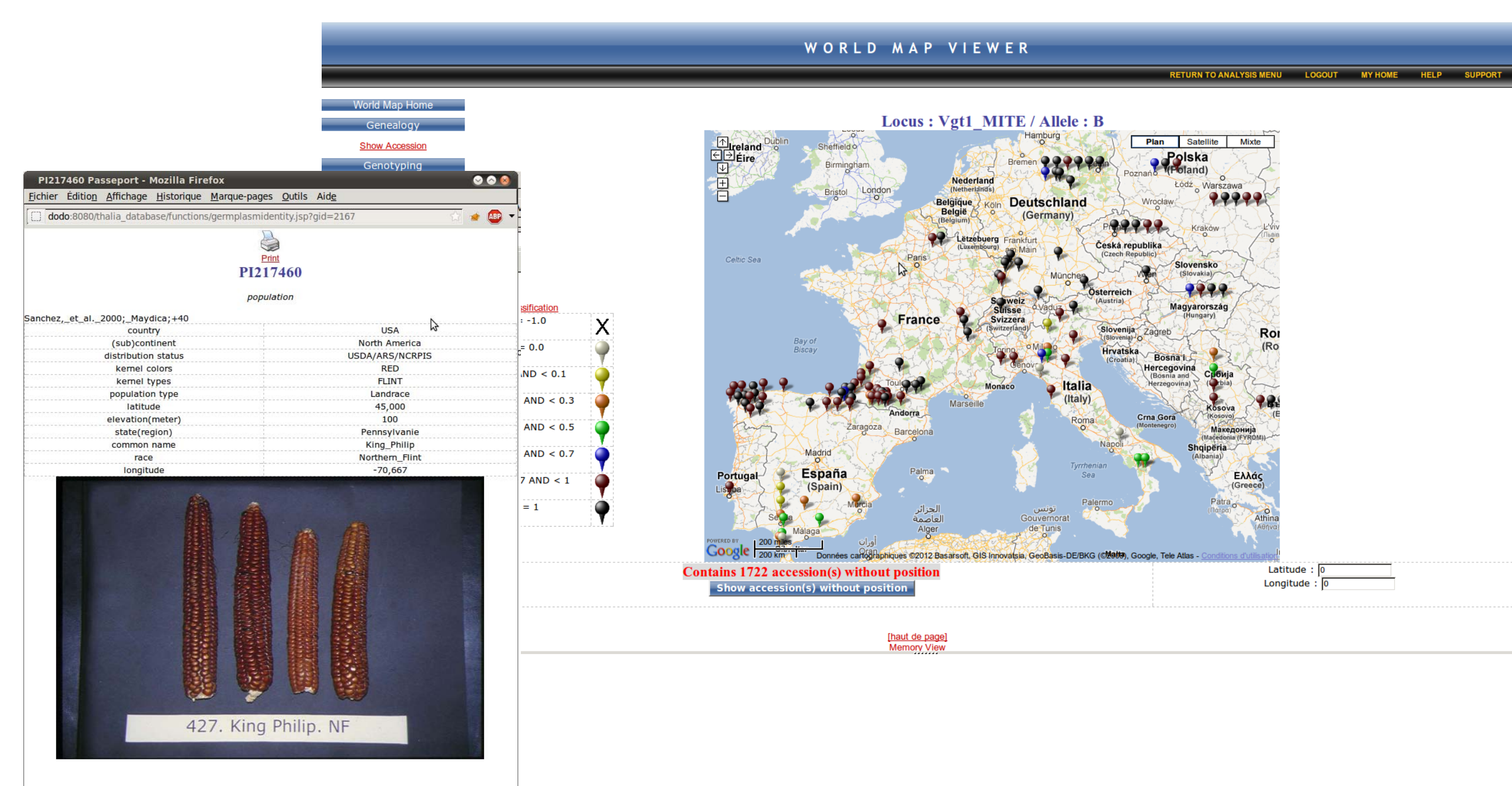
The Thaliadb database has been developed with this aim. It manages genetic resources, phenotyping and genotyping data, and also population structure information. Thalia enables data extraction in formats used by genetic association software.

Database structure



Thaliadb has been developed using the Postgresql database management system. Recently, we migrate a part of the database in the MongoDB system. Thus, performances has been improved and Thaliadb is now able to manage the data produced by high-throughput technologies.

Data export and visualization



Seed lot	Vgt1 MITE A	Vgt1 MITE B	Days to pollen	Northern flint	Andes	Caribbean	Corn belt dent	Europe	Mexican	Italian
PPS48	0,00	1,00	57,9	0,9140	0,0104	0,0088	0,0092	0,0162	0,0106	0,0308
P1217460	0,21	0,79	72,3	0,9156	0,0056	0,0088	0,0412	0,0084	0,0096	0,0110
P1218143	0,52	0,48	85,8	0,0852	0,0178	0,0164	0,6516	0,0178	0,2016	0,0088
P1213719	0,80	0,20	81,8	0,0810	0,0136	0,0268	0,8024	0,0150	0,0204	0,0410
PPS774	1,00	0,00	77,2	0,0090	0,0664	0,2494	0,0174	0,0410	0,0898	0,5274

Main projects involving Thaliadb



A. Dereeper & al.

The aim of GnpAsso project was to develop generic tools to manage and exploit results from association genetics studies. Thaliadb is integrated with the SNIPlay pipeline to process genotype datasets. New functionalities has been developed in Thaliadb to export data to GnpAsso database.



F. Tardieu & al.

DROPS aims at developing novel methods and strategies for genetic yield improvement under dry environments and for enhanced plant water-use efficiency.

In this project, we are developing a web service to allow user request Thaliadb through DROPS application



A. Charcosset & al.

Amaizing aims to identify markers and candidate genes responsible for traits of agronomic interest such as yield, quality and tolerance to abiotic stress. The project aims also to develop breeding methods for creation of new varieties of improved maize. The web service of Thaliadb and the functionalities dedicated to high-throughput data will be improved