



**HAL**  
open science

## **SHiNeMaS : a database dedicated to seed lots genealogy, phenotyping and cultural practices.**

Yannick de Oliveira, Laura Burlot, Marie Lefebvre, Darkawi Madi, Pierre  
Riviere, Mathieu Thomas

### ► **To cite this version:**

Yannick de Oliveira, Laura Burlot, Marie Lefebvre, Darkawi Madi, Pierre Riviere, et al.. SHiNeMaS : a database dedicated to seed lots genealogy, phenotyping and cultural practices.. Journées Ouvertes en Biologie, Informatique & Mathématiques 2015, Jul 2015, Clermont-Ferrand, France. hal-02417547

**HAL Id: hal-02417547**

**<https://hal.science/hal-02417547>**

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.

# SHiNeMaS : a database dedicated to seed lots genealogy, phenotyping and cultural practices

Yannick De Oliveira<sup>1</sup>, Laura Burlot<sup>1,2</sup>, Marie Lefebvre<sup>1</sup>, Darkawi Madi<sup>1</sup>, Pierre Rivière<sup>1,3</sup> and Mathieu Thomas<sup>1</sup>

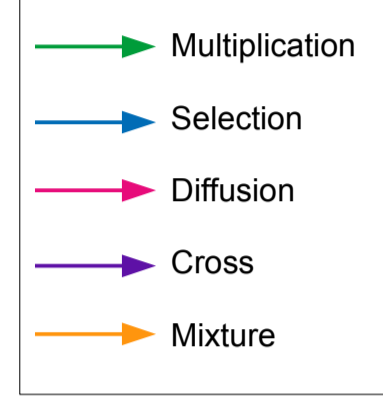
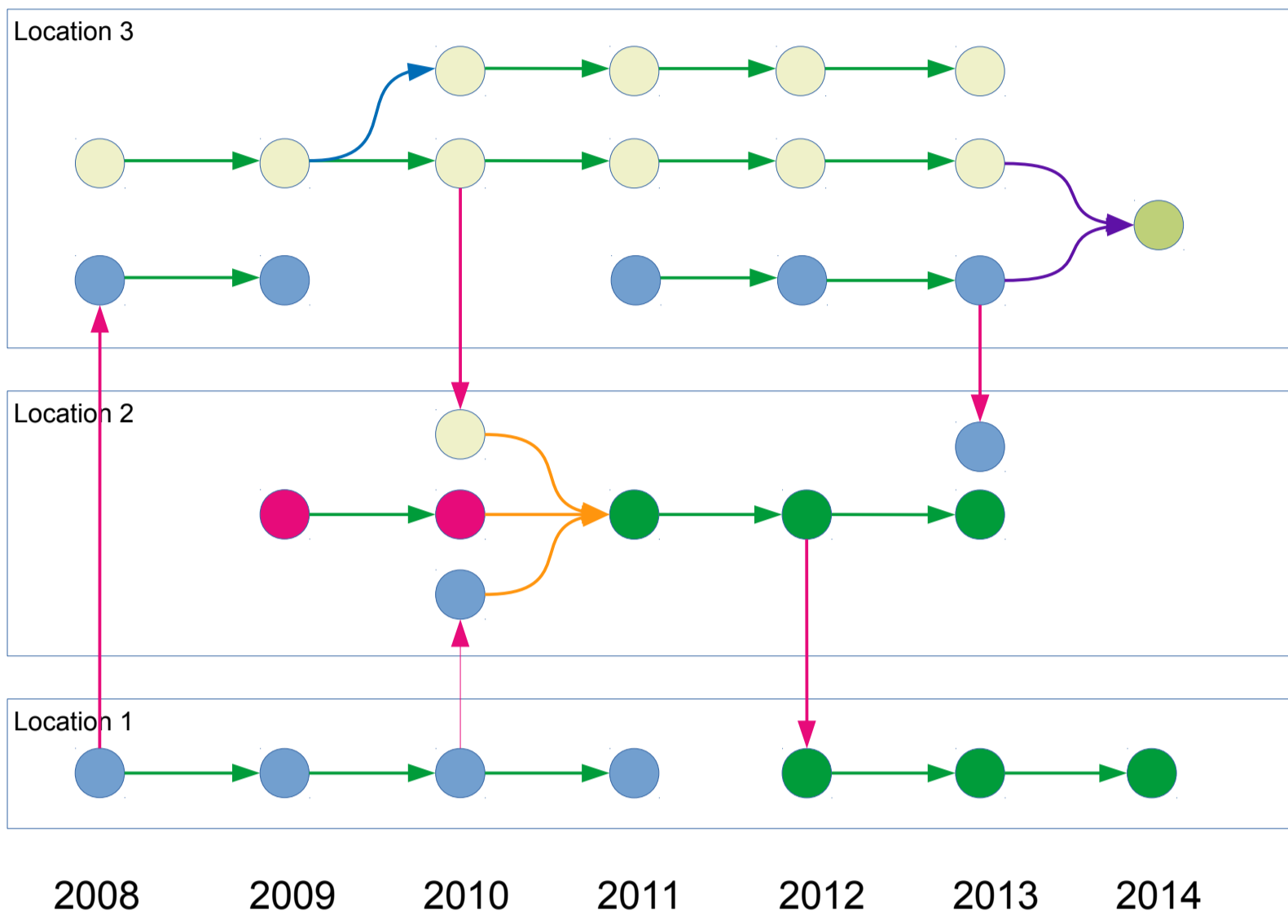
<sup>1</sup> INRA, UMR Génétique Quantitative et Evolution - Le Moulon, 91190, Gif-sur-Yvette, France

<sup>2</sup> AgroBioPerigord, 20 rue du Vélodrome, 24000, Périgueux, France

<sup>3</sup> Réseau Semences Paysannes, 3 avenue de la Gare, 47190, Aiguillon, France

In 2005, a collaboration started between the French National Institute for Agricultural Research (INRA) and the farmer organization Réseau Semences Paysannes (RSP). The aim was: (1) to study on-farm management of crop diversity, cultural and selection practices; (2) to develop population-varieties adapted to organic and low inputs agricultures in the context of a participatory plant breeding program involving farmers, NGOs' facilitators and researchers. In these projects, it is needed to map the life cycle of seed lots using the network formalism. To reach this aim, we developed SHiNeMaS (Seeds History and Network Management System) a database with its web interface, dedicated to the management of the history of seed lots and the associated data.

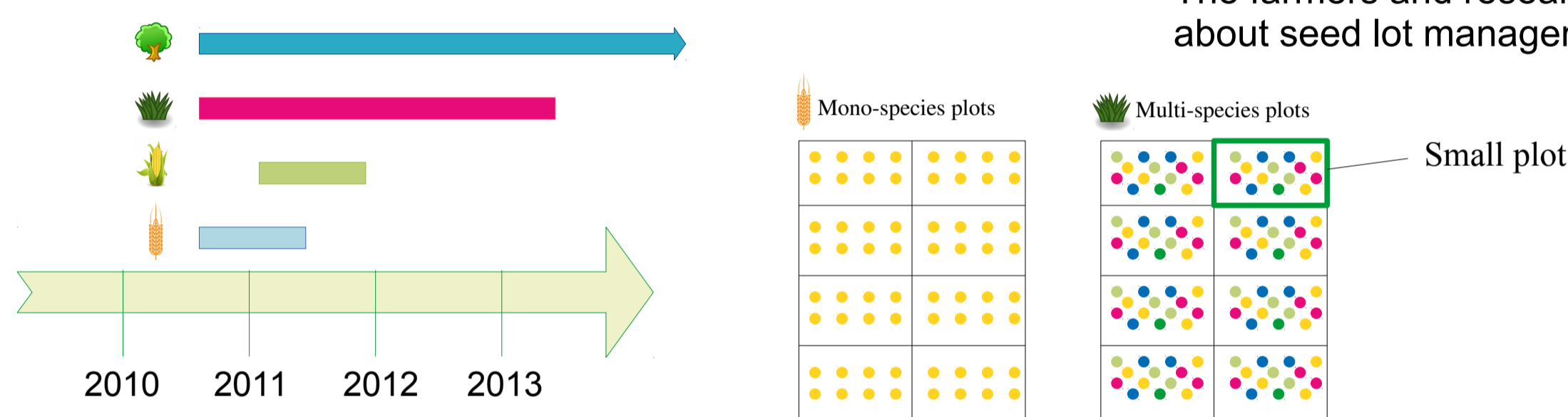
## Participatory plant breeding programs



The graph on the left shows a classic network coming from a participatory plant breeding program. Each circle represents a seed lot, and a same color for a seed lot points out a same population-variety.

Such networks aims to provide answers about :  
- Which population-variety are the most diffused ?  
- Which population-variety are conserved from one year to another ?  
- Do farmers select same varieties ?  
...

The farmers and researchers aim also to get information about seed lot management (stock and location).



SHiNeMaS is designed to support multiple species, in the same instance of the database, dealing with different life cycles of plants from annual to perennial. The field trials are located in farms or in experimental stations and split into small plots. A mixture of species can be sown in a single small plot, SHiNeMaS can therefore deal with species association like prairie trials.

SHiNeMaS is also designed to collect data dealing with cultural practices and phenotyping. The user can define dynamically the variables he/she wants to measure. The data can be linked at the small plot, the seed lot or at the individual level. Each variable is measured according to a dedicated method defined by the user.

## Data organization & technologies

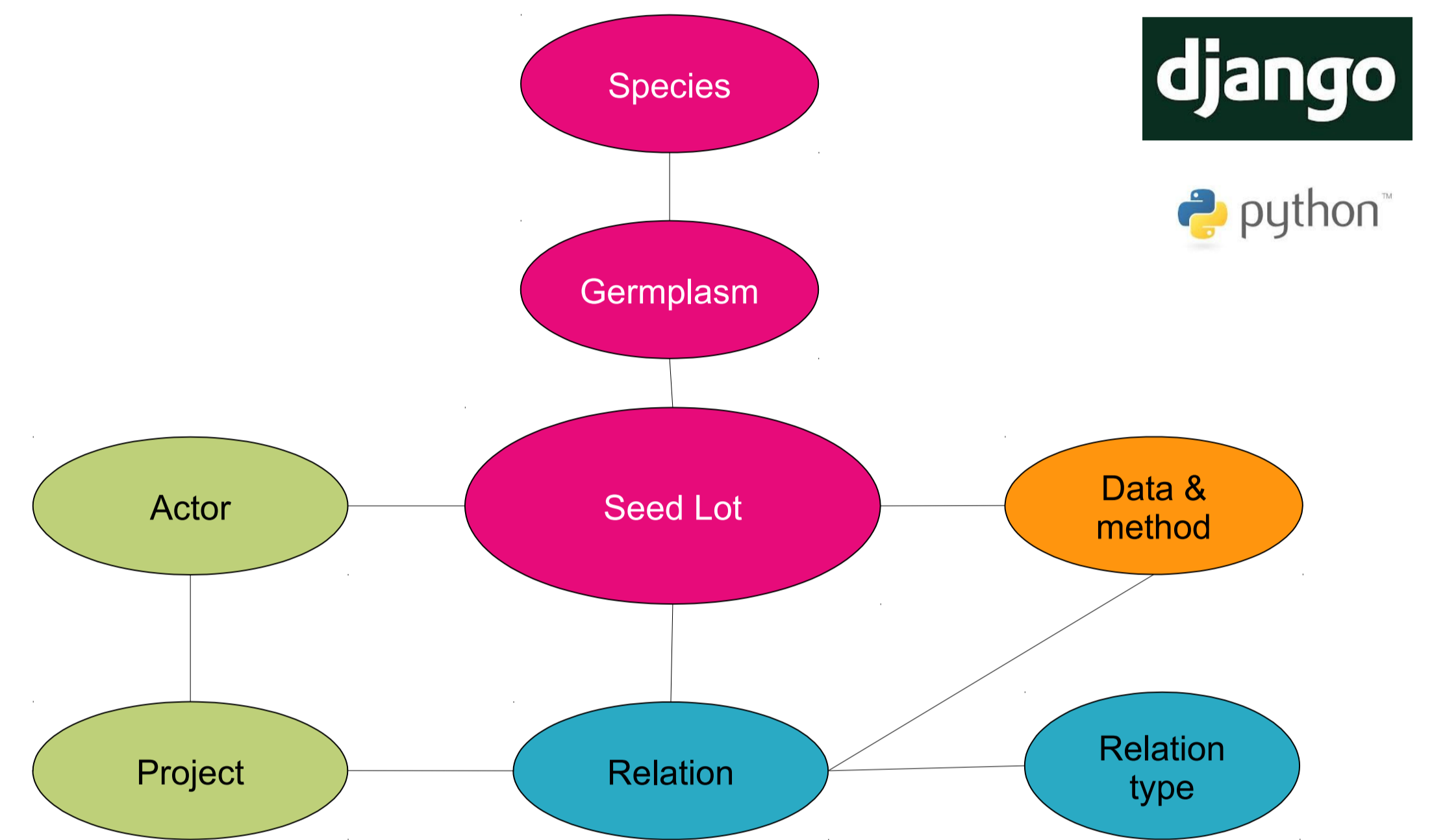
The figure below is a simplified representation of the database schema of SHiNeMaS. The schema is centered around the **Seed Lot**. A seed lot is characterized by a **Germplasm** which is characterized by a **Species**. Users can describe dynamically germplasm's type like cross, mixture, local variety, inbred lines, etc.

**Actors** describes the people involved in the seed lot management. They can be farmers, researchers, or whoever is implied in an action of diffusion, selection, mixture, etc. All the data are organized in **Projects**.

The **Relation** describes the genealogical bonds between **Seed Lots**. Basically a relation implies a parent seed lot and a child seed lot. A **Relation** is linked to a **Project**. Each **Relation** refers to a **Relation type** among multiplication, diffusion, selection, mixture.

All the **Data & Method** collected are stored in the same part of the schema. Each **Data & Method** are linked to a **Relation** or a **Seed Lot**.

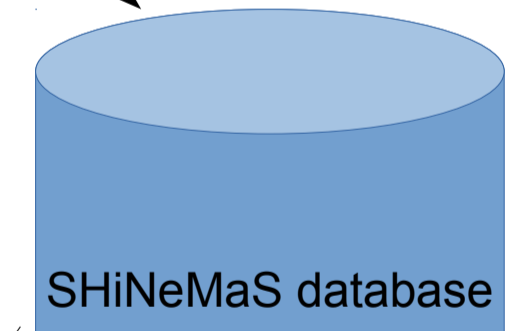
SHiNeMaS is developed using the python technologie with django framework. The database management system used is PostgreSQL but SHiNeMaS should be able to run with MySQL, Oracle or SQLite



## Web interfaces

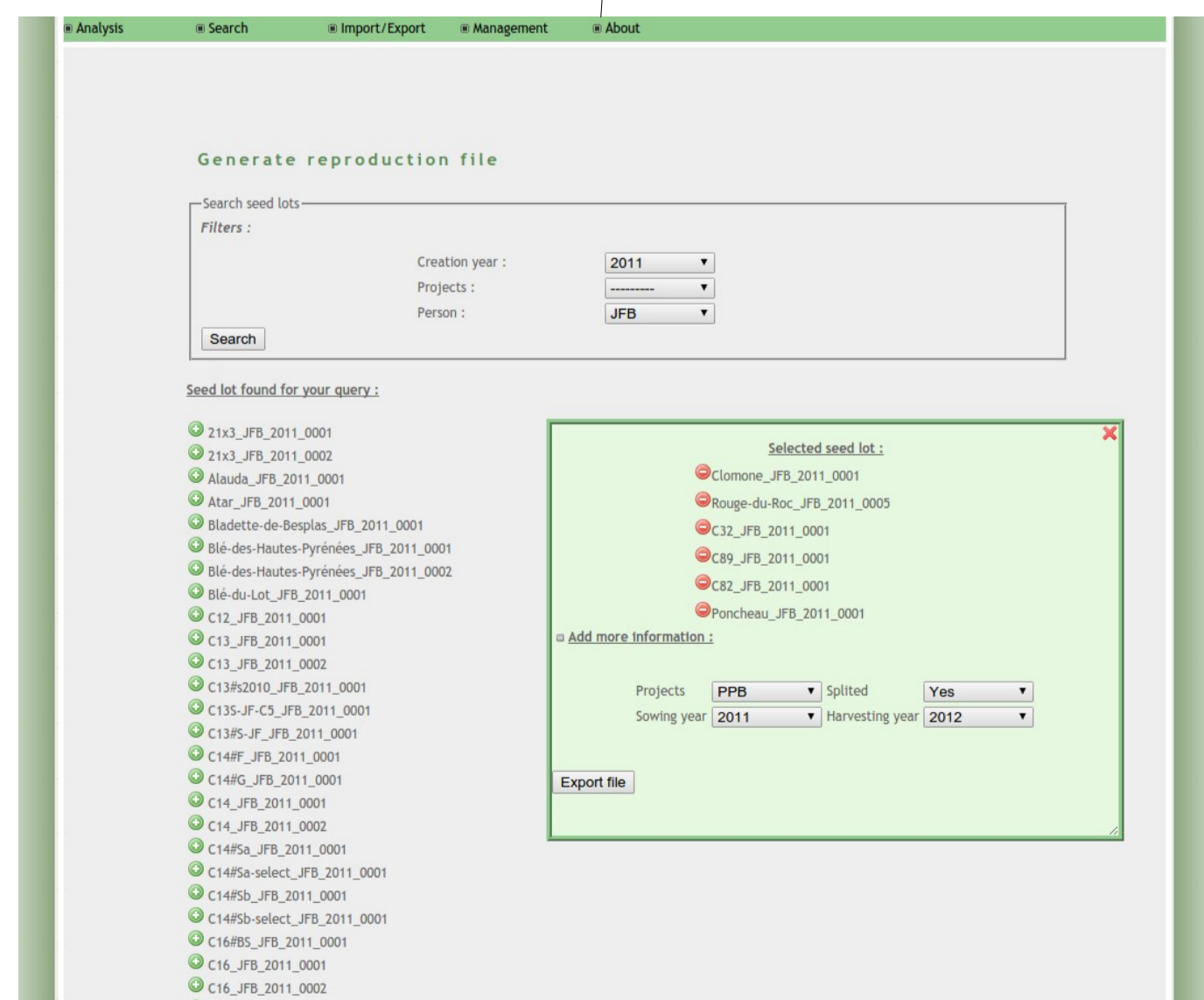
CSV files containing raw data

1- SHiNeMaS provides interfaces to massively load data using CSV formatted files

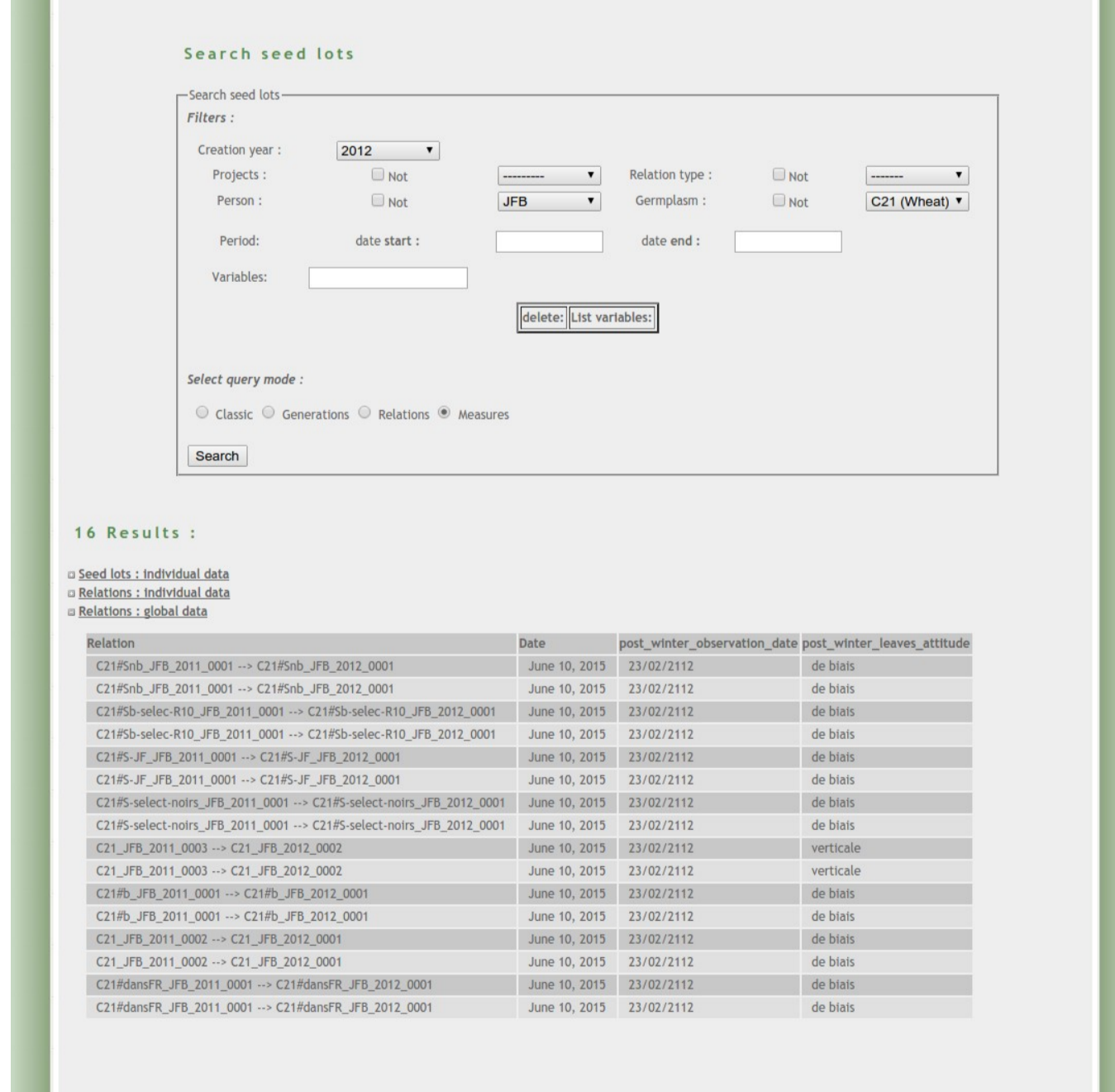


2- SHiNeMaS provides assistant interfaces to prepare files with seed lots already existing in the database

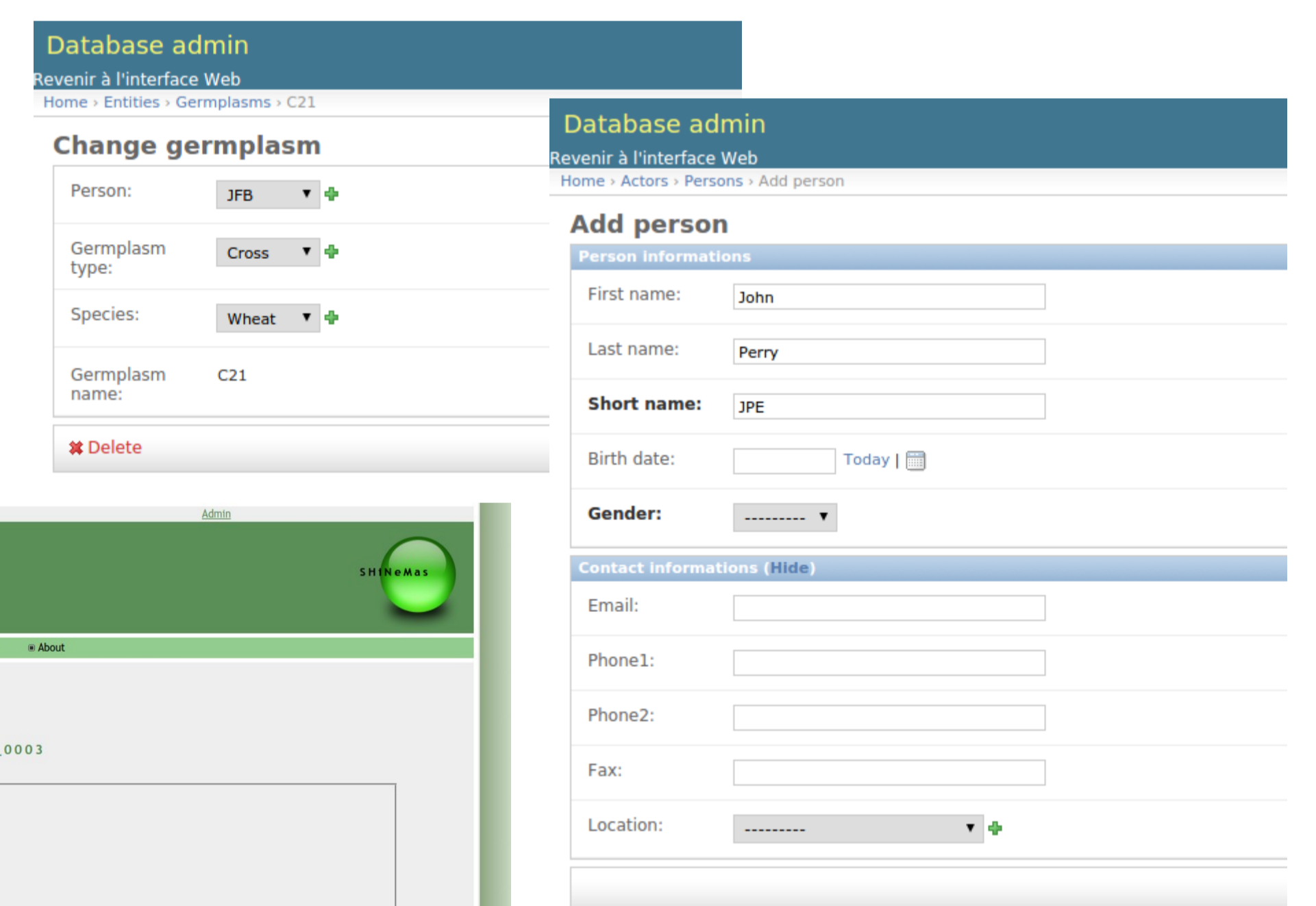
3- The user can export this prefilled file and finish to fill the remaining fields



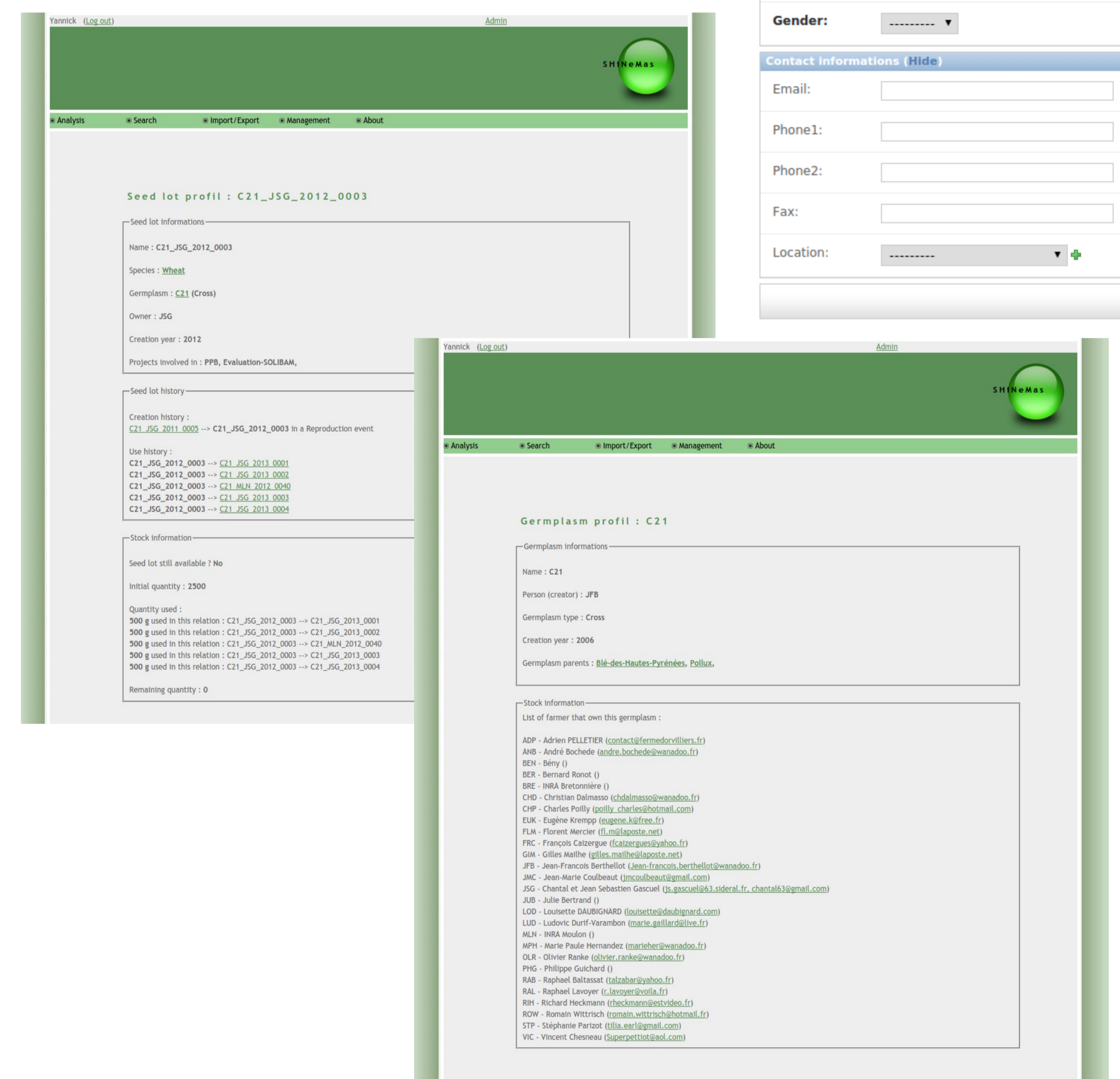
Query interfaces



Administration interfaces



Profile interfaces



## Perspectives and next developments

An instance of SHiNeMaS is in production at the GQE laboratory (<http://moulon.inra.fr>), managing around 700 germplasm, 10 000 seed lots, 10 000 relations and more than 1 000 000 records of collected data on the network.

SHiNeMaS will be freely available under the AGPL licence (<http://www.gnu.org/licenses/agpl-3.0.fr.html>). We aim to distribute a first stable and packaged version late in 2015.

For the next developments we plan to implement interfaces to manage single data, events, seed lots etc. We aim also to improve query interfaces, management of user rights, quality control during data loading. New data types need to be managed like climatic data and transplantation of trees. Data analysis features will be developed, integrating the R package (shinemas2R) that is under development.

**Acknowledgments** : Isabelle Goldringer, Sophie Jouanne-Pin and the DEAP team. Johann Joets from the ABI team.

**Web Site** : <http://moulon.inra.fr/index.php/en/traverse-team/atelier-de-bioinformatique/projects/181>