



HAL
open science

SHiNeMaS: A database and its web interfaces dedicated to seed lots history, phenotyping and cultural practices

Yannick de Oliveira, Laura Burlot, Isabelle Goldringer, Darkawi Madi, Pierre Riviere, Delphine Steinbach, Gaëlle Van Frank, Mathieu Thomas

► To cite this version:

Yannick de Oliveira, Laura Burlot, Isabelle Goldringer, Darkawi Madi, Pierre Riviere, et al.. SHiNeMaS: A database and its web interfaces dedicated to seed lots history, phenotyping and cultural practices. XXVIIIth Plant and Animal Genome Conference (PAG 2019), Jan 2019, San Diego, CA, United States. . hal-02417574

HAL Id: hal-02417574

<https://hal.science/hal-02417574v1>

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.

Yannick De Oliveira¹, Laura Burlot^{1,3}, Isabelle Goldringer¹, Darkawi Madi¹, Pierre Rivière^{1,2}, Delphine Steinbach¹, Gaëlle Van Frank¹ and Mathieu Thomas⁴

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

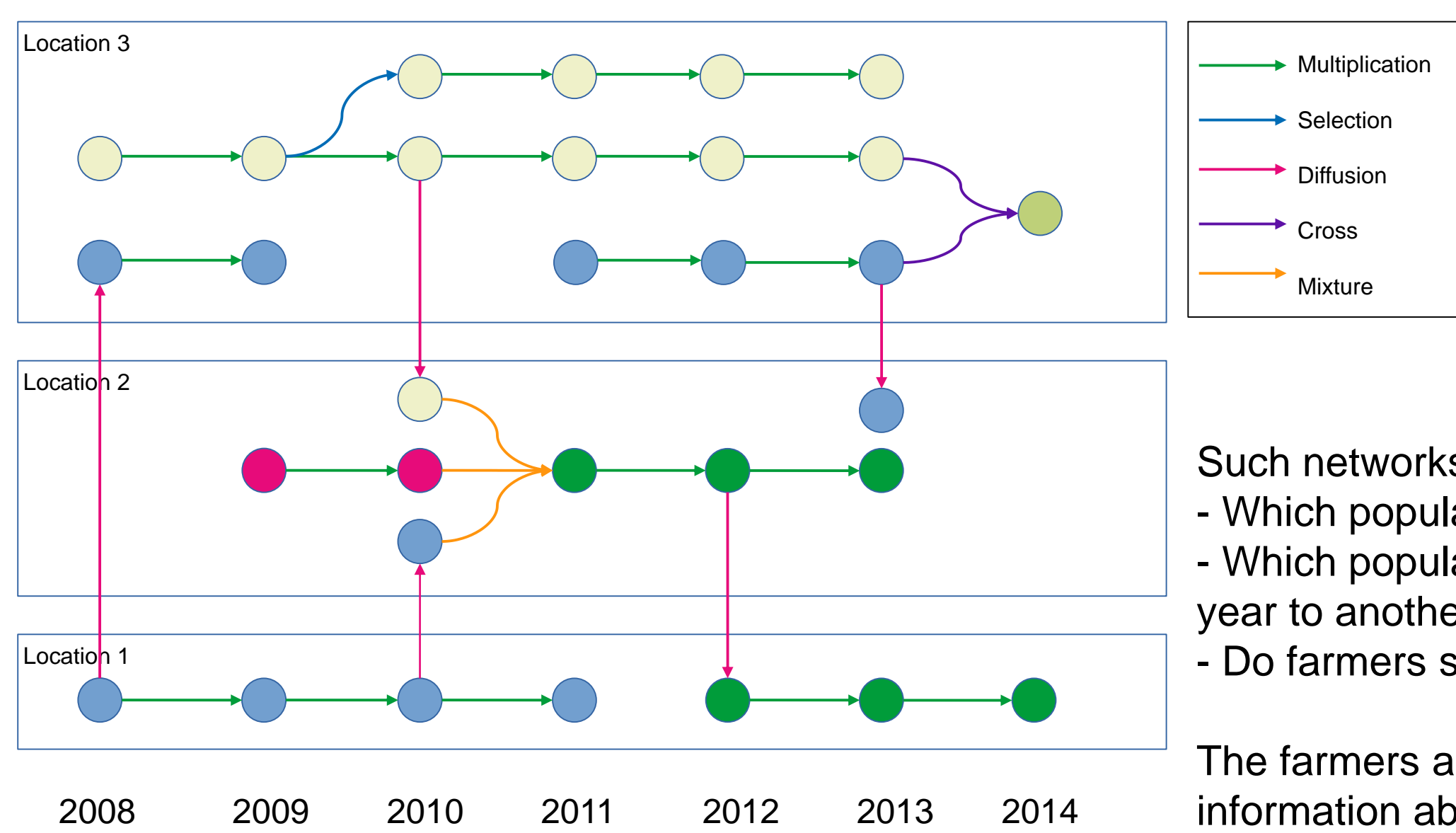
² Réseau Semences Paysannes, 10 Place Clémenceau, 47190, Aiguillon, France

³ AgroBioPérigord, 7, impasse de la Truffe, 24430, Coursac, France

⁴ CIRAD, UMR AGAP, Avenue Agropolis, 34398 Montpellier Cedex 5, 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 (**S**eeds **H**istory and **N**etwork **M**anagement **S**ystem) a database with its web interface, dedicated to the management of the history of seed lots and the associated data.

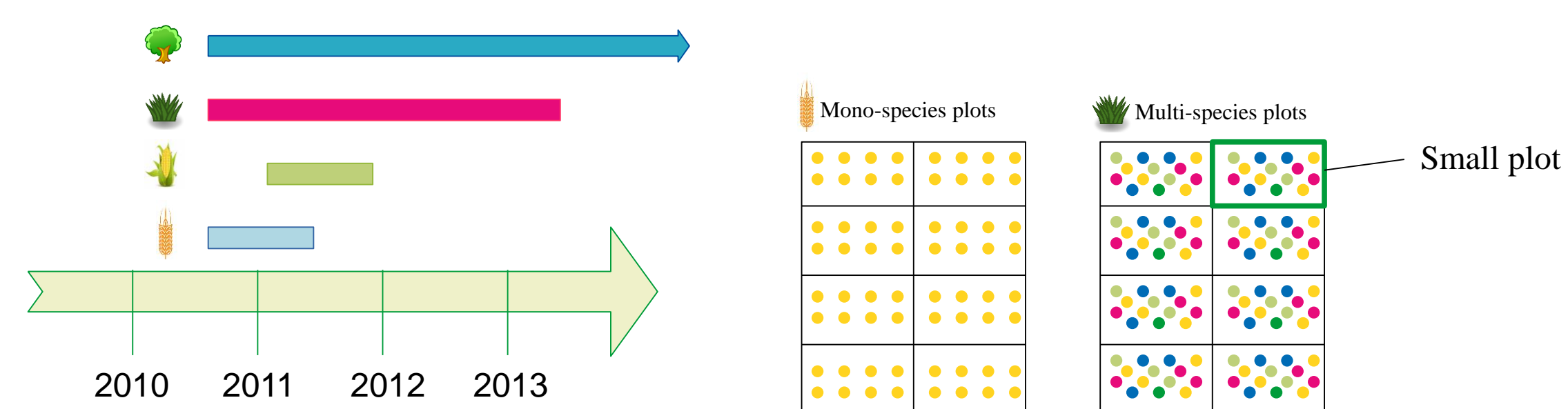
Participatory plant breeding programs



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.

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 ? Etc.

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

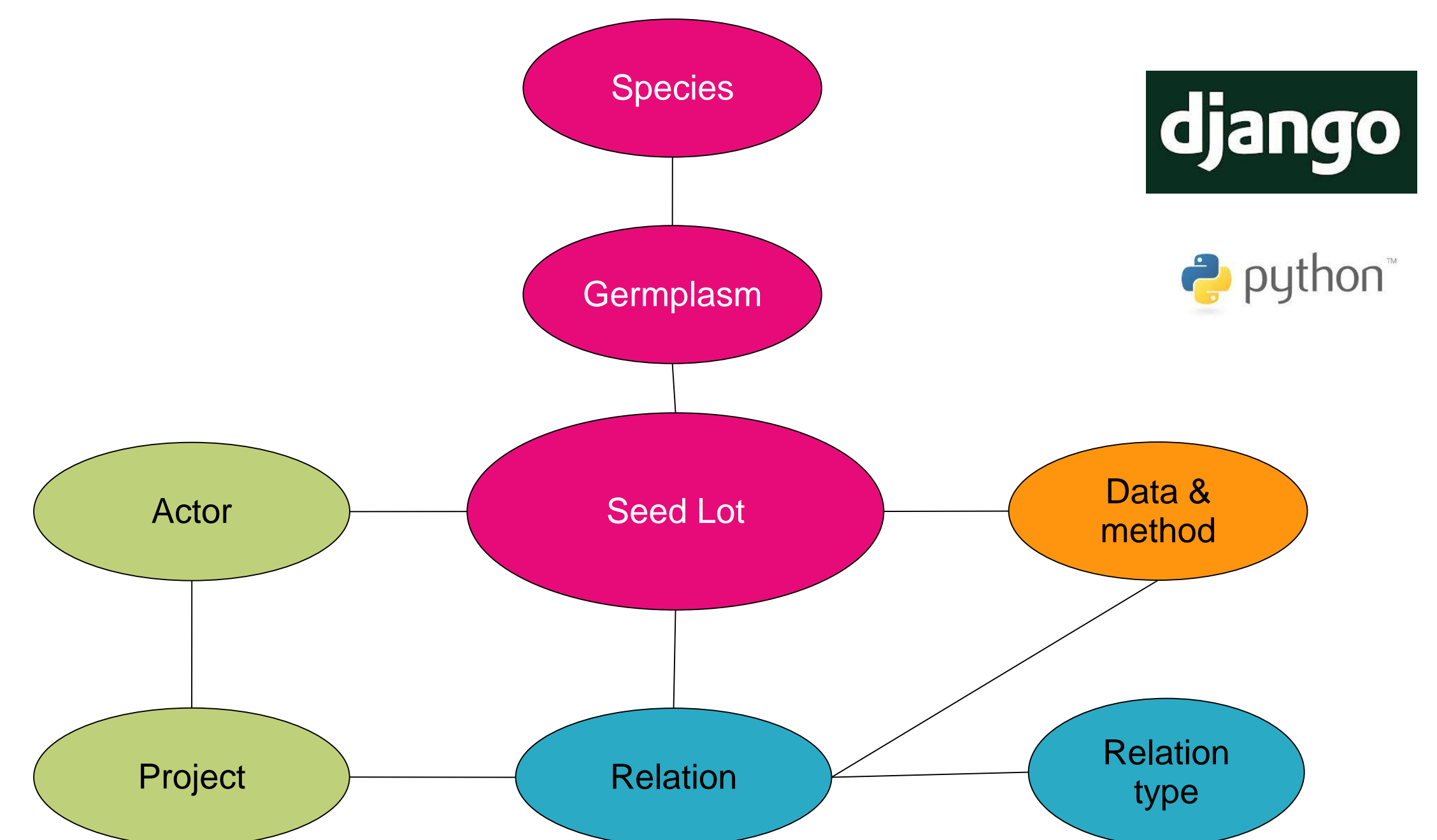
The figure below is a simplified representation of the database schema of SHiNeMaS. The schema is centered around the **Seed Lot**. A seed lot is characterized by a **Germplasm** which is characterized by a **Species**. Users can describe dynamically germplasm's type like cross, mixture, population,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

Actors describes the people involved in the data production. They can be farmers, researchers, or whoever is implied in a plant breeding program. All the data are organized in **Projects**.

The **Relation** describes the genealogical relation 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 technology with django framework. The database management system used is PostgreSQL but SHiNeMaS should be able to run with MySQL, Oracle or SQLite.



Web interfaces

Data Submission

1- SHiNeMaS provides interfaces to load massively the data with tabulated text files.

These files contains information like :
- Sowed seeds lots list, year, quantity
- Plot position
- Data measured

— insertion
— export
— deletion
— sheet

3- User fill the file with plot location, quantity, data measured

2- export **pre-filled** file with :
- sowed seeds lots list (previous harvested seeds lots)
- sowing/harvesting year

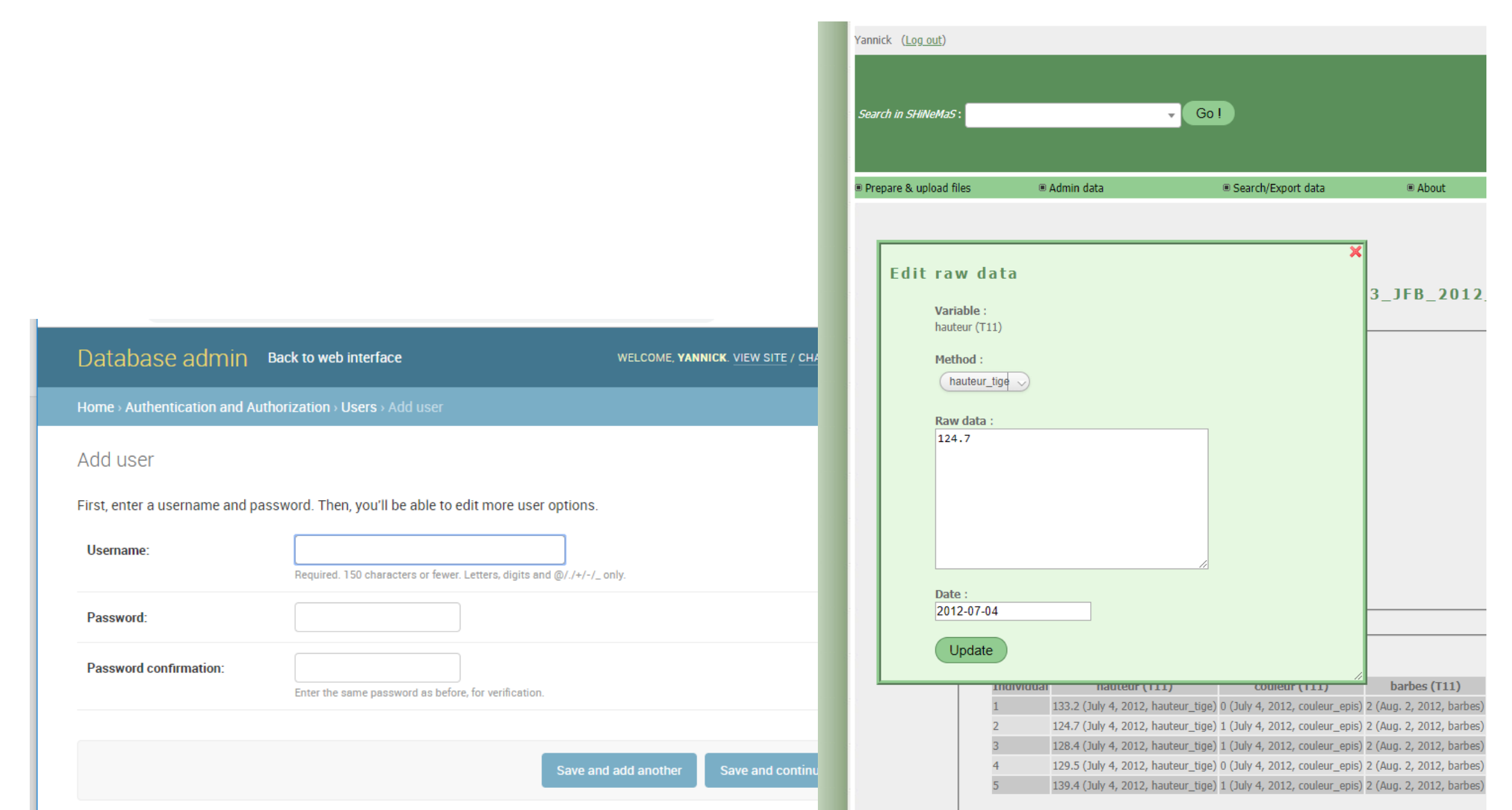
4- the exported file can contains **existing relations** and can be used to **delete** data

SHiNeMaS also provide some forms to create new event (Multiplication, Diffusion, Selection, Cross)

Data Management

The Django framework provides an automatic admin interface that we decided to use in SHiNeMaS. We enriched it with administration features. Other features were developed to manage data through the web interface.

A user can edit/delete data on a profile card (see below) or delete a whole series of data if he decides to reload it through a data file.

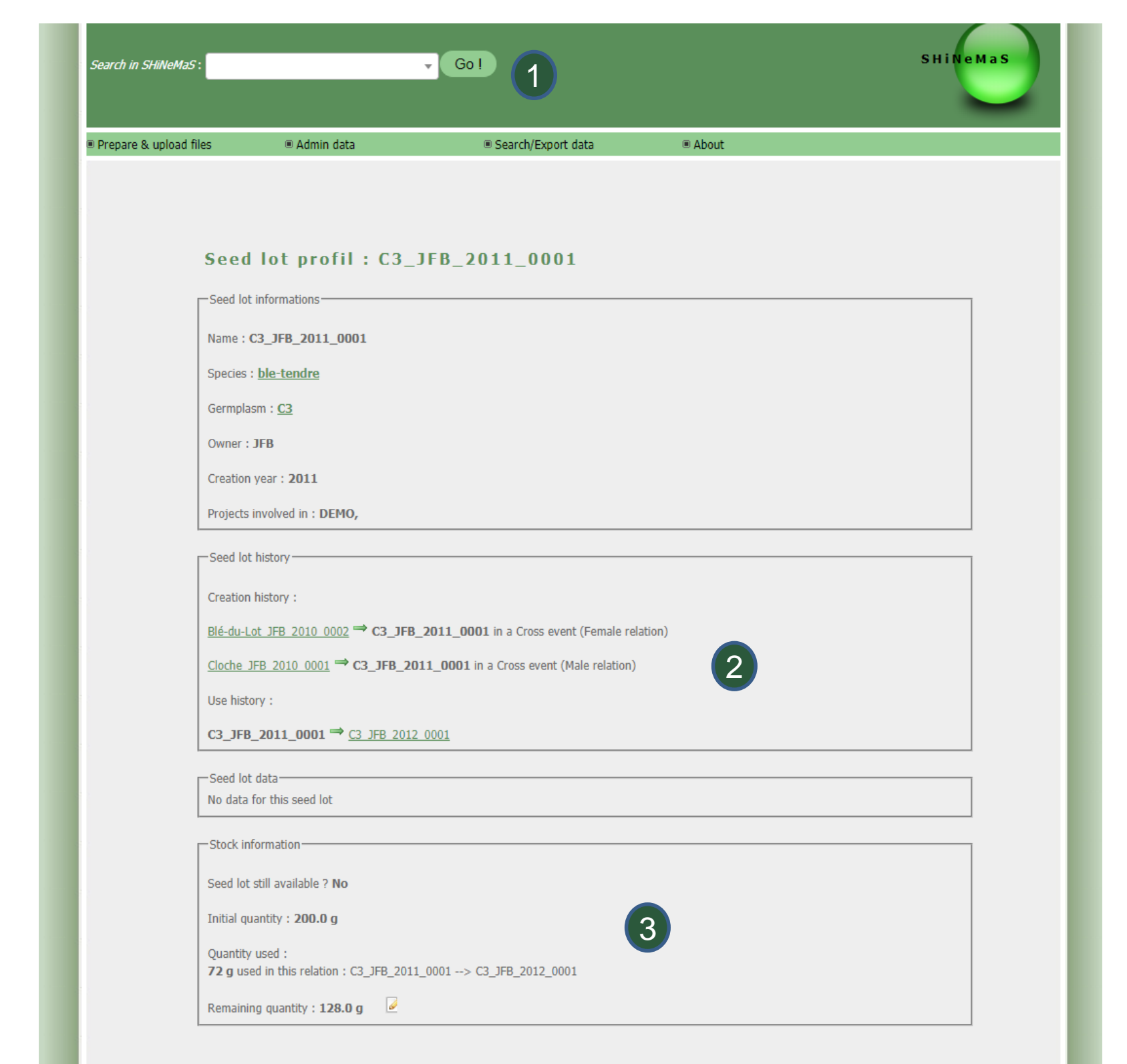


Data Querying

SHiNeMaS provides powerful query forms to search data. Information can be searched by two way. The first one is a global form (1) that searches over all the seed lots, germplasm, relation, and gives a direct access to their profile card (the picture on the right shows a seed lot card). The second is an advanced query form with much filters (year, germplasm, location, variables measured etc.) that displays a data table with a list of seed lots resulting of the query.

A card shows all the information and data related to an entity (seed lot, germplasm, relation). For example, a seed lot card will show its basic information, history (2) and stock information (3). A relation card will also show global data as well as individual data as a table

From any screen, the data table can be downloaded as a text file



Perspectives and next developments

An instance of SHiNeMaS is in production at the GQE laboratory. It stores about 2 000 germplasm, 19 000 seed lots, 18 000 relations and more than 1 600 000 records of collected data on the network. SHiNeMaS 1.0 is already available under the AGPL license and SHiNeMaS 1.1 will be available in early 2019. Development of version 2.0 is ongoing and will include new data types management (germplasm level data, images and climatic data). Features will be available to export data in PPBStat format, a R package to analyze data stored in SHiNeMaS. In the next versions we aim also to improve management of user access, quality control on imported files and use ontologies. This software have been developed with the collaboration of DEAP team (GQE-Le Moulon).