Breeding Management System

Version 3.0

Complete User Manual

September 2014

ibp@cgiar.org

Funding & Acknowledgements

The Integrated Breeding Platform (IBP) is jointly funded by: the Bill and Melinda Gates Foundation, the European Commission, United Kingdoms Department for International Development, CGIAR, the Swiss Agency for Development and Cooperation, and the CGIAR Fund Council. Coordinated by the Generation Challenge Program the Integrated Breeding Platform represents a diverse group of partners; including CGIAR Centers, national agricultural research institutes, and universities.

Table of Contents

Get Started

Introduction

About the Breeding Management System

Breeding Management System (BMS) is an information management system developed by the Integrated Breeding Platform to help breeders manage the breeding process, from program planning to decision-making. The BMS suite of tools supports multiple breeding strategies: fmns

- Conventional breeding
- Marker-assisted selection
- Marker-assisted recurrent selection
- Marker-assisted backcrossing

The BMS package contains

- Software needed to run BMS
- Crop databases preloaded with germplasm characterization for priority crops and empty databases without preloaded information for additional crops
- Instructions, tutorials, and other documents to help you install and use the BMS

System Requirements

At present you can download the Breeding Management System (BMS) package from the internet or you can request a copy on DVD. To run the BMS, you will need a personal computer running the Microsoft Windows operating system (XP or a more recent version). The Workbench runs in a web browser (e.g. Internet Explorer or Firefox), but you do not need to be connected to the internet to run the BMS. A web-based version will soon be available for users who have high-speed access to the internet.

Minimum system requirements

- 2 GB RAM
- 1 GHz dual core
- 250 GB hard disk

Recommended system requirements

- 4 GB RAM
- 1 GHz dual core
- 500 GB hard disk

Supported Browsers

- Internet Explorer 8 and above
- Firefox 20 and above
- Google Chrome 27 and above

Workbench Overview

The Workbench is the Breeding Management System (BMS) user interface, and can launch all breeding tools and access associated crop databases. The home page of workbench is where new programs can be created and previously started programs browsed and opened. The Workbench runs in a web browser, but a connection to the Internet is not required to run the Workbench.

Tools Overview

Breeding Management System (BMS) tools are applications that run on your local computer to assist and inform program development. Tools can be categorized by (1) database integration and (2) stand-alone capabilities. All tools can be launched from the Workbench, but some tools function as stand-alone applications independent of the Workbench and the databases.

Database Integrated Tools

The Breeding Management System (BMS) contains database-integrated tools that seamlessly connect to the crop database when launched from an established Workbench program. Database integration minimizes the need to manually create, save, and import files as you move through a breeding program. When these tools are launched through the Workbench, selected data moves seamlessly from the database to the application.

- Germplasm List Manager: Browse, search, and filter germplasm to create custom lists
- Breeding Manager: Design crosses, advance generations, and select progeny
- Genotype Database Manager: Connect germplasm to genotype data
- Nursery Manager: Design, manage, and advance nurseries
- Trial Manager: Design and manage field trials
- Statistical Analysis (Breeding View): Analyze phenotypic data, single site, and multi-site analyses
- Molecular Breeding Design Tool: Select target genotypes for marker-assisted backcross breeding (MABC) and determine optimal population sizes for each generation

Tools without Database Integration

These tools, even when launched from the workbench, are not automatically preloaded with selected data. Data files must be created, saved, and imported into these tools. Expect these tools to be database integrated in future versions of the Breeding Management System (BMS).

- Multi-Site Multi-Year Analyses
- QTL Analysis (Breeding View): Identify quantitative trait loci
- Molecular Breeding Planner: Match breeding goals and crop genetic considerations to marker-assisted breeding programs
- Decision support tool for marker-assisted selection (OptiMAS): Generate a target genotype by predicting the recombination of favorable QTL into a target genotype

Crop Database Overview

The BMS contains databases for nine crops (bean, cassava, chickpea, cowpea, groundnut, maize, rice, sorghum and wheat). You can also use BMS for other crops. BMS contains empty databases without preloaded information, other than a few very basic traits, such as yield.
When you install a crop database, two databases are created:

- Public database
- Program database

Public Database

The public database contains publicly available pedigree, phenotypic and genotypic data curated by a nominated center for each crop. For example, the International Institute of Tropical Agriculture (IITA) curates the cowpea database. Additionally the public database includes genetic maps for each crop based on single nucleotide polymorphism (SNP) markers, trait linkage information, and some public fingerprinting data. Genotyping with SNP markers is available from the Integrated Breeding Platform. The public database is preloaded with traits of relevance to each crop to help design field and nursery trials and to assist with the design and analysis of marker assisted breeding programs. Users cannot make changes to the public database, but data can be exported as spreadsheet (.xls) files.

Program Database

Program databases are accessible to program members, and data generated are automatically saved to the program database. BMS tools can only access one local program database at a time. Data stored in separate programs cannot be queried or browsed simultaneously. Single crop breeding programs will generally require only a single program.

We invite users to publish program data to the public crop database; sharing pedigree and performance information on your germplasm accessions. If you are interested in publishing program data to the public database please contact support@vsni.co.uk, and you will be put in touch with the appropriate crop-center curator.

Release Notes

Overview

Version 3.0 of the BMS is now available. This version offers significant improvements to the installation process, completely redesigned Fieldbook tools that are now browser-based, the introduction of seed inventory management features, and a range of improvements to our existing tools.

Installation

With version 3.0 of the BMS, we have streamlined the installation process by offering a single installer option. Users will now be able to download one installer file instead of separate installers for infrastructure, application, and crop data. After downloading, the user can select options to install, and then the installer downloads all needed components and executes the installation. The single installer can be used for both new installations of the BMS and for updating existing installations.

New Browser-Based Fieldbook Tools

In version 3.0 of the BMS, the Nursery Manager and Trial Manager components of the BMS have been redesigned as browser-based applications that are fully integrated with the other core BMS tools in the workbench.

The redesigned tools offer the following key benefits:

- Users can move seamlessly between Fieldbook tools and others in the workbench environment without launching a standalone application.The user interface is consistent with other BMS tools.
- Nursery and trial template management has been significantly improved:
- Any existing nursery or trial in the database can be used as the template for a new nursery or trial; users no longer need to customize and manage Excel-based template files that reside outside the BMS. Basic nursery templates will be included in the database to help new users get started
- The trial manager now supports an expanded range of experimental design types:
    - Randomized complete block designs
    - Resolvable incomplete block (also referred to as alpha lattice) designs
    - Incomplete block designs
    - Resolvable row-and-column designs (arranged in replicates)
    - Row-and-column designs (not arranged in replicates)
- Enhanced options for exporting measurement sheets including
    - Choice of either plot or serpentine order
    - Formats compatible with a range of handheld data collection apps

Seed Inventory Management

In version 3.0 of the BMS we have introduced several features allowing breeders to manage seed inventory:

- Ability to add inventory for seed when a when a nursery is advanced
- Ability to view available seed inventory when working with lists and making crosses, including:
- Display of inventory levels in the list view and the individual germplasm view
- Display of inventory locations in the list view and the individual germplasm view
- Ability to reserve seed inventory when working with lists and making crosses
- The BMS is compatible with legacy ICIS tools for seed storage management. Interested users should contact the IBP team for additional information and assistance.

List Manager

In version 3.0 of the BMS we have made some improvements to the List Manager, including:

- Refinements to the user interface the List Manager to improve usability and maximize the space available to work with new and existing lists.
- Ability to overwrite an existing list with changes or updates, e.g. if a user has modified an existing list she can now save it with the same name as the original list.
- Seed inventory features have been integrated into the List Manager to allow the user to see stock levels and reserve seed for a list (see above).

Crossing Manager

In version 3.0 of the BMS we have made a number of minor functional improvements to the Crossing Manager, including:

- Ability to select a complete existing list to use as either a male or female parent list
- Ability to delete individual entries from the male and female parent lists, and save modified parent lists with the same name as the original list.
- The display of male and female parent lists now shows the total number of entries and the number of currently selected entries.
- Seed inventory features have been integrated into the Crossing Manager to allow the user to see stock levels and reserve seed for parent lists while planning crosses.

Germplasm Import

In version 3.0 of the BMS we have made improvements to the Germplasm Import tool, including:

- Streamlining the import process to two steps instead of three and updating the the screen design to match other BMS tools
- The ability to import Attributes and Stock/Inventory levels
- The ability for users to select pedigree options when importing a list with an advanced import template (with GIDs included):
    - Add all entries with new records connecting to existing sources
    - Select existing germplasm wherever found
- The user is now offered new options If the system finds matches during the import:
    - The option to ignore the match and add a new record for the germplasm instead
    - The option to remember the selected match and apply it for additional instances of the same GID in the import list.
    - The user can also choose to have the system accept single matches automatically if they are found during the import.

Analytical Pipelines

The Breeding View analytical pipelines have been enhanced with the following in version 3.0:

- The Single Site Analysis now includes new functionality to assist users with identifying outliers and allowing these to be excluded from the analysis
- Summary statistics from the Single Site Analysis are now saved together with means data in the BMS database
- Output reports for the Single Site and Multi-site analysis pipelines now include summary statistics. The BMS now includes R-AP and REML scripts that can be used for Single Site and Multi-site Analysis.

GDMS

In version 3.0, we have added the following new features to the GDMS tools:

- Ability to handle duplicate germplasms while uploading and retrieving the genotyping data genotyping data upload can be done through a importing a germplasm list
- Ability to upload MTA
- Ability to create haplotypes and search for lines with particular haplotypes
- Ability to query for a region of interest e.g. all markers in the QTL region or a chromosome
- Ability to query for a particular allele
- Improved validation and error reporting

MBDT

The MBDT application is now integrated with the BMS IBDBv2 database, so that users no longer need to work with flat files when using this tool.

Program Administration & General Updates

In version 3.0, we have made the following improvements in these areas:

- The height of the header bar has been reduced in order to maximize working space for the BMS tools
- Users can now open nurseries and trials directly from the Home screen program folders inside the BMS installation folder are now named with the program name instead of a number, making it easier to find analysis reports and other output files
- The Add Program process has been expanded to allow users to optionally set up favorite methods and locations when creating a new program
- The Program Details, Members, Locations and Methods management screens have been consolidated into a single tabbed screen under the Manage Program Settings link
- Backup files automatically stored in the Backup folder are no longer removed during the un-installation process

Data Import Tool

Version 3.0 of the BMS includes several improvements to the data import tool:

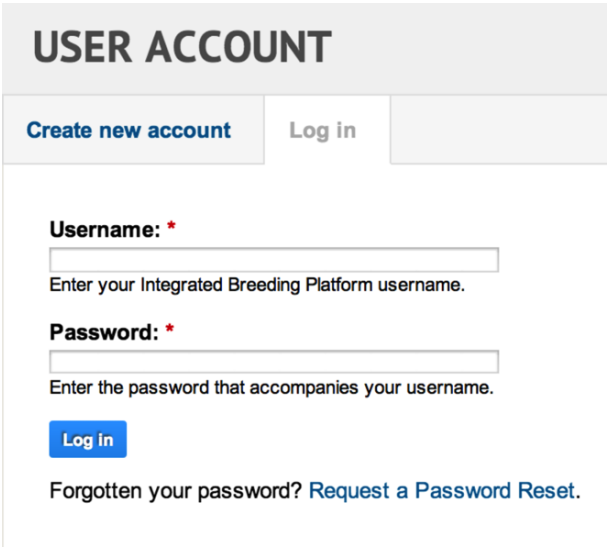
- Nurseries and trials imported through the Data Loader tool can now be opened in the new Nursery and Trial Manager tools
- The Data Loader now automatically reads details about the nursery or trial from the cover sheet during import
  Users can now import means datasets using the Data Import tool.

Registration, Download, & Install

Access Integrated Breeding Platform via the iPlant Collaborative

The Integrated Breeding Platform (IBP) uses iPlant Collaborative to manage clients. Gain access to the Breeding Management System software, and all features of the Integrated Breeding Platform by obtaining a username and password at iPlant Collaborative (https://user.iplantcollaborative.org/register/).

Login to the Integrated Breeding Platform with iPlant Collaborative username and password (https://www.integratedbreeding.net/user/login).



Establish an IBP user profile and join community forums.

Download and Install the Breeding Management System

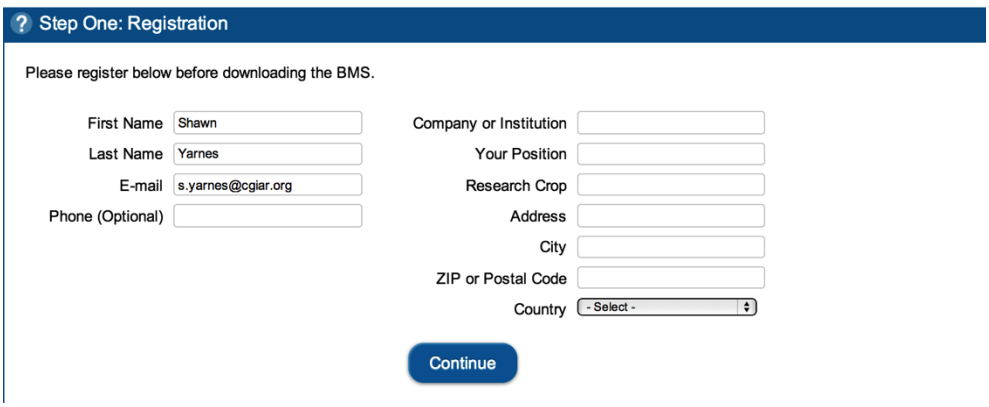
Download the BMS package after logging into the Integrated Breeding Platform with iPlant Collaborative username and password.

Installation of the BMS involves:

- Register for software
- Download and Install BMS infrastructure files
- Download and Install BMS Applications
- Download and Install Selected Crop Databases
- Download and Install Documentation Files

Alternatively, if you have a slow or unreliable internet connection you can request a DVD of the BMS by contacting ibp@cgiar.org. A DVD will be mailed to you within 24 hours of your request.

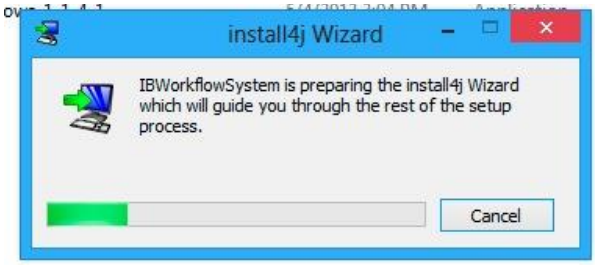
Register to Begin Download



Download and Install BMS Infrastructure Files

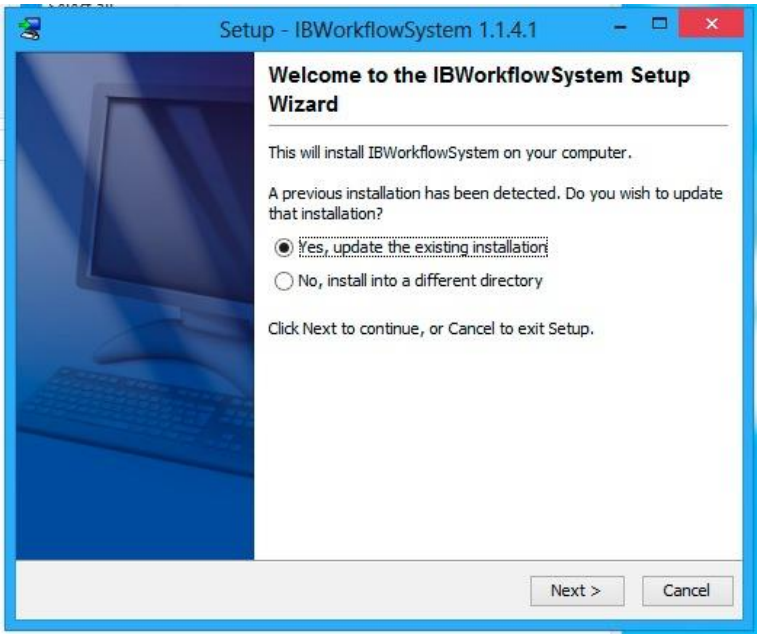
Click the Download button to download the infrastructure installer. You may be asked to select a file directory for the installer or it may automatically be saved in your Downloads folder.

Locate the installer file on your hard drive and double click on it to run it. A pop-up window will indicate the initiation of the setup.



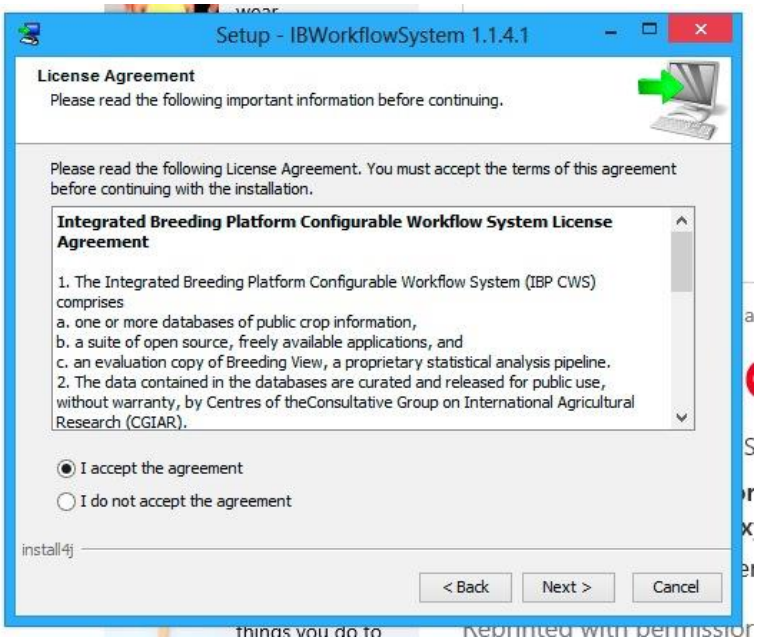
Select your preferred language from the drop-down menu and click OK.

If you have previously installed the BMS, you will be asked to hoose whether you would like to update your existing installation or create a new installation in a different directory.

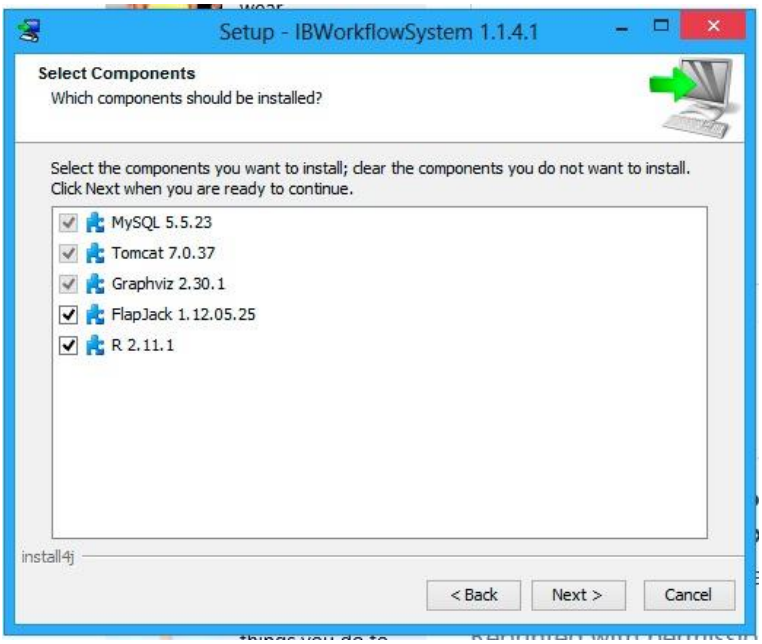


Click Next to proceed.

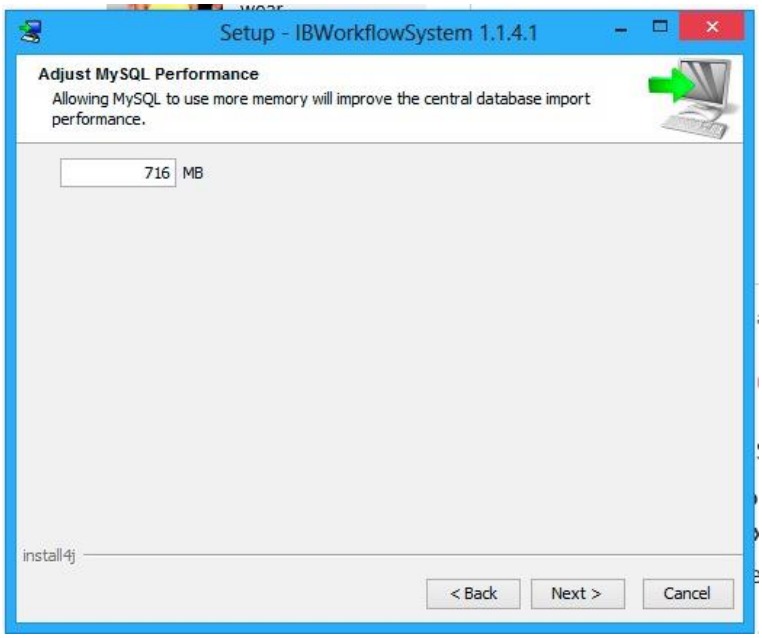
Read the license agreement shown in the next screen. Make sure that the 'I accept the agreement' option is selected.



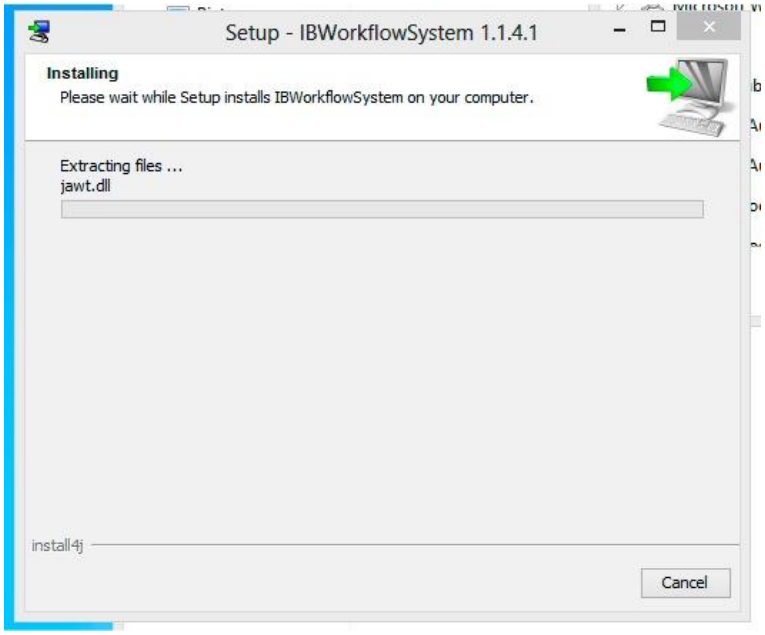
Select the components you wish to install.



Advanced users may choose to increase the amount of memory allocated to the installation process. Increasing the memory reduces the amount of time it takes for the installation to complete. If you choose this option, do not enter a value that exceeds the amount of RAM available on your system. Otherwise, leave the default value as 716 MB.



You will see the following screen while the installation is in progress. If you need to stop the installation, click Cancel; otherwise wait for the installation to complete.



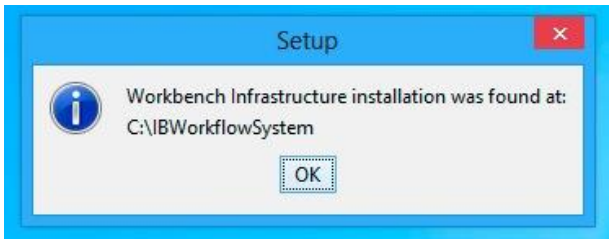
Click Finish to complete the installation

Download and Install BMS Application

Once you have downloaded and installed the BMS infrastructure, as described above, the next step is to download and install the BMS application. Click the Download button to start the download of this installer. Depending on your browser configuration, you may be asked where you would like to save the installer, or it may automatically be saved in your Downloads folder. Locate the installer file on your hard drive and double click on it to run it. Once the installer begins, you will see a pop-up indicating that it is preparing to do the setup.

Select your preferred language from the drop-down menu and click OK.

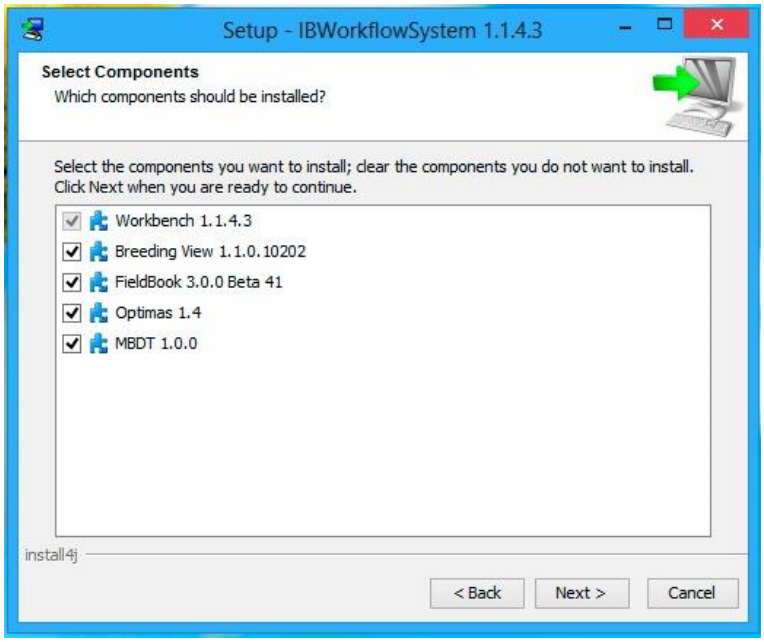
Click OK to acknowledge the location of previously installed files.



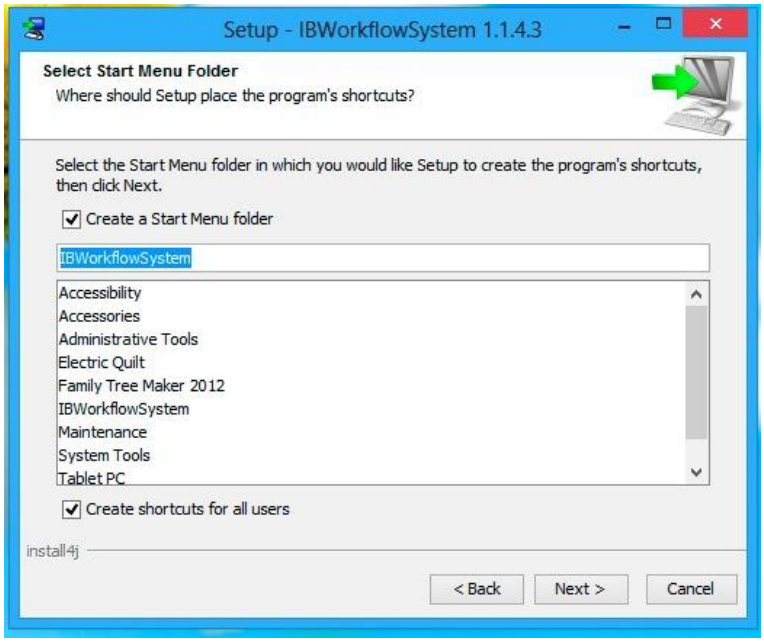
You will now see the first screen of the setup wizard. Click Next to proceed.

Read the license agreement. Make sure that the 'I accept the agreement' option is selected, and then click Next.

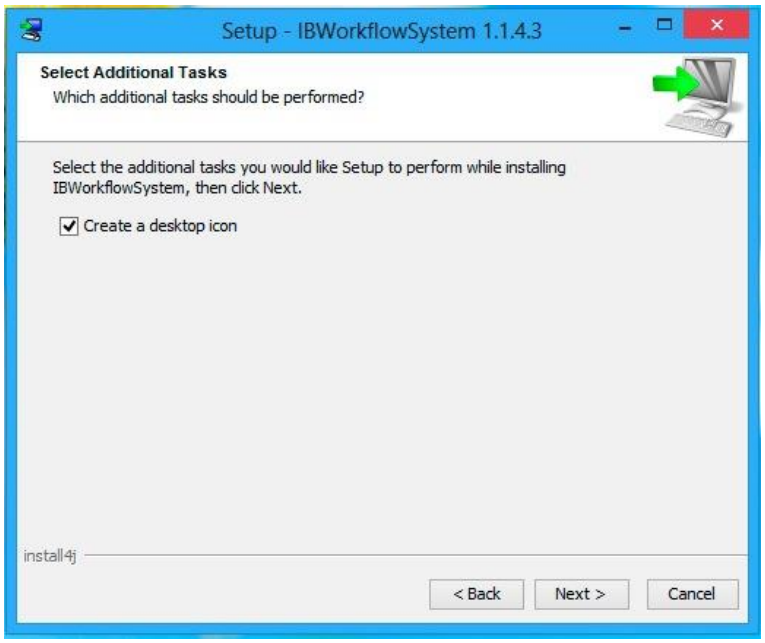
Select the components you wish to install, then click Next.



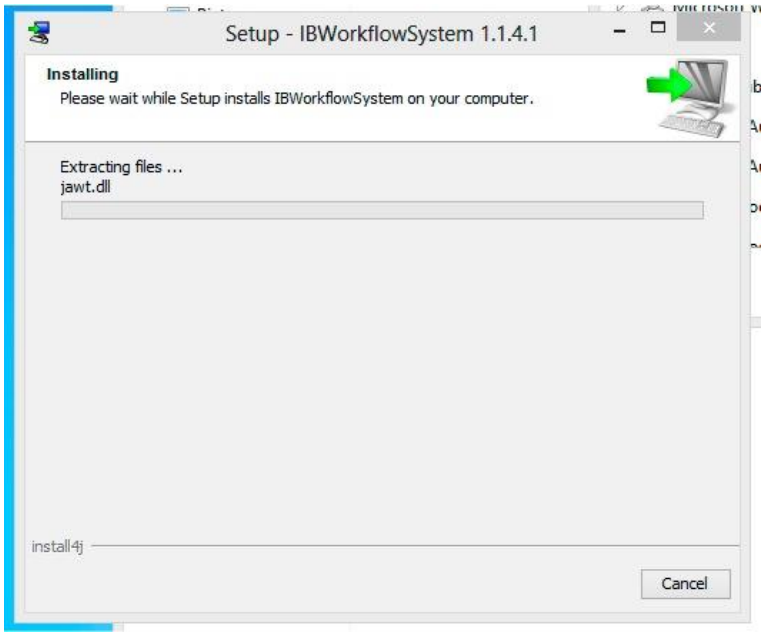
Choose whether or not you would like the installer to create a Start menu folder and shortcuts, then click Next to proceed. You may change the name of the folder from the default if you wish.



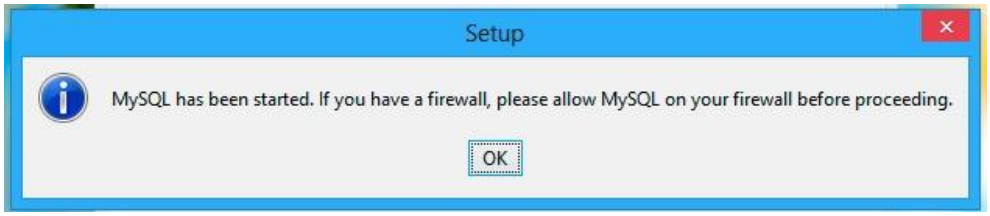
Choose whether or not you would like the installer to create a desktop icon for you (one will be created by default).



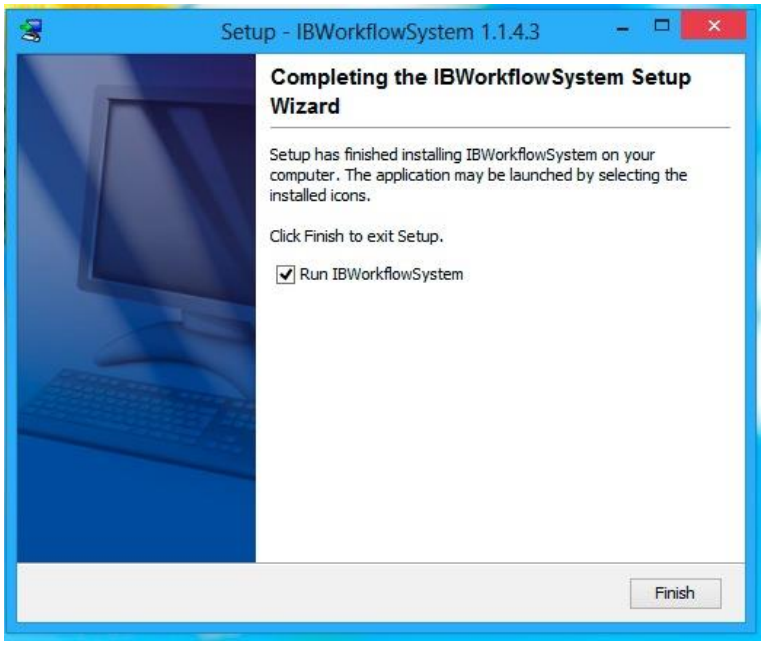
If you need to stop the installation, click Cancel; otherwise wait for the installation to complete.



When the installation is nearly complete, you will see the following notification that MySQL has been started, and that you may need to adjust your firewall settings. Click OK to continue.

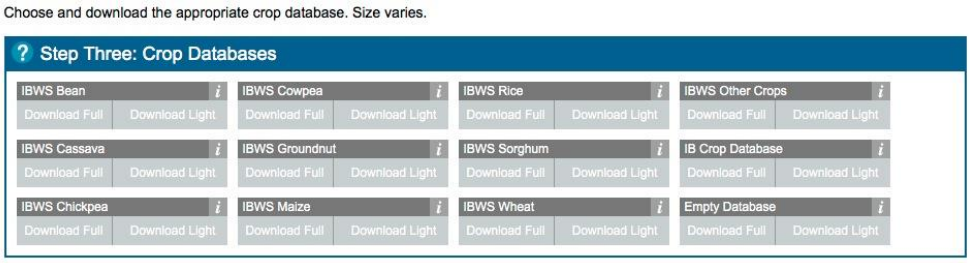


Click Finish to complete the installation.



Download and Install Crop Databases

Once you have downloaded and installed the BMS application and the BMS support files, the next step is to select the crop database(s) you wish to work with, then download and install.



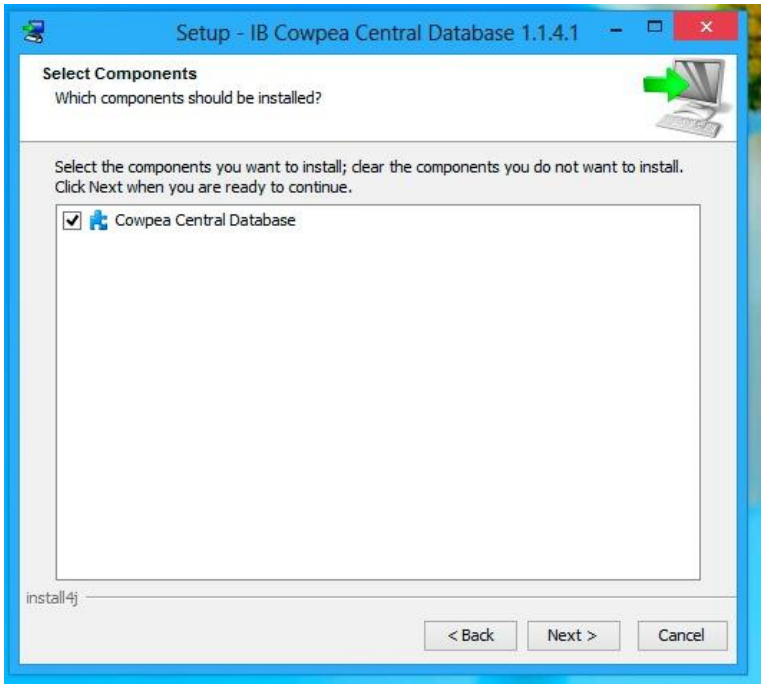
Depending on your browser configuration, you may be asked where you would like to save the installer, or it may automatically be saved in your Downloads folder.

Locate the installer file (BMS_[crop]_central_database_windows_[version].exe) on your hard drive and double click to run it. Once the installer begins, you will see a pop-up indicating that it is preparing the setup.

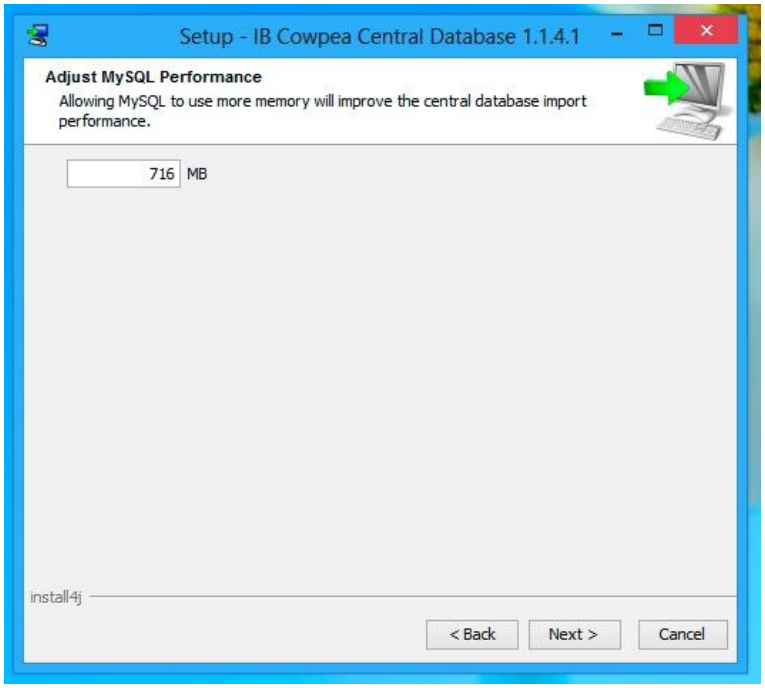
Select your preferred language from the drop-down menu. Click OK to acknowledge the location of previously installed files. Click Next to proceed.

Read the license agreement. Make sure that the 'I accept the agreement' option is selected, then click Next.

Select the components you wish to install. Click Next.



Advanced users may want to increase the amount of memory allocated to the installation process. Increasing this allocation will make the installation quicker. If you choose to use this option, do not enter a value that exceeds the amount of RAM available on your system. Otherwise, leave the default value as is and click Next to proceed.



If you need to stop the installation, click Cancel; otherwise wait for the installation to complete and select Finish. You have now completed the installation of your selected crop database, and you are ready to begin using the BMS. You can install additional crop databases to use with the BMS if you need them for your program. To do this, download the appropriate installer and go through this process again.

Download and Install Documentation

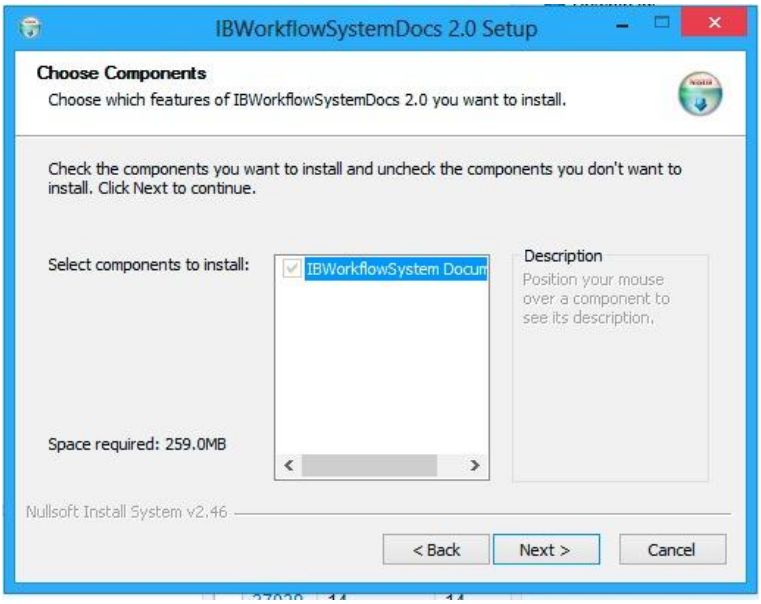
The Documentation installer is a self-extracting archive that contains tutorials and sample files for use with the BMS.

Locate the Documentation Installer file on your hard drive and double click on it to run it.

Select your preferred language from the drop-down menu and click OK.

You will now see the first screen of the setup wizard. Click Next to proceed with your installation.

Select the components you wish to install, then click Next.



Choose the location where you would like to install the documentation, then click Next. By default, it will install into the same directory as the BMS.

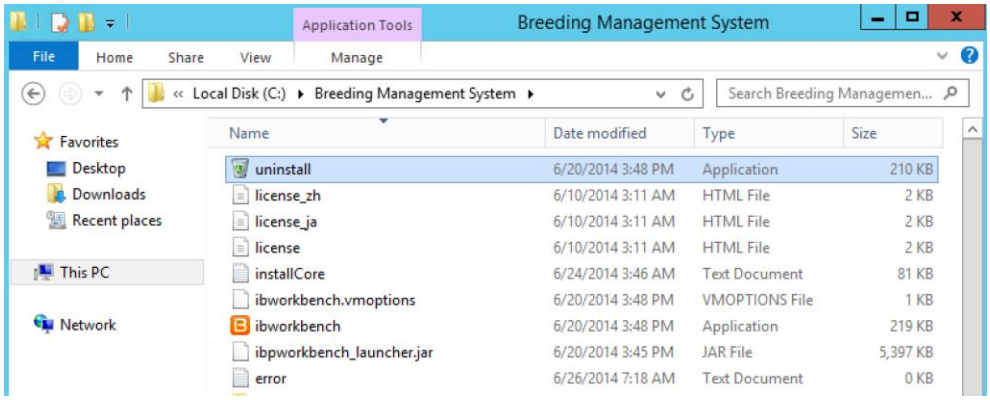
If you need to stop the installation, click Cancel. Otherwise wait for the installation to complete.

When the installation process is complete, click Finish to complete the installation.

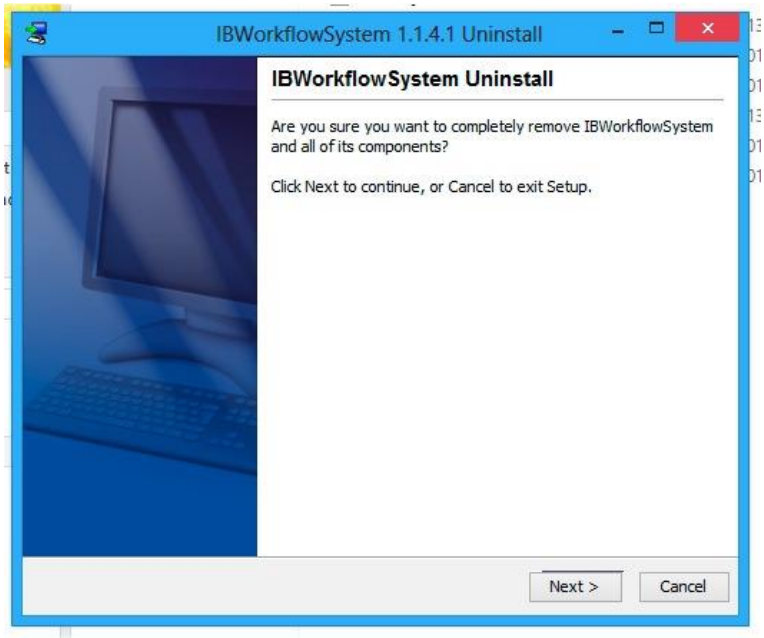
Uninstall the Breeding Management System

The Breeding Management System (BMS) provides tools to partly or completely uninstall the application, the infrastructure, and the databases. Make sure to back up your program database in another directory before uninstalling the BMS.

Locate the BMS installation folder. Double click on the Uninstall file to begin the uninstall process.

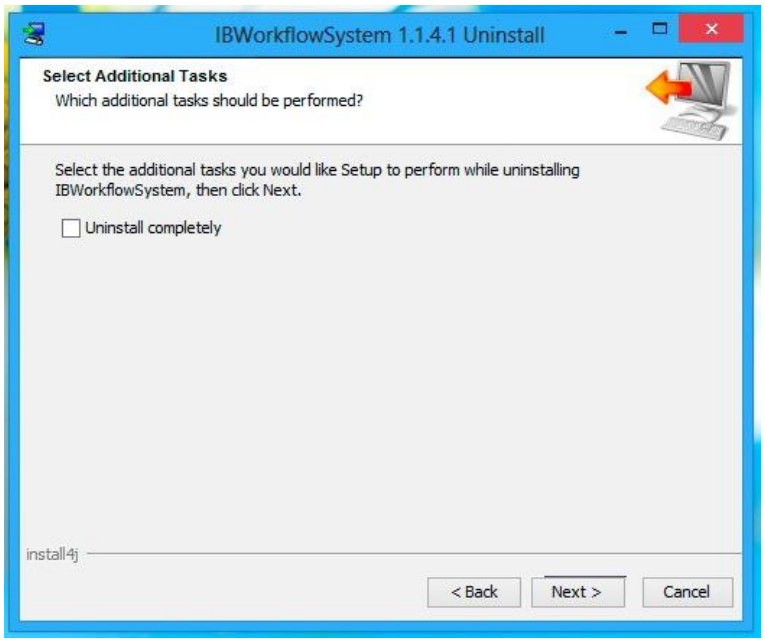


Click Next on the first screen of the uninstall wizard to removing the BMS.

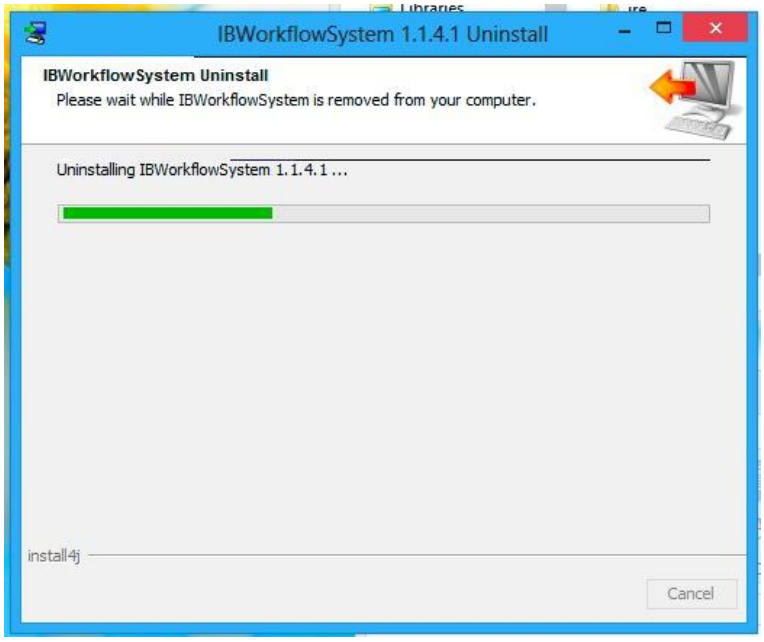


Choose Between Complete or Partial Uninstall

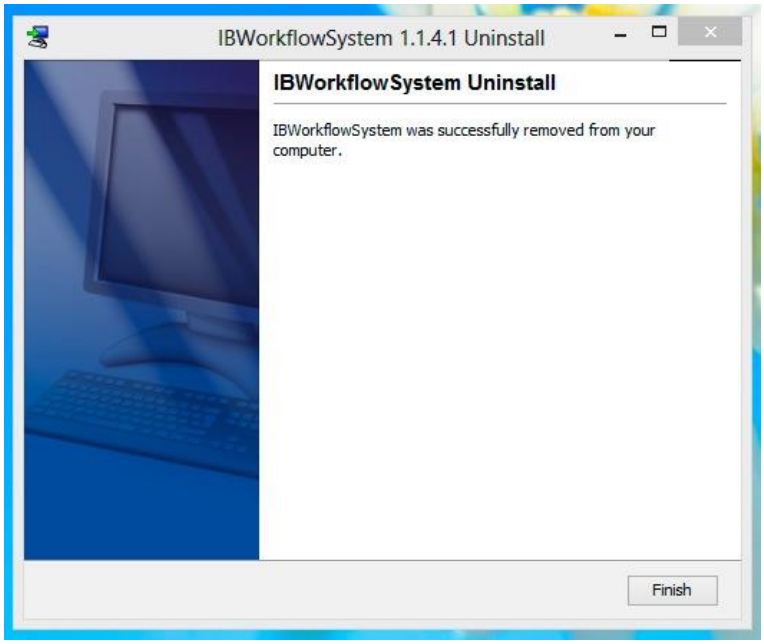
A partial uninstall will remove the BMS application and infrastructure, but will not remove any of the crop databases you have installed. This allows you to continue to use your crop databases if you reinstall the BMS or if you upgrade to a new version of it. The uninstall wizard will perform a partial uninstall by default, and all you need do is click Next. A complete uninstall will remove the application, infrastructure and any crop databases you installed. To choose this option, check the 'Uninstall completely' box and click Next.



The uninstallation process may take a few moments to complete.

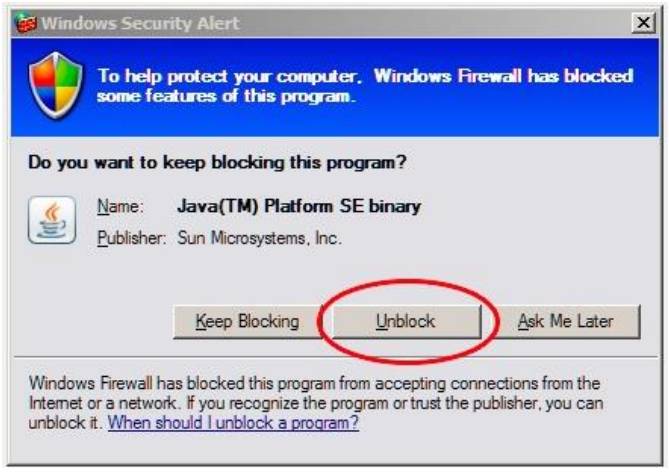


Once the uninstallation is done, you will see a notification that the software has been removed from your computer. Click the Finish button to complete the process.



Launch BMS for the First Time

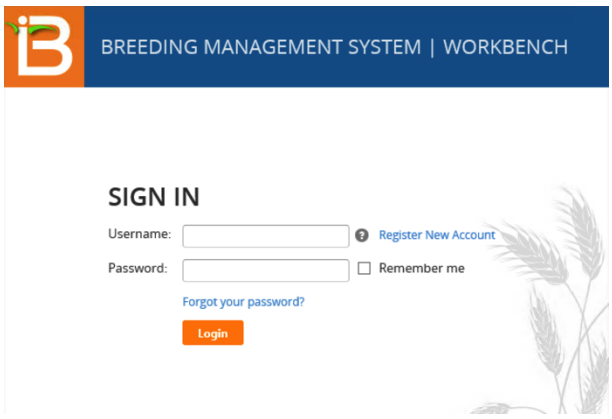
After Breeding Management System (BMS) installation, launch the BMS from the Start menu or the desktop icon if you opt to have the installer create one for you. Unblock MySQL in your firewall if you have one installed. Click the OK button to launch the BMS.



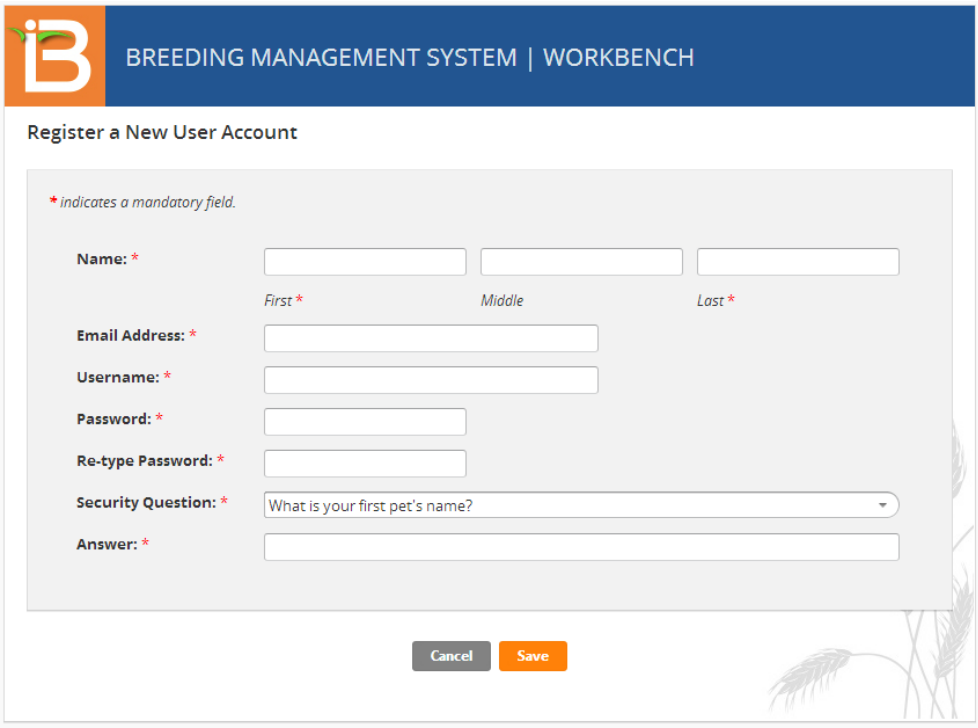
Next, you will see the BMS start-up screen. Please note that when you launch the BMS for the first time, it can take several minutes for the initialization to complete move you to the log in screen. This may also occur the first time you launch the BMS after you restart your computer.

Register as User

The first time you use the BMS register as a network user. This identifies you from other program members that collaborate on shared programs. Click on Register New Account.

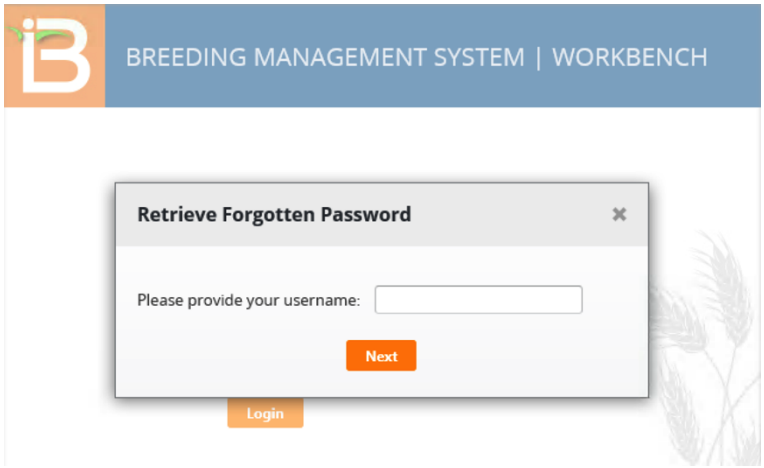


Complete the Register User Account Form. Click the Save button to create your user account.



Logging Into the Workbench

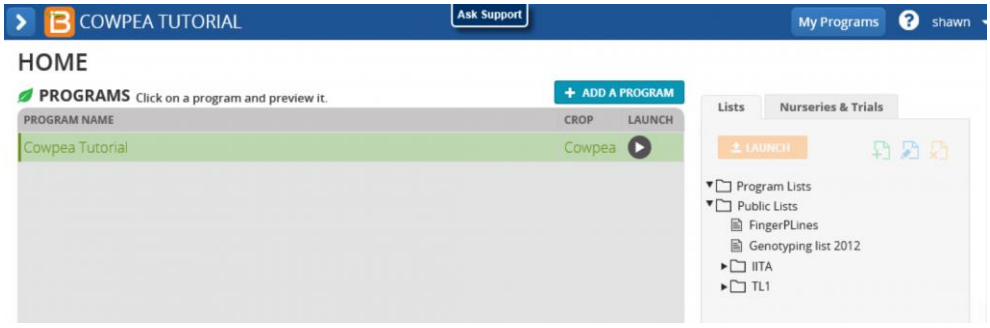
After launching, enter your username and password then click Login to enter the BMS.
Retrieve forgotten password by selecting Forgot Your Password. Enter your user name when prompted. Click Next.



Answer security question and click Retrieve Password. If you have answered the security question correctly, the system will display your password. If you do not enter it correctly, you will see an error message and will need to try again.

Home Screen

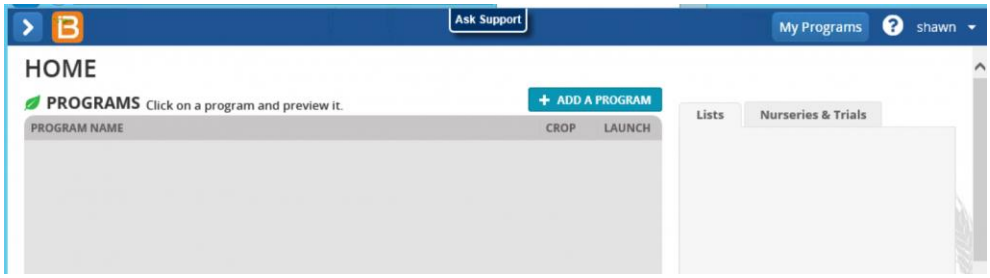
The BMS opens to the Home screen in your default browser. The home screen of the BMS is where programs are added, selected, and program file directories can be viewed and modified.



Home screen showing two program databases, cowpea and maize. The program, Cowpea Tutorial, is selected and the public list files are visible.

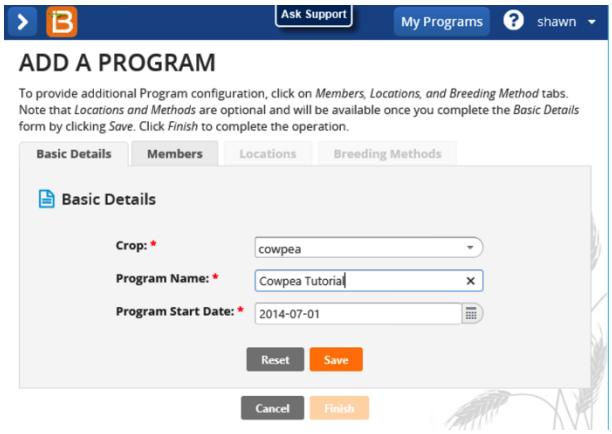
Create a Program

The Home Screen will be empty until you create a new Program database by clicking the Add a Program button at the top right of this panel.

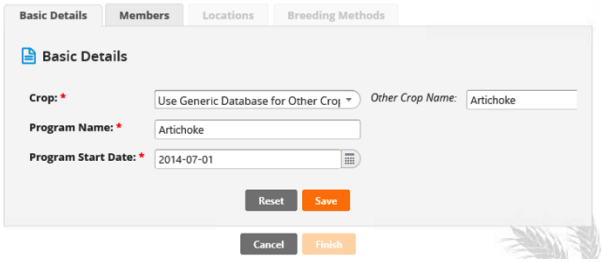


Basic Details

The Basic Details tab is the only mandatory information needed to create a program. Select the desired crop from the drop-down menu, and enter a name for your program. The start date is set to the current date by default. All of the public crop databases downloaded and installed will appear in the Crop dropdown menu. If a public crop database was installed after launching the Workbench, the crop will not appear in the drop-down menu until the Workbench is re-launched.



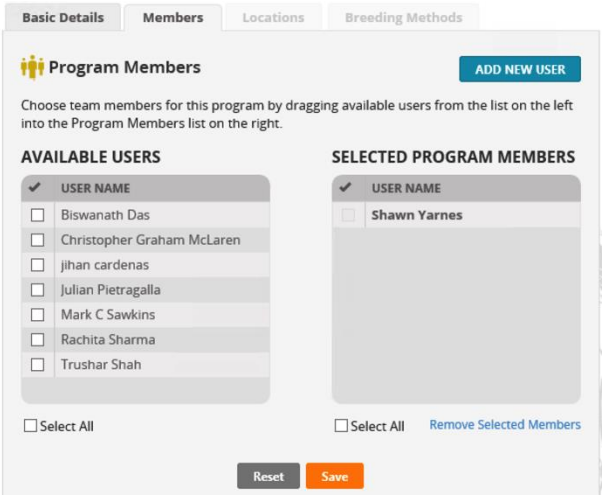
If you work with a crop not supported with a public database, the BMS provides an empty database for you to use. Choose to Use the Generic Database and type the name of the crop you wish to work with.



Save Basic Details. Select Finish to open the program or enter optional information about Members, Locations, and Breeding Methods.

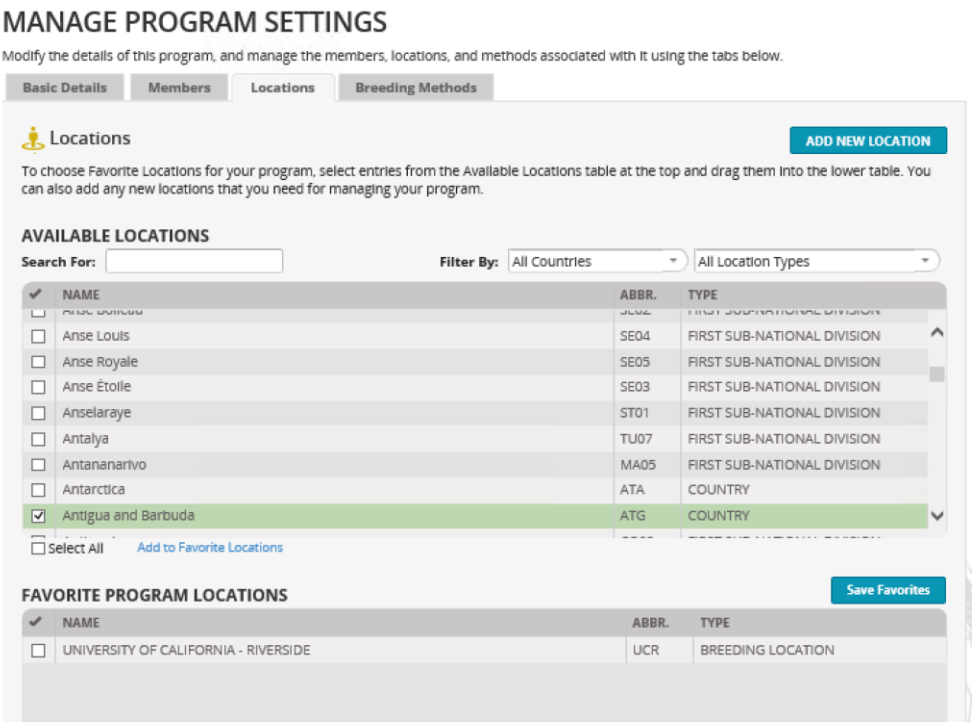
Manage Program Members

Add program members by clicking on Program Members. A list of current registered Workbench users (will be displayed with a field to add new project members. This is an optional step. If you do not need to include additional members, click Finish and open your program.



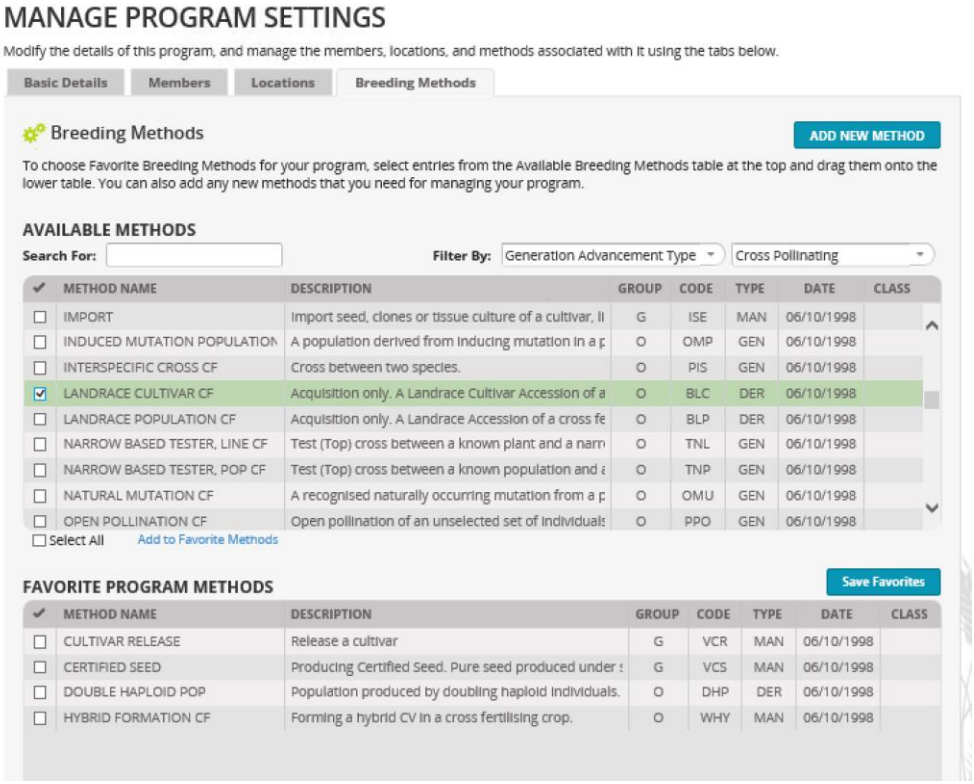
Manage Locations

Choose the locations used in the breeding program from the list of available locations. Add new locations as needed. Highlight available locations to populate the program favorites lists. Select Add to Favorite Locations and Save Favorites. Selected breeding locations will be available for selection during nursery and trial design.



Manage Breeding Methods

Choose the breeding methods used in the breeding program from the list of available methods. Add new custom methods as needed. Highlight available methods to populate the program favorites lists. Select Add to Favorite Methods and Save Favorites. Selected breeding methods will be available for selection during nursery design.



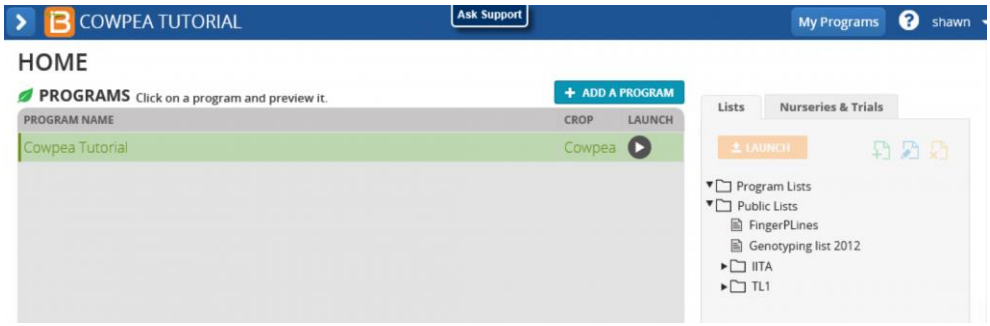
Finish

Selecting finish after establishing the program settings will automatically launch the BMS workbench from the List Manager.

Launch Tools and Activities for Established Programs

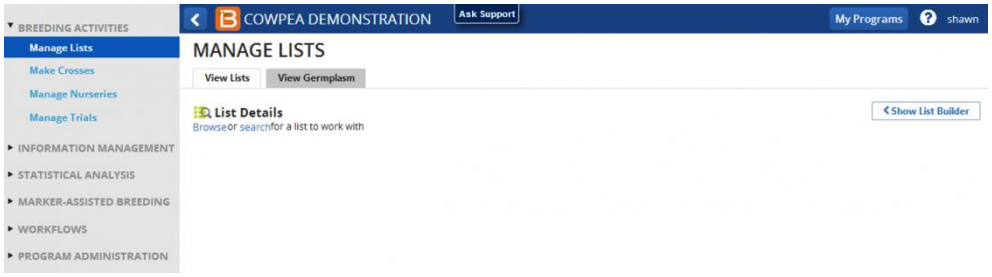
Three ways to launch a program from the Home Screen.

- Double click the selected program.
- Double click on a specific file from the file directory
- Select launch to open a specific selected file.



Workbench Organization

Launching a program opens the BMS Workbench where the BMS tools and administrative functions are located in the left hand menu. The List Manager is the default tool that opens when a program is launched.



Workbench functions are grouped by:

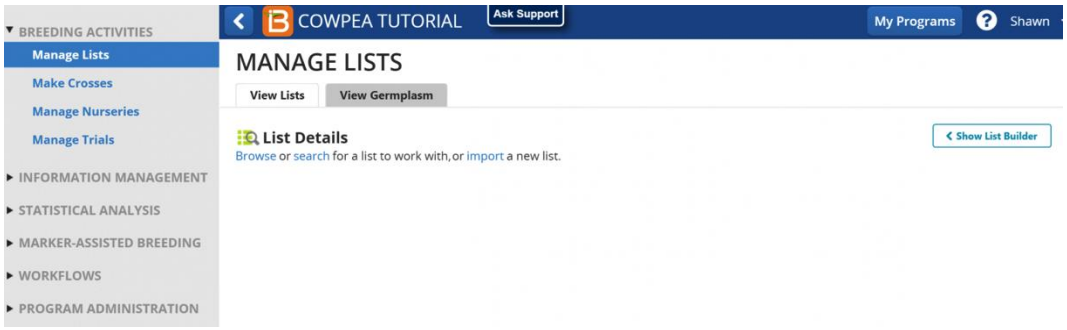
- Breeding Activities
- Information Management
- Statistical Analysis
- Marker-Assisted Breeding
- Workflows
- Program Administration

Collapse the left side menu by selecting the **<** icon next to Menu.

Breeding Activities

Manage Lists

Browse and search existing lists from the program and public databases with the List Manager. Use the List Manger to create and save new germplasm lists to the program database. See more on importing a new germplasm list.

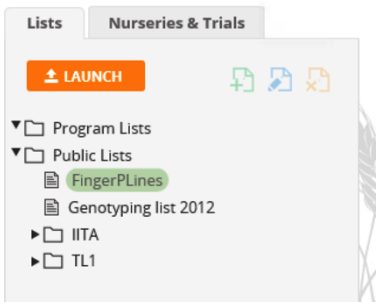


Search Lists

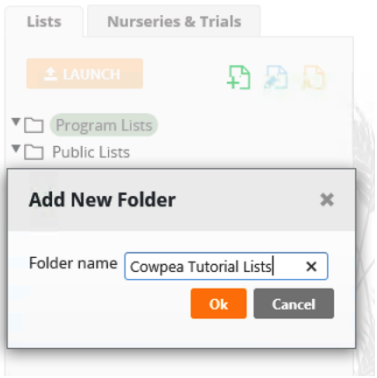
Selecting Search will reveal a popup window. Enter partial or exact list or germplasm information to find matching lists.

Browse & Manage Files

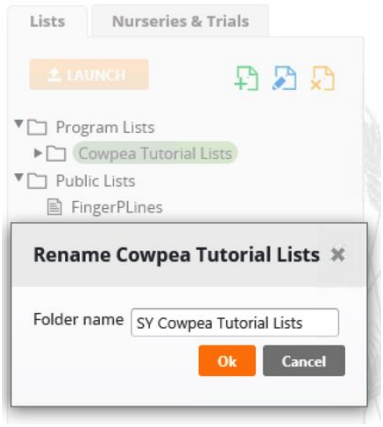
Selecting Browse will reveal a popup window containing a collapsible file directory of all lists, germplasm lists, as well as lists of Nurseries and Trials. Files are split between program database and public database sections. You cannot modify files in the public database.



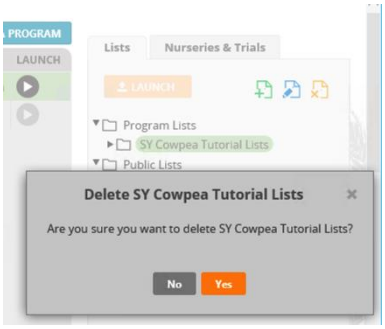
Add a new folder to program lists by selecting Program Lists and the green + icon and entering a new folder name.



Rename a program file by selecting the blue pen icon. Click OK.

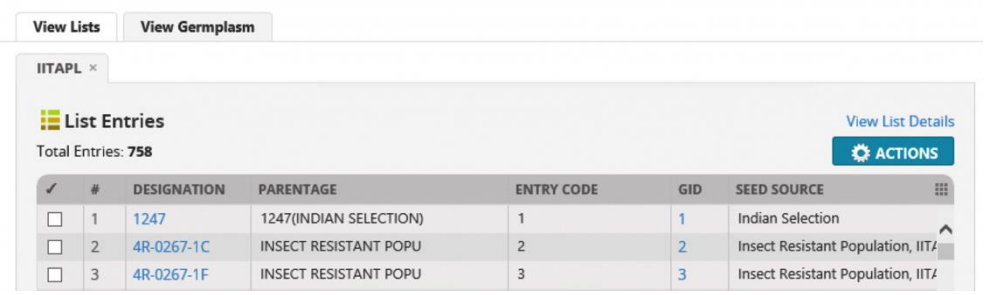


Delete a file from the program directory by selecting the yellow X icon. Click Yes.



Review List Details

Highlight a list and double click to reveal contents.



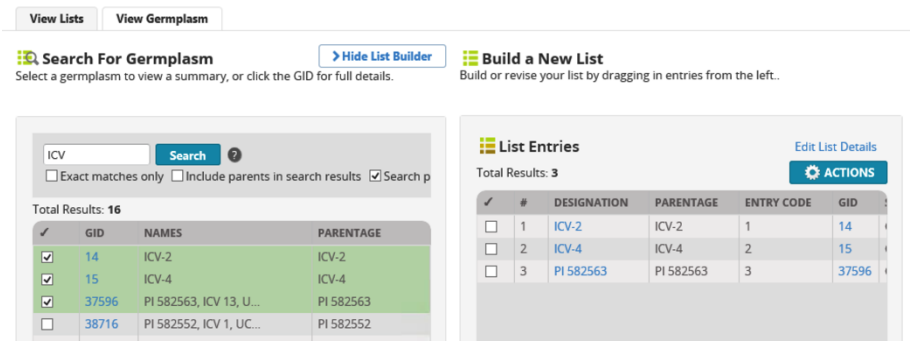
Default columns of data define Germplasm. Double click a column header to sort the list by column values.

- Designation
- Parentage
- Entry Code
- GID: Germplasm ID
- Seed Source

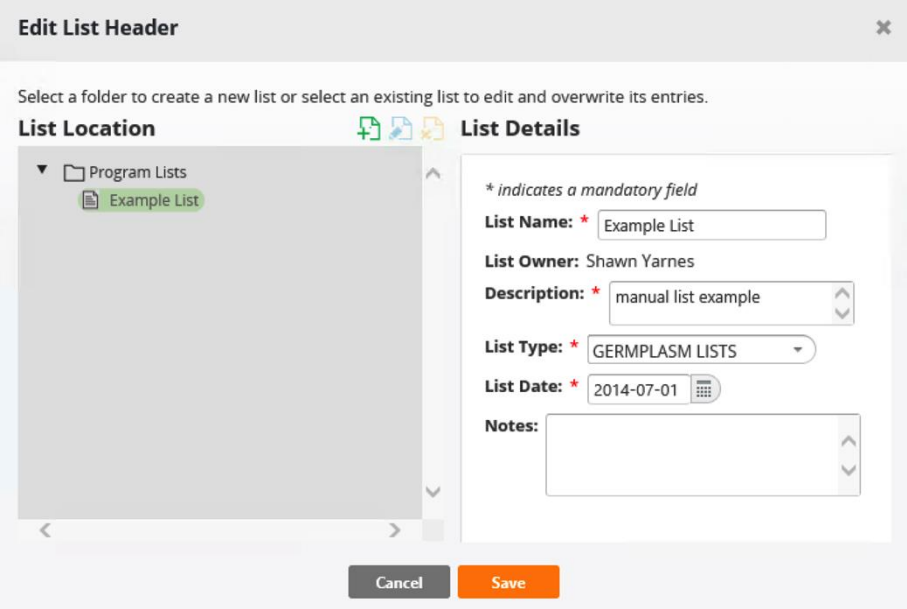
Select individual Designations or GIDs (germplasm identifiers) to reveal germplasm details.

Build a New List

Select germplasm of interest in and drag and drop the selections into the build a new list window.



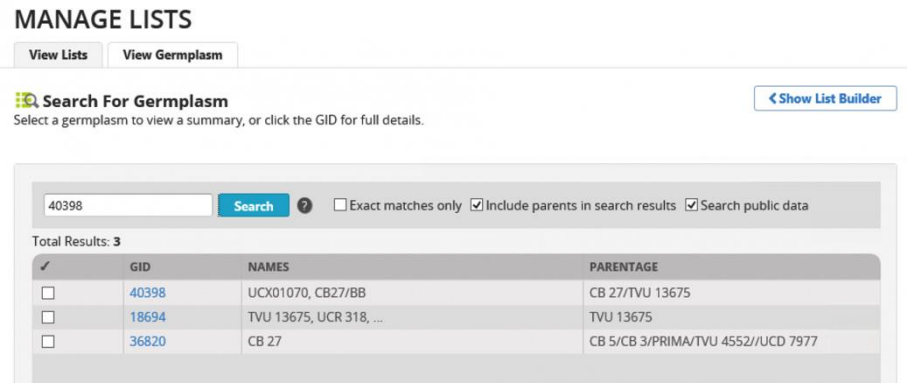
Enter the list details. The list name must be unique to the program database.



The newly created list now appears in the Program Lists file folder. The list can be renamed by selecting the pen icon and locked against further editing by selecting the lock icon.

Search Germplasm

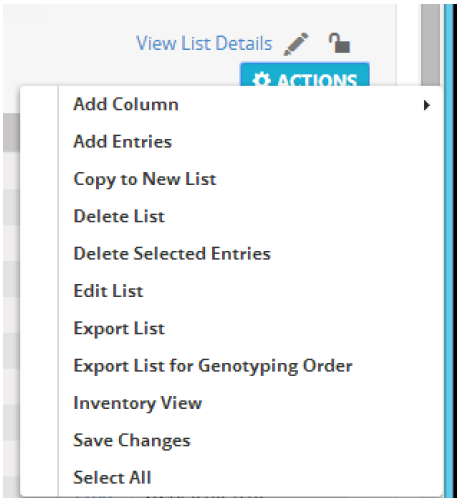
Search by exact matches or partial matches to list names, germplasm designations, or GIDs (germplasm IDs). The search feature can also provide parental information. Searches are not case-sensitive. Notice that the Build a New List pane started in the browse list tab remains open when you move to the Search tab.



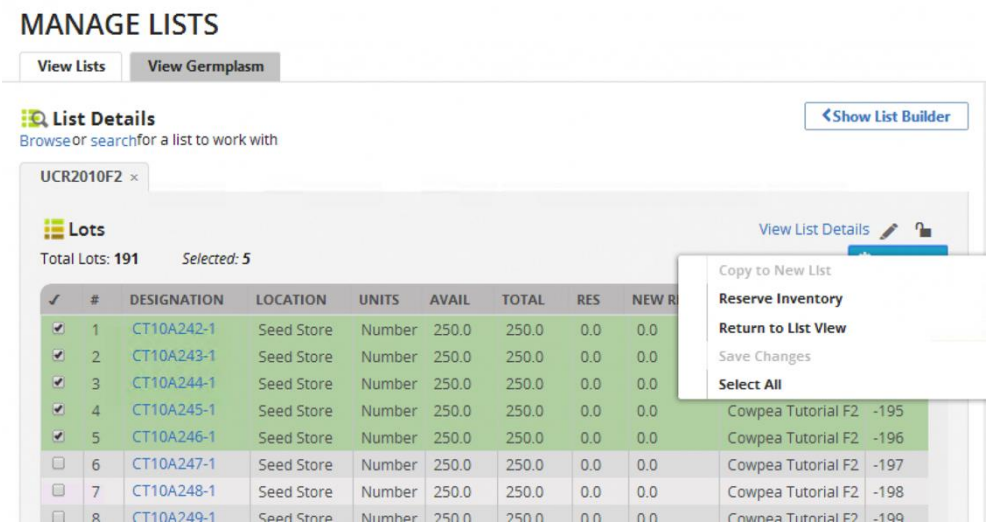
A search of the cowpea database for 40398 reveals germplasm matching GID 40398 as well as the parents of this germplasm. Notice that GI 40398 was not found associated with any existing program database lists, only the public database.

Seed Inventory

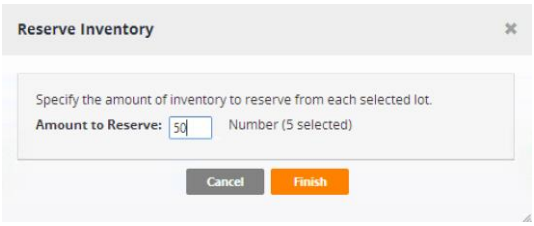
Review the seed inventory by selecting Inventory View from the Actions dropdown menu.



Reserve seed from selected germplasm inventory view by choosing Reserve Inventory from the Actions dropdown menu.

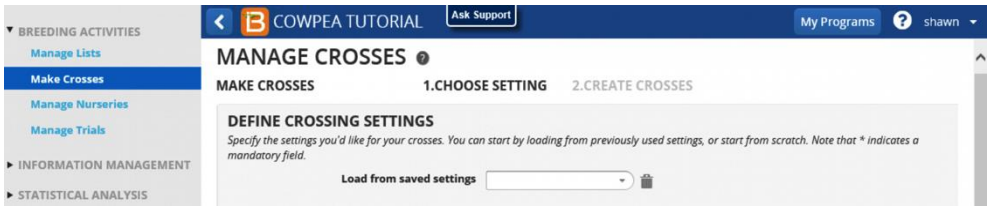


Specify the amount of seed to reserve.



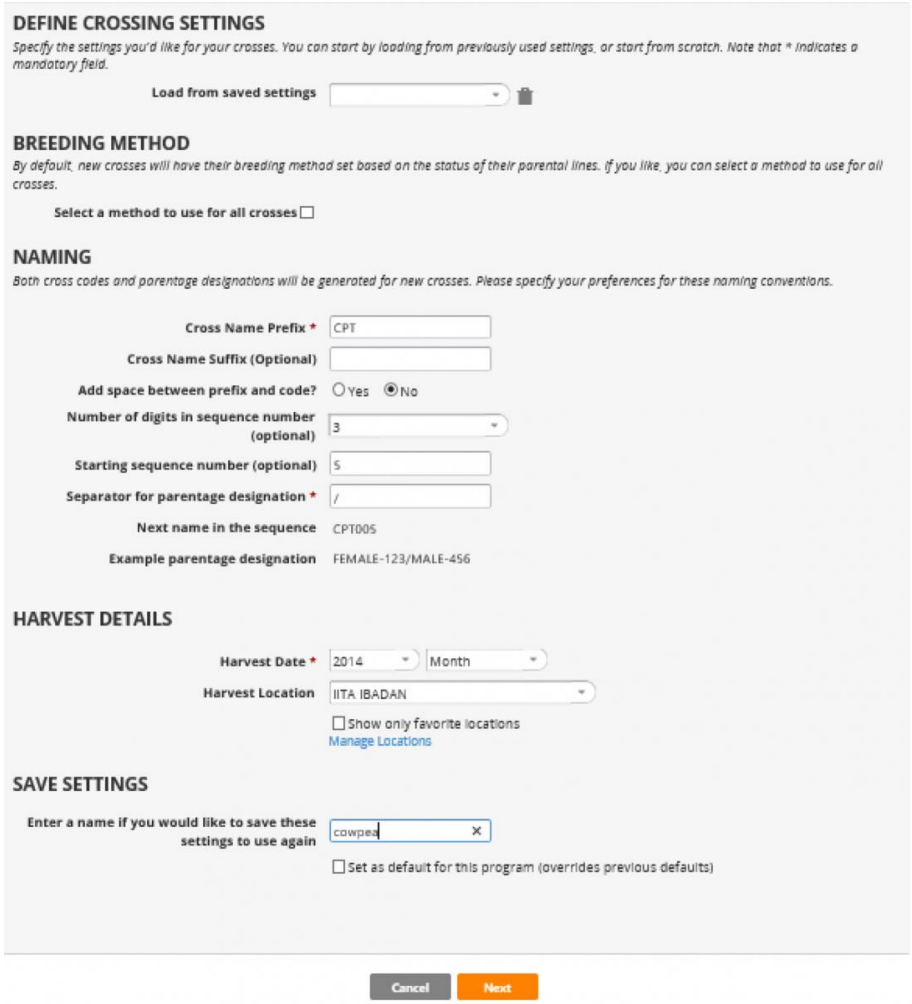
Make Crosses

The crossing manager allows users to define the cross settings, choose parents, review seed inventory, reserve seed, and make crossing lists..



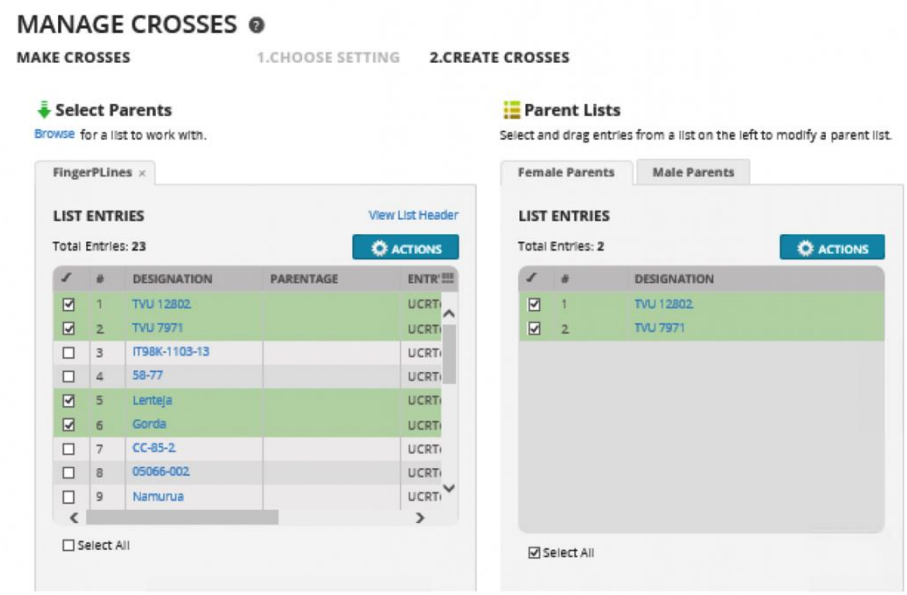
Define Cross Settings

Defining the crossing settings can be done every time a cross is made. However, most breeding programs only use a few crossing methods. The BMS allows you to save commonly used methods and create a list of favorite methods. You are also able to set a default method.

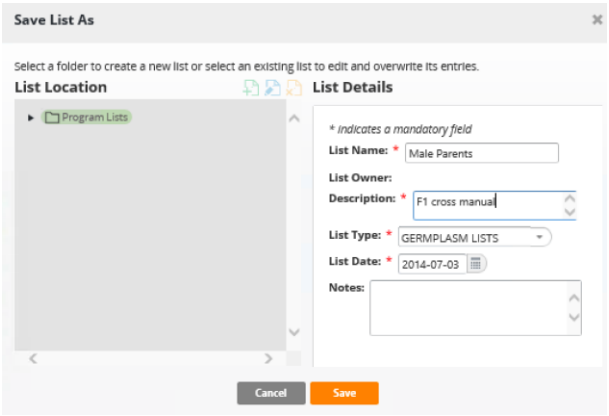


Design Crosses

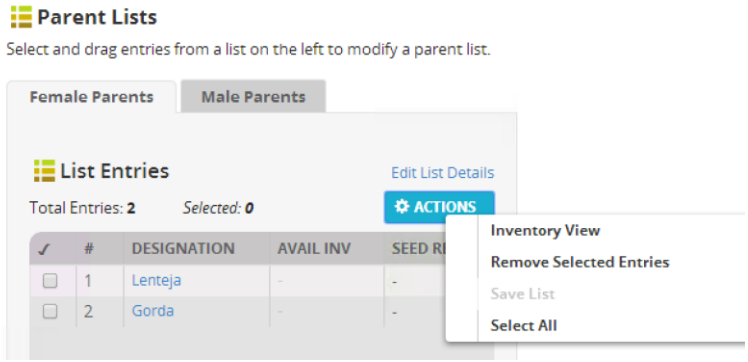
Select parents from established lists by dragging and dropping to the appropriate female and male parental lists.



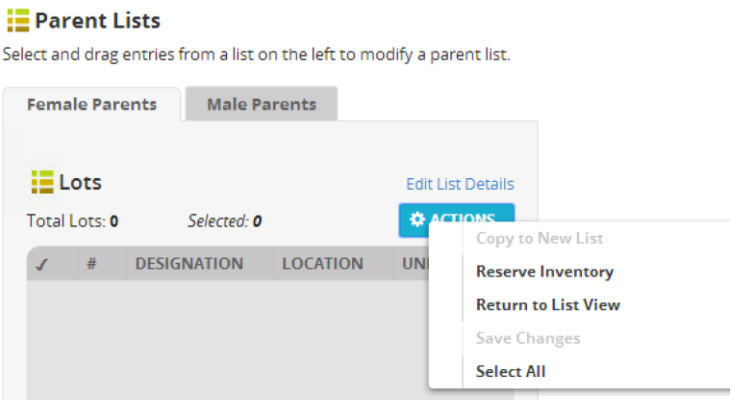
Save the parental lists to activate the save button on the crossing list.



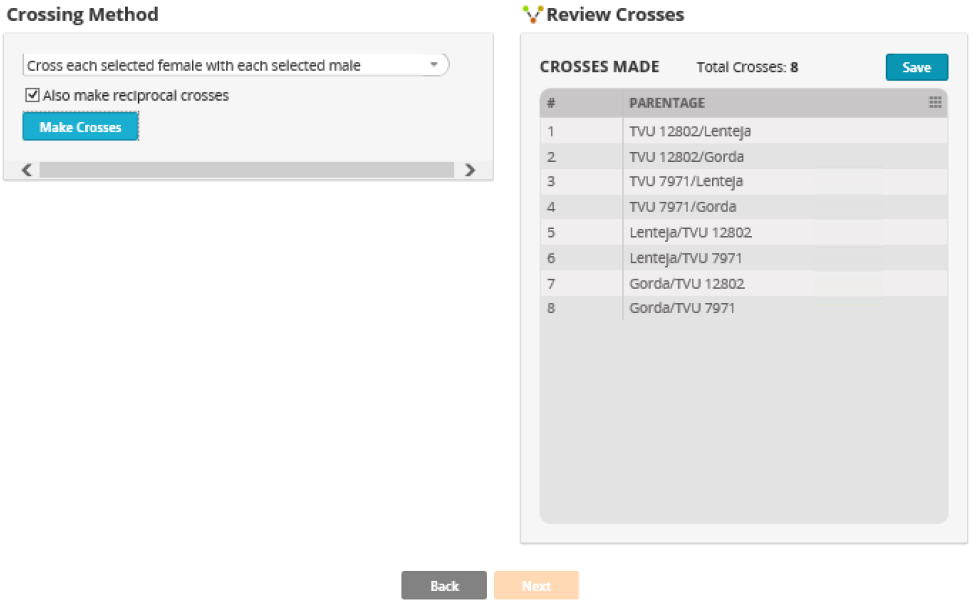
Review the seed inventory of the parental lists by selecting Inventory View from the Actions menu.



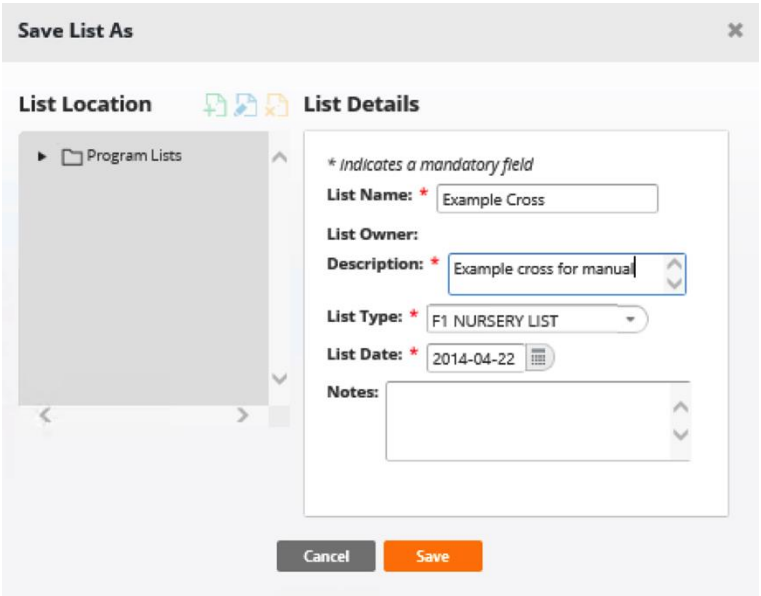
Reserve parental seed by selecting Reserve Inventory from the Actions menu in Inventory View.



Choose a crossing method from the drop down menu and select whether or not to make reciprocal crosses. Select Make Crosses to review the crossing list.

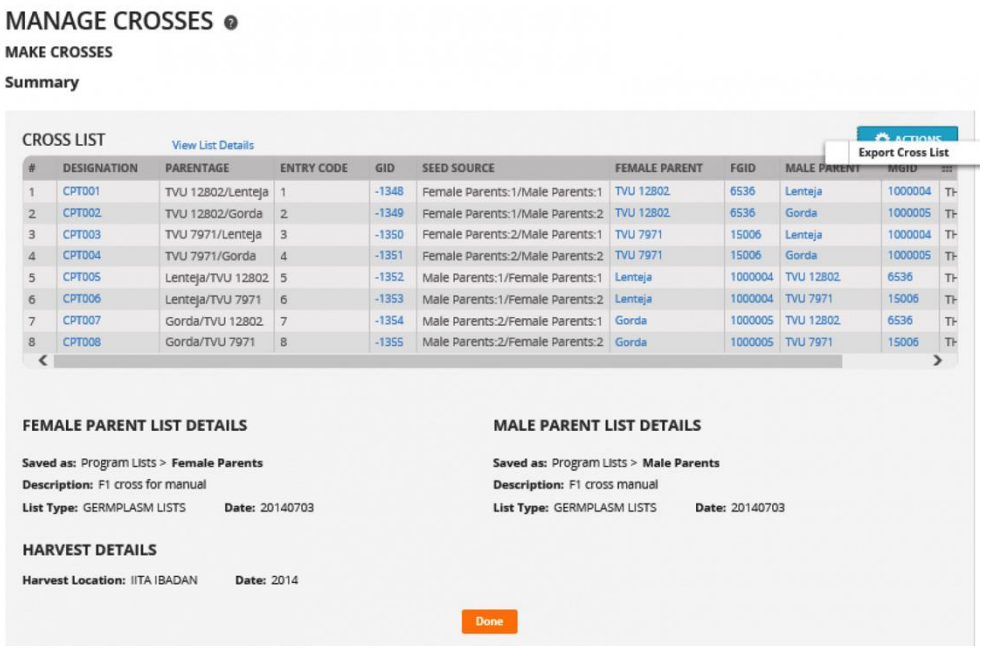


Save the crossing list and select Next.



Cross Summary & Export

Select Export Cross List from Actions Menu to export crossing list as a spreadsheet (.xls) file. Select Done to exit this cross.



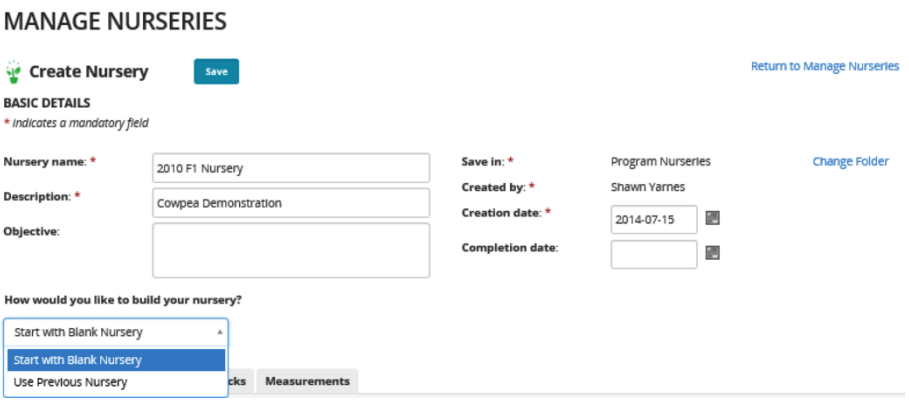
Overview of Nursery Manager

The nursery manager is a tool for designing nurseries. The nursery manager adds nurseries to the program database and provides the basis for developing field maps and labels to assist planting. Create data collection worksheets in a variety of formats, including digital recording and barcode scanning, and import back to the database directly.

Create Nursery

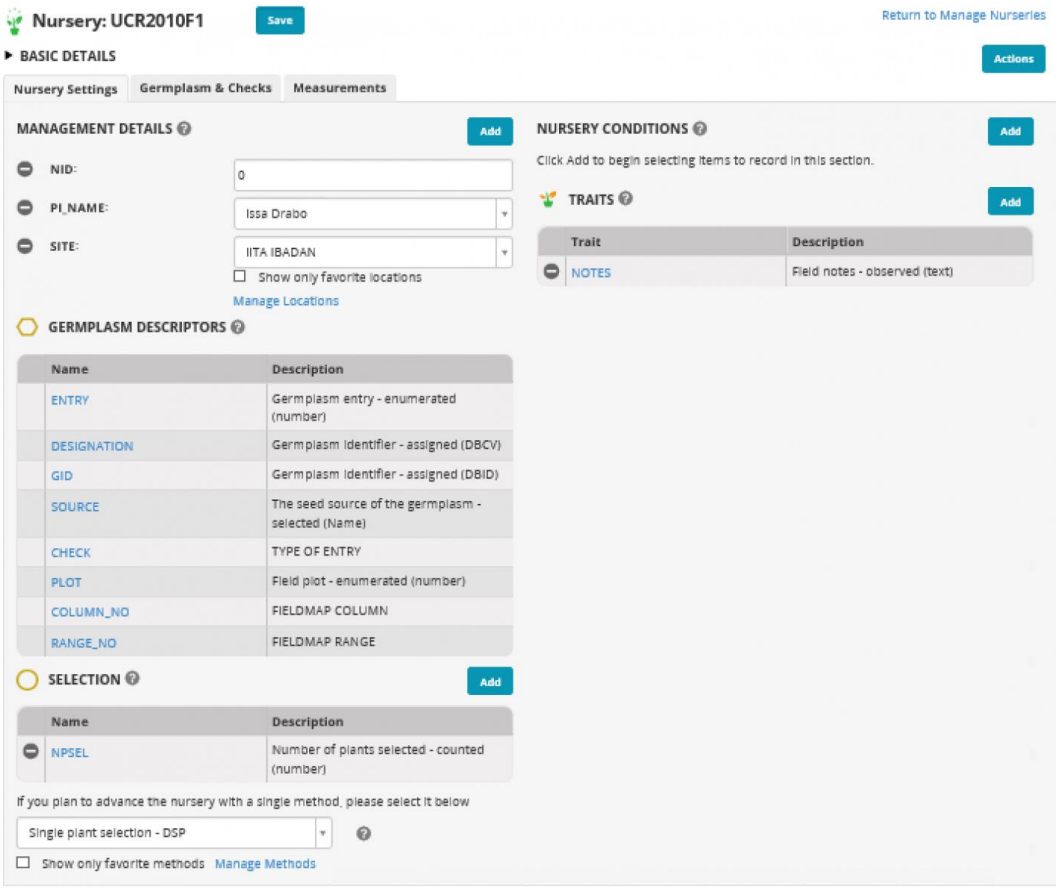
Launch Nursery Manager from the left side menu. Select Add a New Nursery.

Enter the basic details of the nursery. Start with a blank nursery or retrieve nursery information from a previous nursery. Once the nursery has a name and description, the nursery details can be saved.



Select nursery settings.

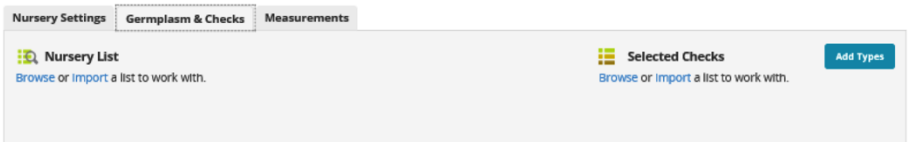
- Management details
- Germplasm descriptions
- Nursery conditions
- Traits
- Selection Criterion
- Breeding Method for Advance



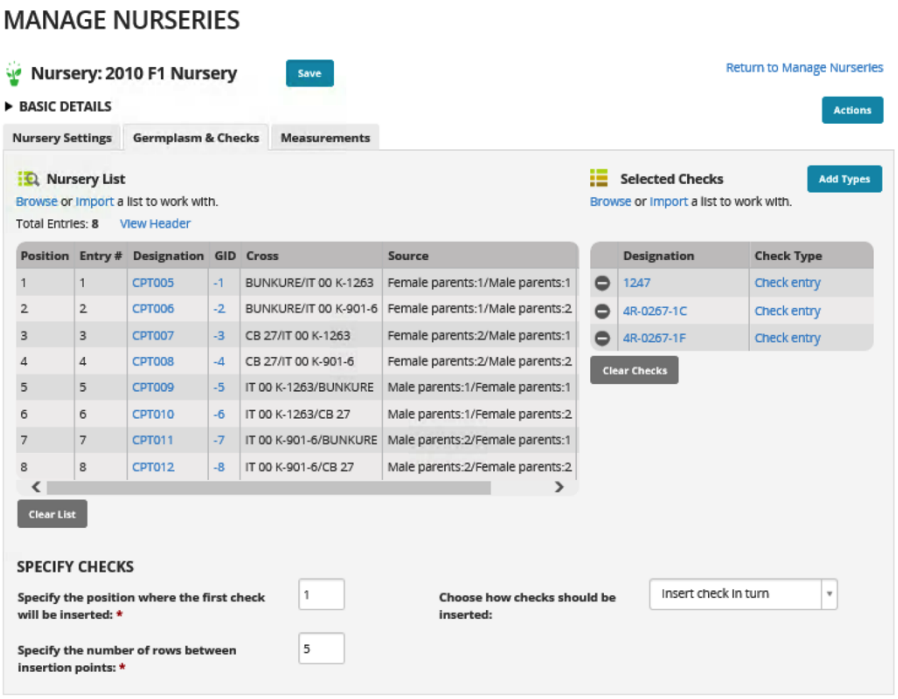
In this nursery no nursery conditions are specified. Plants will be described with field notes. Selections will be described by number of plants selected. The breeding method, single plant selection, has been chosen.

Assign Germplasm & Checks

Select the Germplasm & Checks tab to browse or import nursery lists and checks.

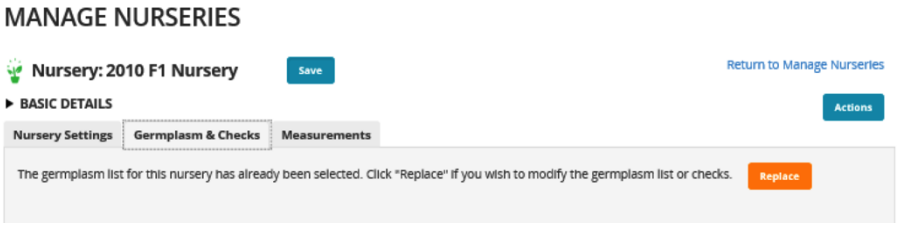


Specify the details of check positions.



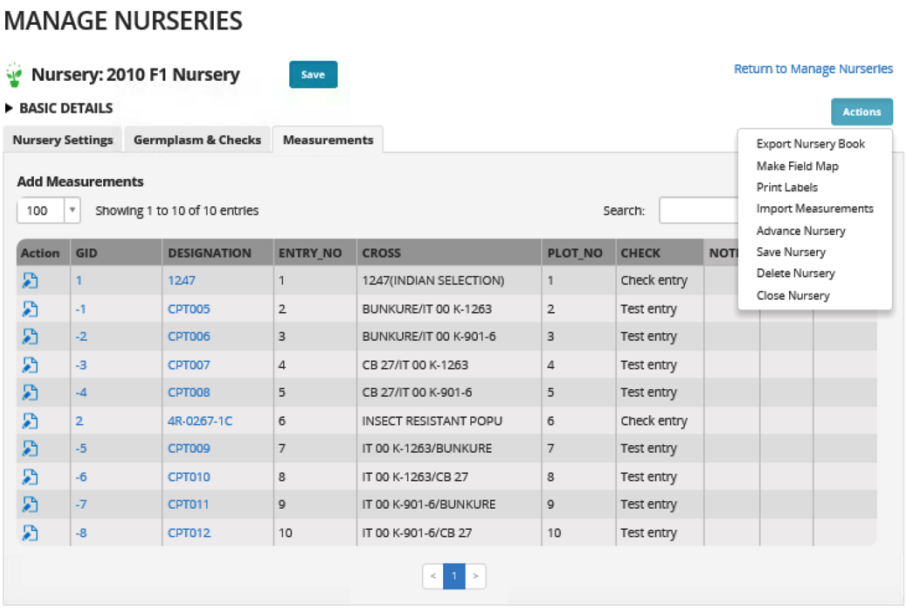
Save Nursery

Once germplasm and checks are saved to the nursery, field layout and labels can be designed. Data collection sheets are also available for printing or for download to a handheld data logger. Any changes to the germplasm or checks at this point will over write the associated nursery design elements.



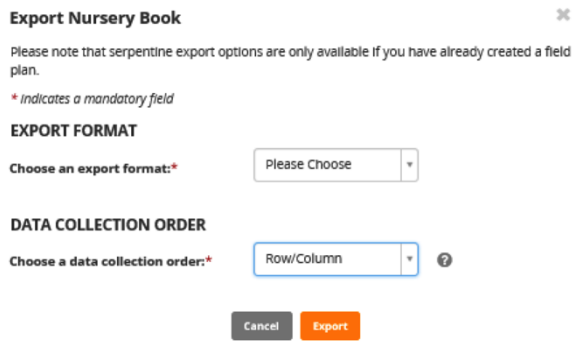
Nursery Management

After a nursery has been saved with germplasm, the Actions button reveals a dropdown menu of nursery management tasks.



Export Nursery Book

Select Export Nursery, and choose export format and data collection order from the dropdown menus.



Export Formats

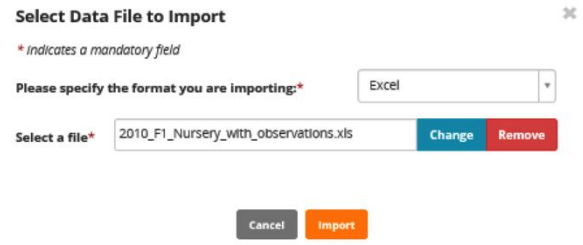
- Excel
- Data Kapture
- Fieldlog/Fieldriod
- KSU Fieldbook CSV
- KSU Fieldbook Excel
- R

Data Collection Orders

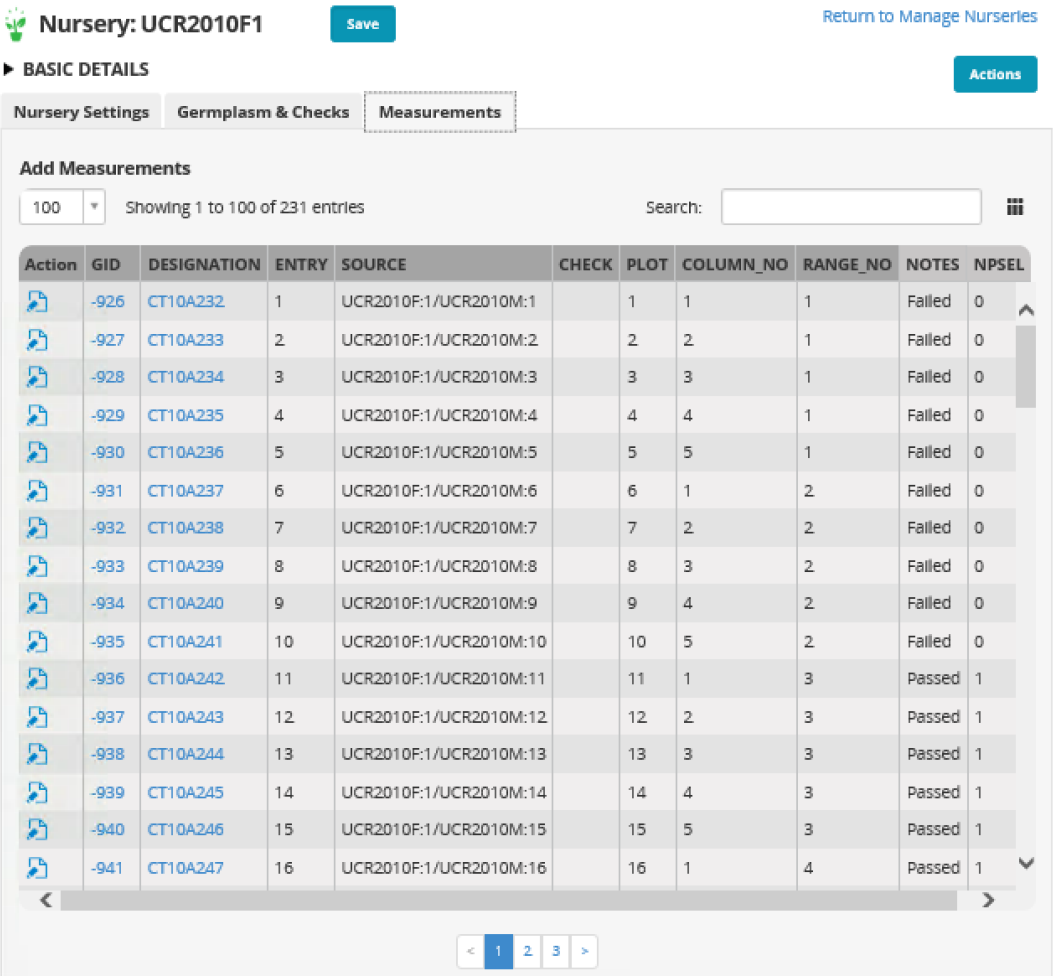
- Row/Column
- Serpentine Along Rows
- Serpentine Along Columns
- Import Measurements
- Select data file to import.

Import Selections & Observations

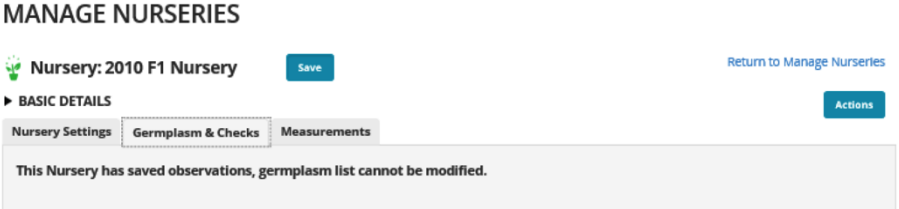
Select data file to import.



The data will load into the BMS. Select Save to upload the data to the BMS database.

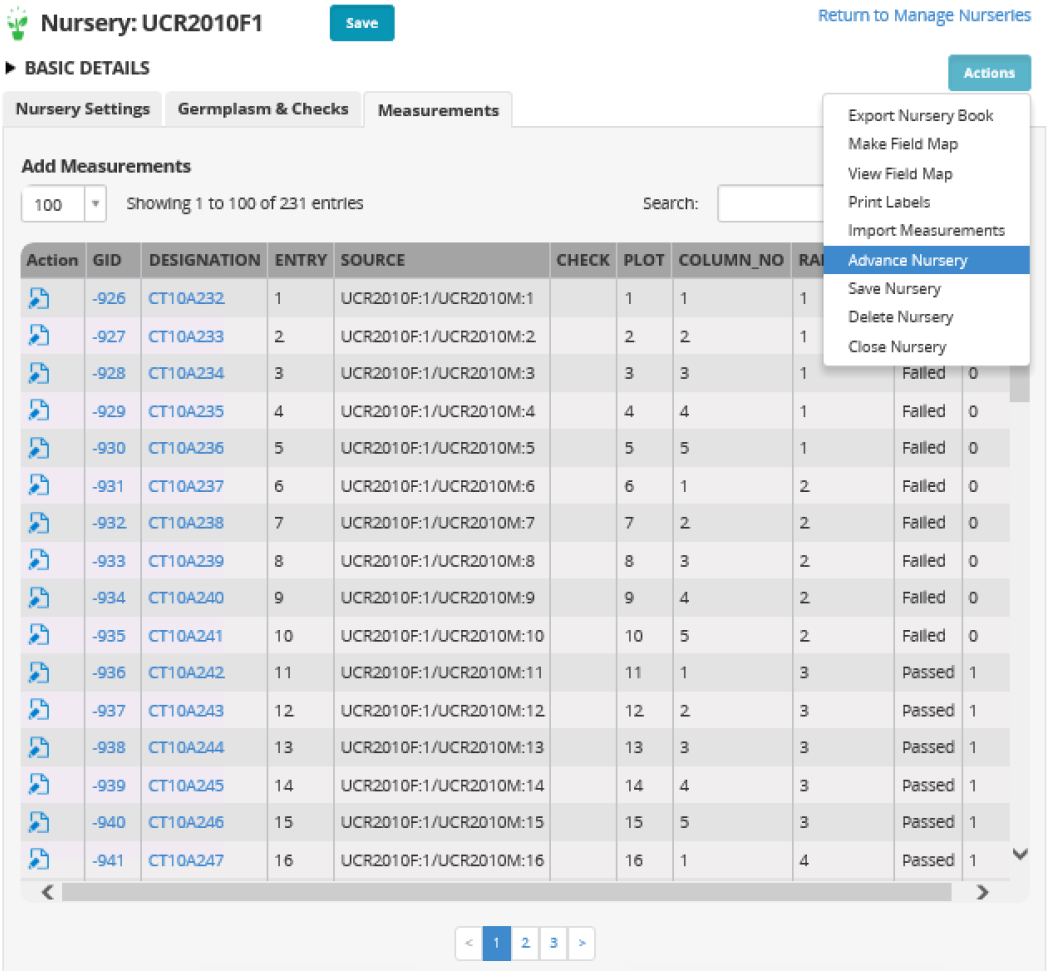


Once nursery data is loaded into the database, the associated germplasm and checks become locked.

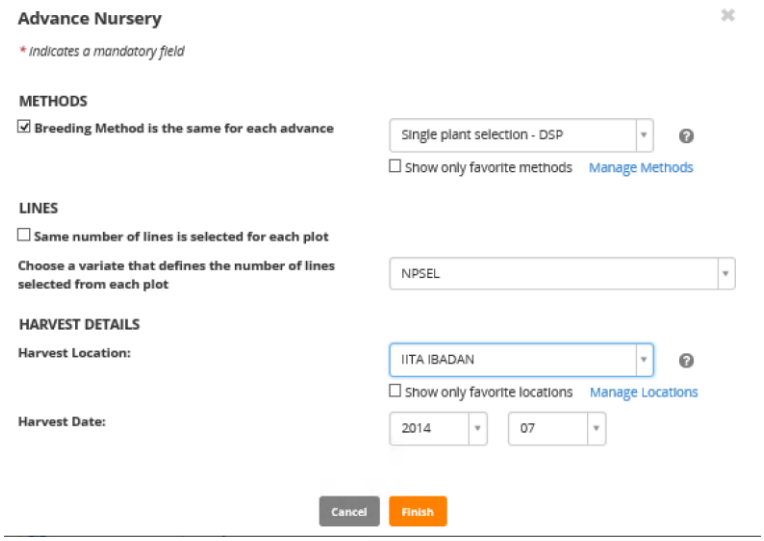


Advance Nurseries

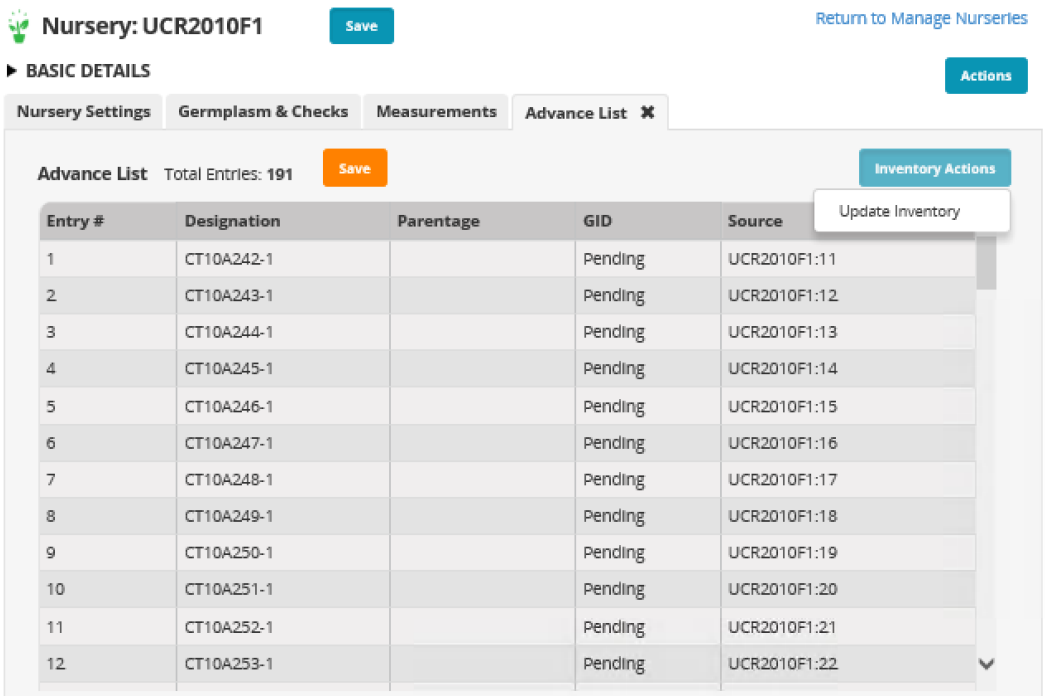
Select Advance Nursery from the Actions button dropdown menu.



Enter details about the nursery advancement.

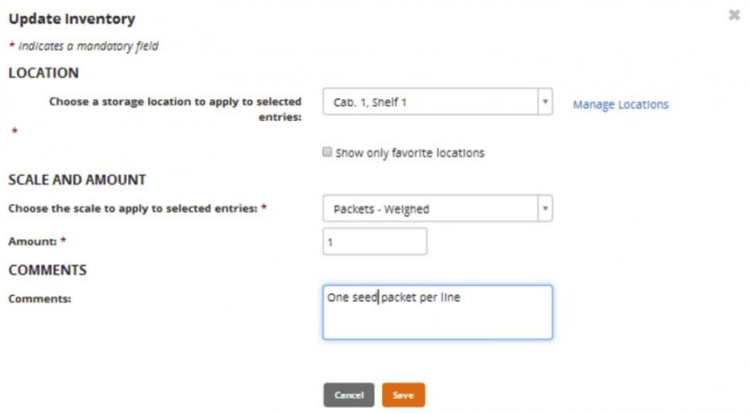


Save the Advance List, which in this case is the F2 list.



Update Seed Inventory

Highlight lines of interest for seed inventory and select Update Inventory from the Inventory Actions dropdown menu. Enter the seed inventory details for the selected germplasm.



Spreadsheet files can be exported from and imported into the Breeding Management System (See more on importing germplasm lists and trial and nursery data). BMS files (.xls) files contain description and observation worksheets corresponding to germplasm lists and data from nurseries and trials.

Trial Management: Cowpea Tutorial

Contributors

Shawn Yarnes, The Integrated Breeding Platform

Summary

This tutorial describes the design of a three-site field trial using a randomized complete block design with three replicates, the export of data collection sheets, and import of field observations back to the database.

Restore from Previous Tutorial

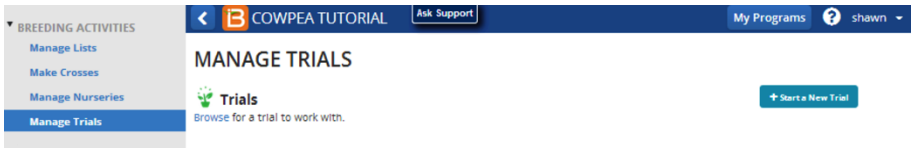
Screenshots and activities in this tutorial build upon work preformed in previous tutorials. If you are not following the cowpea tutorials in sequence, restore the Cowpea Tutorial database (.sql) to the end of the previous tutorial to match database contents with current tutorial.

BMS File Directory

C:\Breeding Management System\Documents\BMS Workbench_Nursery & Trial Managers\Tutorials\Cowpea\6.0Cowpea Tutorial.sql

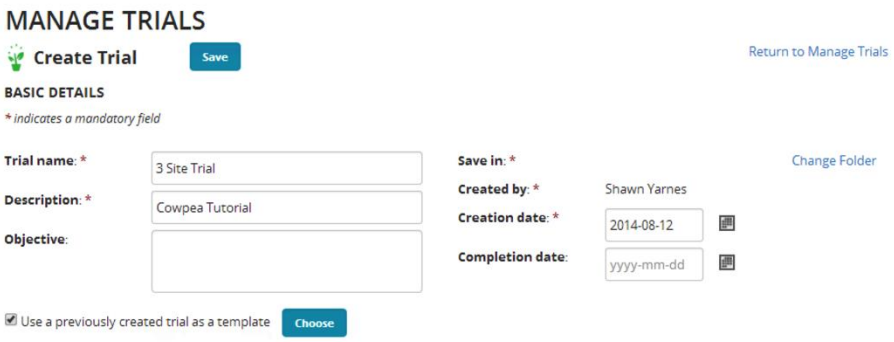
Overview of the Trial Manager

The trial manager is a tool for designing field trials. The Breeding Management System supports randomized complete block, incomplete block, and row-column experimental designs. The Trial Manager uses standardized crop ontology terms to describe treatments, factors, and trait measurements. Access the Trial Manager from Breeding Activities in the workbench menu. Select Start a New Trial.

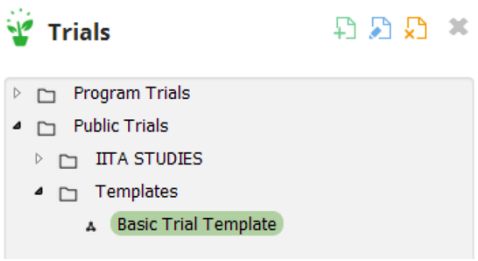


Basic Details

Enter the Trial Name, 3 Site Trial, and a Description. Leave the default settings for the other Basic Details. Select Use a previously created trial as template. At any point in trial design save progress by selecting the Save button.

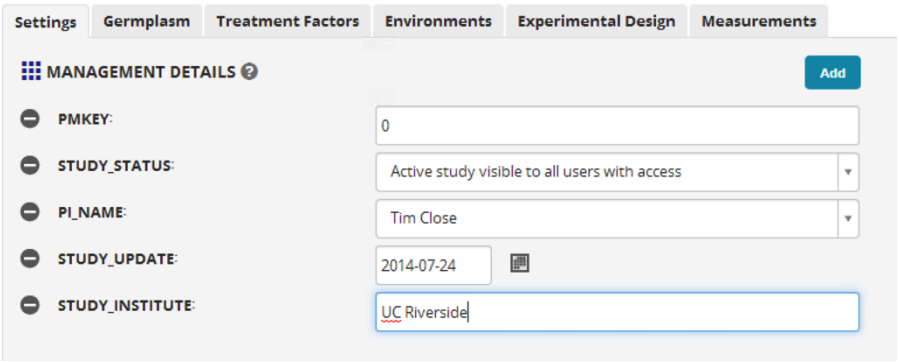


Choose the Basic Trial Template from the Public Trials folder. This will populate the trial with customizable trial design.



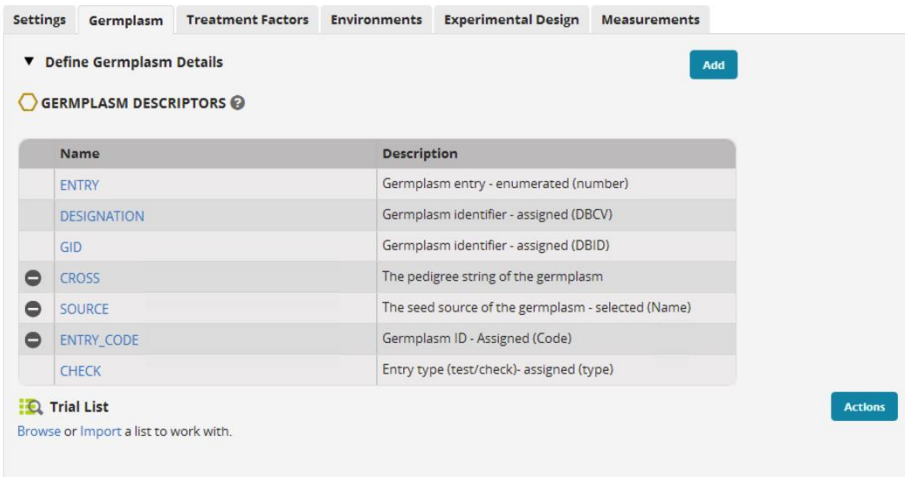
Settings

Under Trial Settings add a management detail not included in the template, STUDY_INSTITUTE, and type in UC Riverside. Select Tim Close from the dropdown menu as the principle investigator (PI_NAME).

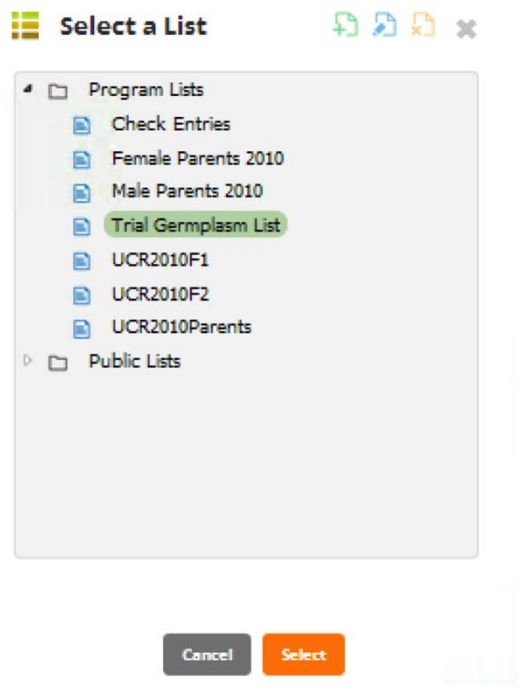


Germplasm

Browse for the Trial Germplasm List.



Select Trial Germplasm List.

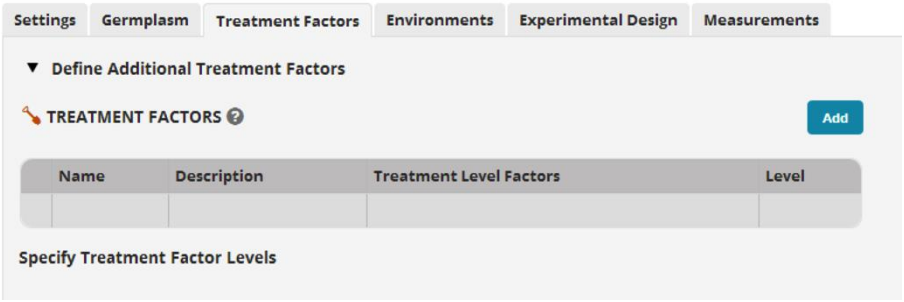


Review the list of 15 germplasm that will be included in the trial.

ENTRY	DESIGNATION	GID	CROSS	SOURCE	ENTRY_CODE	CHECK
1	ART 91-1	768	-		1	Test entry
2	ART 91-2	769	-		2	Test entry
3	IAR-48	13	-		3	Test entry
4	IT 87 D-1629	908	-		4	Test entry
5	IT 87 D-590-5	228	-		5	Test entry
6	IT 87 D-885	237	-		6	Test entry
7	IT 88 D-867-11	35617	-		7	Test entry
8	IT 89 KD-107-5	35859	-		8	Test entry

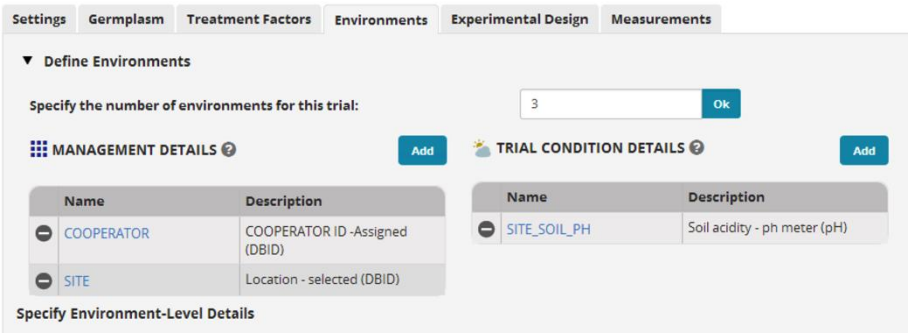
Treatment Factors

This trial will not have treatment factors, so leave empty.

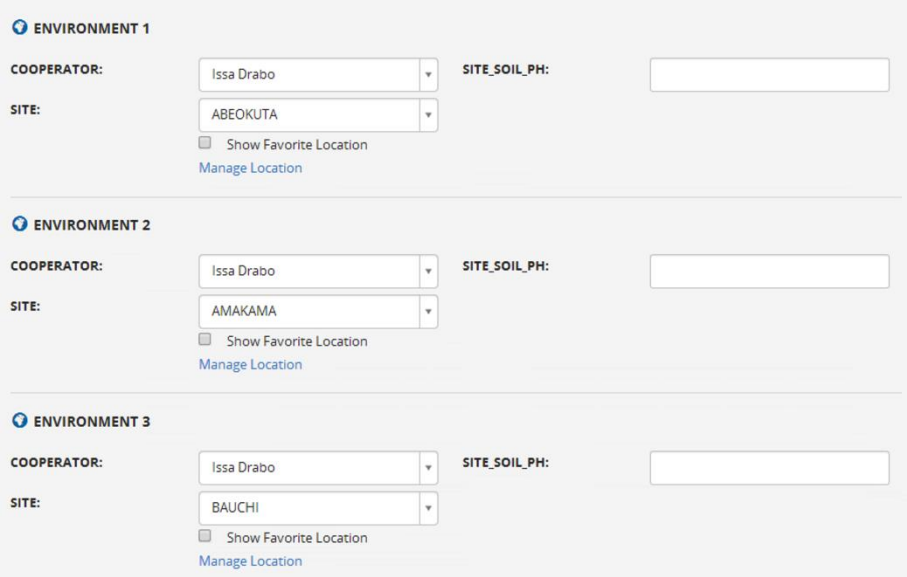


Environments

This is a three location Nigerian field trial. Specify 3 environments for this trial.

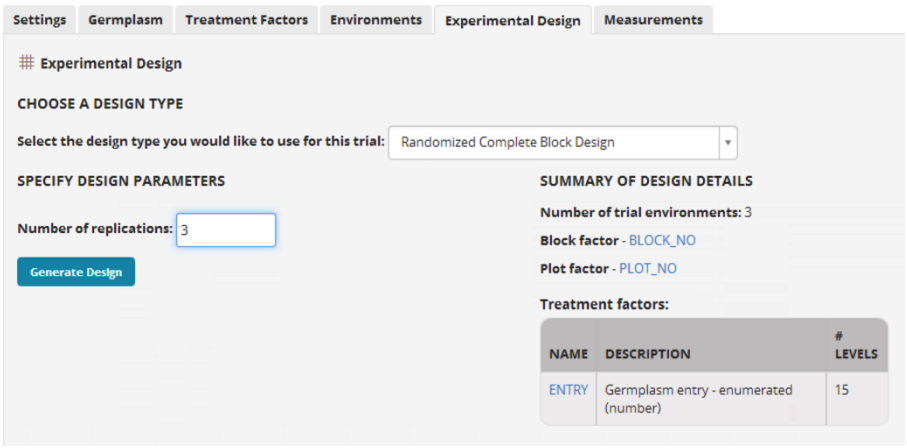


Add the three location names, Abeokuta, Amakama, and Bauchi, and the name of the cooperating breeder for these locations as Issa Drabo.



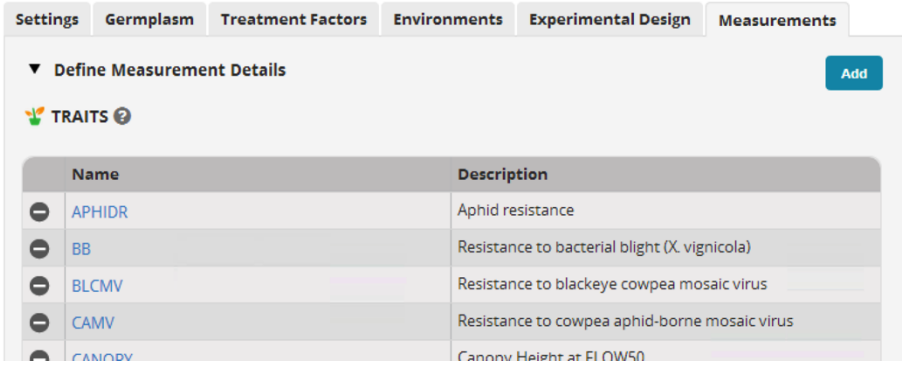
Experimental Design

Choose the Randomized Complete Block Design with 3 replications for this tutorial. Select generate design to populate the measurements tab.

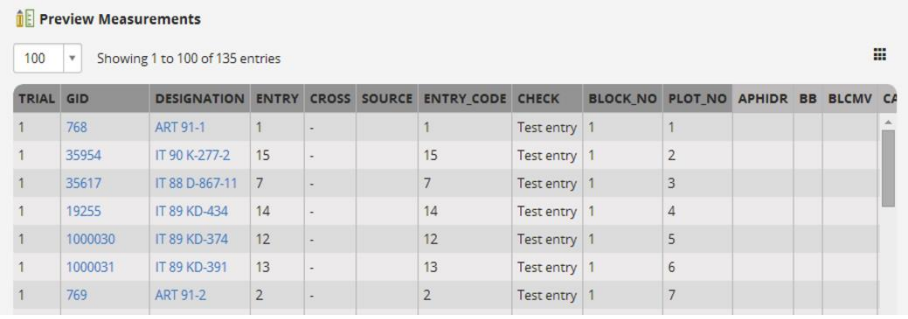


Measurements

The traits of interest in this trial are the weight of 100 seeds (g) and 50% maturity (days) (SDWT100 and MAT50). Review the list of traits preselected in the template. Notice that SDWT100 and MAT50 are preselected.



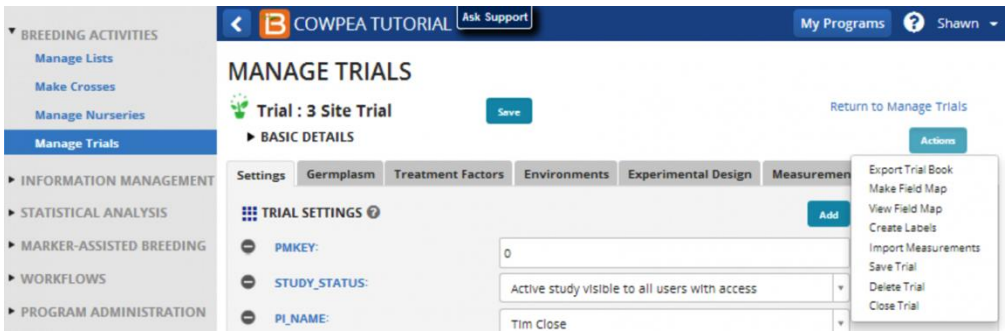
At this point the trial is now ready for field mapping, label printing, and file export for data collection. Select Save at the bottom of the page to save the empty trial design.



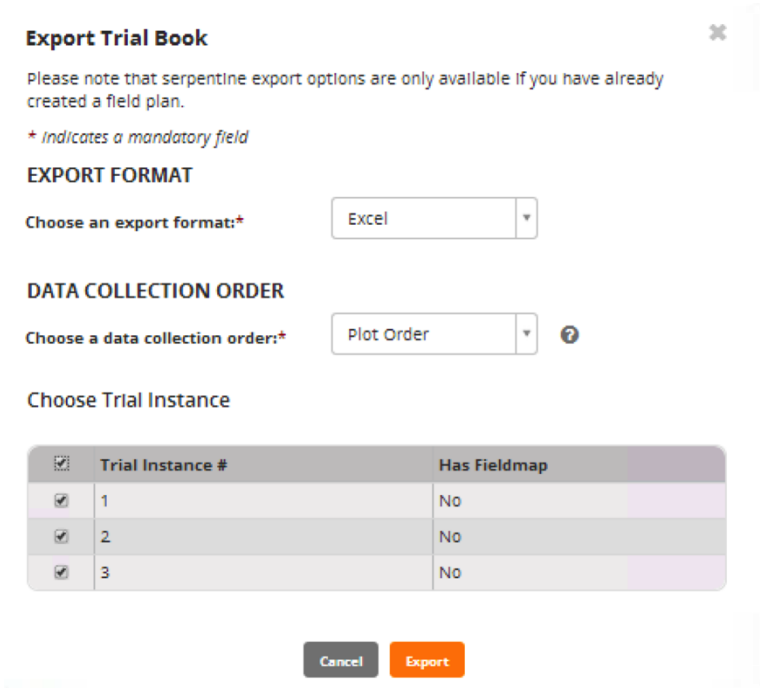
TRIAL	GID	DESIGNATION	ENTRY	CROSS	SOURCE	ENTRY_CODE	CHECK	BLOCK_NO	PLOT_NO	APHIDR	BB	BLCMV	CA
1	768	ART 91-1	1	-		1	Test entry	1	1				
1	35954	IT 90 K-277-2	15	-		15	Test entry	1	2				
1	35617	IT 88 D-867-11	7	-		7	Test entry	1	3				
1	19255	IT 89 KD-434	14	-		14	Test entry	1	4				
1	1000030	IT 89 KD-374	12	-		12	Test entry	1	5				
1	1000031	IT 89 KD-391	13	-		13	Test entry	1	6				
1	769	ART 91-2	2	-		2	Test entry	1	7				

Export Trial Book

Select Export Trial Book from the Actions button dropdown menu to export files for field data collection.



Export Excel files corresponding all three locations to a know location on you computer and open.



Restore Program

Since each randomized trial will be different due to randomization, you must restore your program so that the trial design matches the tutorial's pre-formated observation files.

BMS File Directory

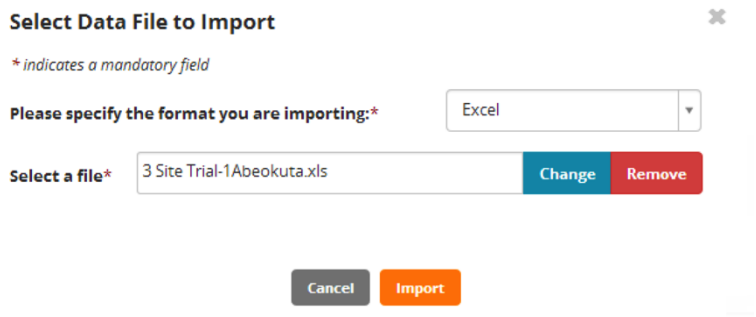
C:\Breeding Management System\Documents\BMS Workbench_Nursery & Trial Managers\Tutorials\Cowpea\6.1Cowpea Tutorial.sql

Import Trial Data

Select Import Measurements from the Actions button dropdown menu to import field observations. Browse and select to the preformatted Abeokuta data. Select Import. Save the trial after import.

BMS File Directory

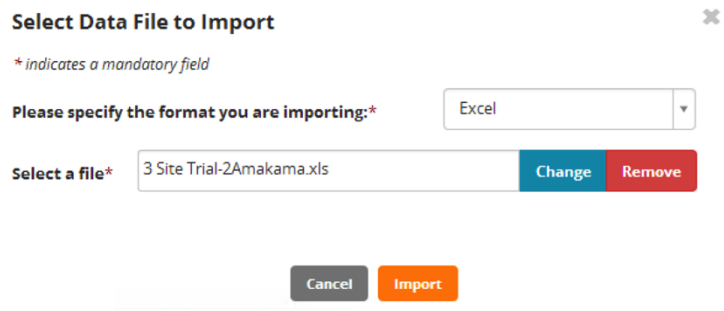
C:\Breeding Management System\Examples\Cowpea\Sample Data\3 Site Trial1-Abeokuta.xls



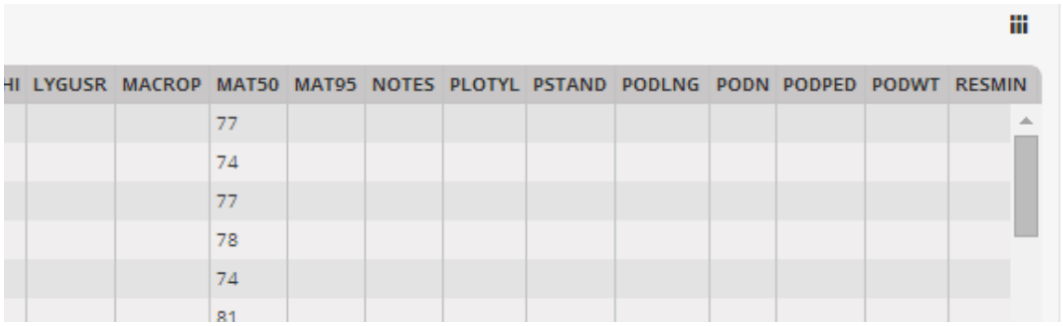
Import of the Amakama and Bauchi location data. Save the trial after import of each data file.

BMS File Directory

C:\Breeding Management System\Examples\Cowpea\Sample Data\3 Site Trial2-Amakama.xls
C:\Breeding Management System\Examples\Cowpea\Sample Data\3 Site Trial3-Bauchi.xls



Confirm that the data has been upload by scrolling the to the MAT50 and SDWT100 columns in the measurements table and select Save.



Acknowledgements

The statistical algorithms in the Breeding View were developed by VSN International Ltd in collaboration with the Biometris group at University of Wageningen. Cowpea demonstration data was provided by Jeff Ehlers, Tim Close, Philip Roberts, Bao Lam Huyuh at the University of California Riverside and Issa Drabo at the Institut de l'Environnement et de Recherches Agricoles in Burkina Faso. These data may have been adapted for training purposes. Any misrepresentation of the raw breeding data is the solely the responsibility of the IBP.

File Format

Description Sheet

The first worksheet contains a description of the germplasm list, cross, nursery, or trial.

	A	B	C	D	E	F	G	H
1	STUDY	F109A						
2	TITLE	EARLY-F1 population development						
3	PMKEY	0						
4	OBJECTIVE	EARLY-F1 population development						
5	START DATE	20120903						
6	END DATE	20170927						
7	STUDY TYPE	N						
8								
9	CONDITION	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	LABEL
10	NID	NURSERY SEQUENCE NUME	TRIAL INSTANCE	NUMBER	ASSIGNED	N		STUDY
11	PI_NAME	Name of the Principal Investiga	PERSON	DBCV	ASSIGNED	C	EHLERS JEFF	STUDY
12	PI_ID	ID of the Principal Investigator	PERSON	DBID	ASSIGNED	N	4	STUDY
13	BREEDING_METHOD	Breeding method to be applied	METHOD	DBCV	APPLIED	C	RANDOM BULK CF	STUDY
14	BREEDING_METHOD_ID	ID of Breeding method	METHOD	DBID	APPLIED	N	507	STUDY
15	SITE	NURSERY SITE NAME	LOCATION	DBCV	ASSIGNED	C	IITA IBADAN	STUDY
16	SITE_ID	NURSERY SITE ID	LOCATION	DBID	ASSIGNED	N	20054	STUDY
17								
18	FACTOR	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	LABEL
19	ENTRY	ENTRY NUMBER	GERMPLASM ENTRY	NUMBER	ENUMERATED	N	STUDY	ENTRY
20	DESIGNATION	ENTRY DESIGNATION	GERMPLASM ID	DBCV	ASSIGNED	C	STUDY	ENTRY
21	GID	GERMPLASM ID	GERMPLASM ID	DBID	ASSIGNED	N	STUDY	ENTRY
22	SOURCE	The seed source of the germpl	SEED SOURCE	NAME	SELECTED	C	STUDY	ENTRY
23	CHECK	TYPE OF ENTRY	CHECK	CODE	ASSIGNED	C	STUDY	ENTRY
24	PLOT	PLOT NUMBER	FIELD PLOT	NUMBER	ENUMERATED	N	STUDY	PLOT
25								
26	CONSTANT	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	LABEL
27								
28	VARIATE	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE	LABEL
29	NOTES	BREEDERS NOTES	Comment	Text	Observed	C	STUDY	STUDY
30	NPSEL	NUMBER OF PLANTS SELEC	PLANTS SELECTED	Number	Counted	N	STUDY	STUDY
31								

Description Sheet: Describes a cowpea nursery

Observation Sheet

The second worksheet, titled Observation, contains the observation data. Column headings are specified by ontology terms recognized by the database.

	A	B	C	D	E	F	G	H
1	ENTRY	DESIGNAT	GID	SOURCE	CHECK	PLOT	NOTES	NPSEL
2	1	CT10A001	-4	UCR2010F:1/UCR201(		1	Failed	0
3	2	CT10A002	-5	UCR2010F:1/UCR201(		2	Failed	0
4	3	CT10A003	-6	UCR2010F:1/UCR201(		3	Failed	0
5	4	CT10A004	-7	UCR2010F:1/UCR201(		4	Failed	0
6	5	CT10A005	-8	UCR2010F:1/UCR201(		5	Failed	0
7	6	CT10A006	-9	UCR2010F:1/UCR201(		6	Failed	0
8	7	CT10A007	-10	UCR2010F:1/UCR201(		7	Failed	0
9	8	CT10A008	-11	UCR2010F:1/UCR201(		8	Failed	0
10	9	CT10A009	-12	UCR2010F:1/UCR201(		9	Failed	0
11	10	CT10A010	-13	UCR2010F:1/UCR201(		10	Failed	0
12	11	CT10A011	-14	UCR2010F:1/UCR201(		11	Passed	1
13	12	CT10A012	-15	UCR2010F:1/UCR201(		12	Passed	1
14	13	CT10A013	-16	UCR2010F:1/UCR201(		13	Passed	1
15	14	CT10A014	-17	UCR2010F:1/UCR201(		14	Passed	1
16	15	CT10A015	-18	UCR2010F:1/UCR201(		15	Passed	1

Observation Sheet: Contains the nursery information and observations

Information Management

Import Germplasm

File Format

All BMS import and export files contain two spreadsheets: (1) Description and (2) Observation. Germplasm import templates require that the columns ENTRY and DESIGNATION are filled on the observation sheet. ENTRY is a numerical sequence and DESIGNATION is the name of the germplasm in your breeding program. The basic template only contains these two columns. The advanced template contains more columns for the import of optional geneaology data.

BMS File Directory
C:\Breeding Management System\Examples\Cowpea\templates\GermplasmImportTemplate-basic-rev4b.xls
C:\Breeding Management System\Examples\Cowpea\templates\GermplasmImportTemplate-advanced-rev4.xls

Advanced template optional data:

GID: Specific to the BMS database

CROSS: Parental pedigrees (female parent/male parent)

SOURCE

ENTRY CODE

LIST NAME	Parental Germplasm MABC_Lam
LIST DESCRIPTION	MABC project on aphid resistance at UCR
LIST DATE	20140225
LIST TYPE	LST

CONDITION	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	VALUE
LIST OWNER	Name of the Principal Investigator	PERSON	DBCV	ASSIGNED	C	
ID OF LIST OWNER	ID of the Principal Investigator	PERSON	DBID	ASSIGNED	N	

FACTOR	DESCRIPTION	PROPERTY	SCALE	METHOD	DATA TYPE	NESTED IN
ENTRY	The germplasm entry number	GERMPLASM ENTRY	NUMBER	ENUMERATED	N	
DESIGNATION	The name of the germplasm	GERMPLASM ID	DBCV	ASSIGNED	C	
GID	The GID of the germplasm	GERMPLASM ID	DBID	ASSIGNED	N	
CROSS	The pedigree string of the germplas	CROSS NAME	NAME	ASSIGNED	C	
SOURCE	The seed source of the germplasm	SEED SOURCE	NAME	Seed Source	C	
ENTRY CODE	Germplasm entry code	GERMPLASM ENTRY	CODE	ASSIGNED	C	

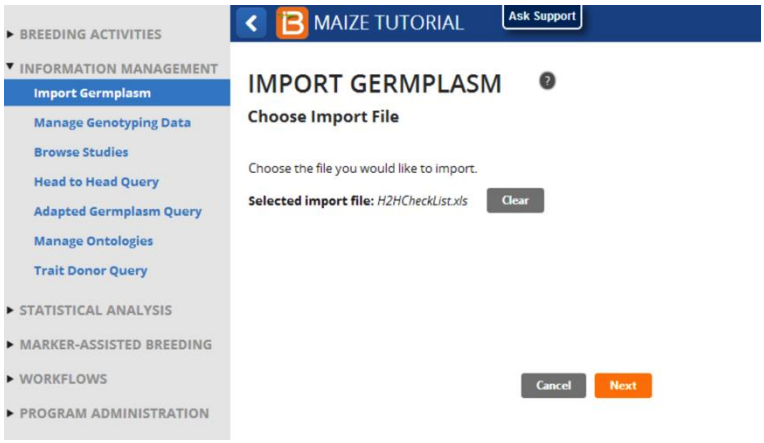
Description sheet customized from the advanced template to include list name, description, date, and type

ENTRY	DESIGNATION	GID	CROSS	SOURCE	ENTRY CODE
1	CB27			UCR	
2	RIL39		CB27/IT97K-556-6	UCR	
3	2012-077		CB27/RIL39	UCR	

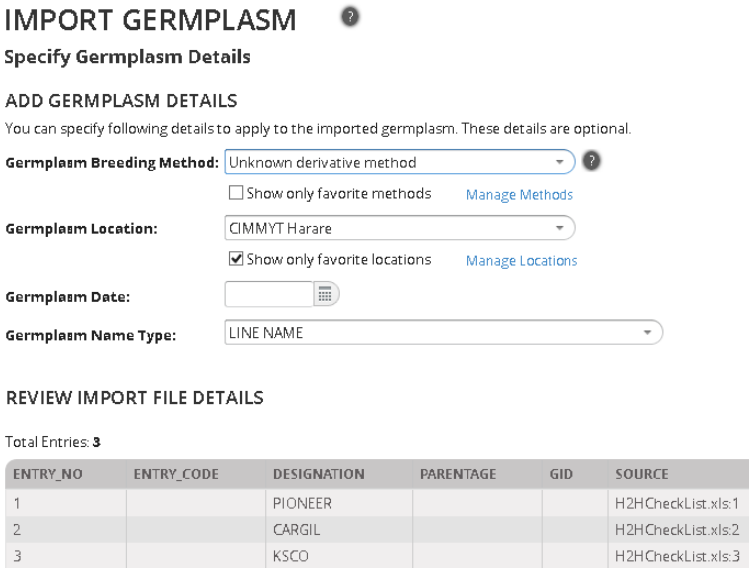
Observation sheet with mandatory data, ENTRY and DESIGNATION, and optional CROSS and SOURCE information

Import Germplasm Tool

The Import Germplasm tool adds germplasm lists from Excel file, with or without additional genealogy data, to the database by matching to existing GIDs or assigning new GIDs. (See more on importing phenotypic and maker data.) Importing a germplasm list ensures that each germplasm entry is assigned a GID (germplasm ID), which will allow pedigree and genotype tracking by the database. Browse a properly formatted Excel Germplasm List. Select Upload, and click Next.



Review the import file data and specify the germplasm details.



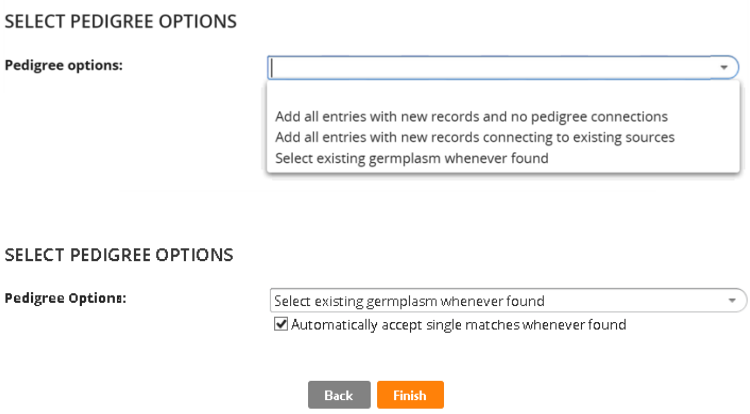
Select pedigree options.

Add all entries with new records and no pedigree connections: All entries receive new GIDs with no pedigree connections. This option is only available using the basic germplasm template.

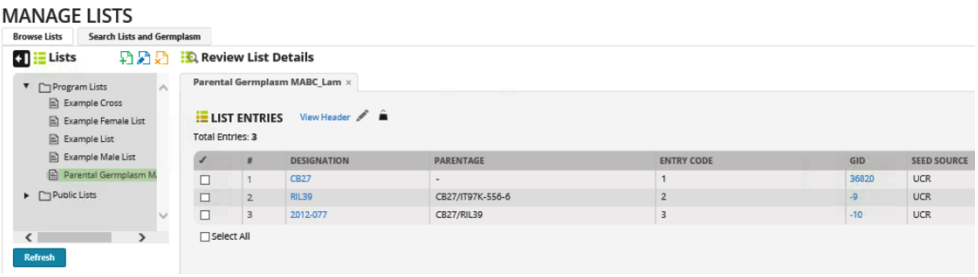
Add all entries with new records connecting to existing sources: All entries receive new GIDs with pedigree connections where appropriate

Select existing germplasm whenever found: Entries will be matched to the GIDs and pedigrees of existing germplasm whenever found.

Automatically accept single matches whenever found: Permit the BMS to automatically import single database matches without user confirmation.



When you return to the List Manager the new list will now appear with GIDs. Positive GIDs represent matches to germplasm in the database. Negative GIDs represent program-specific germplasm.

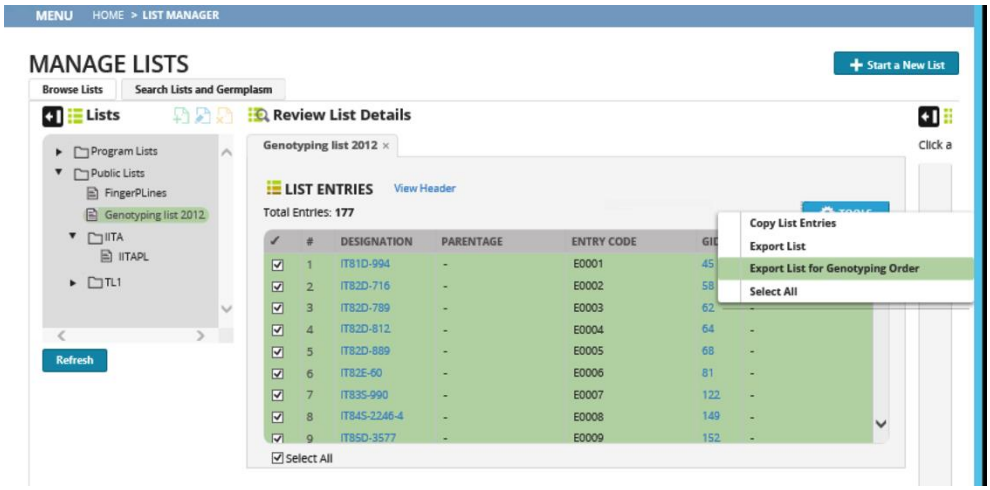


One germplasm designation, CB27, matches germplasm in the database and is represented by the positive GID, 36820. The other germplasm entries are new and represented by newly assigned negative GIDs.

Manage Genotyping Data

Export Genotyping Order Form

Export an Excel formatted genotyping order form from the Tools button.



The genotyping order form will provide a plate map (96 well) of the samples that need to be prepared and sent to the genotype service provider. 'Subject ID' corresponds to the Entry # associated with germplasm in the BSM database. 'Plate ID' is the list name plus plate number. 'Well' refers to grid coordinates on a 96 well plate.

	A	B	C	D	E	F	G	H
1	Subject ID	Plate ID	Well	Sample type	96	Primer	Subject BC	Plate BC
2	1	Genotyping list 2012-1	A01					
3	2	Genotyping list 2012-1	A02					
4	3	Genotyping list 2012-1	A03					
5	4	Genotyping list 2012-1	A04					
6	5	Genotyping list 2012-1	A05					
7	6	Genotyping list 2012-1	A06					
8	7	Genotyping list 2012-1	A07					
9	8	Genotyping list 2012-1	A08					
10	9	Genotyping list 2012-1	A09					

SNP Genotype Data Format

Sequence data will be reported back in spreadsheet format (example file .xls) where Subject IDs (first columns of data) are matched to marker sequence in a genotype by marker matrix.

	A	B	C	D	E	F	G	H	I
1	KBiosciences grid report								
2	Grid version 1.02								
3	More information is available in the Genotyping-1265.038-01.csv file. This file lists only the calls for each SNP on each well with a subject ID. When a subject ID is duplicated and the calls don't match the keyword DUPE is used.								
4	Project number	1							
5	Order number	1							
6	Plates	1							
7									
8	DNA \ Assay	10969_452	3194_319	8944_1233	894_153	4780_374	16822_160	17067_979	4146_1588
9	1	G:G	G:G	C:C	A:A	T	C:C	?	A:A
10	2	A:A	A:A	C:C	A:A	C	T:T	A:A	T:T
11	3	A:A	A:A	A:A	G:G	T	T:T	A:A	T:T
12	4	A:A	A:A	A:A	G:G	C	T:T	A:A	T:T
13	5	A:A	A:A	A:A	G:G	T	C:C	A:A	T:T
14	6	G:G	G:G	A:A	G:G	T	T:T	C:C	A:A
15	7	G:G	G:G	C:C	A:A	C	T:T	A:A	T:T

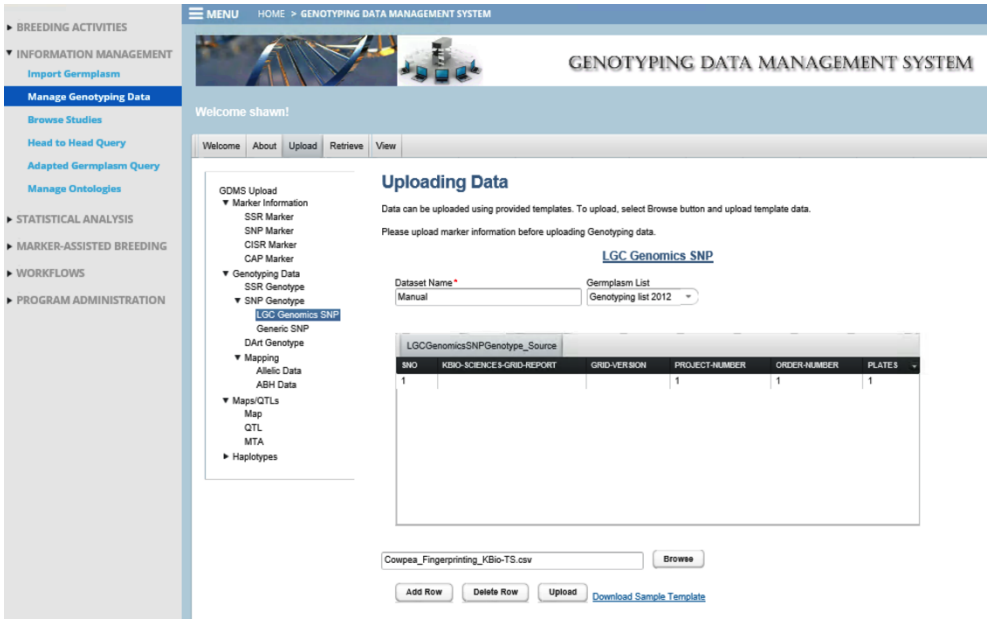
This file of sequence data contains the sequence data from 177 germplasm accessions and 1059 SNP markers (marker IDs are highlighted grey).

The DNA Assay Column will need to be more specific to upload to the database. In Excel, type the name of the list exactly as it appears in the BMS database followed by a dash and the Source ID. Copy cell and paste down the entire column. Excel will continue the Source IDs sequentially. The DNA Assay column is now replaced with the genotyping list name followed by dash Source ID. Save the modified Excel file.

	A	B	C	D	E
7					
8	DNA \ Assay	10969_452	3194_319	8944_1233	894_153
9	Genotyping list 2012-1	G:G	G:G	C:C	A:A
10	Genotyping list 2012-2	A:A	A:A	C:C	A:A
11	Genotyping list 2012-3	A:A	A:A	A:A	G:G
12	Genotyping list 2012-4	A:A	A:A	A:A	G:G
13	Genotyping list 2012-5	A:A	A:A	A:A	G:G
14	Genotyping list 2012-6	G:G	G:G	A:A	G:G
15	Genotyping list 2012-7	G:G	G:G	C:C	A:A
16	Genotyping list 2012-8	A:A	A:A	C:C	A:A
17	Genotyping list 2012-9	G:G	G:G	A:A	G:G
18	Genotyping list 2012-10	G:G	G:G	C:C	A:A
19	Genotyping list 2012-11	A:A	A:A	C:C	A:A

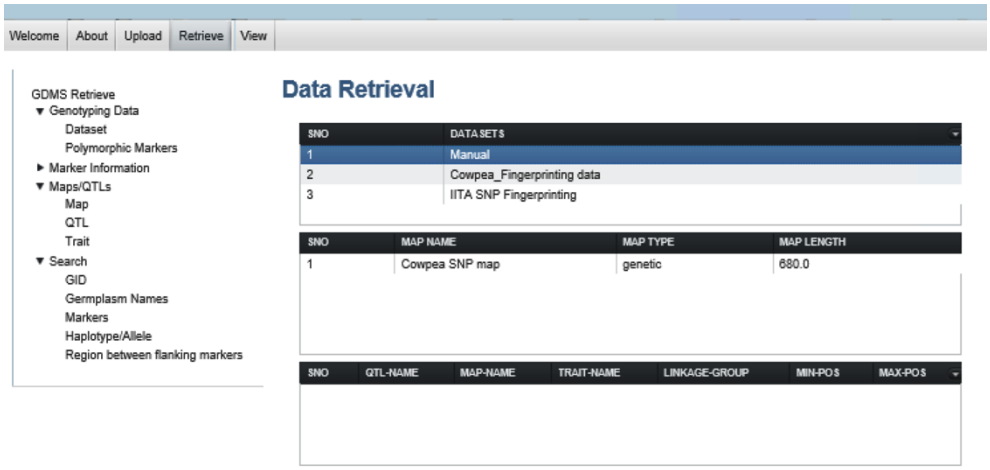
Import SNP Genotype Data

Open Manage Genotyping Data, and select the Upload tab. Choose the LGC Genomics SNP option. Keep in mind that other DNA service providers will provide similarly formatted sheets of data, so you are not limited to LCG Genomics data. Select the Genotyping List from the drop down menu. Choose a Dataset Name. Browse for the formatted file to import and select Upload.

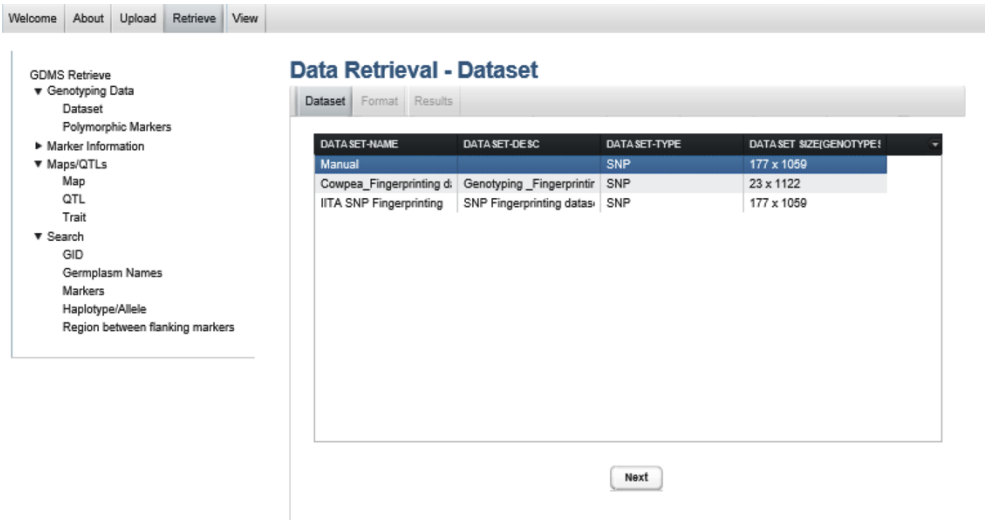


Retrieve Dataset

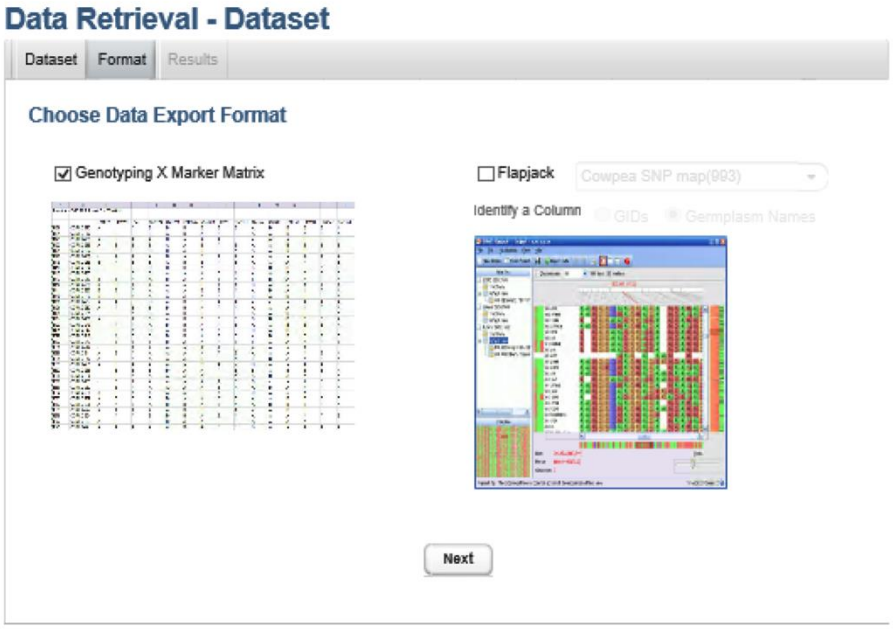
Newly imported datasets can be found under the Retrieve tab.



Export genotyping data by selecting Dataset. Highlight the dataset of interest, and click Next.



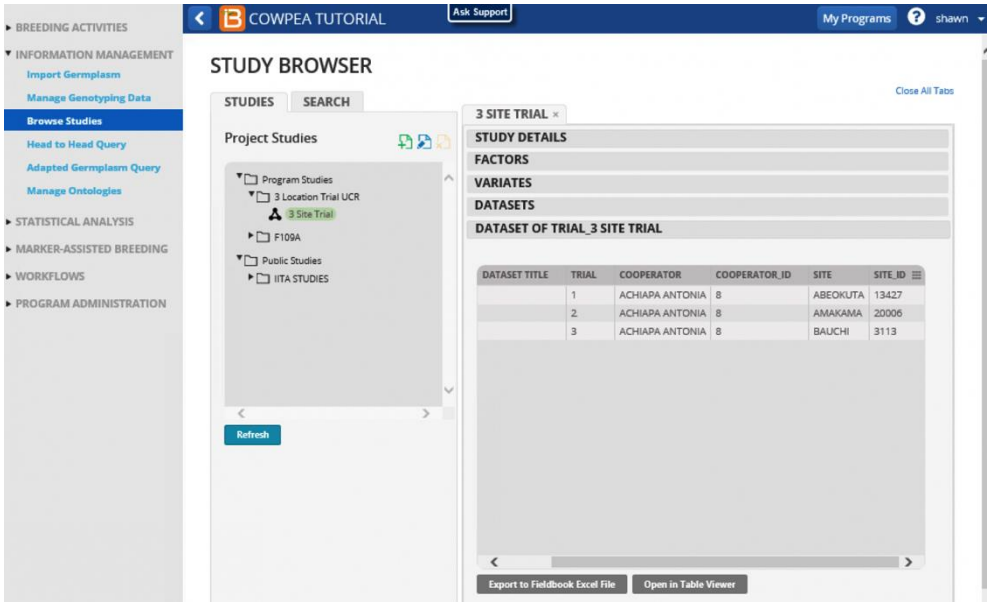
Choose to export the dataset as a genotyping by marker matrix or as a Flapjack formatted file.



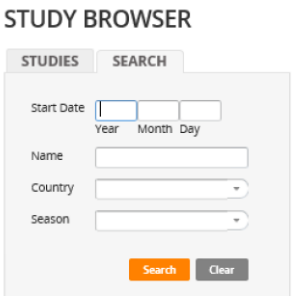
Download the selected export file.

Browse Studies

Select Browse Studies to browse and search the details of trials and nurseries.

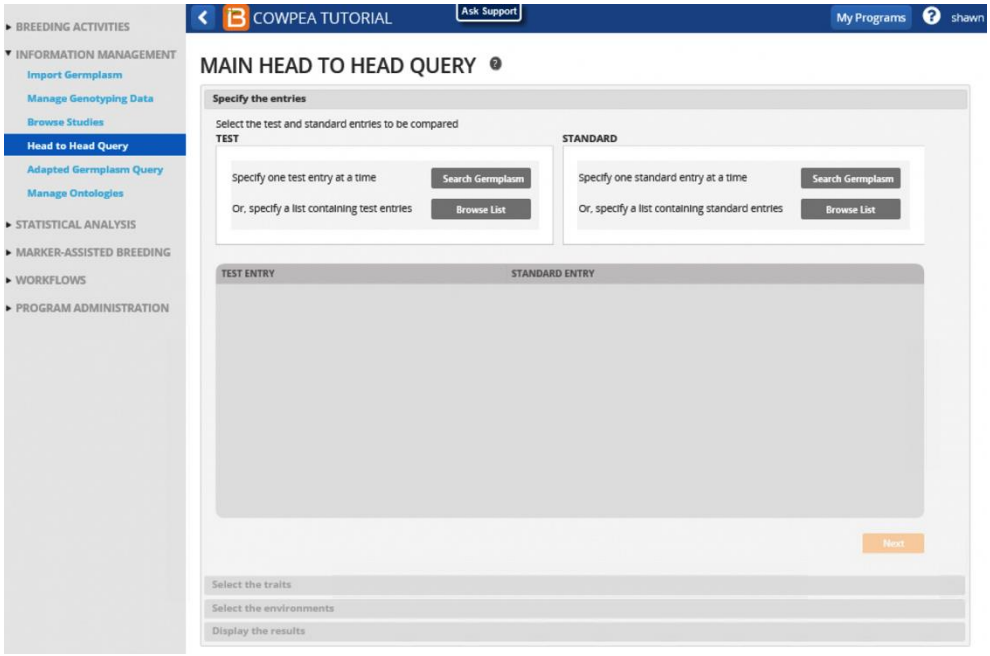


Use the Search tab to find studies based on date, name, country, and season. Double click on the appropriate search result to open the details in the right hand pane.



Head to Head Query

A head to head query allows users to compare germplasm grown in the same environment based on performance for a trait of interest. The query does a pairwise comparison of the test and standard germplasm. The result is based on the proportion of environments where the test germplasm surpassed the standard germplasm for a particular trait.

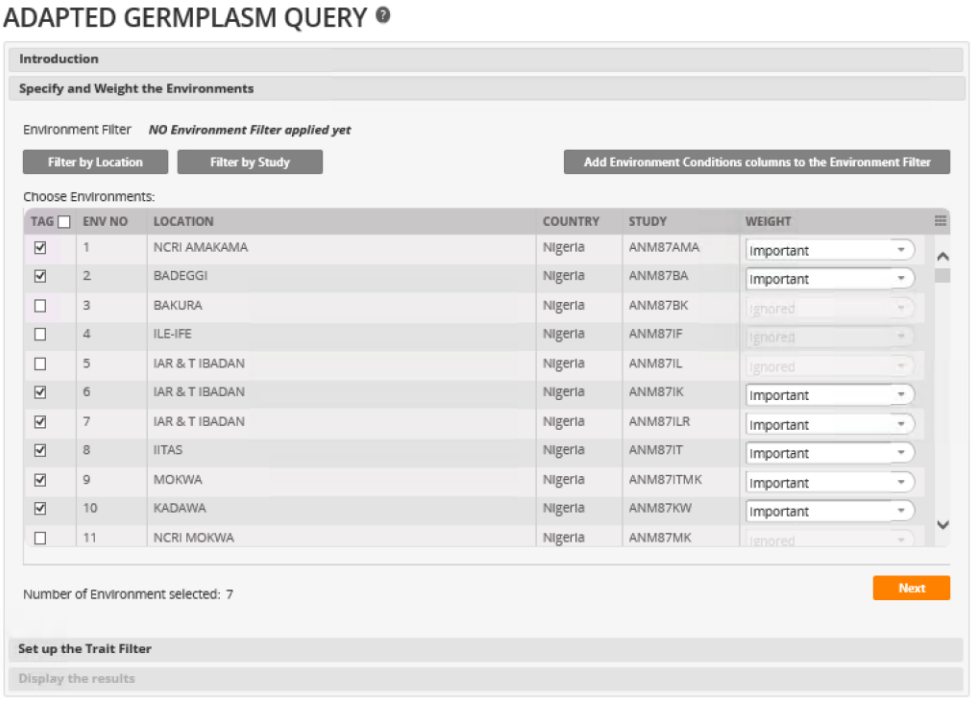


Adapted Germplasm Query

Find germplasm suitable for different environmental conditions using the adapted germplasm query. The query filters trial environments and finds the best performing germplasm for selected traits. Select Next.

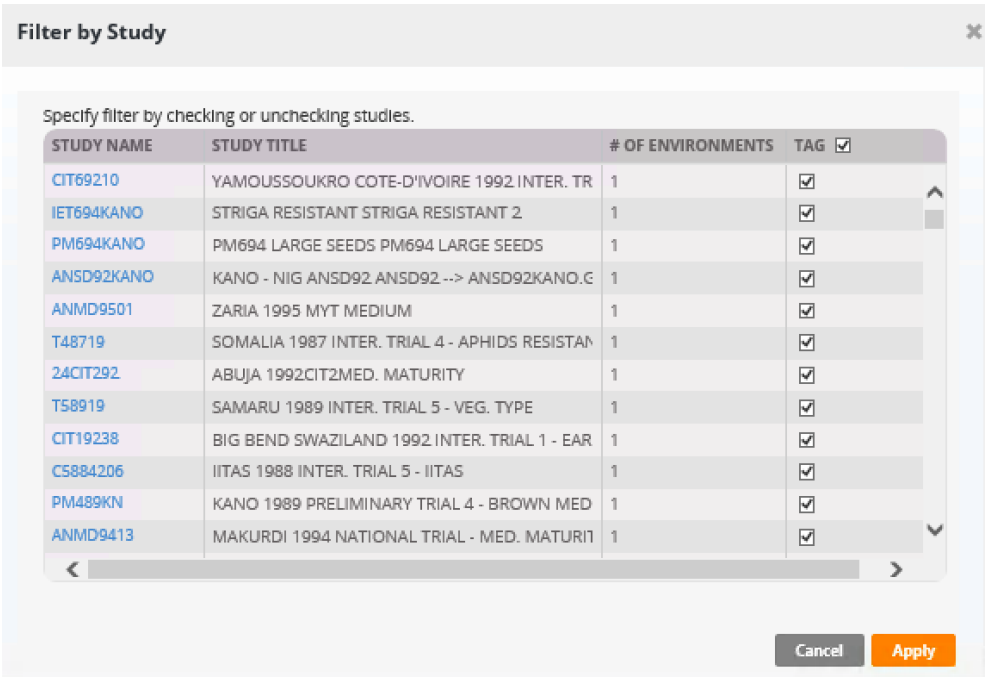
Weight & Environments

Choose and weight environments.

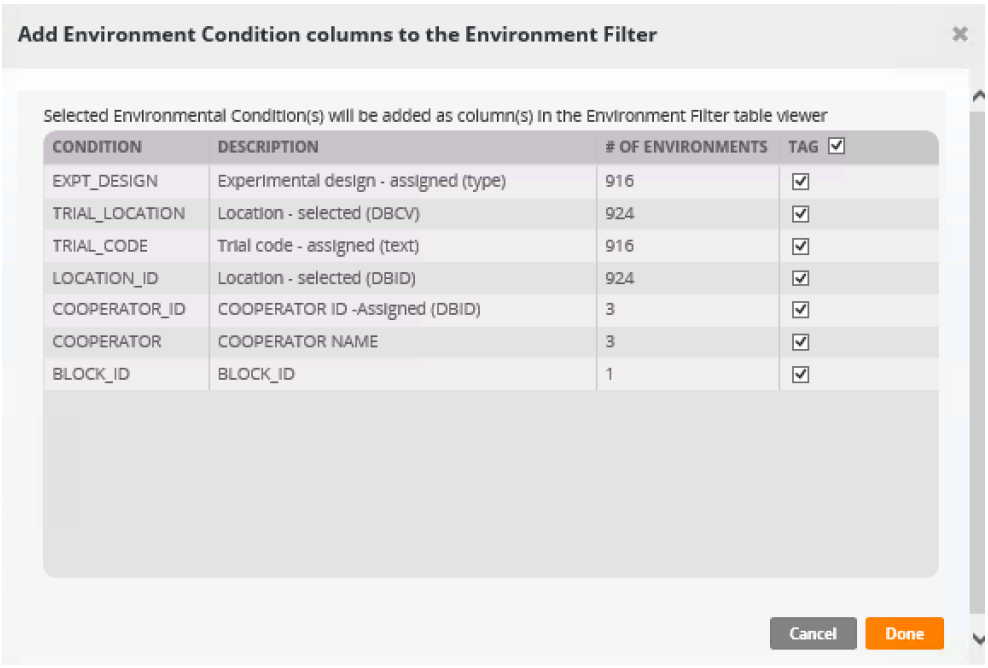


Select Filter by Location to narrow environments by location or country.

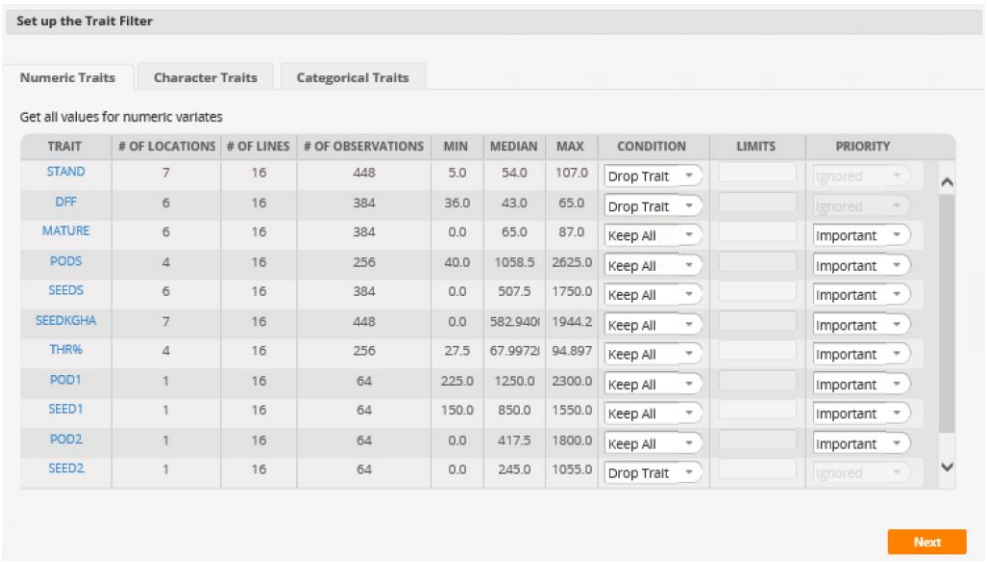
Select Filter by Study to narrow environments by study.



Add Environment Condition columns to filter table by selecting the button.

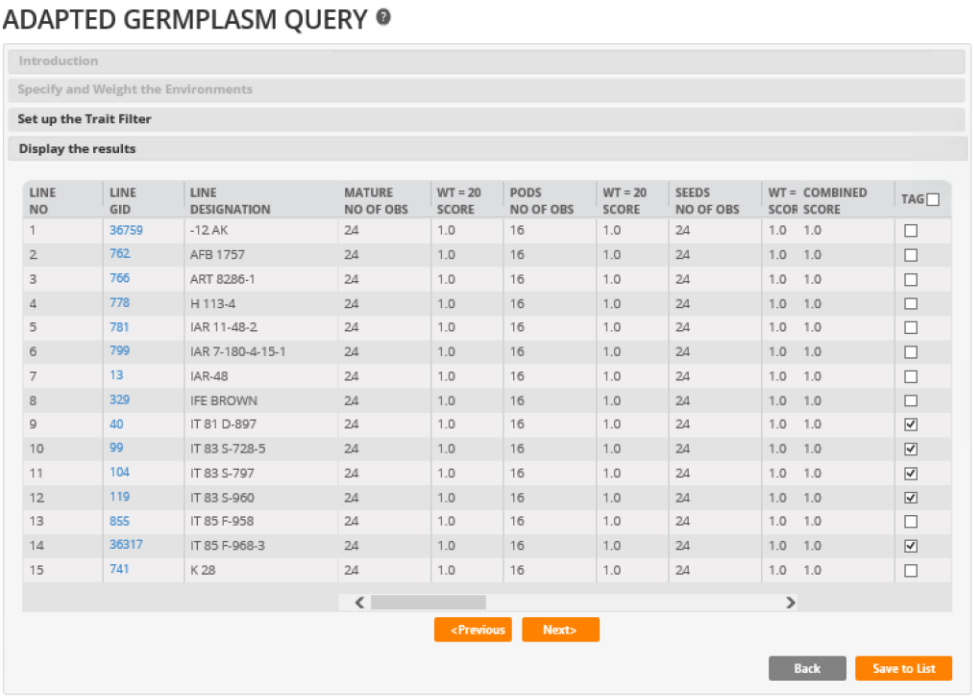


Select the trait filters.

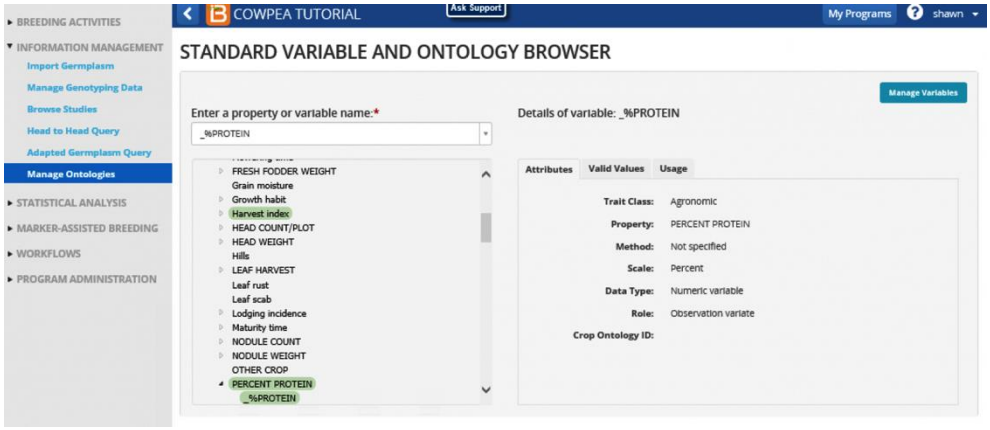


Display Results

Display the results and make selections. Click Save to List to save adapted germplasm selections in the database.

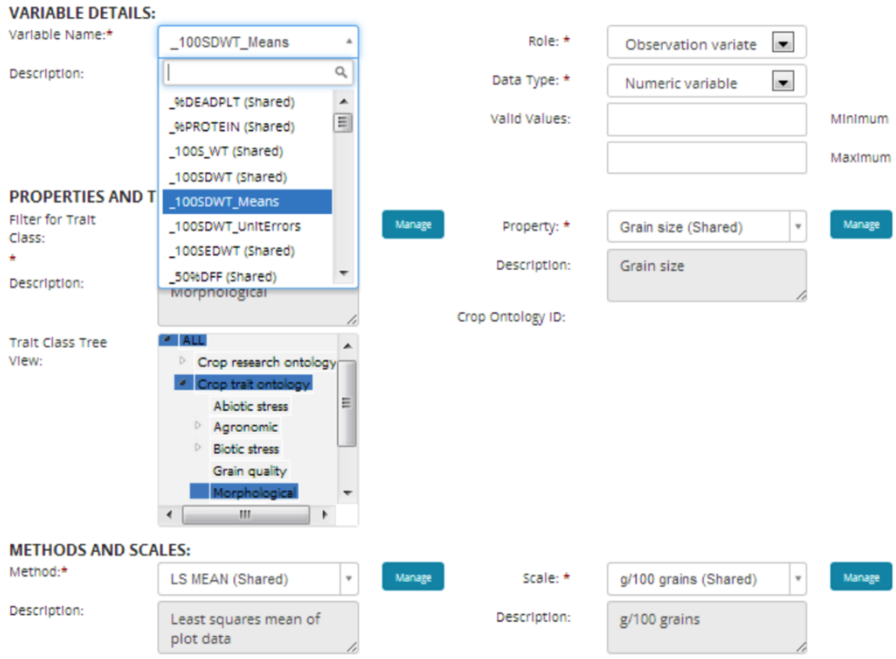


The Breeding Management System's crop ontology management tool uses a structured vocabulary to allow descriptors to act as variables in database queries and statistical analyses. Crop ontology terms for research and traits can be browsed from the Ontology Browser.

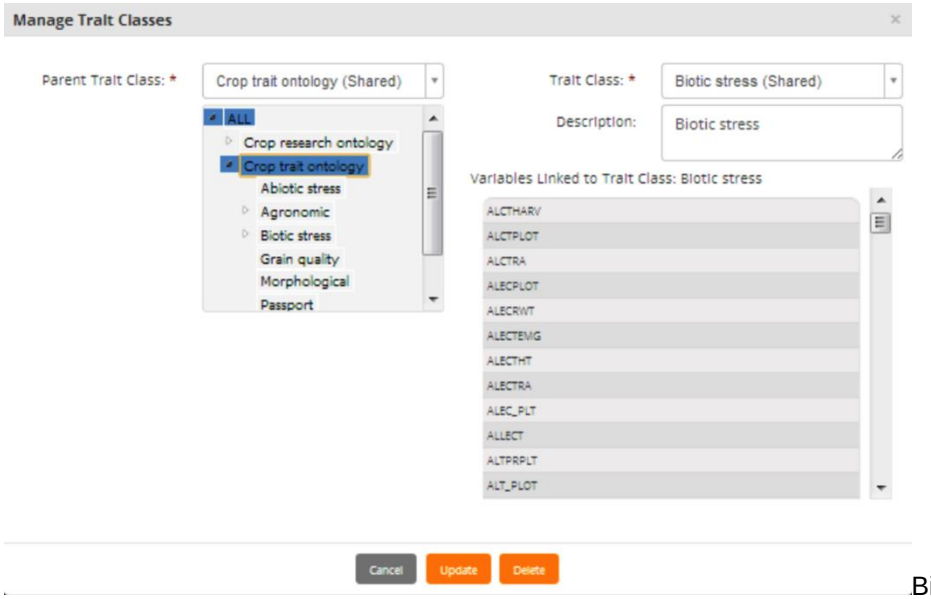


Ontology Term Details

Select Manage Variables to examine variables further. Chose variable from the drop down menu and the screen will populate with variable details.



Select the Manage button associated with Properties and Traits to associate traits to related groupings. Choose a trait class from the dropdown menu to obtain a list of associated variables.



Biotic Stress is a crop trait ontology term associated with the list of variables the lower right hand window.

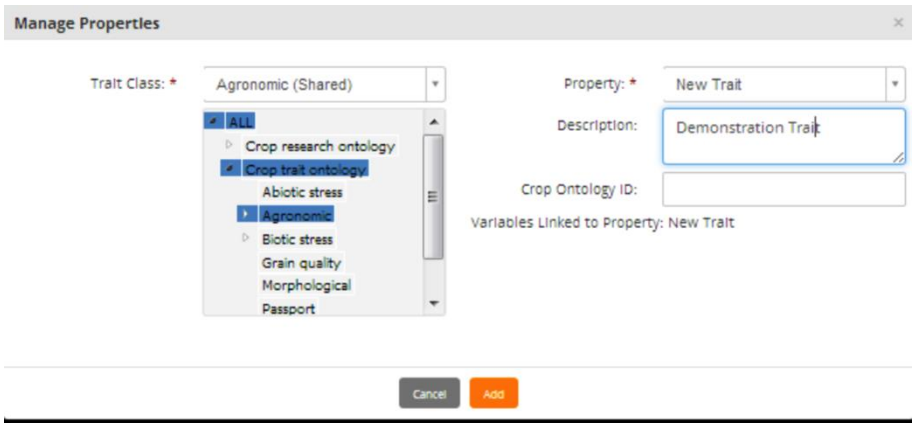
Manage Ontologies

Modify Existing Terms

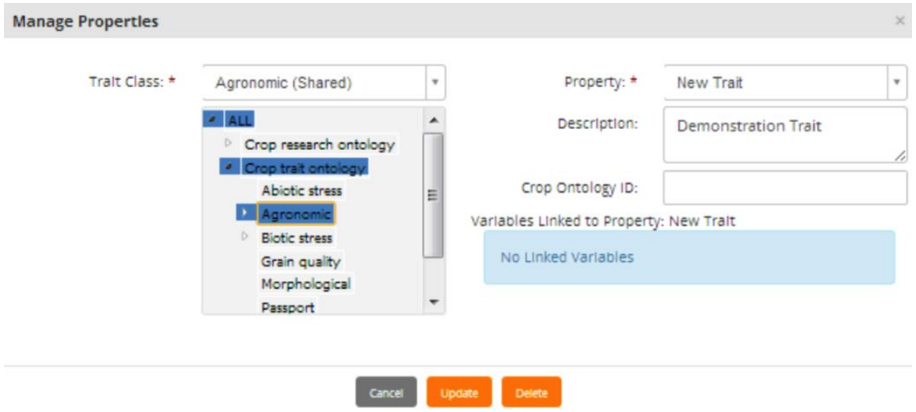
Variables with the (Shared) designation are associated with the public database and cannot be modified. Ontology terms from the program database, those without the shared designation, can be customized in numerous ways from the drop down menu. Select Update to save changes to the program database.

Add New Term

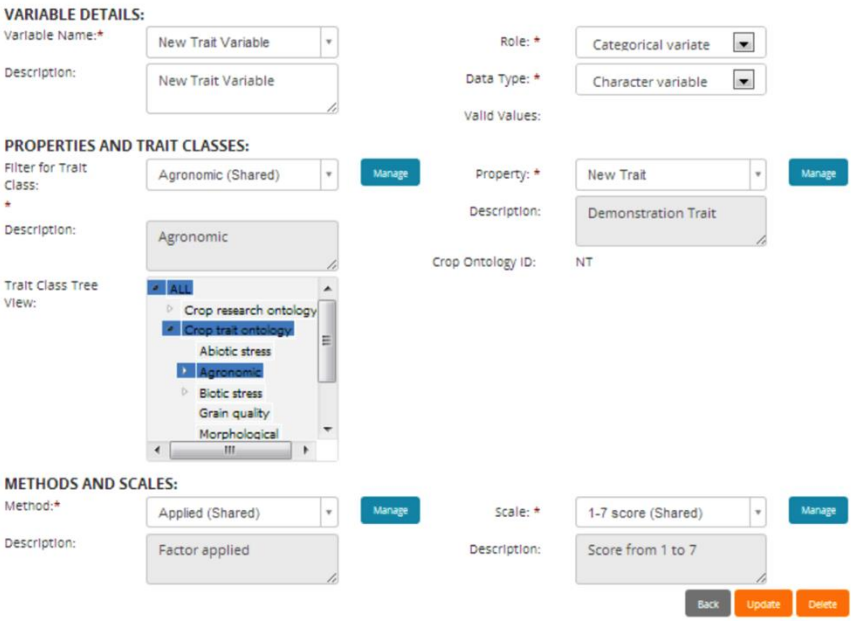
With Variable Details blank, select Manage Properties. Type a new property name and description. Select the appropriate trait class from the dropdown menu and click Add.



A message will notify that there are no linked variables. Select Update to save the new property and Cancel to close the window and return to the Manage Variables screen.



Set the details for Properties and Traits as well as Methods and Scales, updating each.



Once all of the details are set for the new ontology term, type a new variable name and description. Select Update to add the new variable to the program database.

Be aware that the combination of variable details must be unique. No two variables can have the same parameters.

Statistical Analyses

Single Site Analysis: Cowpea Tutorial

Contributors

Shawn Yarnes [a], Darren Murray [b], Roger Payne [b] & Zhengzheng Zhang [b]
[a] *The Integrated Breeding Platform,* [b] *VSN International Ltd*

Summary

The tutorial describes the single site analysis of three locations of data from a multisite cowpea field trial using a completely randomized block design with 3 replicates per location.

- Restore from previous tutorial
- Introduction
- Select Dataset to Analyze
- Run Analysis
- Analysis Results
- References

Restore from Previous Tutorial

Screenshots and activities in this tutorial build upon work preformed in previous tutorials. If you are not following the cowpea tutorials in sequence, restore the Cowpea Tutorial database (.sql) to the end of the previous tutorial, Design and Management of Field Trials, to match database contents with current tutorial.

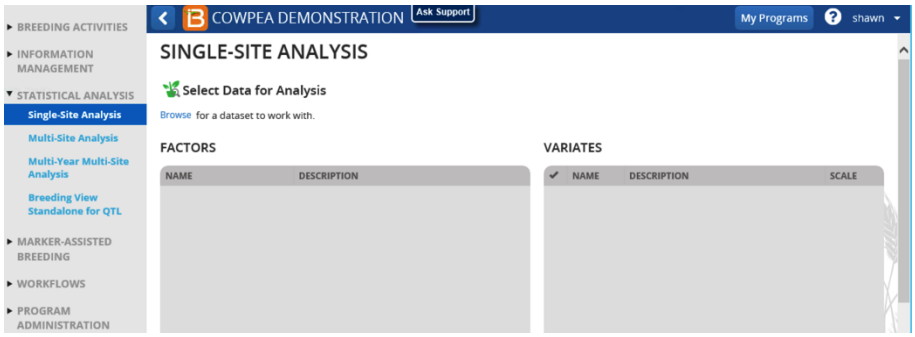
BMS File Directory
C:\Breeding Management System\Documents\BMS Workbench_Nursery & Trial Managers\Tutorials\Cowpea\7.0Cowpea Tutorial.sql

Introduction

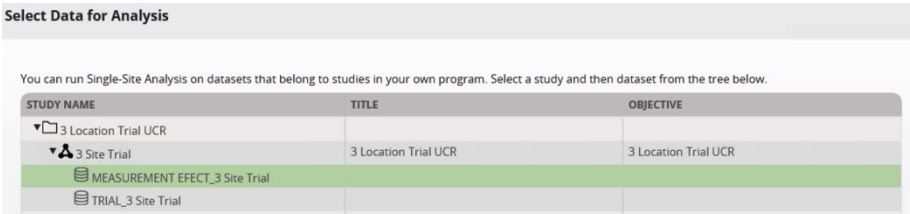
Breeding Views single site analysis uses mixed models to account for extraneous sources of variation in breeding trials, including replicates and incomplete blocks. The single site analysis produces adjusted means, best linear unbiased estimators and best linear unbiased predictors (BLUEs and BLUPs) per genotype. These adjusted means can be used within a genotype by environment (GxE) analysis or QTL (quantitative trait loci) analysis pipeline.

Select Dataset to Analyze

Open Single Site Analysis from the Statistical Analysis menu of the Workbench. Browse for to find the trial data.

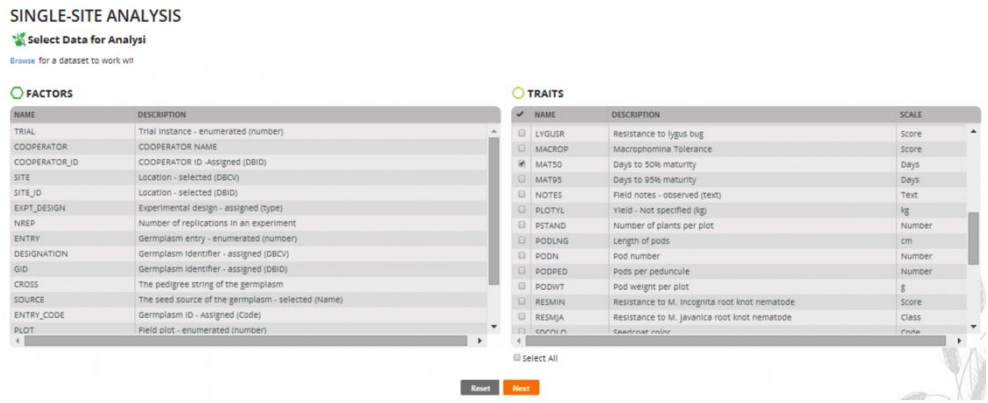


Navigate to the Measurement Effect file within the 3 Locations Trial UCR folder. Highlight the Measurement Effect dataset. There is no Measurement Effect, only ENVIRONMENT and PLOTDATA.



Specify Variables

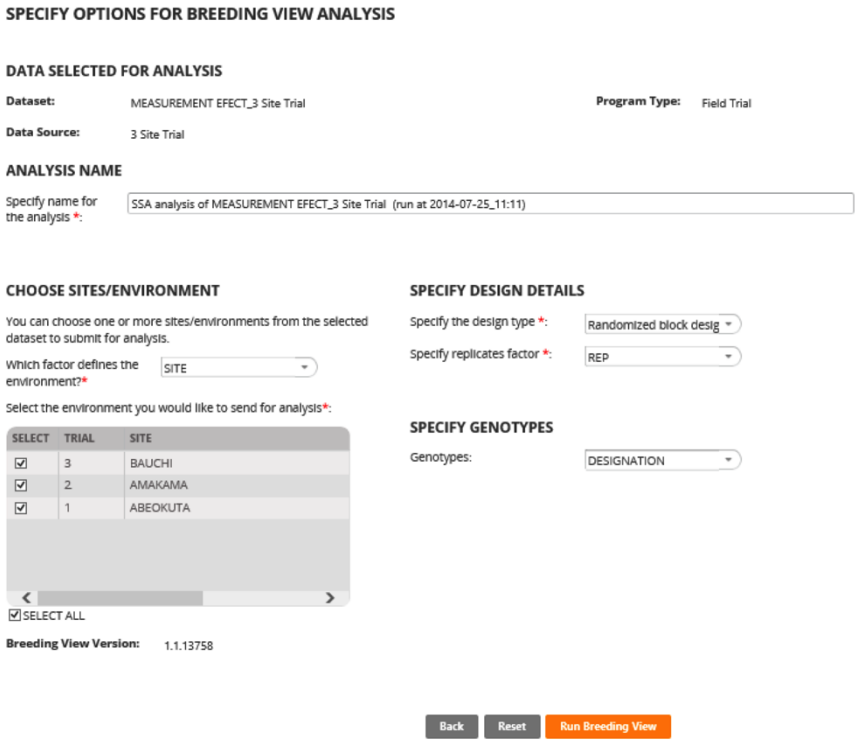
Review the factors and traits. You will recall that only two traits, MAT50 and SDW100, were measured for this trial, so deselect all traits represented by empty data columns. With only these two traits checked, select next.



Specify Analysis Conditions

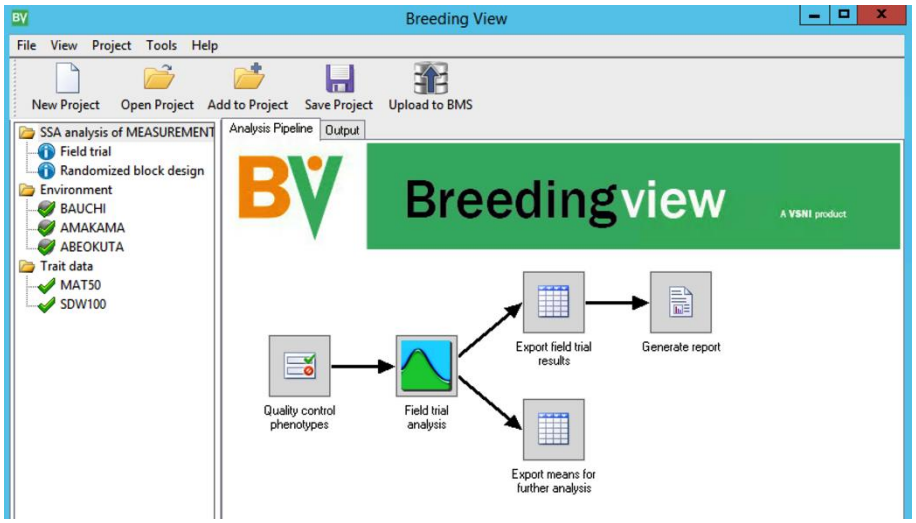
- Use the default analysis name.
- Select Randomized Block Design for the design type.
- The factor that defines environment is SITE.
- Select all three locations to preform individual single site analysis on each location.
- REP is the factor in this dataset that defines replications. DESIGNATION defines the germplasm factor to be used in the analysis.

Click Run Breeding View to launch the breeding view application.



Run Analysis

Single site analysis is database integrated. When the Breeding View application launches the analysis conditions and data are loaded. Notice that the three locations and both traits are selected. Right click on any trait or location to deselect from the analysis. When a project has been created or opened, a visual representation of the analytical pipeline is displayed in the Analysis Pipeline tab. The analysis pipeline includes a set of connected nodes, which can be used to run and configure pipelines.



Description of Nodes

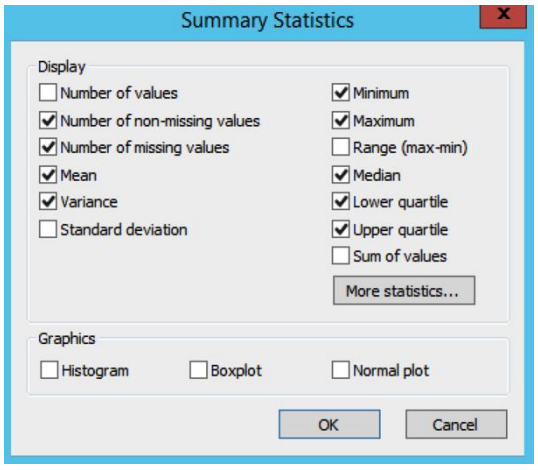
- Quality control phenotypes: Summary statistics for the trait(s)
- Field trial analysis: Mixed model analysis of field trial for the trait(s)
- Export field trial results:Stores results in external files
- Generate report:HTML report of results including means and summaries for the trait(s)
- Export means for further analysis: Stores adjusted means in an external file using a format ready to use in a GxE or QTL analysis pipeline

Analysis Options

Some of the nodes have options to control they way an analysis is performed and output that is displayed. To access the options, right-click on a node and select the Settings item from the shortcut menu. The changes to the options are retained during the current session and are saved to the project file.

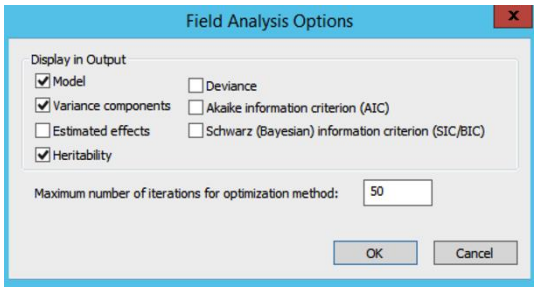
Summary Statistics

Default settings under Quality Control Phenotypes define the summary statistics that are displayed in the output.



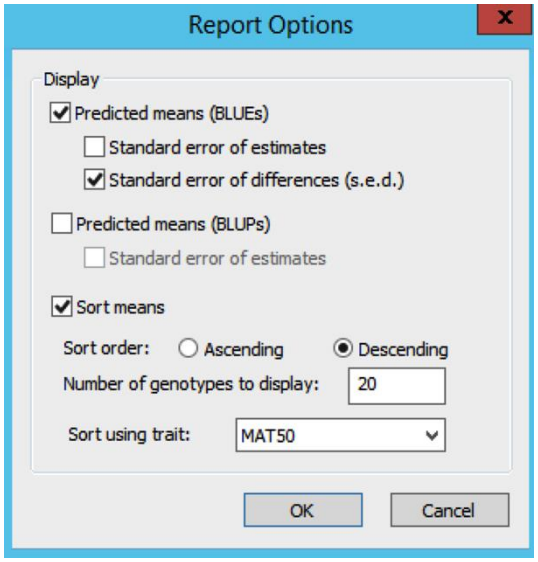
Field Analysis Options

Default settings under Field Trial Analysis define the output and iterations for optimization.



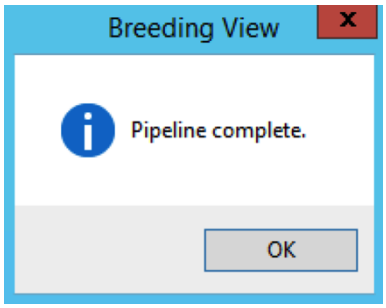
Report Options

Default settings under Generate Report define the reporting options.



Analysis Results

In this example leave the default settings and run the entire pipeline. Right click the first box and Run Selected Environment Pipelines. When the analysis is complete a popup notifies the user.

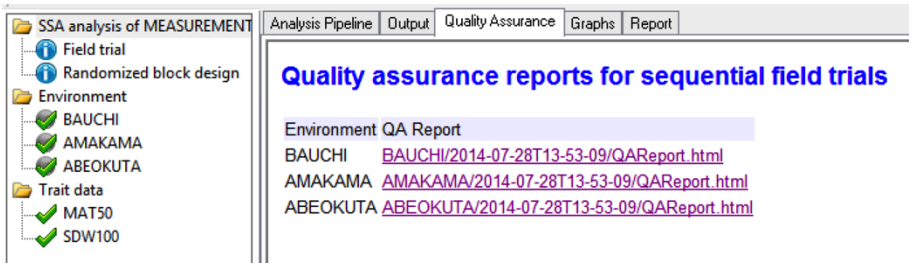


Quality Assurance

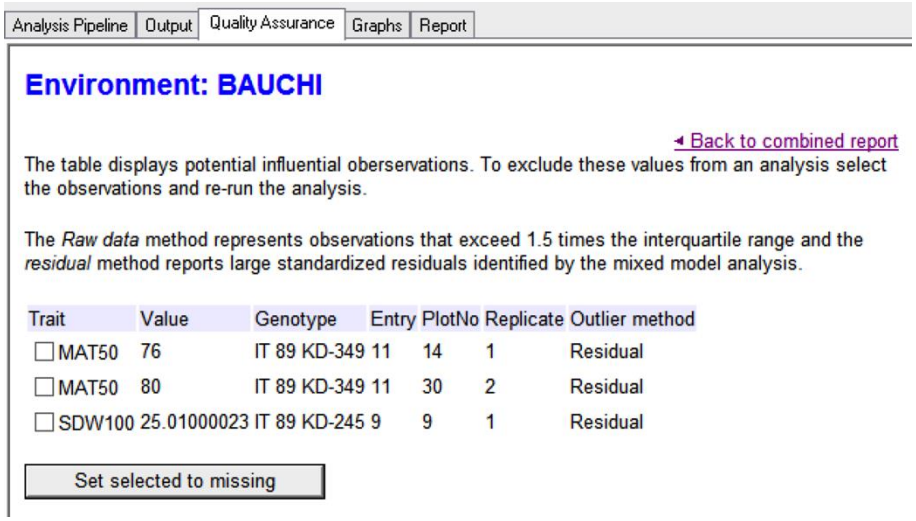
When the analysis pipeline is complete a Quality Assurance tab is generated. The Quality assurance tab displays reports of potentially influential (outlier) values for each location. Select the Bauchi environment to view details.

Breeding View identifies two types of outliers:

- Raw data outliers are observations that exceed 1.5 times the interquartile range, and can be seen on the accompanying boxplot.
- Residual outliers are observations that have been reported as a large standardized residual by mixed model analysis.



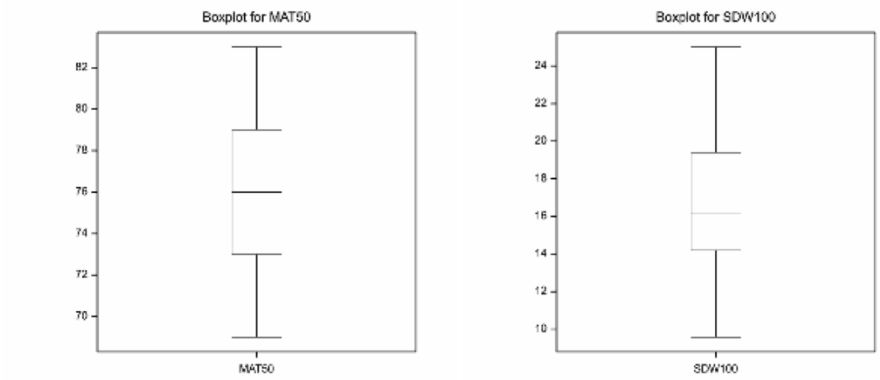
The Quality Assurance report can be used to change observations to become missing for the analysis. To do this, select the observations to set as missing from the list of traits, and click the Set Selected as Missing. The next time the analysis is run those observations will be excluded from the analysis. For this tutorial, do not exclude any values from the analysis.



Breeding View presents box plots of the variation found in the raw data. Diagnostic plots for the mixed model analysis can be viewed in the Graph or Report tabs

Boxplots of raw data

Boxplots of the raw data displaying individual obervations which are 1.5 times greater than the interquartile range.



Review the table of entries/genotypes that have potentially influential observations with their associated values, as well as other plots within the field design.

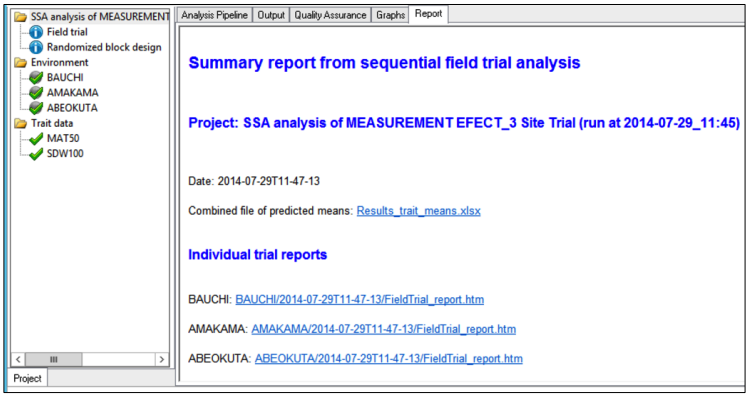
Influential observations and associated entry numbers

Trait	Value	Genotype	Entry	PlotNo	Replicate	Type
MAT50	76	IT 89 KD-349	11	14	1	Outlier
MAT50	80	IT 89 KD-349	11	30	2	Outlier
MAT50	72	IT 89 KD-349	11	36	3	
SDW100	25.01000023	IT 89 KD-245	9	9	1	Outlier
SDW100	19.85000038	IT 89 KD-245	9	24	2	
SDW100	24.60000038	IT 89 KD-245	9	32	3	

Report

The report tab includes:

- Combined file of predicted means: Excel file of BLUEs and BLUPs
- Links to individual trial reports
- Heritability Table: Broad-sense mean line heritability, derived from an estimate of the correlation between the genotype BLUPs and their unknown true value (Cullis, Smith & Coombes, 2006).

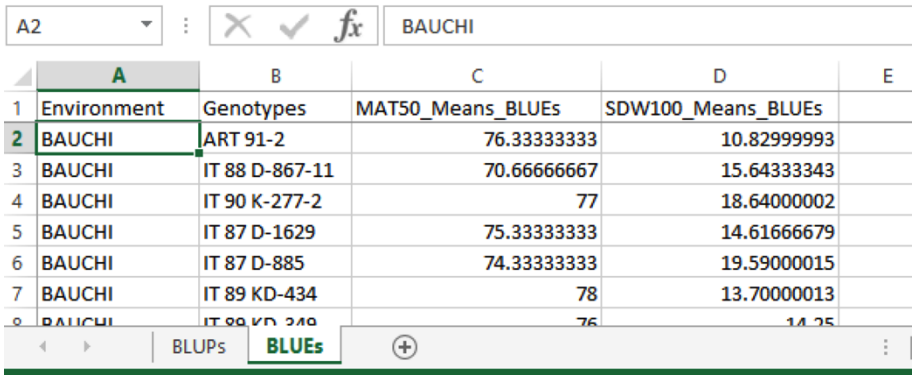


Heritability values

SITE	Trait	Heritability
BAUCHI	MAT50	0.8027
BAUCHI	SDW100	0.8836
AMAKAMA	MAT50	0.0167
AMAKAMA	SDW100	0.1197
ABEOKUTA	MAT50	0.0000
ABEOKUTA	SDW100	0.7788

Combined File of Predicted Means

Open the combined file of predicted means in Excel or other spreadsheet program. The file contains two worksheets with BLUPs and BLUEs.



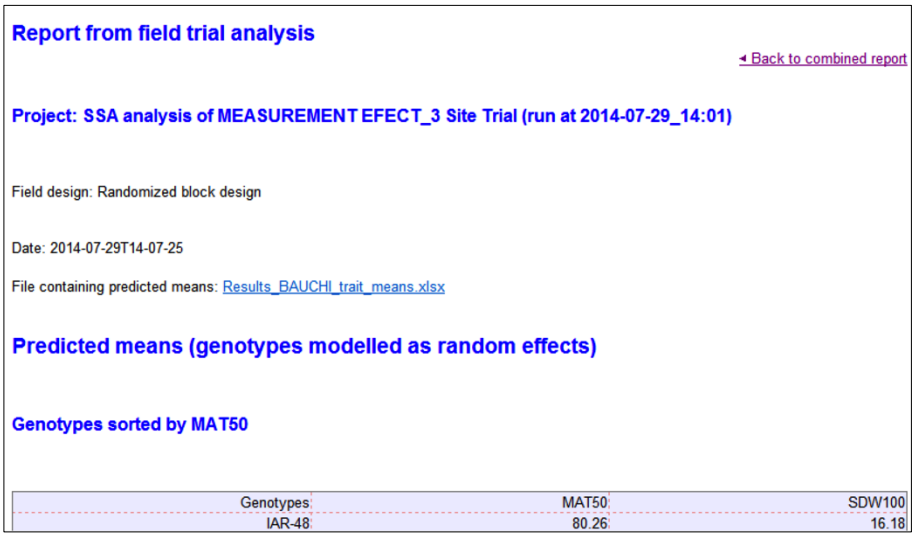
A2 | BAUCHI

	A	B	C	D	E
1	Environment	Genotypes	MAT50_Means_BLUEs	SDW100_Means_BLUEs	
2	BAUCHI	ART 91-2	76.33333333	10.82999993	
3	BAUCHI	IT 88 D-867-11	70.66666667	15.64333343	
4	BAUCHI	IT 90 K-277-2	77	18.64000002	
5	BAUCHI	IT 87 D-1629	75.33333333	14.61666679	
6	BAUCHI	IT 87 D-885	74.33333333	19.59000015	
7	BAUCHI	IT 89 KD-434	78	13.70000013	
8	BAUCHI	IT 89 KD-249	76	14.25	

BLUPs | BLUEs

Individual Trial Reports

View individual trial reports with summary statistics and diagnostic plots by selecting the location of interest, in this case Bauchi.



Report from field trial analysis

◂ Back to combined report

Project: SSA analysis of MEASUREMENT EFECT_3 Site Trial (run at 2014-07-29_14:01)

Field design: Randomized block design

Date: 2014-07-29T14-07-25

File containing predicted means: Results_BAUCHI_trait_means.xlsx

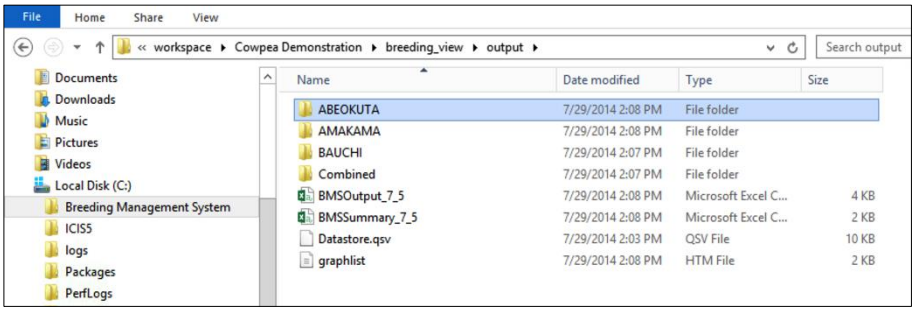
Predicted means (genotypes modelled as random effects)

Genotypes sorted by MAT50

Genotypes	MAT50	SDW100
IAR-48	80.26	16.18

Save Analysis Results

When an analysis pipeline is run, the associated files are automatically saved into the BMS workspace folder.



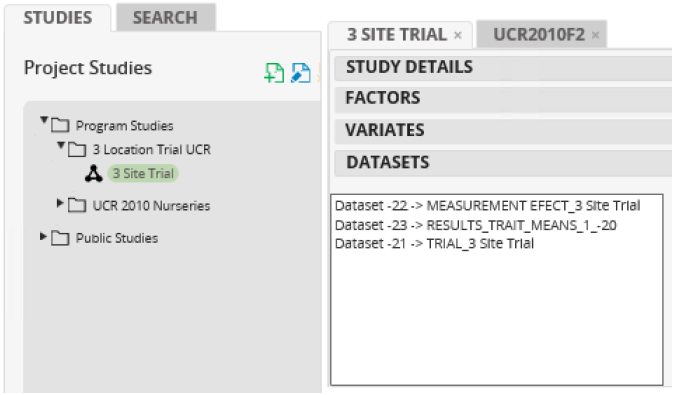
File | Home | Share | View

« workspace ▸ Cowpea Demonstration ▸ breeding_view ▸ output ▸ | Search output

Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Breeding Management System
ICIS5
logs
Packages
PerfLogs

Name	Date modified	Type	Size
ABEOKUTA	7/29/2014 2:08 PM	File folder	
AMAKAMA	7/29/2014 2:08 PM	File folder	
BAUCHI	7/29/2014 2:07 PM	File folder	
Combined	7/29/2014 2:07 PM	File folder	
BMSOutput_7_5	7/29/2014 2:08 PM	Microsoft Excel C...	4 KB
BMSSummary_7_5	7/29/2014 2:08 PM	Microsoft Excel C...	2 KB
Datastore.qsv	7/29/2014 2:03 PM	QSV File	10 KB
graphlist	7/29/2014 2:08 PM	HTM File	2 KB

Upload the results of the analysis to the Workbench database by selecting Upload to the BMS. Once saved to the database, select Browse Studies from the Information Management tools on the Workbench menu to view study data. Notice that the trait means have now been added to the database.



References

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. Journal of Agricultural, Biological, and Environmental Statistics, 11(4), 381-393.

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity, 95(3), 221-227.

Murray, D. Payne,, R,, & Zhang, Z. (2014) Breeding View, a Visual Tool for Running Analytical Pipelines: User Guide. VSN International Ltd. (.pdf) (associated sample data .zip)

Acknowledgements

The statistical algorithms in the Breeding View were developed by VSN International Ltd in collaboration with the Biometris group at University of Wageningen. Cowpea demonstration data was provided by Jeff Ehlers, Tim Close, Philip Roberts, Bao Lam Huyuh at the University of California Riverside and Issa Drabo at the Institut de l'Environnement et de Recherches Agricoles in Burkina Faso. These data may have been adapted for training purposes. Any misrepresentation of the raw breeding data is the solely the responsibility of the IBP.

Multi-Site (G x E) Analysis: Cowpea Tutorial

Contributors

Shawn Yarnes [a], Darren Murray [b], Roger Payne [b] & Zhengzheng Zhang [b]
[a] *The Integrated Breeding Platform,* [b] *VSN International Ltd*

Summary

This tutorial provides instruction on the analysis of genotype by environment (GxE) interactions within a three-location cowpea field trial. This tutorial builds upon the adjusted means calculated for the individual locations in the previous tutorial, Single Site Analysis: 3 Location Batch.

- Restore from Previous Tutorial
- Introduction
- Select Data from Database
- Run Analysis
- Results
- References

Restore from Previous Tutorial

Screenshots and activities in this tutorial build upon work preformed in previous tutorials. If you are not following the cowpea tutorials in sequence, restore the Cowpea Tutorial database (.sql) to the end of the previous tutorial, Single Site Analsysis, to match database contents with current tutorial.

BMS File Directory
C:\Breeding Management System\Documents\BMS Workbench_Nursery & Trial Managers\Tutorials\Cowpea\8.0Cowpea Tutorial.sql

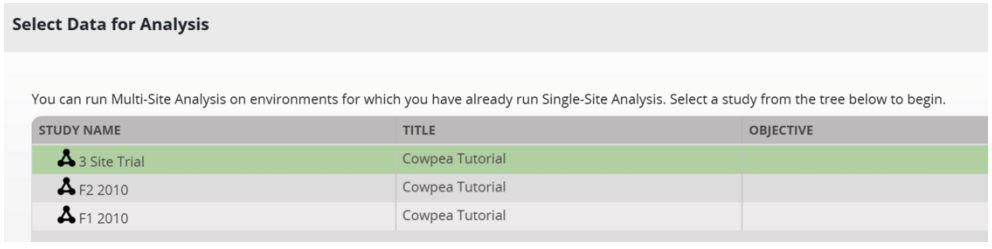
Introduction

After single site analysis has been preformed for each location and the adjusted means are saved to the database, a multi-site analysis is possible to further investigate genotypic and environmental (GxE) interactions.

Select Data from Database

Open Multi-Site Analysis from the Statistical Analysis menu of the Workbench. Select Browse.

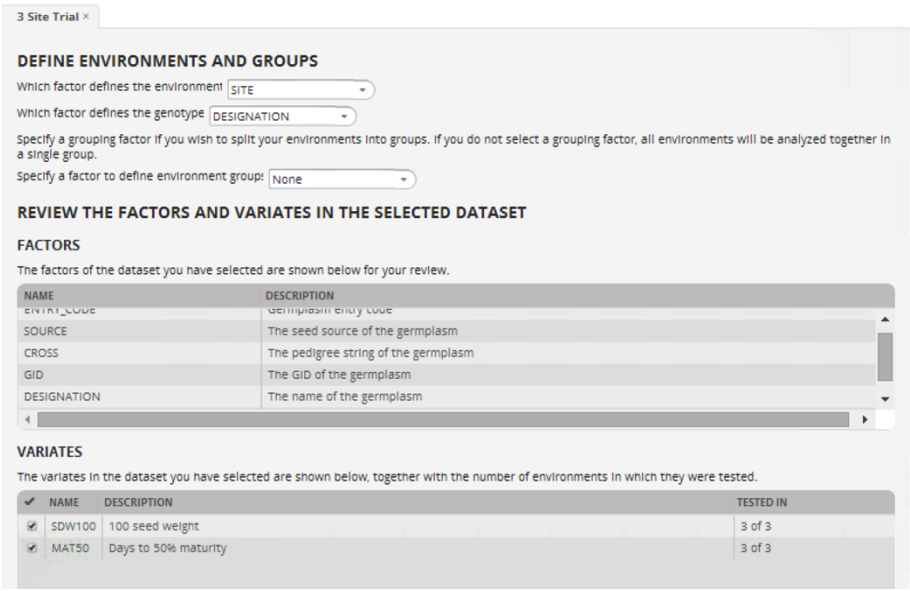
Highlight and Select 3 Site Trial, which houses the previously calculated adjusted means.



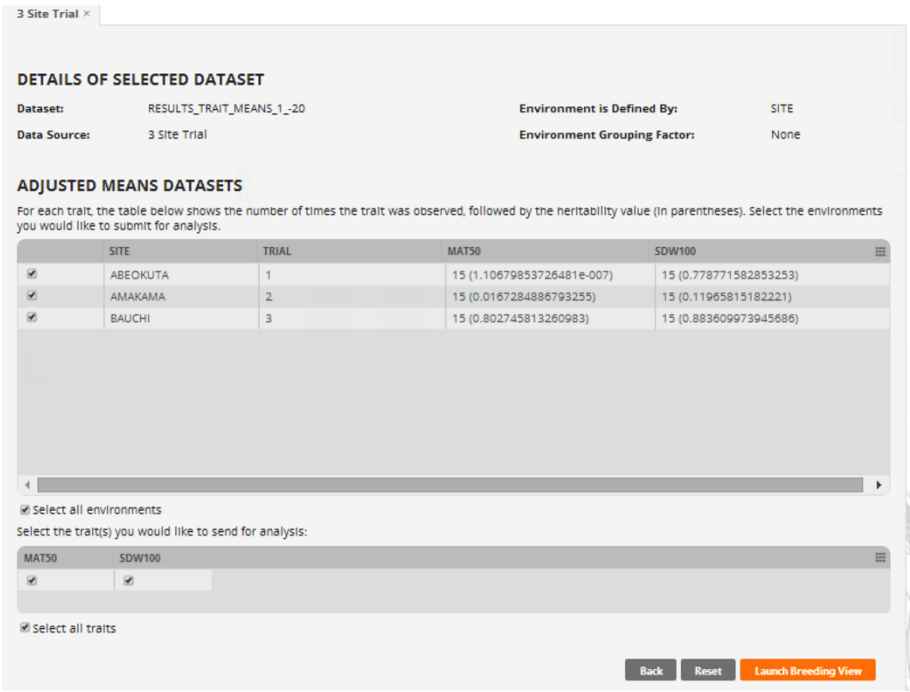
Define the environments and groups and review the factors and variates.

- Factors
- Environment: SITE
- Genotype: DESIGNATION
- Environment Group (Mega-Environments): NONE

Select both variates; seed weight and 50% maturity.

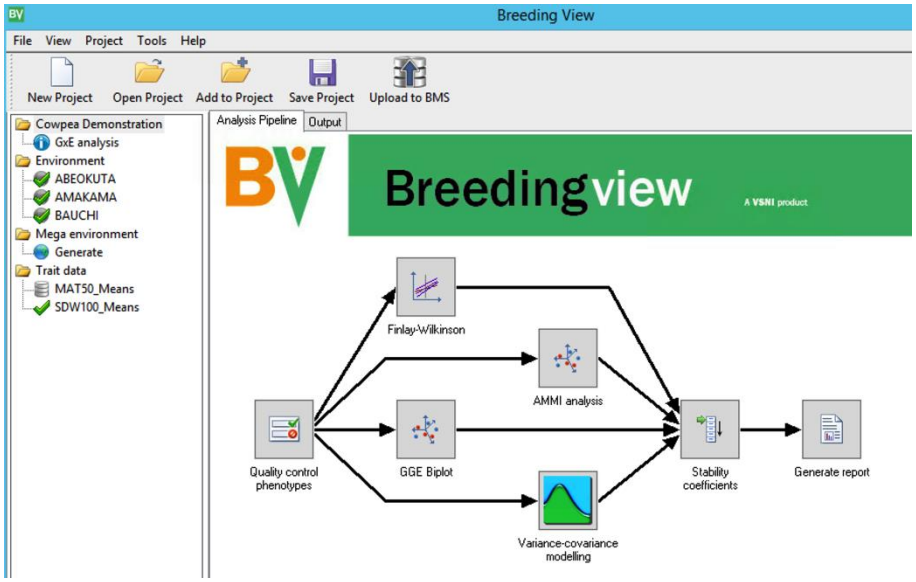


Review the details of the dataset and launch Breeding View. Click "Next" By the way, this process takes way too long (both generating the table and launching Breeding View).



Run Analysis

Mulit-Site analysis is database integrated. When the Breeding View application launches, the analysis conditions and data are automatically loaded. Notice that all locations and traits are selected by default. Right click on any trait or location to deselect from the analysis. When a project has been created or opened, a visual representation of the analytical pipeline is displayed in the Analysis Pipeline tab. The analysis pipeline includes a set of connected nodes, which can be used to run and configure pipelines. Right click on MAT50_Means and exclude this trait from the analysis.



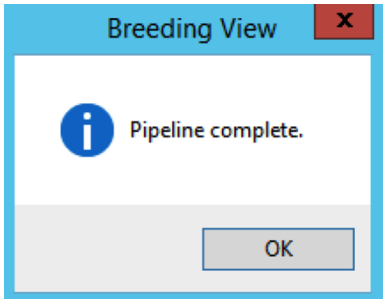
Node Descriptions:

- Quality Control Phenotypes Summary statistics within and between environments for the trait(s)
- Finlay-Wilkinson: Performs a Finlay-Wilkinson joint regression
- AMMI Analysis:
- GGE Biplot
- Variance-Covariance Modeling: Selects the best covariance structure for genetic correlations between environments
- Stability Coefficients: Estimates different stability coefficients to assess genotype performance and generate HTML report of the results

Analysis Options

Some of the nodes have options to control they way an analysis is performed and output that is displayed. To access the options, right-click on a node and select the Settings item from the shortcut menu. Changes to the options are retained during the current session and are saved to the project file.

Leave the default analysis settings. Right click on Quality Control Phenotypes to run pipeline. When the analysis is complete a popup window will notify you that the results are ready to view.



Results

Analysis results can be found in the Output, Graphs, and Results tab. Results are also saved in the Breeding Management Systems workspace folder.

Analysis Results

- Summary statistics for seed weight within and between the environments
- Sensitivity Calculations: Finlay-Wilkinson Joint Regression, AMMI, GGE,
- Stability statistics: Different variance-covariance models.

Summary Statistics

Analysis Pipeline | Output | Graphs | Report

Trait: SDW100_Means

Summary statistics for SDW100_Means

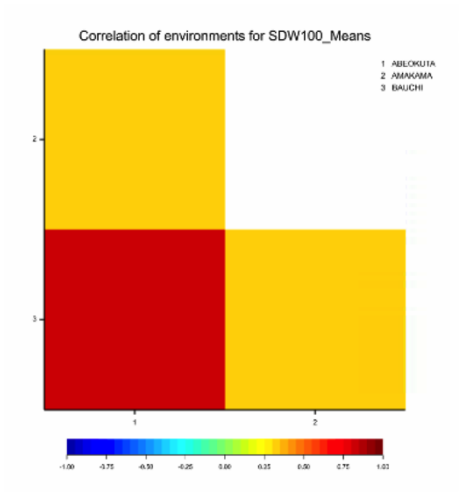
	No. of observations	No. of missing values	Mean	Median
AMAKAMA	15.00	0	20.33	20.00
ABEOKUTA	15.00	0	19.41	19.97
BAUCHI	15.00	0	16.86	16.14

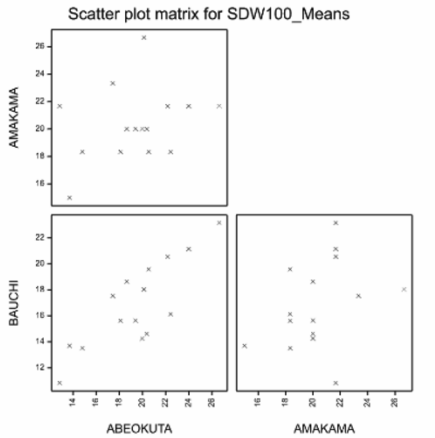
	Min	Max	Lower quartile	Upper quartile	Variance
AMAKAMA	15.00	26.67	18.33	21.67	7.22
ABEOKUTA	12.83	26.63	17.60	21.76	14.02
BAUCHI	10.83	23.15	14.34	19.35	11.22

Correlations between environments

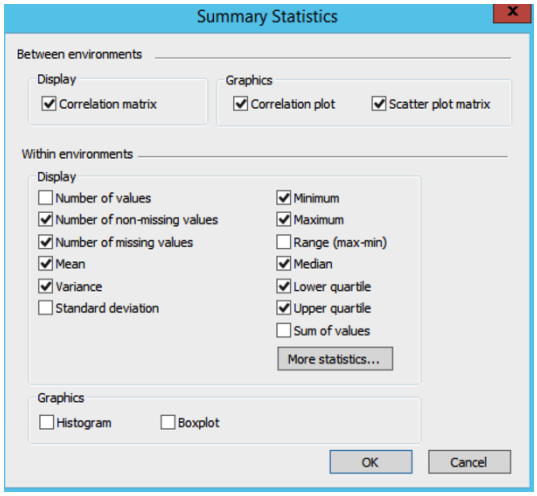
ABEOKUTA	-		
AMAKAMA	0.3057	-	
BAUCHI	0.8289	0.3483	-
	ABEOKUTA	AMAKAMA	BAUCHI

Summary statistics from the Output tab showing that the Amakama site has the highest mean seed weight and that the Abeokuta and Bauchi sites are highly correlated.





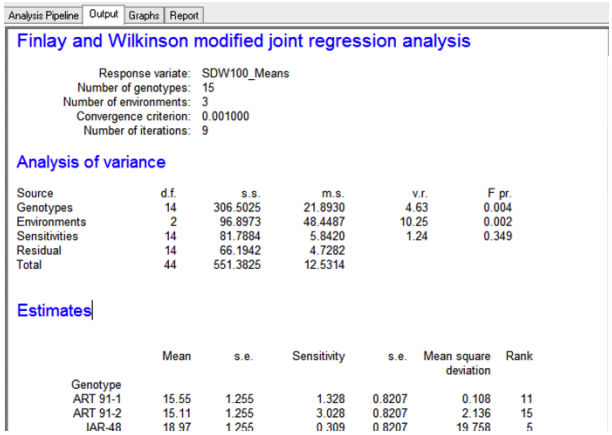
Graphical display of environmental correlations: The red coloration on the heat map indicates a high correlation (0.8289 from table in Output) between Abeokuta and Bauchi, and small correlations for the other two sites.



Default summary statistics options

Finlay-Wilkinson Joint Regression

Finlay-Wilkinson Joint Regression characterizes the sensitivity of each genotype to environmental effects by fitting a regression of the environment means for each genotypes on the average environmental means. In the estimates a value of 1 represents the average sensitivity and genotypes with a value greater than one exhibit higher than average sensitivity and genotypes with a value less than 1 are less sensitive than average.


Analysis Pipeline | Output | Graphs | Report

Finlay and Wilkinson modified joint regression analysis

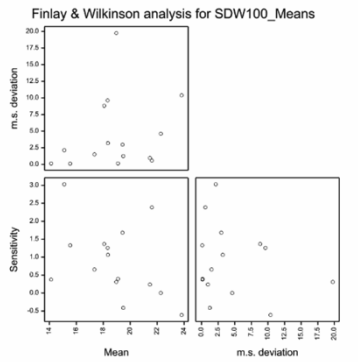
Response variate: SDW100_Means
Number of genotypes: 15
Number of environments: 3
Convergence criterion: 0.001000
Number of iterations: 9

Analysis of variance

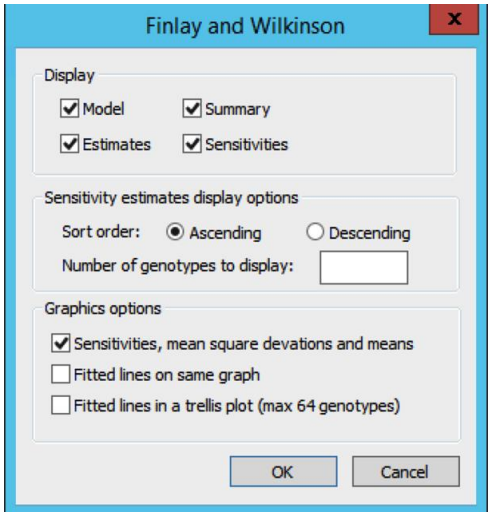
Source	d.f.	s.s.	m.s.	v.r.	F pr.
Genotypes	14	306.5025	21.8930	4.63	0.004
Environments	2	96.8973	48.4487	10.25	0.002
Sensitivities	14	81.7884	5.8420	1.24	0.349
Residual	14	66.1942	4.7282		
Total	44	551.3825	12.5314		

Estimates

Genotype	Mean	s.e.	Sensitivity	s.e.	Mean square deviation	Rank
ART 91-1	15.55	1.255	1.328	0.8207	0.108	11
ART 91-2	15.11	1.255	3.028	0.8207	2.136	15
IAR-48	18.97	1.255	0.309	0.8207	19.758	5



The Finlay Wilkinson analysis indicates that genotypes and environments are significant sources of variation (p = 0.004 and 0.002 respectively), but that G x E sensitivity is not significant (p=0.349). The most sensitive genotype is ART 91-2.



Default Finlay & Wilkinson options

AMMI Analysis

The AMMI analysis node fits a model which involves the Additive Main effects of ANOVA with the Multiplicative Interaction effects of principal components analysis (PCA). The AMMI model is more flexible than Finlay-Wilkinson, in that more than one environmental quality variable is explained using multiplicative terms.

Analysis Pipeline | Output | Graphs | Report

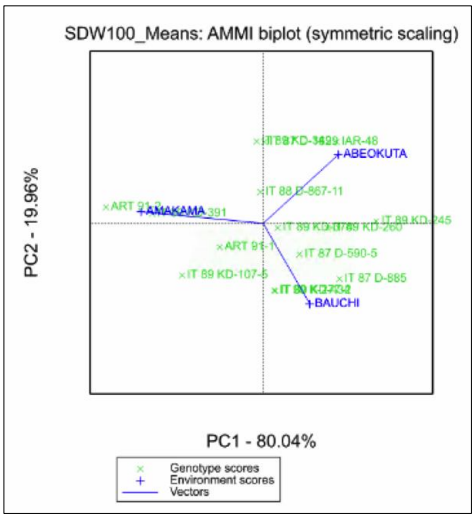
AMMI Analysis

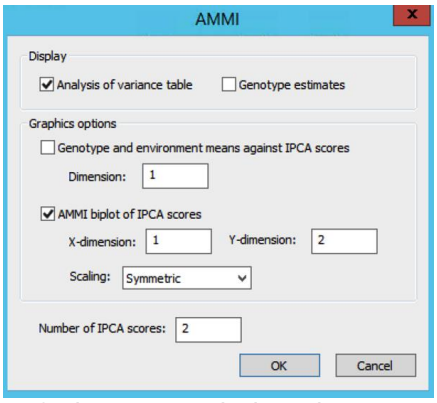
ANOVA table for AMMI model

Source	d.f.	s.s.	m.s.	v.r.	F pr
Genotypes	14	306.5	21.89		
Environments	2	96.9	48.45		
Interactions	28	148.0	5.29		
IPCA 1	15	118.4	7.90		
IPCA 2	13	29.5	2.27		
Residuals	0	0.0			

The ANOVA table for the AMMI analysis cannot be completed with variance and P values, because degrees freedom is limited by too few environments.

A desirable property of the AMMI model is that genotype and environmental scores can be used to construct biplots to help interpret genotype-by-environment interaction. In the biplot, genotypes that are similar to each other are closer in the plot than genotypes that are different. Similarly, environments that are similar will group together as well. When environment scores are connected to the origin of the plot, an acute angle between lines indicate a positive correlation between environments. A right angle between lines indicates low or no correlation between environments, and an obtuse angle indicates negative correlation. The projection of a genotype onto the environmental axis reflects performance in that particular environment.



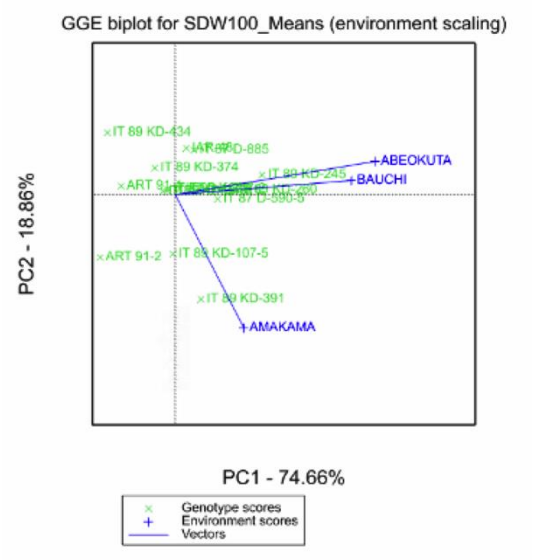


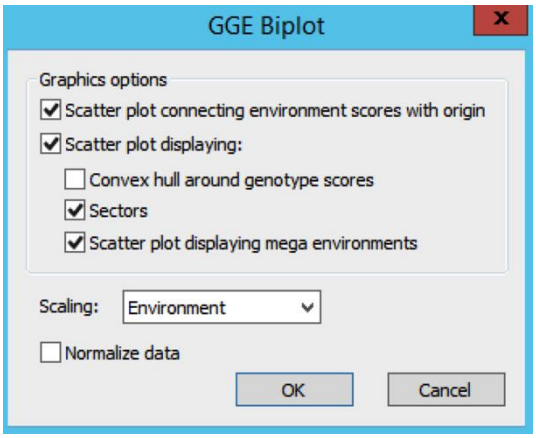
Default AMMI Analysis options

The obtuse angle between Amakama and both Bauchi and Abeokuta, indicate that Amakama is negatively correlated to the other two environments. The acute angle between Abeokuta and Bauchi reiterates the summary statistics result, that phenotypic responses at these two locations are correlated.

GGE Biplot

The GGE (genotype main effect (G) plus genotype-by-environment (GE) interaction) biplot is a modification of the AMMI model, which joins the effects of the genotypic main effects and the genotype-by-environment interaction (Yan & Kang, 2003). As this describes both the genotypic main effects and genotype-by-environment interaction together, it is known as a GGE model, and the biplots are called GGE biplots. The interpretation of the GGE biplot is similar to the AMMI biplot, but now the genotypes are distributed according to overall performance in each environment, rather than just genotype-by-environment interaction. In the GGE biplot, the best performing genotypes are on the right-hand side of the plot.

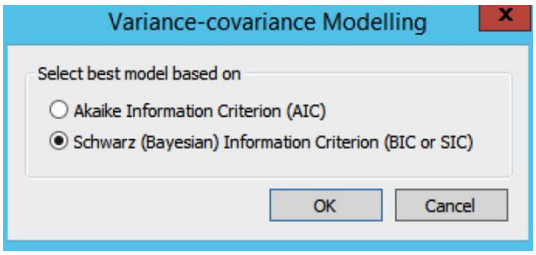




Default GGE Biplot options

Variance-Covariance Modeling

Variance-covariance modeling is a mixed model approach to examination of genotype-by-environment interactions in terms of heterogeneity of variances and covariances. Breeding View evaluates a range of possible variance-covariance models and selects the best based on an information criterion.

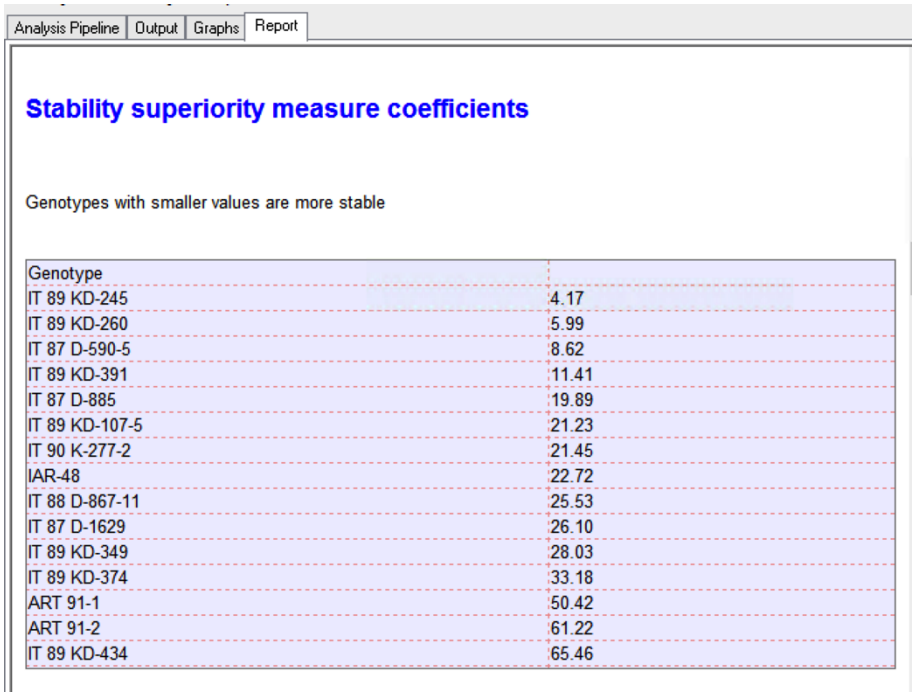


Stability coefficient measures for the genotypes:

Cultivar-Superiority Measure (Lin & Binns, 1988): The sum of the squares of the difference between genotypic mean in each environment and the mean of the best genotype, divided by twice the number of environments. Genotypes with the smallest values of the superiority tend to be more stable, and closer to the best genotype in each environment.

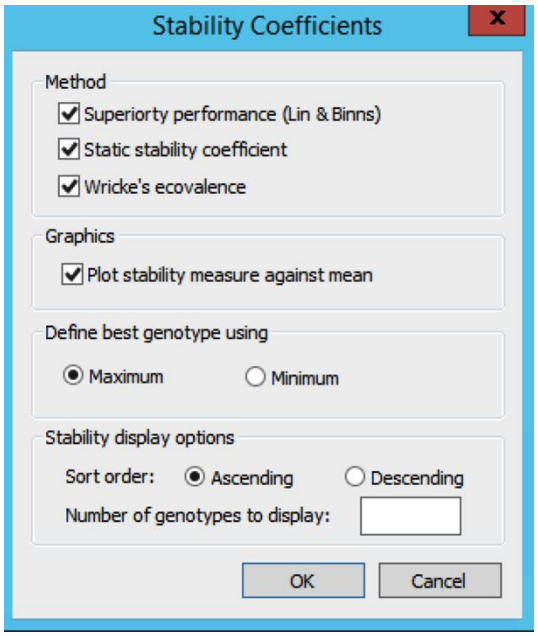
Static Stability Coefficient is defined as the variance between its mean in the various environments. This provides a measure of the consistency of the genotype, without accounting performance

Wrickes Ecovalence Stability Coefficient (Wricke, 1962): The contribution of each genotype to the genotype-by-environment sum of squares, in an un-weighted analysis of the genotype-by-environment means. A low value indicates that the genotype responds in a consistent manner to changes in environment.



Genotype	
IT 89 KD-245	4.17
IT 89 KD-260	5.99
IT 87 D-590-5	8.62
IT 89 KD-391	11.41
IT 87 D-885	19.89
IT 89 KD-107-5	21.23
IT 90 K-277-2	21.45
IAR-48	22.72
IT 88 D-867-11	25.53
IT 87 D-1629	26.10
IT 89 KD-349	28.03
IT 89 KD-374	33.18
ART 91-1	50.42
ART 91-2	61.22
IT 89 KD-434	65.46

The most stable genotype identified by the Cultivar-Superiority Measure is IT 89-KD-245.



When Breeding View is run, the data files are automatically saved to your computer in the Workspace folder.

References

Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. Journal of Agricultural Biological and Environmental Statistics 11(4), 381-393.

Gauch, H.G. (1992). Statistical Analysis of Regional Yield Trials AMMI analysis of factorial designs. Elsevier, Amsterdam.

Finlay, K.W. & Wilkinson, G.N. (1963). The analysis of adaptation in a plant-breeding programme. Australian Journal of Agricultural Research, 14, 742-754.

Murray, D. Payne,, R,, & Zhang, Z. (2014) Breeding View, a Visual Tool for Running Analytical Pipelines: User Guide. VSN International Ltd. (.pdf) (associated sample data .zip)

Lin, C.S. & Binns. M.R. (1988). A superiority performance measure of cultivar performance for cultivar x location data. Canadian Journal of Plant Science, 68, 193-198.

Wricke, G. (1962). Uber eine method zur erfassung der okogischen streubreite in feldversuchen. Zeitschriff Fur Pflanzenzuchtung, 47, 92-96.

Yan, W. & Kang, M.S. (2003). GGE Biplot Analysis: a Graphical Tool for Breeders, Geneticists and Agronomists. CRC Press, Boca Raton.

Yan, W., Kang, M. S., Ma, B., Woods, S., & Cornelius, P. L. (2007) GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Science, 47(2), 643-653.

Acknowledgements

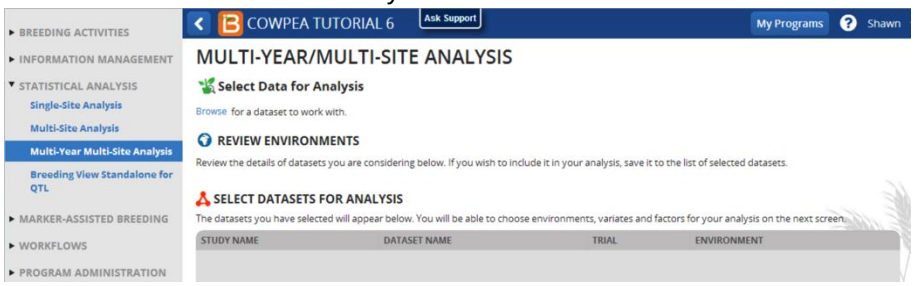
The statistical algorithms in the Breeding View were developed by VSN International Ltd in collaboration with the Biometris group at University of Wageningen. Cowpea demonstration data was provided by Jeff Ehlers, Tim Close, Philip Roberts, Bao Lam Huyuh at the University of California Riverside and Issa Drabo at the Institut de l'Environnement et de Recherches Agricoles in Burkina Faso. These data may have been adapted for training purposes. Any misrepresentation of the raw breeding data is the solely the responsibility of the IBP.

Multi-Site Multi-Year Analysis

Introduction

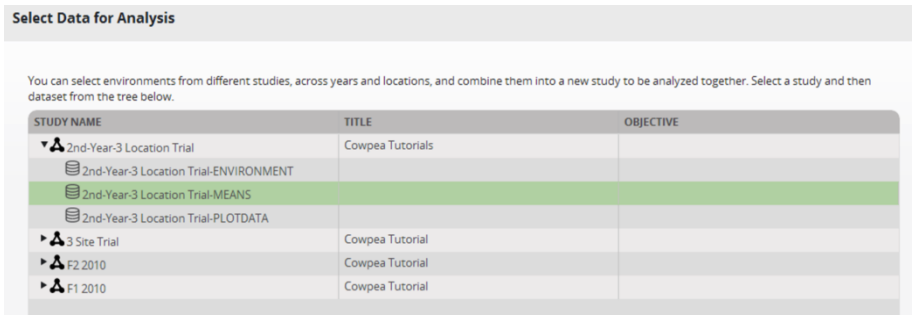
The Breeding Management System (BMS) does not yet have the capability to preform multi-site multi-year statistical analysis. Expect this functionality in future versions. However, the BMS does support the export of multi-location trial data, including multiple years, as a merged single data file. The observation sheet of this merged data file contains all of the factors needed to support analysis of trial data by external statistical packages, such as R.

Select Multi-Year Multi-Site Analysis from the workbench menu. Browse for a dataset of interest.

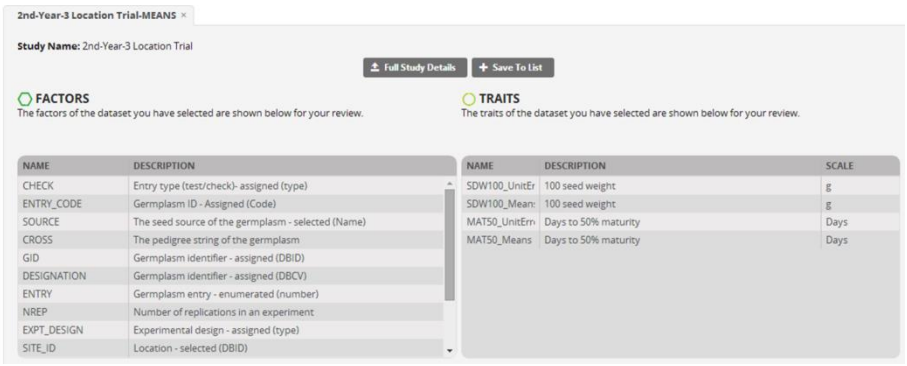


Select Datasets of Interest

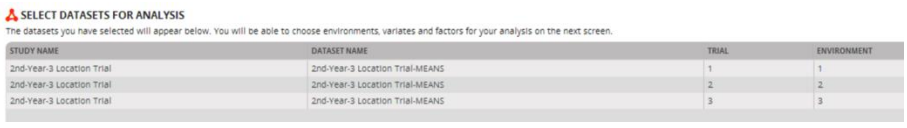
Select the raw data (PLOTDATA) for a trial or the means data (MEANS) created from the single site analysis to include in the export file.



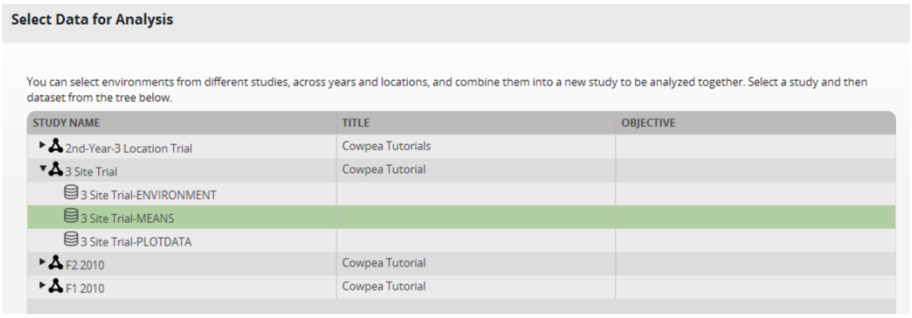
Review the data from a single trial and Save to List.



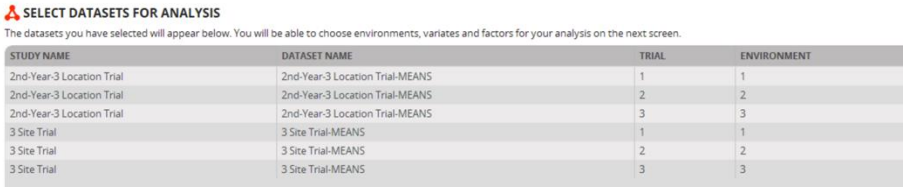
Multiple locations within the selected trial will display in the review window.



Select additional trial datasets.

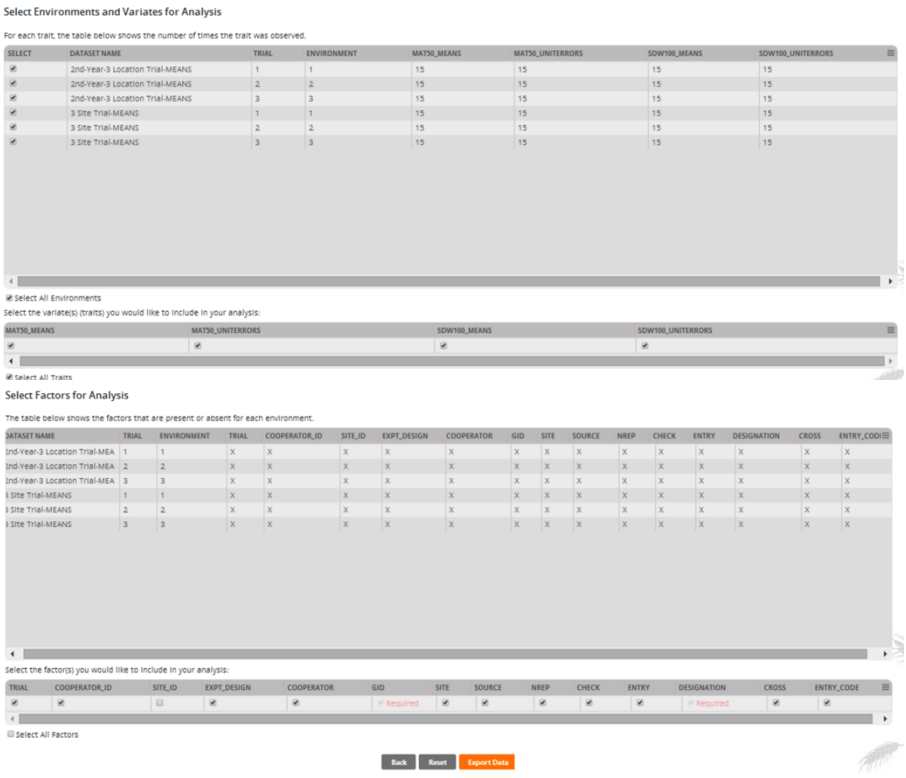


Review the content of selected datasets and click Next.

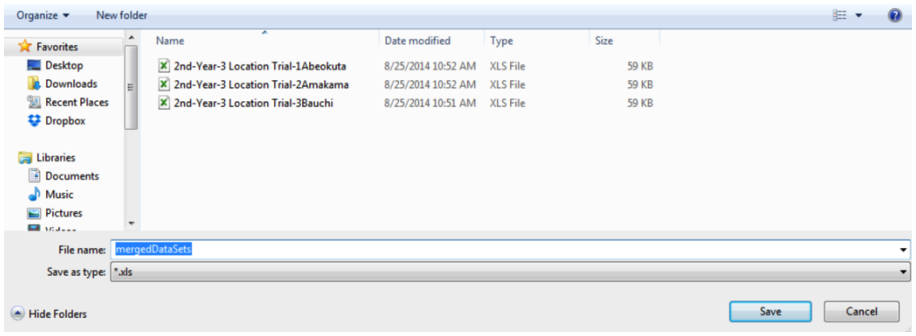


Specify Data to Export

Select environments, traits, and factors of interest to export.



The default name for the export file is mergedDataSets.xls. Choose a location for the export file and select save. The merged file contains combined data from all selected trials.



QTL Analysis: Cowpea Tutorial

Contributors

Darren Murray [a], Roger Payne [a] & Zhengzheng Zhang [a], Shawn Yarnes [b]
[a] *VSN International Ltd,* [b] *The Integrated Breeding Platform*

Summary

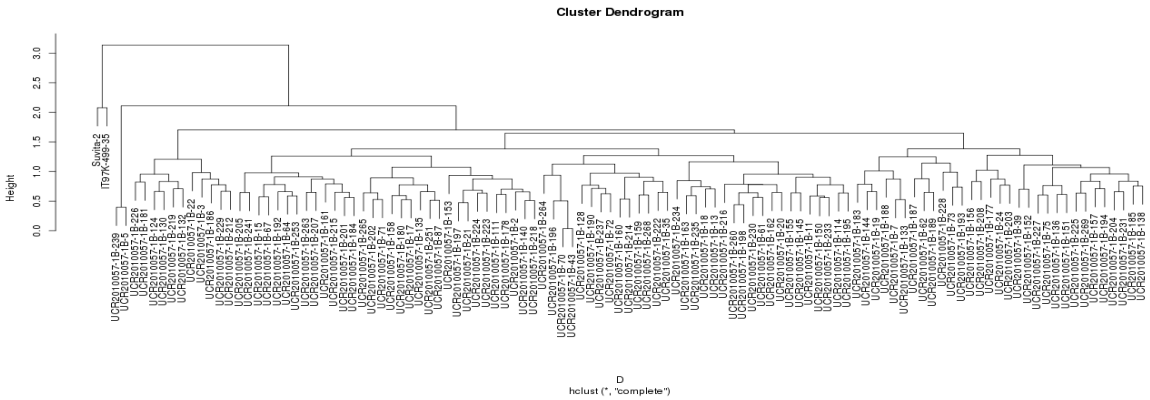
Use the standalone Breeding View application to correlate population genotypic variance and population phenotypic variance to identify quantitative trait loci (QTL). Visualize QTL, map, and genotypes using Flapjack.

Introduction

Breeding View uses a combination of simple interval mapping (SIM) and composite interval mapping (CIM) to test for significant correlations at each of the marker positions as well as between markers. The SIM method is followed by CIM to increase the power of the genome-wide QTL search. CIM includes candidate QTL identified by SIM as cofactors in the QTL scan model to control for variation caused by QTL outside the test region.

QTL Demonstration Data

Cowpea (*Vigna unguiculata*) genotypic and phenotypic data are from a population (N=111) of 109 F2 individuals and 2 parental lines, Suvita-2 and IT 97K-499-35. The F2 population has been genotyped with 164 markers. The map file contains 159 of the 164 markers mapped to 11 linkage groups.



Genotypic relationships in cowpea QTL demonstration data

Three data files are needed to perform a QTL analysis:

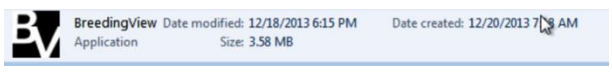
- Map file
- Genotype file
- Phenotype file

The map file specifies the marker locations by linkage group and position (cM) within linkage group. The genotype file contains the marker scores of each individual in the population. The map and genotype files are plain text files in Flapjack format. The phenotype data file contains the observations of one or more traits of each individual in the population.

Start New QTL Project

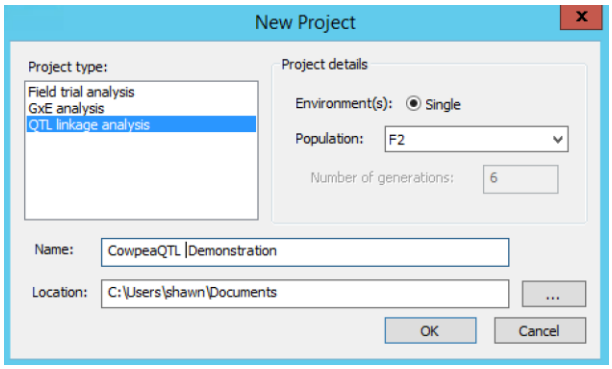
The standalone Breeding View application can be launched from the left hand Workbench menu.

Alternatively, Breeding View can be launched independently of the Workbench by accessing the application within the Tools folder > Breeding_View.

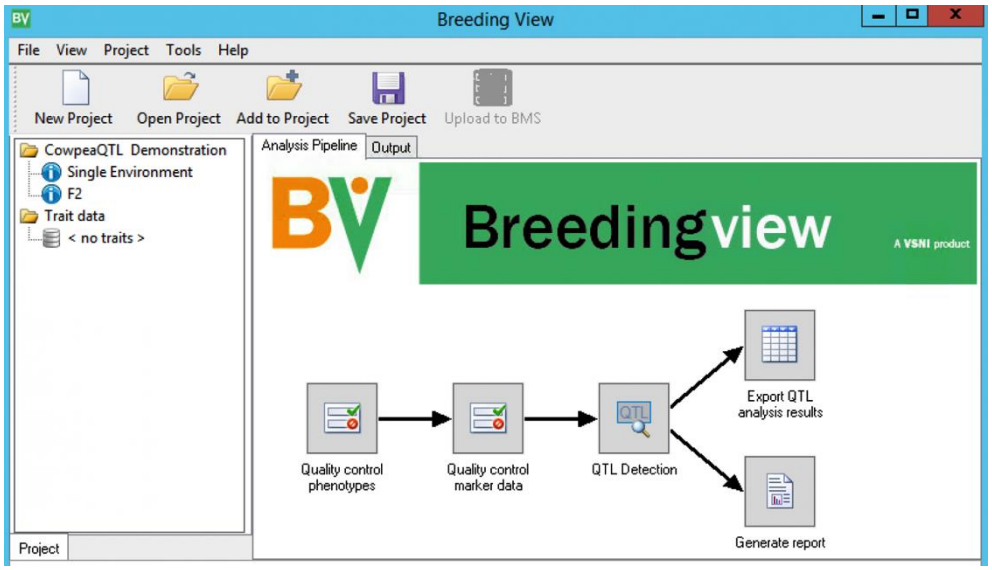


When Breeding is launched, the Project window and Analysis Pipeline will not contain any information.

Create a new project, by selecting File > New Project from the menu bar, or click on the New Project tool button. In the New Project dialog, select the QTL linkage analysis item within the Project type list. Enter project details. The current version only supports single site analysis, so select Single for the Environment and choose the appropriate Population description from list. Select F2 population. Name the project and select the location of the working directory. Click OK to create the project.

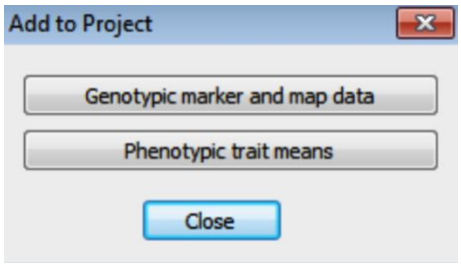


When a project has been created, the Analysis Pipeline is displayed along with project details on the left. The Analysis Pipeline includes nodes, which are connected by arrows. Each node represents a different task that is performed within the QTL analysis pipeline.



Import Data Files

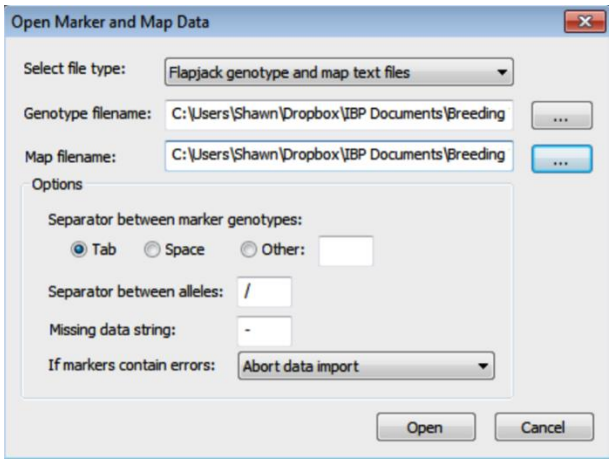
The genotypic and phenotypic data need to be imported into the project before the analysis pipeline can be run. To import the genotypic data, select Project > Add Data from the menu, or click on the Add to Project tool button. On the Add to Project dialog, click on the Genotypic and Map Data button.



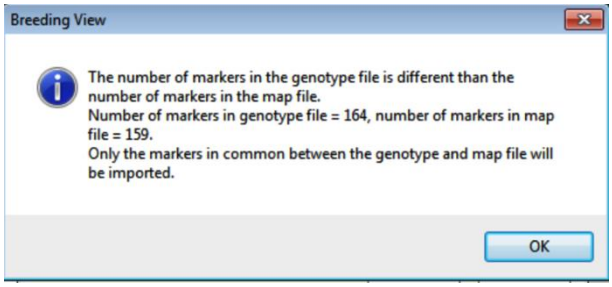
BMS File Directory

C:\Breeding Management System\Documents\Breeding View\Sample Files\data1\Cowpea_genotypes.txt
C:\Breeding Management System\Documents\Breeding View\Sample Files\data1\Cowpea_map.txt
C:\Breeding Management System\Documents\Breeding View\Sample Files\data1\Burkina_trait_means_for_qtl_analysis.txt

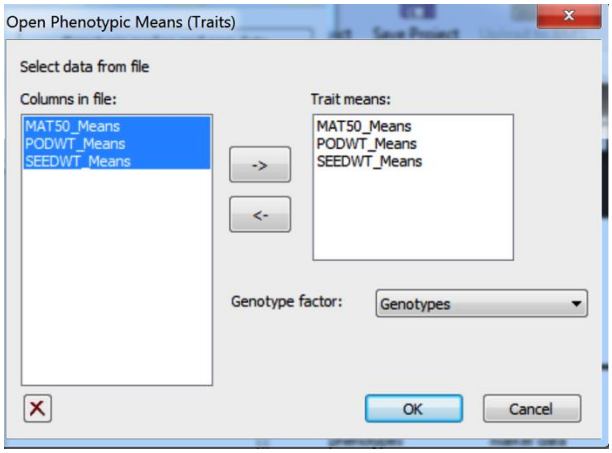
In the Open Marker and Map Data dialog, select the Flapjack formatted genotype and map .txt files. Browse for files by clicking on the [...] button. Click Open to import the data.



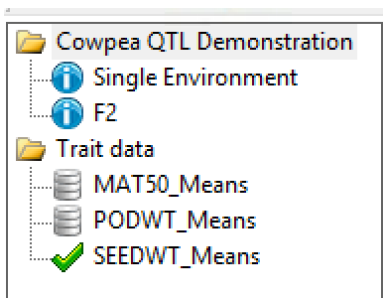
Breeding View will only import information on markers common to genotype and map files. Select OK.



Import the phenotypic data file by selecting Phenotypic Trait Means from the Add to Project menu. Import all three traits, and choose Genotypes as the genotype factor. Click OK to import the data.

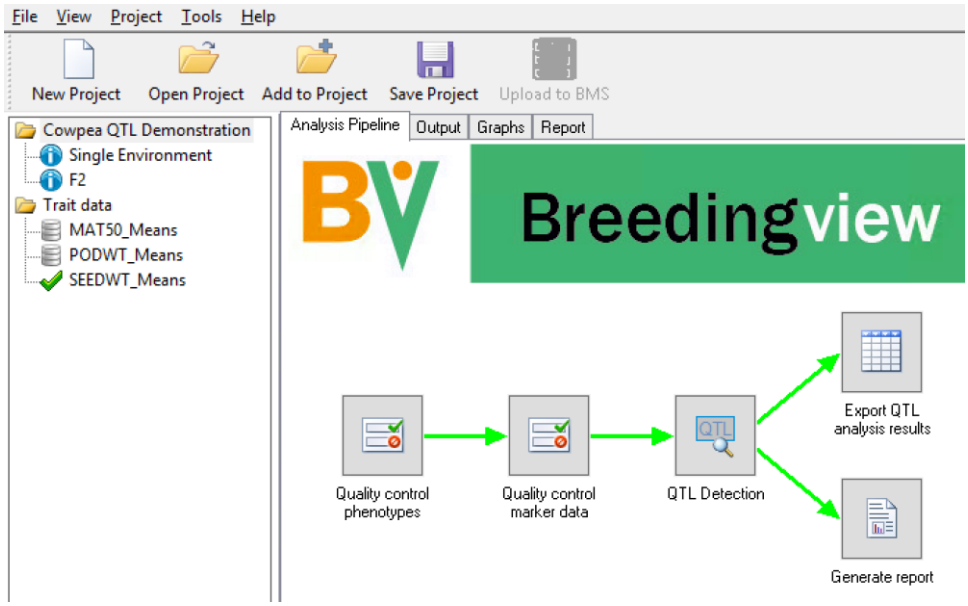


After the data have been imported, the selected traits appear within Project tab on the left side of the screen. Set Seed Weight as the active trait by right-clicking SEEDWT_Means and selecting Set As Active Trait. The green check mark indicates that it is the trait that will be used within the analysis pipeline.



Run Analysis

To run the QTL analysis pipeline, click on the right-mouse button on the first node, Quality Control Phenotypes, and select Run Pipeline. This will run the whole analysis pipeline, by performing the task at each node in turn. As the task at each node is completed, the connecting arrow will change color, to indicate that the pipeline is moving onto the next task.

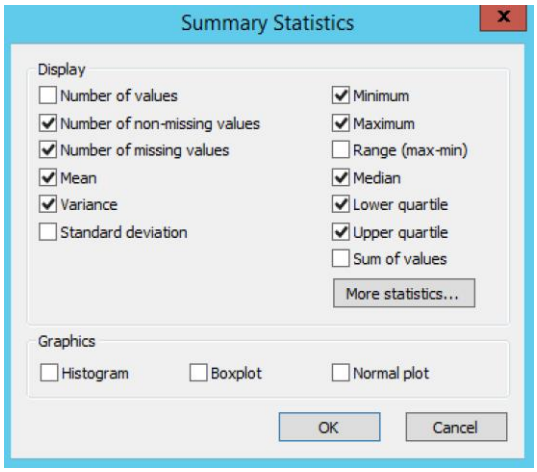


Analysis Options

Some of the nodes have options to control they way an analysis is performed and output that is displayed. To access the options, right-click on a node and select the Settings item from the shortcut menu. The changes to the options are retained during the current session and are saved to the project file.

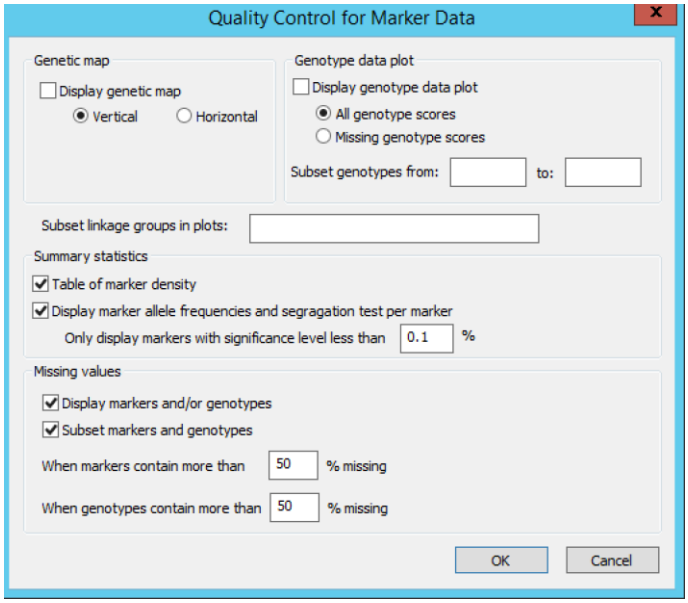
Summary Statistics

Settings under Quality Control Phenotypes define the phenotypic summary statistics that are displayed in the output.



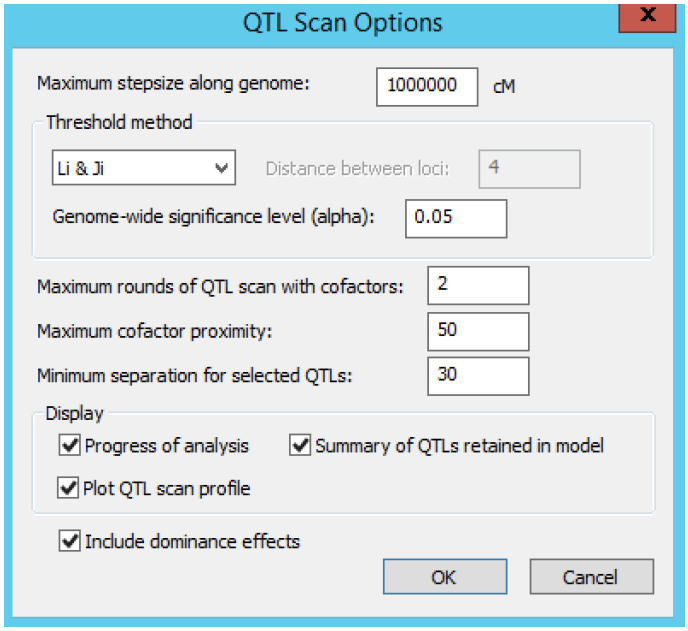
Quality Control Marker Data

Settings under Quality Control Marker Data define the display, summary statistics, and missing value thresholds for marker data.

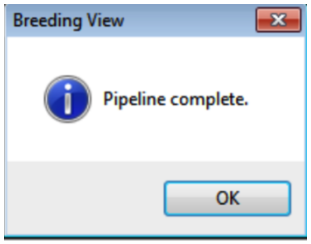


QTL Scan Options

Settings under QTL Scan Options define the QTL scanning conditions. By default Breeding View performs a single SIM followed by two rounds of CIM scanning. The number of CIM scans can be increased until the list of detected QTLs does not change. In CIM scanning, to avoid co-linearity between cofactors and tested positions, cofactors are removed temporarily from the model when testing for QTLs close to cofactor positions. The window within which cofactors are removed is set by default to 50 cM.



Leave the default settings for this tutorial and run the entire analysis pipeline by right clicking the first node and selecting Run the Pipeline. When the analysis completes, a prompt will appear to indicate it has finished.

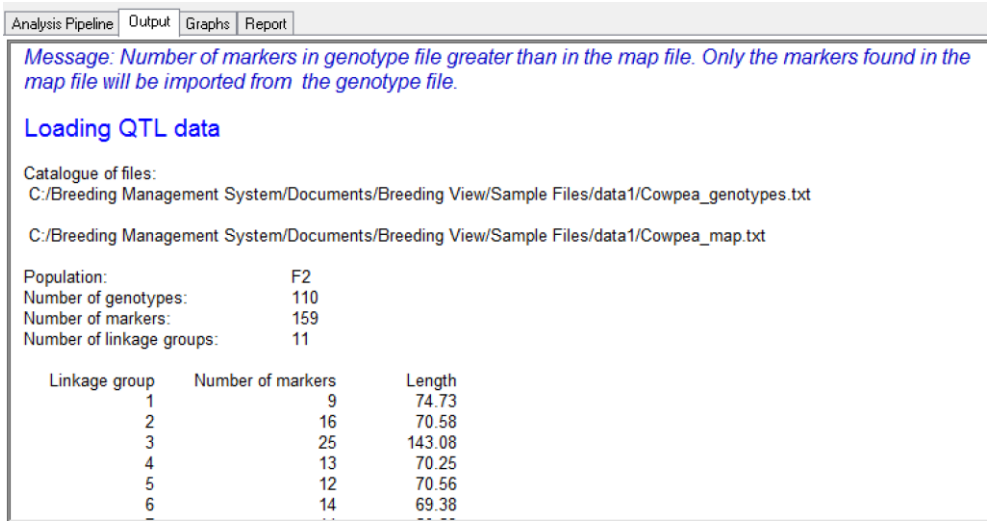


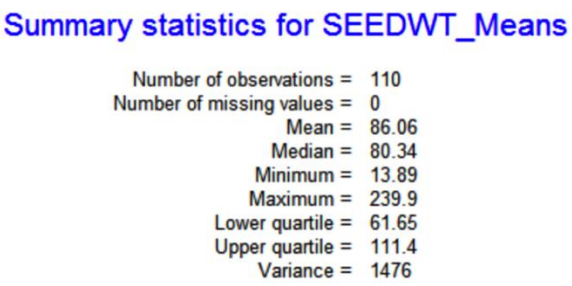
Restart Analysis

After a pipeline has been run, you can change settings and start again from any point along the pipeline.

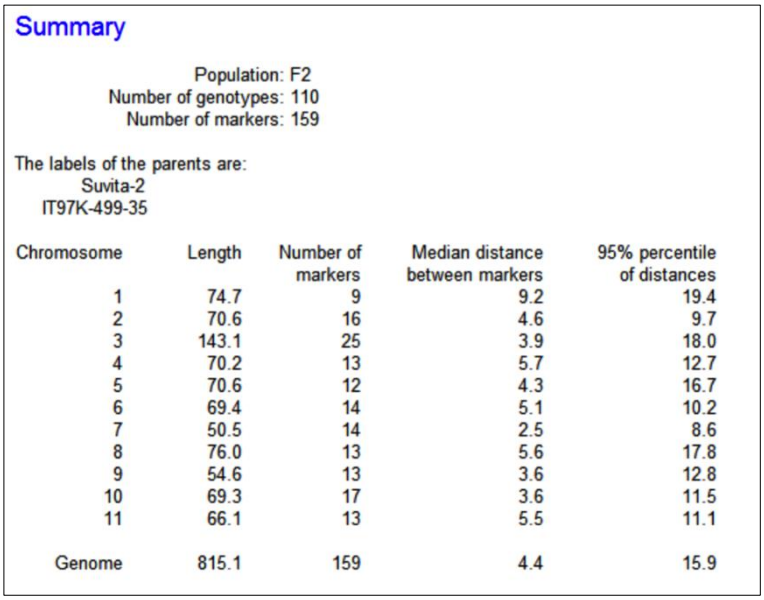
Results

The output tab contains the summary statistics for seed weight means, marker data, and detected QTL.





Summary statistics seed weight Cowpea QTL



Summary statistics marker data Cowpea QTL

The summary for the marker data, displayed in the Output tab, provides details of the numbers of markers within each linkage groups and the number of missing observations.

Missing values

There are 526 scores missing. This is 3.007% of the 17490 scores.

There are 156 markers with missing values. This is 98.11% of the 159 markers.

There are no markers with more than 50% missing values.

There are 87 genotypes with missing values. This is 79.09% of the 110 genotypes.

The 1 genotypes with more than 50% missing values over the 159 markers are:

Genotype	Number of missing values	Percentage missing values
UCR2010057-1B-40	84	52.8

There are no markers with 50% or more missing observations. However, there is one genotype, UCR201005-1B-40, reported that contains greater than 50% missing observations, so this has been removed from the QTL detection.

Output is produced for each round of SIM and CIM, with a summary of the loci greater than the detection threshold and candidate QTL.
After the last round of CIM, the effects of significant QTL are estimated.

Summary

Trait: SEEDWT_Means
Population type: F2
Number of genotypes: 109
Number of linkage groups: 11
Number of markers: 159

List of QTLs

Locus no.	Locus name	Linkage group	Position	-log10(P)
78	1_0074	6	11.45	3.629
106	1_0771	8	26.75	5.916

QTL effects

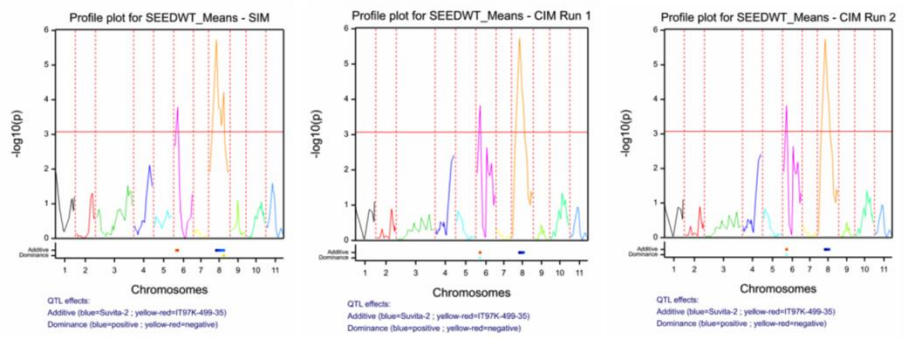
Locus no.	Locus name	%Expl. Var.	Add. eff.	High value allele	s.e.	Dom. eff.	Dominant allele	s.e.
78	1_0074	11.908	18.783	IT97K-499-35	4.576	12.903	Suvita-2	6.396
106	1_0771	17.237	22.599	Suvita-2	4.386	*		*

Estimated lower and upper bounds of QTL positions

Locus no.	Locus name	Lower bound	Position	Upper bound
78	1_0074	0.000	11.450	69.380
106	1_0771	0.000	26.750	69.350

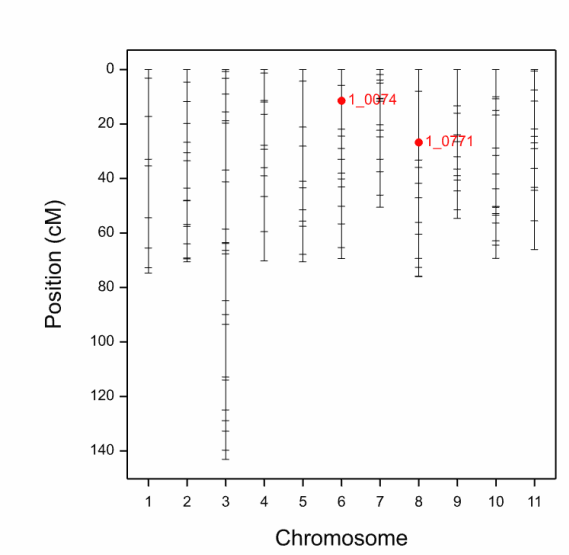
Two QTL associated with seed weight identified on chromosomes 6 and 8 explain 11.9 and 17.2% of the variance in seed weight respectively.

The QTL detection produces graphs of the scan profiles for the simple interval mapping and each round of composite interval mapping. Each graph displays the profile within each linkage group, the threshold of detection, and parent contributor of the high value allele. All figures are saved as image (.png) files within the working project folder, in a subfolder using the date and time. The phenotypic identifier will begin the file name.



Seed weight QTL profiles for SIM and two rounds of CIM

A figure where QTL are overlayed on the genetic map is automatically saved an image (.png) file to the working directory when the pipeline is run.

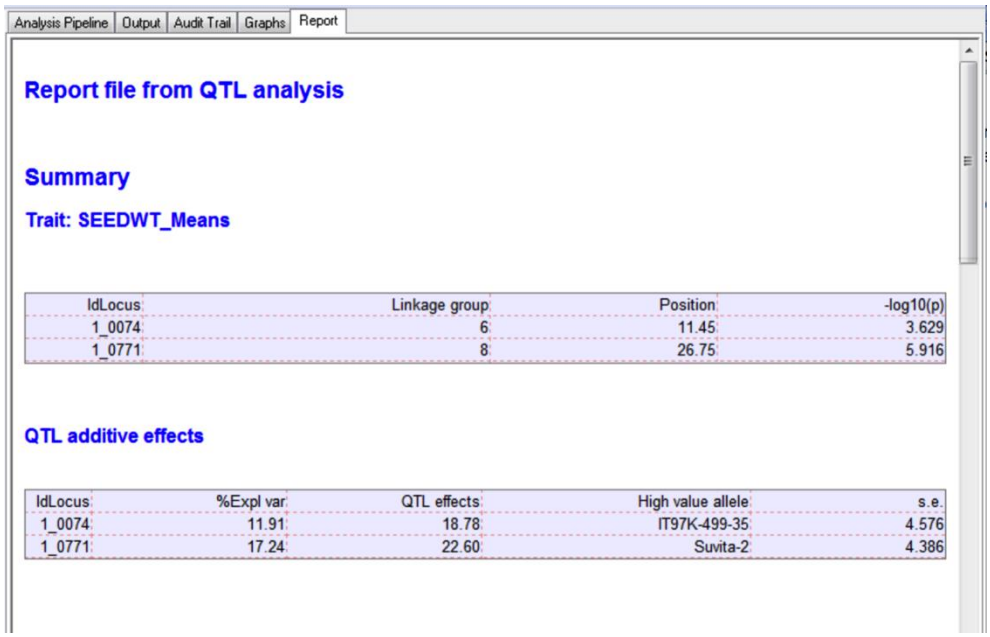


SEEDWT_Means_Report_GeneticMap001.png from the cowpea QTL demonstration data

Report

The final results are collated and displayed in the Report tab within an HTML document. The HTML report (Report.htm) is automatically saved within the working project folder, in a subfolder using the date and time.

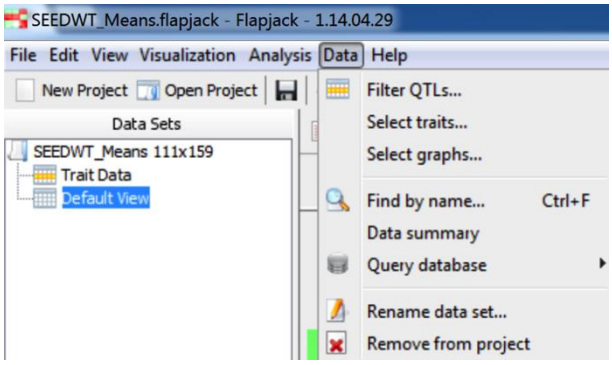
The seed weight report and genetic map graphs are automatically named from the column heading in the phenotype file (SEEDWT_Means_Report.htm and SEEDWT_Means_Report_GeneticMap001.png).



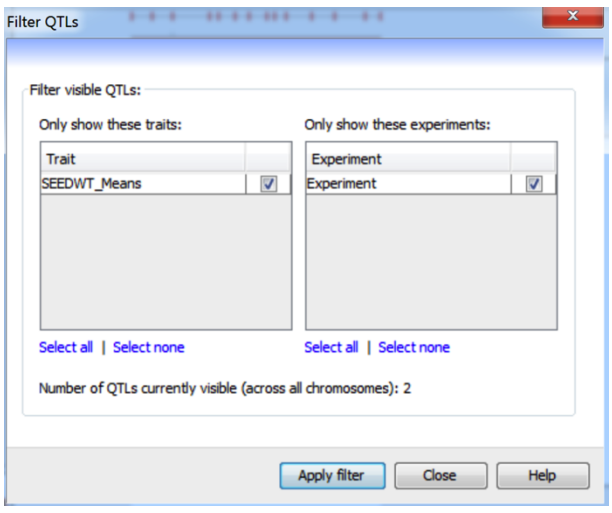
Visualize QTL in Flapjack

When the QTL analysis pipeline is complete, Flapjack (Milne et al., 2010) automatically opens a graphical genotype display. The Flapjack project file is saved in a file within the working project folder in a subfolder using the date and time of the run. The file is automatically named SEEDWT_Means.flapjack. You may be asked to update your version of Flapjack before proceeding.

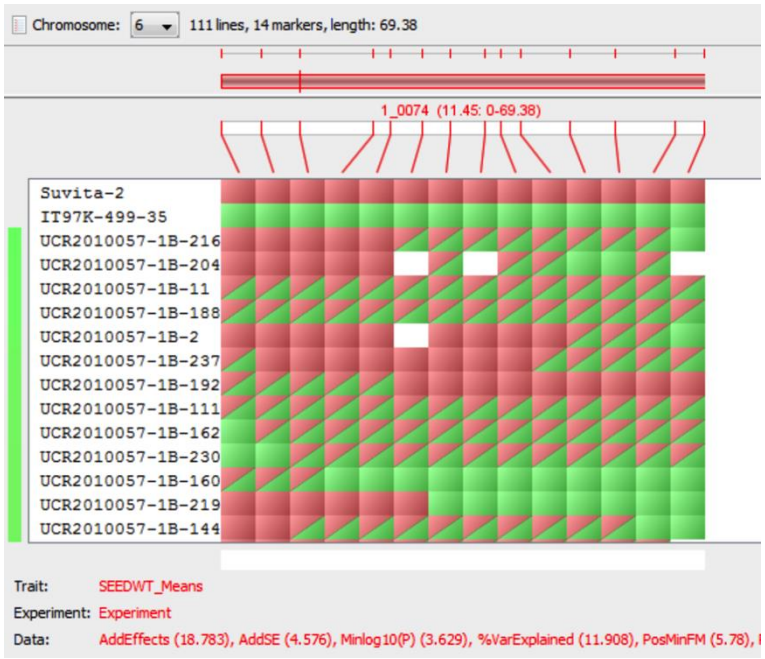
Overlay QTL on the genotypic display by selecting Filter QTLs from the Data menu.



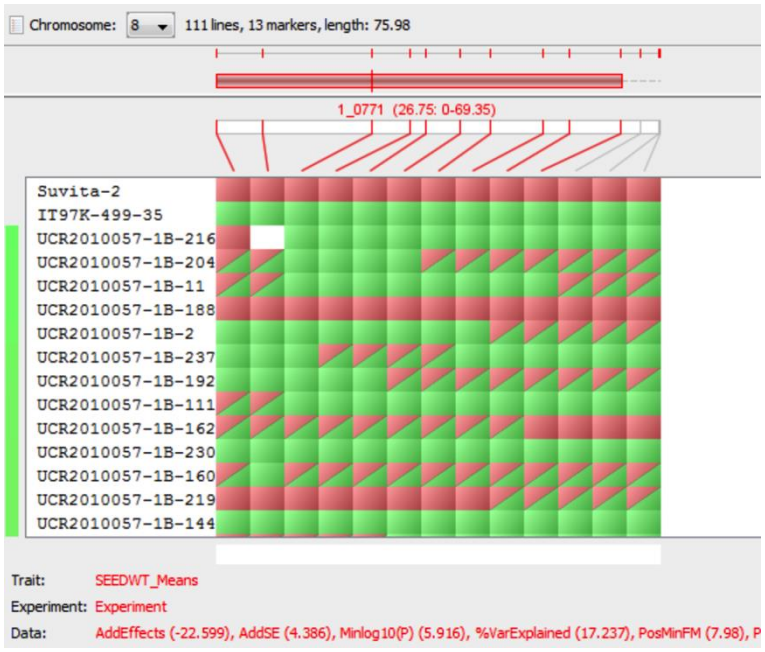
Select the trait(s) of interest and the relevant experiment.



QTL are now visible on their respective chromosomes.



QTL associated with seed weight on chromosome 6 at 11.45 cM



QTL associated with seed weight on chromosome 8 at 26.75 cM

References

Roberts, P. A., J. D. Ehlers, T. J. Close, N. Cisse, B. L. Huynh, I. Drabo, M. R. Lucas, and P. da Silva Vinholes. 2013.Frontiers in Plant Science. Association studies and legume synteny reveal haplotypes determining seed size in *Vigna unguiculata*. 4: 95.

Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WTB, Flavell AJ, and Marshall D. 2010. Flapjack graphical genotype visualization. Bioinformatics 26(24), 3133-3134.

Murray, D. Payne, R, & Zhang, Z. (2014) Breeding View, a Visual Tool for Running Analytical Pipelines: User Guide. VSN International Ltd. (.pdf) (associated sample data .zip)

Acknowledgements

The statistical algorithms in the Breeding View were developed by VSN International Ltd in collaboration with the Biometris group at University of Wageningen. Cowpea QTL demonstration data were provided by Issa Drabo at the Institut de l'Environnement et de Recherches Agricoles in Burkina Faso and Jeff Ehlers, Tim Close, Philip Roberts, Bao Lam Huyuh at the University of California Riverside. These data may have been adapted for training purposes. Any misrepresentation of the raw breeding data is the solely the responsibility of the IBP.

Marker-Assisted Breeding

Molecular Breeding Planner

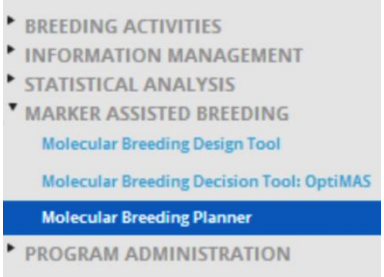
The Molecular Breeding Planner facilitates planning for population development, selections, and crosses using molecular marker data. Marker-assisted selection (MAS) can speed the breeding process by increasing the frequency of favourable alleles and decreasing the frequency of unfavourable alleles in breeding populations. The Molecular Breeding Planner helps breeders develop breeding schemes, including timelines and optimal population sizes, to improve the efficiency of genetic gains using genetic markers.

Planning tools support

- Marker-Assisted Recurrent Selection (MARS)
- Marker-Assisted Backcross (MABC)
- Marker-Assisted Gene Pyramiding

Launch

The Molecular Breeding Breeding Planner is a standalone application that can be launched from the Workbench menu.

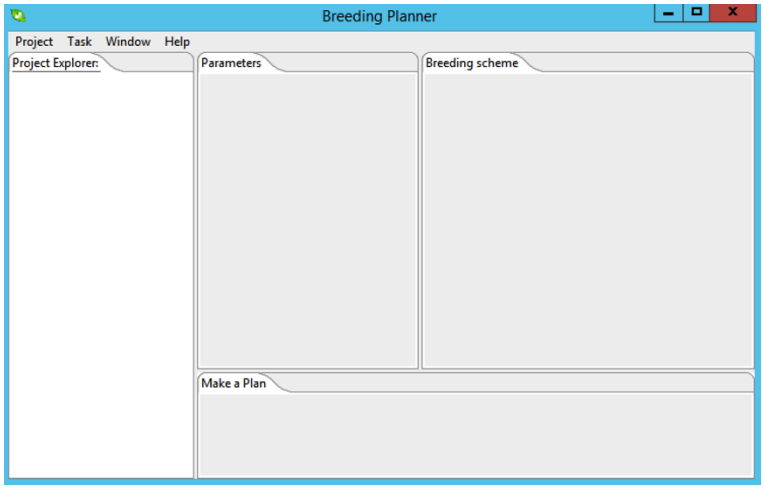


Alternatively, the Molecular Breeding Planner can be launched independently of the Workbench by accessing the application within the Tools folder.

Interface

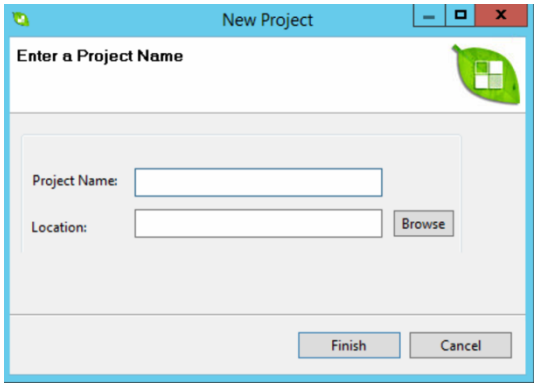
The interface contains the following windows, which will be empty until you establish a project:

- Project Window: Lists all molecular breeding programs within an open project- three distinct programs can be considered: MARS, MABC and MAS for gene pyramiding.
- Parameter Setting/Viewing Window: You can edit/view your breeding parameters in this window.
- Breeding Scheme Window: Once the breeding parameters are specified, a breeding flowchart will be demonstrated in this window.
- Plan-Making Window: You can select the current stage/generation of your breeding programs in this window to formulate a detailed plan for the near future.



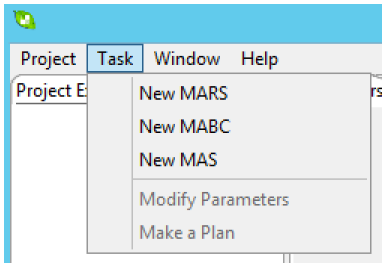
Create a New Project

Select New Project from the Project dropdown menu. Give your project file a name and select a destination folder for the output. Multiple breeding plans can be saved within a single project file.

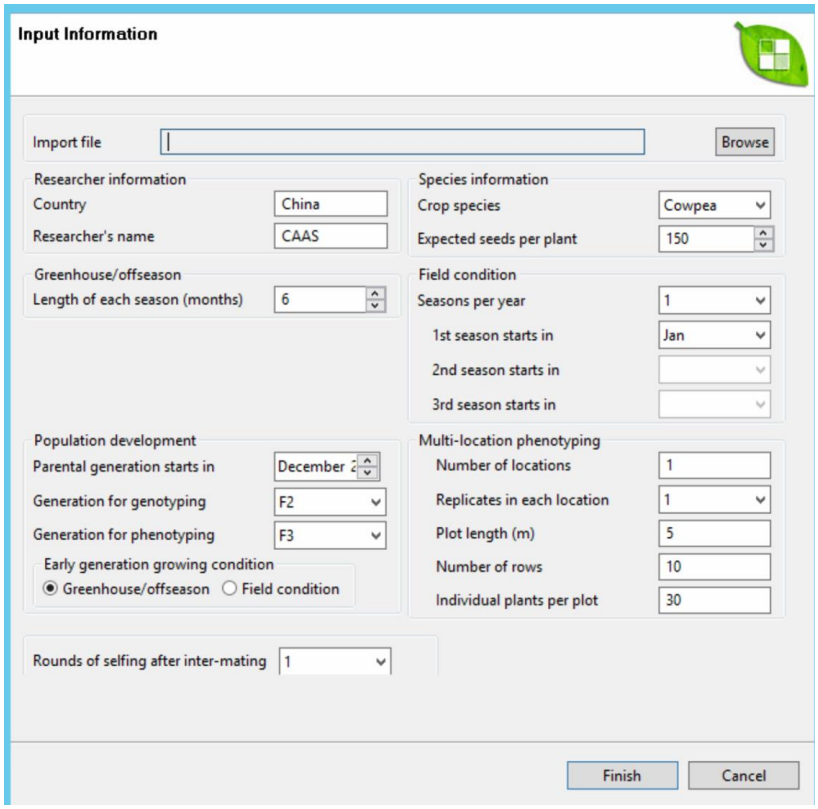


From the task menu, select one of the new breeding plans:

- MARS: Marker-Assisted Recurrent Selection
- MABC: Marker-Assisted Backcross Breeding
- MAS: Marker-Assisted Selection for Gene Pyramiding



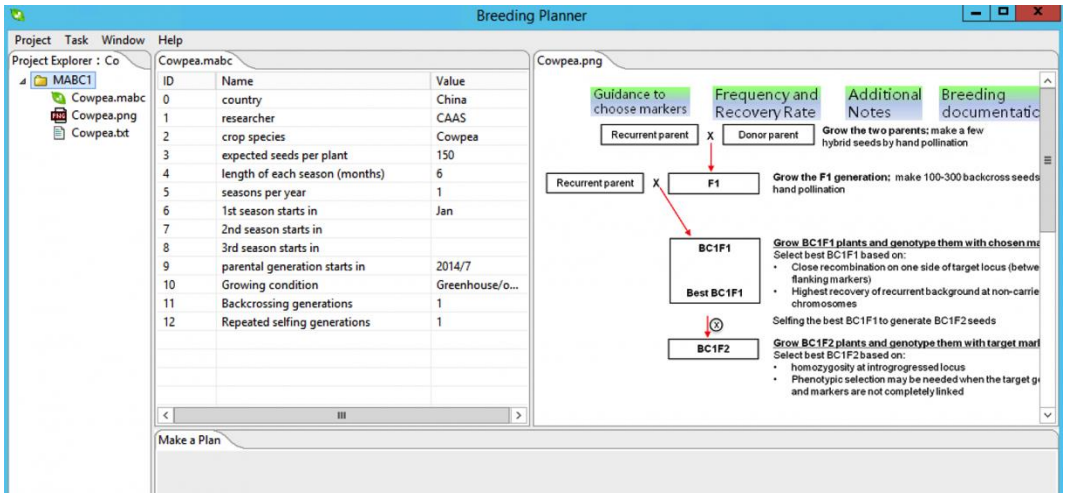
Add details of the breeding plan to the Input Information screen. Users can import the parameters from an external file (file extensions .mars, .mabc, or .mas) or set the parameters by hand.


Input Information
Import file
Browse
Researcher information
Country
China
Researcher's name
CAAS
Species information
Crop species
Cowpea
Expected seeds per plant
150
Greenhouse/offseason
Length of each season (months)
6
Field condition
Seasons per year
1
1st season starts in
Jan
2nd season starts in
3rd season starts in
Population development
Parental generation starts in
December
Generation for genotyping
F2
Generation for phenotyping
F3
Early generation growing condition
Greenhouse/offseason
Field condition
Multi-location phenotyping
Number of locations
1
Replicates in each location
1
Plot length (m)
5
Number of rows
10
Individual plants per plot
30
Rounds of selfing after inter-mating
1
Finish
Cancel

About Expected Seeds Per Plant

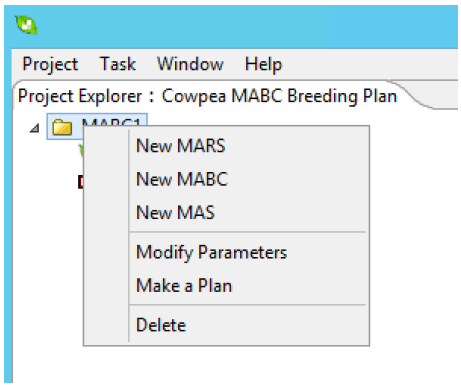
- Expected Seeds Per Plant is set to a crop specific default. The user can modify this number, but the input must fall into the min-max range for the selected crop. If input number is smaller than the minimum number, the minimum number will be assumed. If the input number is greater than the maximum number, the maximum number will be assumed.
- Expected seeds per plant is used to determine seed availability for phenotyping. If seed inventory is too low , additional seed increase by selfing will be requested.
- The number of seeds required for phenotyping is calculated from settings for Multi-Locational Phenotyping. For example, multi-locational phenotyping is only possible when there are enough F2 plants to produce the expected amount of seed. Otherwise, phenotyping will be delayed until the required seeds are produced.

Select Finish to populate the parameters and breeding scheme.



Make a Breeding Plan

Once the parameters are set, right click the breeding method folder, and select Make a Plan from the dropdown menu.



Marker-Assisted Recurrent Selection (MARS) Plan

The marker-assisted recurrent selection (MARS) planning tool schedules events, models QTL to include as selection criteria, proposes selections, plans crosses, and predicts homozygosity in fixed line development.

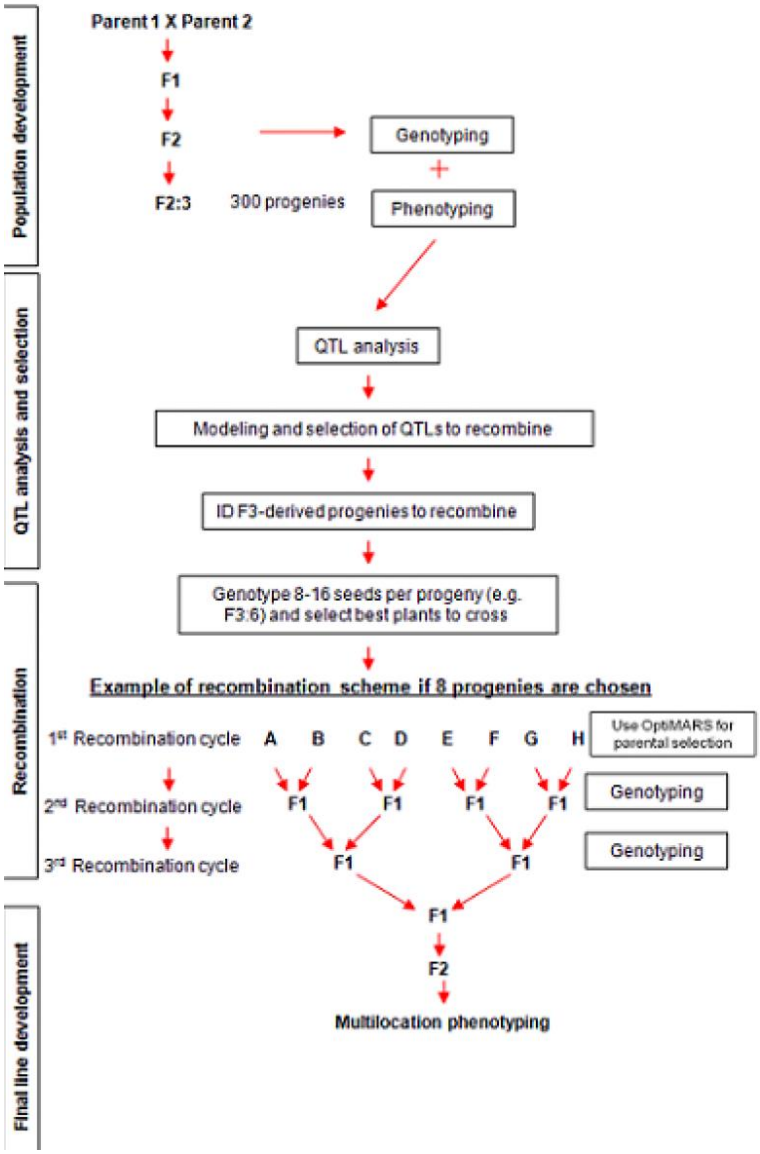
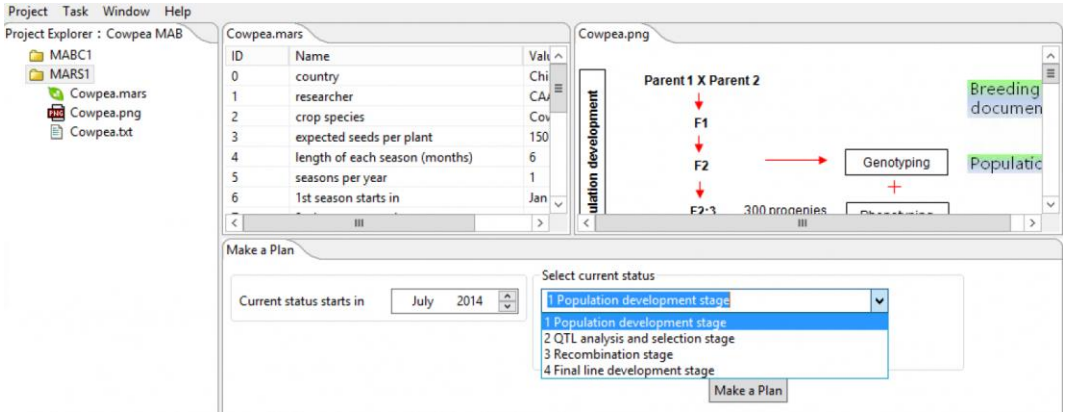


Diagram of a MARS breeding program in cowpea, a self-pollinated species

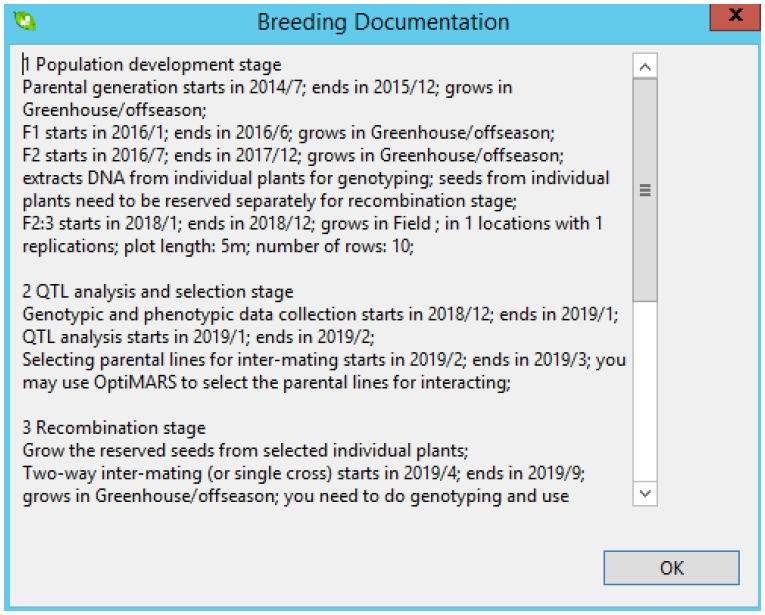
MARS consists of four main stages:

- Training population development: A training population is evaluated for phenotype in field trials and genotyped using mapped genetic markers. The MARS planning tool will assist in scheduling training population events to meet anticipated deadlines.
- QTL analysis and selection: QTL analysis correlates phenotypic variance to genotypic variance, and identifies putative genetic regions involved in quantitative traits. The MARS planning tool will schedule QTL analysis and model a selection of QTL that optimize the traits of interest for recombination. QTL analyses can be performed with the Breeding Management Systems Breeding View statistical software (see section 5.8) or with other external software applications.
- Recombination: After QTL are identified another Breeding Management Tool, OptiMAS, can be used to identify progeny from the training population for inter-mating. The MARS Planning Tool develops a plan for crossing progeny selected using OptiMAS.
- Fixed line development: The MARS Planning Tool will propose optimum selection scheme for each generation of selfing and predict the homozygosity of the breeding population.

Choose start date and the set the current status of the breeding program. Select the Make a Plan button.

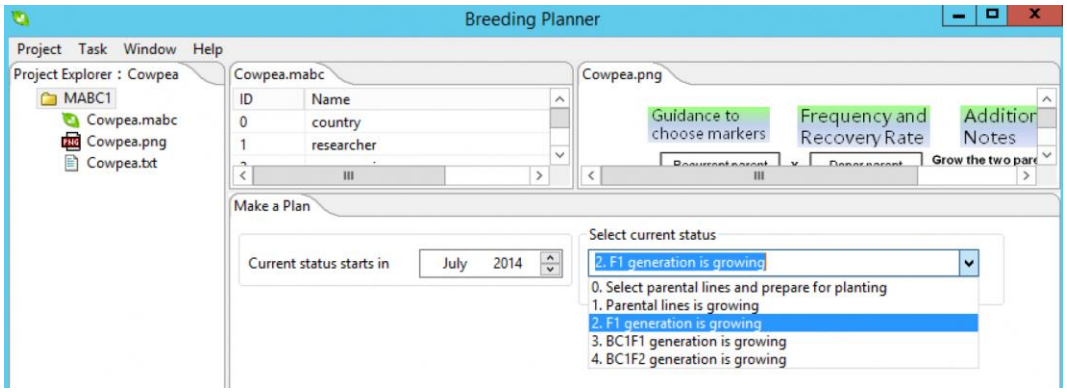


The breeding documentation window will appear with a timeline for the project accessible as a .txt file in the project folder.

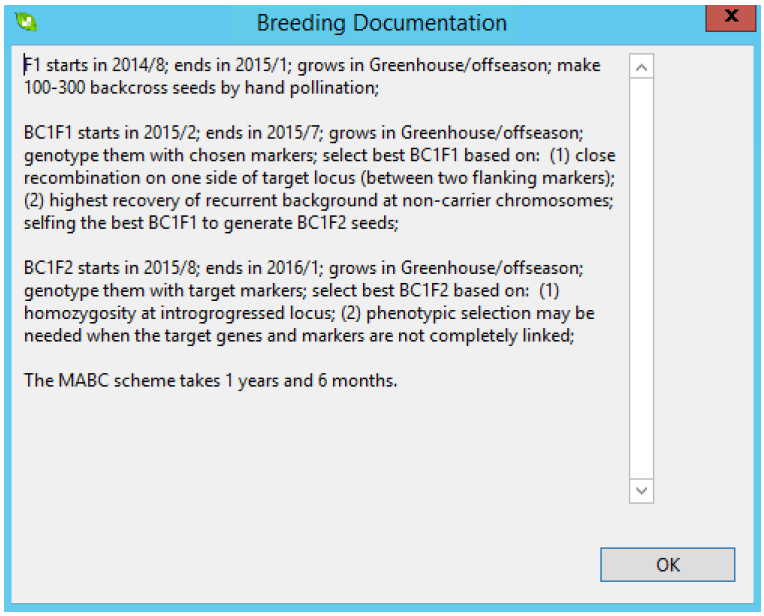


Marker-Assisted Backcross Breeding (MABC) Plan

Choose start date and the current status of the breeding program. Select the Make a Plan button.



The breeding documentation window will appear with a timeline for the project accessible as a .txt file in the project folder.

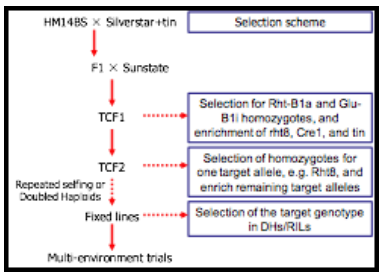


Marker-Assisted Selection (MAS) for Gene Pyramiding Plan

Selection for two or more genes at a time, or gene pyramiding, can be greatly enhanced by marker-assisted section (MAS). Gene pyramiding requires a complex design strategy that outlines the number of markers required, best crossing strategies, and the level of inbreeding to maximize the efficiency of marker implementation. The Molecular Breeding Planner does a simulation analysis to develop breeding plans for gene pyramiding.

Gene	Rht-B1	Rht-D1	Rht8	Sr2	Cre1	VPM	Glu-B1	Glu-A3	tin
Chr.	4BS	4DS	2DL	3BS	2BL	7DL	1BL	1AS	1AS
Marker	Codom	Codom	Codom	Codom	Dom	Dom	Codom	Codom	Codom
MK-gene distance	0	0	0.6	1.1	0	0	0	0	0.8
HM14BS	Rht-B1a	Rht-D1a	Rht8	sr2	cre1	vpm	Glu-B1a	Glu-A3e	Tin
Sunstate	Rht-B1a	Rht-D1b	rht8	Sr2	cre1	VPM	Glu-B1l	Glu-A3b	Tin
Silverstar +tin	Rht-B1b	Rht-D1a	rht8	sr2	Cre1	vpm	Glu-B1l	Glu-A3c	tin
Target	Rht-B1a	Rht-D1a	Rht8	Sr2	Cre1	VPM	Glu-B1l	Glu-A3b	tin

Nine major genes and their distribution in three wheat parental lines



Optimum crossing and selection scheme for gene pyramiding identified by simulation

Save a Project

When the Molecular Breeding Planner is closed, projects are automatically are saved as .ibp files. A folder titled with your project name will appear in the destination folder you established when you created the project. This project folder will automatically be populated with the .ibp project file and subfolders relating to different marker-assisted breeding plans.

Individual breeding plan files that are displayed in the Project window can also be saved independently of the project folder by right clicking and selecting the Save As option.

Open Existing Project

From Open in the Project dropdown menu, you can browse for existing .ibp project files.

Alternatively, from New Project in the Project dropdown menu, you can search for existing projects if you know the name and can specify the root directory.

Molecular Breeding Design Tool for Backcrossing

The Molecular Breeding Design Tool can be used in backcross breeding applications to design ideotypes based on QTL target regions (foreground markers) and to recover recurrent parent genome (background markers). The graphical display facilitates the comparison of germplasm based on genotype.

Graphical display assists with:

- The selection of donor and recurrent parents
- Assign foreground markers associated with QTL
- Assign background markers to recover recurrent parent genome
- Design a target genotype (ideotype)

Additional tools to:

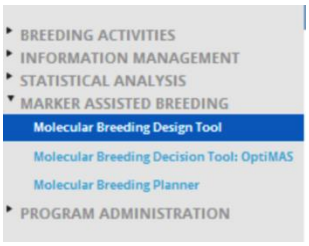
- Determine the minimum number of individuals that need to be genotyped in each BC generation to recover at least one double homozygote for the markers flanking target QTL at the end of the program
- Identify polymorphic makers between any two accessions

Acknowledgement

The Molecular Breeding Design tool was developed by Trushar Shaw, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) Patancheru, India with funding from the Integrated Breeding Platform.

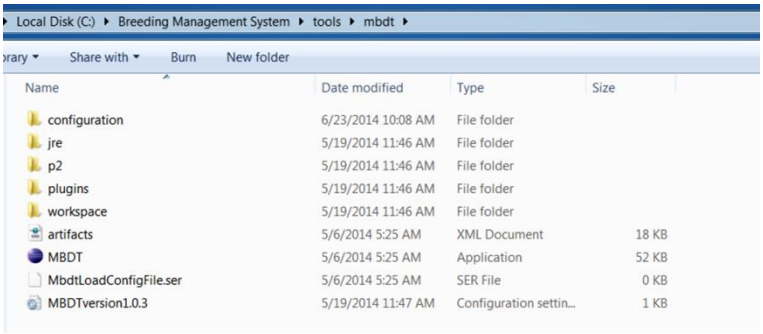
Launch

The Molecular Breeding Design Tool (MBDT) is a standalone application that can be launched from the left hand Workbench menu.



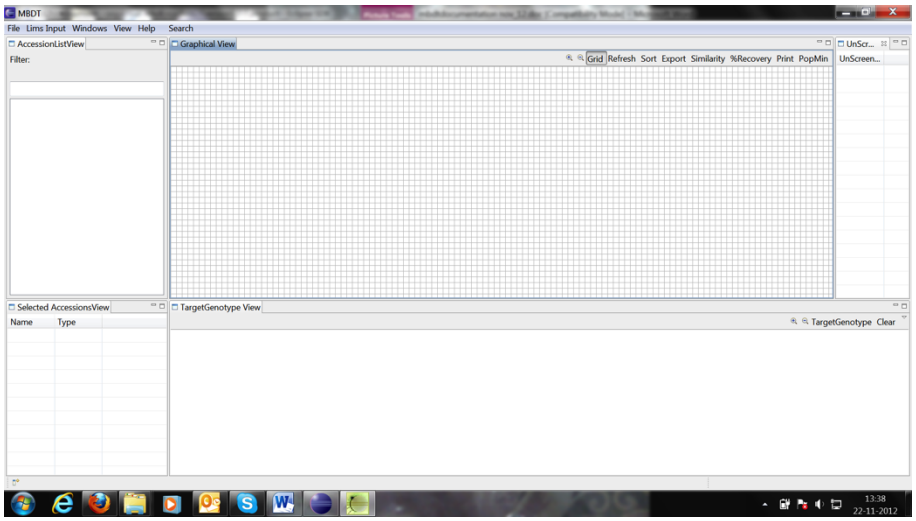
Alternatively, the Molecular Breeding Design Tool can be launched independently of the Workbench by accessing the application within the Tools folder.

BMS File Directory
C:\Breeding Management System\Tools\MBDT



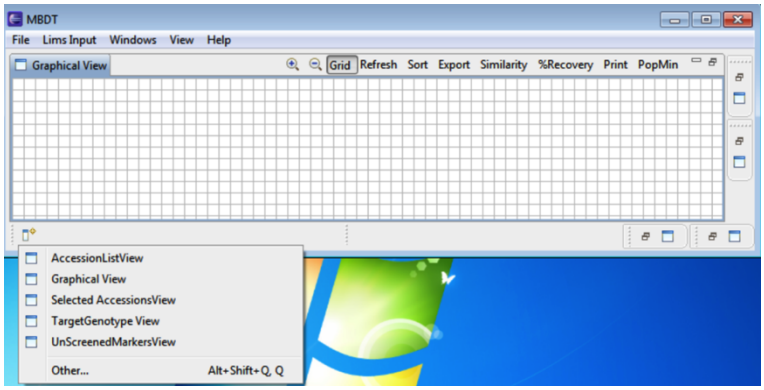
Navigating the Workbench

The application opens with the MBDT workbench, where you can create and manage your projects.

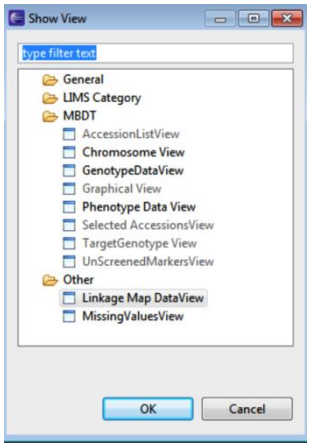


Default workbench view

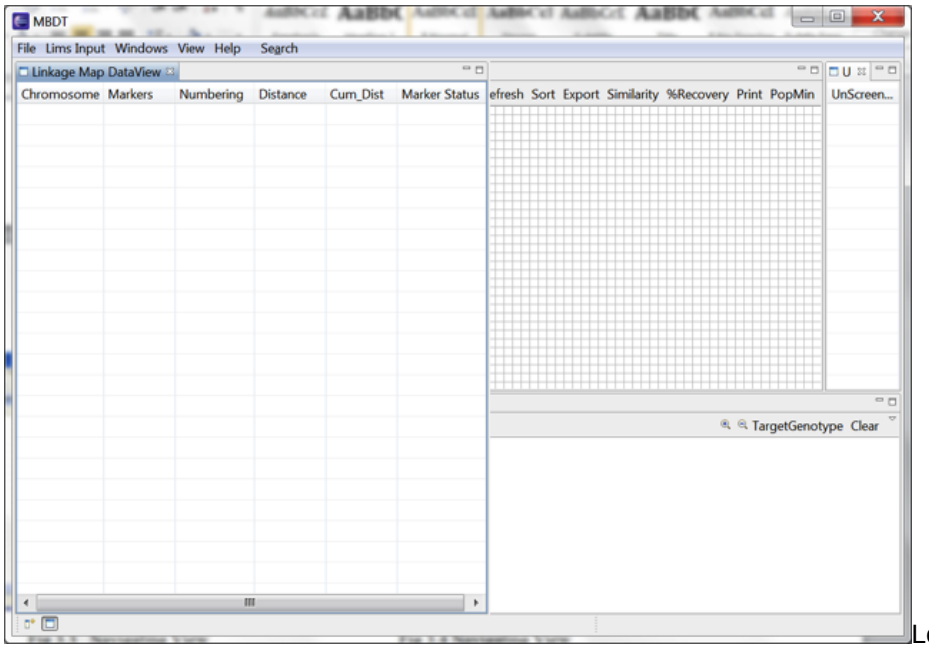
Clicking the button in the bottom left of the window displays the other tab options.



Selecting ‘Other’ allows for additional windows to be selected and opened by selecting OK.



Double clicking any tab brings the window to full screen. The windows on the workbench can be dragged and dropped to customize the workbench.



Left: Linkage Map Data View is opened in the workbench at the starting position of the workbench. Right: Linkage Map Data View repositioned in right corner of the workbench

Input Files

MBDT input files include tab delimited text files:

- Genotype data in GCP genotyping data template,
- Linkage Map data in CMTV readable file format,
- QTL information from iMAS
- Phenotype Data

BMS File Directory

C:\Breeding Management System\Documents\MBDT\Input\ Files\Sample Data

Genotype File

Prepare the genotype data file as a .txt file. When saving in Microsoft Excel select the Tabs Delimited Text option. The DatasetID column identifies the dataset. The Genotype/Marker identifies the lines or accessions. Subsequent columns contain the marker data .

	A	B	C	D	E	F
1	**DatasetID**	**Genotype/Marker**	**umc83a**	**umc83a**	**bnl6.29b**	**bnl6.29b**
2	2	IS8603	191	191	110	110
3	2	IS8607	191	191	125	125
4	2	IS8608	191	191	110	110
5	2	IS8610	191	191	113	122

Prepare advanced generation genotype data as follows:

	A	B	C	D	E	F
1	C1:Ch1	bnl5.62a	0	0	0	0
2	C1:Ch1	umc164c	1.0	7.4	7.4	0
3	C1:Ch1	gsr1	0	2.0	33.9	41.3
4	C1:Ch1	UMC11a	3.0	32.1	73.4	0
5	C1:Ch1	umc53c	0	4.0	10.0	83.4

Linkage Map File

Prepare the Genotype data file as a .txt file. When saving in Microsoft Excel select the Tabs Delimited Text option. The DatasetID column identifies the dataset. The Genotype/Marker identifies the lines or accessions. Subsequent columns contain the marker data.

	A	B	C	D	E	F
1	**DatasetID**	**Genotype/Marker**	**umc83a**	**bnl6.29b**	**umc53a**	**UMC80b**
2	2	Xisep0101	A	B	B	A
3	2	Xisep0107	A	B	B	A
4	2	Xisep0108	A	B	B	B
5	2	Xisep0114	A	B	B	A

QTL File

Prepare the QTL file as a .txt file with the following column headings:

- Name: QTLID
- Chromosome: Chromosome ID
- Position: QTL position with highest LOD score (cM)
- Pos-min: Start position of the QTL (cM)
- Pos-max: End position of the QTL
- Trait: Trait ID
- Experiment: Environment ID
- CLEN: Chromosome length (cM)
- LFM: Left flanking QTLmarker
- PLFM: Position of left flanking QTLmarker
- RFM: Right flanking QTLmarker
- PRFM: Position of right flanking QTLmarker
- Effect
- LOD
- R2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	**Name**	**Chromosome**	**Position**	**Pos-min**	**Pos-max**	**Trait**	**Experiment**	**CLEN**	**LFM**	**PLFM**	**RFM**	**PFRM**	**Effect**	**LOD**	**R2**
2	QTL1	Chrom1	22	20	26	FT	1	38.5	MK171	9.4	MK172	26.4	0.578	5.76	30.1
3	QTL2	Chrom1	5	2	14	FT	2	38.5	MK171	9.4	MK172	26.4	0.578	5.76	30.1
4	QTL3	Chrom1	25	24	28	BDW	1	57.3	MK830	48.5	MK186	50.2	-0.7	7	37.3

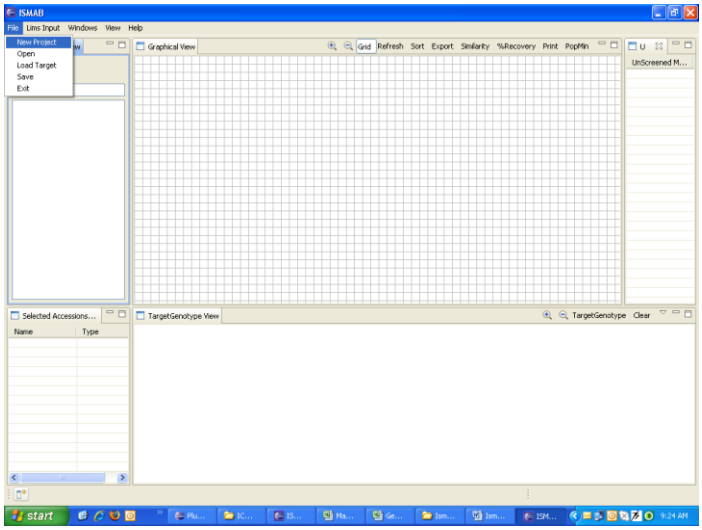
Phenotype File

Prepare the phenotype file as a .txt file with the first column heading, Genotype, identifying lines or accessions, followed by columns of trait measurements.

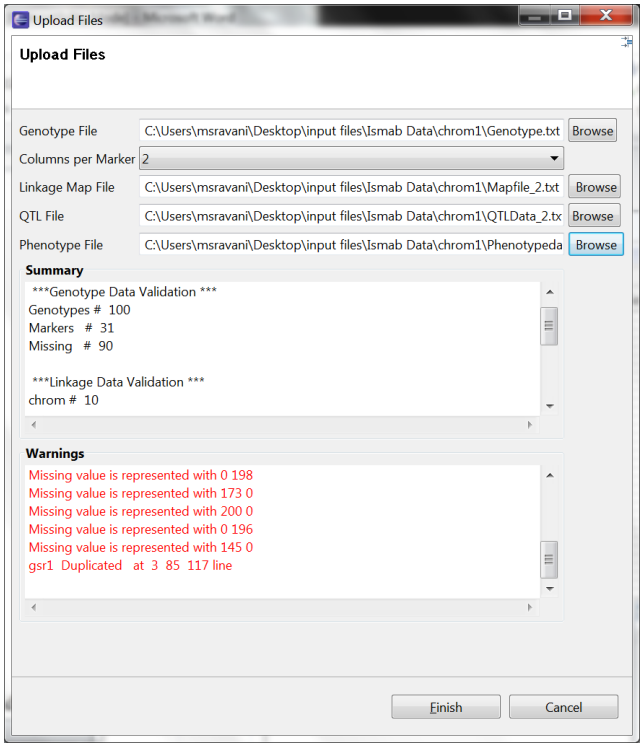
	A	B	C
1	**Genotype**	**yield**	**heading**
2	IS8603	7.232	200
3	IS8607	6.326	193
4	IS8608	6.533	197
5	IS8610	7.469	200.5

Create New Project

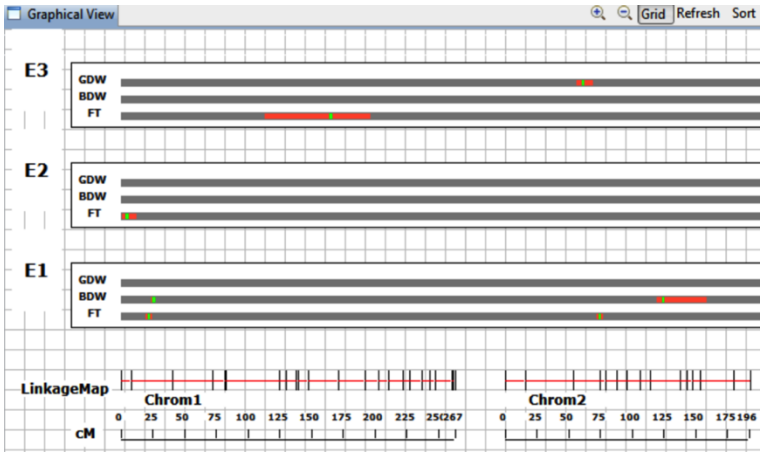
Select New Project action from the File menu to load data.



Browse the input files: genotype, linkage map, QTL, phenotype. Genotype and linkage map file are mandatory. Once you browse the files, the system validates the input files and prompts the summary, which includes marker type, number of genotype, markers, traits, and QTLs. The warning log will display information about missing data, duplicated markers and other relevant warnings.



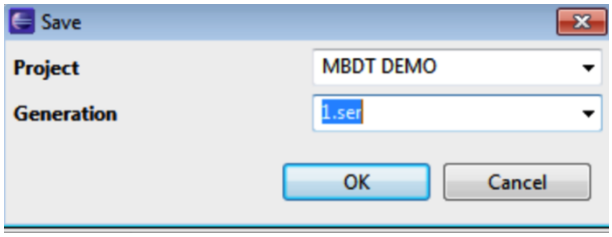
Click Finish to upload files to the workbench and observe that Graphical View overlays QTL onto the Linkage Map.



E1, E2, E3 indicates that phenotypic data was collected from three environments.

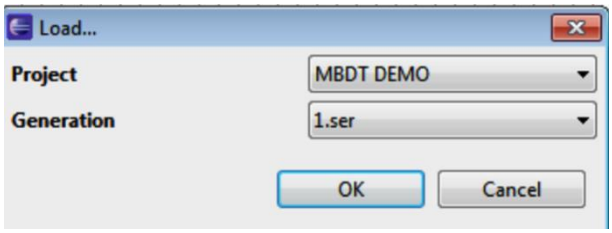
Save Project

Select File>Save. Enter and project name and generation. Click OK. The project will be saved as a .ser file within a MBDT_PROJECTS files folder.



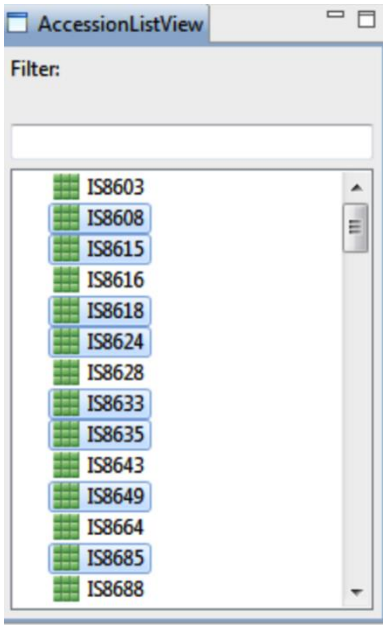
Open Project

Select File>Open. Select the project name and generation. Click OK to load the project.

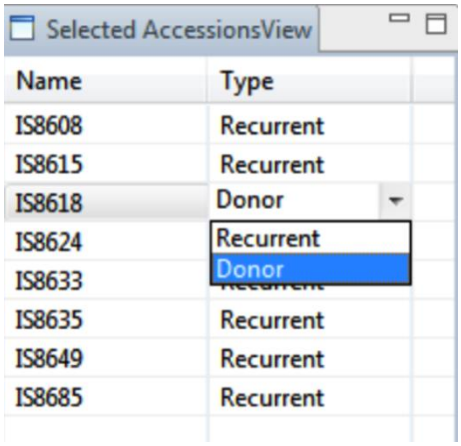


Accession List & Selected Accession View

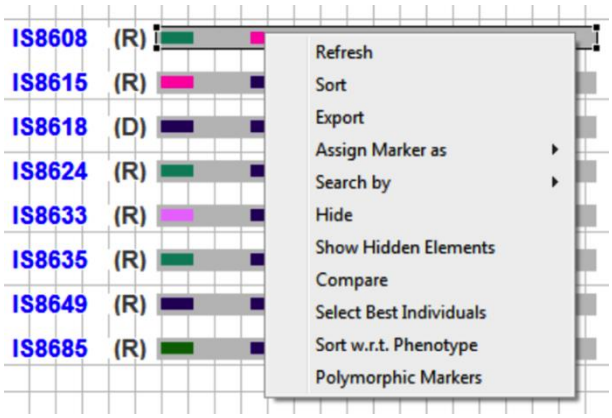
Select accessions of interest as probable donor or recurrent parents. Use Ctrl/ Shift Key for multiple selections.



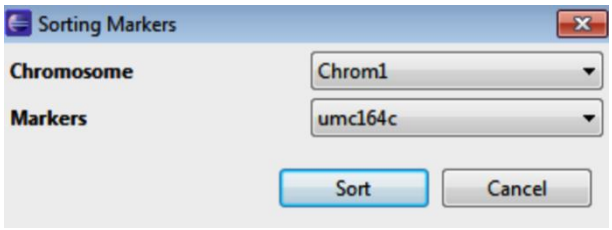
These selections will now appear in the Selected Accessions View. Selections are by default set to Recurrent. Change the default by choosing Donor from the drop down menu.



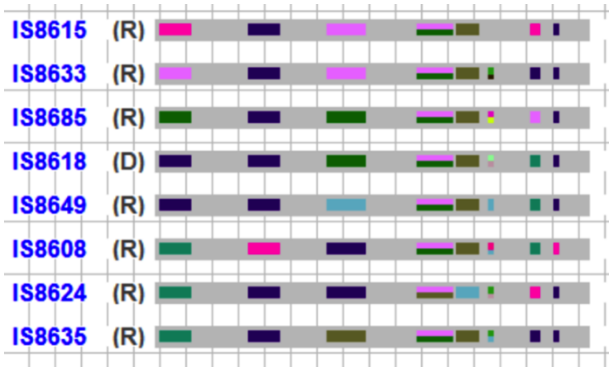
Sort individuals based on the frequency of alleles for the selected marker. Select marker of interest and click on, or right click on Graphical view. Select Sort.



Alternatively, go to the Windows menu option and select Sorting Markers. Choose chromosome and marker, and click Sort.



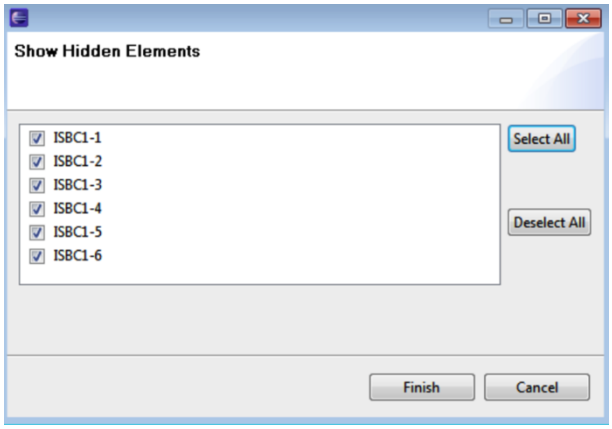
Sorting displays accession list with rare alleles at top and more common alleles at bottom.



Selected accessions sorted by first marker, umc164c, on chromosome 1

Hide Accessions

Hide the heat maps of selected accessions in Graphical View. Select the appropriate heat maps. Right click on selections and click Hide.

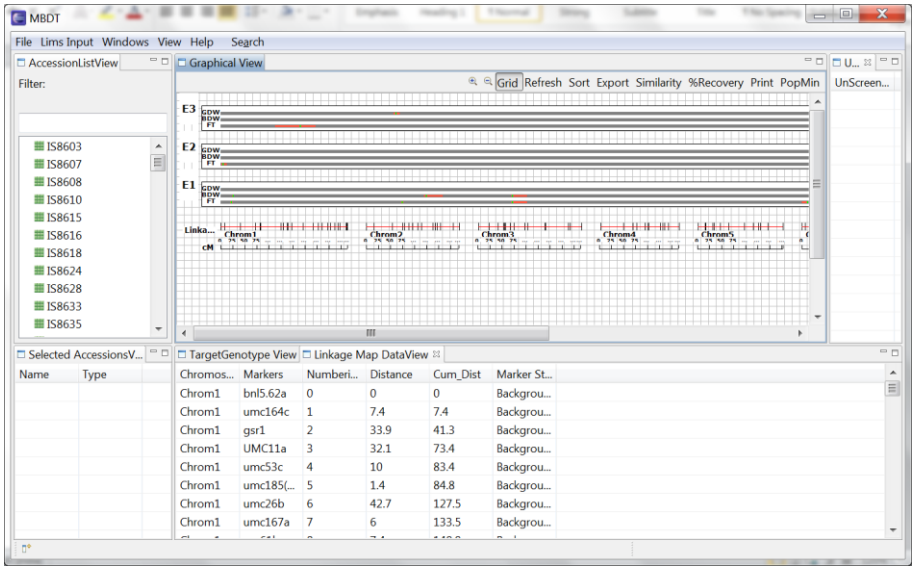


Reveal Hidden Accessions

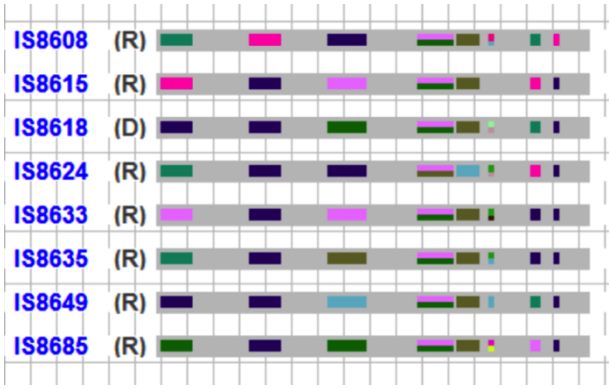
Right click on selection and click Show Hidden Elements. Choose the hidden accessions that you would like to reveal and click Finish.

Graphical View

Select Refresh to update changes made in the Selected Accession View. Heat maps of selected accessions display genotype information as coloured boxes allowing comparison of genotypes with linkage and QTL map.

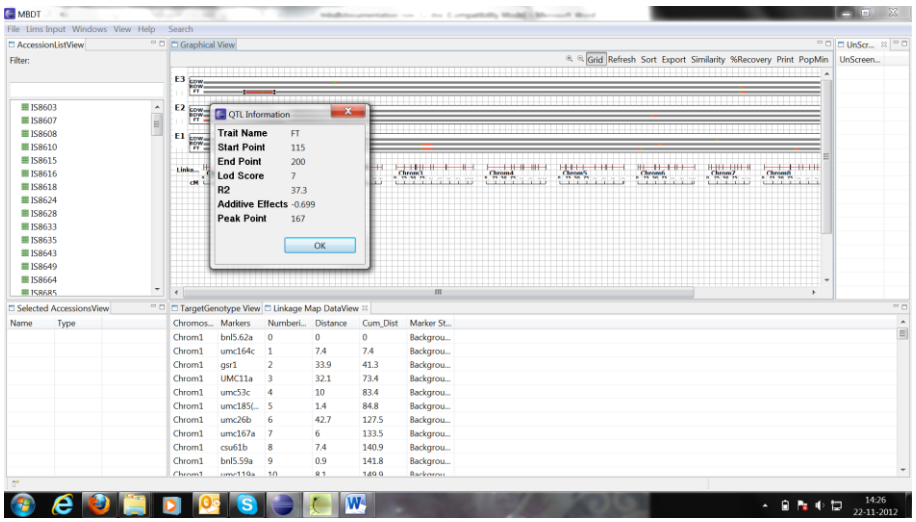


The accessions are now visible from the graphical view.



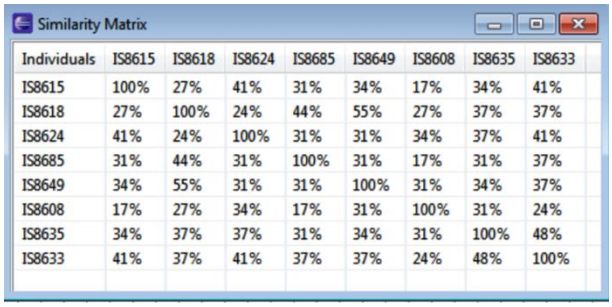
R and D represent recurrent and donor parent assignment. Colored boxed on the grey bar represent the screened marker for which genotype data is available. The markers are coloured based on the allele values. The missing values are coloured in background colour, light grey to represent un-genotyped sections of the chromosome.

Click on QTL region to view QTL information in pop-up window.



Similarity Matrix

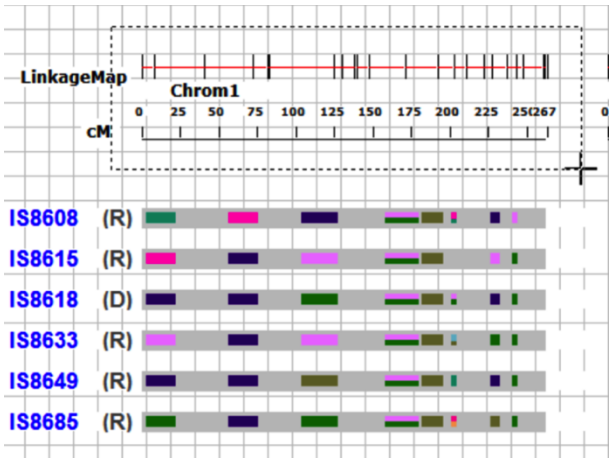
Select Similarity to generate a similarity matrix to visualize the percentage of similarity between selected parents based on the number of shared alleles.



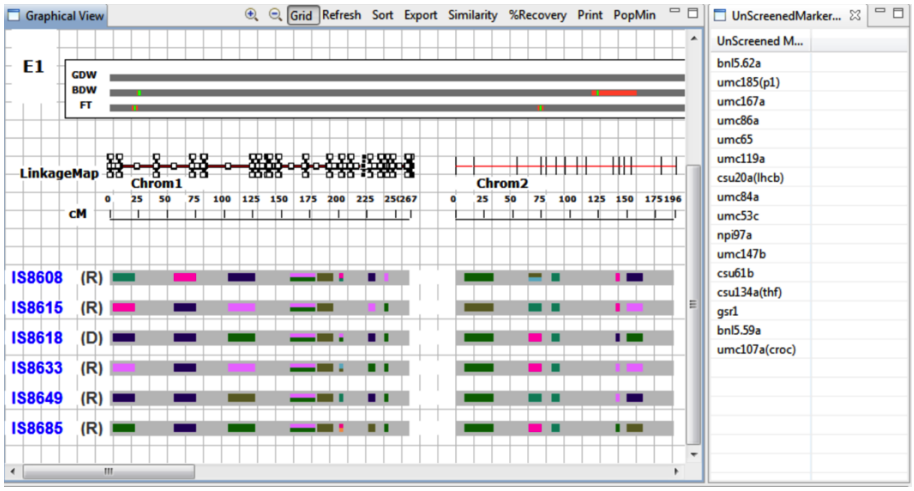
Individuals	IS8615	IS8618	IS8624	IS8685	IS8649	IS8608	IS8635	IS8633
IS8615	100%	27%	41%	31%	34%	17%	34%	41%
IS8618	27%	100%	24%	44%	55%	27%	37%	37%
IS8624	41%	24%	100%	31%	31%	34%	37%	41%
IS8685	31%	44%	31%	100%	31%	17%	31%	37%
IS8649	34%	55%	31%	31%	100%	31%	34%	37%
IS8608	17%	27%	34%	17%	31%	100%	31%	24%
IS8635	34%	37%	37%	31%	34%	31%	100%	48%
IS8633	41%	37%	41%	37%	37%	24%	48%	100%

Unscreened Markers

Select an area of linkage map to populate the Unscreened Marker view.

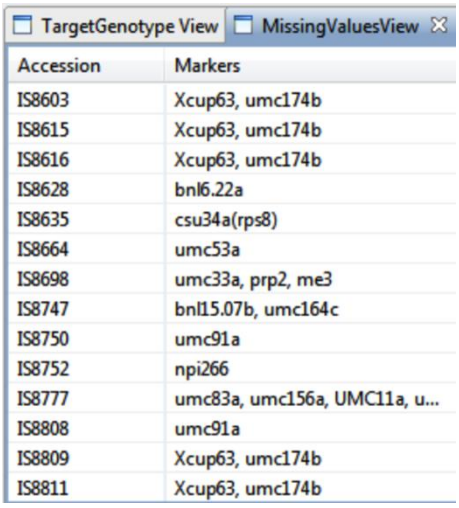


The unscreened markers display as a list.



Missing Markers

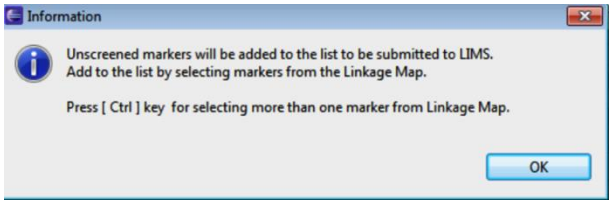
View the missing markers for each accession.



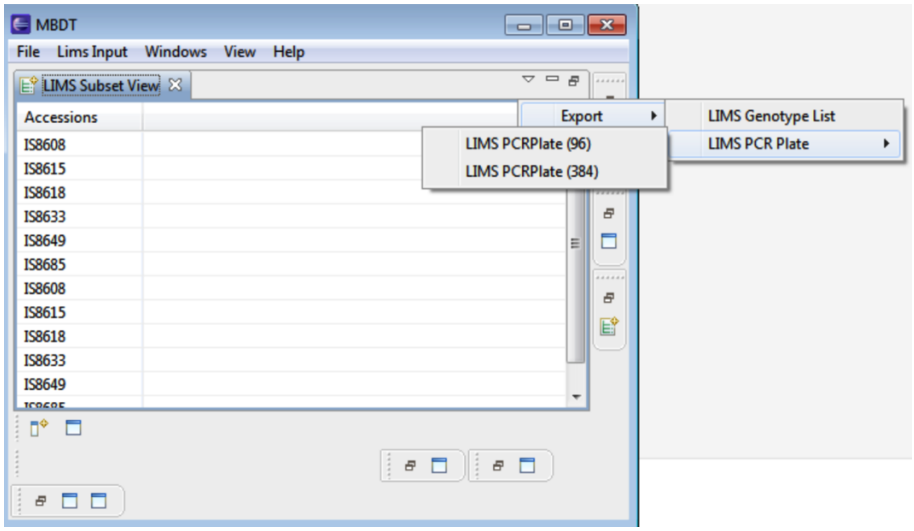
Accession	Markers
IS8603	Xcup63, umc174b
IS8615	Xcup63, umc174b
IS8616	Xcup63, umc174b
IS8628	bnl6.22a
IS8635	csu34a(rps8)
IS8664	umc53a
IS8698	umc33a, prp2, me3
IS8747	bnl15.07b, umc164c
IS8750	umc91a
IS8752	npi266
IS8777	umc83a, umc156a, UMC11a, u...
IS8808	umc91a
IS8809	Xcup63, umc174b
IS8811	Xcup63, umc174b

Lims Markers

Identify genotyped markers that match markers within the linkage map. Select Lims Input menu > Lims Markers. Select OK.

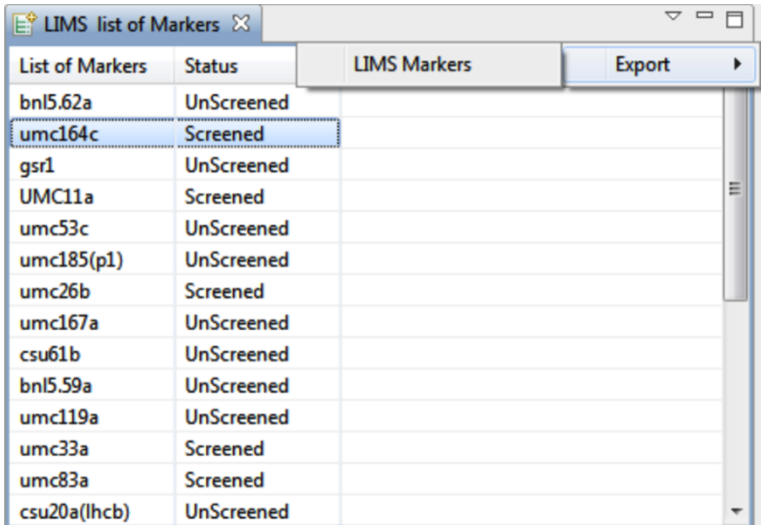


Delete markers from the list by right clicking selections and choosing Delete Selected Marker. Select Export > Lims Markers to save this list as a .txt file.



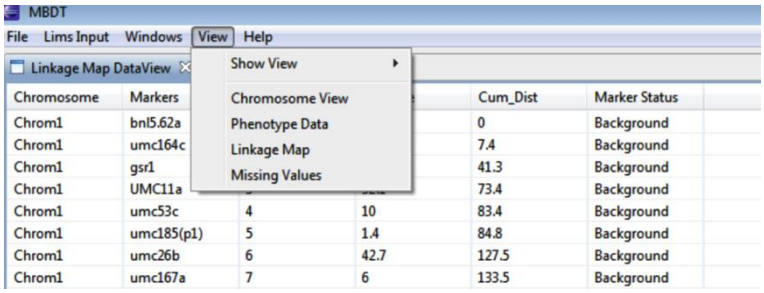
Lims Subset

Select Lims Input > Lims Accessions. The Lims Subset View allows export of the genotype list as well as a file outlining 96 or 384 well PCR plates.



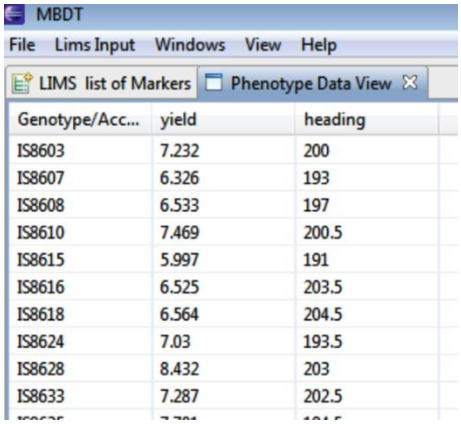
Linkage Map Information

View information about the linkage map by selecting this view.



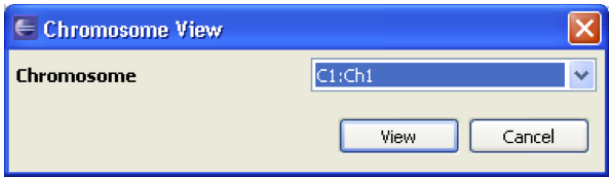
View Phenotypic Data

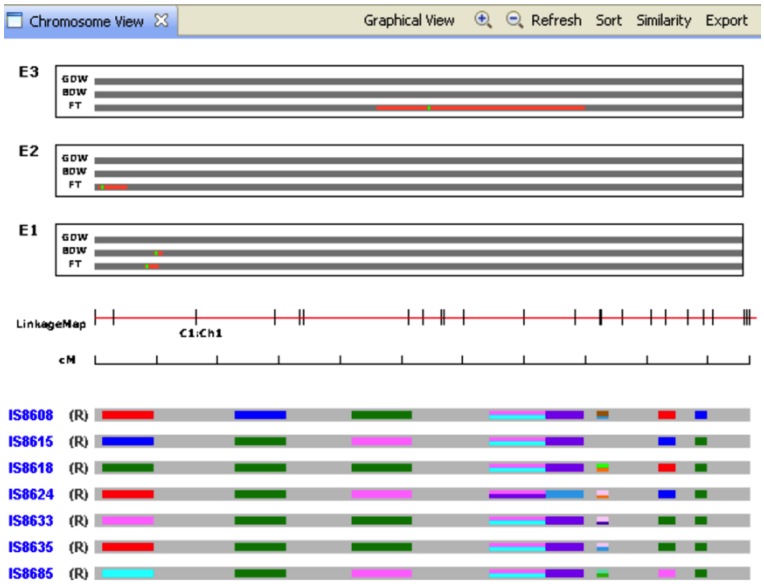
Sort phenotypic data by double clicking the trait name.



Chromosome View

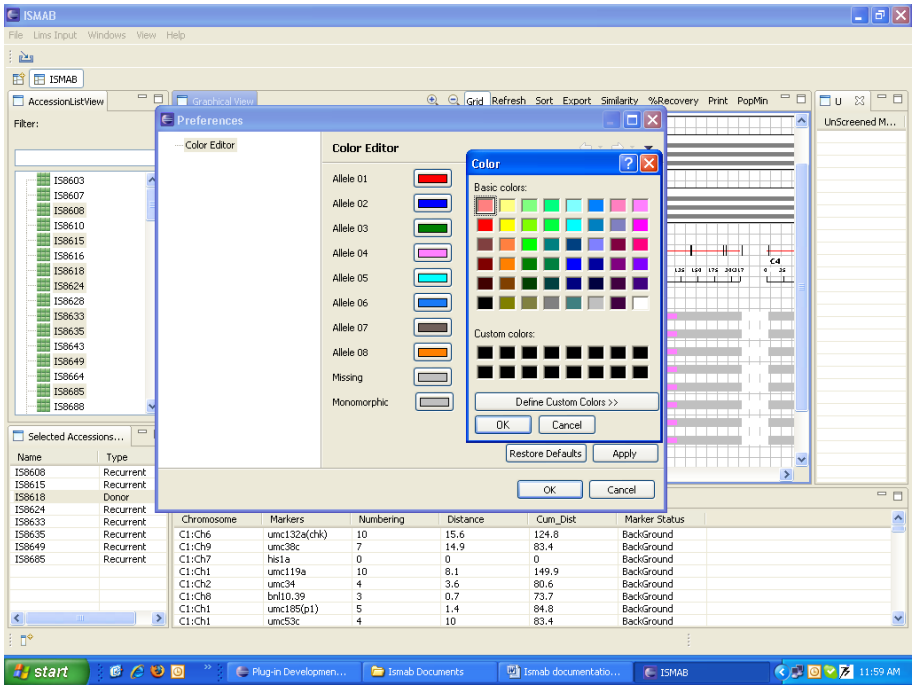
Limit graphical view to a single chromosome by selecting Chromosome View and choosing a chromosome of interest.





Customize Marker Colors

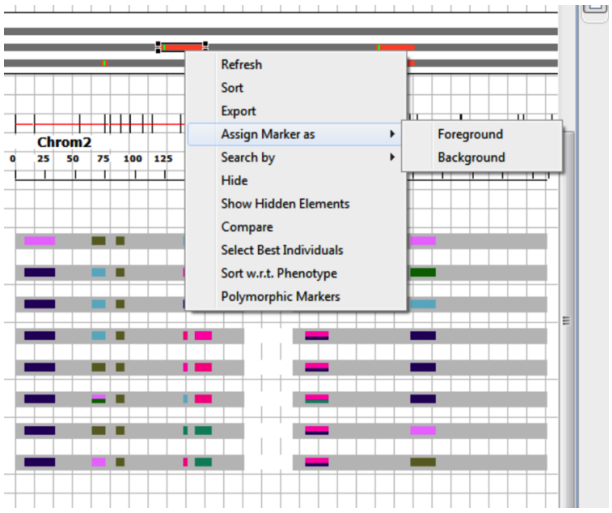
Change the default colors by select Preferences under the Windows menu. Change the color of any allele using the color editor.



Assign Foreground & Background Markers

Foreground Markers are the markers linked to QTL from the donor parent. Background markers are unique to the recurrent parent and are used to track the recovery of the recurrent parent genome. By default all markers are defined as background markers.

Assign markers as Foreground markers by selecting a QTL on the QTL map or a portion of the linkage map. Right click and choose Assign Marker as > Foreground.



Assigning a foreground marker will result in a broadening the representative rectangular boxes. The status of marker in the LinkageMap data view will be changed accordingly.

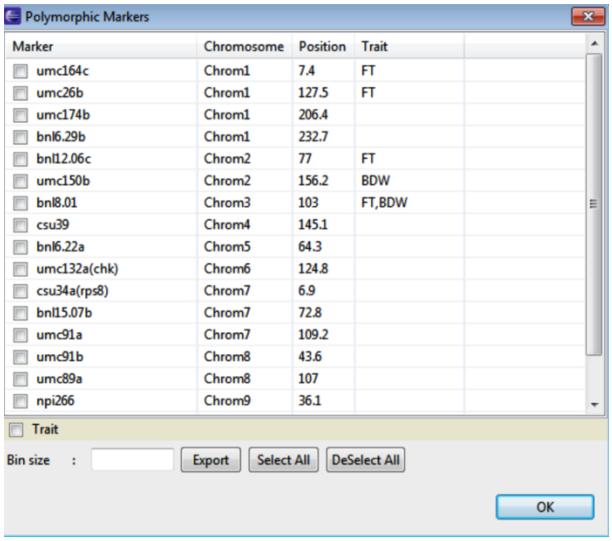
Revert foreground markers to background markers by selecting the QTL or a portion of the linkage map. Right click and choose Assign Marker as > Background. The status of marker in the LinkageMap data view will change accordingly.

Find Polymorphic Markers

Compare potential parents based on their polymorphic markers by selecting two accessions and right clicking to selecting Compare from the drop down menu.

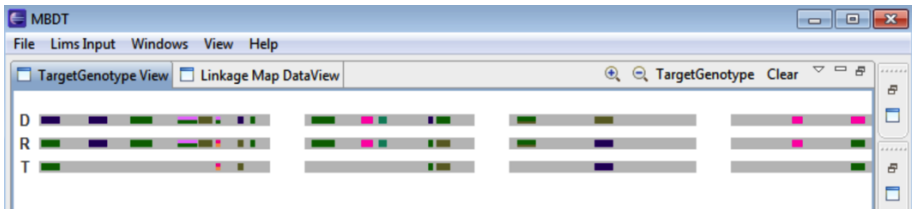
The graphical field will narrow to the two selected accessions.

Select both accessions. Right click to choose polymorphic markers from a list.



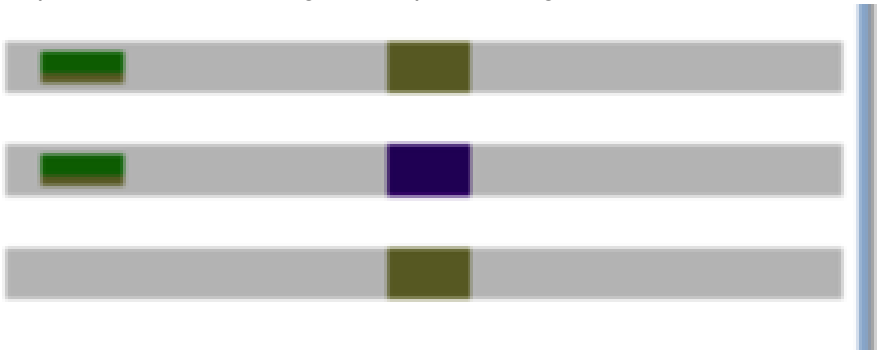
Create Target Genotype

Select the donor and recurrent parents in graphical view. Select the Target Genotype button from Target Genotype View. The target genotype illustrates polymorphic markers between donor and recurrent parent. Markers in the target genotype represent the recurrent parent.



D donor parent (IS86818) and R recurrent parent (IS8685). The T target represents the markers that differ between the parents

Replace recurrent parent marker alleles with the desired forward markers from the donor parent. Select the donor parent marker and drag and drop to the target marker location.

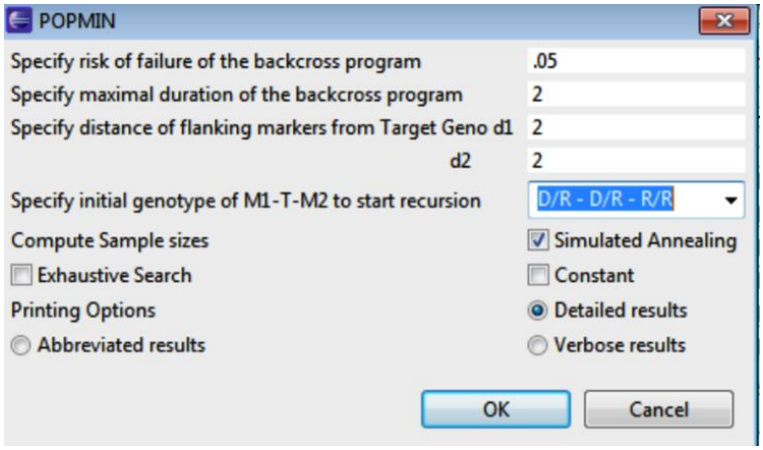


Chromosome 3: Donor parent marker (top) dropped on the target genotype (bottom)

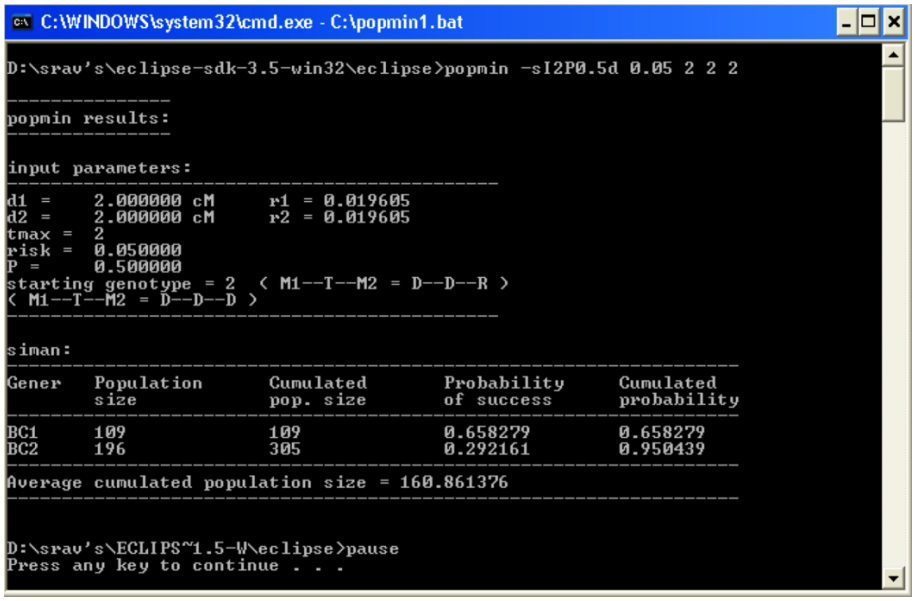
Save the movements of markers while preparing Target Genotype by clicking on drop down arrow and choosing Action and then History of movements. A click on History of movements will open a Save As File Dialog to save the list of moved markers into text file with *.txt file.

Determine Population Size for Next Generation

Determine the minimum number of individuals that must be genotyped in successive backcross generations. The PopMIN tool estimates the population size required to obtain at least one double homozygote for the two markers flanking target QTL. Select PopMIN, enter the appropriate values and click OK.

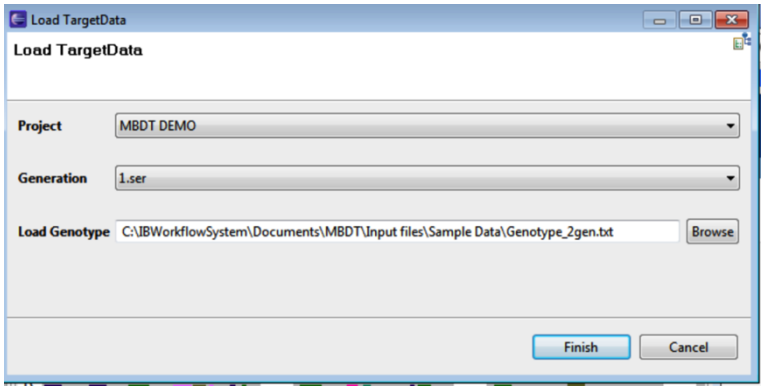


PopMIN results are displayed in the terminal.



Advanced Generations

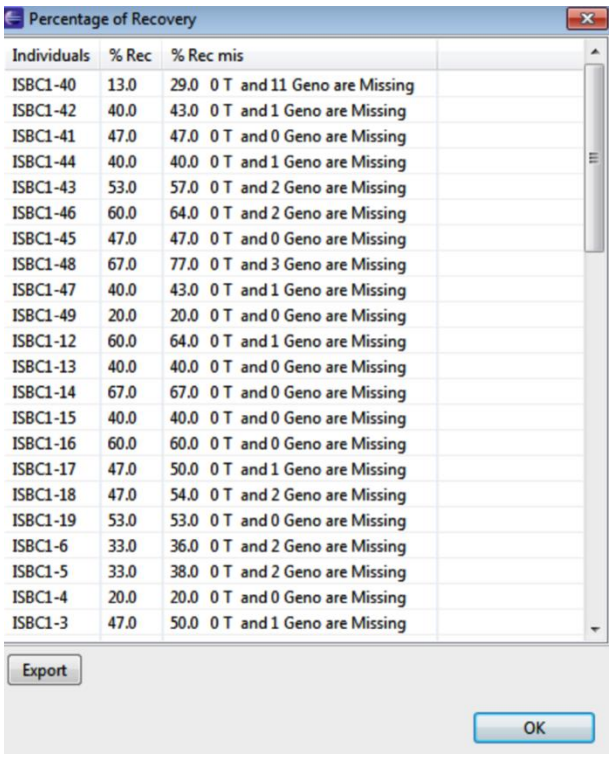
As subsequent generations are genotyped they can also be viewed with Molecular Breeding Design Tool. Select File > Load Target Data. Select the project and the previous generation from the drop down menu. Load the current generations genotype data, and click Finish.



The second generations genotype is graphically displayed beneath the previously established target genotype.

Percentage Recovery

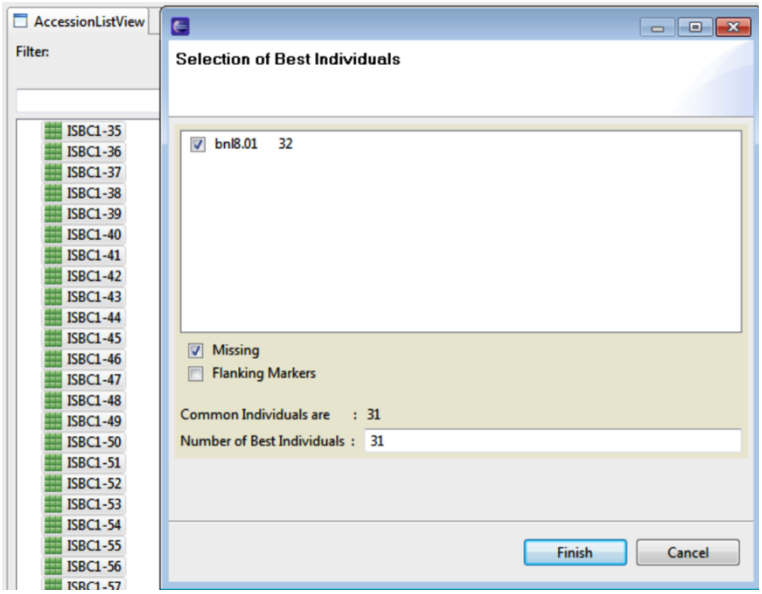
Select Percentage of Recovery to view the proportion of recurrent parent markers in the second generation accessions.



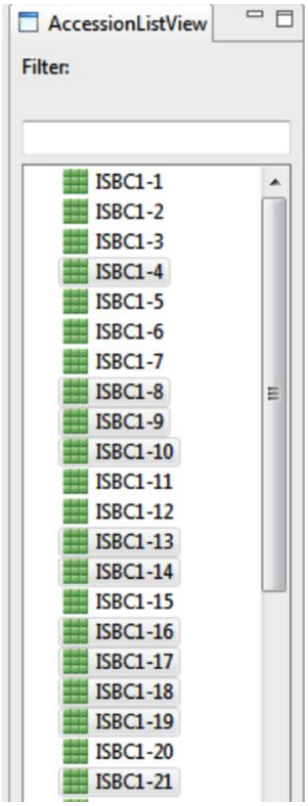
Individuals	% Rec	% Rec mis
ISBC1-40	13.0	29.0 0 T and 11 Geno are Missing
ISBC1-42	40.0	43.0 0 T and 1 Geno are Missing
ISBC1-41	47.0	47.0 0 T and 0 Geno are Missing
ISBC1-44	40.0	40.0 0 T and 1 Geno are Missing
ISBC1-43	53.0	57.0 0 T and 2 Geno are Missing
ISBC1-46	60.0	64.0 0 T and 2 Geno are Missing
ISBC1-45	47.0	47.0 0 T and 0 Geno are Missing
ISBC1-48	67.0	77.0 0 T and 3 Geno are Missing
ISBC1-47	40.0	43.0 0 T and 1 Geno are Missing
ISBC1-49	20.0	20.0 0 T and 0 Geno are Missing
ISBC1-12	60.0	64.0 0 T and 1 Geno are Missing
ISBC1-13	40.0	40.0 0 T and 0 Geno are Missing
ISBC1-14	67.0	67.0 0 T and 0 Geno are Missing
ISBC1-15	40.0	40.0 0 T and 0 Geno are Missing
ISBC1-16	60.0	60.0 0 T and 0 Geno are Missing
ISBC1-17	47.0	50.0 0 T and 1 Geno are Missing
ISBC1-18	47.0	54.0 0 T and 2 Geno are Missing
ISBC1-19	53.0	53.0 0 T and 0 Geno are Missing
ISBC1-6	33.0	36.0 0 T and 2 Geno are Missing
ISBC1-5	33.0	38.0 0 T and 2 Geno are Missing
ISBC1-4	20.0	20.0 0 T and 0 Geno are Missing
ISBC1-3	47.0	50.0 0 T and 1 Geno are Missing

Select Best Individuals

Select Accessions of interest and right click to select the Best Individuals. Choose the foreground marker(s) to base the selection. Include missing information by selecting Missing. Click Finish.



The best individuals are highlighted in the Access List View.



OptiMAS: Maize Tutorial

Contributors

Fabio Valente [a], Johann Joets [a], Alain Charcosset [a], Laurence Moreau [a], and Shawn Yarnes [b]

Institut National de la Recherche Agronomique (INRA) [a], The Integrated Breeding Platform [b]

Summary

Use OptiMAS to select and cross the best of 297 genotyped individuals for the next round of marker-assisted recurrent selection (MARS). Calculate the predicted genetic value of genotyped individuals based on 11 target QTL. Design crosses based on the likelihood of combining all favorable alleles into a single individual.

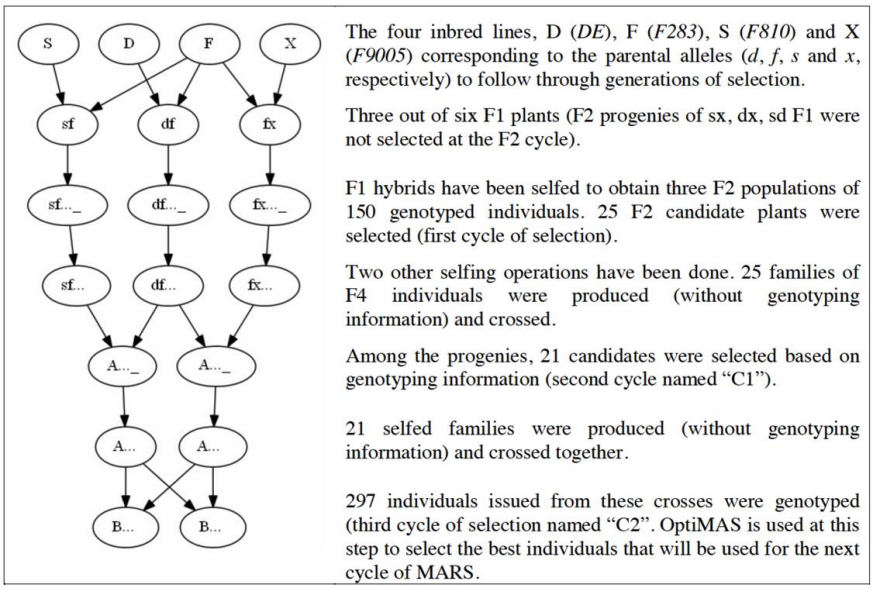
Introduction

OptiMAS (Blanc et al. 2013) predicts crossing strategies that optimize the likelihood of assembling favourable alleles into a target genotype. Molecular markers in the vicinity of favourable parental QTL are used to trace the movement of QTL between generations. OptiMAS uses genotype data to predict the probability of allele transmission in different marker-assisted selection (MAS) schemes and mating designs (intercrossing, selfing, backcrossing, double haploids, RIL), allowing some generations to be considered without the need to genotype. Selection and crossing strategies are based on expected progeny genotypes. OptiMAS supports decision-making associated with the marker-assisted breeding plan generated by the Molecular Breeding Planner.

OptiMAS is a Breeding Management System standalone tool that can be launched from the workbench or from a desktop icon. OptiMAS has a graphical user interface, but can also operate from command line. For more information on command line operation and additional demonstration examples see the OptiMAS development site or read the full manual (pdf).

Demonstration Data

The maize (*Zea mays*) data set comes from a mulitparental population described by Blanc et al. 2006 and 2008. Eleven QTL were detected for silking date based on six connected F2 populations, based on 150 individuals each. The F2 populations were obtained from a half-diallel design between four unrelated maize inbred lines (DE, F283, F810 and F9005). A set of 34 markers was selected with at least three microsatellite markers following each QTL. Two cycles of marker-assisted recurrent selection (MARS) were performed with a step of selfing occuring before each intermating. In this example, OptiMAS is used at the last cycle to select the best individuals among 397 genotyped plants for the next cycle of MARS.



MARS Breeding Scheme (Reprinted with permission from Blanc et al. 2006, 2008)

Genetic Map File (.map)

The map file is supplied by the user and specifies QTL of interest, flanking makers, and favourable alleles. QTL identified by the Breeding Management Systems Breeding View application, and other external applications, can be used to create the genetic map file for OptiMAS.

BMS File Directory

C:Breeding Managment System\Tools\OptiMAS\input\blanc.map

	A	B	C	D	E
1	**Locus**	**Chr**	**QTL**	**Pos**	**All+**
2	Pr443	1	1	0	
3	qtl1	1	1	23.3	x
4	Pr256	1	1	28.5	
5	Pr299	1	2	0	
6	qtl2	1	2	8	f
7	Pr454	1	2	14.5	
8	Pr1204	1	2	36.3	
9	Pr1211	2	3	0	
10	Pr96	2	3	4.5	
11	qtl3	2	3	10.5	d
12	Pr1165	2	3	12.7	
13	Pr309	2	3	23.3	
14	Pr796	3	4	0	
15	Pr311	3	4	11	
16	qtl4	3	4	25	s/f
17	Pr506	3	4	38.8	
18	Pr534	3	4	41.2	
19	Pr510	3	5	0	
20	qtl5	3	5	6	s
21	Pr1265	3	5	14.7	

5 of the 11 QTL and flanking markers described by their map position (pos) in cM

Map File Columns

- Locus: Mame of markers and QTL. QTL name are prefaced with "qtl".
- Chr: Chromosome
- QTL: QTL numbered in ascending order
- Pos: (cM) For QTL, this is the estimated QTL position coming from the QTL detection results. Positions of the different loci must be obtained using the Haldanes mapping function (i.e. by assuming no interference).
- All+: Parental allele considered favourable for the QTL

Genotype/Pedigrees File (.dat)

The genotype/pedigree file is in plain-text tab-delimited format (no space between fields). Column headings A-F should not be changed even if the 2 optional columns, Cycle and Group, are left blank (or -). Column headings begining with G describe the 34 microsatalite markers. This format is very close to the input format of the Flapjack software.

BMS File Directory
C:Breeding Managment System\Tools\OptiMAS\input\blanc.dat

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
1	Id	P1	P2	Step	Cycle	Group	Pr4	Pr2	Pr2	Pr4	Pr1	Pr1	Pr9	Pr1	Pr3	Pr7	Pr3	Pr5	Pr5	Pr5	Pr1	Pr5	Pr3	Pr5	Pr1	Pr8	Pr1	Pr5	Pr5	Pr1	Pr1	Pr8	Pr5	Pr3	Pr6	Pr1	Pr6	Pr9	Pr9	Pr93
2	D	d	d	IL			1	3	3	2	2	1	1	3	2	1	1	2	3	3	2	2	2	2	2	3	2	3	3	2	1	2	1	2	3	1	3	1	2	2
3	F	f	f	IL			1	3	2	1	3	3	1	1	3	2	2	1	3	3	1	1	3	4	1	2	1	2	1	3	1	1	2	4	1	1	2	3	2	1
4	S	s	s	IL			1	1	1	1	1	2	3	2	1	3	1	1	2	2	4	2	1	3	3	4	1	1	3	1	1	2	2	3	2	2	2	1	2	3
5	X	x	x	IL			2	2	2	2	2	1	2	2	3	1	2	3	1	1	3	1	1	1	2	1	1	3	2	3	2	1	2	1	2	1	1	2	1	3
6	df	D	F	CR																																				
7	fx	F	X	CR																																				
8	sf	S	F	CR																																				
9	df19_	df	df	S1	F2		1	3	2	1	3	1	1	1	2	1/2	1/2	1	3	3	1	1/2	2/3	2	1/2	2/3	2	3/-	1/3	2/3	1	1/2	1/2	2/4	1	1	2/3	1/3	2	1/2
10	df21_	df	df	S1	F2		1	3	2/3	1/2	2/-	3/-	1	3	2/3	1/2	1/2	1/2	3	3	2	2	2/-	2	1/2	2/3	1/2	3/-	1/3	2/3	1	1	2	2/4	1/3	1	2	1/3	2	1/2
11	df24_	df	df	S1	F2		1	3	2	1	3	1	1	3	2	2	2	1	3	3	1/2	1/2	2/3	2/4	1/2	2/3	1	3/-	1/3	2/3	1	1/2	1/2	2/4	1/3	1	2/3	3	2	1
12	df37_	df	df	S1	F2		1	3	2/3	1/2	2/-	1	1	1/3	2/3	1/2	1/2	1/2	3	3	1/2	1/2	2/3	2/4	1	2/3	1	3/-	3	2	1	1/2	2	4	3	1	3	1/3	2	1/2
13	df58_	df	df	S1	F2		1	3	2	1	2/-	3/-	1	3	3	1/2	1/2	1/2	3	3	1	2	2/-	2	1	2	1/2	3/-	1/3	2/3	1	1/2	2	4	1/3	1	2/3	1/3	2	2
14	df59_	df	df	S1	F2		1	3	2/3	1/2	2/-	3/-	1	1/-	3	1/2	1/2	1	3	3	1/2	1	2/3	2/4	1	2	1/2	3/-	3	2	1	1	2	4	3	1	3	1/3	2	1
15	df64_	df	df	S1	F2		1	3	2/3	1/2	2/-	3/-	1	1	2	1/2	1/2	1/-	3	3	1/2	1	3/-	4	1/2	2/3	1/2	3/-	3	2/3	1	1	2	4	1/3	1	2/3	1	2	1/2

15 of the 279 genotyped lines: The first four rows are inbreed parental lines (IL) homozygous for all 34 markers. Rows 6-8 are heterozygous F1s, and are not genotyped. Rows 9-15 are genotyped F2 individuals. Markers are numerically coded; A=1, C=2, G=3, T=4.

Genotype/pedigree file column descriptions

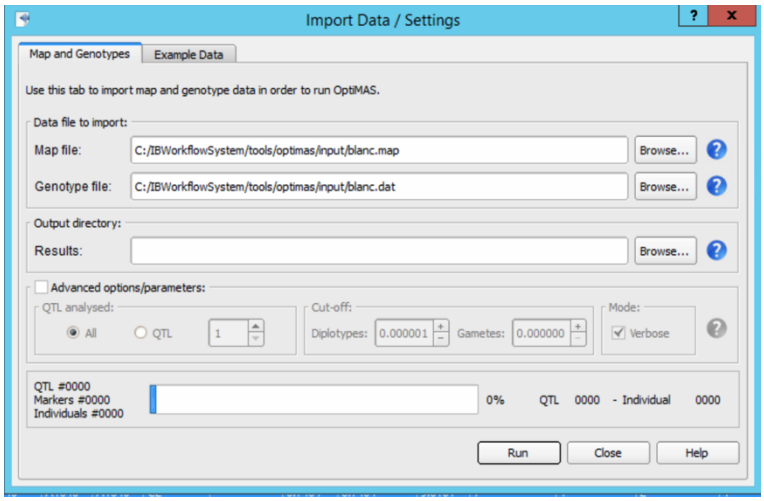
- Id: The name of each genotyped individual. Individual IDs must unique and ranked according to generations (from oldest to most recent). The first individuals of this file must correspond to the founder parents of the program (here D,F,S, & X) that are assumed to be homozygous lines.
- P1 & P2: Correspond to the name(s) of the parent(s) of the individual (must exist as individuals in above in the file). The pedigree of the parental lines is assumed to be unknown. For a founder line P1 and P2 columns indicate the name given to the allele (e.g. A, B or any other character) coming from this line that will be traced through generations.
- Step: Corresponds to the pedigree relationship between the individual and its parent(s):
- CR: Cross indicates that the individual results from a cross between its two parents.
- Sn: Selfing indicates that the individual results from n generations of selfing of its parent (in this case the two parent Ids must be identical).
- RIL: Recombinant inbred lines, assumes that the individual results from an infinite number of selfing generations from an initial F1 hybrid. In this case, parent 1 and parent 2 must be identical and correspond to the F1 hybrid.
- DH: Double haploids assumes that the individual results from haplo-diploidisation from an initial F1 hybrid
- IL: Inbreed line indicates founder inbred status
- Cycle: Optional information regarding the generation in the program (e.g. F2, F4, C1, etc.).
- Group: Optional information regarding another classification criterion (e.g. subprograms, families, etc.).

- Mk1-Mkn: These are the genotyping results. The software can deal with SNPs, microsatellites, and any bi/multi allelic marker genotyping technique with either dominant or codominant scoring. The markers present in the genotypes/pedigree file must be ordered and match those in the map file (same number of markers). Homozygous genotypes for an allele (e.g. A) can be scored either as A or A/A. Heterozygotes are expected to be separated by a / (e.g. A/G). Heterozygous genotypes are assumed to be unphased (i.e. A/B equivalent to B/A). Missing data at marker loci are allowed and must be entered as - (or can be left blank and corrected later see below). For dominant markers, assuming A dominant vs. a recessive, genotypes presenting allele A must be coded A/-. Parental inbred lines should not contain missing data. Given the possibility to include non-genotyped individuals, this makes it possible to analyse most common marker-assisted selection schemes and mating designs.

Run OptiMAS

Launch OptiMAS from the workbench or from the OptiMAS executable file located in the Tools folder.

Select File > Import Data from the menu bar and browse for the genetic map file (.map) and genotypes/pedigree file (.dat) Set the location of the output directory. Results from each run will be stored within a new dated directory created automatically within this folder. Note that your output directory should not be in the Program Files folder or other specific directories with administrator privileges. Select Run button to analyse the data set. Close the Import Data window when the progress bar displays 100%.

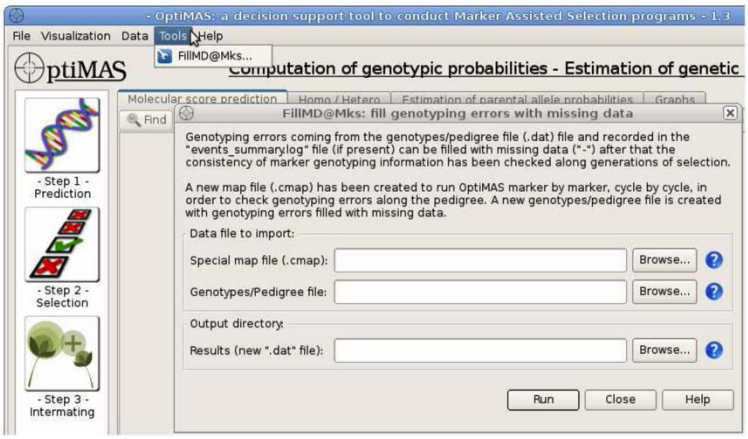


Advanced options/parameters

- QTL Analysed: By default all the QTL present in the input files will be analysed. You can also choose to select a specific QTL to run the analysis.
- Cut-off -Diplotypes: Genotypic probability below which a rare phased genotype (diplotype) is removed and not considered in subsequent computations (default value = 0.000001), to reduce memory and time needed for computation
- Cut-off-Gametes: Gametic probability (default value = 0.000000) that corresponds to the probability that the number of crossovers expected in the region between flanking markers exceeds a given value. Thus, unlikely haplotypes, or gametes, with number of crossovers over this value are removed and not considered in subsequent computations. Use of this option with values up to 0.01 is recommended in case many flanking markers per QTL lead to high computation time with default option.
- Verbose: Verbose mode creates two files per QTL position, reporting respectively gamete and diplotype probabilities for all individuals (default value = ON).

Correct Missing Genotype Data

A .cmap file has been created for import into the FillMD@Mks tool, which will fill in missing data with a placeholder and reconfigure the genotype/pedigree (.dat) file for reanalysis. A genotype/pedigree file may contain missing data due to genotyping errors. If a warning message appears after running OptiMAS, missing data may need to be replaced with a placeholder (-). Load input (.cmap and .dat) and set output directory, the path to the folder where the results will be stored. The .cmap file created at the end of a previous run of OptiMAS is used to re-run OptiMAS marker by marker and localize genotyping errors at individual marker position. The (.dat) file is the same genotype pedigree file used to run OptiMAS the first time. The output directory sets the destination of a new genotype/pedigree file (.dat) missing data is filled with a placeholder (-).

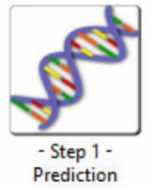


Note: It is also possible to directly display the results of previous analyses by selecting File > Reload data. You can also display results from the two examples data sets provided with the program, that are located in File > Example Data > Multiparental from the menu bar.

Prediction

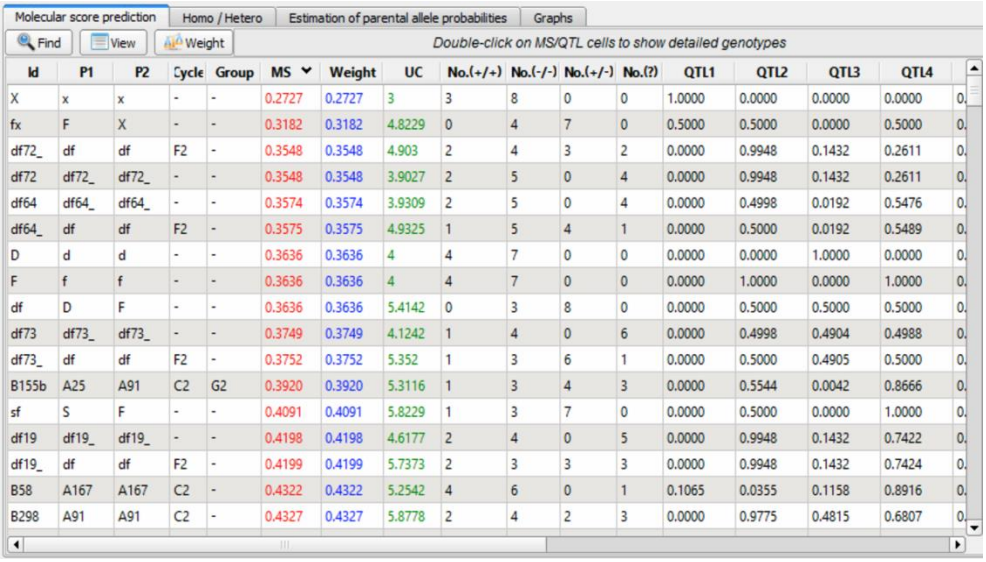
Each QTL is given a score based on its probability of transmitting favourable alleles to gametic genotypes, or haplotypes. Genotyping data is phase ambiguous when markers are heterozygous, so OptiMAS calculates all possible phased diplotypes along with their probabilities, and then determines all of the possible gametes (haplotypes) and their probabilities at each QTL. OptiMAS summarizes the predicted genetic value of genotyped individuals into several metrics that inform selection.

Access Prediction Functions by clicking Step 1 from the left menu.



Molecular Score Prediction Table

Genetic values and QTL scores are summarized in the Molecular Score (MS) Prediction window. Data can be sorted by MS or other criterion by double clicking the column heading.



Id	P1	P2	Cycle	Group	MS	Weight	UC	No.(+/+)	No.(-/-)	No.(+/-)	No.(?)	QTL1	QTL2	QTL3	QTL4	
X	x	x	-	-	0.2727	0.2727	3	3	8	0	0	1.0000	0.0000	0.0000	0.0000	0.
fx	F	X	-	-	0.3182	0.3182	4.8229	0	4	7	0	0.5000	0.5000	0.0000	0.5000	0.
df72_	df	df	F2	-	0.3548	0.3548	4.903	2	4	3	2	0.0000	0.9948	0.1432	0.2611	0.
df72	df72_	df72_	-	-	0.3548	0.3548	3.9027	2	5	0	4	0.0000	0.9948	0.1432	0.2611	0.
df64	df64_	df64_	-	-	0.3574	0.3574	3.9309	2	5	0	4	0.0000	0.4998	0.0192	0.5476	0.
df64_	df	df	F2	-	0.3575	0.3575	4.9325	1	5	4	1	0.0000	0.5000	0.0192	0.5489	0.
D	d	d	-	-	0.3636	0.3636	4	4	7	0	0	0.0000	0.0000	1.0000	0.0000	0.
F	f	f	-	-	0.3636	0.3636	4	4	7	0	0	0.0000	1.0000	0.0000	1.0000	0.
df	D	F	-	-	0.3636	0.3636	5.4142	0	3	8	0	0.0000	0.5000	0.5000	0.5000	0.
df73	df73_	df73_	-	-	0.3749	0.3749	4.1242	1	4	0	6	0.0000	0.4998	0.4904	0.4988	0.
df73_	df	df	F2	-	0.3752	0.3752	5.352	1	3	6	1	0.0000	0.5000	0.4905	0.5000	0.
B155b	A25	A91	C2	G2	0.3920	0.3920	5.3116	1	3	4	3	0.0000	0.5544	0.0042	0.8666	0.
sf	S	F	-	-	0.4091	0.4091	5.8229	1	3	7	0	0.0000	0.5000	0.0000	1.0000	0.
df19	df19_	df19_	-	-	0.4198	0.4198	4.6177	2	4	0	5	0.0000	0.9948	0.1432	0.7422	0.
df19_	df	df	F2	-	0.4199	0.4199	5.7373	2	3	3	3	0.0000	0.9948	0.1432	0.7424	0.
B58	A167	A167	C2	-	0.4322	0.4322	5.2542	4	6	0	1	0.1065	0.0355	0.1158	0.8916	0.
B298	A91	A91	C2	-	0.4327	0.4327	5.8778	2	4	2	3	0.0000	0.9775	0.4815	0.6807	0.

Molecular score prediction table sorted in ascending order by MS

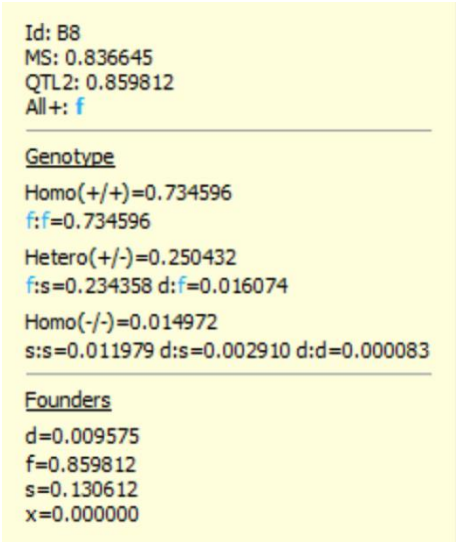
Column Descriptions

- The first five columns correspond to the first five columns of the genotype/pedigree (.dat) file (see Genotype/Pedigree File Structure section 7.5.1)
- MS: Molecular Score is the additive probability of transmitting favourable alleles over all QTL. MS varies between 0, for an individual who does not carry any of the favourable alleles, to 1, for an individual which is homozygote for the favourable alleles. Individuals with a MS of 1 correspond perfectly to the target genotype.
- Weight: Breeder designated scaling of QTL importance in the breeding scheme (see QTL Weights section 7.7.5)
- UC: The utility criterion (UC) combines molecular score of an individual with the expected variance of gametic molecular scores possible from that individual. UC is based on the estimation of the expected number of favourable alleles carried by the superior 5% gametes produced by the individual. This criterion favours individuals with no fixed unfavourable alleles. This score ranges from 0 to the number of QTL. Note that present version of UC estimation assumes independence between QTL and should be considered as only indicative in the case of linked QTL. UC also assumes that the distribution of scores can be approximated by a normal distribution, which is not a valid assumption in case of small number of heterozygous QTL.
- No.(+/+): Number of QTL homozygous for favourable allele(s). A given QTL is considered homozygous for favourable allele(s) when the probability (+/+) exceeds a default threshold value of 0.75 (see Change Default p. 151).
- No.(-/-): Number of QTL homozygous for unfavourable allele(s). A given QTL is considered as homozygous for unfavourable allele(s) when prob (-/-) exceeds a default threshold value of 0.75 (See Change Default).
- No.(+/-): Number of QTL heterozygous with both favourable and unfavourable allele(s). A given QTL is considered to belong to this category when prob (+/-) exceeds a default threshold value of 0.75 (See Change Default).

- No.(): Number of QTL defined as uncertain. Concerns QTL which are not attributed to any of the three previous categories.
- QTL: Individual molecular scores corresponding to each QTL

Molecular Score Details

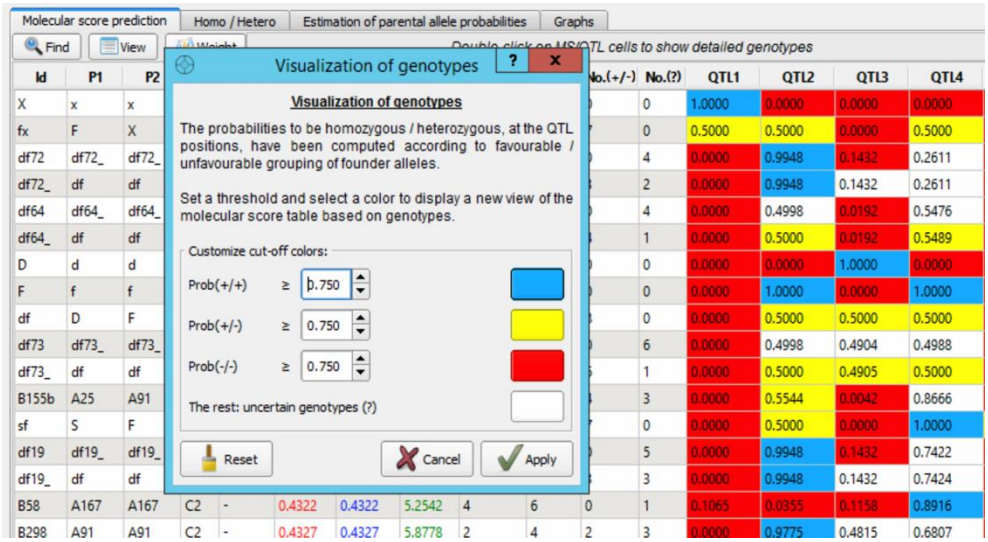
More detailed genotypes can be displayed by double clicking on MS or QTL cells. This view summarizes and aggregates the information presented in the Homo/Hetero and Estimation of Parental Allele Probabilities tables. Double clicking the cell representing QTL 2 of individual B8 to reveal a yellow popup box.



B8 has a molecular score of 0.8366451 at QTL 2, which has a probability of 0.734596 to be homozygous for the favourable allele f (i.e. Homo(+/+)). This score corresponds to the sum of the probabilities of the genotypes f:f=0.734596, f:s=0.234358 and d:f=0.0167074. QTL 2s MS of 0.995832 corresponds to the expected proportion of favourable allele(s) (i.e. Homo(+/+) + Hetero(+/)). Founders, represented by the d,f,s, and x alleles, indicate the expected proportion of parental alleles.

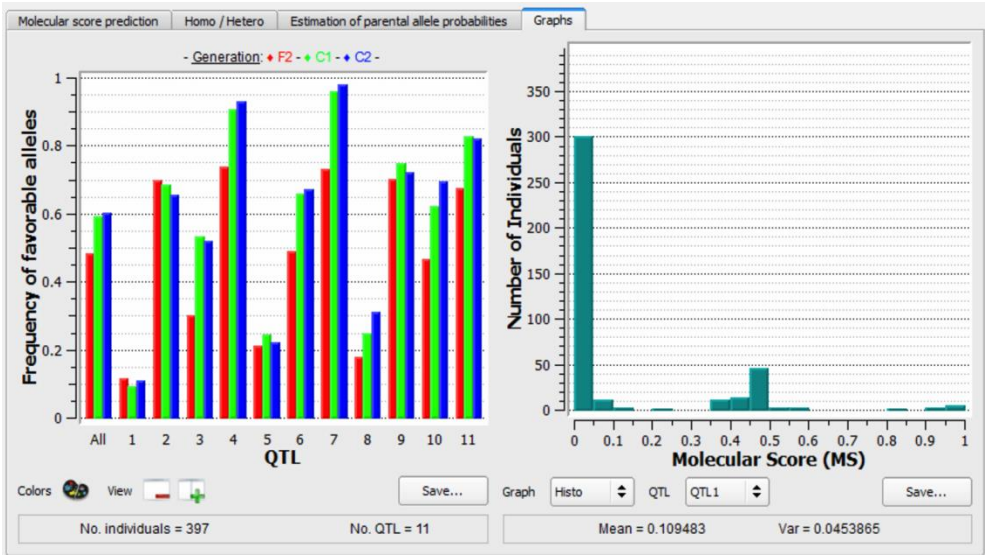
Visualize Genotypes

The Visualization of Genotypes window creates a custom color-coded view of the molecular score table to more easily identify fixed QTL. The default value of 0.750 for cut-off colours can be altered from this window. When you apply a new set cut-off/colour parameters, the four corresponding columns [No.(+/+), No.(-/-), No(+/-), No.()] on the MS table are updated. Uncertain genotypes can be due to (1) recombination near the QTL position making one flanking marker heterozygous and the other being homozygous favourable or (2) due to missing data in the genotypes/pedigree (.dat) file.



Graphical Output

Select graphs to visualize the tabular information. All the graphs can be exported as .png, .svg or .eps formats by selecting Save.



The left graph illustrates the frequency of favourable alleles at the different generations of selection. Note that no genetic gain is expected for the last generation (C2, in blue), because individuals have not yet been selected. The right graph is a histogram of the number of individuals with different molecular scores (MS) for QTL1 .

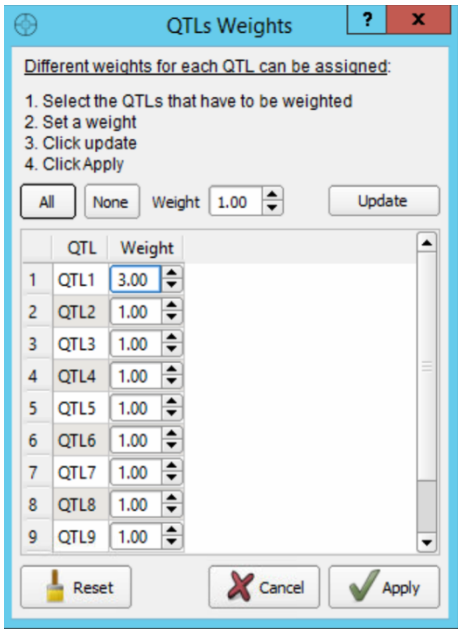
QTL Weights

The favorable allele at QTL 1 is in danger of being lost, because the 11 best individuals as defined by overall molecular score (MS) are fixed at 0 for this QTL.

Id	P1	P2	Cycle	Group	MS	Weight	UC	.(+/	›.(-/	›.(+	lo.(:	QTL1
B8	A1005	A1005	C2	-	0.8366	0.8366	9.2031	8	1	0	2	0.0000
B158	A251	A1005	C2	G1	0.8024	0.8024	9.3268	7	1	1	2	0.0000
B28	A1006	A251	C2	G1	0.7740	0.7740	9.2215	6	1	1	3	0.0000
B13	A1006	A1005	C2	-	0.7609	0.7609	9.0767	6	1	2	2	0.0000
B38	A1040	A1005	C2	-	0.7494	0.7494	9.1095	6	1	1	3	0.0000
B37	A1040	A1005	C2	-	0.7433	0.7433	8.6768	7	1	1	2	0.0000
B40	A1040	A1040	C2	-	0.7404	0.7404	9.0101	7	1	2	1	0.0000
B242	A37	A1005	C2	-	0.7305	0.7305	8.7424	6	1	1	3	0.0000
B246	A37	A1040	C2	-	0.7303	0.7303	8.8995	6	1	3	1	0.0000
B293	A9	A1040	C2	-	0.7268	0.7268	8.8608	6	1	3	1	0.0000
B7	A1005	A1005	C2	-	0.7238	0.7238	8.4621	6	2	0	3	0.0000
B124	A23	A167	C2	-	0.7223	0.7223	8.8114	6	1	1	3	0.4812

Top 12 germplasm based on molecular score (MS): Not weighted for QTL 1

OptiMAS attributes the same weight to all QTL in the map file for molecular score (MS) estimation. However is possible to discard QTL and/or to attribute weights defined by the breeder by applying a Weight index to the MS calculation. Select the Weight button to open the QTL Weights dialog window. Assign QTL 1 a weight of 3 to produce a new classification of individual based on the new weighted MS.



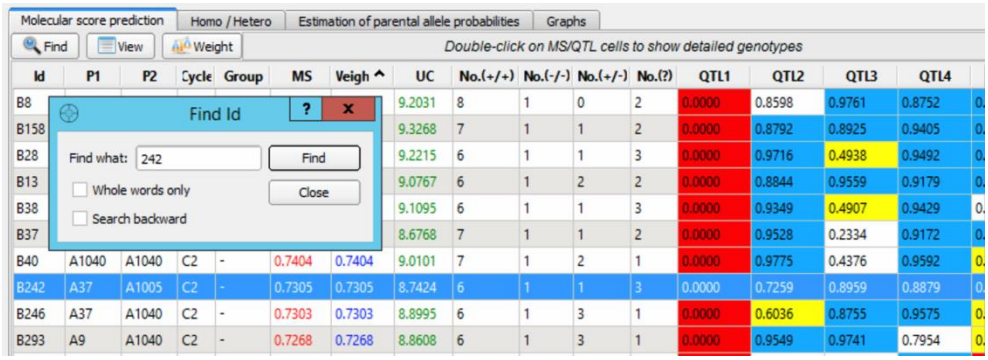
Increasing the weight of QTL 1 allows germplasm to be sorted by weighted molecular score, and ensures that some of the top scoring germplasm can transmit the favorable allele.

Id	P1	P2	Cycle	Group	MS	Weight	UC	.(+/	›.(-/	›.(+	lo.(?	QTL1
B8	A1005	A1005	C2	-	0.8366	0.7079	9.2031	8	1	0	2	0.0000
B124	A23	A167	C2	-	0.7223	0.6852	8.8114	6	1	1	3	0.4812
B158	A251	A1005	C2	G1	0.8024	0.6790	9.3268	7	1	1	2	0.0000
B125	A23	A167	C2	-	0.7032	0.6691	8.4426	6	2	1	2	0.4812
B47	A166	A167	C2	-	0.6309	0.6570	7.8056	4	2	2	3	0.8010
B28	A1006	A251	C2	G1	0.7740	0.6550	9.2215	6	1	1	3	0.0000
B57	A167	A1040	C2	-	0.6860	0.6545	8.7713	4	0	6	1	0.4812
B110	A212	A1005	C2	-	0.6863	0.6481	8.6678	4	0	4	3	0.4376
B13	A1006	A1005	C2	-	0.7609	0.6438	9.0767	6	1	2	2	0.0000
B123	A212	A91	C2	-	0.6717	0.6390	8.5066	4	1	3	3	0.4590
B93	A211	A212	C2	G3	0.6543	0.6380	8.0634	6	2	3	0	0.5484
B109	A212	A1005	C2	-	0.6675	0.6355	8.3429	4	1	4	2	0.4590

Top 12 germplasm based on weighed molecular score: QTL 1 weighted 3

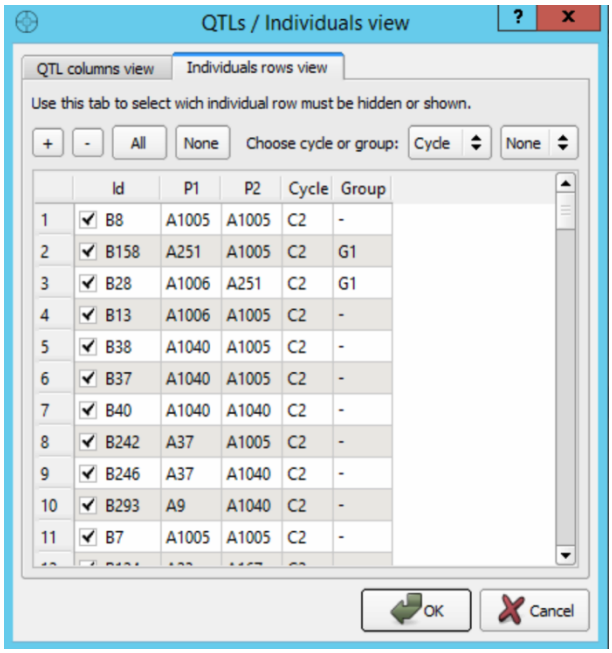
Find Individuals on Molecular Score Table

The Find ID dialog box can be used to search and locate a specific individual in the panel. Select Find and enter 242.



Filter Molecular Score Table

The QTL/Individuals dialog is used to enable or disable the display of QTL and/or individuals on the MS table. Press the View button to display the filter dialog. Select and check QTL and/or individuals to follow. Select OK to apply the corresponding filter and immediately refresh the MS table. This new view of the MS table can be useful if you are working with a large number of QTL and/or individuals and you want to focus on specific QTL/plants.



Selection

The selection features are accessed by choosing Selection. Genetic value predictions can be used in three different ways to inform selections:

- Manually select individuals based on own judgment
- Truncated selection based on molecular score (MS), weighted MS, or the utility criterion (UC).
- QTL complementation selection which aims to prevent the loss of rare favourable alleles

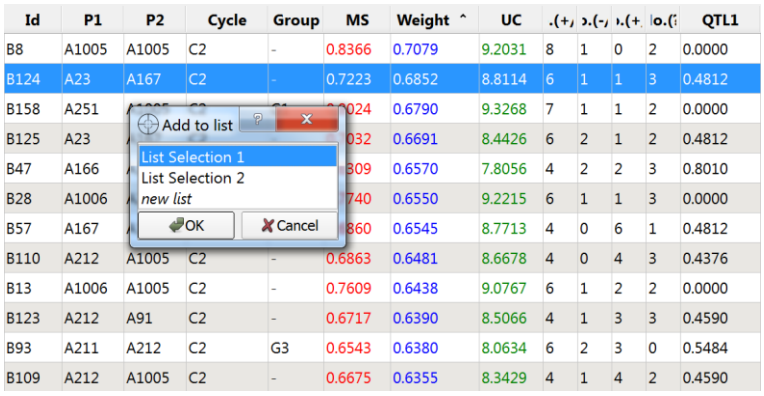
Access Selection Functions by clicking Step 2 from the left menu.

Manual Selection

Create a custom list of germplasm with high weighted molecular scores, but without fixation at QTL1.

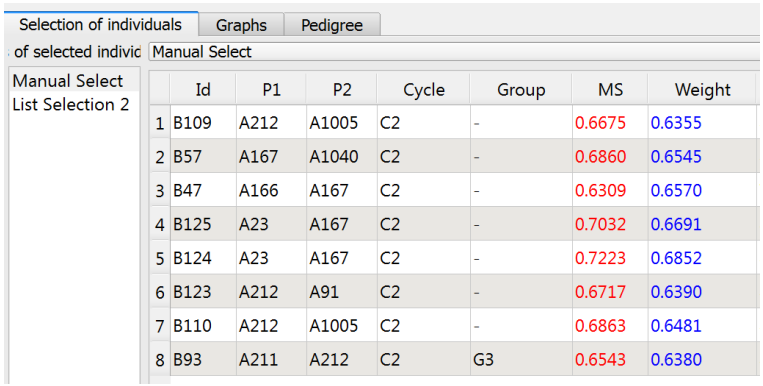
Double click the following lines and add to List Selection 1:

- B124
- B125
- B47
- B57
- B110
- B123
- B93
- B109



Id	P1	P2	Cycle	Group	MS	Weight	UC	.(+/	›.(-/	›.(+	lo.(:	QTL1
B8	A1005	A1005	C2	-	0.8366	0.7079	9.2031	8	1	0	2	0.0000
B124	A23	A167	C2	-	0.7223	0.6852	8.8114	6	1	1	3	0.4812
B158	A251				024	0.6790	9.3268	7	1	1	2	0.0000
B125	A23				032	0.6691	8.4426	6	2	1	2	0.4812
B47	A166				309	0.6570	7.8056	4	2	2	3	0.8010
B28	A1006				740	0.6550	9.2215	6	1	1	3	0.0000
B57	A167				860	0.6545	8.7713	4	0	6	1	0.4812
B110	A212	A1005	C2	-	0.6863	0.6481	8.6678	4	0	4	3	0.4376
B13	A1006	A1005	C2	-	0.7609	0.6438	9.0767	6	1	2	2	0.0000
B123	A212	A91	C2	-	0.6717	0.6390	8.5066	4	1	3	3	0.4590
B93	A211	A212	C2	G3	0.6543	0.6380	8.0634	6	2	3	0	0.5484
B109	A212	A1005	C2	-	0.6675	0.6355	8.3429	4	1	4	2	0.4590

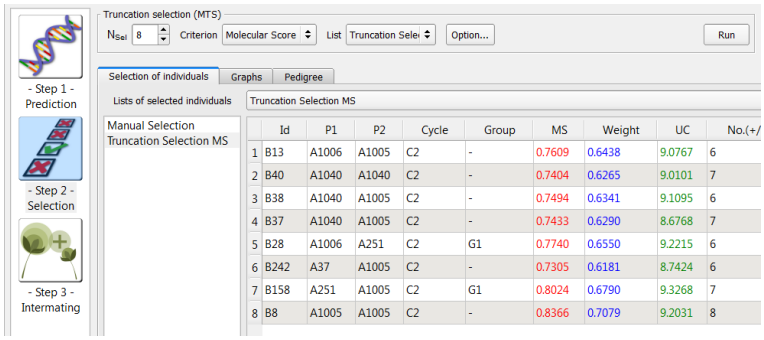
Access List Selection 1 through the Selection Window. Double click on "List Section 1" and rename "Manual Selection"



	Id	P1	P2	Cycle	Group	MS	Weight
1	B109	A212	A1005	C2	-	0.6675	0.6355
2	B57	A167	A1040	C2	-	0.6860	0.6545
3	B47	A166	A167	C2	-	0.6309	0.6570
4	B125	A23	A167	C2	-	0.7032	0.6691
5	B124	A23	A167	C2	-	0.7223	0.6852
6	B123	A212	A91	C2	-	0.6717	0.6390
7	B110	A212	A1005	C2	-	0.6863	0.6481
8	B93	A211	A212	C2	G3	0.6543	0.6380

Truncation Selection

Truncated selections are automatically made bases on Molecular score (MS), weighted (MS) or utility criterion (UC). Set the truncation selction paramenters. Select 8 lines (N_{sel}) based on molecular score (MS). Rename "List Selection 2", "Truncation." Select the Truncation list and click run.



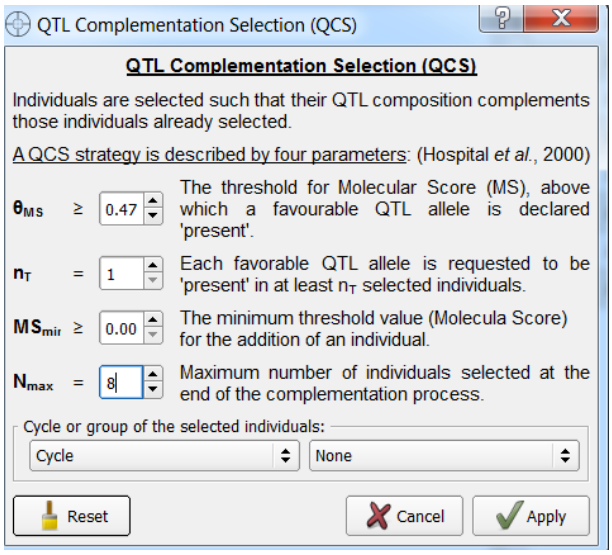
	Id	P1	P2	Cycle	Group	MS	Weight	UC	No.(+/
1	B13	A1006	A1005	C2	-	0.7609	0.6438	9.0767	6
2	B40	A1040	A1040	C2	-	0.7404	0.6265	9.0101	7
3	B38	A1040	A1005	C2	-	0.7494	0.6341	9.1095	6
4	B37	A1040	A1005	C2	-	0.7433	0.6290	8.6768	7
5	B28	A1006	A251	C2	G1	0.7740	0.6550	9.2215	6
6	B242	A37	A1005	C2	-	0.7305	0.6181	8.7424	6
7	B158	A251	A1005	C2	G1	0.8024	0.6790	9.3268	7
8	B8	A1005	A1005	C2	-	0.8366	0.7079	9.2031	8

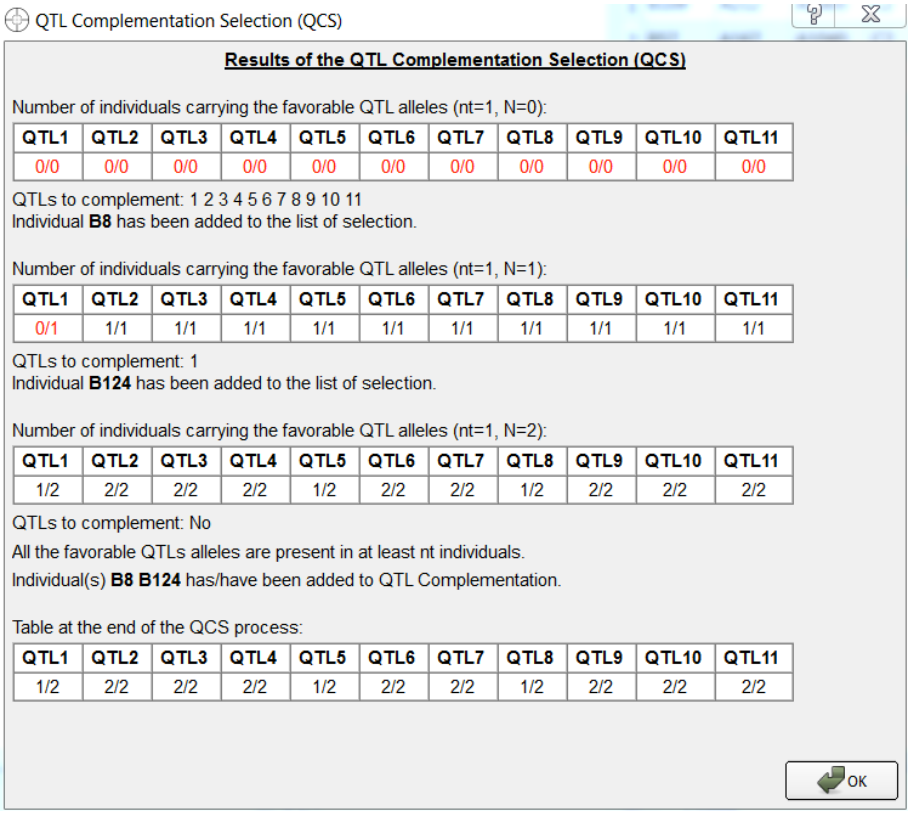
QTL Complementation Selection

QTL complementation selection aims to prevent the loss of rare favourable alleles, and is recommended when high numbers of QTL are being considered. QTL complementation adds individuals that complement a list previously created by truncation selection.

From the Selection window add a new list titled, "QTL Complementation". Select this list for QTL complementation and click Option.

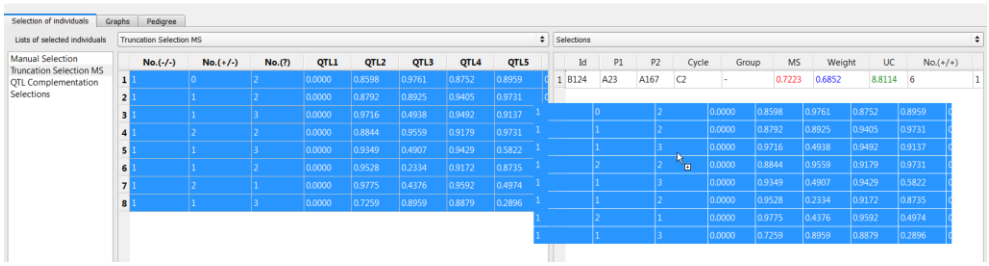
Set N_{max} to 8. Leave the other options in their default settings. Select Apply and Run QCS.





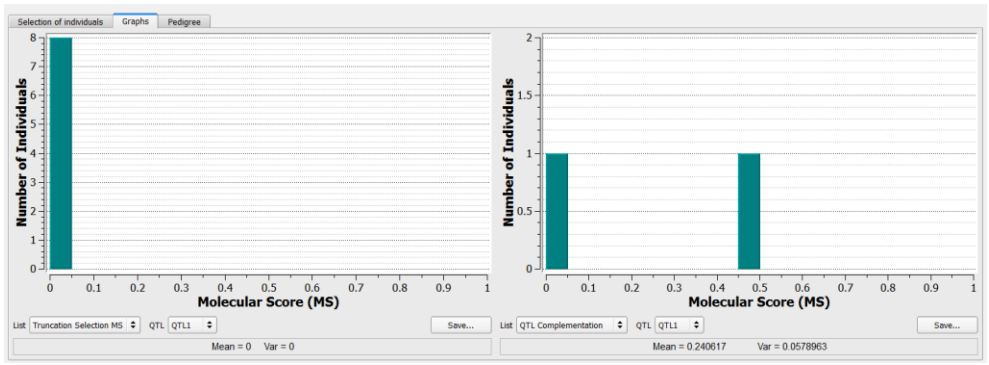
Two individuals, B8 and B124 have been identified by QTL complementation selection. B124 was not identified by truncation selection.

Create a combined list of the 9 individuals selected by truncation selection and QTL complementation. Add a new list and name it "Selections." Drag and drop the unique lines to the new list.



Graphs

The Graphs tab allows for further comparisons of selection criteria.

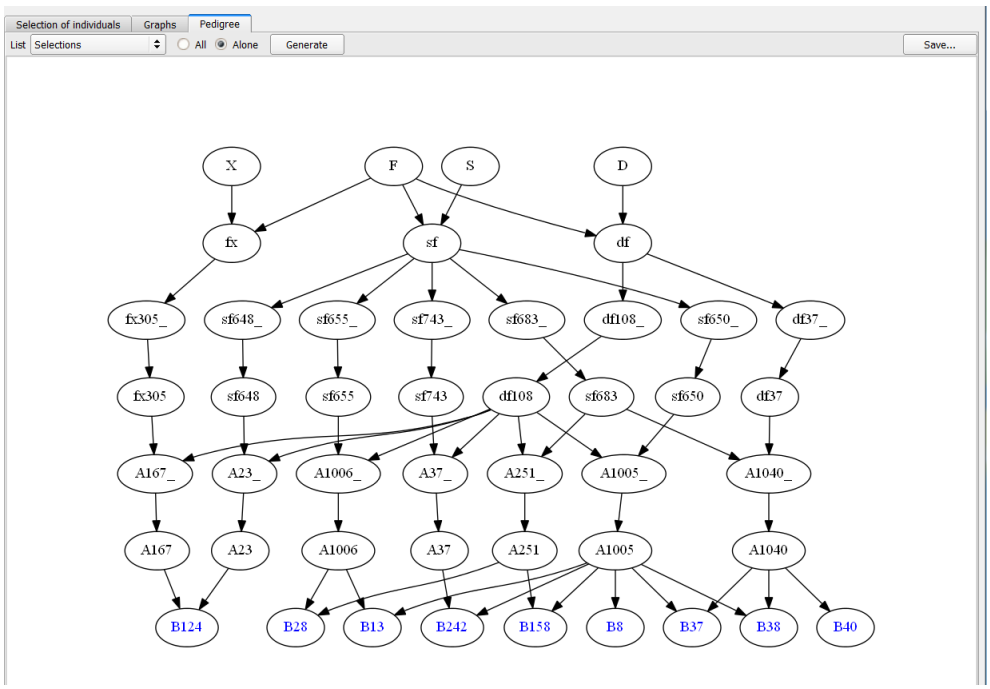


The left graph shows that all 8 individuals selected through truncation will carry the unfavourable allele at QTL1. The right graph on the includes the one individual, B124, added via the QTL complementation procedure that carries the favorable allele.

Selection Pedigrees

Pedigree illustration allows for easy visualization of the genetic contribution of selected individuals over generations.

To display the pedigree illustration, choose the Selections list and indicate Alone to only display selected individuals. Click Generate.



Pedigree of 9 Individuals Selected based on truncation and QTL complementation selection.

Intermating

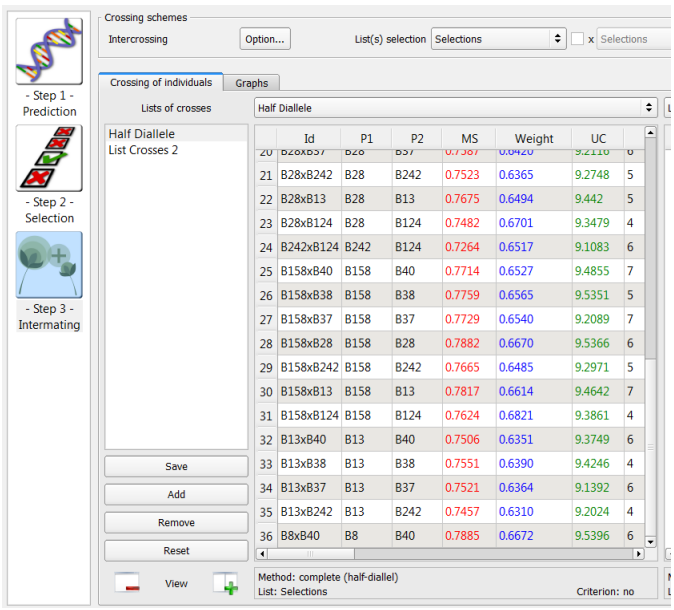
Identify crosses to initiate the next round of marker-assisted selection using one of three design methods.

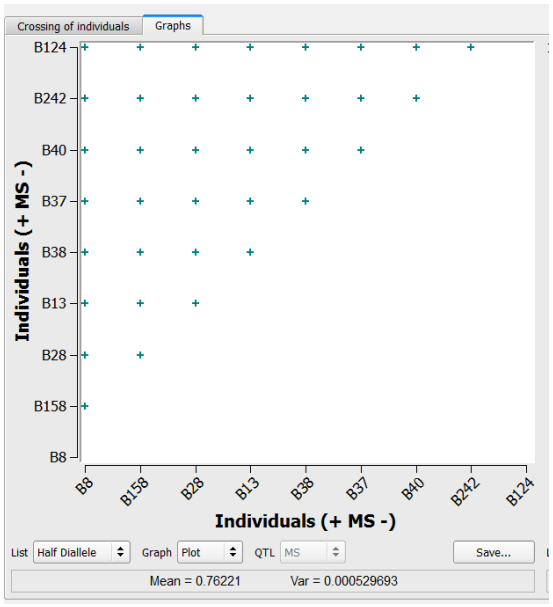
- The half-diallel complete method identifies crosses among all selected plants in a list
- The better half method identifies crosses within a list, while avoiding crosses between selected individuals with low genetic values
- Factorial Design Method: identifies crosses between two selection lists

Select the Intermating functions by selecting Step 3 from the left menu.

Half-Diallel Complete Method

The half-diallel complete method identifies crosses among all selected plants in a list. Name an empty intermating list, Half-diallel. Choose the Selections list of 9 individuals selected based on truncation and QTL complementation selection. Select the Complete method option and click Apply. Assign the output to the Half Diallel list. Select Run.

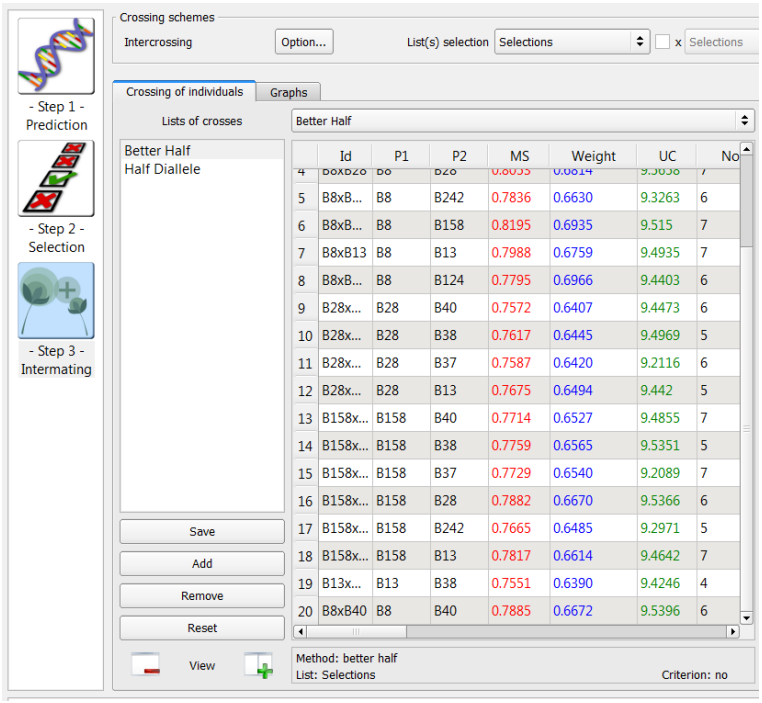


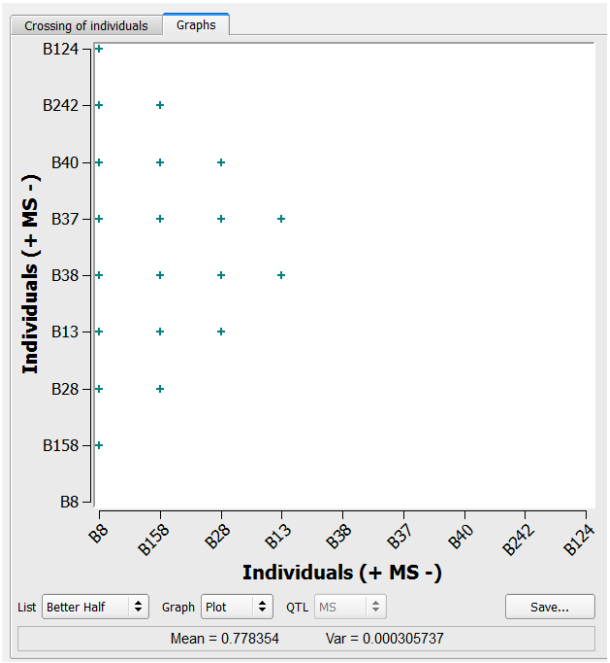


36 crosses generated by the half-diallele method on the list of 9 individuals selected based on truncation and QTL complementation selection

Better Half Method

The better half method identifies crosses within a list, while avoiding crosses between selected individuals with low genetic values. Name an empty intermating list Better Half. Choose the Selections list of 9 individuals selected based on truncation and QTL complementation selection. Select the Better Half method option. Select Run.

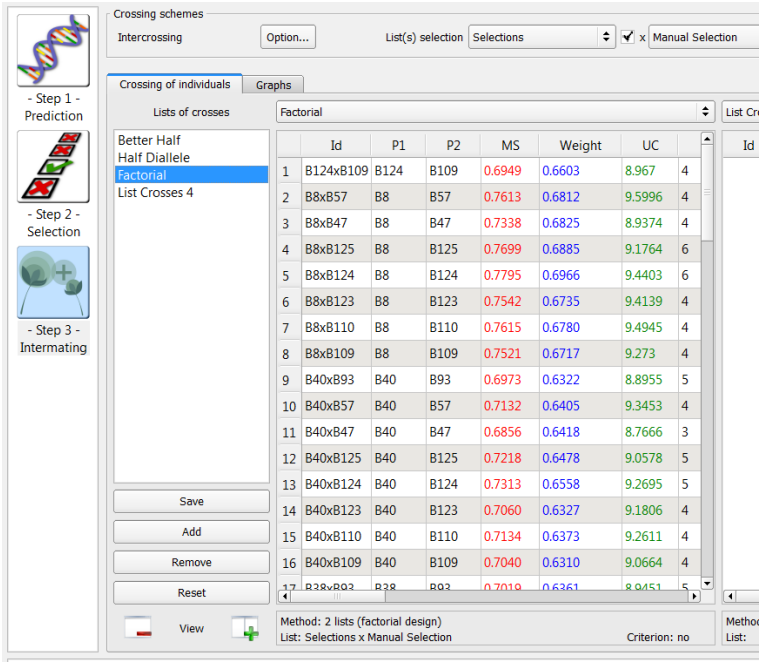




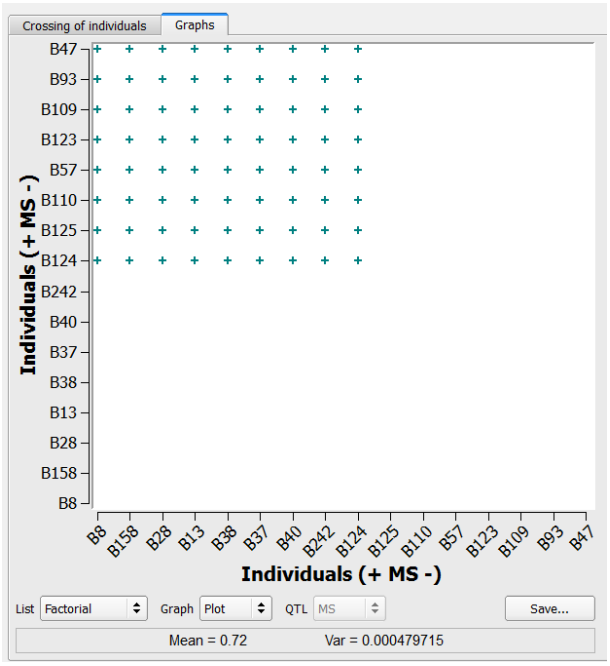
20 crosses generated by the better half method on the list of 9 individuals selected based on truncation and QTL complementation selection

Factorial Intermating Design

Intermating designs can be created using two selection lists by populating the additional list option. Cross the Manual selection (n=8) list with the Selection (n=9).



	Id	P1	P2	MS	Weight	UC	
1	B124xB109	B124	B109	0.6949	0.6603	8.967	4
2	B8xB57	B8	B57	0.7613	0.6812	9.5996	4
3	B8xB47	B8	B47	0.7338	0.6825	8.9374	4
4	B8xB125	B8	B125	0.7699	0.6885	9.1764	6
5	B8xB124	B8	B124	0.7795	0.6966	9.4403	6
6	B8xB123	B8	B123	0.7542	0.6735	9.4139	4
7	B8xB110	B8	B110	0.7615	0.6780	9.4945	4
8	B8xB109	B8	B109	0.7521	0.6717	9.273	4
9	B40xB93	B40	B93	0.6973	0.6322	8.8955	5
10	B40xB57	B40	B57	0.7132	0.6405	9.3453	4
11	B40xB47	B40	B47	0.6856	0.6418	8.7666	3
12	B40xB125	B40	B125	0.7218	0.6478	9.0578	5
13	B40xB124	B40	B124	0.7313	0.6558	9.2695	5
14	B40xB123	B40	B123	0.7060	0.6327	9.1806	4
15	B40xB110	B40	B110	0.7134	0.6373	9.2611	4
16	B40xB109	B40	B109	0.7040	0.6310	9.0664	4



72 crosses generated by the factorial method on the list of 9 individuals selected based on truncation and QTL complementation selection and 8 individuals manually selected

References

Bernardo, R., Moreau, L. and Charcosset, A. (2006) Number and fitness of selected individuals in marker-assisted and phenotypic recurrent selection, Crop Science. 46: 1972-1980.

Blanc, G., Bardol, N., Charcosset, A., Gauthier, F., Joets, J., Moreau, L., & Valente, F. (2013) OptiMAS: A decision support tool for marker-assisted assembly of diverse alleles. Journal of Heredity. 104(4):586-590.

Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L. (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theoretical and Applied Genetics. 113:206224.

Blanc G, Charcosset A, Veyrieras JB, Gallais A, Moreau L. (2008) Marker-assisted selection efficiency in multiple connected populations: a simulation study based on the results of a QTL detection experiment in maize. Euphytica. 161:7184.

Hospital, F., Goldringer, I., & Openshaw, S. (2000). Efficient marker-based recurrent selection for multiple quantitative trait loci. Genetics Research. 75(03): 357-368.

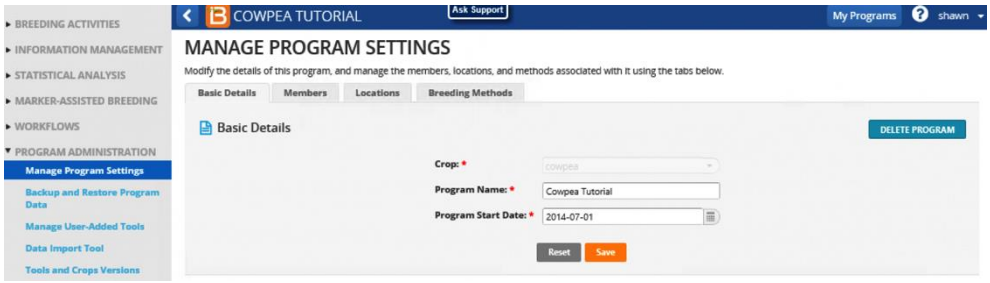
Acknowledgements

Maize mulitparental demonstration data were provided by Alain Charcosset. These data may have been adapted for training purposes. Any misrepresentation of the original data is the solely the responsibility of the IBP.

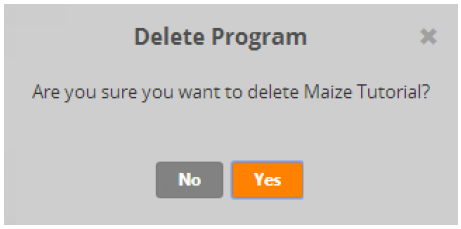
Program Administration

Manage Program Settings

Manage Program Settings allows you to edit the program settings established when the program was added (see Add & Manage Programs) and to delete existing programs.

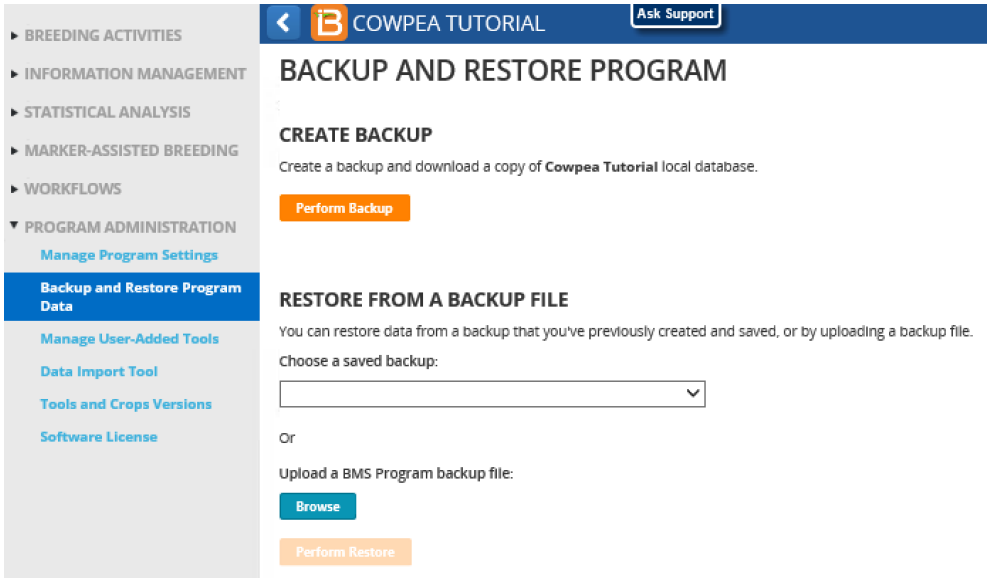


Deleting a program will result in permanent data loss. Backing up the program before deleting will ensure that the data can be restored.



Backup Program

Frequently backup your programs to protect against accidental data loss. Use the restore function to restore your actual data or to follow along with tutorial instruction.



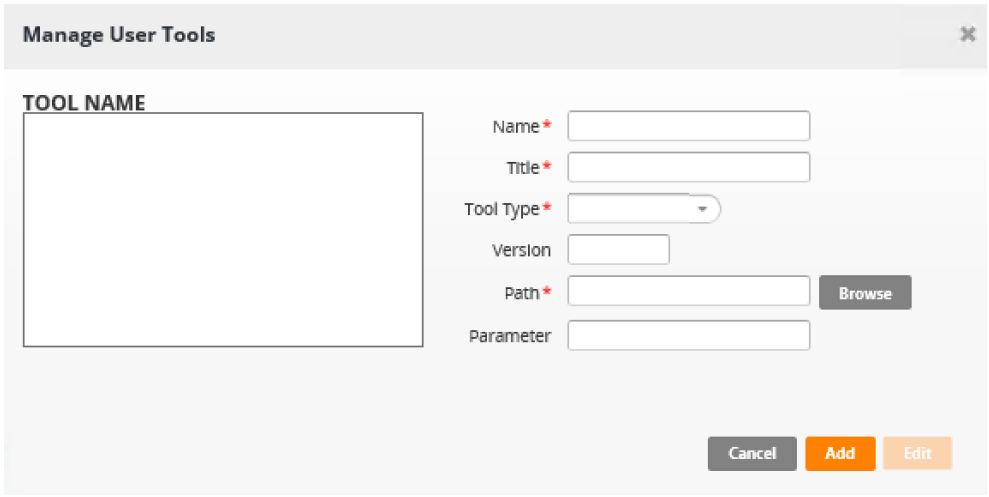
Backup programs will be automatically saved as .sql files within C:\Breeding Management System\ backup folder. However, uninstalling the BMS will result in the loss of all files within C:\Breeding Management System. Regularly copy backup files to additional locations and computers, particularly before updating or uninstalling the software.

Restore Program

Select a program database (.sql file) and select Preform Restore.

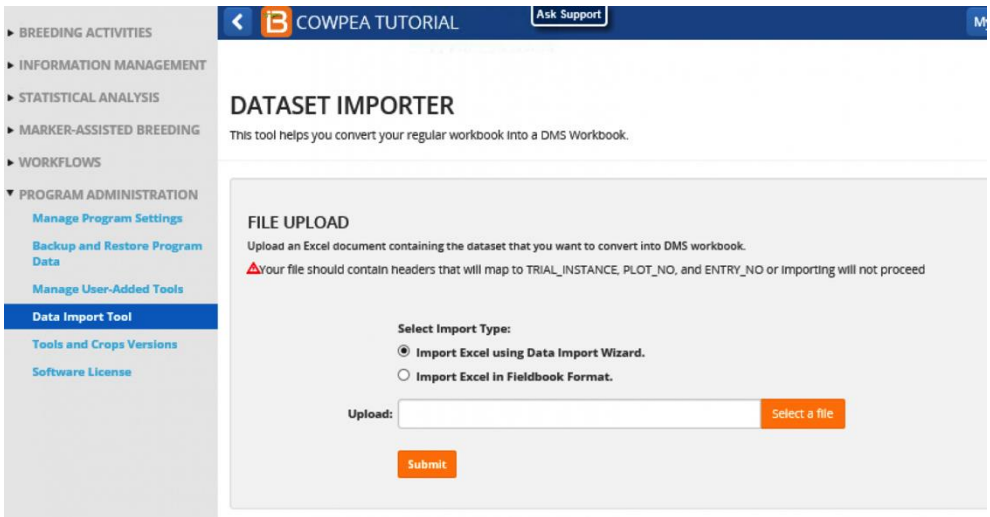
User-Added Tools

The Breeding Management System allows users to add custom user tools to the workbench. Please contact support@vsni.co.uk for support using this function.



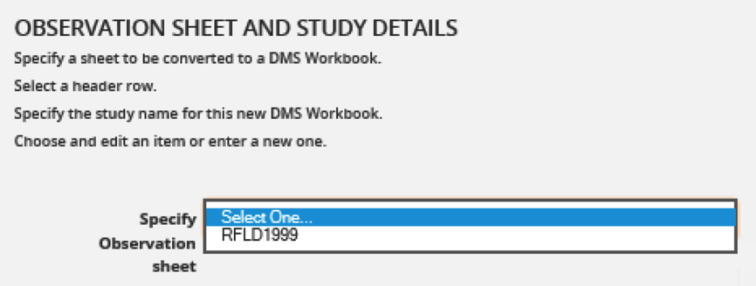
Data Import Tool

The data import tool allows users to upload historical field observations from Excel files into the Breeding Management System. The Data Import Wizard is the most flexible way to import historical data from trials and nurseries. Importing from Fieldbook formatted Excel files is the simplest way to import, but only useful if you have been using a standalone Fieldbook to manage your trials and nurseries.

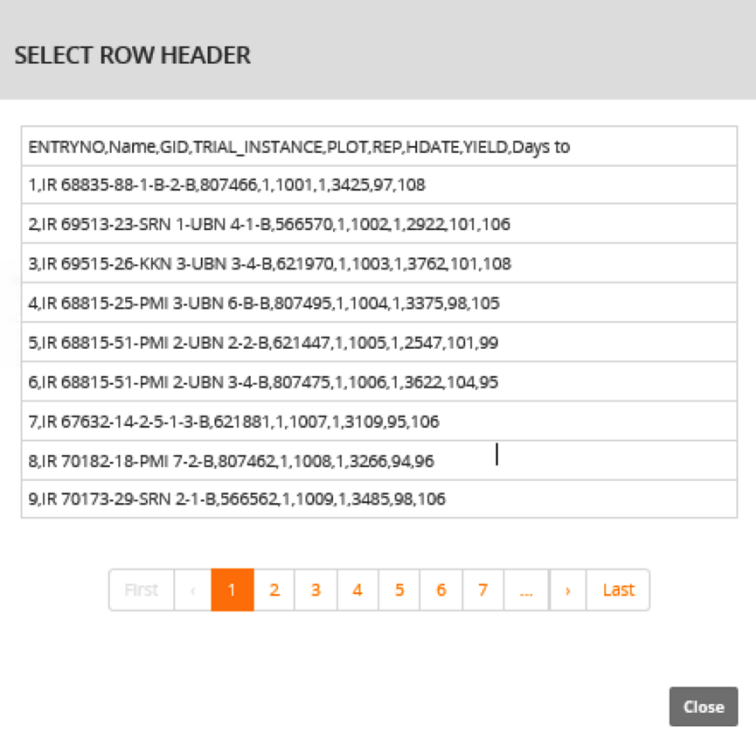


Data Import Wizard

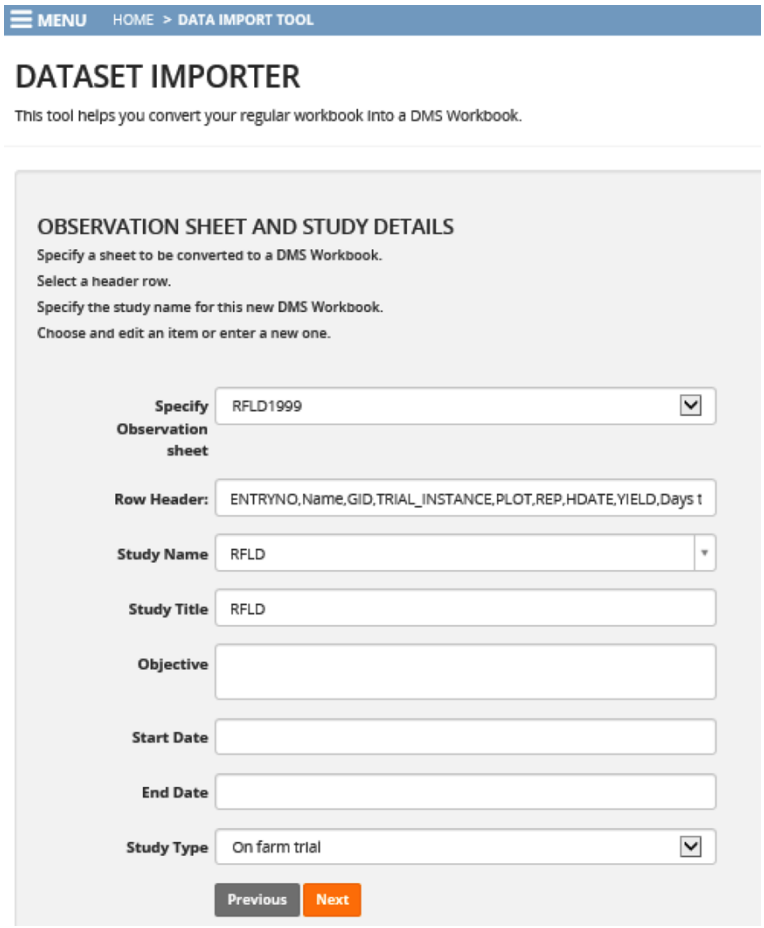
Select the Data Import Wizard, and choose an excel file to import. The rice example (.xls) illustrated here can be replicated within a rice program. See below for file format details. Specify the observation sheet containing the phenotypic data within the Excel file. If the file only has one sheet of data there will be only one observation sheet option.



Select the appropriate row of column headings and Close.

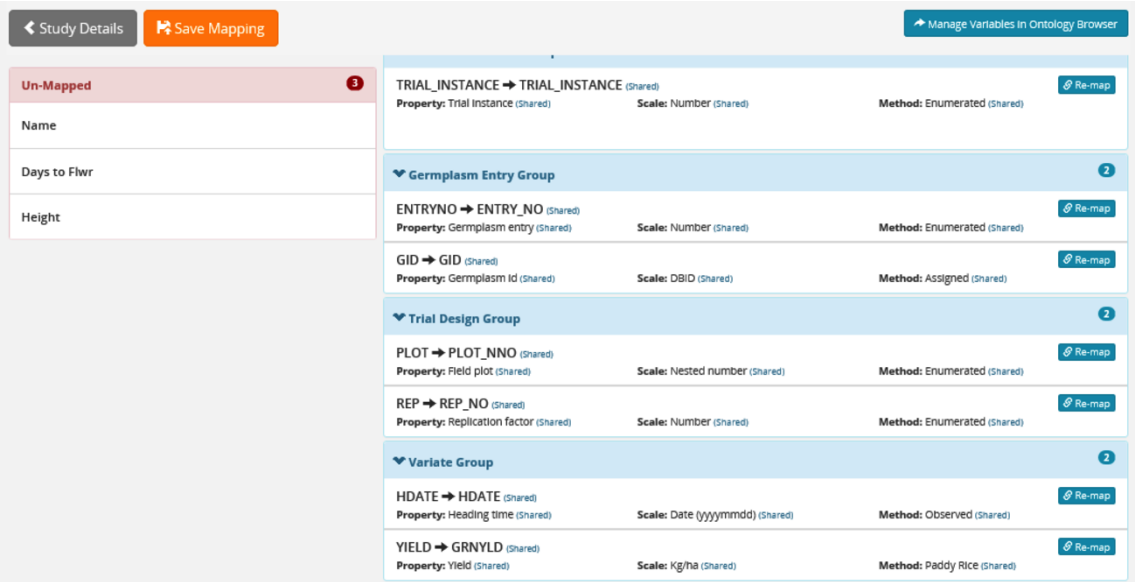


Enter a Study Name, Title, and Type. Select Next. Expect to wait a few seconds for initial mapping to complete.



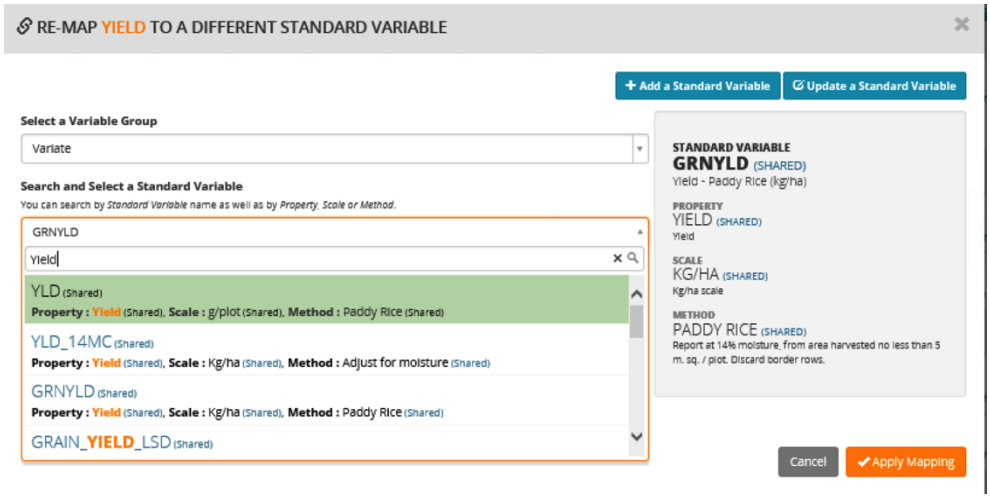
Map Excel Column Headers to Database Ontology

The BMS will attempt to map the Excel sheet column headers to their corresponding database terms. Carefully review mapped selections, highlighted in blue, to ensure that ontology term chosen by the system truly matches the meaning and measurement of the imported data.

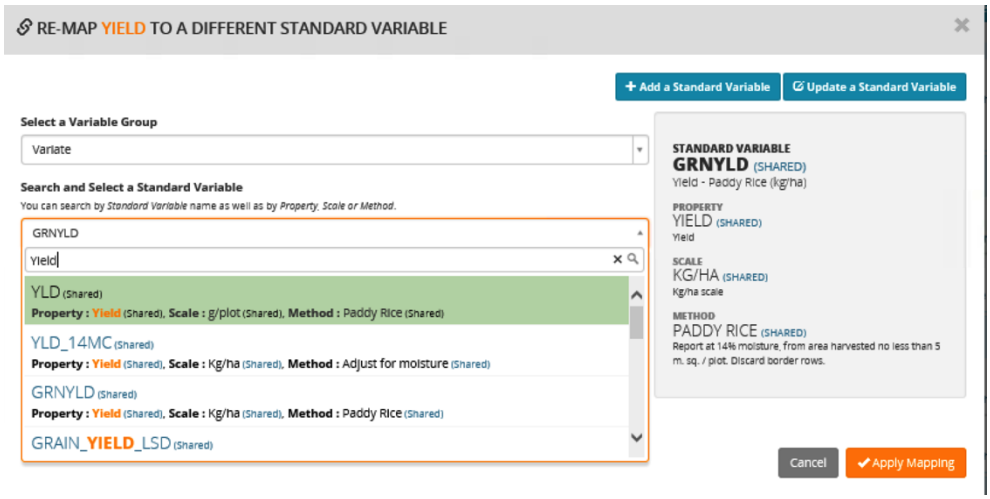


In this example, the BMS has mapped 7 of the 10 columns of data. These 7 will need to be carefully reviewed, and the 3 unmapped columns of data will need to be manually matched to the database ontology.

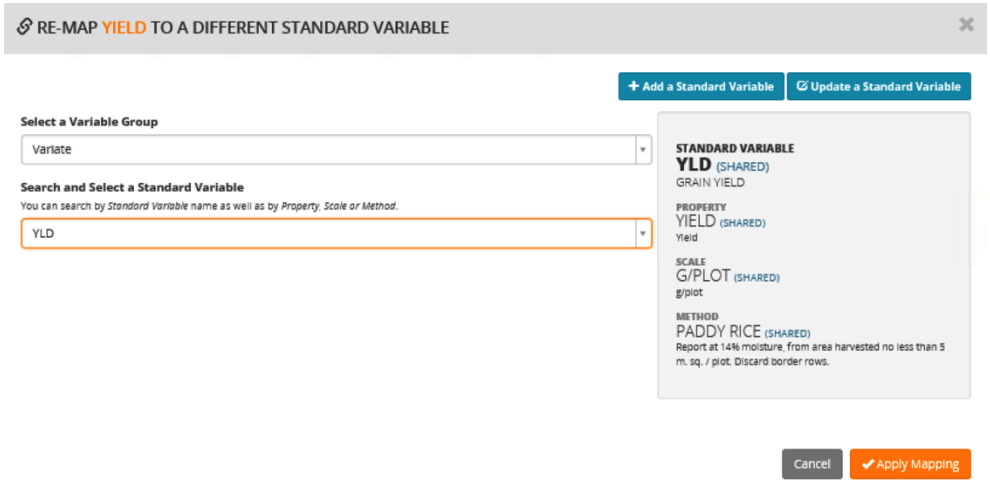
Review the details of the suggested matches by selecting the Re-Map icon. Search for alternative matches if the scale or method for the selected match is different than the actual experimental design. Obtain the details of an alternative variable name by highlighting and selecting.



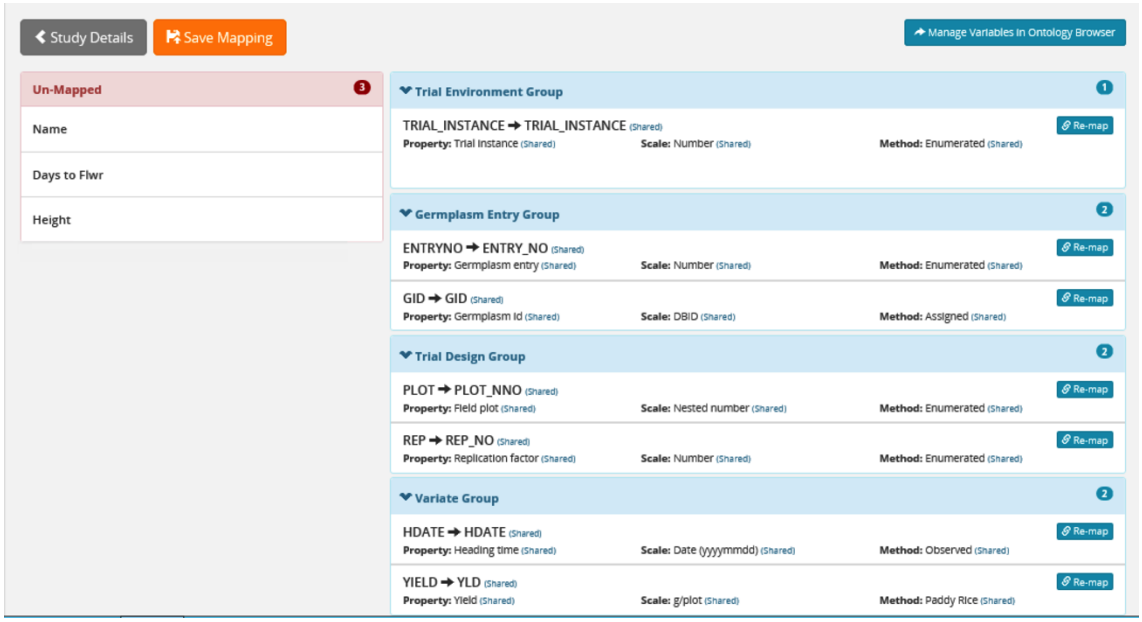
If the alternative term is an actual match, select Apply Mapping. If a match is not found, select Add a Standard Variable, which takes you to the Ontology Manager.



If the alternative term is an actual match, select Apply Mapping. The change will be reflected on the main screen.

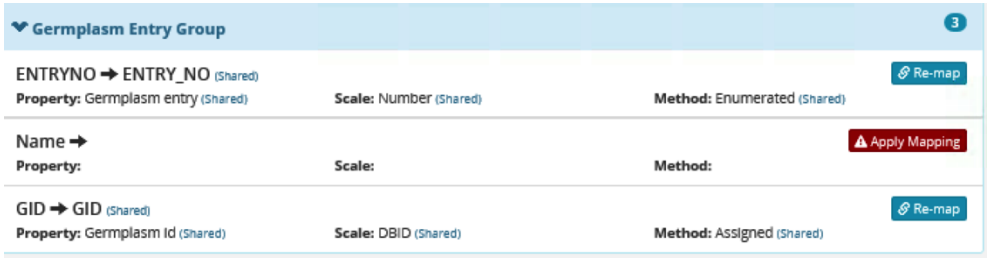


Drag and Drop selected Un-Mapped terms to their appropriate group: Trial Environment, Germplasm Entry, or Variate.

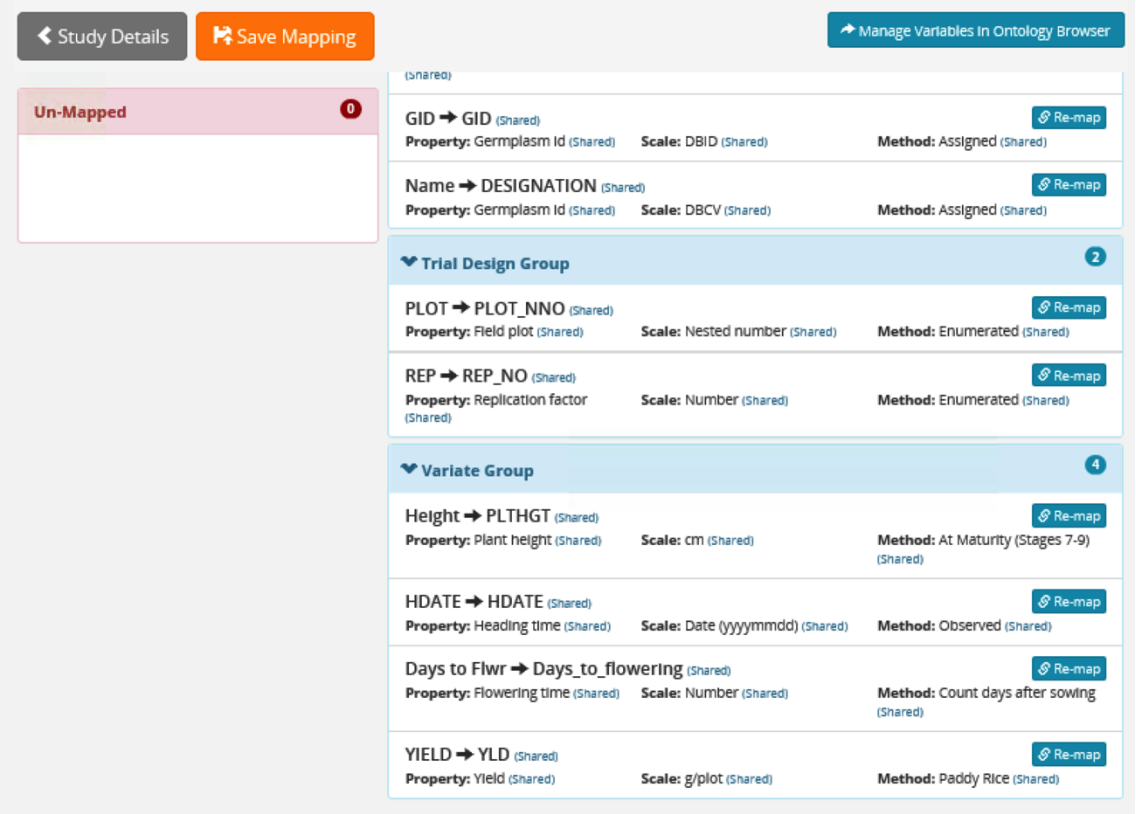


In this example, the Un-Mapped term, Name, describes germplasm entries and should be included with the Germplasm Entry group on the right. Days to Flower and Height are variates (phenotypic observations).

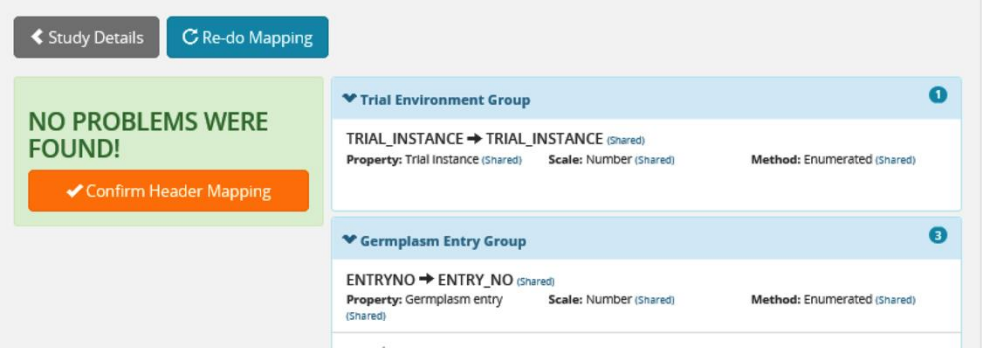
Apply mapping to the un-mapped terms by manually searching for the appropriate term.



Select Save Mapping when mapping is complete or if the user will complete mapping at a later time.

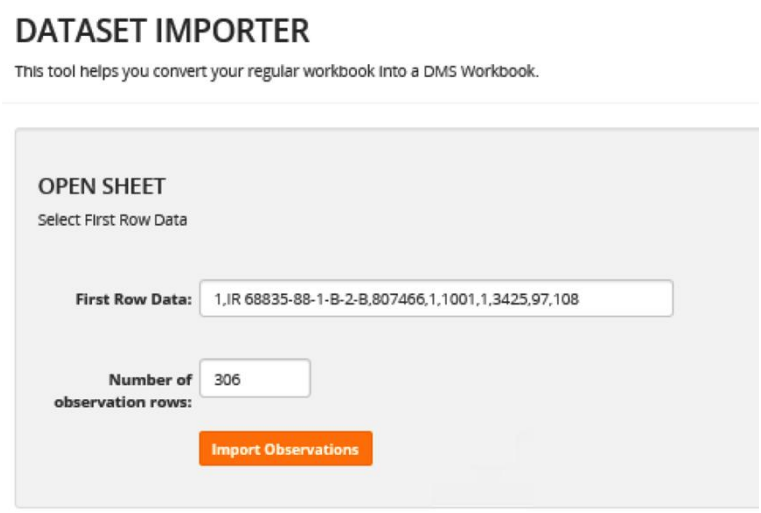


When mapping is complete, select Confirm Header Mapping.

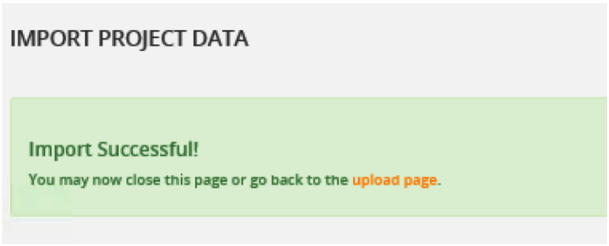


Ontology mapping is complete. Import the project data by selecting continue or return to the upload page to add more sheets of data to the trial.

Select the first row of data and import the observations. The phenotypic observations for this trial are now saved in the program database.



The phenotypic observations for this trial are now saved in the program database.



Excel File Format for Import Wizard

Only one sheet of observation data can be read at a time. Two columns of descriptive data are required to import phenotypic observations: TRIAL_INSTANCE and ENTRY_NO. Trial instance is a numerical identification of an individual experiment within a multisite or multiyear trial. For example, data from a single site or nursery will all have the same trial instance value. Entry number is a digit unique to each entry. For example, a trial containing three replicates of 50 accessions, will have the entry numbers 1-50 replicated three times. However you will probably also want to include additional descriptive data. If you plan to complete a single site phenotypic analysis you will need to include replication (REP). You will probably also want to import a column of germplasm names that are meaningful to your program to ease the interpretation of results.

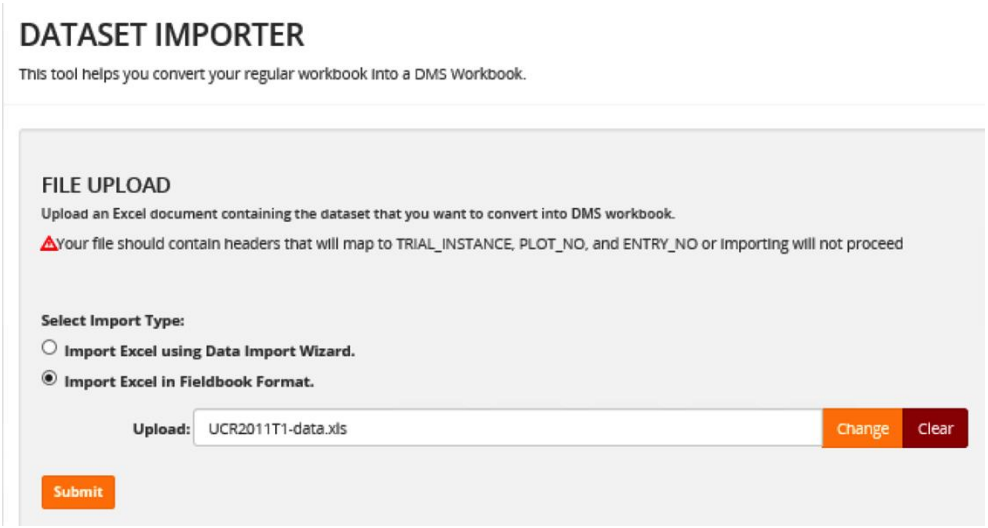
Full database integration requires that every germplasm have a GID. Without GIDs, the imported phenotypic data will not relate to phenotypic data outside of this trial or to other database information, like genotype and pedigrees. **Full database integration requires that you first import the germplasm list into the database and assign GIDs to all germplasm.**

	A	B	C	D	E	F	G	H	I	J
1	ENTRYNO	Name	GID	TRIAL_INSTANC	PLOT	REP	HDATE	YIELD	Days to Flwr	Height
2	1	68835-88-1-B-2-	807466	1	1001	1		3425	97	108
3	2	13-23-SRN 1-UBN	566570	1	1002	1		2922	101	106
4	3	15-26-KKN 3-UBN	621970	1	1003	1		3762	101	108
5	4	15-25-PMI 3-UBN	807495	1	1004	1		3375	98	105
6	5	15-51-PMI 2-UBN	621447	1	1005	1		2547	101	99
7	6	15-51-PMI 2-UBN	807475	1	1006	1		3622	104	95
8	7	67632-14-2-5-1-3	621881	1	1007	1		3109	95	106
9	8	70182-18-PMI 7-2	807462	1	1008	1		3266	94	96
10	9	70173-29-SRN 2-	566562	1	1009	1		3485	98	106
11	10	502-1-SRN 3-UBI	204226	1	1010	1		3833	103	114
12	11	502-29-SRN 1-UB	807438	1	1011	1		3069	99	109
13	12	513-11-SRN 1-UB	95426	1	1012	1		2538	107	102
14	13	0207-15-CPA 6-1	694030	1	1013	1		3304	99	120
15	14	0215-70-CPA 3-4	807437	1	1014	1		3052	101	103
16	15	R 71505-22-B-1-E	807436	1	1015	1		4218	107	114
17	16	R 71506-27-2-1-E	807435	1	1016	1		3799	101	109
18	17	68835-28-2-B-1-3	688470	1	1017	1		3393	111	125
19	18	68835-28-2-B-1-4	576763	1	1018	1		3242	107	116

This data import sheet (.xls) contains the required entry number and trial instance columns, but also contains additional descriptive information: germplasm name, GID, plot number, and replication. This sheet also contains four columns of phenotypic data: heading date, yield, days to flower, and plant height.

Data Import from Fieldbook Formatted File

Excel files generated by exporting from the BMS or the standalone IB Fieldbook are formatted for import into the BMS. Select a Fieldbook formatted Excel file and Submit.



Fieldbook file import is complete. Return to the upload page to upload additional files.