



Wang Lab

Seed Database Manager

Ver. 1.1 Manual

The ultimate guide to the Wang Lab Seed Database Manager

Database tool and manual written by Sam Schafer



Table of Contents

Seed Database Manager Overview	3
Introduction	3
Database Format.....	3
Accessing the Application and Database	5
Adding new users into the database	5
Logging into the database.....	5
Navigating the Database.....	5
Search Queries.....	6
Downloading the Database.....	7
The Toolbar	7
The Operations Deck.....	7
Function Guide.....	8
Register new seeds	8
Register multiple seeds (template upload).....	10
Add a note.....	13
Organize: create germplasm sets.....	13
Organize: remove germplasm groups.....	14



Seed Database Manager Overview

Introduction

The Seed Database Manager (SDM) was designed to centralize and organize all germplasm used within the Wang lab by generating a new and unified entry code, paired with entries in an ever-growing, easily accessible database. Previous projects within the lab have all used proprietary codes for seed genotypes that were only meant to be used temporarily, which complicates analysis of data between experiments. Additionally, seed amplified from a contaminated batch of seed could expose issues with previous seed collections, to which we would want a centralized location to take note.

The SDM is a simple series of tools that can be accessed via the ShinyR application that allows users to add entries into the database and explore existing entries. With each new entry into the database, a new code is generated (*TAG####...*) by incrementing the previous code by 1. This system allows seeds belonging to an experiment to be grouped and organized chronologically. To learn more about the entry codes, please view the README on Github!

We hope that by using this tool, we can achieve maximum efficiency as a lab!

Database Format

The database is organized into 10 columns, which are described in the following table:

Column Name	Description
Code	The newly generated entry code for the entry
Accession	Plant common accession name
Source	Internally produced, or externally received?
Prox_Source	Name of location from which seed was retrieved (if <i>externally</i> sourced) <u>OR</u> entry code from which the seed was produced (if seed was <i>internally</i> amplified)
Add_ID_1	Additional seed identification
Add_ID_2	Additional seed identification
Species	Scientific name of plant species
Researcher	Wang lab researcher whom the seed is relevant to (Purdue career ID)
DoE	Date of entry into the database
Desc	A description of the seed, project, or any important information



[Download database \(.csv\) \[6 entries\]](#)

Select search type:
☒ Simple search ☐ Range search (by code) ☐ Multi-search (by code) ☐ Group search

Showing 1 to 6 of 6 entries

	Code	Accession	Source	Prox_Source	Add_ID_1	Add_ID_2	Species	Researcher	Harvest	DoE	Desc
1	OTTER001	IR64	External	USDA ARS DBNRR	GOSR 301401		Oryza sativa	schafer12	2024-09-1	2025-03-31	Used for an exp...
2	OTTER002	IR64	Internal	OTTER001	Exp_Genotype_01		Oryza sativa	schafer12	2025-04-1	2025-03-31	Seed produced i...
3	OTTER003	Jefferson	External	Cornell University	GSOR 000000	PI000000	Oryza sativa	schafer12		2025-04-18	This is a rice ...
4	OTTER004	Scarlett	External	USDA ARS DBNRR	PI111111		Oryza sativa	schafer12		2025-04-18	Testing group i...
5	OTTER005	IR64	Internal	OTTER002	Exp_Genotype_01		Oryza sativa	schafer12		2025-04-18	Grouping test...
6	OTTER006	ZEBU	External	USDA ARS DBNRR	PI222222	GSOR 213423	Oryza sativa	schafer12		2025-04-18	Group testing...

Show 10 entries

Previous 1 Next

Sign in

Enter username

Enter password

Authenticate

Username

Password

It is not always likely that the same information is available with all seeds we store in the lab. Many seeds we receive or amplify will have different naming schemes for plant identification. Because of this, the format of the database was designed to be extremely flexible and mainly requires that **one ID is known and that the proximal source is specified**. Information on all the entry requirements are detailed in the [Register section](#) in this manual.



Accessing the Application and Database

The database is deployed and available online at <https://sdm.up.railway.app/>. By accessing this link, you can explore the current, up-to-date information present in the database. However, modifying and adding data into the database **requires that you become an authorized user**. The following subsections below outline the steps required to become an authorized user.

Adding new users into the database

New users are added to the database manually on the backend, which requires access to the GitHub files and the Railway account where the SDM application and database are hosted. Please reach out to the owner for assistance accessing these files, or assistance becoming an authorized user.

New users are created through the use of the **add_user()** function within the **RailwayScripts.R** file within the SDM project files. The new user must supply a username and password that will be stored within the Railway database.

Please note: the username and password cannot be changed later.

Logging into the database

To be able to register new seeds and update any information, you must first log into the database using your username and password. Following a successful login, you will automatically be logged out after 10 minutes. Refreshing the application will also log you out.

Login

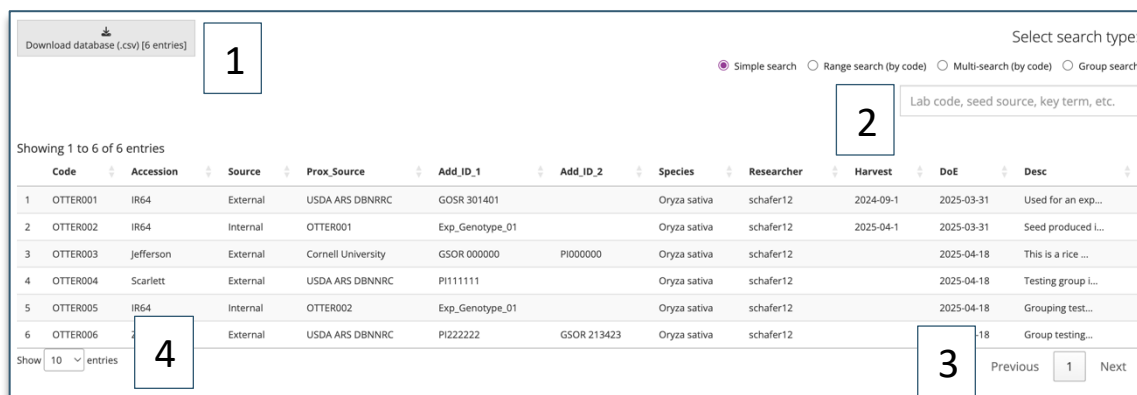
Enter username Enter password Authenticate

Username Password

Navigating the Database

The database will be displayed in the center of the application upon opening the application and will display entries chronologically by entry code. The ordering of the table can be quickly altered simply by **clicking the desired column**, to which it will reorder all lab entries in alphabetical or numerical order (based on the data present in the column). The number of entries visible on one screen can also be adjusted by **changing the number in the drop-down box** in the bottom left corner of the database. Additionally, you can cycle between entries in

the database by using the **buttons in the bottom right corner** of the database. If you wish to subset the database viewer to specific keywords, the **search box in the top right** can be used to quickly filter the database. Filtered results of the database can also be saved as a copy to your personal device by simply clicking the **download buttons** in the top left corner of the database.



Download database (.csv) [6 entries]

Select search type:
☒ Simple search ☐ Range search (by code) ☐ Multi-search (by code) ☐ Group search

Lab code, seed source, key term, etc.

Showing 1 to 6 of 6 entries

Code	Accession	Source	Prox_Source	Add_ID_1	Add_ID_2	Species	Researcher	Harvest	DoE	Desc
1 OTTER001	IR64	External	USDA ARS DBNRRRC	GOSR 301401		Oryza sativa	schafer12	2024-09-1	2025-03-31	Used for an exp...
2 OTTER002	IR64	Internal	OTTER001	Exp_Genotype_01		Oryza sativa	schafer12	2025-04-1	2025-03-31	Seed produced L...
3 OTTER003	Jefferson	External	Cornell University	GSOR 000000	PI000000	Oryza sativa	schafer12		2025-04-18	This is a rice ...
4 OTTER004	Scarlett	External	USDA ARS DBNRRRC	PI111111		Oryza sativa	schafer12		2025-04-18	Testing group L...
5 OTTER005	IR64	Internal	OTTER002	Exp_Genotype_01		Oryza sativa	schafer12		2025-04-18	Grouping test...
6 OTTER006		External	USDA ARS DBNRRRC	PI222222	GSOR 213423	Oryza sativa	schafer12		-18	Group testing...

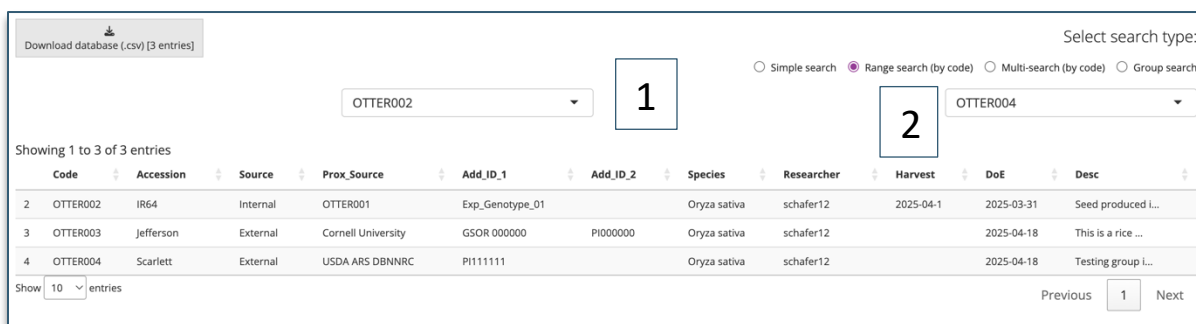
Show 10 entries

Previous 1 Next

Screenshot of the database in the application. [1] Download button used to save a copy of the current filtered database. [2] Search bar used to filter the database. [3] Buttons used to cycle through pages in the database. [4] Drop-down box for changing number of rows shown.

Search Queries

By default, the table on screen will show the complete, unfiltered database containing all seed information available at that point. To filter down to fewer results, the search features in the top right corner of the table can be used to reduce the number of rows on screen. The **simple search** will show only rows that contain specified keywords, which can include and entry codes, accession names, researcher username, etc. The **range search** is a more advanced search that will return all rows between two specified entry codes (see below screenshot). The **multi-search** allows users to select specific entries to visualize on the database. Lastly, the **group search** will filter the database for all entries that have been assigned to germplasm set groups (see Create germplasm sets).



Download database (.csv) [3 entries]

Select search type:
☐ Simple search ☒ Range search (by code) ☐ Multi-search (by code) ☐ Group search

OTTER002

OTTER004

Showing 1 to 3 of 3 entries

Code	Accession	Source	Prox_Source	Add_ID_1	Add_ID_2	Species	Researcher	Harvest	DoE	Desc
2 OTTER002	IR64	Internal	OTTER001	Exp_Genotype_01		Oryza sativa	schafer12	2025-04-1	2025-03-31	Seed produced L...
3 OTTER003	Jefferson	External	Cornell University	GSOR 000000	PI000000	Oryza sativa	schafer12		2025-04-18	This is a rice ...
4 OTTER004	Scarlett	External	USDA ARS DBNRRRC	PI111111		Oryza sativa	schafer12		2025-04-18	Testing group L...

Show 10 entries

Previous 1 Next

Example of a range search. [1] Lower end of the search is entry code OTTER002 and [2] the upper end is OTTER004.



Downloading the Database

To download the database, simply click the **Download dataset (.csv)** button, and a new window will appear on screen prompting you for a location to save the file to. If no search query has been completed, the download button will download the entire, unfiltered database by default.

Tip: you can download a filtered version of the database simply by completing a search query and then clicking the download button!

The Toolbar

The left-hand side of the application contains the Toolbar, which is used to navigate to different functions of the SDM. All four functions of the SDM tool are accessible using the buttons on the toolbar. Additionally, a simplified description of the function will be shown under the **Operation Description** header. By default, the *Register new seeds* function will be selected.

What would you like to do?

- ☒ Register new seeds
- ☐ Register multiple seeds (template upload)
- ☐ Add a note
- ☐ Organize: create germplasm sets
- ☐ Organize: remove germplasm sets

Operation Description

Register new seeds by providing accession name, seed source, associated IDs, species, relevant researcher, harvest date information, and a short description of the seed entry. Running this function will automatically generate a new lab code for the new entry (WL####). Please make sure to double check correct information before submitting data!

The Operations Deck

Located directly below the database (you may need to scroll down) is the operations deck, where all user inputs for functions are done. This section of the application will change depending on the function that is selected in the toolbar. The text boxes, buttons, and drop-down tabs featured in this section are designed to make entering new data very user friendly but are functionally identical to manually programming in R. Notes and other important information are also present within the **note text** in the bottom right corner of the operations deck. As always, follow the guidelines of this note when entering new data, exporting query results, and adding information to the database.

Note: Examples for what to enter for each function parameter are shown in the Script Functions section below.

Note: The operation deck will only display the login page until users have logged in.

Batch seed registration

Upload a template file

Browse...

No file selected

Please provide a description for all entries*

Seed info, project, etc.

Submit batch registration

Note: For this function to work properly, you must download and fill out the template sheet. The description is filled out separately and appended to all new entries. Please be aware, this process may take a while. Go grab yourself a cup of coffee while you wait!

Function Guide

The current build of the SDM contains five total functions:

- ➔ Register new seeds
- ➔ Register multiple seeds (template upload)
- ➔ Add a note
- ➔ Organize: create germplasm sets
- ➔ Organize: remove germplasm sets

These functions are described below in detail:

Register new seeds

This function is used to register new seeds into the database, individually. When selecting this option, you will be greeted with the following view in the **operations deck** below the data table.

Register new seeds

Enter common accession name

e.g. Cup Leaf

Was seed produced by the lab (Internal) or retrieved (External)?*

Enter the proximal source*

e.g. GRIN, WL####

Enter an ID*

e.g. PI529014, SA-0944

Enter an additional ID, if applicable

e.g. PI529014, SA-0944

Select a species*

Year of harvest? (if known)

Month of harvest? (if known)

Day of harvest? (if known)

Who is the researcher?*

Please provide a description for the entry*

Seed info, project, etc.

Submit information

Note: Before submitting, please make sure all the above information is correct and free of typos. Boxes can be left empty, but it is not recommended to do so. The application will refresh upon registration. * = Required field.

Operation deck view when registering new seeds individually. The asterisk (*) present indicates that the field is required to be entered into the database. After the desired information has been filled in, you may click the *Submit information* button, which will refresh the application to show changes.

Not all seeds are created equal, and thus, information may be different between entries. For these reasons, some entries are not required for new seeds to be entered into the database. However, you may still find that some seed may not have even the required information to fill out in the database. In these events, it is up to the user to make **the best judgement** for how to enter in the required fields (such as using a common accession name for the required ID, if no IDs for the seed are known).



The table below provides an overview of each input and some examples of proper inputs, please refer to it when entering new seeds into the database.

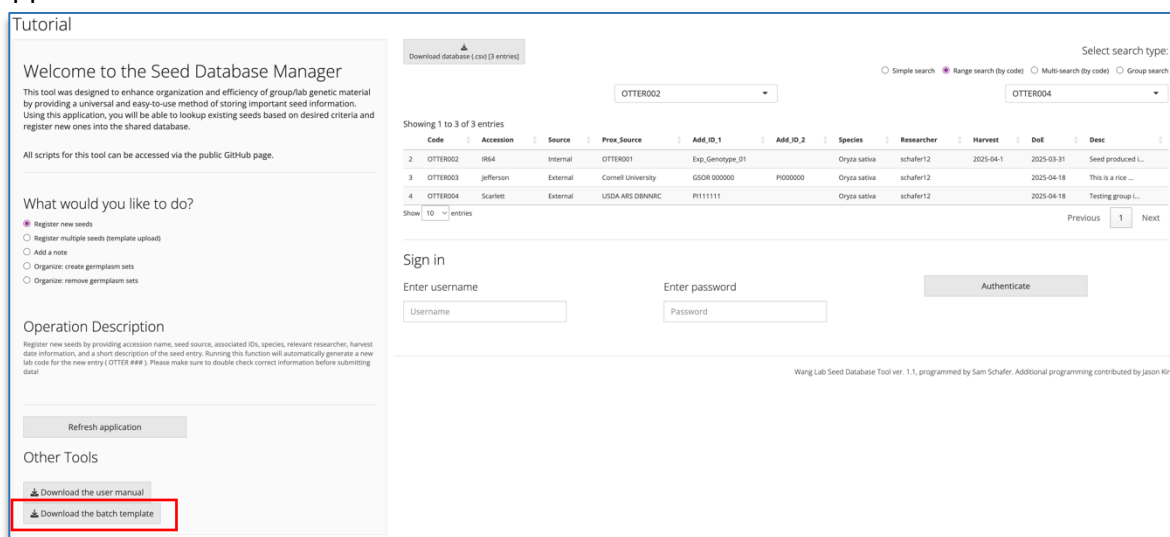
Input name	Description	Example	Required?
Accession	Plant common accession name	"Cup Leaf"	No
Source	Must either be <i>Internal</i> (seed came from our lab) or <i>External</i> (received from outside source)	"Internal"; "External"	Yes
Prox_Source	Name of location from which seed was retrieved (if <i>externally</i> sourced) <u>OR</u> entry code from which the seed was produced (if seed was <i>internally</i> amplified)	"WL0004"; "University of Arizona"	Yes
Add_ID_1	Additional seed identification	"PI123456"	Yes
Add_ID_2	Additional seed identification	"PI123456"	No
Species	Scientific name of plant species	"Glycine max"; "Oryza sativa"	Yes
Harvest year, month, day	Enter as YYYY-MM-DD format.	"2025-03-27"	No
Researcher	Wang lab researcher whom the seed is relevant to (Purdue career ID)	"schafe43"; "sjamshidi"	Yes
Desc	A description of the seed, project, or any important information	"Part of a project"	Yes
Notes	May be used to add information on individual entries, rather than the entire group. Found in the template only.	"Seed packet indicates contamination"	No

Note: The table above also applies to template uploads for batch seed registration (see guide below)

Register multiple seeds (template upload)

When registering many seeds at the same time, it is recommended to enter seeds in the database using a template upload rather than doing so individually. Essentially, this function runs multiple iterations of the *Register* function, but instead uses a user-modified template as an input parameter. The following guide outlines the steps for successfully creating and uploading the template file.

1. Download the **Batch_Template_Example.csv** file found at the **bottom left** of the application.



Tutorial

Welcome to the Seed Database Manager

This tool was designed to enhance organization and efficiency of group/lab genetic material by providing a universal and easy-to-use method of storing important seed information. Using this application, you will be able to lookup existing seeds based on desired criteria and register new ones into the shared database.

All scripts for this tool can be accessed via the public GitHub page.

What would you like to do?

- ☒ Register new seeds
- ☐ Register multiple seeds (template upload)
- ☐ Add a note
- ☐ Organize: create germplasm sets
- ☐ Organize: remove germplasm sets

Operation Description

Register new seeds by providing accession name, seed source, associated IDs, species, relevant researcher, harvest date information, and a short description of the seed entry. Running this function will automatically generate a new lab code for the new entry (OTTER ###). Please make sure to double check correct information before submitting data.

Refresh application

Other Tools

- [Download the user manual](#)
- [Download the batch template](#)

Download database (csv) (3 entries)

Select search type: ☐ Simple search ☒ Range search (by code) ☐ Multi-search (by code) ☐ Group search

Showing 1 to 3 of 3 entries

Code	Accession	Source	Prox. Source	Add_ID_1	Add_ID_2	Species	Researcher	Harvest	DoE	Desc
2	OTTER002	IR64	Internal	OTTER001	Exp_Genotype_01	Oryza sativa	schuler12	2025-04-1	2025-03-31	Seed produced L...
3	OTTER003	Jefferson	External	Cornell University	GSOR 000000	Oryza sativa	schuler12	2025-04-18		This is a nice ...
4	OTTER004	Scarlett	External	USDA ARS DBNHR	PH111111	Oryza sativa	schuler12	2025-04-18		Testing group L...

Sign in

Enter username:

Enter password:

Authenticate

Wang Lab Seed Database Tool ver. 1.1, programmed by Sam Schuler. Additional programming contributed by Jason King.

2. Open the downloaded **Batch_Template_Example.csv** file in excel. The file contains 7 columns that are to be edited with new information. Make sure to not overwrite the header column, or else this could cause issues with the application. The template includes examples of how each column should work, observe and delete these rows as needed.

	A	B	C	D	E	F	G	H	
1	Plant Name	Source (Internal or External)	Proximal Source	ID	Additional ID	Species	Researcher	Harvest	Notes
2	PD3-example c	External	University of Arizona-	PI511353-identificat	SA3778-optional space	Gossypium hirsuti, schafe43	Purdue career	2022-01-01-Date of harvest, if known	
3	PD3	Internal	WL0001-internal dat	Gh22_18-example short term identification		Gossypium hirsuti, schafe43		4/20/22	
4	Pronto	External	University of Arizona	PI529594	SA1540	Gossypium hirsuti, schafe43		1/1/22	Potentially contaminated seed based on pheno
5	Pronto	Internal	WL0003	Gh22_19		Gossypium hirsuti, schafe43		4/20/22	Plant had unusual leaf shape when grown out
6	delete all rows above except the header								



3. Once the template sheet is filled out with the desired information, return to the SDM application and click the **Register multiple seeds (template upload)** button. An upload box will now be visible in the operations deck below the database following a successful login, as well as a text box for entering in a description to be given to all new entries (see below screenshot).

Batch seed registration

Upload a template file

Browse...

No file selected

Please provide a description for all entries*

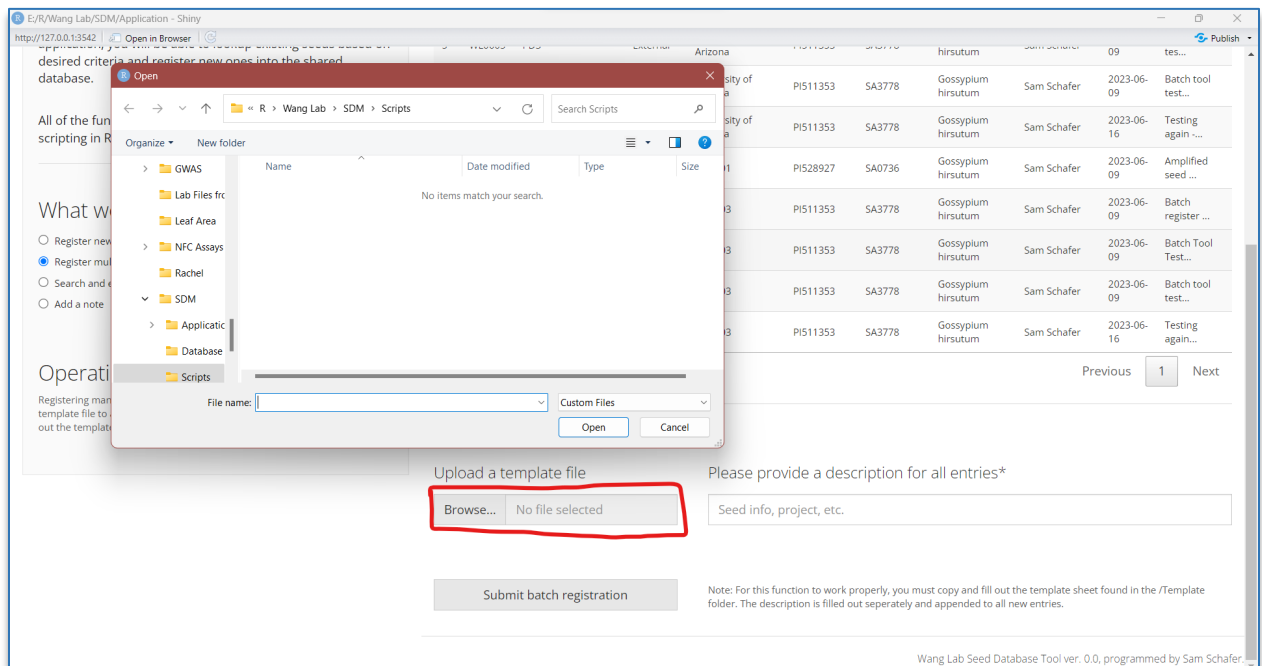
Seed info, project, etc.

Submit batch registration

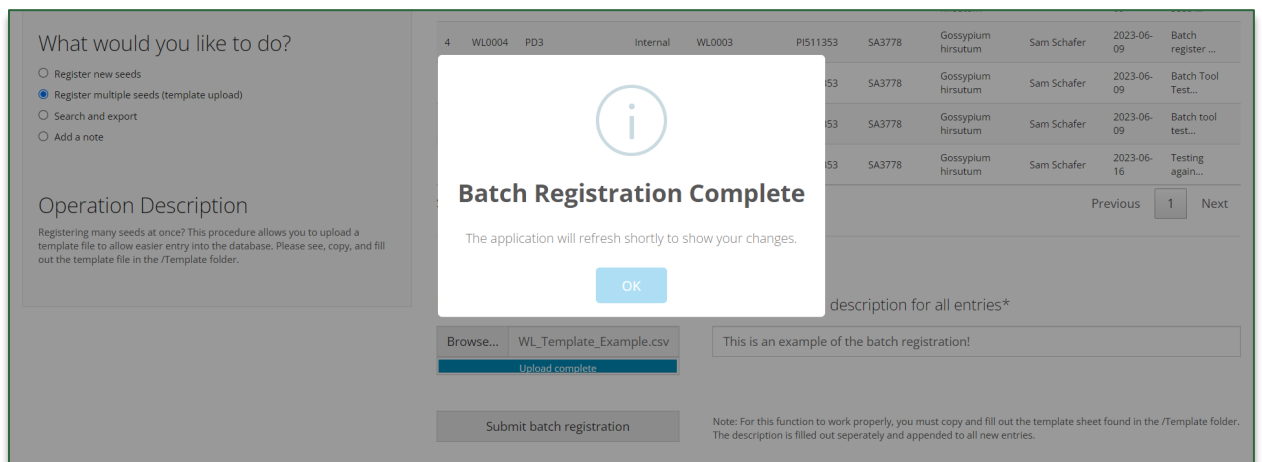
Note: For this function to work properly, you must copy and fill out the template sheet found in the /Template folder. The description is filled out separately and appended to all new entries.

Note: any text written in the description box will be applied to every new seed being added into the database. This helps identify batches of seeds easily. Individual information can be added to individual seeds later by adding a note.

- Click the **Browse...** button and navigate to the location where your completed template is saved in the pop-up window. Select your file, and it will now be uploaded onto the app.



- Finally, type in a description for the seeds into the text box. This description will be appended to **all new entries added to the database**. Once this is done, simply click the **Upload batch registration** button!



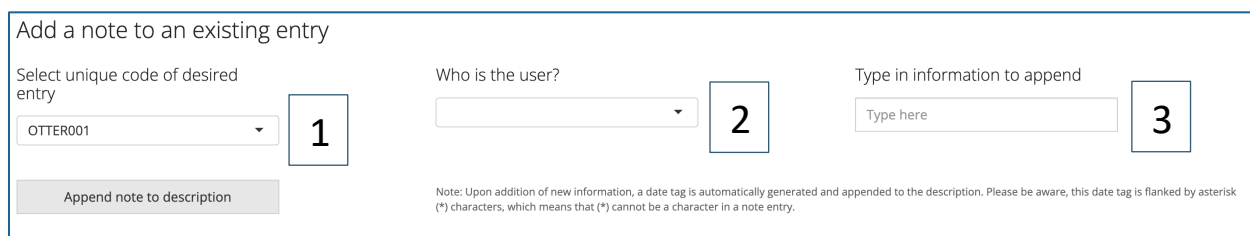
Note: the upload will only accept .csv files, so be sure to save your modified template as a .csv file.

Add a note

This function is used to append new information to the description of an existing database entry. When new information is added, an auto-generated date is appended to the end of the note, to keep track of when the note was added. To add a note, navigate to the **Add a note** button on the left-hand toolbar, and you will be able to select an entry to modify (see below screenshot).

Important note: The auto-generated date is continually updated as new information is appended to the description. Because of this, the asterisk * symbol is used by the function as an identifier for the date information and must not be used in the note.

Note: If new information is available on older seeds, it is highly recommended to append new information to both the proper database entry and the seed packet in question.



Adding notes to existing entries in the database. [1] Drop-down box is used to select an existing entry to modify, [2] the second drop-down box selects the authorized user who is making the new note, and [3] the last box is where the new note is written.

The following table gives examples of acceptable inputs for this function:

Input name	Function Description	Example	Required?
Code	The entry code for the entry that you wish to change	"WL0001"; "OTTER032"	Yes
User	User who is adding the note.	"schafe43"; "sjamshidi"	Yes
Note	The note you wish to append.	"Contaminated seed"	Yes

Organize: create germplasm sets

This function allows users to organize entries into groups of germplasm sets. By default, every entry not assigned to any groups will be placed in a germplasm set titled **ungrouped**, which can be seen by viewing the database with the **group search** setting *before* any germplasm sets are



created. To create a new germplasm set, navigate to the **Organize: create germplasm** on the left-hand toolbar and should see a new set of options appear in the operations deck below the database. From here, you will be prompted to select entries to be placed into a new germplasm collection and then specify a name for the group (see screenshot and caption below). Entries may be provided manually via the drop-down list, or may be selected as a range between two entry codes. Additionally, entries can be added to existing groups at any time by selecting from the drop-down list of existing entries. Once the proper information has been given to the fields, you may click the **Create group with entries** button.

Note: If adding many entries to one group, the process will take a long time!

Note: Entries cannot be added to ungrouped!

Note: [1] Entries can be selected manually via a drop-down box of all existing entries OR via a range between two entries in the database, which can be toggled using the buttons below the entry field. [2] Users may manually type a name for new groups OR can choose to add entries to existing groups via the drop-down box, which can be toggled using the buttons below the entry field.

Organize: remove germplasm groups

Similar to the previous function, germplasm sets may be removed if they are no longer needed or desired. To do this, navigate to the **Organize: remove germplasm sets** option on the left-hand toolbar and you will be able to select existing groups from a drop-down box in the operations deck for removal (see below screenshot).

Note: germplasm sets cannot be recovered once removed, so be careful! Entries themselves are not impacted by the creation or removal of collection groups.



Remove germplasm sets

Select an existing germplasm set:

ungrouped ▼

Remove specified group

Note: Make sure you really want to remove the germplasm set before removing! Otherwise, you have to manually recreate the set again.

Note: By default, “ungrouped” will be selected if no groups exist. This cannot be removed, even if you try!