Wang Lab

Seed Database Manager

Ver. 1.0 Manual

The ultimate guide to the Seed Database Manager

Database tool and manual written by Sam Schafer
in collaboration with the Diane Wang Lab
(https://www.dianewanglab.com/)

Wang Lab

Table of Contents

Wang Lab

Seed Database Manager Overview

Introduction

The Seed Database Manager (SDM) was designed to centralize and organize all germplasm used within the lab/group by generating a new and unified lab code, paired with entries in an ever-growing, easily accessible database. Every lab and work group have various naming conventions and proprietary codes used to refer to germplasm information, which becomes complicated when working with germplasm between multiple groups. Additionally, more organization between seed information allows for easy trace-back if issues are noted with seeds.

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 unique entry 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.

We hope that by using this tool, your group can achieve maximum efficiency!

Database Format

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

Column Name	Description
Code	The newly generated lab 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 *externally* sourced) OR lab code from which the seed was produced (if seed was *internally* amplified)
Add_ID_1	Additional seed identification
Add_ID_2	Additional seed identification
Species	Scientific name of plant species
Researcher	Researcher whom the seed is relevant to (Career ID, name, depends on setup)
Harvest	Harvest date (YYYY-MM-DD format)
DoE	Date of entry into the database
Desc	A description of the seed, project, or any important information

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.

Wang Lab

Accessing the Application and Database

The files for the tool are available at https://github.com/DingoGoesBingo/Seed-Database-Manager-GE. The tool must first be set up by your group by following the instructions in the README file, afterwards will be available via a weblink. By accessing the weblink, 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 by running the **AddingNewUsers.R** file within the Setup folder of 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.
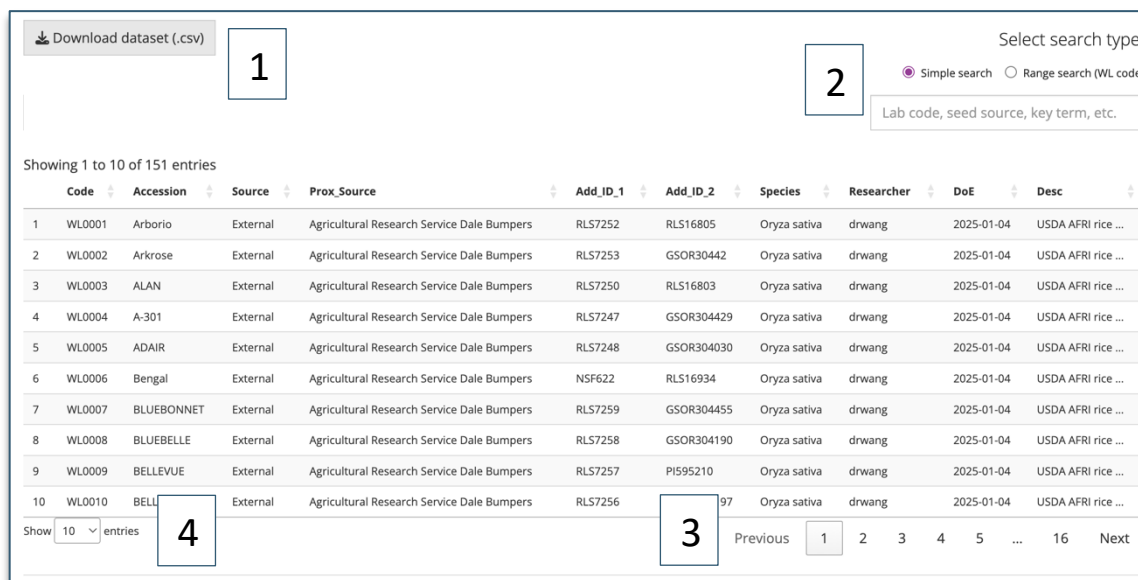
Login

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

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 lab codes, accession names, researcher username, etc. The **range and multi searches** are more advanced search types that will return a specified set of lab codes (see below screenshots).

Example of a range search. Lower end of the search is unique code WL0012 (1) and the upper end is WL0014 (2).

Example of a multi search. Each unique lab code is specified and separated by a comma. Notice how only the two specified rows are returned in the database.

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!

![Wang Lab logo]

The Toolbar

The left-hand side of the application contains the Toolbar, which is used to navigate to different functions of the SDM. All three 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.

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 or adding information to the database.

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

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

![Wang Lab logo]

Function Guide

The current build of the SDM contains three total functions:

- ➔ **Register new seeds**
- ➔ **Register multiple seeds (template upload)**
- ➔ **Add a note**

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.

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).

Wang Lab

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.
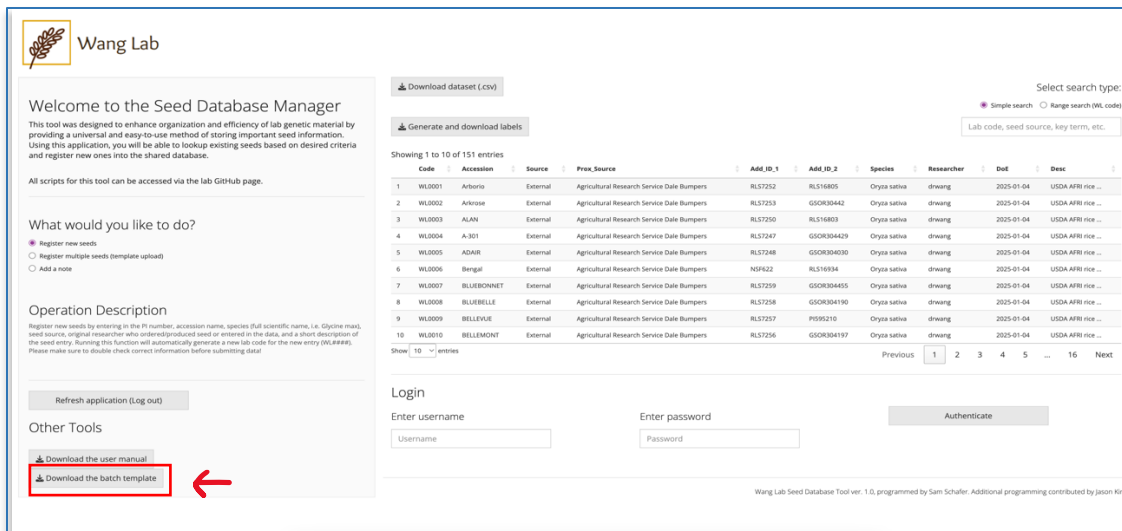
Input name	Description	Example	Required?
Accession	Plant common accession name	"Cup Leaf"	No
Source	Must either be *Internal* (seed came from our lab) or *External* (received from outside source)	"Internal"; "External"	Yes
Prox_Source	Name of location from which seed was retrieved (if e*xternally* sourced) <u>OR</u> unique entry code from which the seed was produced (if seed was *internally* 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)	[App] All fields can be left blank, but *year* is required for any date to print. If *month* and *day* are not entered, it will assume 01 for both. [Template] Can be left blank. If entered, use the <u>YYYY-MM-DD</u> format.	"2025-04-20"; "2016-01-01"	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

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** file found at the **bottom left** of the application.

2. Open the downloaded **csv** file in excel. The file contains seven columns that are to be edited with new information, with an additional **Note** column used to add information relevant to a single entry, if applicable. Make sure to not overwrite the header column, or else this could cause issues with the application.

An example of what a filled-out table looks like is shown below:

	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	Pronto	External	University of Arizona	PI529594	SA1540	Gossypium hirsutum	schafe43	1/1/22	Potentially contam
3	Pronto	Internal	WL0003	Gh22_19		Gossypium hirsutum	schafe43	4/20/22	Plant had unusual l

Note: Columns in the template are the same as the parameters for the register function. You may view examples of the register function parameter in the corresponding section above.

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).

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.

4. 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.
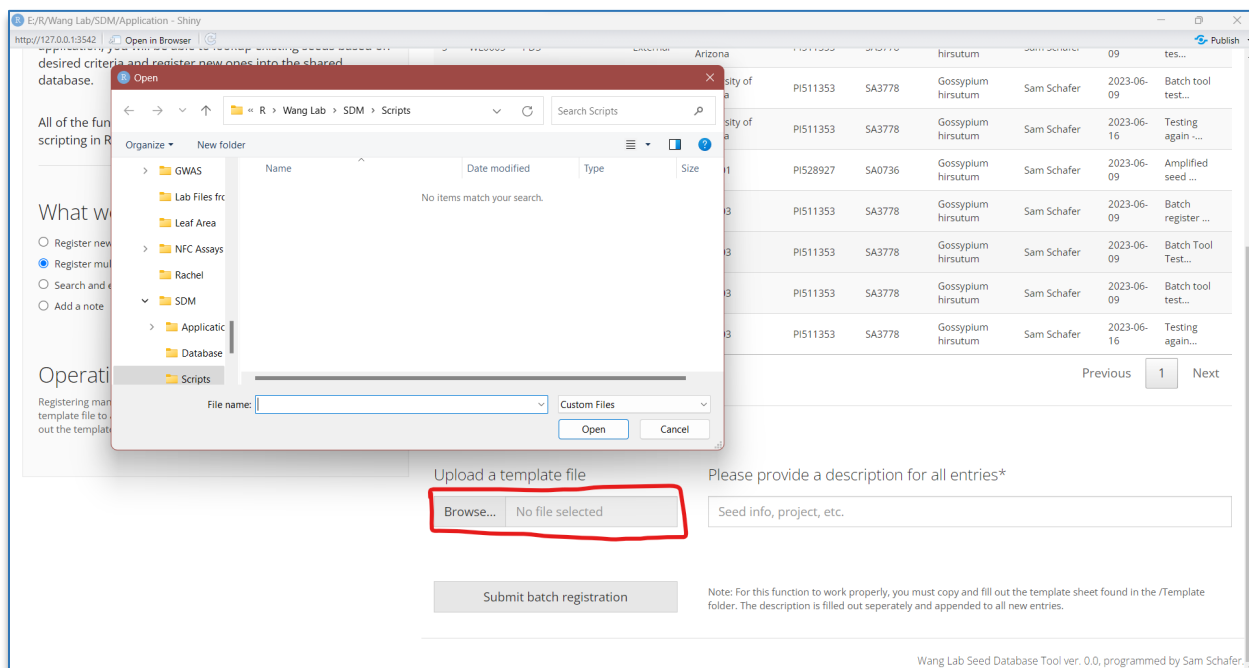
5. Finally, type in a description for the seeds into the text box. This description will be appended to **all new entries added to the database**. Once this is done, simply click the **Upload batch registration** button!
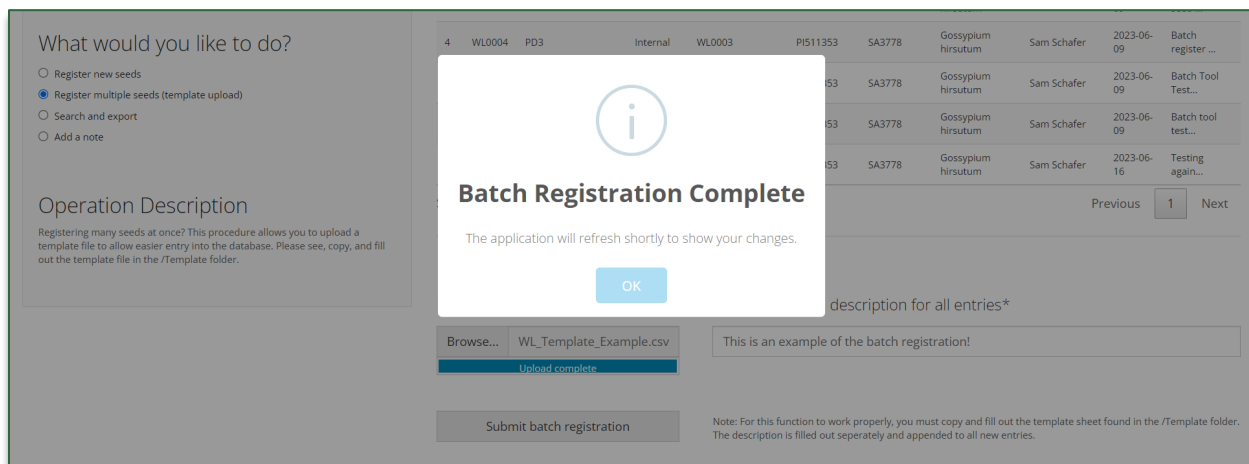
Note: the upload will only accept .csv files, so be sure to save your modified template as a .csv file.

![Wang Lab logo]

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.

*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. Please view the best practices section below.

The following table gives an overview of the parameters for this function:

Input name	Function Description	Example	Required?
Code	The unique code for the entry that you wish to change	"WL0001"; "WL0032"	Yes
User	User who is adding the note.	"schafe43"; "sjamshidi"	Yes
Note	The note you wish to append.	"Contaminated seed"	Yes