Importing Accession Data

Revision Date
November 23, 2020

Importing Accession Records into GRIN-Global
- Using the Curator Tool

In this step-by-step exercise, you will learn how to copy genebank data originally stored in a spreadsheet into the GRIN-Global Curator Tool. The premise being used here is that a genebank has been maintaining data in spreadsheets, and now intends to use GRIN-Global to manage their data. We are using sample "Vicuna Beans" data, but the same steps followed here can be used for any plant genetic resource data.

This exercise illustrates how to import genebank data by *matching your spreadsheet's column headings* with the column names used in the Curator Tool.

Note that your organization's GRIN-Global administrator establishes the Curator Tool column names, based on the organization's procedures. In some cases, you may be the administrator as well as the user of the Curator Tool.

Note	There are several methods for getting existing accession records into GRIN-Global. This exercise illustrates just one of the methods. Here, you will prepare a spreadsheet with data and then copy (drag) the data into the Curator Tool. A second method uses the Admin Tool's Import Wizard to load data from a spreadsheet (actually a .csv file) into GRIN-Global. In this second method, the Curator Tool is not involved. Both of these methods are effective for bringing many records into GRIN-Global simultaneously. The third method involves using the Curator Tool's Accession wizard to walk through inputting *one* accession record at a time; in this lesson we will not be looking at that method either.

Prerequisites:
In addition to this document, you will also need:
- a spreadsheet with the sample Vicuna Beans data. This can be downloaded from the wiki Training page:
  http://www.ars-grin.gov/npgs/gringlobal/training/exr_vicuna_data.xlsx
- spreadsheet software, such as Microsoft's Excel, google docs, OpenOffice.org, etc.
- the Curator Tool User Guide (for reference, if desired)

Directions:

Step1: Open the GRIN-Global Curator Tool

Step 2: For this exercise, we have some sample data "almost" ready to be imported. Open the file exr_vicuna_data.xlsx. The spread looks like this:

|	A	B	C	D	E	F	G	H	I	J	K	L	M	
1		Prefix	Num.	Suffix	Genus	Species		Accession Name	Origin	Maintenance Site	Is Core?	Status	Initial Material Type	Fc
2		Use initials for the prefix; ex: MR	95001	Something uinique; ex: REI	Phaseolus	vulgaris	hint: =E2&" "&F2	(leave blank)	(leave blank)	(leave blank)		ACTIVE	SD	
3			95005		Phaseolus	vulgaris						ACTIVE	SD	
4			95006		Phaseolus	vulgaris						ACTIVE	SD	
5			95008		Phaseolus	vulgaris						ACTIVE	SD	
6			95009		Phaseolus	vulgaris						ACTIVE	SD	
7			95010		Phaseolus	vulgaris						ACTIVE	SD	
8			95011		Phaseolus	vulgaris						ACTIVE	CT	
9			95012		Phaseolus	vulgaris						ACTIVE	CT	
10			95017		Phaseolus	vulgaris						INACTIVE		
11			95020		Phaseolus	vulgaris						INACTIVE		
12			95021		Phaseolus	vulgaris						ACTIVE	CT	
13			95023		Phaseolus	vulgaris						ACTIVE	CT	
14														
15														
16														

Step 3: In the Curator Tool, ensure that the **Accessions** tab is the active tab. Click the **Edit Data** button to initiate "Edit" mode.
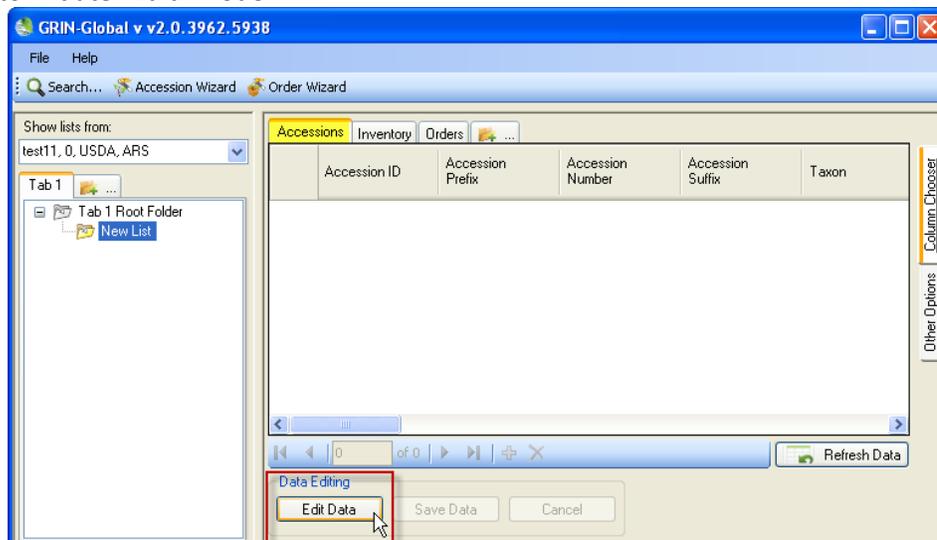

Step 4: Click the **Add new** button.

Step 5: Click the row header cell to highlight all of the row's cells.
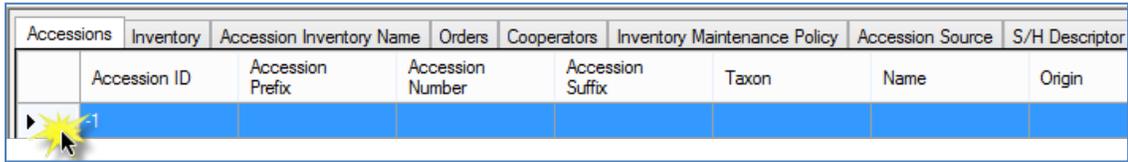

Step 6: Click in the blue area and drag the highlighted block into the Vicuna spreadsheet (from step 2) – below the existing data, say, in rows 20 or so. (Be patient – sometimes there is a slight delay before the spreadsheet accepts the data. A video on the GRIN-Global training page illustrates how to drag and drop.)
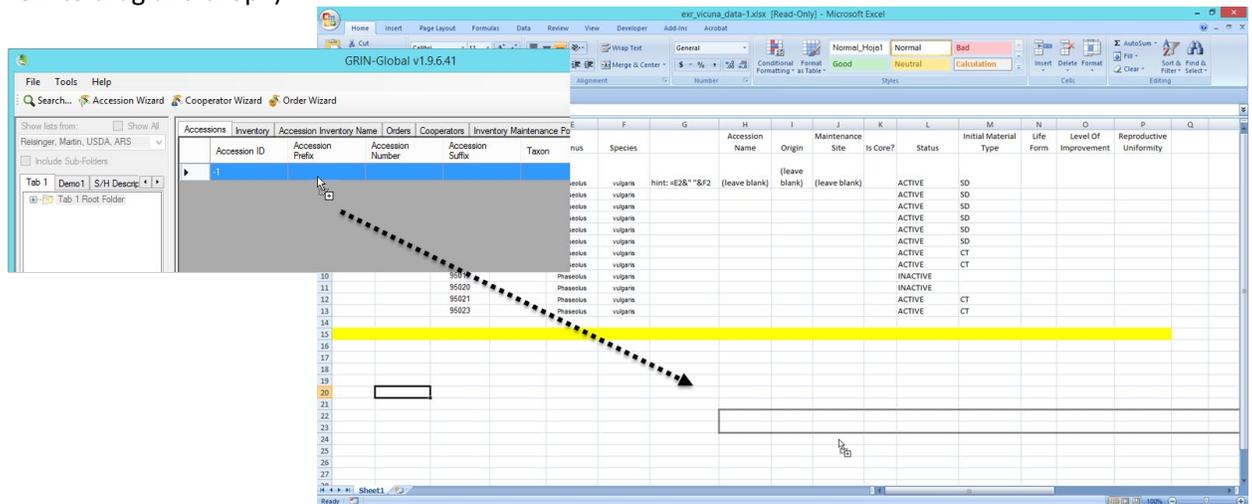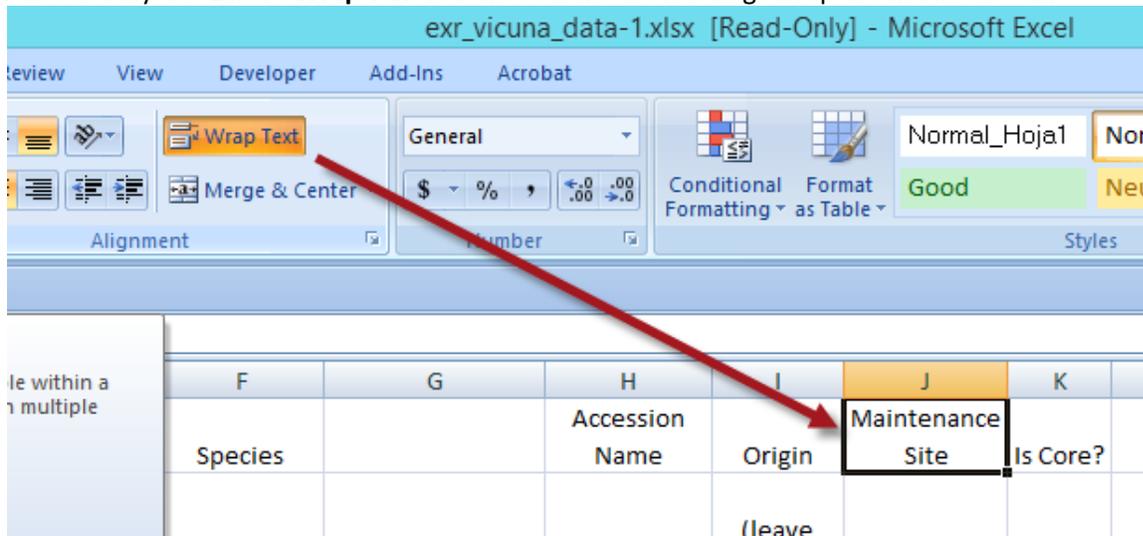

Step 7: Save the sheet; name it "**transfer**."

Step 8: Widen the columns in the **transfer** spreadsheet: click the top, left header cell; then double-click on the vertical column separator between the A and B column headings.
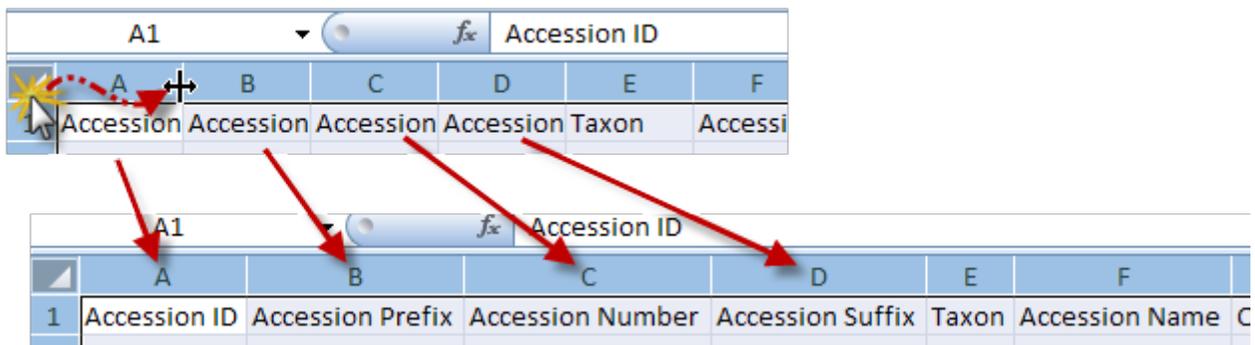
I also usually use Excel's **Wrap Text** feature to have the headings wrap around in the cell.

Step 9: In this step, you ultimately want your source data (**exr_vicuna_data.xlsx**) to have column headings that match the exact spelling of the columns from the Curator Tool. You can drag the relevant headings from the CT on top of the original data headings.

Note: In order to upload Accession data into the Curator Tool, two fields are required:

- **Accession Prefix**
- **Taxon**

Although **Accession Number** and **Accession Suffix** are optional, they are frequently used by organizations. As you create accession records, the combination of the Prefix, Number, and Suffix fields must be unique to each record.

Your new **transfer** spreadsheet currently should look similar to the following:

	A	B	C	D	E	F	G	H	I	J
1	Accession ID	Accession Prefix	Accession Number	Accession Suffix	Taxon	Accession Name	Origin	Maintenance Site	Is Core?	Is Backe
2	-1								N	N
3										
4										

Since you will not be using all of these columns, at this point, you can remove the unnecessary columns and rearrange the **transfer** sheet's columns, but you don't have to. The column order is not critical– *only the spelling is critical*.

Note that the combined fields **Accession Prefix**, **Number**, and **Suffix** comprise each record's key; the combination must be unique for each record. If your organization does not use the **Suffix** field, then the combination of the **Prefix** and **Number** fields must be unique. Similarly, if only the **Prefix** is used, then each Prefix must be unique.

So that multiple students can complete this same exercise on the same server, include your initials or some other unique text for the **Accession Prefix** and **Suffix** fields. For instance, you could use your initials, such as "MR."

	A	B	C	D	E	F	
1		Prefix	Num.	Suffix	Genus	Species	
2		MR	95001	Something uinique; ex: REI	Phaseolus	vulgaris	hint: =E:
3		MR	95005		Phaseolus	vulgaris	
4		MR	95006		Phaseolus	vulgaris	
5		MR	95008		Phaseolus	vulgaris	
6		MR	95009		Phaseolus	vulgaris	
7		MR	95010		Phaseolus	vulgaris	
8		MR	95011		Phaseolus	vulgaris	

The Taxon column is something you need to create using an Excel formula because in the original **vicuna** spreadsheet, genus and species data are in columns **E** and **F**, but the CT expects the Taxon field to have the genus and species combined.

D	E	F	G
Suffix	Genus	Species	Taxon
Something uinique	Phaseolus	vulgaris	hint: =E2&" "&F2
REI	Phaseolus	vulgaris	
	Phaseolus	vulgaris	
	Phaseolus	vulgaris	
	Phaseolus	vulgaris	
	Phaseolus	vulgaris	
	Phaseolus	vulgaris	
	Phaseolus	vulgaris	

In this scenario, since all of the accessions have the same taxon (genus + species in this case), you could simply enter *Phaseolus vulgaris* into column **G** of the spreadsheet. However, using the formula works when the genus and species data differ. The function in column G can then be copied down for the remaining records.

Complete two other columns:
- for the **Accession Suffix**, use several letters from your surname or anything to ensure unique suffixes. In the example below, both the Prefix and Suffix were altered. (By having a unique prefix and suffix, you will be ensuring that unique Accessions are added. This is important if

multiple people attempt to do this exercise on the same database.)

	A	B	C	D	E	F	G	H
1		Prefix	Num.	Suffix	Genus	Species		Accessi Name
2		MR	95001	REI	Phaseolus	vulgaris	hint: =E2&" "&F2	(leave bla
3		MR	95005	REI	Phaseolus	vulgaris		
4		MR	95006	REI	Phaseolus	vulgaris		
5		MR	95008	REI	Phaseolus	vulgaris		
6		MR	95009	REI	Phaseolus	vulgaris		
7		MR	95010	REI	Phaseolus	vulgaris		
8		MR	95011	REI	Phaseolus	vulgaris		
9		MR	95012	REI	Phaseolus	vulgaris		
10		MR	95017	REI	Phaseolus	vulgaris		
11		MR	95020	REI	Phaseolus	vulgaris		
12		MR	95021	REI	Phaseolus	vulgaris		
13		MR	95023	REI	Phaseolus	vulgaris		

The data in the **transfer** spreadsheet is ready to be copied or dragged into the Curator Tool. However, before doing so, in step 10, you will return to the Curator Tool and create a **Tab** and **List**. In the List Panel, you can establish tabs and lists with meaningful names which will assist you in organizing your data.

Step 10: In this step, using the Curator Tool, you will set up a tab and a list which will ultimately point to the records. (In step 11 you will drag the records into the Curator Tool.)

(Earlier, in step 3, you were in Edit mode. If you are still in Edit mode, temporarily exit from Edit mode: click the **Cancel** button under the Data Grid.)

In the List Panel, create your Tab and List names as desired. To rename a Tab, right-click on the tab name and select **Properties**, then rename; to rename the **NewList**, right-click on **NewList** and select **Rename**. Create names similar to the following:
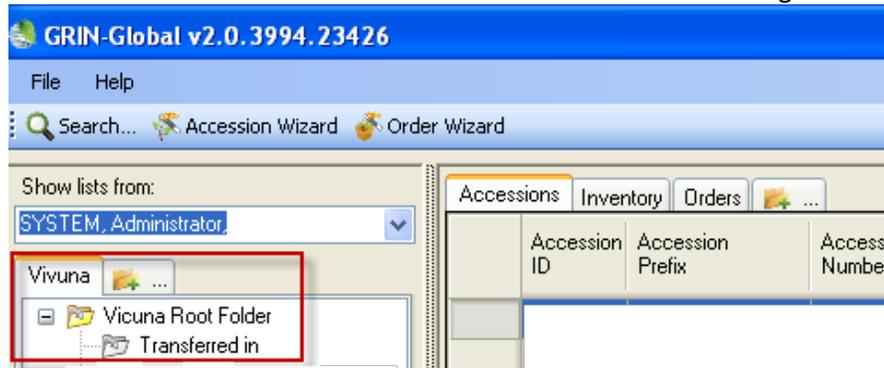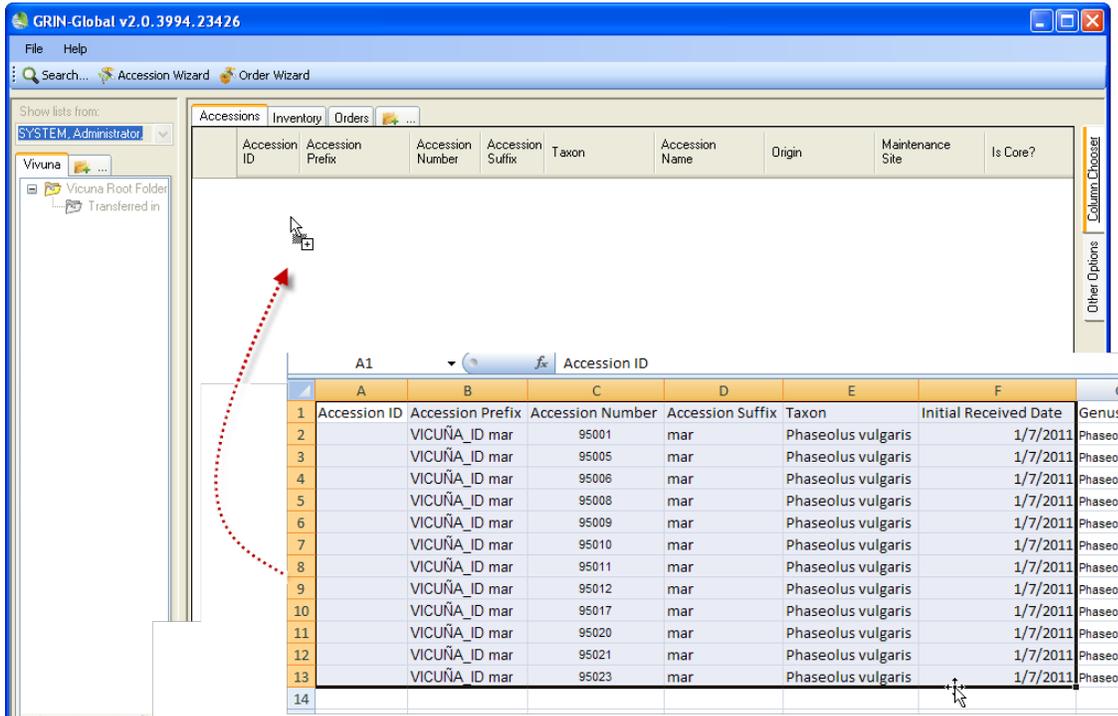
Step 11:

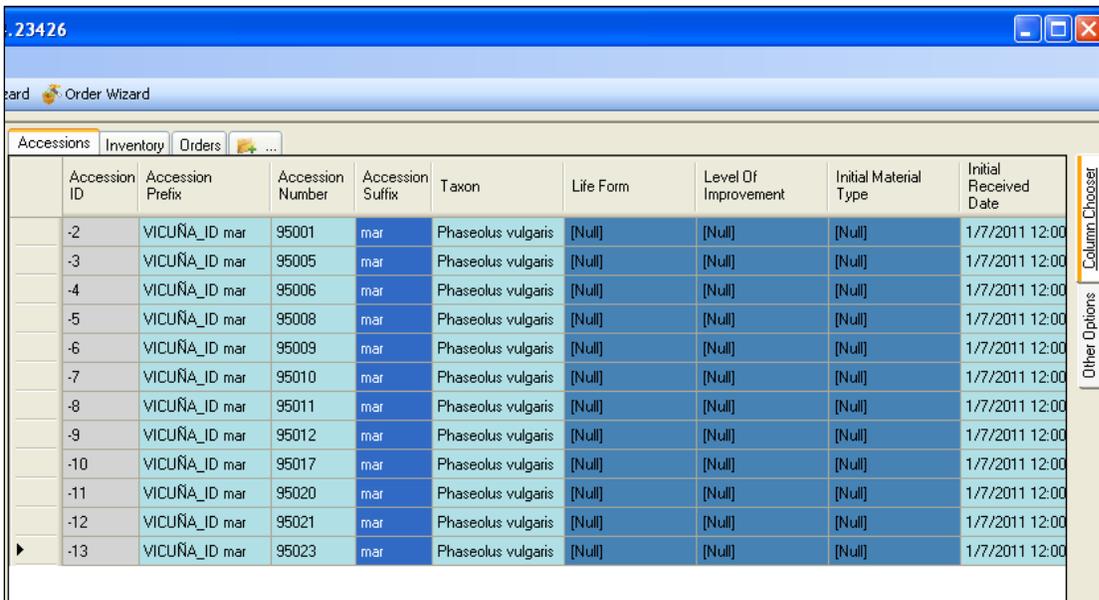Copy data from the transfer sheet to the Curator Tool.

Drag and drop or copy the block of data from the spreadsheet into the Curator Tool Data Grid. Remember:
- the Data Grid must be in Edit mode (click the **Edit** button if necessary)

The Curator Tool should look similar to the following window:

Notice that GRIN-Global temporarily assigned Accession IDs to the new records. The minus (-) signs indicate that the IDs are temporary; when the data is saved, permanent IDs are assigned.

Skip to Step 12 if your window was similar to the previous page's.

Your GG database does not have the relevant Taxonomy records needed to handle Phaseolus vulgaris If the Curator Tool window displays the following window:
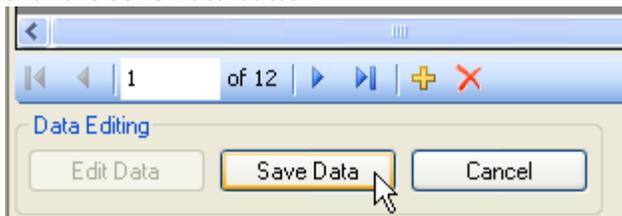

The Taxon field is a required field, expecting a match for the records being transferred with the existing Taxon data in the GG database. The message is indicating that the Taxon record for Phaseolus vulgaris doesn't exist.
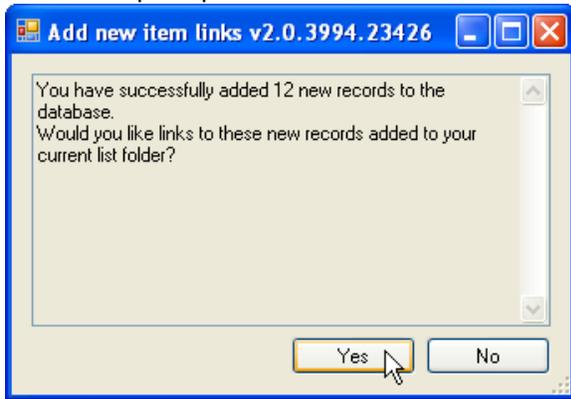
Step 12:
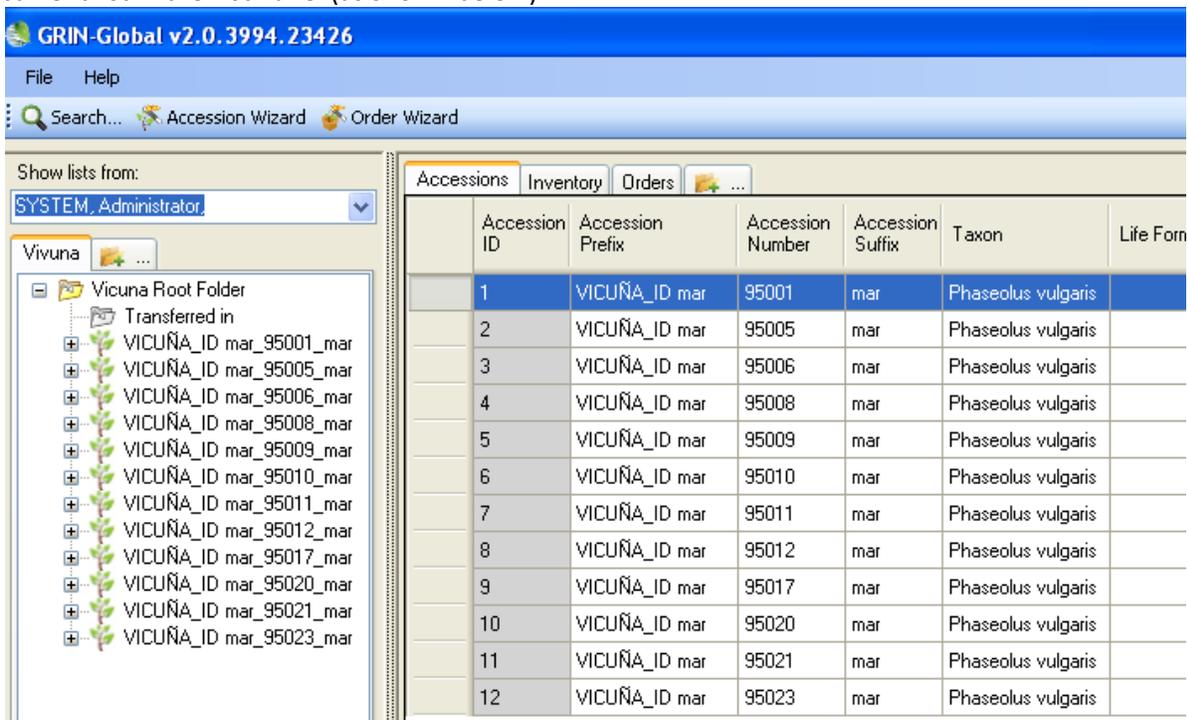Save the imported data now in the Curator Tool.

Click the **Save Data** button.

You will be prompted to add the links to these records; click **Yes**.

Note: If you had clicked **No**, the records would be added, but pointers would not be created under the current list in the List Panel (as shown below).

Other Data

Several Concluding Notes

In this exercise, you added accession records to the Curator Tool from an existing spreadsheet. To do this, you matched your spreadsheet headings with the Curator Tool's column headings.

In some cases the data stored in an organization's source files may not be in the same exact format as expected by GRIN-Global. For example, you saw in this exercise that the GRIN-Global **Taxon** field

consisted of more than one data field; the source data consisted of genus and species. In the exercise you combined source data from two cells into one.

The accession records depend on existing taxonomy information to be in the system. If the accession records have a reference to taxonomic data ("Taxon") not already in GRIN-Global, you must first add the taxonomy records before you can add the accession records. (Taxonomy uses three interdependent dataviews to relate the taxonomy family, genus, and species data.)

If you have time, you might try repeating this exercise. If you do, will need to use different Accession Numbers or Accession Prefixes, otherwise you will be attempting to add duplicates (which the Curator Tool does not allow). We made the assumption that these records did not already exist in the Curator Tool. If they had, you would have received an error message indicating that these records were duplicates of existing records. In real life, as a standard procedure, or when you are uncertain that the records exist or not, first search in the Curator Tool before attempting to add new records. (We didn't show how to use Search in this exercise in order to simplify the exercise.)

Obviously, we copied just a few fields to simplify this exercise. We could have had a more elaborate exercise and transferred other accession fields, but ideally you understand the basics and can do that.

Earlier we mentioned that there are two other methods for importing accession data into GRIN-Global. One method involves the GRIN-Global Admin Tool, which not every user has access to. That method involves several additional tables and fields and is a handy method for importing data. The second method is in the Curator Tool. The Curator Tool has an Accession Wizard which makes it possible to add not only accession data but also other related dataviews. However, the wizard doesn't handle the importing of many accession records from a spreadsheet as we did in this exercise.

Appendix: Revisions

November 23, 2020
- corrected link to the CT User Guide

Original October 2, 2015
-