

Checklists – Adding species.

This document assumes you have already created the metadata for a checklist in OpenHerbarium but have not yet added any names to it. The instructions for adding names to a list that does have species in it are similar.

Checklists in OpenHerbarium are for use with scientific names.

There are basically three ways to add names to a species checklist:

- Add names individually.
- Batch upload a csv file of names.
- Add names found by searching OpenHerbarium for species that match its geographic and taxonomic selection criteria.

The first steps are the same for all three approaches.

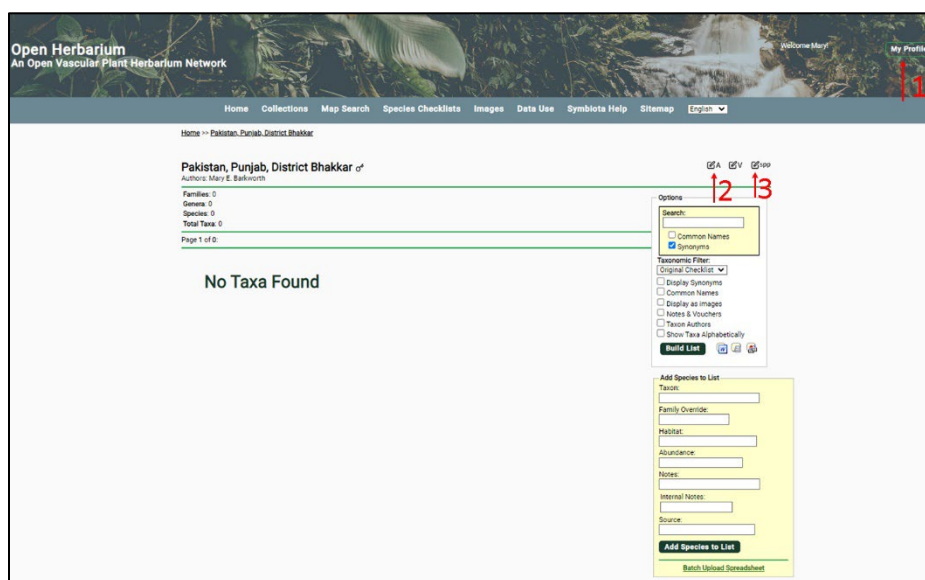


Figure 1. Opening a private checklist

1. Open the checklist. At present your checklist is private, meaning it can only be viewed by its editors. To open it, go to “My Profile” (see Fig. 1, **Arrow 1**), then look under the “Specimen checklists” tab. It will show you all the checklists you have permission to edit. Only individuals listed as editors for a checklist can add or delete names in it. Click on its name. To assist in preparing this document, I have created another district checklist for Pakistan. Its name is Pakistan, Punjab, District Bhakkar. Once I have completed this document, I shall make it public.

By default, the link will open the *Admin* tab for the checklist but, **to add species**, click “Return to checklist” just below the banner (Figure 2). This will bring up the screen shown in Figure 3. (For more information about *Admin* functions, see **Checklists – Administration**).

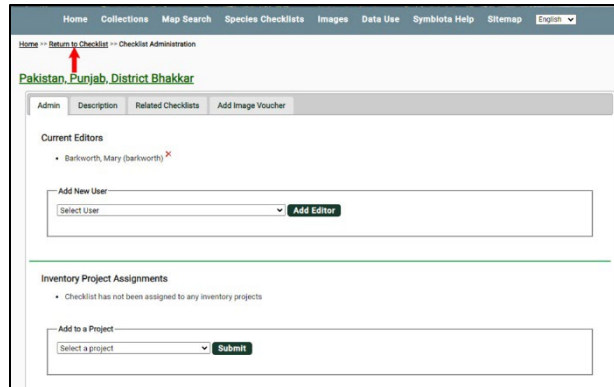


Figure 2 "Return to checklist" link.

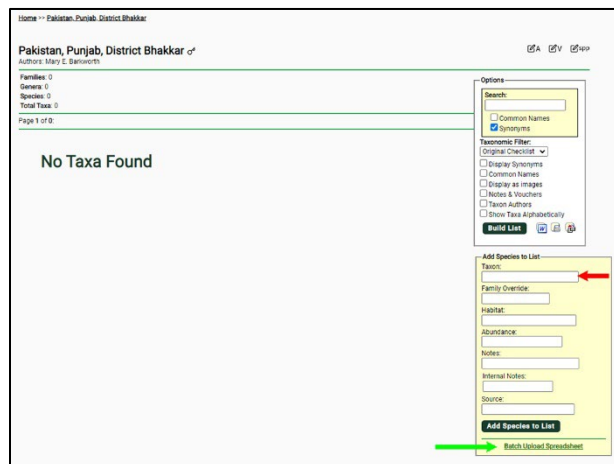


Figure 3. Adding names to a checklist.

Adding names individually

1. Start typing a name into the *Taxon* box (Figure 3 Red arrow). As you type, names in OpenHerbarium's taxonomic table will be suggested. Click on the one you wish to add. Note that you must not include the authors of a name, only the name itself.

If no names are suggested, there are two possibilities:

- you have not spelled it correctly or
- the name has not been entered into the taxonomic table yet.

Check your spelling first; we all make mistakes. If that is not the problem, send the name to [Mary Barkworth](#) together with its family (if you know it). I shall investigate it and, if appropriate, add it to the tables. If it is not a valid name, I shall email you a brief explanation of why I have not added it and the probable acceptable name.

- a. Family: do **not** put anything in here. OpenHerbarium will enter the currently accepted family name. We need to encourage and help the next generation of field botanists to become familiar with the current interpretation of families.
- b. Habitat: Optional, useful if developing a teaching checklist that is intended for use in a class which requires learning the usual habitat of the species on the list.
- c. Abundance: Optional. Again, might be useful in a teaching checklist "Common weed", "Disturbed areas", "In Pinus forests".

- d. Notes: Optional. Pretty well anything can go in here but try to be consistent. For instance, you could decide to insert the lifeform of the taxon or whether it is cultivated or wild, or “cultivated, ornamental”, “cultivated, food”, “wild, food”.
- e. Source. This is valuable information to add. It could be anyone of a specimen’s Catalogue Number, a doi (if from a published paper), a publication citation or a GBIF identifier.

Uploading multiple names

1. To add several names at once, first create a csv file with at least one field, *Sciname*, and up to 5 others. The five other fields permitted are listed in Figure 4. *Family* has a line through it because OpenHerbarium will complete it automatically, based on current usage. The easiest way to create a csv file is to create a spreadsheet and then save it as a csv file.

Pakistan, Punjab, District Bhakkar

Authors: Mary E. Barkworth

Checklist Upload Form

Checklist File: No file chosen

Taxonomic Resolution: ▼

Input file must be a CSV text file containing the following columns. Column order does not matter, though the first row should contain column names in accordance with the names in bold listed below. Note that Excel spreadsheets (xlsx) can be saved as a CSV file via the "Save as..." option.

- **sciname** (required)
- ~~family~~ (optional)
- **habitat** (optional)
- **abundance** (optional)
- **notes** (optional)
- **internalnotes** (optional) - displayed only to editors
- **source** (optional)

Figure 4. Field (column) names to use in spreadsheet for uploading multiple names

The only required field (column) is *sciname*. This is the scientific name without the authors. All the other columns are optional. They are explained below.

- a. Family: do **not** put anything in here. OpenHerbarium will enter the currently accepted family name. We need to encourage and help the next generation of field botanists to become familiar with the current interpretation of families.
- b. Habitat: Optional, useful if developing a teaching checklist that is intended for use in a class which requires learning the usual habitat of the species on the list.
- c. Abundance: Optional. Again, might be useful in a teaching checklist “Common weed”, “Disturbed areas”, “In Pinus forests”.
- d. Notes: Optional. Pretty well anything can go in here but try to be consistent. For instance, you could decide to insert the lifeform of the taxon or whether it is cultivated or wild, or “cultivated, ornamental”, “cultivated, food”, “wild, food”.

- e. Source. This is valuable information to add. It could be anyone of a specimen's Catalogue Number, a doi (if from a published paper), a publication citation or a GBIF identifier.

Here is a sample file in which only three columns are used:

Sciname	Habitat	Source
Pistia stratiotes	Aquatic	10.5829/idosi.wasj.2013.28.06.13824
Eichhornia crassipes	Aquatic	10.5829/idosi.wasj.2013.28.06.13824

2. To upload your file, click "Batch You can upload several different files, one after the other. Names will be added only if they are not already in the checklist.
If your checklist includes names that are misspelled, OpenHerbarium will reject them.
If you try to upload names that are not in OpenHerbarium's taxon tables, it will reject them.
After uploading a checklist, you will be seen a summary of the results. You will also be shown a list of the names that were rejected and the reason they were rejected.
3. Names that are rejected because they are not in OpenHerbarium's taxon tables, you will see a table listing them. First check their spelling. A good place to do this is at Plants of the World Online (POWO; <https://powo.science.kew.org/>). Try typing the name into POWO. As you do so, it will suggest possible names. Perhaps this is helpful in locating typing errors. It is also possible the name is not in the Taxon tables. If you think that is the case, , send the name to [Mary Barkworth](#) together with its family (if you know it). I shall investigate it and, if appropriate, add it to the tables. If it is not a valid name, I shall email you a brief explanation of why and the probable acceptable name.

WRITING A PAPER OR THESIS?

If you are preparing a paper or thesis that will include tables listing information by species, consider creating a csv file of the species list and uploading it to create a private checklist.

Adding names from [GBIF](#)

GBIF is the global aggregator of biodiversity occurrence records. Currently, it provides access to more than 2.5 billion records. It accepts many kinds of record (preserved specimens such as herbarium and museum specimens, fossils, freeze dried plates of fungi or bacteria, microscopic slides, images, audio files, material samples, etc.) of all kinds of organisms (viruses, bacteria, archaea, algae, fungi, plants, animals, etc.). The data can be downloaded as csv files. Because of the many different organisms involved, the diversity of record types, and the many users that GBIF serves, each record includes many more fields than OpenHerbarium and far more than needed for a simple checklist.

GBIF is also careful to keep the taxon name used by the collection submitting a record, but it will also show what name it gives it considers the acceptable name for the taxon. It relies on the Catalogue of Life and other major global groups for determining which names are accepted and which are synonyms.

Because it is funded by national governments, it only accepts country names accepted by the United Nations. All that by way of background. This impacts Somaliland but not Pakistan. A problem for both areas is that GBIF's geographic files are not up to date. The problem may lie with the country involved, not GBIF.

To add names from GBIF to a checklist:

1. Enroll for a GBIF account. click the *Signin* button (Figure5, **Arrow 1**) to get started. There is no charge for having an account. The figure shows that I am signed in. When you open GBIF, the location for creating an account will be where Arrow 1 points.
2. To find the records you need, click "Occurrences" (**Arrow 2**). This will bring up Figure 6.

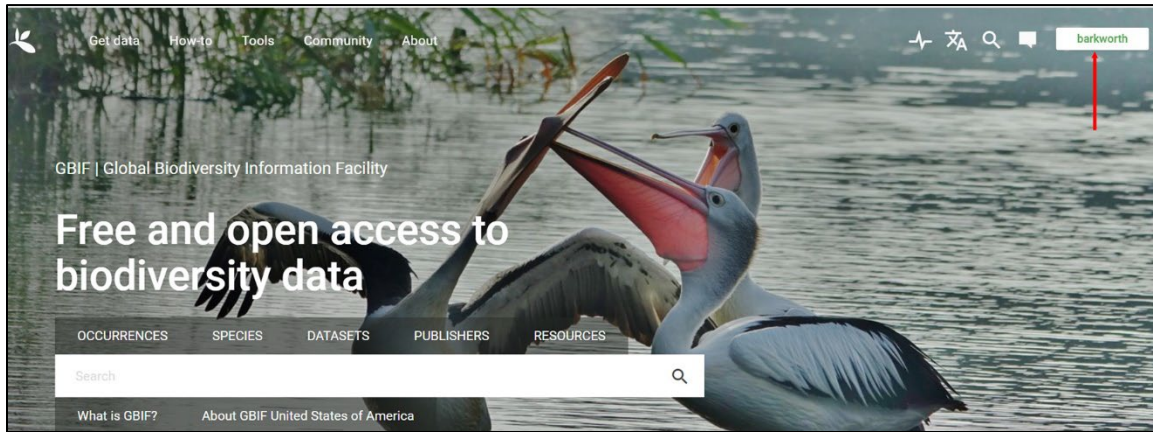


Figure 5 GBIF home page

3. Selecting records. Clicking the *Occurrences* link will bring up a screen like that shown in Figure 6.

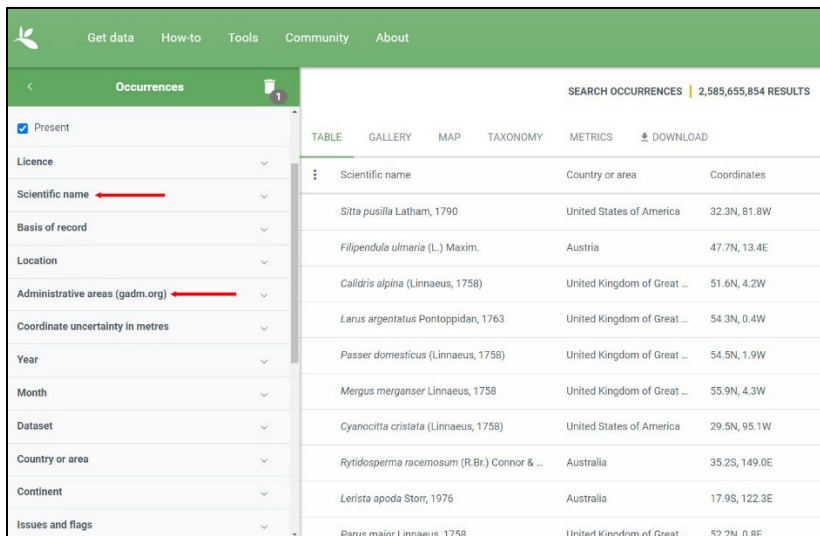


Figure 6. Selecting occurrence records from data in GBIF

The arrows point to the two sets of choices that are needed to generate a checklist of taxa present in a specific area, in the example used in the document, taxa present in District Bhakkar in Pakistan.

- a. *Scientific name(s)*. This must be selected from the names in GBIF's taxonomic backbone. It will suggest names once you start typing.

You can search for more than one name and for higher level taxa. Start typing the name of the first one, hit enter, and then enter the name of the next one, etc.

The biggest problem in searching for upper level taxa is often determining the rank used by GBIF (which takes its classification from the [Catalogue of Life](#), (CoL)) because there is often disagreement on names and ranks. The Catalogue of Life also emphasizes monophyletic groups.

Selecting the taxonomic group can be difficult, particularly at higher taxonomic levels, because there is often disagreement on their names and ranks. The Catalogue of Life emphasizes monophyletic groups so groups such as “algae”, “archegoniates”, “gymnosperms” and “dicots” are not recognized even if written in Latin. They persist because they are convenient but evidence (from studies or micromorphology, anatomy, paleobotany, and study of life cycles, as well as molecular analyses) shows that each of these groups is poly- or paraphyletic. *Plantae* includes all the ancestors of plants, not just land plants. This means it includes some algal groups (*Rhodophyta*, *Chlorophyta*, *Charophyta*, and *Glaucophyta*) and a fossil group, (*Langiophytrophyta*).

I have listed below some traditionally recognized groups and the name or names corresponding to them in Catalogue of Life and hence, GBIF. OpenHerbarium uses a mix of systems. If you have questions about a particular group or groups, send it to [me](#).

Traditional name	Current (2023-09-21) Catalogue of Life Names name(s)
Algae	Ochrophyta, Cyanophyta
Bryophytes	Bryobiotina (a subkingdom of Plantae); CoL does not provide name(s) for other subkingdom(s) which makes this name almost useless; it includes the Anthocerotophyta, Bryophyta, Marchantiophyta, and the fossil division Langiophytrophyta.
Archegoniates	Bryobiotina, Polypodiopsida
Vascular plants	Tracheophyta
Fern allies	Lycopodiopsida
Ferns	Polypodiopsida
Pteridophytes	Lycopodiopsida, Polypodiopsida
Gymnosperms	Cycadopsida, Ginkgoopsida, Gnetopsida, Pinopsida
Angiosperms	Liliopsida, Magnoliopsida
Monocots	Liliopsida
Dicots	Magnoliopsida

- b. *Area*: GBIF offers three approaches for filtering by area. They are listed below in the order used in the panel.

Location: This filter allows filtering by marking an area on a map. The area can be a rectangle or a polygon. The map has no boundary lines. Two alternative approaches are listed below the map. One allows specifying an area by giving the geographic coordinates of four corners, the other by a WKT or JSON file. OpenHerbarium can be used to generate a WKT file.

Administrative areas: An administrative area is a political unit. GBIF recognizes four levels, but this filter term is only used for levels *below* the level of country. GBIF obtains the geographic files it used for determining the geographic boundaries of administrative levels from gadm.org. I do not know how GADM obtains them. They are not always current. For example, the GADM choices for Pakistan continue to list “Federally Administered Tribal Areas” and includes older District boundaries, such as Chitral and Dir, rather than relatively recently recognized districts such as Lower Chitral, Upper Chitral, Lower Dir, and Upper Dir.

Country or area: This is the place to search for records from a country or an area formally recognized by the United Nations. Palestine is recognized as a non-member observer, so its boundaries are included. Somaliland is not recognized. To search within it, one has to recognize the regions but using the names and boundaries set before 1990.

You can, of course, explore the other filters. The four discussed above are the most useful for creating a basic checklist.

- c. *Basis of Record*: GBIF accepts many kinds of evidence for the existence of a taxon at a given time and place. To make it easier to find a specific kind of evidence, it groups the different kinds into relatively few categories. These categories were agreed on by people in charge of large collections in many different countries. “Herbarium specimen” is not a category because herbaria may have specimens stored in liquid (“pickled”) or as freeze dried cultures, as well as the commonly encountered “dried plant material and label attached to stiff paper” type of specimen. To find records that are mostly in this last category, choose “Preserved specimen” from the options listed.

After making your selection, the number of records found will be shown (see image below, above the Purple arrow).

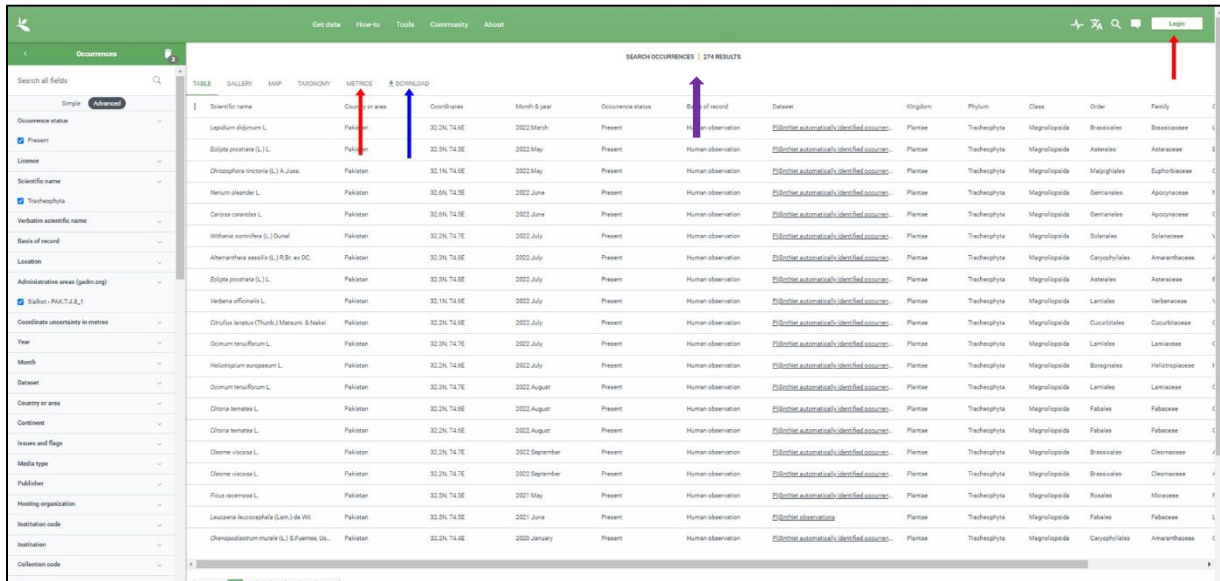


Figure 7. GBIF display of records meeting selection criteria.

Click the “Basis of Record” down arrow to see what kind of records are included (see image on right). They will be highlighted in the display as shown on the right.

To see where the records came from, click “Metrics” Above the long red arrow to the left of the blue arrow. That will bring up graphs summarizing the characteristics of the records that satisfied your filter.

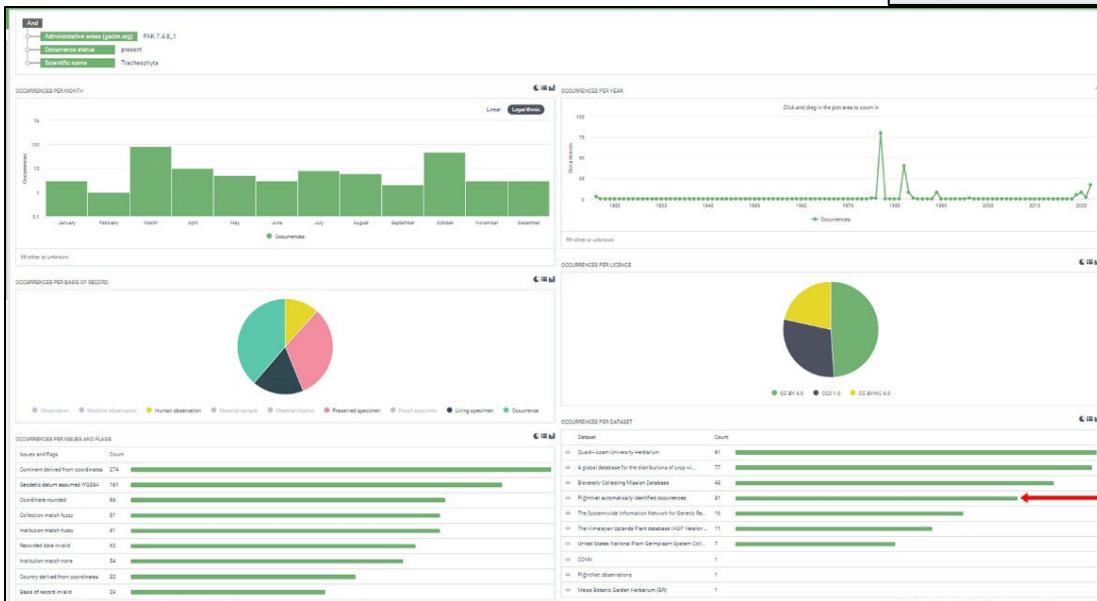
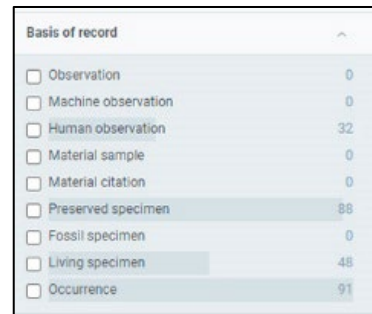


Figure 8. Metrics for records found.

- To download records, click “download” (Blue Arrow). You will be asked to login at that point if you have not already done so. Once you login, you will see three choices (see next page). Select the third choice, *Species list*. It will contain all you need and take less time to generate than the other two file options.

	Raw data	Interpreted data	Multimedia	Coordinates	Format	Estimated data size
<input type="button" value="SIMPLE"/>	X	✓	X	✓ (if available)	Tab-delimited CSV (for use in Excel, etc.) ⓘ	148 KB (33 KB zipped for download)
<input type="button" value="DARWIN CORE ARCHIVE"/>	✓	✓	✓ (links)	✓ (if available)	Tab-delimited CSV (for use in Excel, etc.) ⓘ	451 KB (100 KB zipped for download)
<input type="button" value="SPECIES LIST"/>	X	✓	X	X	Tab-delimited CSV (for use in Excel, etc.) ⓘ	

- GBIF will send you an email when your download is ready. RECORD THE CITATION INFORMATION:

GBIF.org (21 September 2023) GBIF Occurrence Download <https://doi.org/10.15468/dl.3kstzt>

You should add it to the description of your checklists.

- Opening the GBIF download: The downloaded file is a zip file containing a single csv file. Microsoft Windows will extract the csv file so you can save it to the folder of your choice.

To view the file:

- Open a new Excel spreadsheet.
- insertClick “Data” in the top tabs, and then go to the left side of the banner and click “From Text/CSV” (red arrows below)

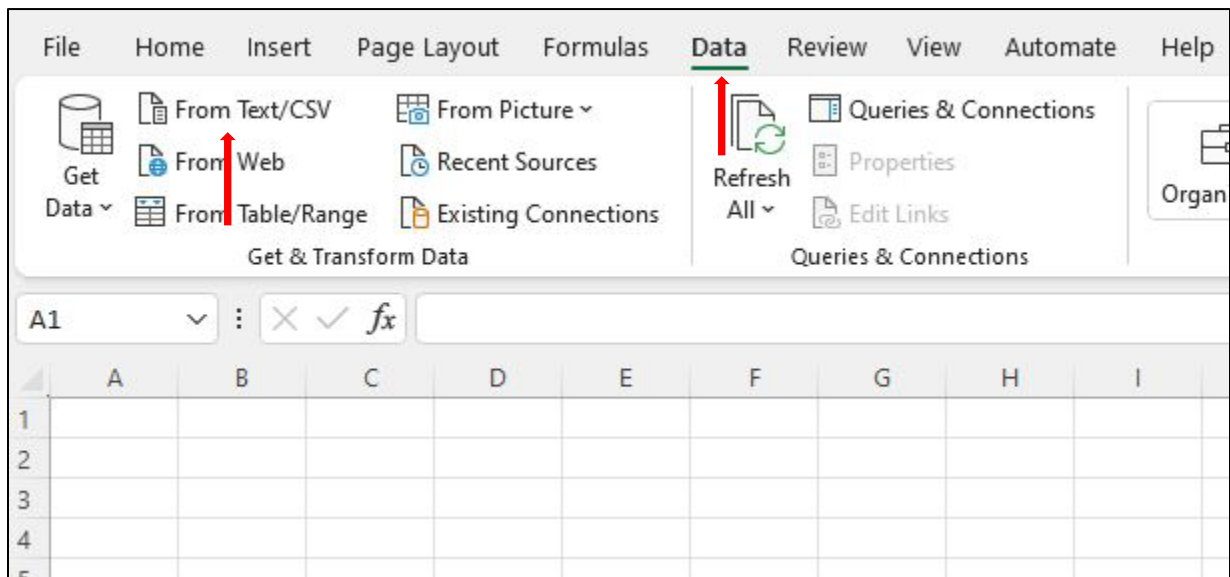


Figure 9. Importing a text/csv file into Excel

- Excel will load some of the records into the file and then pause for you to look at them. If you see no problems, click “Load” and then save the file.

- Creating a file for uploading the list of names. The GBIF download will have multiple columns containing scientific names. There are three that are important for creating a checklist: *ScientificName*, column B. This is the submitted name. *AcceptedScientificName*, column D. This is what the Catalogue of Life) considers the Accepted Name. *Species*, Column T. **This is the column to paste into the upload file as Sciname.** It contains the accepted names of the species *without their authors*. To add infraspecific names (subspecies and varieties, it is easiest to copy them from column D (the “accepted names” column) and delete the authors.
- Create the upload file. It must be a csv file. Give it two columns/fields: Sciname and Source. Paste column T, species, plus cleaned infraspecific names into the Sciname column and the DOI into the Source column for each record. Save the file as a csv file.
- Go to your checklist (MyProfile/Checklists/ChecklistName). Click the “spp” in the top right hand corner and then scroll down until you see “Batch Upload Spreadsheet”. Click there. Click the “Choose file” button, navigate to your newly created upload file and upload. If any names do not go up, please email them to me ([Mary Barkworth](mailto:Mary.Barkworth@openherbarium.org)).

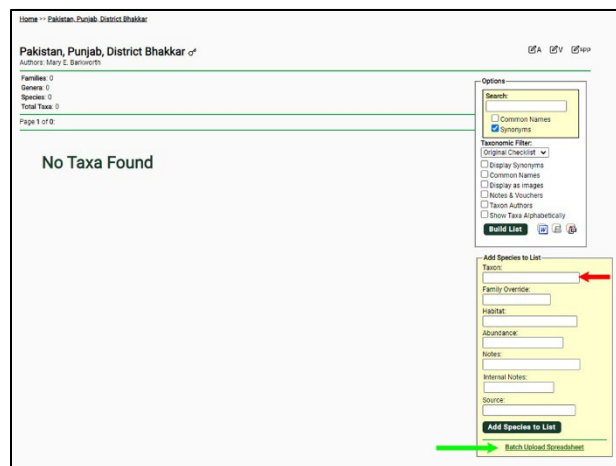


Figure 10. Click the green arrow to upload names in the csv file

As always, misspelled and duplicate names will not be added to your checklist. If there are correctly spelled names that OpenHerbarium will not upload, send me a list of them so I can identify and correct the problem. It may be that they are not yet in OpenHerbarium’s files.