U.S. National Library of Medicine
National Center for Biotechnology Information

GenBank Submissions Handbook
NCBI Help Manual

National Center for Biotechnology Information
U.S. National Library of Medicine

The "Source Modifiers" Page

Michael Fetchko and Adrienne Kitts

Purpose

The BankIt "Source Modifiers" page is the place where you will give us information about the organism(s) from which you obtained the nucleic acid sequence(s) you are submitting. The information we ask for in this page includes:

- The cellular location (if applicable) of the nucleic acid sequence
- Descriptive information about the organism
- Any information about how or where the organism was obtained
- Any additional descriptive information about the organism(s) that will serve to more specifically identify the nature of the organism(s).

Bacterial/Archaeal Sequences

This section will appear on the "Source Modifiers" page only if you identified the organism of the sequence(s) you are submitting as being either bacteria or archaea on the "Nucleotide" (or "Organism") page.

This section will help you determine the required source modifiers you need to submit. The required source modifiers depend on the method you used to obtain the sequence from the bacteria or archaea. When you select one of the listed methods, the required source modifiers that you need to provide will appear in a light blue box just below (Figure 11).

You must select one of the methods listed in order to proceed with you submission. If you do not select one of the methods, when you click the "Continue' button, you will get an error message and will not be able to go to the next page of the form.

In addition to the required source information for your bacterial/archaeal sequence(s), make sure to submit any other source information that you have about your bacterial/archaeal sequence(s).

Mouse and Rat Submissions

Provide the strain of the mouse or rat from which you obtained your sequence(s) if it is available. If you do not provide the strain, you will receive a warning asking you to do so when you click on the "Continue" button to go to the next page of the form.

Once you enter your strain information, click the "Continue" button. This will update the static table at the bottom of the page with your source modifier selections. Review the table to make sure your source modifiers are correct, and then click the "Continue" button again to move to the next page of the form.
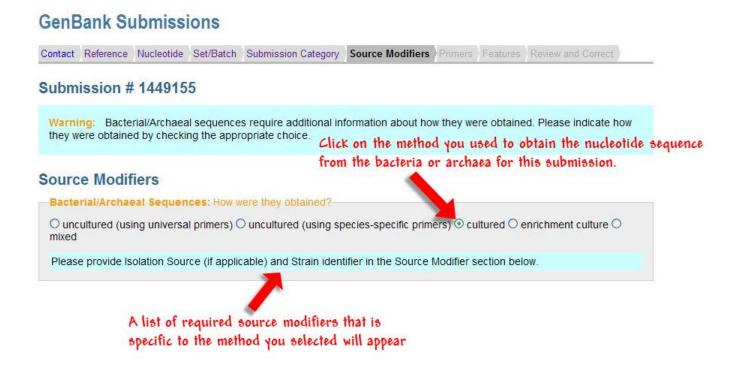
Figure 11: The Source Modifiers page when bacteria/archaea is submitted. This view shows the required source modifiers for the submission of bacteria/archaea sequences once a method is selected. Figure text gives information on how to get the list of required source modifiers.

Rice Submissions

Provide the rice cultivar from which you obtained your sequence(s) if it is available. If you do not provide the cultivar, you will receive a warning asking you to do so when you click on the "Continue" button to go to the next page of the form.

Once you enter your cultivar information, click the "Continue" button. This will update the static table at the bottom of the page with your source modifier selections. Review the table to make sure your source modifiers are correct, and then click the "Continue" button again to move to the next page of the form.

Virus Submissions

You must provide the following information for virus submissions:

- A virus strain name or isolate name

- The country where the virus was isolated
  A country modifier must be from GenBank's *approved list*
- The date that the virus was collected
  A "collection_date" must be in the format DD-Mmm-YYYY.
  Examples: 12-Sep-2002 or Jun-1999 or 2010

If you do not provide one or more of these pieces of information, you will receive a warning message asking you to do so when you click the "Continue" button to go to the next page of the form.

If a country name is unrecognized or a collection date is provided in the wrong format, you will receive an error message when you click the "Continue" button to go to the next page of the form. The error message explains the error and instructs you how to correct it. Once you correct the error, click the Continue button again to move to the next page.

Once you enter the requested virus information, click the "Continue" button. This will update the static table at the bottom of the page with your source modifier selections. Review the table to make sure your source modifiers are correct, and then click the "Continue" button again to move to the next page of the form.

The "Source Information" Section

Click the "Organelle/Location text box to release a drop-down menu of organelles and locations to choose from. (Figure 12). Select an organelle or location from the list only if it applies to your sequence.

Note:

- You do not have to select any of the listed organelles or locations if none of them applies to the sequence(s) you are submitting.
- The terms "nucleomorph" and "macronuclear" do NOT mean "nuclear". Do not choose either of these organelles unless they are appropriate to your specific organism

The "Source Modifiers" Section

Single Sequence Submissions

If you submitted a single sequence in the "Nucleotide" page of the form, when you get to the "source modifier" section of the "Source Modifiers" page, you will see a "source modifier" text box and a "value" text box.

Select source modifier(s) for your sequence by clicking in the "source modifier" text box to release a drop-down menu of source modifiers you can select from (Figure 13), and then enter the information that describes that source modifier in the "value" text box.

For example, if you select "collection date" as the source modifier, you would enter the date that the organism (from which you obtained the sequence) was collected in the "value" text box.

- If you are submitting a single sequence, you must enter the source modifiers for your sequence using the "source modifier" and "value" text boxes in this section of the page. You cannot upload a pre-made source modifier table.
- For each additional source modifier you have, click the "Add" button to generate additional "source modifier and "value" boxes you need to enter the information you have.

Once you have entered your source modifiers, click the "Continue" button. This will update the static table at the bottom of the page with your source modifier selections. Review the table to make sure your source modifiers are correct, and click the "Continue" button again to go to the next page of the form.
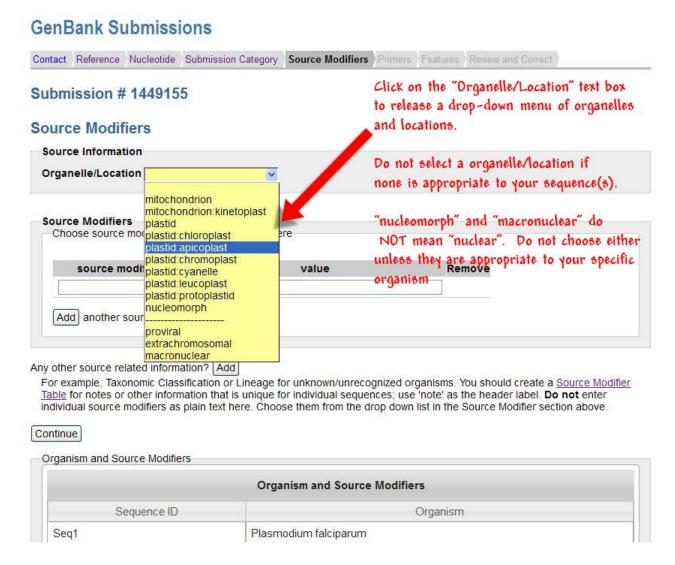
Figure 12: The "Source Information" section of the "Source Modifiers" page showing the drop-down list that appears when you click in the "Organelles/Locations" text box. Figure text gives hints for selecting the correct organelle/location.

Multiple Sequence Submissions

If you submitted more than one sequence in the "Nucleotide" page of the form, when you get to the "source modifier" section of the "Source Modifiers" page, you will see two sub-sections: "Set multiple values for sequences" and "Set one value for all sequences" (Figure 14).

- **If the source modifier(s) you have are unique to each sequence you are submitting**, use the "Set multiple values for sequences" subsection and click on the "Upload source modifiers Table File" button. When you do, a "Browse" button will appear that you can use to select the source modifier table file from your computer (Figure 15).
  Once you select your table file, it will upload when you click the "Continue" button, and the static table at the bottom of the page will update with your source modifier selections. Review the table to make sure your source modifiers are correct, and then click the "Continue" button again to move to the next page of the form.

  If a modifier and/or value must be corrected, make the correction in your source table file and upload the corrected file, which will overwrite your existing selections.
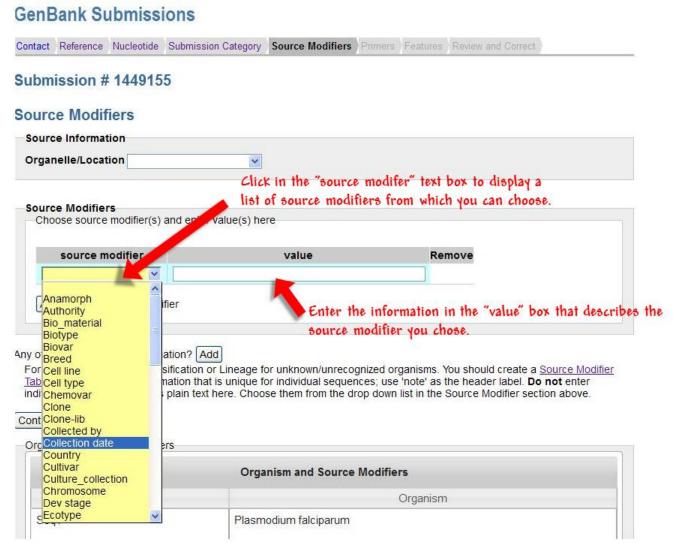
Figure 13: The "Source Modifiers" section of the "Source Modifiers" page when a single sequence is submitted. This view shows the drop-down list that appears when you click in the "source modifier" text box. Figure text gives a description of selecting the source modifier and entering information in the "value' text box.

For instructions on how create a source modifier table file, see the GenBank Submissions Quick Start, or use the BankIt Help documentation.

Note: If the source modifiers you have are unique to each sequence you are submitting you **must** submit the source modifiers for your sequences by uploading a source modifier table.

OR

- **If the same source modifier(s) apply to all the sequences you are submitting**, use the "Set one value for all sequences" subsection and click the "Choose source modifier" button. When you do, "source modifier" and "value" text boxes will appear (Figure 16).

  Use these boxes to enter the source in formation for your sequences. Clicking on the "source modifier" text box will release a drop-down menu of source modifiers you can select from (Figure 17). You can then enter the information that describes each source modifier in the "value" text box.

For example, if you had selected "collected by" as the source modifier, you would put the name of the person who collected the sample from which your sequence came in the "value" text box.

- For each additional source modifier you have, click the "Add" button to generate additional "source modifier" and "value" boxes you need to enter the information you have
- All of the source modifiers you enter will be applied to all of the sequences you submit.

Once you have entered your source modifiers, click the "Continue" button. This will update the static table at the bottom of the page with your source modifier selections. Review the table to make sure your source modifiers are correct, and then click the "Continue" button again to move to the next page of the form.

Common Mistakes Made While Filling Out the "Source Modifiers" Page

- **Mistake: Using the wrong format for the collection date**

  Fix: The "collection date" source modifier needs to be in this format: dd-Mon-yyyy. An example of the correct format is 15-Dec-2011. If you enter the collection date using any other format, you will not be allowed to continue with your submission until you submit the collection date in the correct format.
- **Mistake: Using invalid source modifiers as column heads for your source modifier table**

  Fix: Use only valid source modifiers as column heads for your table. You can find a complete list of valid source modifiers on the BankIt source modifier help page.
- **Mistake: Uploading a Source Modifier table that is not formatted in plain text.**

  Fix: Since BankIt can read only plain text (.txt) files, if you upload a source modifier table in a format other than plain text, it will not load properly and BankIt will display an error message. You will not be allowed to proceed with your submission until you upload a correctly formatted source modifier table.

  See the step-by-step instructions for making a tab-delimited source modifier table for information on making a table file in plain text (.txt) format, or use the BankIt Help documentation.
- **Mistake: Uploading a Flatfile list of Source Modifiers**

  Fix: Flatfile lists cannot be read by BankIt. If you upload a flatfile list of source modifiers, it will not load properly and BankIt will display an error message. You will not be allowed to proceed with your submission until you upload a tab delimited source modifier table in plain text format.

  See the step-by-step instructions for making a table file in plain text (.txt) format, or use the BankIt Help documentation.

  See the step-by-step instructions for making a tab-delimited source modifier table or use the BankIt Help documentation.

  Below are examples of flatfile format source modifiers lists that **cannot** be uploaded as source modifier files:

  source 1..1510
  /organism="Gallus gallus"
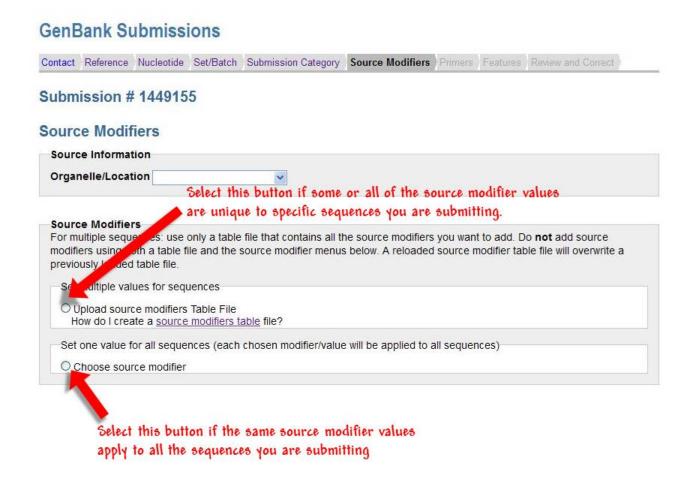 /mol_type="mRNA"
  /breed="Roaster"

Figure 14: The "Source Modifiers" section of the "Source Modifiers" page when multiple sequences are submitted. Figure text gives information about which source modifier loading option to use.

```
 /chromosome="Z"
/map="Zp21"

source 1..549
/organism="Passiflora sprucei"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/country="USA: Okanga National Forest"
/collected_by="John Jones"
```

- **Mistake: Pasting a flatfile list of source modifiers as an additional or corrected source modifier file on the "Review and Correct" page or as the only source information under 'Any other source related information?' on the Source Modifier page**

  Fix: These two text boxes allow a submitter to enter additional or corrected source modifier files or other plain text description for sequence source organisms after clicking a check box. Once you click on the check box, a text box will appear.
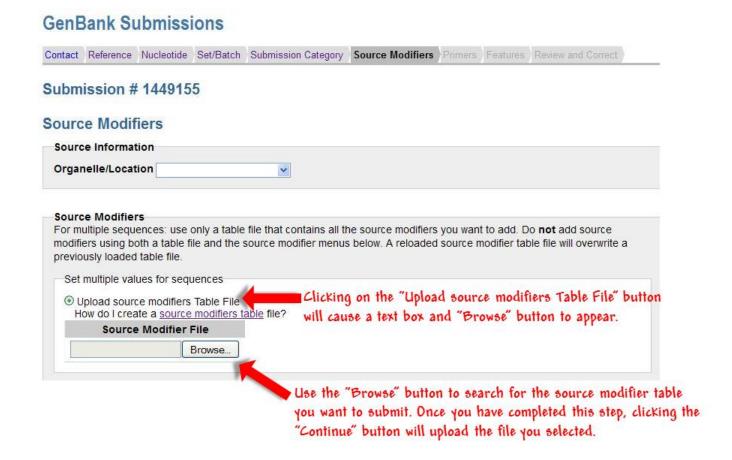
Figure 15: The "Source Modifiers" section of the "Source Modifiers" page when multiple sequences are submitted. This view shows the "Browse" button that appears when you click the "Upload source modifiers Table File" button in the "Set multiple values for sequences" sub-section. Figure text gives a description of selecting and uploading the source modifier table file.

A flatfile format list of source modifiers will be accepted in these text boxes because they are not restricted to the multi-column, tab-delimited format required for source modifiers; they will accept any text you put in them. However, your submission will not be accepted because we cannot upload flatfile format source modifier lists, and you will be requested either to:

○ Resubmit your sequences in a new submission with the source modifier correctly input

   OR

○ Send us your source modifier information in a tab-delimited plain text file.
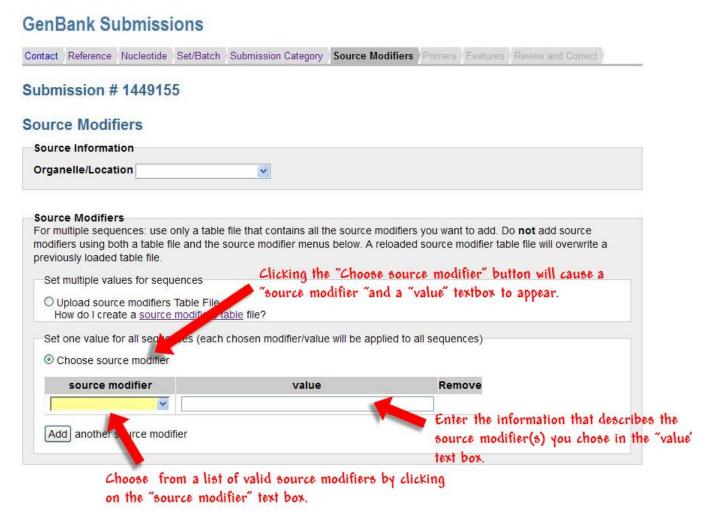
Figure 16: The "Source Modifiers" section of the "Source Modifiers" page when multiple sequences are submitted. This view shows the source modifier and value text boxes that appear when you click the "Choose source modifier" button in the "Set one value for all sequences" sub-section. Figure text gives a description of selecting the source modifier and entering information in the "value' text box.
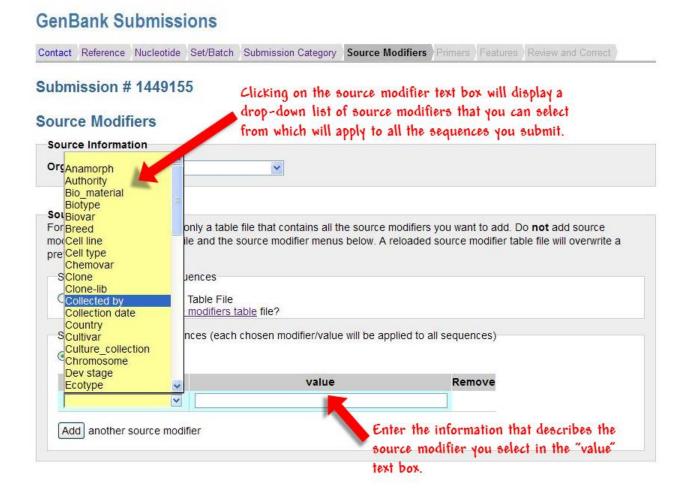
Figure 17: The "Source Modifiers" section of the "Source Modifiers" page when multiple sequences are submitted. This view shows the drop-down list that appears when you click in the "source modifier" text box in the "Set one value for all sequences" sub-section. Figure text gives a description of selecting the source modifier and entering information in the "value' text box.