Submission Wizard for Microsatellite sequences

Created: May 15, 2012; Updated: January 16, 2014.

Purpose

The Microsatellite Sequence Submission Wizard is for submitting microsatellite sequence submissions only. This wizard will guide you in providing all of the necessary source information for microsatellite sequences and will provide assistance and direction with feature annotation. Examples of source information are provided.

Wizard Import Nucleotide Sequences

Requirements: The Microsatellite Submission requirements are listed in the Sequences tab of the Wizard Import Nucleotide Sequences dialog box.

Sequence Format: Use the Import Nucleotide FASTA button to import your properly formatted FASTA file. For help with how to format the FASTA file click the FASTA Format Help button. The Sequences tab will display the information about the imported sequence(s). Please check the number of sequences, Sequence IDs (SeqIDs) and length of each sequence to make sure this information is correct.

If the sequences contain a significant number of ambiguous bases near the 5' or 3' end, you may be prompted to trim or remove these sequences from your submission.

Trim Vector Contamination: It is highly recommended that you perform a vector screen on your sequences and trim vector contamination by clicking the Vector Trim Tool button.

Delete Sequences: You can remove sequences from your submission using the Sequence Deletion Tool under the Edit Menu. This tool will assist you in removing any sequences from your file that you need to delete or that do not meet GenBank minimum sequence length requirements.

Sequencing Method

If you are submitting over 500 sequences or your sequences were generated using next-generation sequencing technology, the information in this form is required.

Sequencing Method: Use the check boxes at the top of the form to select the sequencing technology type(s) used to obtain the sequences. Multiple types can be selected, if appropriate. If you used technology that is not listed in the form, please select other and use the free text box to provide the information.

Assembly Program: After selecting the sequencing technology, select the radio button to indicate if your sequences are raw sequence reads or sequence assemblies. If you are submitting assemblies using next-generation sequencing technology, the name of the assembly program and program version or date the

assemblies were made are required in the free text boxes. If multiple assembly programs were used, Click on Add More Assembly Programs and complete the provided spreadsheet.

Raw sequence reads from next generation sequencing technologies should not be submitted to GenBank.

Microsatellite Wizard Molecule Type

Use this page to select the molecule type that was isolated and sequenced in your experiment. The default selection is nuclear genomic DNA. However, if you did not isolate genomic DNA from the nucleus, you should select no to the first question and then select the appropriate genome location and molecule type.

Microsatellite Wizard Annotation

Use this page to select what type of annotation you are planning to add to your sequence submission.

To apply a single repeat_region across the length of all of your sequences, select the Apply 1 microsatellite across entire sequence(s) radio button. If you have specific information about the repeat, such as the sequence of the repeat (rpt_unit_seq) or the nucleotide location of one repeat unit (rpt_unit_range), select the checkboxes next to those options. The rpt_unit_seq and rpt_unit_range information is optional.

To apply more complex annotation, such as multiple repeat_regions per sequence or repeat_regions with specific nucleotide spans, click Apply multiple microsatellites. After you provide the required source information, you will be prompted to add this information in the record viewer. A text dialog will open with instructions for importing a five-column, tab-delimited table containing the feature locations. You may also apply annotation using the Annotate menu options in the record viewer.

Microsatellite Wizard Information

Requirements: All sequences must have an organism name and a microsatellite name or clone name. After you provide all required information you will be prompted to exit the wizard and the record viewer will open.

How to add information to this table: There are three ways to add the information: 1) directly type into this form, 2) import a tab-delimited table, or 3) automatically populate the source information into the form if source information was included in the FASTA definition lines.

You can set the same qualifier value for all sequences by filling in the top row of boxes and using the appropriate Apply button. Use the Copy from SeqID button to apply the sequence IDs to the qualifier indicated in this table if this information was used as the sequence IDs in the original FASTA file.

Click on the Source Table Help button to open a text dialog with information on making a tab-delimited table. Click the Export This Table button to export a tab-delimited template file. You must maintain the tab-structure of the table in order to correctly import the data back into the submission wizard. Do not use spaces between the columns.

Errors: Any problems or missing information will be listed on the right side of the form. If you have made any changes on this form, please use the Recheck Errors button to validate the new information. Use the Show only sequences with errors radio button to list only those sequences that did not pass the validation.

Are you unable to pass the Information window? If you have not provided some required source information, the issue will be listed in the ***Problems*** column. After fixing any problems, click the Recheck Errors Button to determine if all issues have been fixed. You may display only the entries with problems by selecting the radio button next to Show only sequences with errors.

Do you not see a source qualifier in the table that you want to use in your submission? You may add columns for some commonly added source qualifiers using the buttons below the table. Other optional modifiers can be added to provide additional information using the "Apply/See More Source Information" button or "Import Source Table" button. A window with instructions for creating a source table can be viewed by clicking Source Table Help.

Did you have source information in your FASTA file that is not displayed in this table? This table only displays the required source qualifiers for each type of submission. It does not display all source information. If your FASTA definition lines were correctly formatted, the extra source information you provided in the FASTA definition lines will be imported. You will be able to review this information in the record viewer.