



The “Features” Page

Michael Fetchko and Adrienne Kitts

Purpose

The BankIt submission tool “Features” page is where you will select features for your sequence and provide detailed information about the features you selected. The detailed information we ask for in this section is dependent on the feature you select, but includes:

- The strand on which the feature appears
- Whether the feature is partial or complete. (If it is partial, if the feature is incomplete at the 5’ end or at the 3’ end)
- Whether the nucleotide interval where the feature occurs spans the entire sequence or has a specific span. If it has a specific span you must provide the nucleotide numbers of that span.
- Qualifiers for the feature

Adding Features to your Sequence: Feature Table File vs. Online BankIt Forms

Uploading a feature table file is an efficient method of adding features to sequence(s) if:

- you are adding many different features to a single sequence
- you are adding many different features to a number of different sequences

Complete the BankIt feature form(s) if:

- you are adding a one or a few features to a single sequence submission
- you are adding the same features to all the sequences in a multi-sequence submission
- you are adding the same features to specific sequences in a multi-sequence submission

Adding Features by Uploading a Feature Table

What is a five column feature table and how do I make one?

- A tab-delimited feature table uses a single “Tab” keystroke to delimit (mark the boundary) between one column and the next in a table that contains your feature information.
- BankIt [Help documentation](#) contains information about how to format a feature table and provides examples.
- You can access [step-by-step instructions](#) for creating a feature table in the GenBank Submission Resources Quick Start.
- A list of valid [Features](#) and [Qualifiers](#) you can use in your table are available in the BankIt Help documentation using links found within BankIt’s “Feature” pages.

- A Feature Table file must be saved in plain text format

Note: If you upload a table that includes invalid features or qualifiers, BankIt will tell you to correct your table with valid features/qualifiers and reload it.

See Figure 23 for a sample feature table marked up to show where to place your features and feature modifiers in a feature table.

The Importance of Formatting your Feature Table Correctly

The 5 column feature table format is specific so that BankIt can read the information in the table and put it in the right place within your submission. It is important that you follow the feature table format instructions and examples we provide. If your table differs from the format provided in the instructions and examples, BankIt will not be able to read it, and you will be requested to correct your table and upload it again.

Even if your table strays from the accepted feature table format in a very small way — like using a space between the columns instead of a tab, BankIt will not be able to read your table.

If BankIt does not accept your feature table, recheck your table carefully to see where it differs from the accepted format.

The Feature Table must be saved in a plain text format.

Uploading your Feature Table

Once you select the “Add features by uploading five column feature table file” option, a text box and a “Browse” button will appear on the “Features Overview” page. Click the “Browse” button and select the feature table file that you created and saved on your computer. Then click the “Upload File” button to upload the file.

Once you have uploaded the file, BankIt will read it and generate a list of the features that are in your table followed by a display of how those features will appear in your sequence record at the bottom of the page.

If there are errors in the feature table file, BankIt will display error and/or warning messages that describe the problem and direct you how to fix it.

If you review the table, and find that you need to make some changes, use the “Edit” and “Remove” buttons in the feature list to update the features you added. Once you are satisfied with the features you added, Click the “Continue” button to move on to the next page of the form.

Adding Features using the Online BankIt Forms

Once you select the “Add features by completing input forms” option, a list of 4 feature categories will appear on the “Features Overview” page:

- CDS
- RNA
- Repeat
- Other

When you select one of the categories, you may be directed to select a feature using a drop-down list or a button activated choice. Select one of the features from the drop-down menu or a button activated choice if they appear, then click the “Add” button to go to the “Features Detail” page for the feature. Once you are on the “Feature Details” page, you will give us detailed information about the feature you chose.

Sequence ID Row	Feature Start Value (1st Column of table)	Feature Stop Value (2nd Column of table)	Feature (3rd Column of table)	Modifiers for a feature are in the row below the feature and are in the 4th column of the table.	Modifier values are in the 5th column of the table
>Feature Seq1	<1	>1050	gene		
	<1	1009	CDS	product	acid trehalase
				product	Athlp
				codon_start	2
	<1	>1050	mRNA	product	acid trehalase

Figure 23: A sample feature table. Figure text shows the correct location for features and feature modifiers.

The Coding Region Feature Category

If the sequence you are submitting encodes a protein, select this feature category. When you do, options for how you will add the Coding Region (also called “coding sequence” or “CDS”) feature to your sequence record will appear. Once you select the method by which you will add the CDS feature, click the “Add” button and use the online “Features Detail” page that appears to give us detailed information about the CDS feature you are adding to your submission. For more instructions on how to provide detailed information and qualifiers for the CDS feature, see Box 4.

Box 4: The “Coding Region” Category: Selecting a Feature and Providing Information for it.

1. **Select the “Coding Region” category**
2. **Select an option on how you will add the Coding Region.**
3. **Click the “Add” button.** A “Features Detail” page will appear for you to provide intervals or to provide protein sequence data.

Box 4 continues on next page...

Box 4 continued from previous page.

4. **Answer the remaining questions in the “Feature Detail” page** with the general information requested for the feature.
 5. **Click to activate the “Qualifier” drop down menu and select the qualifiers** for the “Coding region” feature you chose.
 6. **Provide values for the qualifiers you chose.**
 7. **Click the “Add” button to add more qualifiers for your feature or click the “Accept” button** to record the feature details you entered into your submission.
 8. **Clicking the “Accept” button will** make a list of the features you added appear followed by a display of how those features will appear in your sequence record.
 9. **Review the feature display.**
 10. **Go to the “Added Features for editing/removal” list** above the Features display to:
 - a. Remove that feature from your record (click the “X” button located to the right of a feature to remove it).
- OR**
- b. **Change the information you entered for this particular feature** (click the “Edit” button to go back to go back to the “Feature Detail” page where you can change the information you provided for the feature)
11. **Click the “Continue” button** to continue to the next page of the BankIt form once you are satisfied with the features you see in the “Features” display at the bottom of the page.

Adding the CDS feature by providing intervals

If you choose this option to add the CDS feature to your record, in addition to providing qualifiers and general information about the coding sequence and the protein it contains, you will need to provide the coding region spans on the sequence (e.g. multiple intervals if your sequence contains introns, or the entire sequence if it does not). If the coding region spans specific part(s) of the sequence, you will need to give us the nucleotide numbers within the sequence for each span of the coding region.

Adding the CDS feature by providing protein sequence data

If you choose this option to add the CDS feature to your record, you must provide the sequence of the protein encoded by the coding region by either uploading a [protein FASTA file](#), or typing/pasting the sequence in the space provided in the “Features Detail” page.

Coding Region Feature Qualifiers

Once you get to the “Features Detail” page for the CDS feature, in addition to the other information requested on the page, you will also be asked to select feature qualifiers for the coding sequence. A list of valid [Qualifiers](#) you can use are available in the BankIt Help documentation using links found within BankIt’s “Feature” pages.

A product (protein) name (or description) is required for each CDS. If a gene is added, you must also provide a gene name (or description).

The RNA Feature Category

The “RNA” feature category allows you to select an RNA feature for your sequence. Once you select the RNA feature category option and click the “Add” button, a drop-down menu will appear. Click in the text box of the drop-down menu to release a list of RNA features. Once you select the one of the “RNA” features from the drop-down menu and click the “Add” button, you will use the “Features Detail” page to select qualifiers for and give us

more detailed information about the feature you chose. For more instructions on how to select “RNA” features and provide information and qualifiers for the “RNA” feature(s) you select, see Box 5.

Box 5: The “RNA” Category: Selecting a Feature and Providing Information for it.

1. **Select the “RNA” features category to make a dropdown menu appear.**
2. **Click on the drop down menu to release a list of RNA features.**
3. **Select a feature and click the “Add” button.** A “Features Detail” page will appear where you can provide the specific information about the feature you selected.
4. **Answer the questions in the “Feature Detail” page** with information about the feature you selected.
5. **Click to activate the “Qualifier” drop-down menu and select the qualifiers** for the “RNA” feature you chose.
6. **Provide values for the qualifiers you chose.**
7. **Click the “Add” button to add more qualifiers** for your feature **or click the “Accept” button** to record the feature details you entered into your submission.
8. **Clicking the “Accept” button will** make a list of the features you added appear followed by a display of how those features will appear in your sequence record.
9. **Review the feature display.**
10. **Go to the “Added Features for editing/removal” list above the features display to:**
 - a. **Remove that feature from your record** (click the “X” button located to the right of a feature to remove it).

OR

 - b. **Change the information you entered for this particular feature** (click the “Edit” button to go back to go back to the “Feature Detail” page where you can change the information you provided for the feature)
11. **Click the “Continue” button** to continue to the next page of the BankIt form once you are satisfied with the features you see in the “Features” display at the bottom of the page.

RNA Feature Types and Definitions

- **premessage RNA**
is an RNA molecule that was not processed after it was made and therefore contains intervening sequences (introns) in addition to the 5' untranslated region (5' UTR), the coding sequences (CDS, exon), and the 3' untranslated region (3' UTR) normally found in mature (processed) mRNA.
- **messenger RNA (mRNA)**
is RNA that encodes a protein. An mRNA for a protein product includes the 5' untranslated region (5'UTR), the coding sequence (CDS, exon) and the 3' untranslated region (3'UTR).
- **transfer RNA (tRNA)**
is a small RNA molecule (75-85 bases long) that facilitates the translation of a nucleic acid sequence into an amino acid sequence.
- **ribosomal RNA (rRNA)**
is the RNA component of the ribonucleoprotein particle (ribosome) which assembles amino acids into proteins. (e.g. 16S rRNA, 28S rRNA, large subunit rRNA).
- **non-coding RNA (ncRNA)**
is RNA that does not encode a protein. This feature should not be used for ribosomal RNA (rRNA) and transfer RNA (tRNA) as they have their own feature keys (examples of non-coding RNA include scRNA, snRNA, miRNA, and siRNA).

- **transfer messenger RNA (tmRNA)**
acts as a tRNA first, and then as an mRNA that encodes a peptide tag. The ribosome translates the mRNA region of the tmRNA and then attaches the encoded peptide tag to the C-terminus of the unfinished protein. The attached tag targets the protein for destruction or proteolysis.
- **miscellaneous RNA (misc_RNA)**
is any transcript or RNA product that cannot be defined by other RNA feature types listed

RNA Feature Qualifiers

Once you get to the “Features Detail” page for the RNA feature you select, you will be asked to select feature qualifiers for the RNA feature you chose. A list of valid [Qualifiers](#) you can use are available in the BankIt Help documentation using links found within BankIt’s “Feature” pages.

Note: You must provide a product name or description for all RNA features.

The Repeat Region Feature Category

The Repeat region category allows you to choose a repetitive element feature for your sequence. Once you select the Repeat region category button and then click the “Add” button, you will use the online “Features Detail” page that appears to select qualifiers for and give us more detailed information about the feature you chose. For more instructions on how to select a Repeat region feature and provide information and qualifiers for it, see Box 6.

Box 6: The “Repeat region” Category: Selecting a Feature and Providing Information for it.

1. **Select the “Repeat Region” category**
2. **Click the “Add” button.** A “Features Detail” page will appear.
3. **Choose the specific type of repetitive element** (e.g. mobile element, satellite or repeat sequence) you are submitting.
4. **A “Type” dropdown menu will appear** from which you must choose the specific type of mobile element, satellite or repeat sequence you are submitting.
5. **Provide the name you gave to the specific repeat you are submitting.**
6. **Answer the questions in the “Feature Detail” page** with general information about the feature you selected.
7. **Click to activate the “Qualifier” drop down menu** and select the qualifiers for the “Repeat region” feature you chose.
8. **Provide values for the qualifiers you chose.**
9. **Click the “Add” button to add more qualifiers for your feature or you click the “Accept” button** to record the feature details you entered into your submission.
10. **Clicking the “Accept” button** will make a list of the features you added appear followed by a display of how those features will appear in your sequence record.
11. **Review the feature display.**
12. **Go to the “Added Features for editing/removal” list** above the Features display to:
 - a. **Remove that feature from your record** (click the “X” button located to the right of a feature to remove it).

OR

 - b. **Change the information you entered for this particular feature** (click the “Edit” button to go back to go back to the “Feature Detail” page where you can change the information you provided for the feature)

Box 6 continues on next page...

Box 6 continued from previous page.

13. Click the “Continue” button to continue to the next page of the BankIt form once you are satisfied with the features you see in the “Features” display at the bottom of the page.

Repeat Region Feature Definitions

- **Repeat Sequence**

is a specific nucleotide sequence (unit) that recurs multiple times in a genome. A repeat unit can be arranged in any of the following ways:

- **Tandem repeat**

is a repeating nucleotide sequence that exists end-to-end in the same orientation with another copy of that nucleotide sequence.

- **Inverted repeat**

is a repeating nucleotide sequence that normally occurs as part of an end-to-end pair. The first member of the pair is the repeating nucleotide sequence oriented in the forward direction. The second member of the pair is the repeating nucleotide sequence oriented in the reverse direction.

- **Flanking repeat**

is a repeating nucleotide sequence that lies outside the sequence for which it is functionally important (e.g. transposon insertion target sites).

- **Terminal repeat**

is a repeating nucleotide sequence that occurs both:

- At the ends of sequence for which it is functionally important

AND

- Within sequence for which it is functionally important

- **Direct repeat**

is a repeating nucleotide sequence that does not always lie end-to-end with another copy of that nucleotide sequence, but is in the same orientation with it.

- **Dispersed repeat**

is a repeating nucleotide sequence that is found scattered throughout the genome.

- **Other repeat**

is a repeating nucleotide sequence with important characteristics that are not described by the other repeat types listed above.

- **Satellite DNA**

is made of many tandem repeats (identical or related) of a short, basic nucleotide sequence. It is frequently found in the centromere of a chromosome, but can also be found elsewhere. Because of its base composition, satellite DNA’s density is such that it will form bands in a CsCl buoyant density gradient that are “satellite” (separate from but close to) the bands formed by genomic DNA.

- **Mobile Element**

- *is a genetic entity that is capable of movement from one location to another in the genome.*

OR

- *is a genetic entity that is derived from the movement from one location to another in the genome.*

Repeat Region Feature Qualifier Definitions

Once you get to the “Features Detail” page for the Repeat region feature you select, you will be asked to select qualifiers for the repeat region feature you chose. A list of valid [Qualifiers](#) are available in the BankIt Help documentation using links found within BankIt’s “Feature” pages.

The “Other” Feature Category

The “Other” feature category allows you to select features for your sequence that are not shown on the “Feature Overview” page. Once you select the “Other” feature option and click the “Add” button, a drop-down menu will appear. Click on the drop-down menu to display a list of “Other” features. Once you select one of the “Other” features from the drop-down menu and click the “Add” button, you will use the online “Features Detail” page that appears to select qualifiers for and provide more detailed information about the feature you chose. For more instructions on how to select “Other” features and provide information and qualifiers for the “Other” feature(s) you select, see Box 7.

Box 7: The “Other” Category: Selecting a Feature and Providing Information for it.

1. **Select the “Other” features category to make a drop-down menu appear.**
 2. **Click on the drop down menu** to release a list of features not found on the “Feature Overview” page.
 3. **Select a feature and click the “Add” button.** A “Features Detail” page will appear where you can provide the specific information about the Feature Category you selected.
 4. **Answer the questions in the “Feature Detail” page** with information about the feature you selected.
 5. **Click to activate the “Qualifier” drop-down menu and select the qualifiers** for the “other” feature you chose.
 6. **Provide values for the qualifiers you chose.**
 7. **Click the “Add” button to add more qualifiers for your feature or click the “Accept” button** to record the feature details you entered into your submission.
 8. **Clicking the “Accept” button** will make a list of the features you added appear followed by a display of how those features will appear in your sequence record.
 9. **Review the feature display.**
 10. **Go to the “Added Features for editing/removal” list above the Features display to:**
 - a. **Remove that feature from your record** (click the “X” button located to the right of a feature to remove it).
- OR**
- b. **Change the information you entered for this particular feature** (click the “Edit” button to go back to go back to the “Feature Detail” page where you can change the information you provided for the feature)
11. **Click the “Continue” button to continue to the next page of the BankIt form** once you are satisfied with the features you see in the “Features” display at the bottom of the page.

“Other” Feature Definitions

You can see definitions for the [features](#) in the “Other” feature category in the BankIt Help documentation links in the “Features” page.

“Other” Feature Qualifier Definitions

Once you get to the “Features Detail” page, you will be asked to select feature qualifiers for the “Other” feature you chose. A list of valid [Qualifiers](#) are available in the BankIt Help documentation using links found within BankIt’s “Feature” pages.

Common Mistakes Made While Filling Out the “Features” Page

Mistake: Selecting the “Coding Region” feature category and providing only gene information in the “Features Detail” page.

Fix: If you select the “Coding Region” feature category and provide only your gene information and not the actual coding region, BankIt will display an error message that additional information for the CDS is required.

When you select the “Coding Region” category, be sure to provide either the nucleotide interval spans for the CDS or the actual protein sequence data in addition to the product (protein) name and the gene information.

Mistake: Choosing only “Gene” or “Exon” for a sequence that encodes a protein.

Fix: You should select the coding region (CDS) category to add a correct CDS feature. If you do not, your submission may be returned to you.

Mistake: Creating a feature table using features and qualifiers that are not valid.

Fix: Your table must include only those features and qualifiers found in the valid “[Feature](#)” and “[Qualifier](#)” lists that are linked to from the Bankit Help Documentation.

Mistake: Creating a feature table that does not follow the feature table format provided in the instructions and examples.

Fix: The 5 column feature table format is very specific – BankIt can only read a feature table that follows this format exactly. If your table differs from the format provided in the instructions and examples, BankIt will not be able to read it, and you will be requested to correct your table and upload it again.

Even if your table differs from the accepted feature table format in a small way, BankIt will not be able to read your table.

If BankIt does not accept your feature table, recheck your table carefully to see where it differs from the accepted format.