

NLM Citation: The GenBank Submissions Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2011-. Submitting Multiple Sequences as a Set. 2011 Apr 6 [Updated 2014 Nov 3].

Bookshelf URL: https://www.ncbi.nlm.nih.gov/books/



Submitting Multiple Sequences as a Set

Created: April 6, 2011; Updated: November 3, 2014.

What is the difference between a batch and a set?

A Set is a group of sequences of the same gene and/or locus that are part of a population, mutation, phylogenetic or environmental study, and are published together in the same journal article. The sequences in a Set will be processed together as a unit and will be accessible together using Entrez PopSet.

A Batch is a group of sequences that are related in some way, but from different genes and/or loci, such as coming from the ame organism or being published together in the same journal article.

When can multiple sequences be submitted together as a set?

- Members of a sequence set are published together in the same journal article, and are used to analyze the relatedness of a (sequence) population by analysis of the same gene or locus.
- Below is a list of standard set types submitted to GenBank:
 - Population study:

 a set of sequences that were derived by sequencing the same gene from different isolates of the same organism.
 - Phylogenetic study:
 a set of sequences that were derived by sequencing the same gene from different organisms.
 - Mutation study:
 a set of sequences that were derived by sequencing multiple mutations of a single gene.
 - Environmental samples:
 a set of sequences that were derived by sequencing the same gene from a population of unclassified or unknown organisms.

In all the cases listed above, each sequence set is handled as a single submission, but each sequence submitted in a set will receive its own Accession number, and can be annotated independently.

Sequence sets can be submitted to GenBank via Sequin or BankIt.

How do I submit multiple sequences as a set?

To submit a group of sequences as a set, do the following:

- Sequin:
 - 1. Fill out the Sequin Submitting Authors form.
 - 2. Select the set type on the Sequence Format Form (see the answer to "When can multiple sequences be submitted together as a set?).
 - 3. Sequin will prompt you to enter the data appropriate to the set.
- BankIt:

Once you have completed steps 1 (Contact) and 2 (Reference) of the BankIt submission process, step 3 (Submission Type) will give you a choice as to the submission type you intend to submit:

- 1. If you intend to submit a set, select the radio button in front of "Set".
- 2. A question asking the number of sets you intend to submit will appear as will a list of set types (population/phylogenetic/mutation/environmental).
- 3. Enter the total number of sets you intend to submit. (Note: you can only send one set per BankIt submission, so you will have to open a BankIt submission session for each additional set you intend to submit.).
- 4. Select the radio button in front of the set type you intend to submit.
- 5. From this point, BankIt will ask you for appropriate source and feature data for the set type you wish to submit.