



Adding Value to your Submission

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

Information to include with Genomic Sequence Submissions

Information to include with Eukaryotic Protein Coding Gene Sequence Submissions

I have eukaryotic genomic sequence(s) that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your eukaryotic sequence submission:

- CDS feature(s) with product name(s), nucleotide locations, and amino acid translations of all coding regions (showing start and stop codons, if present, and the locations of any exons)
- Gene symbol(s), if known

If any of this information is not known, inform us at the time of your submission

See an online [example](#) of eukaryote genomic sequence submission annotation or perform a BLAST search to find examples of similar sequences with complete annotation

Information to include with Prokaryotic Gene Sequence Submissions

I have bacterial/archaeal genomic sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your bacterial/archaeal sequence submission:

- **Tell us if the bacterium/archaeon is cultured or uncultured:**
 - pure culture: a culture that contains only one microbial species. Include a strain identifier.
 - enrichment culture: use of selective culture media to enrich for a set of microorganisms with a particular phenotypic property, resulting in a partially purified, mixed culture of more than one microbial species
 - uncultured bacteria/archaea are PCR-amplified directly from source/host DNA using universal primers or species-specific primers. Include the isolation_source (environmental conditions) and a unique clone identifier for each sequence.

Information to include with Prokaryotic Protein Coding Gene Sequence Submissions

I have bacterial/archaeal genomic sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your bacterial/archaeal sequence submission:

- CDS feature (s) with product name(s), nucleotide locations, and amino acid translations(s) of all coding regions (showing start and stop codons, if present)
- Gene symbol(s), if known

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of bacterial/archaeal genomic sequence submission annotation.

Information to include with Viral Sequence Submissions

I have viral sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your viral sequence submission:

- A unique name to distinguish each sequence you submit, such as one of the following:
 - Strain
 - Isolate
 - Clone
 - Other unique name
- **Country** where virus was collected (if known)
- Host (scientific/binomial or common name, if known)
- Collection date (if known)
(use three letter abbreviation for month and four digit format for year, e.g. Feb-2001)
- Serotype or genotype (if known)
- CDS feature(s) with product name(s), nucleotide locations, and amino acid translation(s) of all coding regions (showing start and stop codons, if present)
- Gene symbol(s), if known

The information listed above should be applied to any virus submission.

If no coding region is present, provide another description of the sequence

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of viral sequence submission annotation.

Information to include with Genomic Sequence containing Structural RNA and/or Spacers

I have genomic sequence that contains structural RNA and/or spacers that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with the sequence that contains structural RNA and/or spacers:

- The names of any structural RNAs (e.g. tRNA-Ile, 16S ribosomal RNA) present
- The names of any spacer regions (e.g. internal transcribed spacer 1, 16S/23S intergenic spacer)
- The nucleotide spans of each of the above features (if known)

If you do not know the exact nucleotide spans of the above features on your sequence, tell us which of the above components you think exists in each of your sequences using a “misc_feature” or “misc_RNA” feature with the components present listed in a note.

See an online [example](#) of structural RNA and/or spacer annotation.

Information to include with promoter/genomic 5' flanking sequence/genomic 3' flanking Sequence Submissions

I have promoter/genomic 5' flanking sequence/genomic 3' flanking sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your promoter/genomic 5' flanking sequence/genomic 3' flanking sequence submission:

- The protein and/or gene symbol for the sequence to which the promoter or flanking region belongs
- Intervals of any transcribed regions (ie noncoding and/or coding mRNA exons) or coding regions, if present

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of promoter/genomic 5' flanking sequence/genomic 3' flanking sequence submission annotation.

Information to include with Cloning Vector Sequence Submissions

I have cloning vector sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your cloning vector sequence submission:

- Unique name for the vector and type (ie cloning vector, expression vector, shuttle vector, etc.)
- Coding region intervals (if known) including start and stop codons
- Protein names (if known)
- Gene symbols (if known)
- Miscellaneous feature with descriptive note for biologically important regions (multiple cloning site, tags, enhancers, fusions, etc.)

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of cloning vector sequence submission annotation.

Information to include with Transposon/Insertion Sequence Submissions

I have Transposon/Insertion sequence(s) that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Include the following information with your Transposon/Insertion Sequence submission:

- The name of the transposon/insertion sequence
- The nucleotide spans corresponding to the transposon/insertion sequence
- The name of any host gene/product disrupted by the transposon/insertion sequence (if known)
- The name and nucleotide intervals of any gene/product in the transposon/insertion sequence, such as transposase (if known)
- The nucleotide spans any additional features present, such as LTRs or repeat regions (if known)

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of Transposon/Insertion sequence submission annotation.

Information to include with Microsatellite Sequence Submissions

I have microsatellite sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your microsatellite sequence submission:

- A unique microsatellite/clone name for each sequence
- The interval of any repeat region(s) within the microsatellite sequence (if known)

Please make sure to remove all cloning vector contamination as failure to do this will delay the processing of your submission.

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of microsatellite sequence submission annotation.

Information to include with Sequence Submissions containing Repeat Regions

I'd like to submit sequence data that contains repeat regions; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your sequence submission:

- Repeat region intervals
- Repeat family, if known (eg, Alu, Mer)
- Repeat type (tandem, inverted, flanking, terminal, direct, dispersed, or other)
- Repeat unit description/intervals, if region contains more than one repeat

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) repeat region sequence submission annotation.

Information to include with Pseudogene Sequence Submissions

I'd like to submit pseudogene sequence data; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your pseudogene sequence submission:

- gene intervals
- gene symbol
- protein name as a note on the gene feature

If this information is not known, inform us at the time of your submission.

See an online [example](#) of pseudogene sequence submission annotation.

Information to include with RNA/mRNA Sequence Submissions

Information to include with non-coding RNA

I have non-coding RNA sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your non-coding RNA (ncRNA) sequence submission:

- ncRNA intervals
- ncRNA class
You will be able to select a class from a list of options provided in both BankIt and Sequin
- ncRNA product name

If any of the above information is not known, inform us at the time of your submission.

Information to include with mRNA Sequence Submissions

I have mRNA sequence data that I'd like to submit; is there additional information that I should include with the sequence(s) in my submission?

Provide the following information with your mRNA sequence submission:

- Coding region intervals including start and stop codons, if present
- Protein name
- Gene symbol, if known
- Amino acid sequence

If any of this information is not known, inform us at the time of your submission.

See an online [example](#) of mRNA **sequence** submission annotation.

Information to Include with Alternative mRNA Transcript Submissions

I have alternative mRNA transcripts that I'd like to submit; is there additional information that I should include with the sequences in my submission?

- **Each alternative mRNA sequence must be submitted separately.** Each will be processed as a separate submission, and assigned a unique accession number.
- **Provide each alternative mRNA sequence with a specific name.** You can do this by using the name for the transcript's CDS product. For example:
/product=actin isoform A
/product=actin isoform B
- **Provide each alternative mRNA sequence with a note on the CDS feature** identifying the sequence as alternatively spliced:
/note=alternatively spliced