



June 18, 2009

iBOL/GenBank/Genome Canada Letter of Cooperation

BOLD will broker the submission of iBOL barcode sequence entries to GenBank within one week of sequencing. This submission will include:

1. An assembled nucleotide sequence, trimmed according to a common set of criteria to be agreed on by GenBank and BOLD.
2. Trace files.
3. Primers, or a probeid for the primers.
4. Country, and any detailed locality information (which should be available at the time of sequencing)
5. Specimen voucher annotation (which should be available at the time of sequencing) including institution code, optional collection code and specimen id in the Darwin Core triplet format.
6. BOLD organism BIN name (or other more formal taxonomic name). The BIN name should be associated with a taxonomic assignment to order level, e.g. "Lepidoptera sp. BOLD:1A1-8KU"
7. Any other information (apart from images) that is available at the time of submission and is public on the BOLD site will also be submitted.

There are several pieces of metadata that do not have a natural place in the GenBank flatfile annotation scheme – these will be collected in an iBOL structured comment. These data include:

1. iBOL working group (i.e. vertebrate, plant, terrestrial biosurveillance, etc.)
2. Tentative taxonomic identification.
3. BOLD BIN name (included in the iBOL structured comment whether or not the corresponding sequence entry is annotated with a formal taxonomic name or with an informal BOLD BIN name).
4. Any other structured BOLD data that does not have a natural place in the GenBank flatfile (e.g. elevation).

BOLD agrees to set up an update channel to GenBank, to pass along any changes that are made to the barcode entries on the public BOLD site. Contingent on this, GenBank agrees not to process updates to iBOL entries directly; all updates should go through BOLD to GenBank.

GenBank agrees to provide Genome Canada and BOLD with monthly and quarterly reports of the progress of the iBOL project. This report will include counts of

1. Number of traces successfully uploaded
2. Number of GenBank accessions assigned for iBOL entries
 - a. Total number
 - b. By iBOL working group
 - c. By country of sample origin
3. Number of new taxon names
 - a. Total number
 - b. Number of new BINs
 - c. Number of resolved species (with formal taxonomic names)
4. Number of GenBank accessions updated