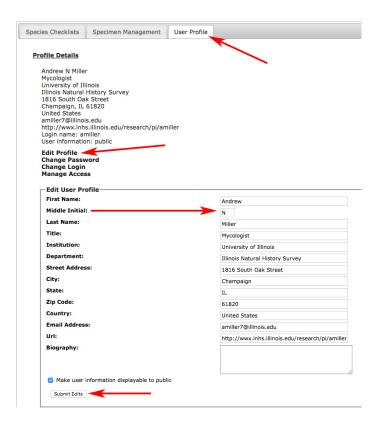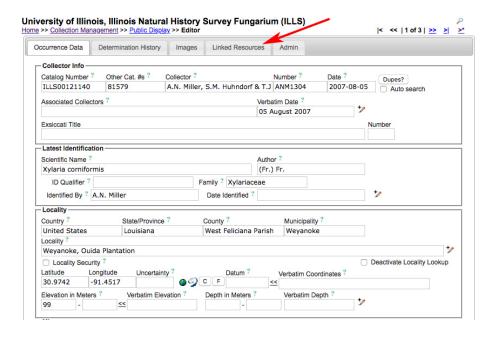# GenBank sequence submission

1. If you have not already done so, go to your User Profile and fill in your middle initial (if you have one).
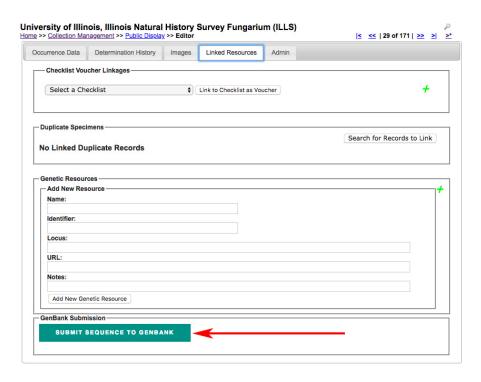
"My Profile" –––> "User Profile" –––> "Edit Profile", enter your middle initial, then click "Submit Edits"

2. Search for the specimen record in the Edit Existing Occurrence Records search tool that you want to submit genetic sequence data for. Confirm all specimen metadata is **complete** and **accurate**! Then, click on "Linked Resources".

3. Click on "Submit Sequence to GenBank".

4. A pop-up window will appear and your username and the Record ID (occid) of the specimen will be automatically displayed.

5. Copy and paste your FASTA formatted sequence into the Sequence box and then chose the Sequence Definition for the gene region you are submitting.

If you hover your mouse over the **?**, the GenBank approved standard definitions will appear in a pop-up box.

6. Select the method used to obtain this sequence. It will nearly always be "Sanger dideoxy sequencing".

7. If this sequence is published in a peer-reviewed journal or as part of a project such as the "North American MycoFlora Project", slide the button to "Yes", enter the title of the publication, and click the "+" to add authors.

8. Add as many authors as you like. They will appear in the order you add them. To remove an author, click the "x".

9. If you want GenBank to embargo (i.e. not release your sequence to the public) for 1 year, then slide the button to "Yes". It is recommended to immediately release your sequences and not embargo them.

10. Finally, click the "Add" button, then the "Generate" button.

You will see a pop-up window that tells you to click the "Open File Manager button". Click "OK", then open the File Manager to see your generated files.

11. Click the "GenBank" file to see your list of files.

12. Click the filename of the scientific name (with date and time stamp) to see your list of sequence files.

13. You will see 2 files: a .gbf file and a .sqn file. The .gbf file will show you exactly what your sequence submission will look like at GenBank (see below). The .sqn file should be downloaded and emailed to GenBank at this email: gb-sub@ncbi.nlm.nih.gov

PLEASE CAREFULLY CHECK YOUR .gbf FILE FOR ANY ERRORS BEFORE EMAILING YOUR SEQUENCE SUBMISSION TO GenBank!!!

```
LOCUS       Xylaria_corniformis      536 bp     DNA      linear      18-JUN-2018
DEFINITION  [gene]=large subunit ribosomal RNA.
ACCESSION
VERSION
KEYWORDS    .
SOURCE      Xylaria corniformis
  ORGANISM  Xylaria corniformis
            Unclassified.
```

publication author(s) →
```
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Miller,A.N. and Bates,S.T.
  TITLE     Fungal DNA Barcoding in the Great Smoky Mountains National Park
  JOURNAL   Unpublished
```
submission author →
```
REFERENCE   2  (bases 1 to 536)
  AUTHORS   Miller,A.N.
  TITLE     Direct Submission
```
institution and address automatically pulled from Symbiota User Profile info →
```
  JOURNAL   Submitted (18-JUN-2018) Illinois Natural History Survey, University
            of Illinois, 1816 South Oak Street, Champaign, IL 61820, United
            States
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
```
Structured Comment specifically developed for Symbiota portals →
```
            ##SymbiotaSpecimenReference-START##
            Source Record URL    :: http://mycoportal.org
            Record ID            :: 782f1928-8e2f-4a49-ad2f-17ef6b9f46f9
            Institution Code     :: ILLS
            Catalog Number       :: ILLS00121140
            Other Catalog Numbers :: 81579
            ##SymbiotaSpecimenReference-END##
FEATURES             Location/Qualifiers
     source          1..536
```
specimen metadata automatically pulled from DwC fields in Symbiota →
```
                     /organism="Xylaria corniformis"
                     /mol_type="genomic DNA"
                     /isolation_source="4 cm. diam. branch on ground"
                     /specimen_voucher="ILLS00121140"
                     /country="USA: Louisiana, West Feliciana Parish, Weyanoke,
                     Ouida Plantation"
                     /lat_lon="30.9742 N 91.4517 W"
                     /altitude="99 m"
                     /collection_date="2007-Sep-05"
                     /collected_by="A.N. Miller, S.M. Huhndorf & T.J. Atkinson"
                     /identified_by="A.N. Miller"
     misc_RNA        <1..>536
                     /product="large subunit ribosomal RNA"
ORIGIN
        1 atctcgccta tcagttggac cggtcgcctg ccgtaaaacc ccccaacttc ttaaagttga
       61 cctcgaatcg gttcagacaa actcgctaaa ttgaagcatc ttagtcagcg aaggaaaaga
      121 aaccaacagg gattgcccta gtaacggcga gtgaagcggc aacagctcaa atttgaaatc
      181 tggccctcgg gtccgagttg taatttgtag aggatgcttt tggcgcggtg ccttccgagt
      241 tccctggaac gggacgcctt agagggtgag agccccgtac ggttggacac caagcctctg
      301 taaagctcct tcgacgagtc gagtagcttg ggaatgctgc tctaaatggg aggtaaattt
      361 cttctaaagc taaatattgg ccagagaccg atagcgcaca agtagagtga tcgaaagatg
      421 aaaagcactt tgaaaagagg gttaaacagc acgtgaaatt gttgaaaggg aagcgtttgc
      481 gaccagacct tttcctagcg gatcatccgg tgttctcacc ggtgcacttc gctagg
//
```