

Training 4: How to download sequences, edit sequences, and upload sequences to BOLD when there isn't one to replace

If BOLD says there are 2 (or 1) trace files but no sequence appears

First look at the trace/chromatogram previews (click box and arrow right)

If you think the sequence may be salvageable, **download the trace files.**

Sequence View for Process ID: NAMPA1302-21

Try **BLASTing** a short cleaned segment 1st to ensure you aren't cleaning a contaminant.



[Upload Traces](#) [Download Traces](#)

[Activity Report](#) [Show Delta](#)

Specimen Details Current

Sample ID: MO393550
Process ID: NAMPA1302-21
Project: NAMPA
Tax Names: Basidiomycota, Agaricomycetes, Boletales, Boletaceae, Xerocomellus
Taxon: Xerocomellus

Marker Summary

Marker Code	Sequence Length	GC	Ambiguous	Trace Count
ITS	744	48.4%	0%	2

If the sequence is recoverable based on quick view(s) of chromatogram(s) & BLAST search of a short segment

To edit the sequence from the chromatogram(s):

- If you can only edit a single sequence, edit the longest, cleanest chromatogram
- If you have Sequencher or similar program:
 - create a consensus sequence if there are both forward (ITS1F) and reverse (ITS4) reads. Use 65% to 85% match
 - If you can't get the two reads to overlap, use BLAST search in GenBank to retrieve a good, long similar sequence to use as a reference sequence.
 - Add the reference sequence to the Sequencher file that has the two uploaded chromatograms and designate the former as **Reference sequence**.
 - Edit the consensus sequence.
 - If there is a small region of non-overlap, insert NNNN for the missing bases.
- Blast the edited sequence again.
- Check the best matched sequence from GenBank (if different) to adjust % similarity values & replace screenshot in slides
- If the best matched sequence is different from the 1st one in GenBank, add this to the Sequencher file, designate as a **Reference sequence** & re-clean the edited sequence.

How to upload sequences to BOLD

Go to the Main Console and look for the box in upper right that says **'Uploads'** and select **'Sequences'**

Welcome to BOLD Systems
Home / BOLD Main Console

Projects [New Project](#)

6
Projects with access

Specimens

3336
Records with access

999
Barcodes with access

Uploads

[Sequences](#) [Traces](#) [Images](#)

[Primers](#) [Publication](#) [Checklist](#)

Your Datasets: 11 [New Dataset](#)

Code	Title	Specimens
DS-FUND0000	Fundis-Trial Plate <small>ITS 853</small>	95
DS-FUND1052	1052-Macrofungi of Lane County (OR) <small>ITS 423</small>	92

Recently Accessed [Top 20](#)

Code	Title	Specimens	Accessed
NAMPA	FunDis Sanger Sequencing (previously North American Mycoflora Project Trial Plate 2019) <small>ITS 9300</small>	1969	today
FDSAF	Igor Safonov- 1021 <small>ITS 58</small>	59	4 days ago

Upload page will look like this:

•Keep Select ID: as Process ID
(the NAMPA####-YR)

•Change Markers: from
COI-SP to ITS

•Enter a Run Site (if it won't
take Counterculture lab,
try typing University of
California Berkeley. It's a
pull-down list

Enter the edited sequence in
Fasta file format in the box, e.g
>NAMPA####-YY
sequence

The screenshot shows a web form titled "Sequence Submission" with a close button in the top right. Below the title, a note states "Required fields for submission are marked in red." The form contains three dropdown menus: "Select ID:" with "Processid" selected, "Markers:" with "COI-5P" selected, and "Run site:" which is currently empty. Below these is a text box with instructions: "Select the institution responsible for generating the sequences listed by beginning to type in the name. BOLD will return matches in a drop down box, and clicking on a name will select it. If the institution is not listed, please register it." At the bottom, there is a section titled "Paste sequences in fasta format:" followed by a large text area for pasting the sequence. Three arrows point from the text on the left to the "Select ID:", "Markers:", and "Run site:" fields.

Filled upload page should resemble this:

Sequence Submission



Required fields for submission are marked in red.

Select ID:

Processid

Markers:

ITS

Run site:

University of Georgia

Select the institution responsible for generating the sequences listed by beginning to type in the name. BOLD will return matches in a drop down box, and clicking on a name will select it. If the institution is not listed, please register it.

Paste sequences in fasta format:

```
>NAMPA1302-21
TAACGAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTATCGAATTCGAGGGGGGAAGATGGAGGGGGAGAGACTGTCGCTGGCTCCACCGGGAAGCATGTGCACGCTTCCITTTTCGTGCACCT
TTCTCTTACTCTCACACCTGTGCACACATTGTAGGTCCTCGAAAGAGGATCTATGATTTATCATCACACATCGTATGTCTAGAATGTATCATGATCGTCGACCGGCGGTCAACAAATAAATAACAATT
TCAGCAACGGATCTTTGGCTCTCGCATCGATGAAGAACGCAGCGAATTGCGATAAGTAATGTAATTGCAGATTTTCAGTGAATCATCGAATCTTTGAACGCACCTTGCCTCTTGGTATTCCGAGGA
GCATGCCCTGTTGAGTGTATTGAATCTCAACCATGTCTTGATCGAATTCGAGGCATGGCTTGGACTTGGGGTGTGCGCGCCGAAAGCTGTCGGCTCTCCTGAAATGCATTAGCAATGGACAGCAA
GTCTGATGTGCACGGCCTTGACGTGATAATGATCGTCGCTGGAACTTGGACGAGCAGGAATGCGTCTGTTTGTCTCAATCCTTGACTTTGGACTTTTTCTTCGAGCGAAGCTTTTAGTACTAGT
TGGTCGTGAGGCTGACGAACGCAAGGCTGGCTTGGAGGTAGAAGGACCTCGTCTTCTTGACAACCTGACCTCAAATCAGGTAGGTAGGTGGT|
```

Submit

Remember to hit
Submit

If you did it correctly, you should see this:

Sequence Submission Report



Submission Status: Success

Total Sequences Submitted: 1

> Sequences Created: 1

Sequences Created: 1

Project Summary

> NAMPA: new

NAMPA1302-21

Search Records

Hit Search Records

To get this

Select	Identification	Specimen Page	Sequence Page	Extra Info	BIN	Record Flags Legend						Bases [Ambig]	Tags	
													ITS	
<input type="checkbox"/>	Xerocomellus	MO393550	NAMPA1302-21	1019			1	2					744[0n]	Jean_Add_Xerocomellus_intermedius_to_BOLD_DB failed

Number of bases in your newly uploaded sequence

