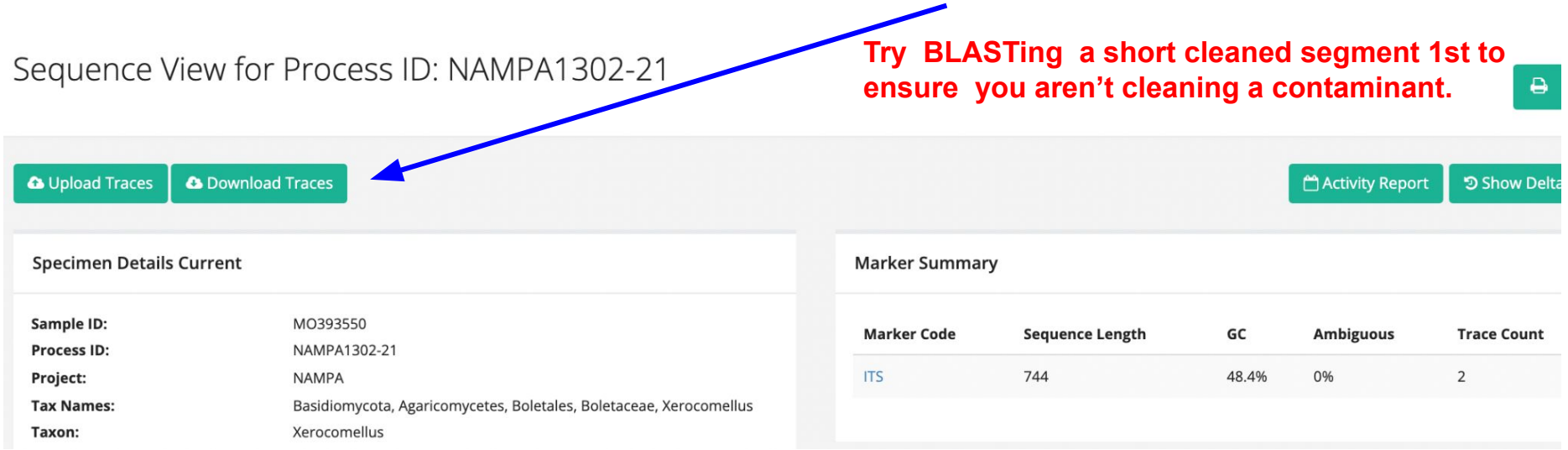
Training 4: How to download sequences, edit sequences, and upload sequences to BOLD when there isn't one to replace

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**.



If the sequence is recoverable

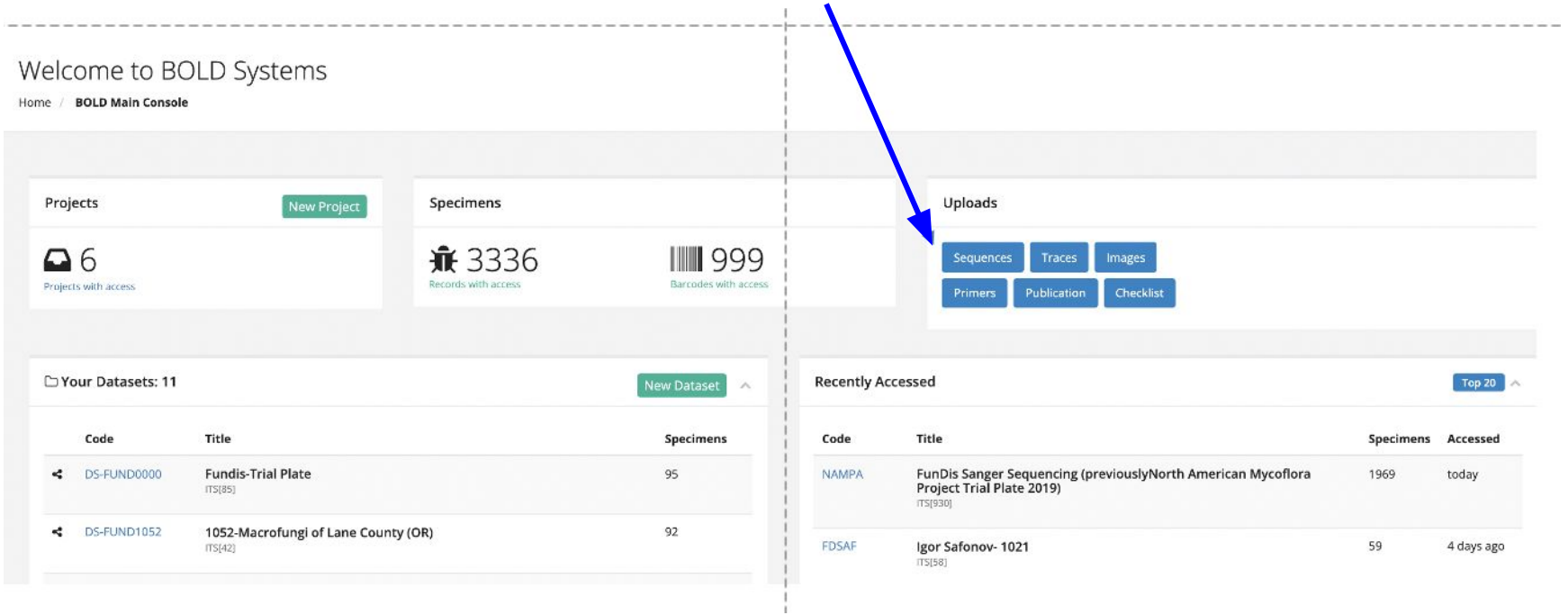
based on quick view(s) of chromatogram(s) & BLAST search of a short segment

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 Seqeuncher file, designate as a **Reference sequence** & re-clean the edited sequence.

How to upload sequences to BOLD

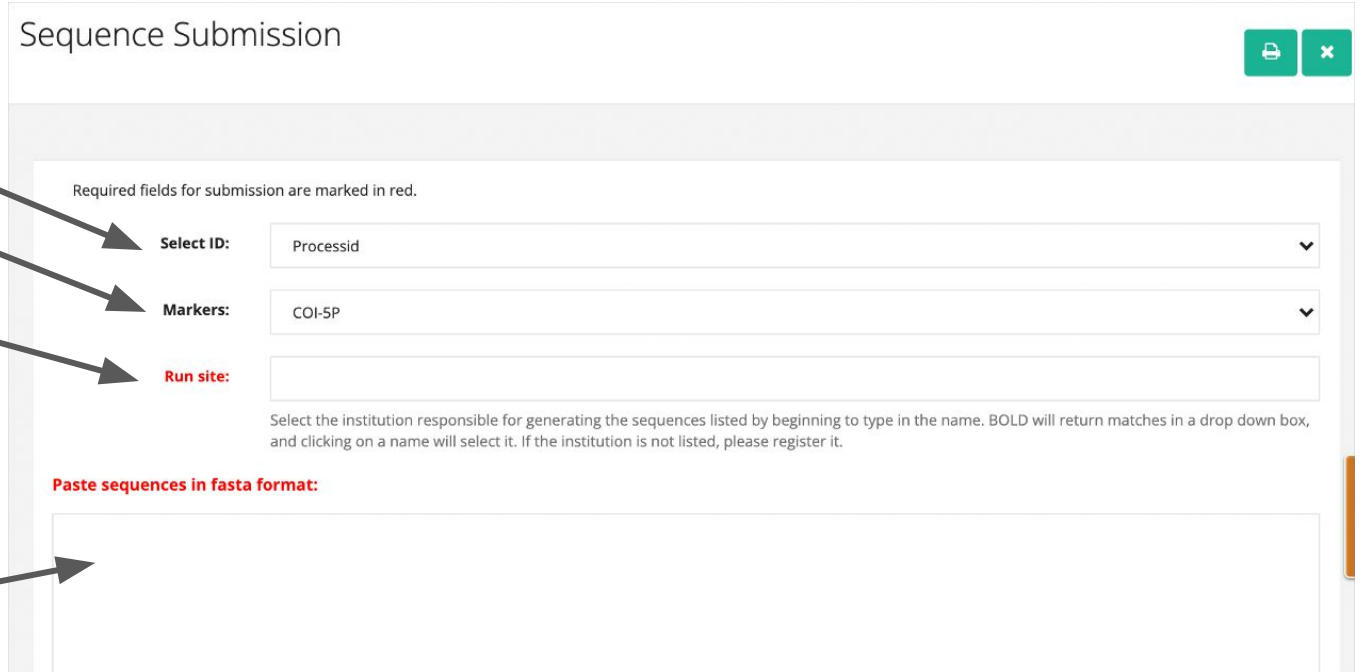
Go to the Main Console and look for the box in upper right that says ‘**Uploads**’ and select ‘**Sequences**’



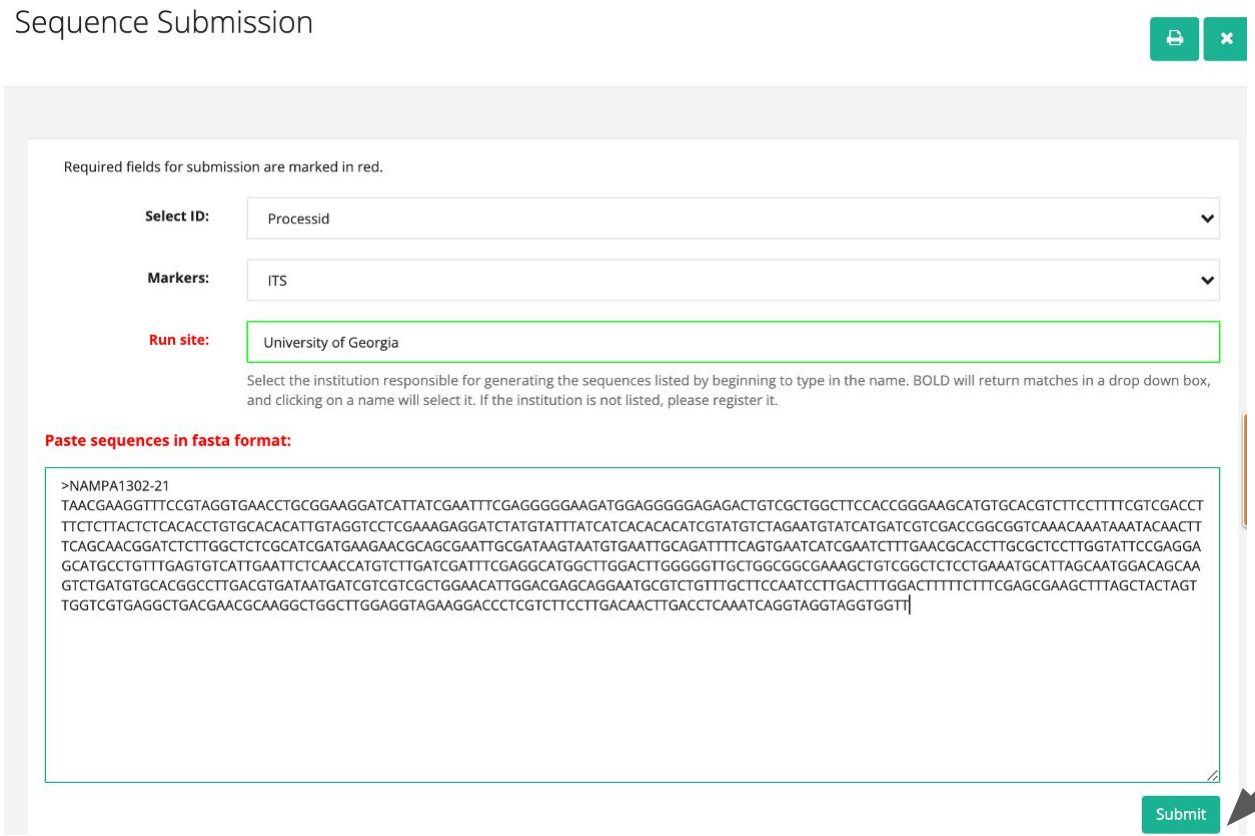
Upload page will look like this:

- 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

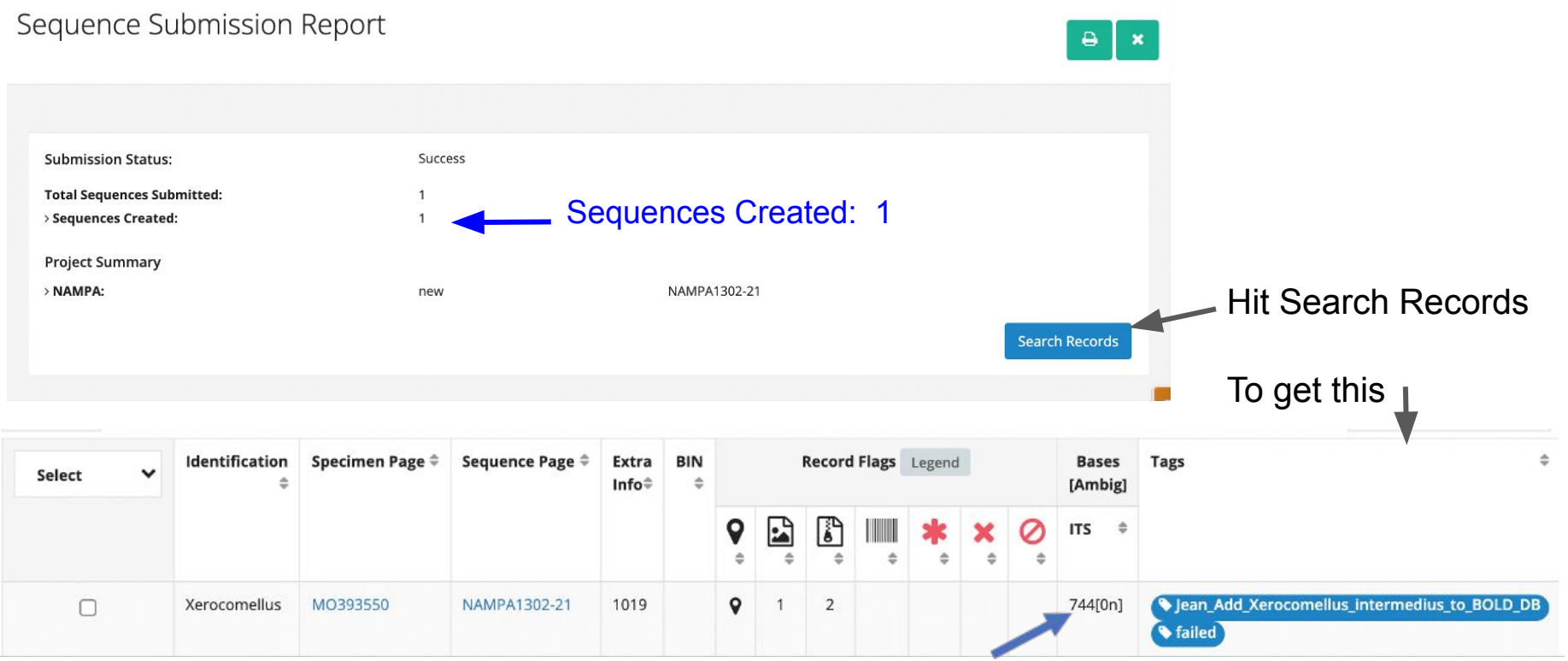


Filled upload page should resemble this:



Remember to hit **Submit**

If you did it correctly, you should see this:



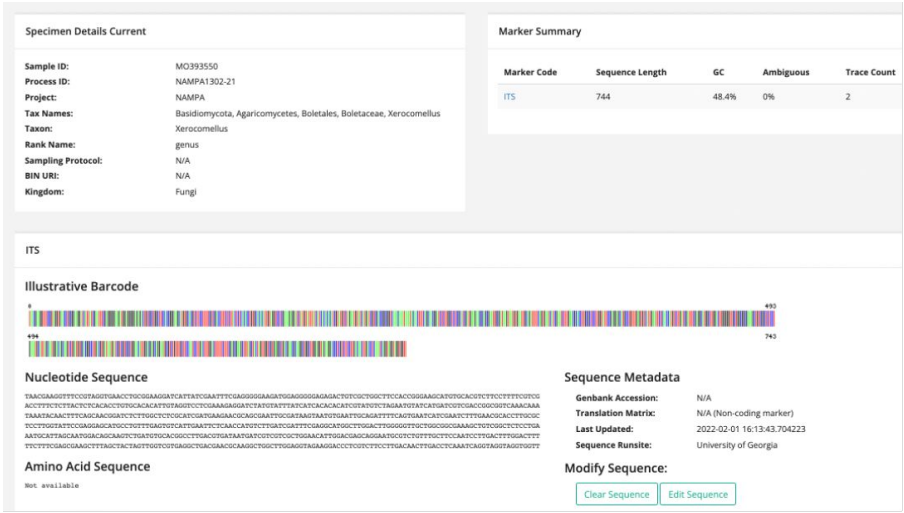
To view the new sequence:

Hit the blue hot-link for the Sequence Process Code (NAMPA####-YY)

Select	Identification	Specimen Page	Sequence Page	Extra Info	BIN	Record Flags Legend							Bases [Ambig]	Tags
												ITS		
☐	Xerocomellus	MO393550	NAMPA1302-21	1019		1	2					744[0n]		Jean_Add_Xerocomellus_intermedius_to_BOLD_DB, failed

To pull up the BOLD Sequence page.

There are Comments boxes to add notes, either to the trace files or to the Specimen page. Use Comments to indicate you created the sequence from the trace files & your name.



Specimen Details Current

Sample ID:	MO393550
Process ID:	NAMPA1302-21
Project:	NAMPA
Tax Names:	Basidiomycota, Agaricomycetes, Boletales, Boletaceae, Xerocomellus
Taxon:	Xerocomellus
Rank Name:	genus
Sampling Protocol:	N/A
BIN URI:	N/A
Kingdom:	Fungi

Marker Summary

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

ITS

Illustrative Barcode

Nucleotide Sequence

Sequence Metadata

Genbank Accession:	N/A
Translation Matrix:	N/A (Non-coding marker)
Last Updated:	2022-02-01 16:13:43.704223
Sequence Runsite:	University of Georgia

Amino Acid Sequence

Not available

Modify Sequence:

Clear Sequence | Edit Sequence