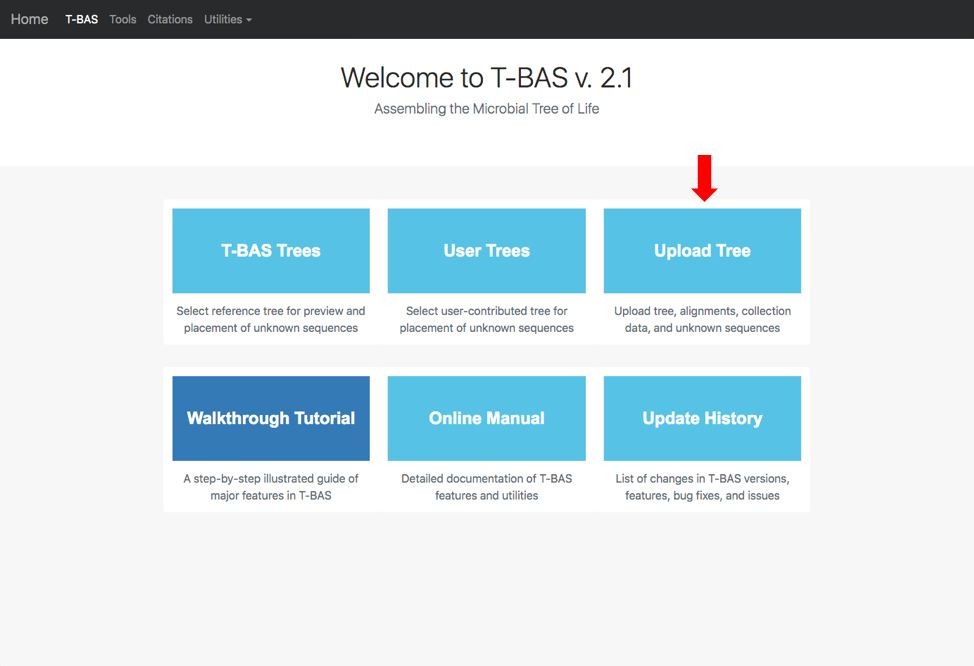
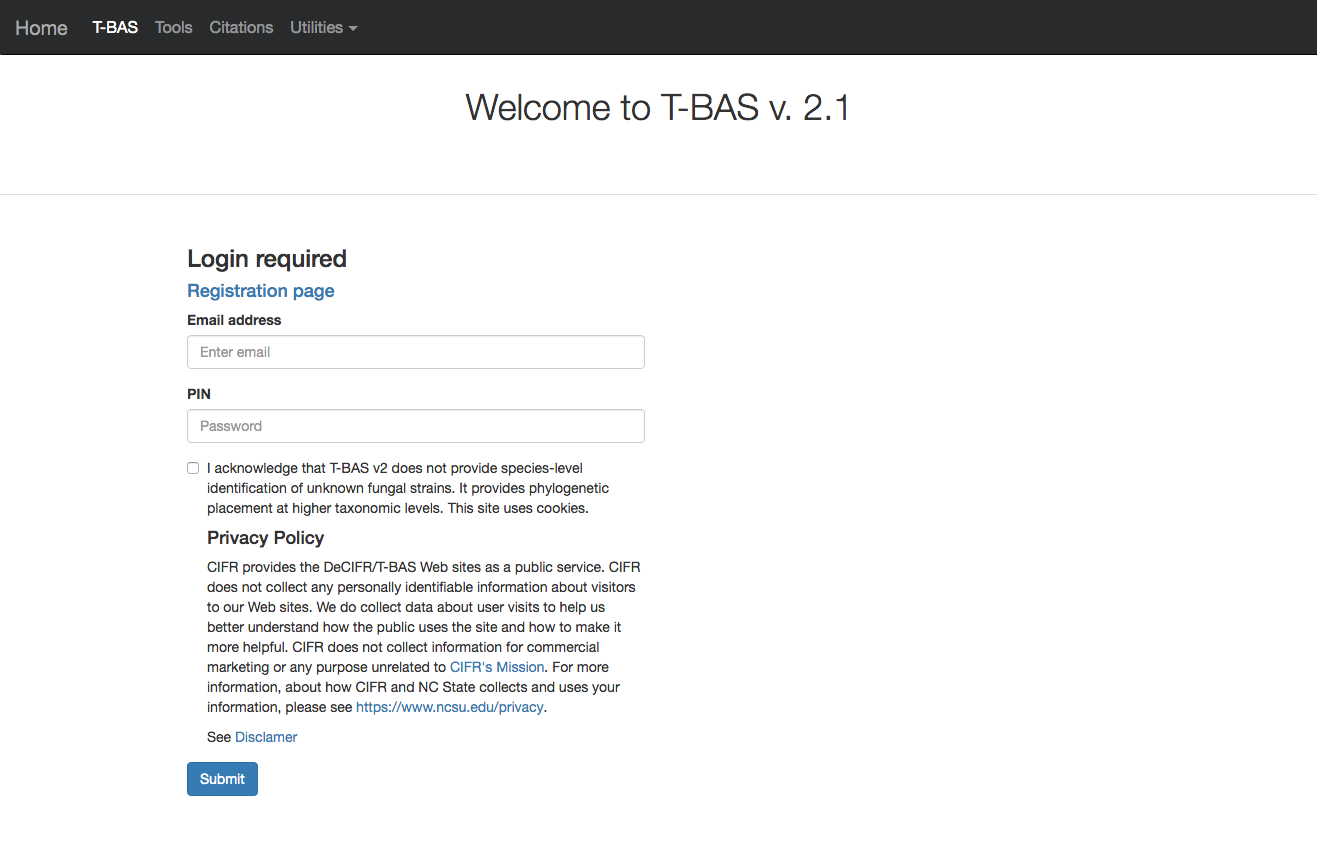
**T-BAS v2.1 Tutorial 4: Using cifr phyloXML to cumulatively update trees**

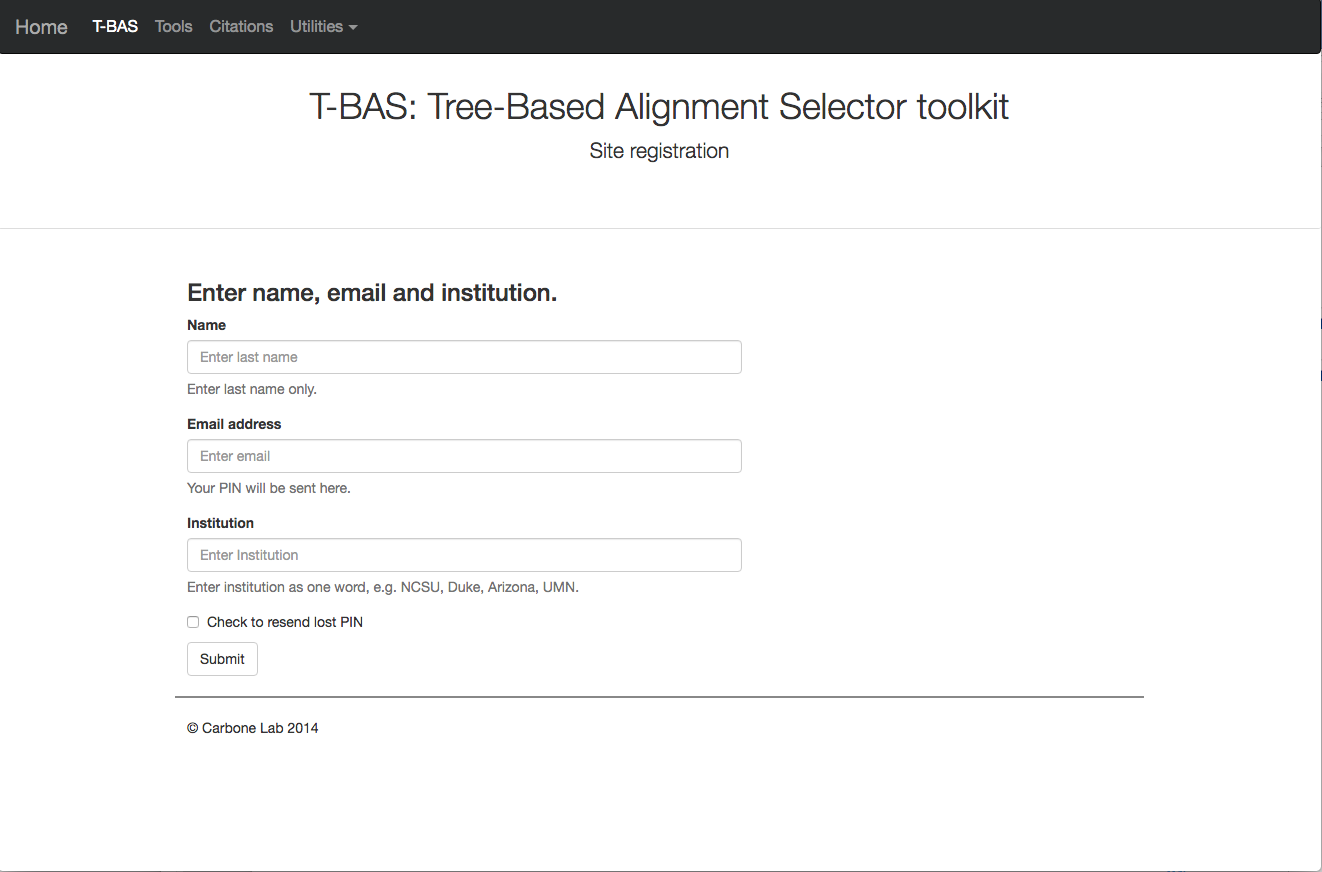
1. In this tutorial we will show how you can perform consecutive placements on the same tree. This is useful to build on a previous placement run with additional sequence data. To start, go to the T-BAS start page (<https://tbas.hpc.ncsu.edu/start>) and click on the Upload Tree button (see arrow).



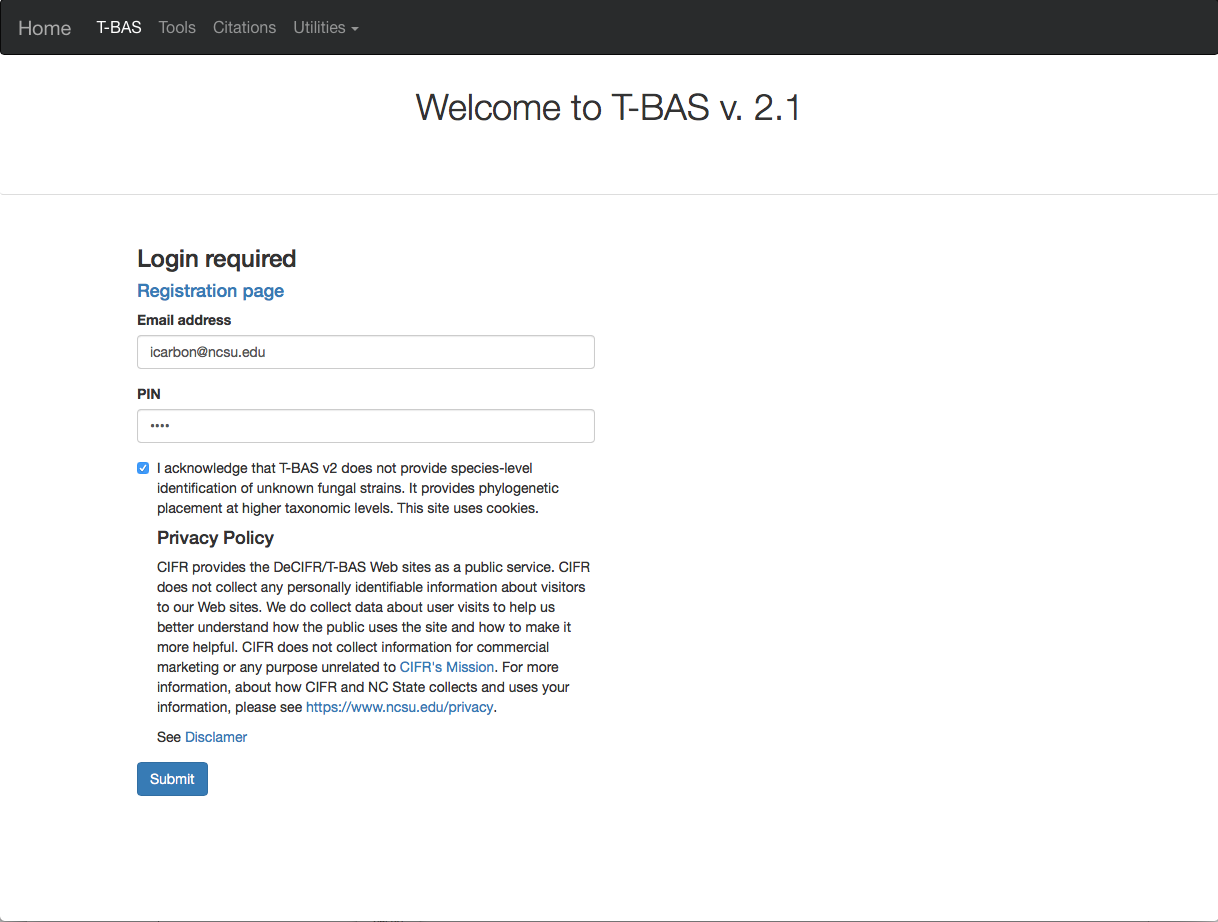
1. You will be required to login. If you are first time user, you will need to register by clicking on the **Registration page** link (see arrow).



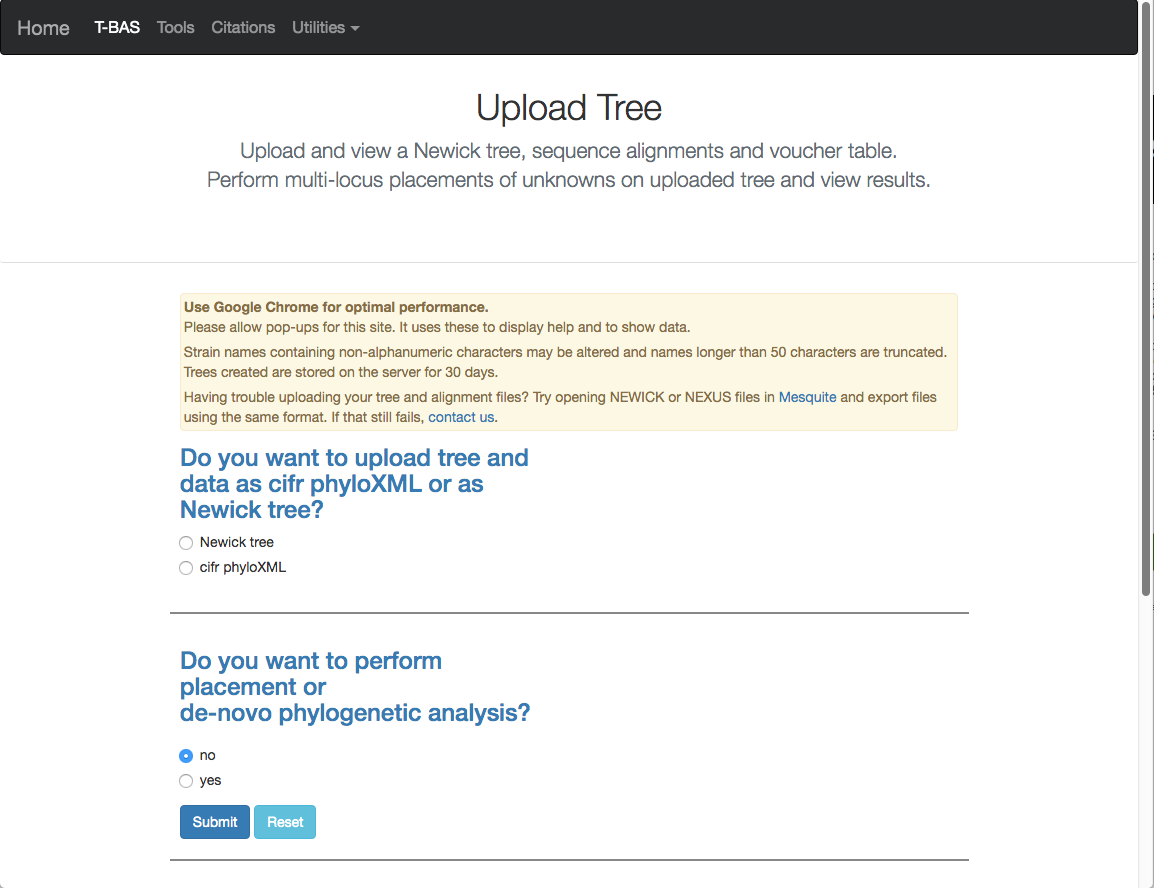
If you already registered but forgot your PIN you can go to the Registration page and **Check to resend lost PIN** (see arrow).



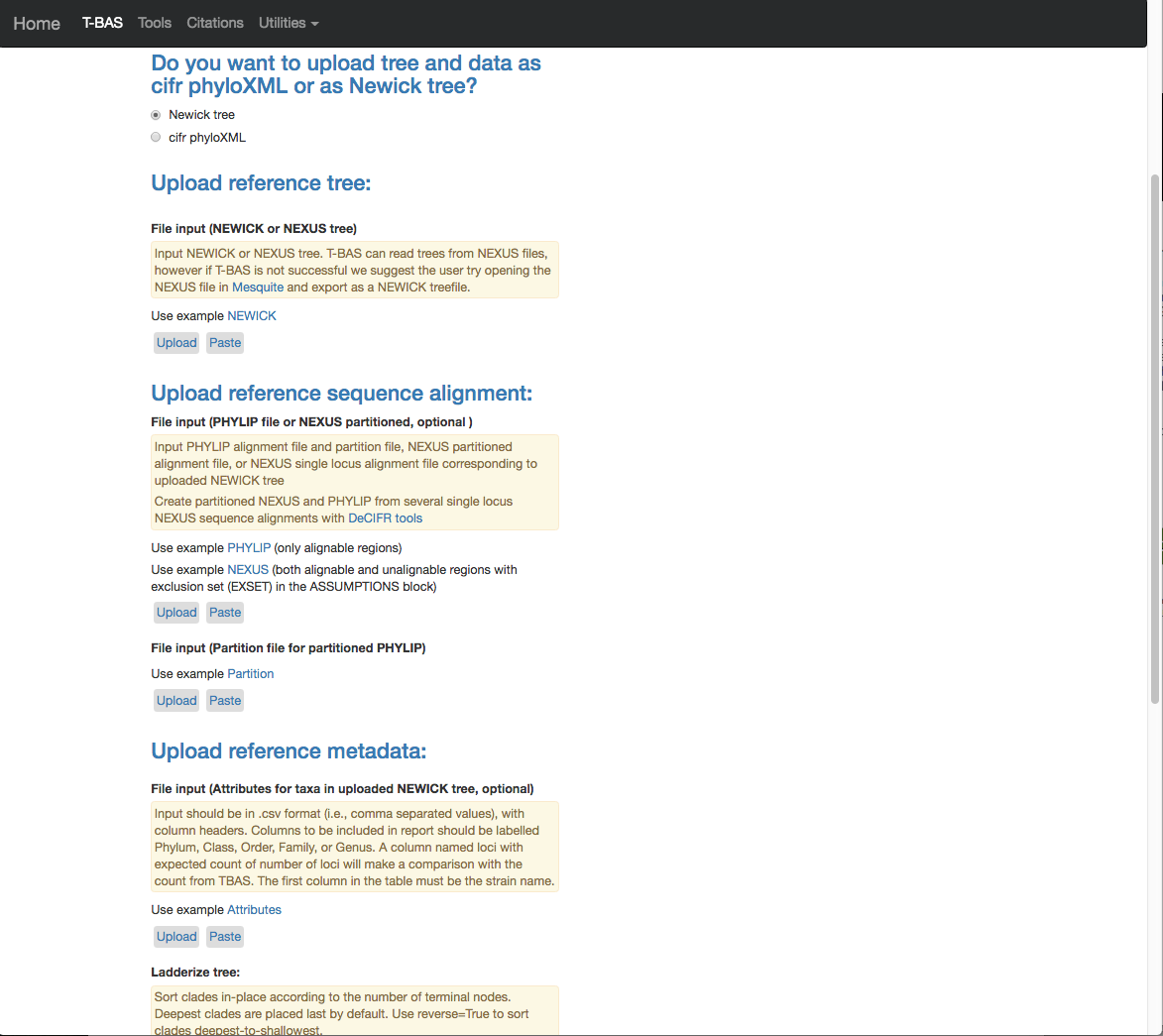
1. After registering you will be returned to the login page. Enter your email address, PIN and check the box acknowledging your acceptance of our disclaimer and privacy policy. Then click **Submit**.



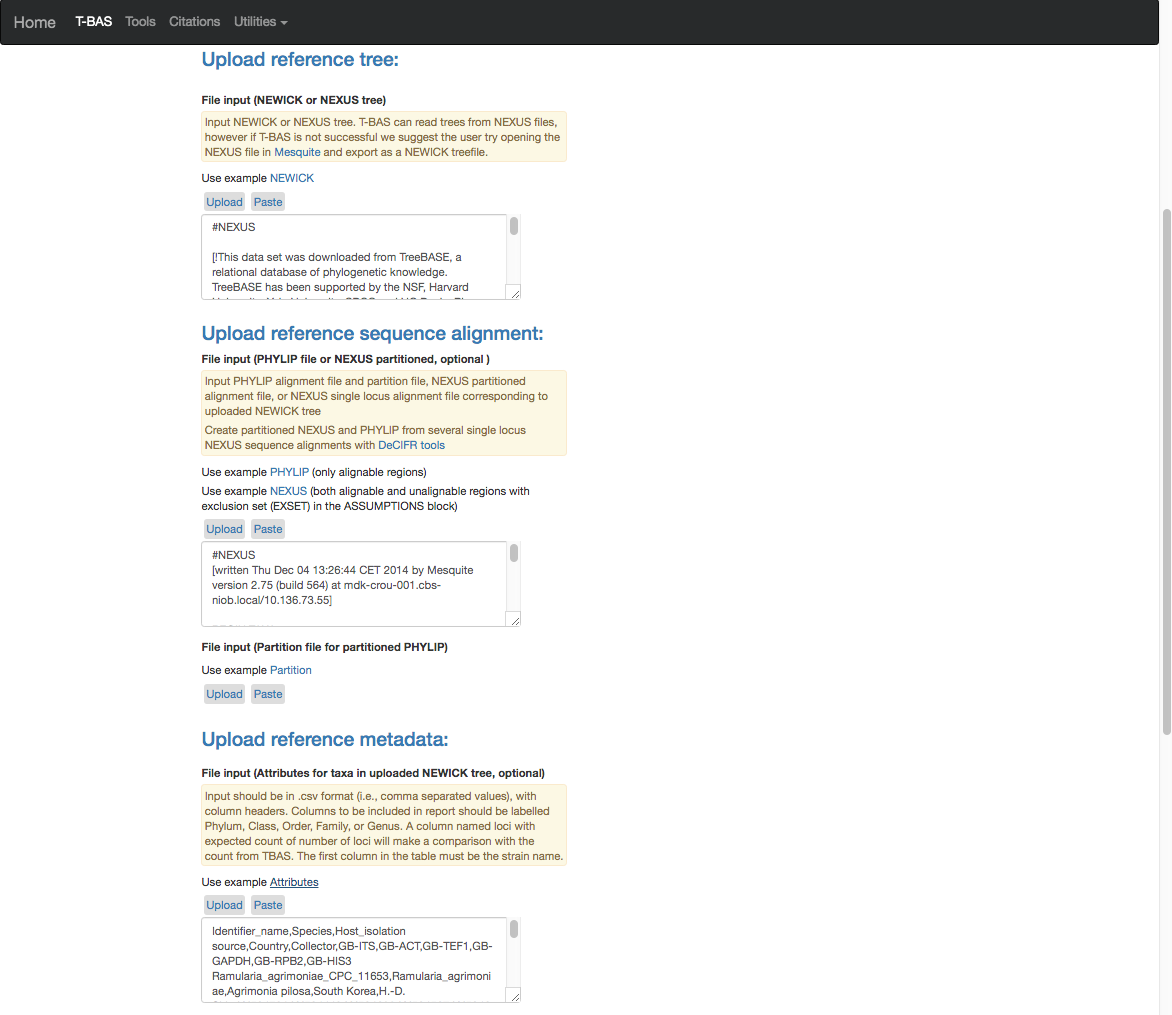
1. After login, select **Newick tree**



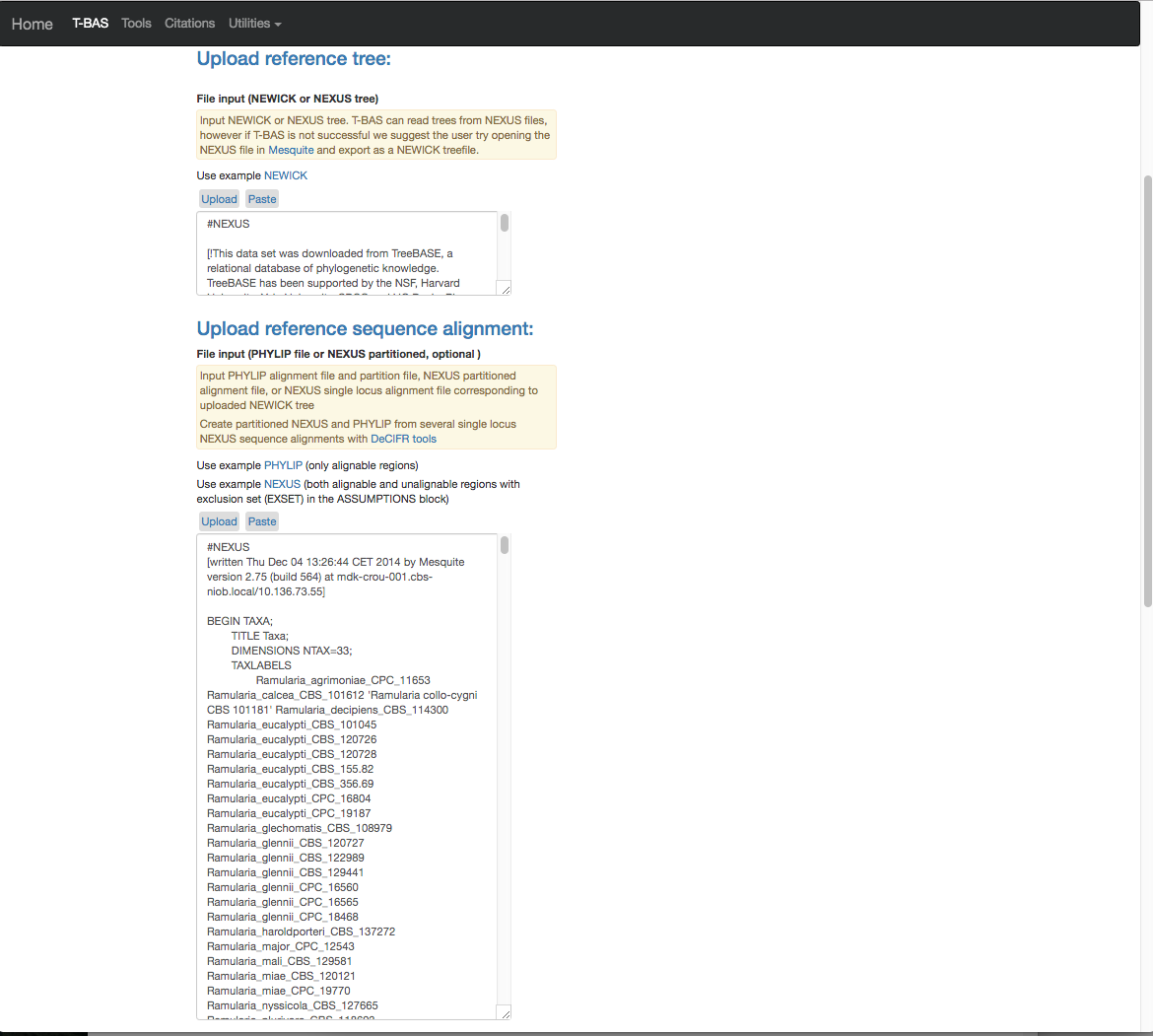
1. This will expand the display to show three upload sections: **Upload reference tree, Upload reference sequence alignment and Upload reference metadata**. Click on the example file links (see arrows) for each section to load the files.



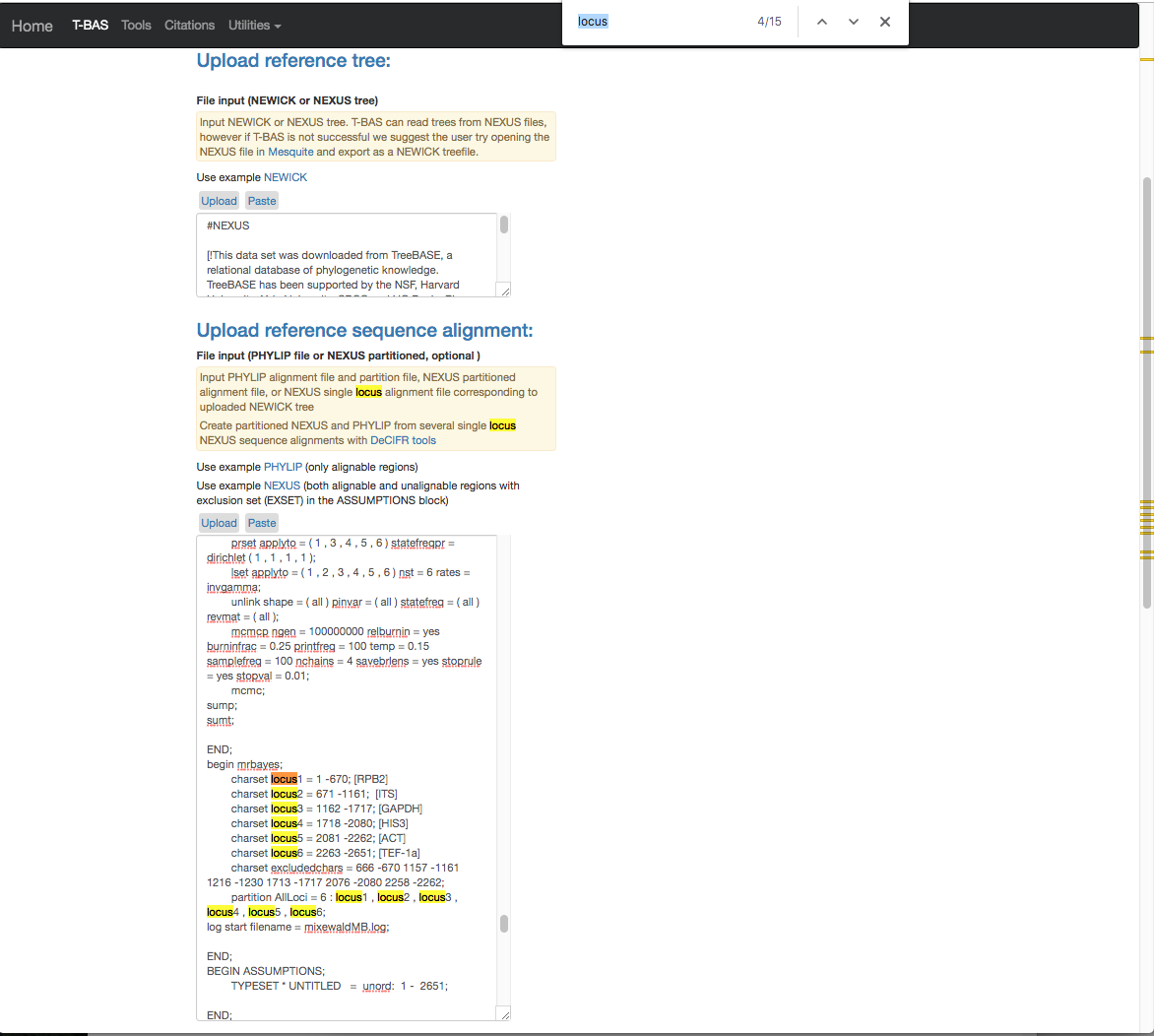
1. The example files will now be visible and can be edited. Some minor edits will be required to the Nexus sequence alignment file. To see more of the alignment file, click and drag down the corner box (see arrow).

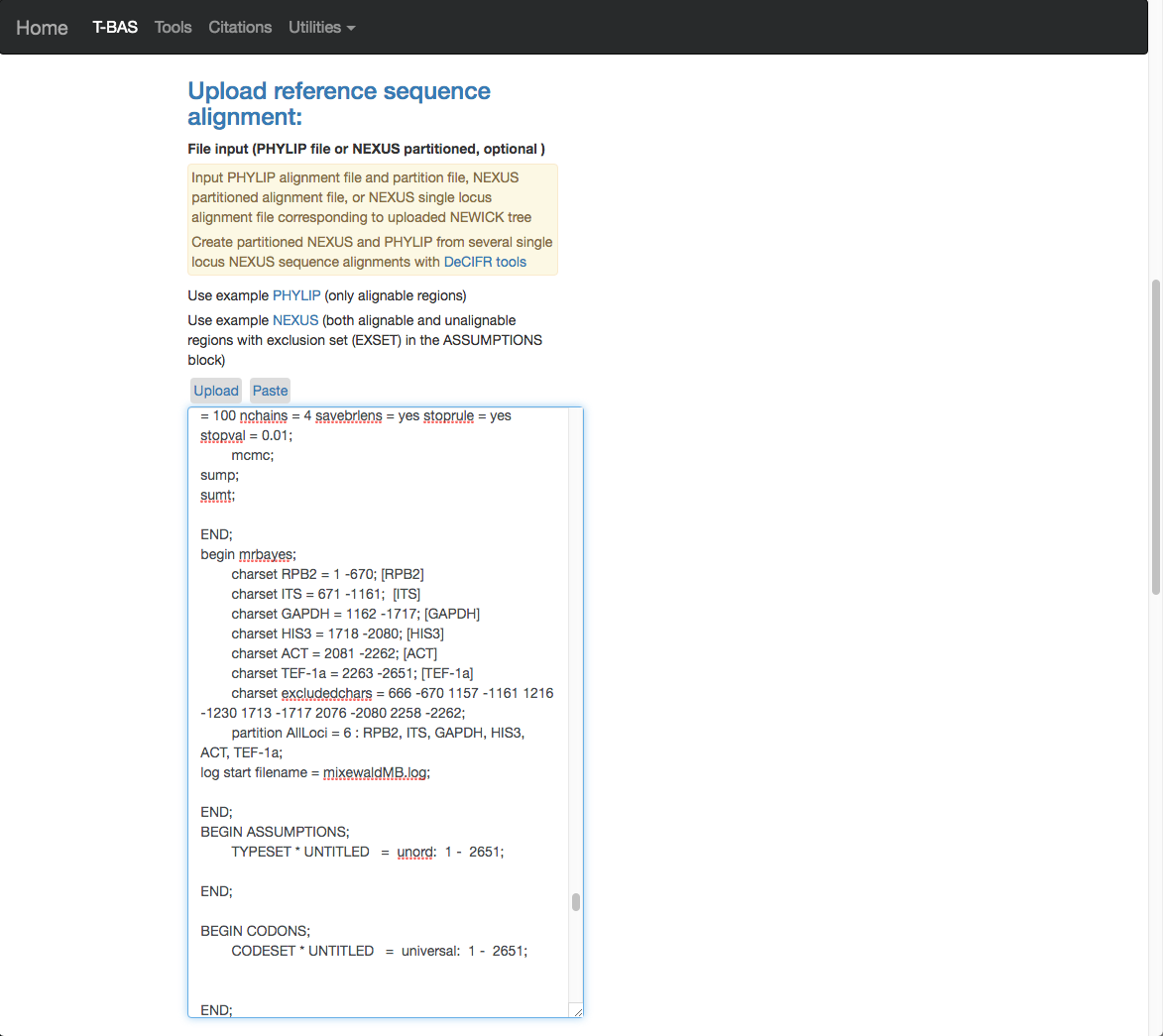
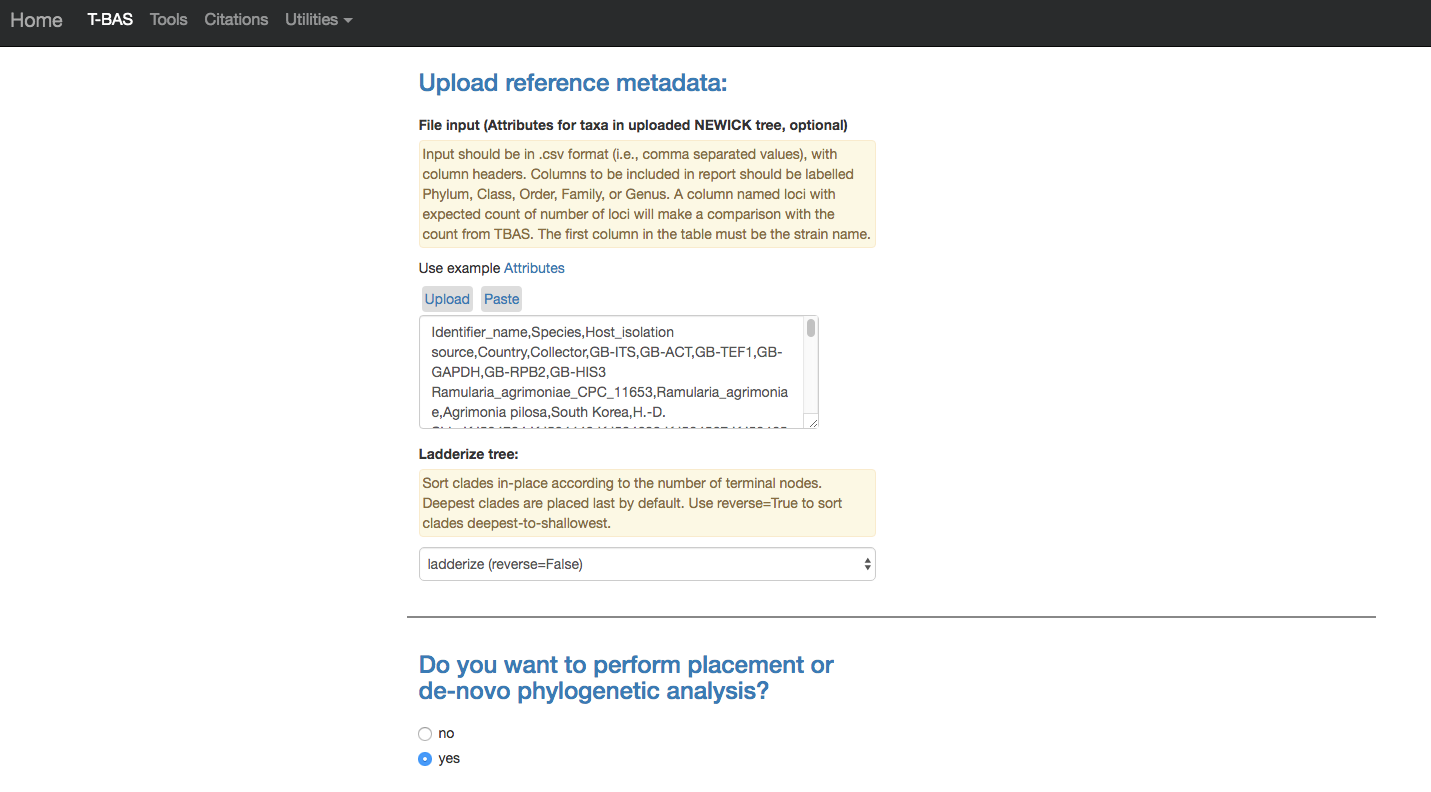


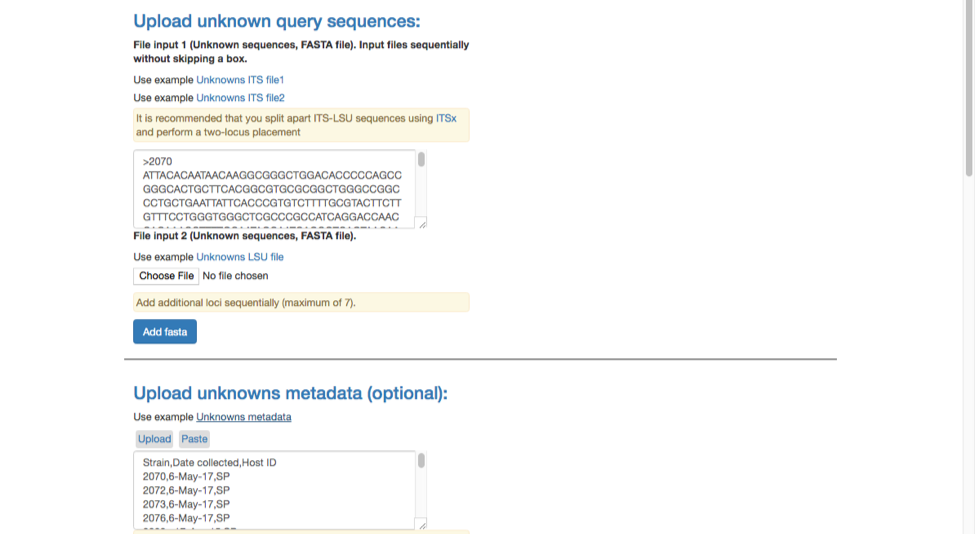
1. The expanded window will look as shown below. Click in the box and search for “locus” in your browser (Ctrl-F).



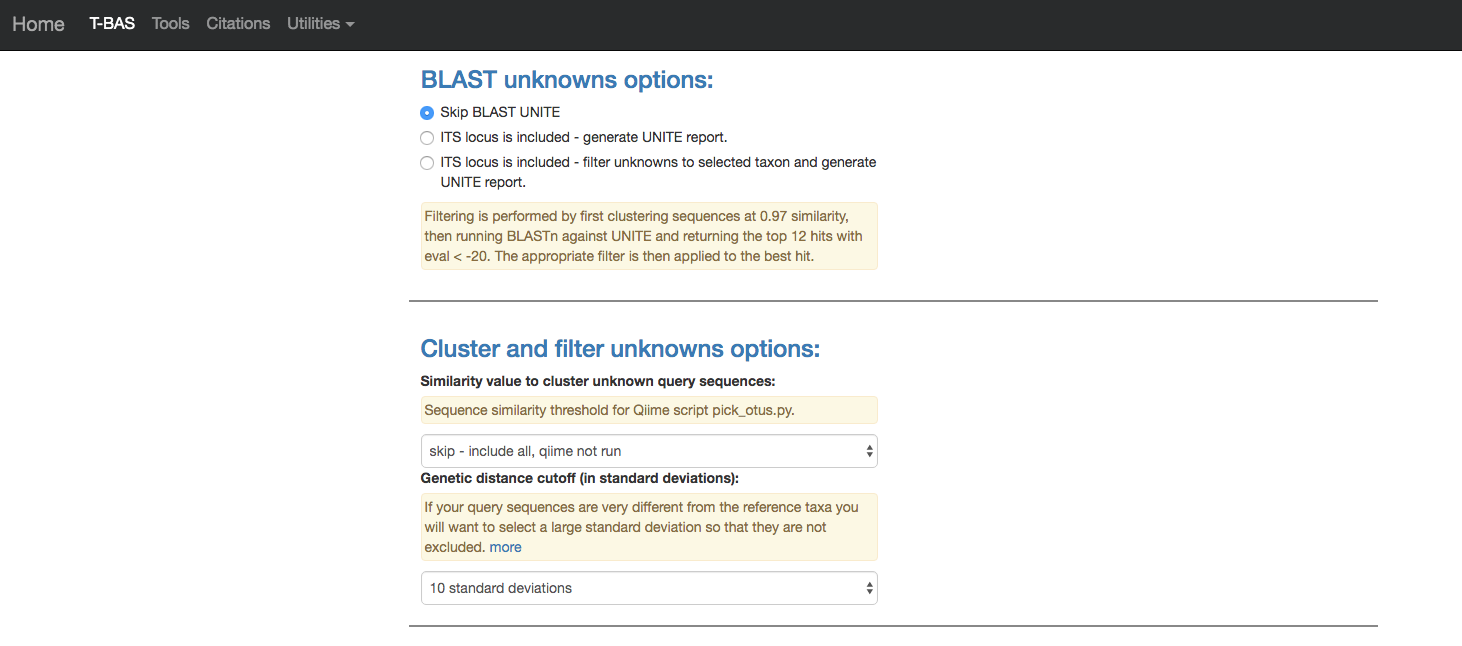
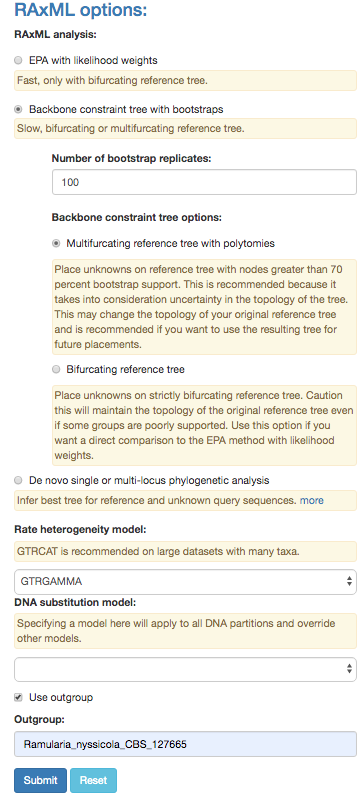
1. The search will highlight the loci that are defined in the character partition block and preceded by “charset” (see arrow). If no changes are made, the loci will be labeled as locus1, locus2, locus3, locus4, locus5 and locus6 in T-BAS. These names make it difficult to track these loci when performing placements and retrieving data. In this example, the actual name of the locus is provided at the end of the charset line. For example, for locus1 the name is RPB2. Note all locus designations in the alignment file (highlighted in yellow and orange) will need to be relabeled – see next step.

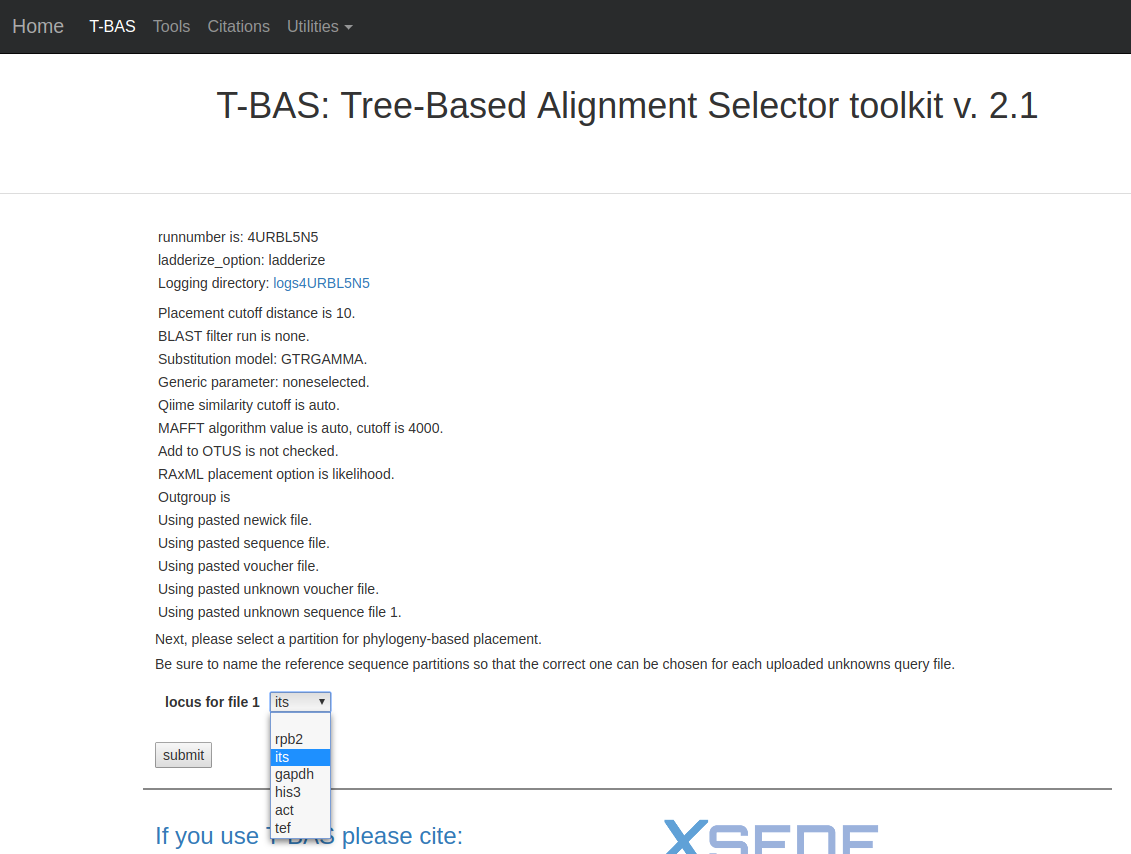
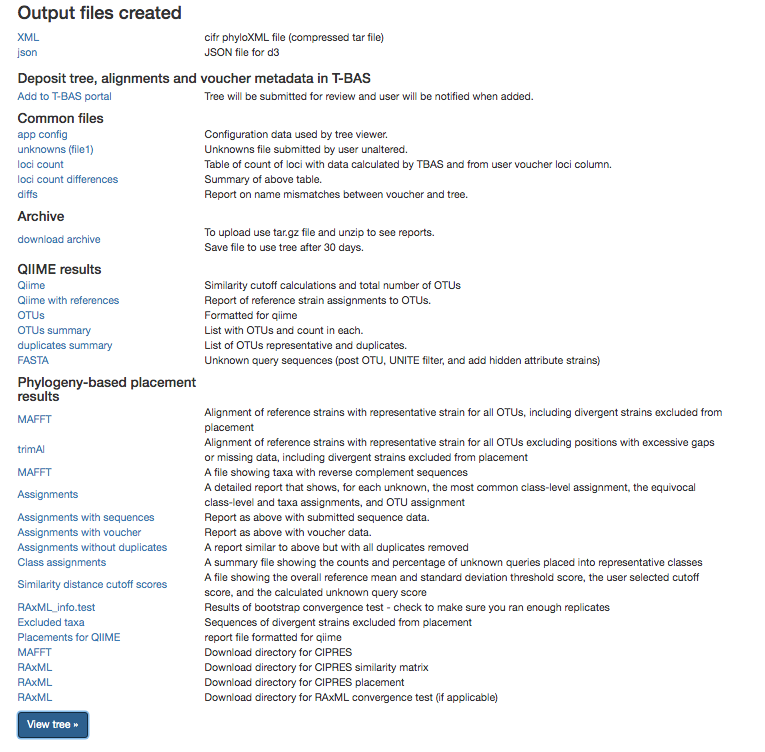


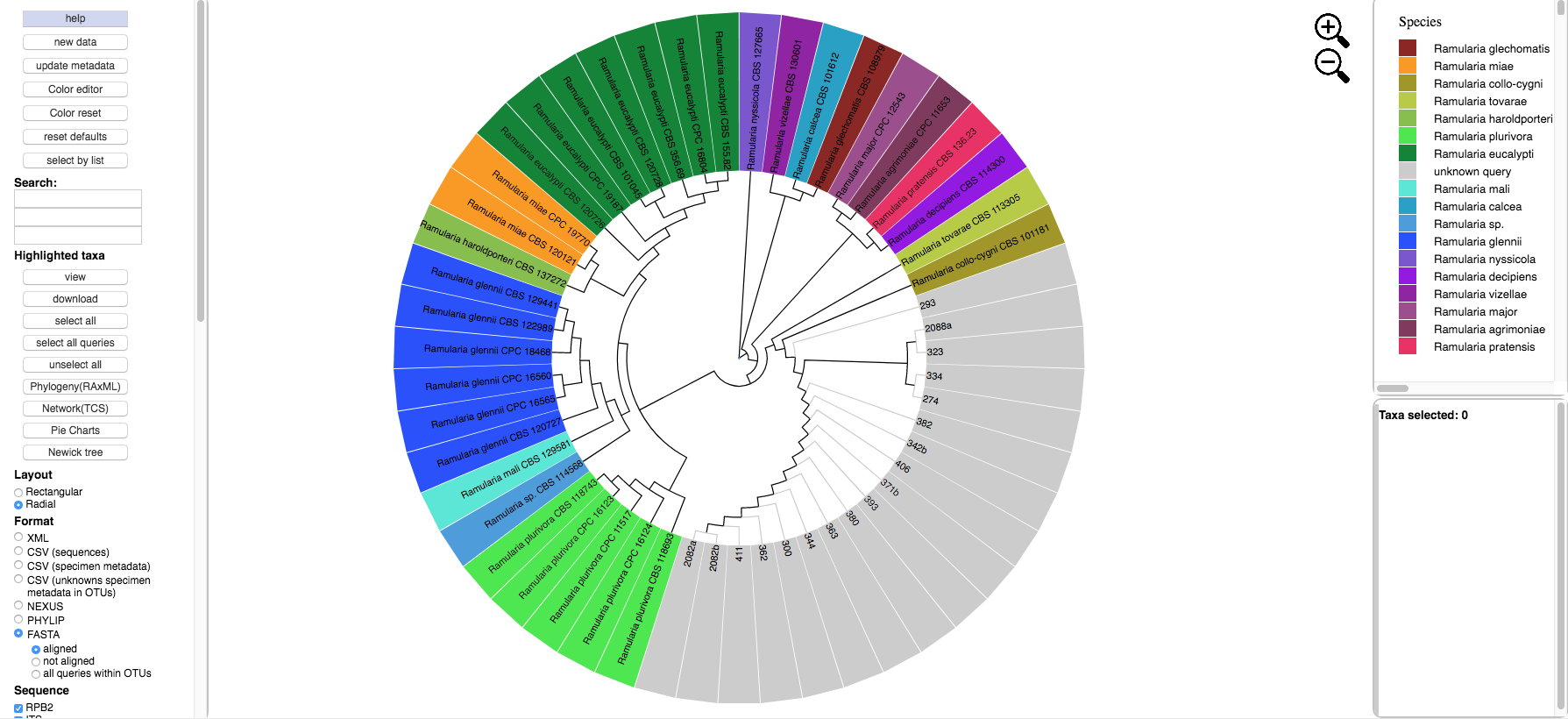
1. Here are the relabeled loci. T-BAS can now correctly track the loci and the user will know which target locus to select when performing a placement.
2. Scroll to the next section where you will be able to view the reference files or perform a phylogeny-based placement using those files. Select **yes**.
3. Select **Unknowns ITS file 1** and **Unknowns metadata** example files.

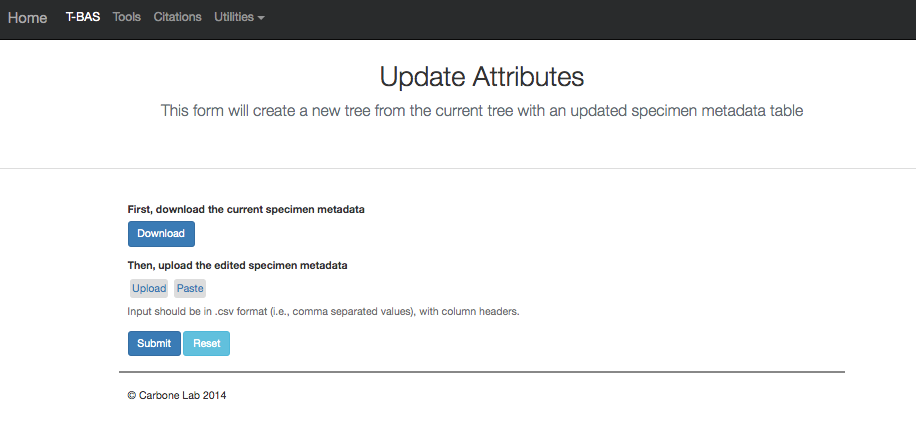
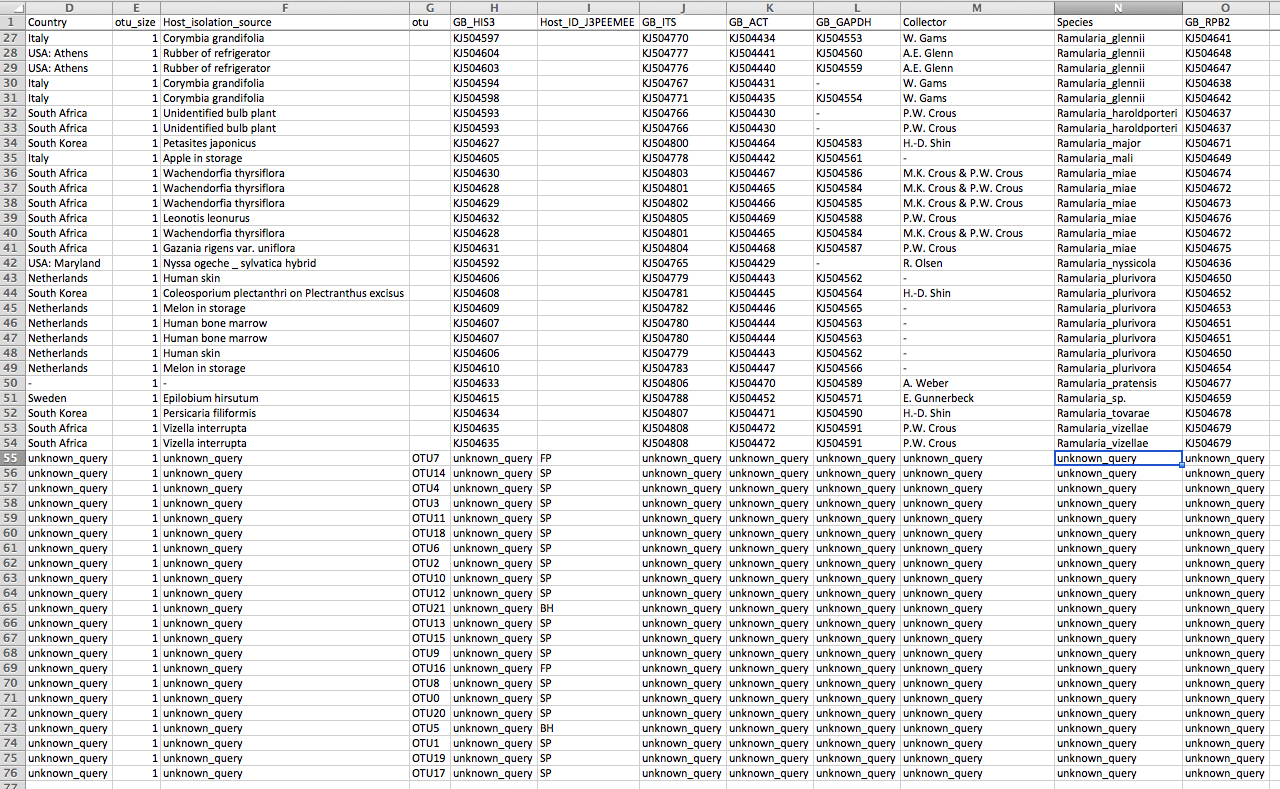
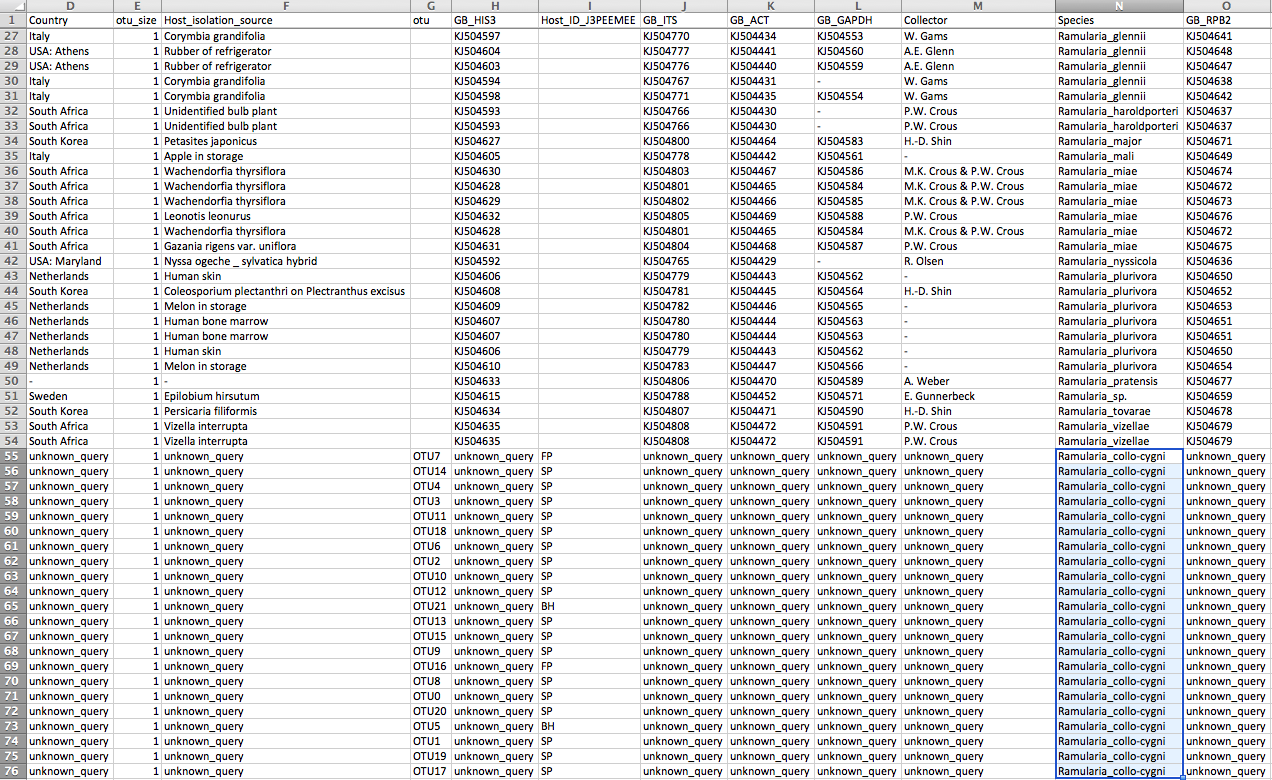
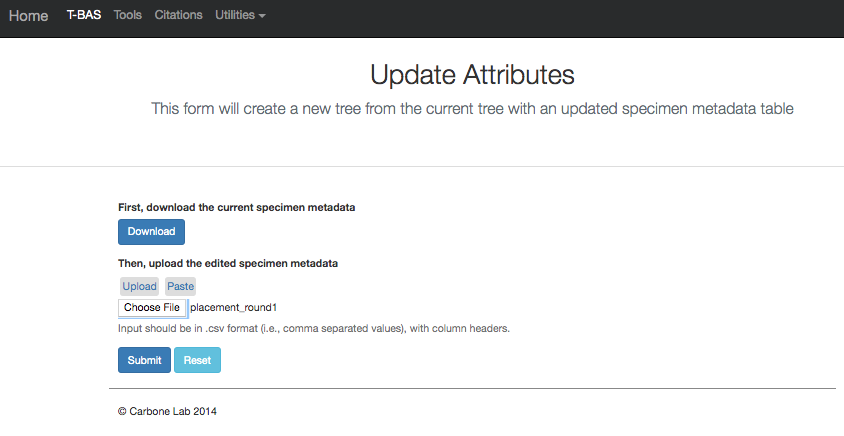
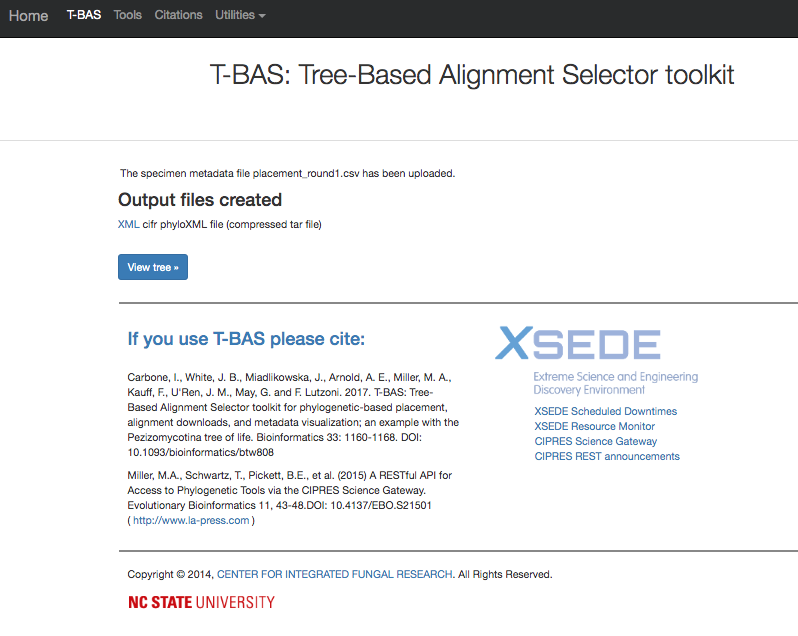
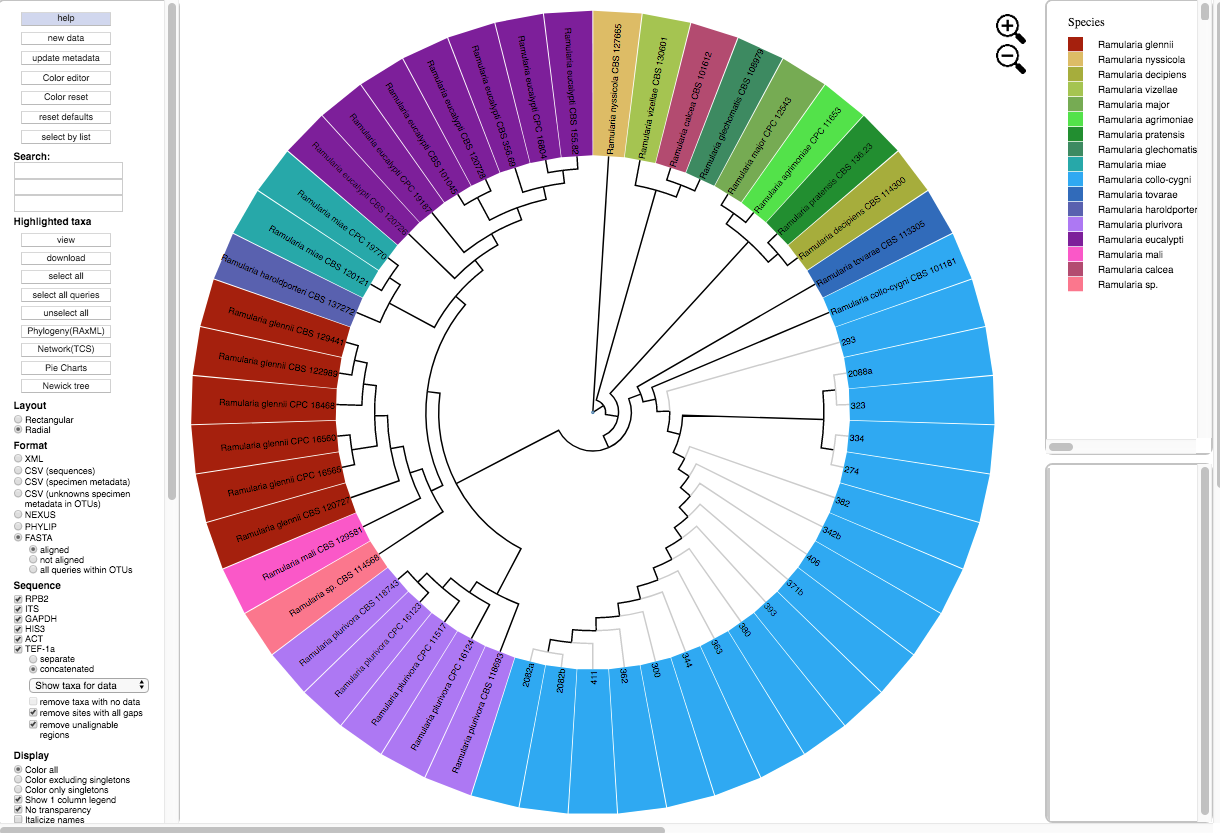


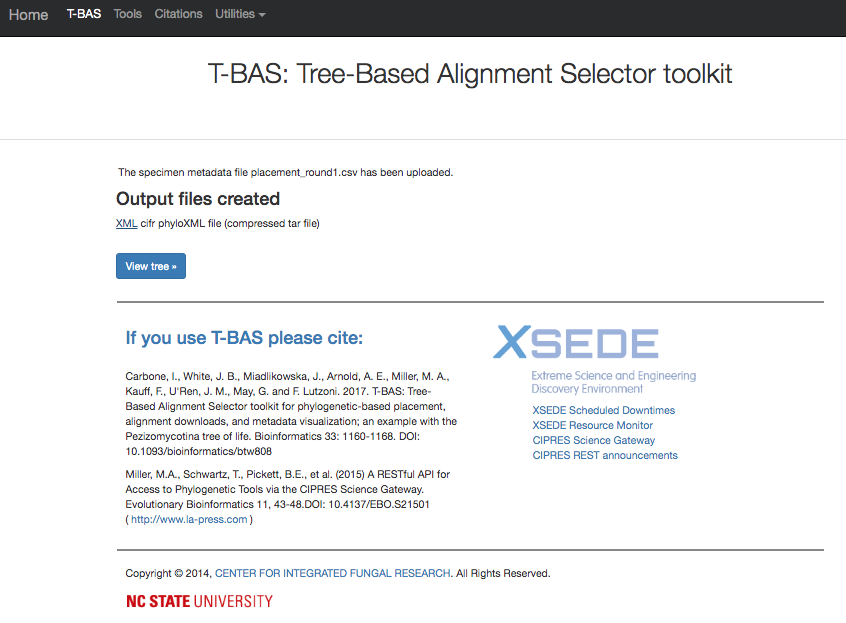
1. Ensure selections are selected as shown:
2. Skip BLAST UNITE
3. Skip-include all, qiime not run
4. Backbone constraint tree with bootstraps
5. Multifurcating reference tree with polytomies
6. Outgroup: Ramularia\_nyssicola\_CBS\_127665
7. Click **Submit**

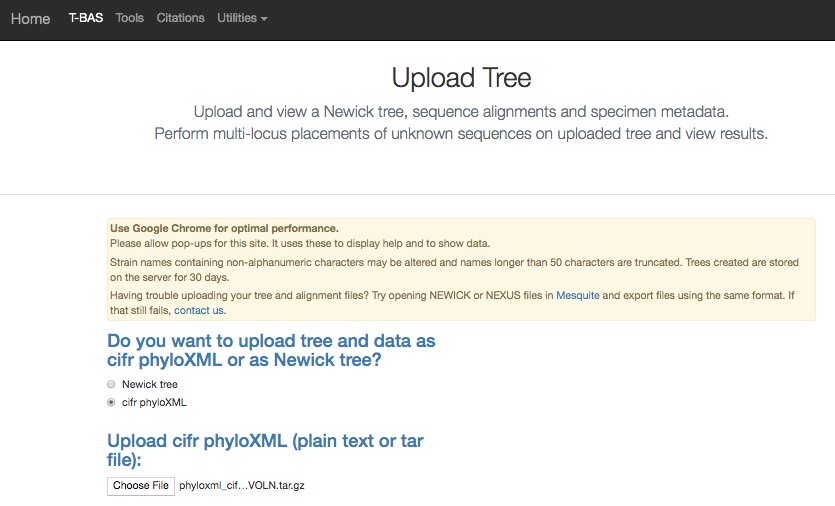
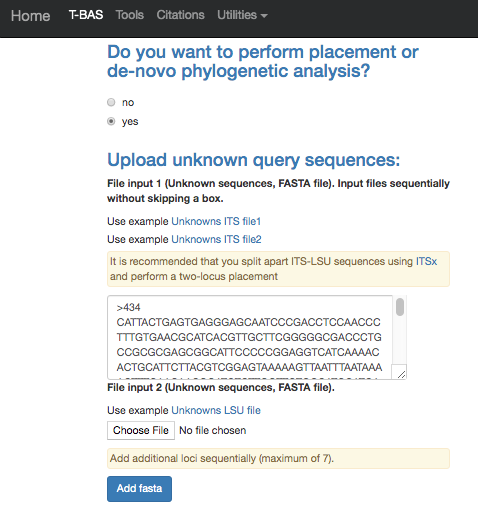
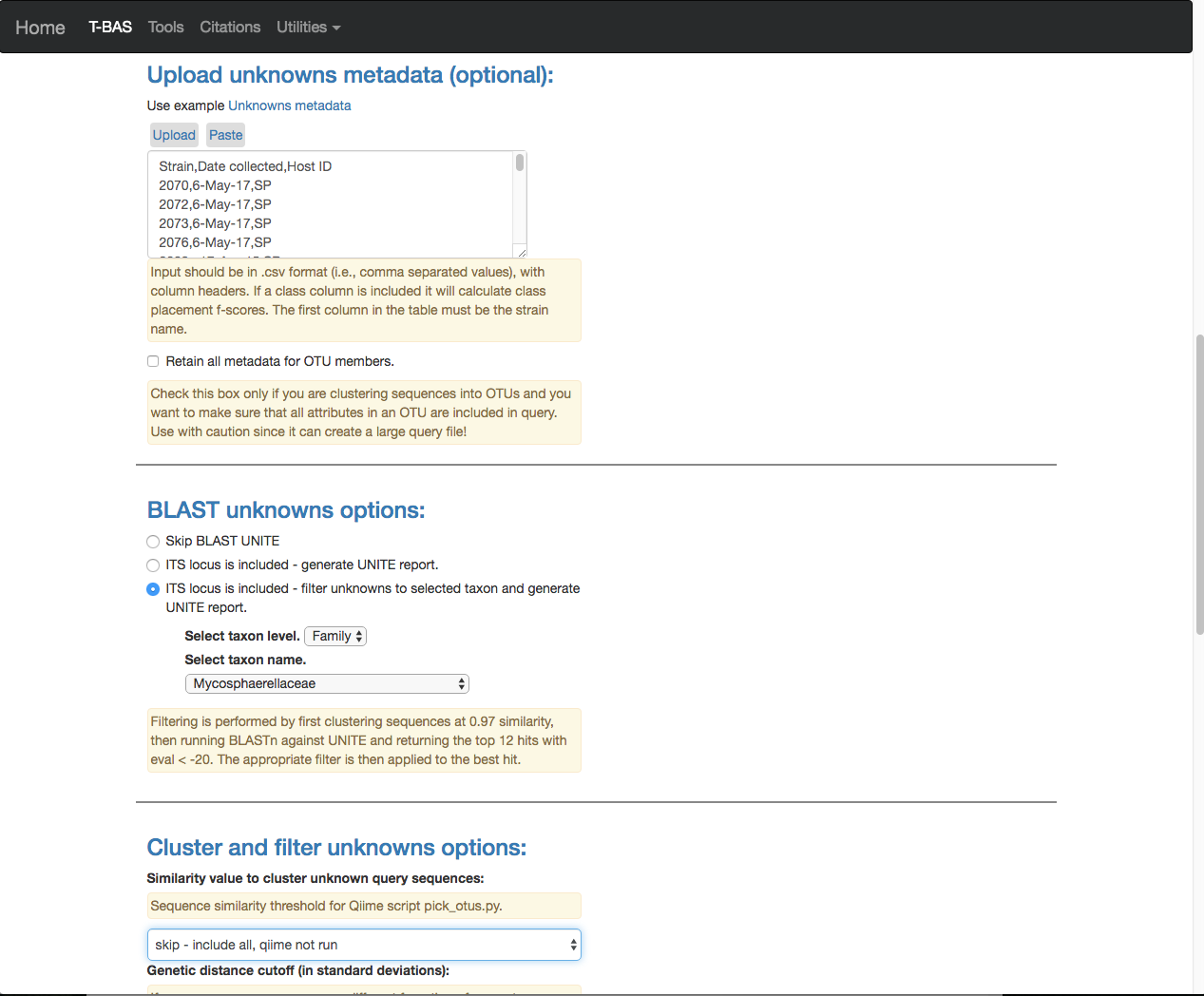
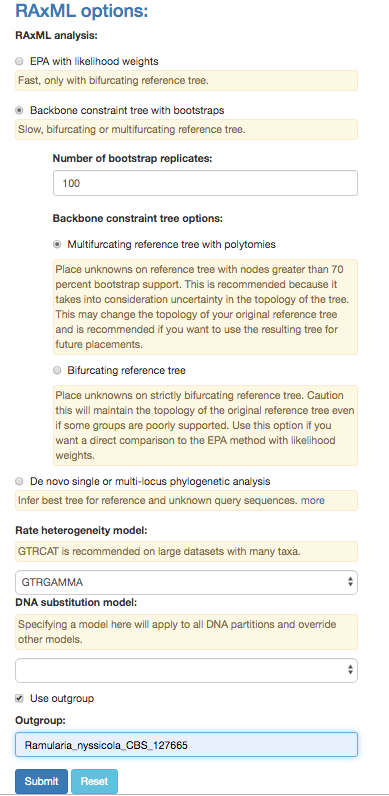
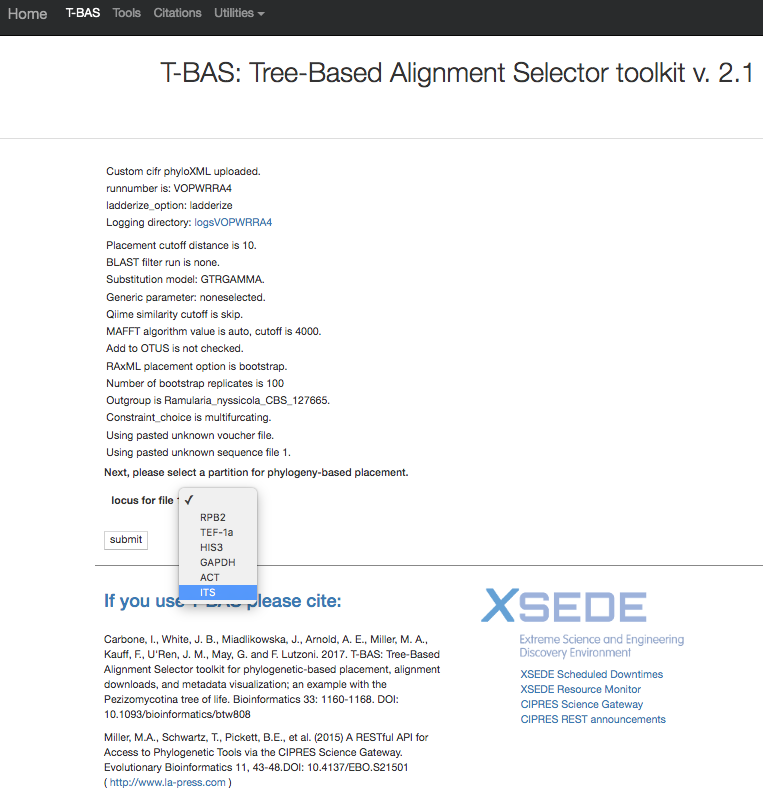
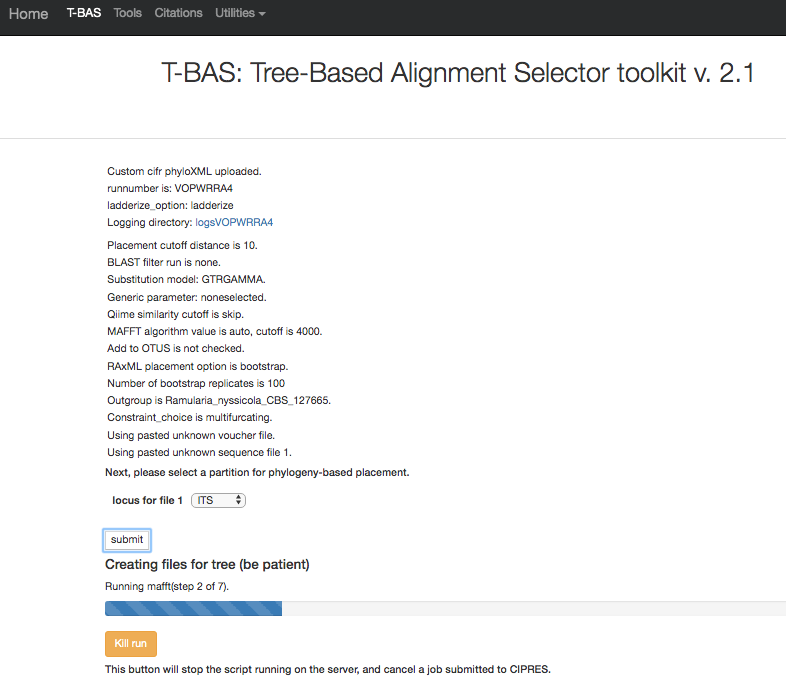
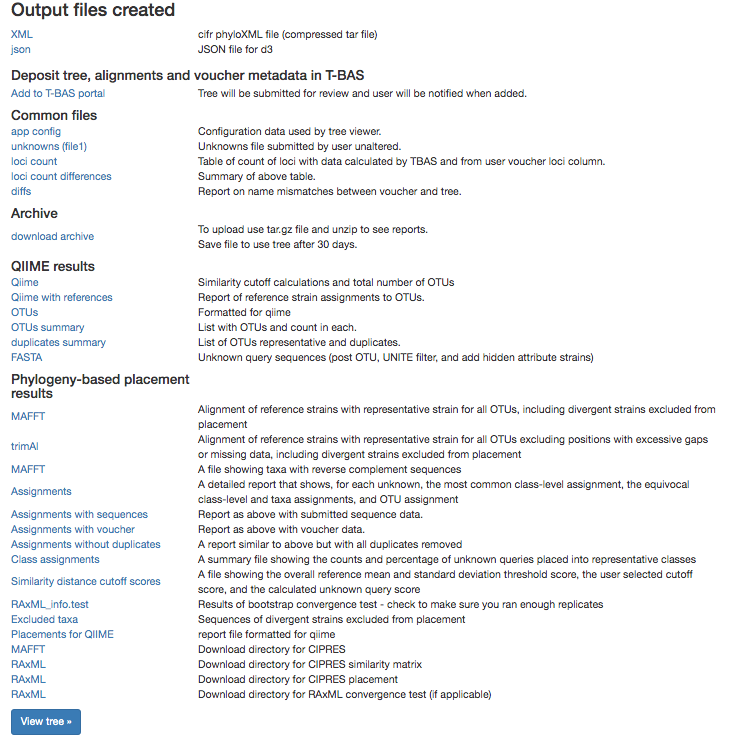
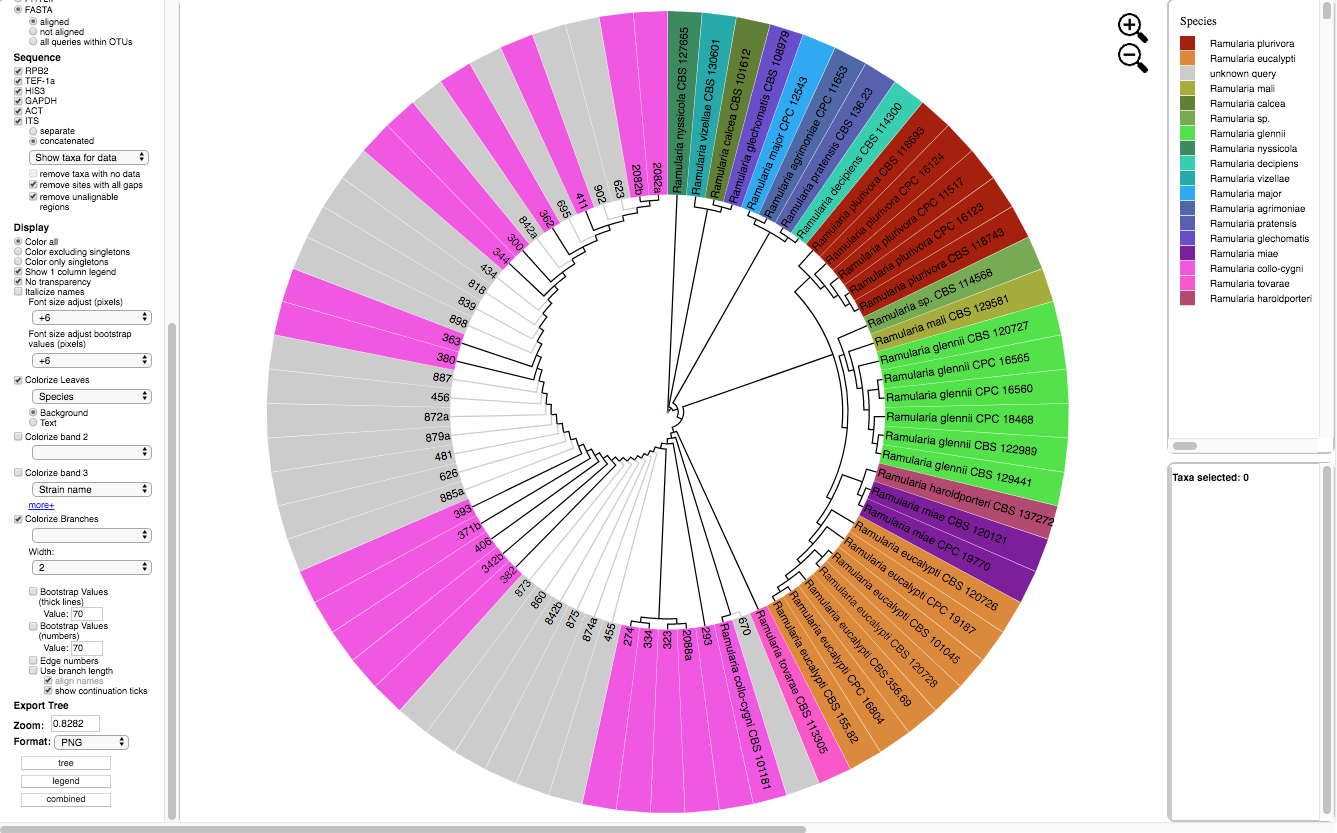


1. On the next page select the **ITS** partition for alignment of unknowns and **submit**.
2. The run takes about 10 minutes to complete.
3. Once the run is complete, click **View tree** to see the placements.
4. 
5. 15) Here is the tree with the following options selected:
6. **Zoom** 0.6103
7. **Font size** +6
8. **Font size adjust bootstrap** +6
9. **Colorize leaves** by Species
10. Branch **Width** 2
11. The **branch lengths** are drawn to scale
12. Next click the **update metadata** button to edit the attributes table.



1. Clicking the download button will download the current specimen metadata. Download and open the spreadsheet. Take note of the filename.
2. 
3. Sort the spreadsheet by Species so that all of the unknowns are grouped together (unknown\_query).
4. 
5. Edit these unknown\_query entries in the species column to match what was shown in the placement tree, which was Ramularia\_collo-cygni. If the unknown sequences were not placed with high confidence (i.e. cumulative likelihood weights > 0.96), then you might want to assign them to a higher taxonomic scale (e.g. genus) where there is more placement confidence. Save this new file with a different name. In this example, the file was saved as “placement\_round1”.
6. 
7. Go back to the Update Attributes page where the current specimen metadata was downloaded from and **upload** the edited specimen metadata. The edited file could also be copied and pasted into the field, if desired. Click **Submit**.
8. 
9. Once the run is complete, click **View tree**.
10. 
11. A new tree is created with the edited values. All of the unknowns are now labeled as Ramularia\_collo-cygni and the leaves are colored accordingly.
12. 
13. Go back to the results page on step 20 and download the cifr phyloXML file. This new download will contain the updated specimen metadata information. Notice the filename of the tar file contains “placement\_round1”, the new name used to rename the specimen metadata with the assignments determined from the first placement.



1. Return to the T-BAS start page (from step 1) and select **Upload Tree**. Click on **cifr phyloXML** and upload the downloaded XML file from step 22. Select **yes** in the placement section and select the **Unknowns ITS file2** example file.
2. 
3. 
4. Click on the **Unknowns metadata** example file. Next, we will illustrate how T-BAS can filter spurious ITS sequences using BLAST of the UNITE database. In the BLAST unknowns options section select **ITS locus is included – filter unknowns to selected taxon and generate UNITE report**. Then select the next highest taxonomic level for filtering – in this example, select the family Mycosphaerellaceae as shown below. In the Cluster and filter unknowns options section, select **skip-include all, qiime not run**.
5. 
6. In the RAxML options section, change selection to **Backbone constraint tree with bootstraps** and select **Multifurcating reference tree with polytomies**. Note that when placements are performed using EPA there will be multifurcations in the resulting tree and in subsequent placements only the backbone constraint method can be used. Use Ramularia\_nyssicola\_CBS\_127665 as the outgroup. Click **Submit**.
7. 
8. 26) Select the **ITS** locus file then select **submit**.
9. 
10. 27) The placement will take about 5-10 minutes to run.
11. 
12. 28) Once the run is complete, the results page will display. Select **View tree** to see the tree.
13. 
14. The tree created by the second run is displayed using the following selections:
15. **Zoom:** 0.8282
16. **Font size:** +6
17. **Font size adjust bootstrap:** +6
18. **Colorize leaves** by Species
19. Branch **Width**: 2
20. The **branch lengths** are drawn to scale
21. 
22. 30) Clicking on update metadata on the tree (as shown in step 15) and clicking download (as shown in step 16) will pull up the new specimen metadata file. Sort the file by specimen. Now the edits made in step 18 are shown here along with the new unknown placements. Note: This view was edited to show the strain name and species column together for illustration purposes. The file has many other attributes in between these two columns.
23. 