

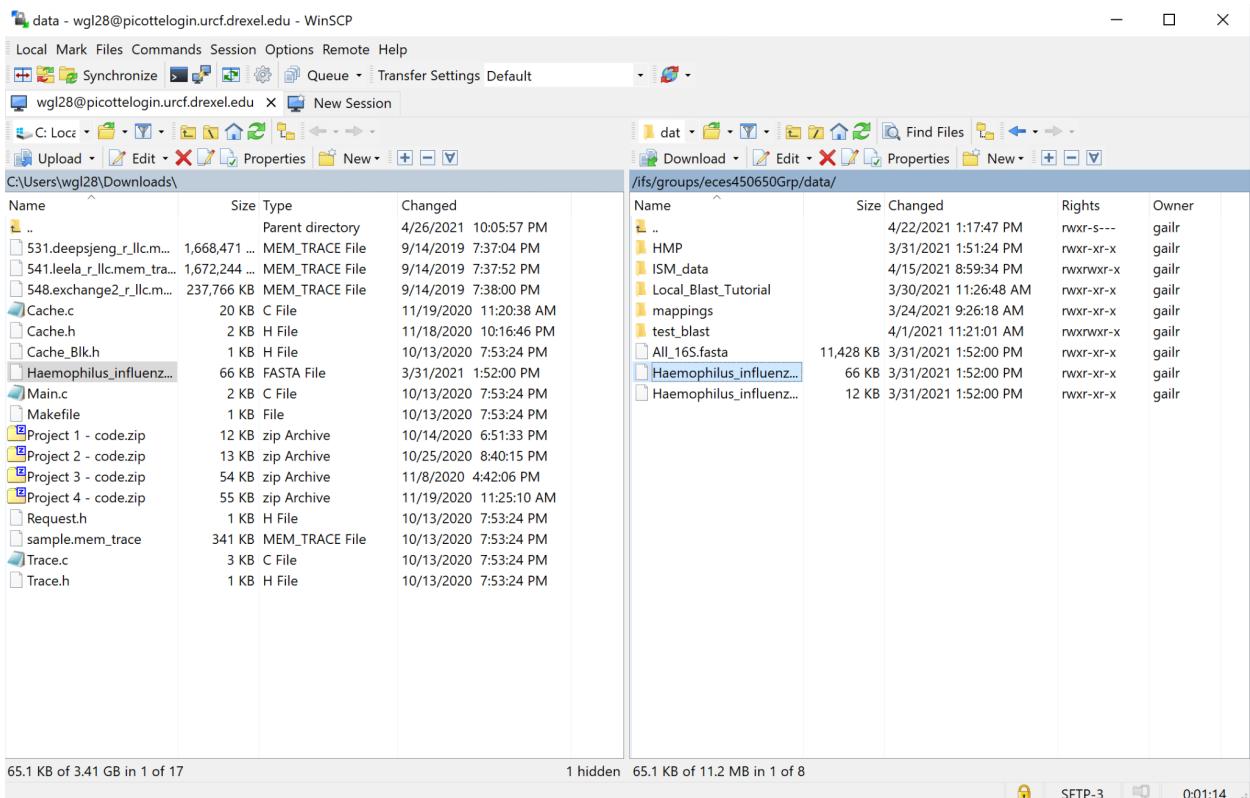
Jonathan Palko and Willow Livengood  
ECES650  
Dr. Rosen  
4/26/2021

## Tutorial 2: Tree Inference

### Procedure to Get Ready

1. Download the file from Picotte stored at

*/ifs/groups/eces450650Grp/data/Haemophilus\_influenzae\_16S.fasta* to your local machine using whichever method you are most comfortable with. In this case I used WinSCP and copied the file into my downloads folder.



2. Log into CIPRES

CIPRES Login:

**Username/Email**

**Password**

**Login**    **Reset**

3. Create a new folder, and enter a label and description.

Folders  
Total Storage: 643.46 KB

- ↳ Tutorial 2
  - ↳ Data (3)
  - ↳ Tasks (2)

Welcome wgl28 to the CIPRES Science Gateway V 3.3. If you would like to view a demo, click [here](#).

**Create New Folder**

**Current Folder Details**

<b>Label</b>	Tutorial 2
<b>Description</b>	MAFFT and RaxML

**Create Subfolder**    **Edit Folder**    **Delete Folder**

Enter Folder Detail

**Label**

**Description**

**Save**    **Cancel**

4. Open the data subfolder by clicking on Data.

The screenshot shows the CIPRES Science Gateway interface. On the left, there is a sidebar titled "Folders" with a "Total Storage: 643.46 KB". Inside, there are three folders: "Tutorial Example" (selected and highlighted in yellow), "Data (0)", "Tasks (0)", and "Tutorial 2". A red box highlights the "Data (0)" folder. To the right, a message says "Welcome wgl28 to the CIPRES Science Gateway V 3.3. If you would like to view a demo, click here." Below this is a "Create New Folder" button. Under "Current Folder Details", it shows "Label: Tutorial Example" and buttons for "Create Subfolder", "Edit Folder", and "Delete Folder".

## 5. Upload data into CIPRES by clicking on Upload/Enter Data.

This screenshot shows the same interface as above, but the "Data (0)" folder is now selected and highlighted in yellow. The message in the center says "There is currently no data in this folder. What would you like to do?". Below are two buttons: "Upload/Enter Data" (highlighted with a red box) and "Create a Task".

Choose the *Haemophilus\_influenzae\_16S.fasta* file and hit save.

This screenshot shows the "Upload File" dialog. On the left, the "Folders" sidebar is visible. The main area has a title "Upload File" and a sub-section "Upload your files" with a "Choose Files" button where "Haemophil...e\_16S.fasta" is selected. It also includes instructions: "You can select multiple files." and "MSIE 9 and below support single uploads only.". Below this is a section "You can also enter your data manually below" with fields for "Label (required)" (with "Label" entered) and "Data:" (with "Enter your data" entered). At the bottom are "Save" and "Cancel" buttons.

This is what it should look like when you have successfully saved the file.

The screenshot shows the 'All Data' page. On the left, a sidebar displays 'Folders' with 'Tutorial Example' expanded, showing 'Data (1)' and 'Tasks (0)'. Below this is 'Tutorial 2'. A message box states 'There is currently 1 data item in this folder.' In the center, a table lists 'User Data' with one entry: User Data ID 2778397, Label Haemophilus\_influenzae\_16S.fasta, Size 65.13 KB, Data Format Unknown, Date Created 4/26/21, 19:10, and Parse Status Queued. At the bottom, buttons include 'Move' (selected to Tutorial 2), 'GO', 'Delete Selected', and 'Transfer Selected to CIPRES Share'.

## Procedure to Align the Sequences using MAFFT

### 6. Open the Tasks subfolder.

The screenshot shows the 'All Data' page. The 'Folders' sidebar now shows 'Data (1)' and 'Tasks (0)' under 'Tutorial Example', with 'Tasks (0)' highlighted by a red box. A message box states 'There is currently 1 data item in this folder.' The central table for 'User Data' shows the same entry as before. Buttons at the bottom include 'Move' (selected to Tutorial 2), 'GO', 'Delete Selected', and 'Transfer Selected to CIPRES Share'.

### 7. Create a new task.

The screenshot shows the 'All Data' page. The 'Folders' sidebar shows 'Tutorial Example' expanded, with 'Data (1)' and 'Tasks (0)' visible. A large message box in the center states 'There are currently no tasks in this folder.' A prominent blue button labeled 'Create New Task' is highlighted with a red box.

**Folders**

Total Storage: 708.59 KB

- Tutorial Example
  - Data (1)
  - Tasks (0)
- Tutorial 2

**Create new task**

Task Summary    Select Data    Select Tool    Set Parameters

You may edit your task using the tabs above.  
Current CPU Hr Usage: 3 [Explain this?](#)

Description

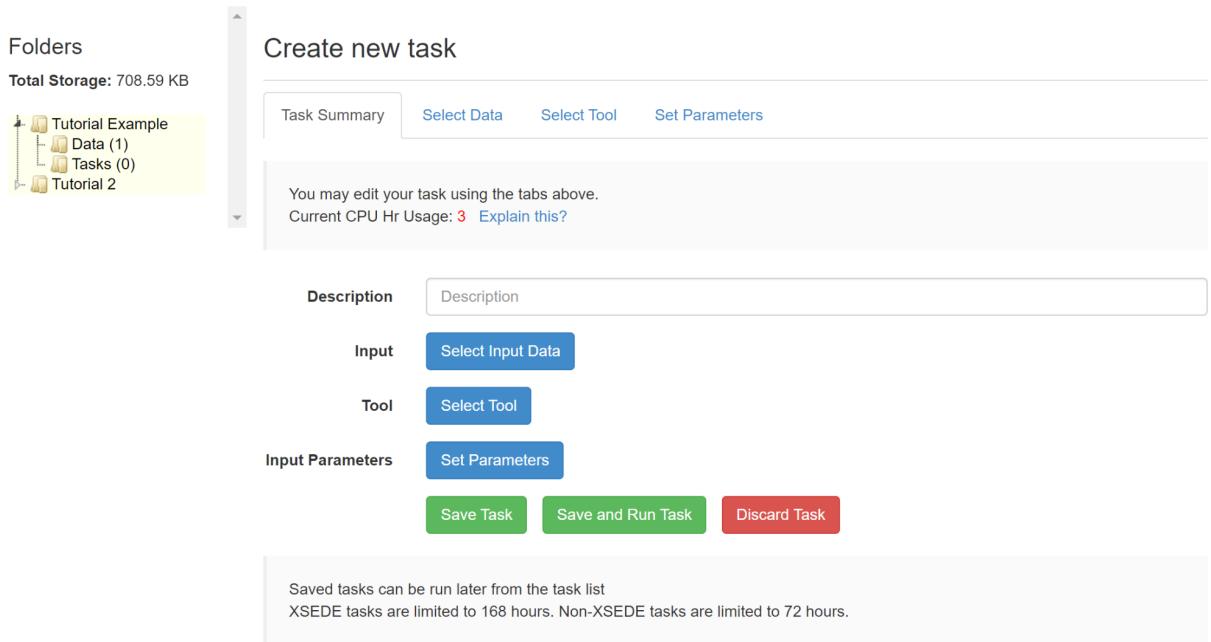
Input [Select Input Data](#)

Tool [Select Tool](#)

Input Parameters [Set Parameters](#)

[Save Task](#) [Save and Run Task](#) [Discard Task](#)

Saved tasks can be run later from the task list  
XSEDE tasks are limited to 168 hours. Non-XSEDE tasks are limited to 72 hours.



## 8. Select the data *Haemophilus\_influenzae\_16S.fasta*

**Folders**

Total Storage: 708.59 KB

- Tutorial Example
  - Data (1)
  - Tasks (0)
- Tutorial 2

**Create new task**

Task Summary    [Select Data](#)    [Select Tool](#)    Set Parameters

You may edit your task using the tabs above.  
Current CPU Hr Usage: 3 [Explain this?](#)

Description

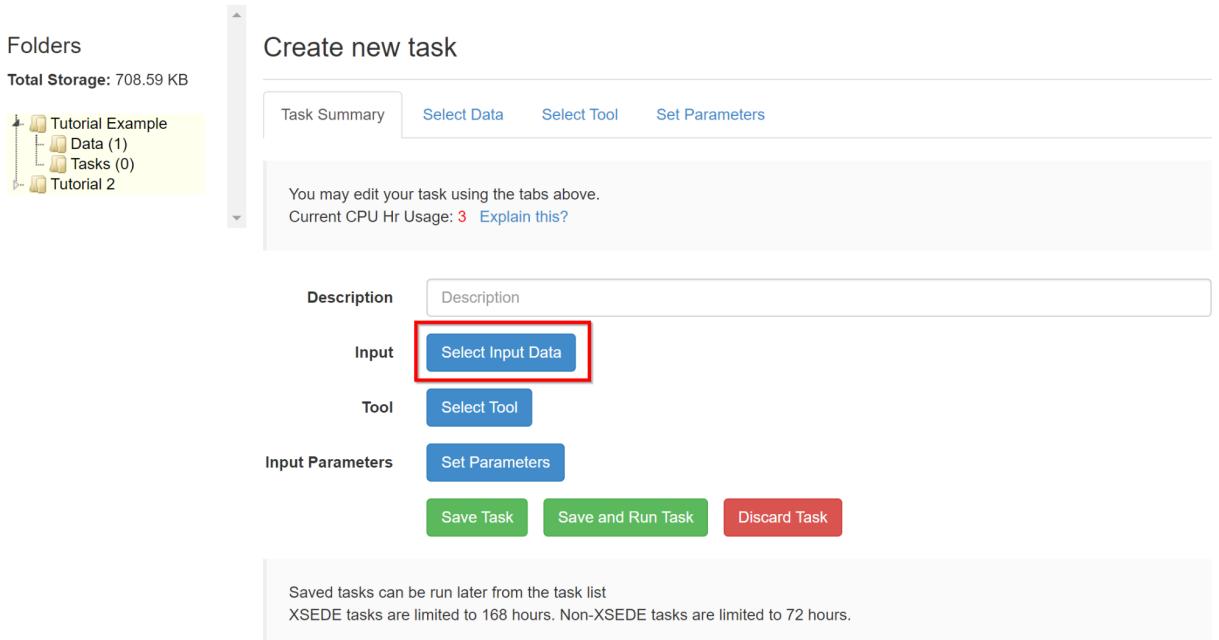
Input [Select Input Data](#)

Tool [Select Tool](#)

Input Parameters [Set Parameters](#)

[Save Task](#) [Save and Run Task](#) [Discard Task](#)

Saved tasks can be run later from the task list  
XSEDE tasks are limited to 168 hours. Non-XSEDE tasks are limited to 72 hours.



The screenshot shows a 'Create new task' interface. On the left, there's a sidebar titled 'Folders' with a total storage of 708.59 KB. It lists 'Tutorial Example' (containing 'Data (1)' and 'Tasks (0)'), 'Tutorial 2', and 'Tutorial 3'. The main area is titled 'Create new task' and has tabs for 'Task Summary', 'Select Data' (which is selected), 'Select Tool', and 'Set Parameters'. A message says 'You can choose the following data.' Below is a table:

Select One	Label	Size	Data Format	Date Created	Parse Status
<input checked="" type="radio"/>	Haemophilus_influenzae_16S.fasta	65.13 KB	Unknown	4/26/21, 19:10	<span style="color: green;">✓ Succeeded</span>

At the bottom are 'Select Data' and 'Cancel' buttons, with 'Select Data' being highlighted by a red box.

- This will take you to selecting a tool. Click on MAFFT to create a MAFFT alignment.

The screenshot shows the 'Create new task' interface with the 'Select Tool' tab selected. The sidebar shows 'Folders' with a total storage of 708.59 KB. The main area displays a list of tools:

- BAli-Phy on XSEDE (3.4.1-3.5.0) - BAli-Phy estimates multiple sequence alignments and evolutionary trees.
- BEAST2 on XSEDE (2.1 - 2.6.3) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- BEAST on XSEDE (1.10.4) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- BEAST (old) on XSEDE (1.8.0 - 1.8.4) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- BlastN (2.2.1) - Search DBs for Nucleotide Sequence similarity
- Clearcut (1.0.9) - Fast Implementation of Relaxed Neighbor Joining
- ClustalW (2.1) - Create Multiple Alignments from Sequences
- Consense (Phylip 3.66) - Find A Consensus Tree
- Cutadapt on XSEDE (2.10) - Remove adapter sequences from high-throughput sequencing reads
- DPPDIV on XSEDE (1.0) - Estimating species divergence times and lineage-specific substitution rates on a fixed topology run on XSEDE
- EPA-NG on XSEDE (0.3.7) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- ExaBayes on XSEDE (1.5.1) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- FastML on XSEDE (3.1) - Fast (Approximate) Maximum Likelihood tree construction - run on XSEDE
- FastTreeMP on XSEDE (2.1.10) - Fast (Approximate) Maximum Likelihood tree construction - run on XSEDE
- GARLI 2.01 on XSEDE (2.01) - Genetic Algorithm for Rapid Likelihood Inference - run on XSEDE.
- GARLI.conf Creator (2.0) - Creates a Garli.conf file for up to five partitions
- GBLOCKS on XSEDE (0.91b) - Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis - run on XSEDE
- G-PhoCS on XSEDE (1.3.2) - A Generalized Phylogenetic Coalescent Sampler
- Guidance2 on XSEDE (2.02) - Accurate detection of unreliable alignment regions - run on XSEDE.
- IMa3 on XSEDE (1.11) - IMa3p - Parallel MCMC and inference of ancient demography under the Isolation with Migration (IM) model
- IQ-Tree on XSEDE (2.0.5) - Efficient phylogenomic software by maximum likelihood, run on XSEDE
- jModelTest2 on XSEDE (2.1.6) - Statistical selection of best-fit models of nucleotide substitution, run on XSEDE
- LogCombiner on XSEDE (1.8.4-2.6.0) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- MAFFT on XSEDE (7.471)** - Multiple alignment program for amino acid or nucleotide sequences; parallel version
- Migrate-N on XSEDE (4.4.4) - Estimation of Population Sizes and Gene Flow using the Coalescent
- ModelTest-NIC on XSEDE (0.1.5) - Statistical selection of best-fit models of nucleotide and protein substitution, run on XSEDE

The screenshot shows the 'Create new task' interface. On the left, there's a sidebar with 'Folders' and 'Total Storage: 708.59 KB'. Under 'Folders', there are two entries: 'Tutorial Example' (with 'Data (1)' and 'Tasks (0)') and 'Tutorial 2'. The main area has tabs: 'Task Summary' (selected), 'Select Data', 'Select Tool', and 'Set Parameters'. Below the tabs, a note says 'You may edit your task using the tabs above.' and 'Current CPU Hr Usage: 3 Explain this?'. The 'Description' field contains 'MAFFT'. Under 'Input', it says '1 Inputs Set'. Under 'Tool', it says 'MAFFT on XSEDE' with a link to 'Click for more info'. Under 'Input Parameters', it says '46 Parameters Set'. At the bottom are three buttons: 'Save Task' (green), 'Save and Run Task' (green), and 'Discard Task' (red).

10. Choose input parameters for MAFFT. The default parameters are good, so hit save. A browser alert may popup warning about CPU usage. Click ok.

The screenshot shows the 'Create new task' interface for the MAFFT tool. The 'Set Parameters' tab is selected. The description is 'MAFFT on XSEDE: Multiple alignment program for amino acid or nucleotide sequences; parallel version (Kazutaka Katoh, Kei-ichi Kuma, Hiroyuki Toh, and Takashi Miyata)'. Under 'Simple Parameters', there are fields for 'Maximum Hours to Run' (set to 0.25) and 'Which version of MAFFT are you using?' (set to 'OV 7.471'). There are several checkboxes: 'I need 60 GB of memory, a previous run ran out of memory' (unchecked), 'There are unusual characters in the dataset (~any symbol)' (unchecked), 'Preserve case (~preservecase)' (unchecked), and 'Output format' (radio buttons for 'FASTA' and 'ClustalW', with 'FASTA' selected). At the bottom, there are buttons for 'Advanced Parameters', 'Save Parameters' (which is highlighted with a red box), 'Reset', and 'Cancel'.

11. Save and run your task. A browser alert may popup a warning about time to execute. Click ok.

**Create new task**

Total Storage: 708.59 KB

Folders

- Tutorial Example
  - Data (1)
  - Tasks (0)
- Tutorial 2

Task Summary    Select Data    Select Tool    Set Parameters

You may edit your task using the tabs above.  
Current CPU Hr Usage: 3 [Explain this?](#)

Description	MAFFT
Input	1 Inputs Set
Tool	MAFFT on XSEDE <a href="#">Click for more info</a>
Input Parameters	23 Parameters Set
<a href="#">Save Task</a> <a href="#">Save and Run Task</a> <a href="#">Discard Task</a>	

Saved tasks can be run later from the task list  
XSEDE tasks are limited to 168 hours. Non-XSEDE tasks are limited to 72 hours.

12. You will be directed to your task list. Refresh the task list in about 30 seconds to see if the MAFFT task is done executing.

**Tasks**

Total Storage: 773.71 KB

Folders

- Tutorial Example
  - Data (1)
  - Tasks (1)
- Tutorial 2

[Refresh Tasks](#)

Current CPU Hrs Usage: 6 [Explain this?](#)  
CPU Hrs Remaining: 49994  
CPU Hrs Expire: 06/30/2021 00:00:00  
Running Tasks: 1  
There is currently 1 data item in this tab.

[Create New Task](#)

Show 20 records on each page

Page 1 of 1

Select All	Label	Tool	Input	Parameters	Date Created	Action
<input type="checkbox"/>	<a href="#">Clone</a>	MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a>	View (22) 4/26/21, 19:12	<a href="#">View Status</a>

Move selected to [Tutorial 2](#) [GO](#)

[Kill Selected](#) [Delete Selected](#)

If it is completed, view the output files.

The screenshot shows the CIPRES Tasks interface. On the left, there's a sidebar titled "Folders" with "Total Storage: 773.71 KB" and a tree view of "Tutorial Example" containing "Data (1)" and "Tasks (1)". The main area is titled "Tasks" with a sub-section "Current CPU Hrs Usage: 5 Explain this?". It displays "CPU Hrs Remaining: 49995", "CPU Hrs Expire: 06/30/2021 00:00:00", and "Running Tasks: 0". Below this, it says "There is currently 1 data item in this tab." A blue button "Create New Task" is at the bottom. Under the table, there are buttons "Move selected to Tutorial 2 GO", "Kill Selected", and "Delete Selected".

<input type="checkbox"/> Select All	Label	Tool	Input	Parameters	Date Created	Action	
<input type="checkbox"/>	Clone	MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a>	<a href="#">View (22)</a>	4/26/21, 19:12	<a href="#">View Output</a>

You can then view or download any of the files, but *output.mafft* is needed for building trees. Select and view *output.mafft*.

The screenshot shows the "Output for MAFFT" interface. On the left, there's a sidebar titled "Folders" with "Total Storage: 998.04 KB" and a tree view of "Tutorial Example" containing "Data (1)" and "Tasks (1)". The main area has a message "Click on an output file below to review its contents." Below is a table of output files:

<input type="checkbox"/> Select all	Tool Output	File Name	File Size	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>	PROCESS_OUTPUT	STDOUT	0.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		STDERR	227.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>	all_outputfiles	stderr.txt	227.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		scheduler.conf	69.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		input.fasta	65.13 KB	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		stdout.txt	0.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		start.txt	40.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		done.txt	49.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		output.mafft	73.67 KB	<a href="#">View</a>	<a href="#">Download</a>
<input checked="" type="checkbox"/>	outputfile	output.mafft	73.67 KB	<a href="#">View</a>	<a href="#">Download</a>

At the bottom, there are buttons: "Download Selected", "Transfer Selected to CIPRES Share", "Return to Task List", and "View Current Task".

### 13. Save *output.mafft*

**View Task Output Details**

**Folders**  
Total Storage: 998.04 KB

- Tutorial Example
  - Data (1)
  - Tasks (1)
- Tutorial 2

**Save To Current Folder** (button highlighted with red box) | **Return** | **Download File**

Tool	MAFFT on XSEDE
File Name	output.mafft
File Size	75438 Bytes

**Show/Hide Output Contents**

```
>Haemophilus_influenzae_86_028NP_uid58093_1
-----aattgaag-----agt
ttgatcatggctcagat----tgaacgctggccggcaggcttaacacatgcgaagtgcgaac
ggtagcaggaggaaacttgtttttctgtcagactggccggacgg-gtgagaatgtcttgg
gaaatctggcttatggatgggataaacgcggaaactgtcgtaataccg---cgat
ttatcgaa---agatgaaatgtcgggactgagagccgcattccataggatgagccaa
gtgggg-----att
agttatgttggggtaatgccttaccaaggctgc-----gatctctagctggctgaa
gaggatgaccacccacactggaaactgagacacggtcagact-cttacggggcggcag
tggggatattgcgaatgggggacccctgacgcagccatggcgatgaaagg
```

**Save Output**

**Folders**  
Total Storage: 998.04 KB

- Tutorial Example
  - Data (1)
  - Tasks (1)
- Tutorial 2

Tool	MAFFT on XSEDE
File Name	output.mafft
File Size	75438 Bytes
Label	output.mafft
Data Format:	<input type="button" value="▼"/>

**Save** | **Cancel**

**Show/Hide Data Contents**

```
>Haemophilus_influenzae_86_028NP_uid58093_1
-----aattgaag-----agt
ttgatcatggctcagat----tgaacgctggccggcaggcttaacacatgcgaagtgcgaac
ggtagcaggaggaaacttgtttttctgtcagactggccggacgg-gtgagaatgtcttgg
gaaatctggcttatggatgggataaacgcggaaactgtcgtaataccg---cgat
ttatcgaa---agatgaaatgtcgggactgagagccgcattccataggatgagccaa
gtgggg-----att
```

It will then be put into your Data folder. You can see it now holds 2 files.

**Folders**  
Total Storage: 1.05 MB

- Tutorial Example
  - Data (2)
  - Tasks (1)
- Tutorial 2

Task output "output.mafft" successfully saved.

**Output for MAFFT**

## Procedure for RaxML

14. Go to the tasks subfolder. Create a new task.

**Folders**  
Total Storage: 1.05 MB

- Tutorial Example
  - Data (2)
  - Tasks (1)
- Tutorial 2

**Tasks**

Current CPU Hrs Usage: 5 Explain this?  
 CPU Hrs Remaining: 49995  
 CPU Hrs Expire: 06/30/2021 00:00:00  
 Running Tasks: 0  
 There is currently 1 data item in this tab.

[Create New Task](#)

Show 20 records on each page

Select All	Label	Tool	Input	Parameters	Date Created	Action
<input type="checkbox"/>	Clone	MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a> <a href="#">View (22)</a>	4/26/21, 19:12	<a href="#">View Output</a>

[Move selected to Tutorial 2](#) [GO](#)

[Kill Selected](#) [Delete Selected](#)

## 15. Select input data and choose *output.mafft*

**Folders**  
Total Storage: 1.05 MB

- Tutorial Example
  - Data (2)
  - Tasks (1)
- Tutorial 2

**Create new task**

Task Summary    [Select Data](#)    [Select Tool](#)    [Set Parameters](#)

You may edit your task using the tabs above.  
 Current CPU Hr Usage: 5 Explain this?

Description

Input [Select Input Data](#) (button highlighted with red box)

Tool [Select Tool](#)

Input Parameters [Set Parameters](#)

[Save Task](#) [Save and Run Task](#) [Discard Task](#)

16. The tool selection screen will show next. Scroll down and choose RAxML-HPC v.8 on XSEDE

- [Phylobayes MPI on XSEDE](#) (1.7b) ⓘ - Phylogenetic reconstruction using infinite mixtures - run on XSEDE
- [Phylobayes MPI on XSEDE \(Restart\)](#) (1.7b) ⓘ - Restart a PhyloBayes MPI run on XSEDE
- [POY](#) (4.1.2) ⓘ - Phylogenetic tree inference using dynamic homologies.
- [Probalign](#) (1.3) ⓘ - Multiple sequence alignment using partition function posterior probabilities.
- [ProbCons](#) (1.12) ⓘ - Probabilistic Consistency-based Multiple Alignment of Amino/Nucleic Acid Sequences
- [RAxML-HPC BlackBox](#) (8.2.12) ⓘ - Phylogenetic tree inference using maximum likelihood/rapid bootstrapping on XSEDE.
- [RAxML-HPC2 on XSEDE](#) (8.2.12) ⓘ - Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE
- [RAxML-HPC2 Workflow on XSEDE](#) (8.2.12) ⓘ - Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE
- [RAxML-HPC v.8 on XSEDE](#) (8.2.12) ⓘ **Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE**
- [RAxML-NG](#) (1.0.1) ⓘ - RAxML-NG offers improvements in speed, flexibility and user-friendliness over previous RAxML versions.
- [Readseq](#) (2.2) ⓘ - Readseq: biological sequence format interconversion utility
- [Seqboot](#) (Phylib 3.66) ⓘ - Bootstrap, Jackknife, or Permutation Resampling
- [TreeAnnotator on XSEDE](#) (1.8.4-2.6.0) ⓘ - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- [TrimAI on XSEDE](#) (1.2.59) ⓘ - Trim alignments using TrimAI - run on XSEDE

17. Select your input parameters. The default parameters are good, so you can just hit Save Parameters. A browser alert may popup, hit ok.

**Folders**  
Total Storage: 1.05 MB

- Tutorial Example
- Data (2)
- Tasks (1)
- Tutorial 2

**Create new task**

Tool "RAXMLHPC8\_XSEDE" successfully set to current task.

Task Summary    Select Data    Select Tool    Set Parameters

You may edit your task using the tabs above.  
Current CPU Hr Usage: 5 [Explain this?](#)

Description	Description
Input	1 Inputs Set
Tool	RAXML-HPC v.8 on XSEDE <a href="#">Click for more info</a>
Input Parameters	36 Parameters Set

[Save Task](#) [Save and Run Task](#) [Discard Task](#)

Saved tasks can be run later from the task list  
XSEDE tasks are limited to 168 hours. Non-XSEDE tasks are limited to 72 hours.

**Create new task**

Task Summary    Select Data    Select Tool    Set Parameters

**RAxML-HPC v.8 on XSEDE: Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE ([Alexandros Stamatakis](#))**

**Simple Parameters**

Maximum Hours to Run (click here for help setting this correctly) \*

How many patterns in your data set?

Set a name for output files (-n)

Enable ML searches under CAT (-F)

Outgroup (-o) (one or more comma-separated outgroups, see comment for syntax)

Specify the number of distinct rate categories (-c) \*

Disable Rate Heterogeneity (-V) \*

Supply a tree (Not available when doing rapid bootstrapping, -x) (-t)

Specify a random seed value for parsimony inferences (-p) \*

Enter a random seed value for parsimony inferences (-p "value" gives reproducible results from random starting tree) \*

Specify an initial rearrangement setting (-i) \*

Specify the distance from original pruning point (-i) \*

Constraint (-g)

Binary Backbone (-r)

Use a mixed/partitioned model? (-q)

My data set has more than 99 partitions

Estimate individual per-partition branch lengths (-M) \*

Specify an ML estimate of base frequencies (GTRGAMMA + X)  no  yes

Estimate proportion of invariable sites (GTRGAMMA + I)  no  yes

Choose an input file that excludes the range of positions specified in this file (-E)

Weight characters as specified in this file (-a)

Disable checking for sequences with no values (-O)

Print output files that can be parsed by Mesquite. (-mesquite)

**Advanced Parameters**

[Save Parameters](#) [Reset](#) [Cancel](#)

18. Enter a description name. Save and run the task. A browser alert may popup, hit ok.

The screenshot shows the 'Create new task' interface. On the left, there's a sidebar titled 'Folders' with 'Total Storage: 1.05 MB' and a tree view showing 'Tutorial Example' with 'Data (2)' and 'Tasks (1)', and 'Tutorial 2'. The main area has tabs: 'Task Summary' (selected), 'Select Data', 'Select Tool', and 'Set Parameters'. Below the tabs, a message says 'You may edit your task using the tabs above.' and 'Current CPU Hr Usage: 5' with a link 'Explain this?'. The 'Task Summary' section contains fields: 'Description' (RaxML), 'Input' (1 Inputs Set), 'Tool' (RAxML-HPC v.8 on XSEDE with a link 'Click for more info'), and 'Input Parameters' (19 Parameters Set). At the bottom are three buttons: 'Save Task' (green), 'Save and Run Task' (highlighted with a red box), and 'Discard Task' (red).

19. Your task is running and you will be directed to the task list. Hit Refresh Tasks when you think the RaxML is done executing, about 2 minutes.

The screenshot shows the 'Tasks' interface. On the left, there's a sidebar titled 'Folders' with 'Total Storage: 1.12 MB' and a tree view showing 'Tutorial Example' with 'Data (2)' and 'Tasks (2)', and 'Tutorial 2'. The main area has a title 'Tasks' and a button 'Refresh Tasks' (highlighted with a red box). Below it, status information is displayed: 'Current CPU Hrs Usage: 6' (with a link 'Explain this?'), 'CPU Hrs Remaining: 49994', 'CPU Hrs Expire: 06/30/2021 00:00:00', and 'Running Tasks: 0'. A note says 'There are currently 2 tasks in this tab. (Items 1 - 2 are shown here.)'. At the bottom is a 'Create New Task' button.

20. View the output of the task in the file called *RAxML\_bestTree.result*

**Folders**  
Total Storage: 1.12 MB

- Tutorial Example
  - Data (2)
  - Tasks (2)
- Tutorial 2

**Tasks**

Current CPU Hrs Usage: 6 Explain this?  
 CPU Hrs Remaining: 49994  
 CPU Hrs Expire: 06/30/2021 00:00:00  
 Running Tasks: 0  
 There are currently 2 tasks in this tab. (Items 1 - 2 are shown here.)

Create New Task

Show 20 records on each page

Select All	Label	Tool	Input	Parameters	Date Created	Action	
<input type="checkbox"/>	Clone	RaxML	RAxML-HPC v.8 on XSEDE	<a href="#">View (1)</a>	<a href="#">View (18)</a>	4/26/21, 19:36	<a href="#">View Output</a>
<input type="checkbox"/>	Clone	MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a>	<a href="#">View (22)</a>	4/26/21, 19:12	<a href="#">View Output</a>

Move selected to [Tutorial 2](#) GO

**Folders**  
Total Storage: 1.12 MB

- Tutorial Example
  - Data (2)
  - Tasks (2)
- Tutorial 2

**Output for RaxML**

Click on an output file below to review its contents.

Select all	Tool Output	File Name	File Size	Action
<input type="checkbox"/>	PROCESS_OUTPUT	STDOUT	5.90 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		STDERR	0.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>	all_outputfiles	RAxML_log.result	176.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		stderr.txt	0.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		RAxML_result.result	3.59 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		infile.txt.reduced	36.24 KB	<a href="#">View</a> <a href="#">Download</a>
<input checked="" type="checkbox"/>		RAxML_bestTree.result	3.59 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		scheduler.conf	85.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		stdout.txt	5.90 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		start.txt	40.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		RAxML_parsimonyTree.result	1.77 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		done.txt	49.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		RAxML_info.result	5.48 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		_scheduler_stderr.txt	938.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		term.txt	316.00 bytes	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		infile.txt	73.67 KB	<a href="#">View</a> <a href="#">Download</a>
<input type="checkbox"/>		_JOBINFO.TXT	338.00 bytes	<a href="#">View</a> <a href="#">Download</a>

[Download Selected](#) [Transfer Selected to CIPRES Share](#)

## 21. Download the file *RAxML\_bestTree.result*

Folders

Total Storage: 1.25 MB

- ↑ Tutorial Example
  - ↓ Data (2)
  - ↓ Tasks (2)
- ↓ Tutorial 2

View Task Output Details

Save To Current Folder    Return

**Download File**

Tool: RAxML-HPC v.8 on XSEDE

File Name: RAxML\_bestTree.result

File Size: 3676 Bytes

Show/Hide Output Contents

```
((Haemophilus_influenzae_86_028NP_uid58093_6:0.0000010000050002909,Haemophilus_influenzae_86_028NP_uid58093_3:0.1
```

## 22. Go to iTOL. Upload *RAxML\_bestTree.result*

**Upload a new tree**

Tree name:  
Tutorial 2

Paste your tree into the box below, or select a file using the **Tree file** selector. You can also simply drag and drop the tree file onto the page (only a regular plain text file, not QIIME QZA files).

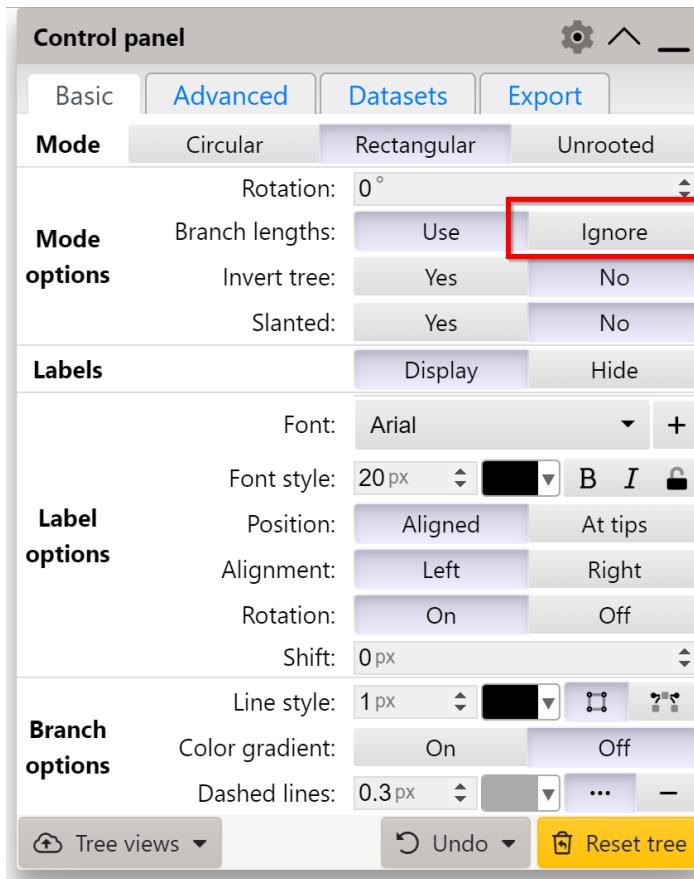
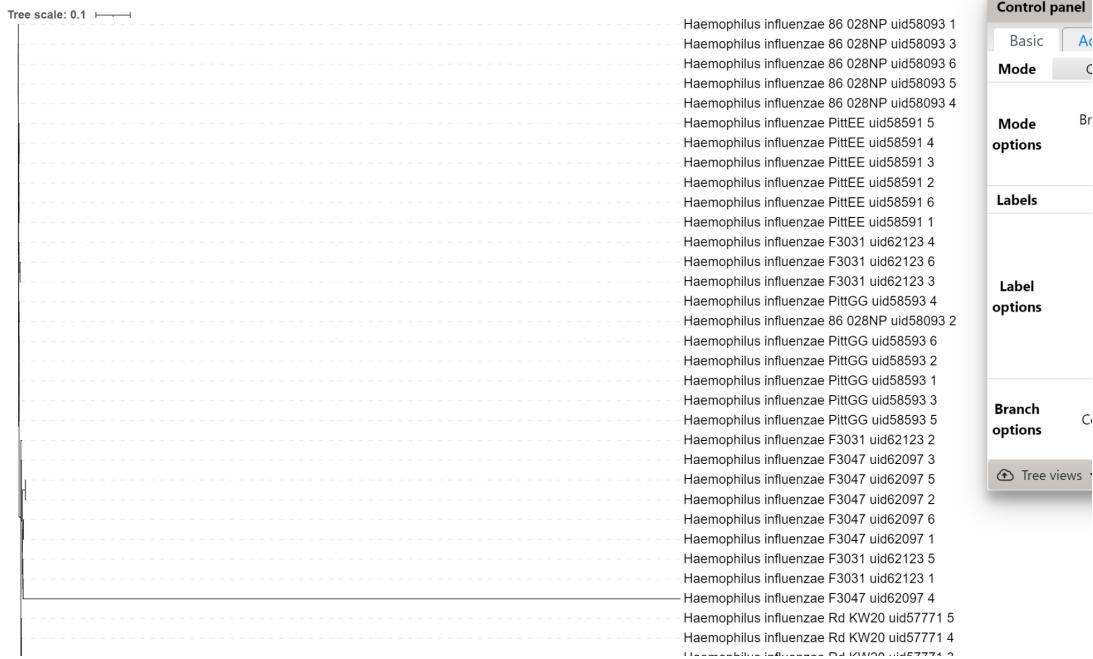
Tree text:

Tree file:

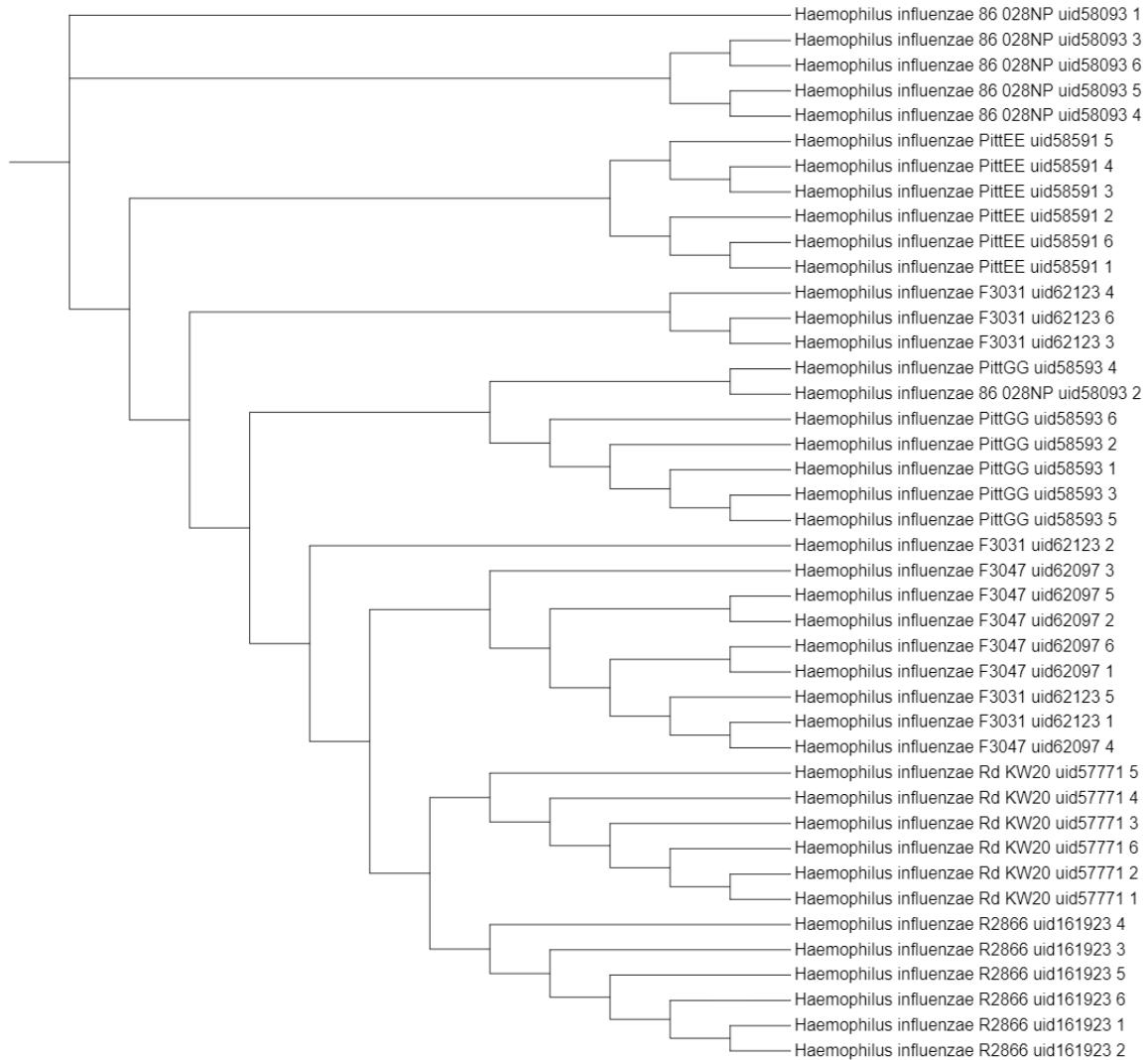
RAxML\_bestTree.result

## 23. It will look a little funky, because one branch is so much longer than the rest.

Choose to ignore branch lengths in the control panel.



You can then view your final tree created using MAFFT and RaxML using CIPRES.



## Procedure for FastTree

24. Create a new task

**Folders**  
Total Storage: 1.26 MB

**Tasks**

Current CPU Hrs Usage: 6 Explain this?  
CPU Hrs Remaining: 49994  
CPU Hrs Expire: 06/30/2021 00:00:00  
Running Tasks: 0  
There are currently 2 tasks in this tab. (Items 1 - 2 are shown here.)

**Create New Task**

Show 20 records on each page

	Select All	Label	Tool	Input	Parameters	Date Created	Action
<input type="checkbox"/>	<input type="checkbox"/> Clone	RaxML	RAXML-HPC v.8 on XSEDE	<a href="#">View (1)</a>	<a href="#">View (18)</a>	4/26/21, 19:36	<a href="#">View Output</a>
<input type="checkbox"/>	<input type="checkbox"/> Clone	MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a>	<a href="#">View (22)</a>	4/26/21, 19:12	<a href="#">View Output</a>

Move selected to  [GO](#)

[Kill Selected](#) [Delete Selected](#)

## 25. Select input data, choose *output.mafft* and hit Select Data

**Folders**  
Total Storage: 1.26 MB

**Create new task**

Task Summary    [Select Data](#)    [Select Tool](#)    [Set Parameters](#)

You may edit your task using the tabs above.  
Current CPU Hr Usage: 6 Explain this?

Description	Description
Input	<b>Select Input Data</b>
Tool	<a href="#">Select Tool</a>
Input Parameters	<a href="#">Set Parameters</a>

[Save Task](#) [Save and Run Task](#) [Discard Task](#)

Saved tasks can be run later from the task list  
XSEDE tasks are limited to 168 hours. Non-XSEDE tasks are limited to 72 hours.

The screenshot shows the 'Create new task' interface. On the left, a sidebar displays 'Folders' with 'Total Storage: 1.26 MB' and a tree view of 'Tutorial Example' containing 'Data (3)' and 'Tasks (2)'. Below this is 'Tutorial 2'. The main area has tabs: 'Task Summary' (selected), 'Select Data' (highlighted in blue), 'Select Tool', and 'Set Parameters'. A message says 'You can choose the following data.' Below is a table:

Select One	Label	Size	Data Format	Date Created	Parse Status
<input type="radio"/>	RAxML_bestTree.result	3.59 KB	Unknown	4/26/21, 19:55	<span>?</span> Unknown
<input checked="" type="radio"/>	output.mafft	73.67 KB	Unknown	4/26/21, 19:33	<span>?</span> Unknown
<input type="radio"/>	Haemophilus_influenzae_16S.fasta	65.13 KB	Unknown	4/26/21, 19:10	<span>✓</span> Succeeded

Buttons at the bottom are 'Select Data' (green) and 'Cancel' (blue).

## 26. It will redirect you to selecting a tool Choose FastTreeMP on XSEDE

The screenshot shows the 'Create new task' interface with the 'Select Tool' tab selected. A green message bar at the top says '1 data item successfully set as input to current task.' Below is a list of tools:

- BALI-Phy on XSEDE (3.4.1-3.5.0) - BALI-Phy estimates multiple sequence alignments and evolutionary trees.
- BEAST2 on XSEDE (2.1 - 2.6.3) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- BEAST on XSEDE (1.10.4) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- BEAST (old) on XSEDE (1.8.0 - 1.8.4) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- BlastN (2.2.1) - Search DBs for Nucleotide Sequence similarity
- Clearcut (1.0.9) - Fast Implementation of Relaxed Neighbor Joining
- ClustalW (2.1) - Create Multiple Alignments from Sequences
- Consense (Phylip 3.66) - Find A Consensus Tree
- Cutadapt on XSEDE (2.10) - Remove adapter sequences from high-throughput sequencing reads
- DPPDIV on XSEDE (1.0) - Estimating species divergence times and lineage-specific substitution rates on a fixed topology run on XSEDE
- EPA-NG on XSEDE (0.3.7) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- ExaBayes on XSEDE (1.5.1) - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE
- FastML on XSEDE (3.1) - Fast (Approximate) Maximum Likelihood tree construction - run on XSEDE
- FastTreeMP on XSEDE (2.1.10)** - Fast (Approximate) Maximum Likelihood tree construction - run on XSEDE
- GARLI 2.01 on XSEDE (2.01) - Genetic Algorithm for Rapid Likelihood Inference - run on XSEDE.

## 27. Choose your input parameters. The default parameters are good, so hit Save Parameters. A browser alert may popup, and select ok.

**Folders**  
Total Storage: 1.26 MB

- Tutorial Example
  - Data (3)
  - Tasks (2)
- Tutorial 2

Tool "FASTTREE\_XSEDE" successfully set to current task.

### Create new task

You may edit your task using the tabs above.  
Current CPU Hr Usage: 6 Explain this?

Description	Description
Input	1 Inputs Set
Tool	FastTreeMP on XSEDE Click for more info
Input Parameters	20 Parameters Set

**Save Task** **Save and Run Task** **Discard Task**

Saved tasks can be run later from the task list

---

**Folders**  
Total Storage: 1.26 MB

- Tutorial Example
  - Data (3)
  - Tasks (2)
- Tutorial 2

### Create new task

Task Summary Select Data Select Tool Set Parameters

FastTreeMP on XSEDE: Fast (Approximate) Maximum Likelihood tree construction - run on XSEDE ([M.N. Price, P.S. Dehal, A.P. Arkin](#))

Simple Parameters

Maximum Hours to Run (up to 168 hours) \* 0.5  
 Please Specify your data type \*  Nucleotide  Amino acid  
 I need more memory \*   
 Starting Tree in Newick Format (-intree)

Advanced Parameters

**Save Parameters** **Reset** **Cancel**

28. Enter a description to name your task. Hit Save and Run to start executing it. A browser alert may popup, hit ok.

Folders  
Total Storage: 1.26 MB

Tutorial Example  
Data (3)  
Tasks (2)  
Tutorial 2

Create new task

Task Summary    Select Data    Select Tool    Set Parameters

You may edit your task using the tabs above.  
Current CPU Hr Usage: 6 [Explain this?](#)

Description	FastTree
Input	1 Inputs Set
Tool	FastTreeMP on XSEDE <a href="#">Click for more info</a>
Input Parameters	28 Parameters Set
<a href="#">Save Task</a> <a href="#">Save and Run Task</a> <a href="#">Discard Task</a>	

29. You will be redirected to your task list. Hit Refresh Tasks when you think the program is done executing, about 1 minute.

Folders  
Total Storage: 1.33 MB

Tutorial Example  
Data (3)  
Tasks (3)  
Tutorial 2

Tasks

Current CPU Hrs Usage: 9 [Explain this?](#)  
CPU Hrs Remaining: 49991  
CPU Hrs Expire: 06/30/2021 00:00:00  
Running Tasks: 1  
There are currently 3 tasks in this tab. (Items 1 - 3 are shown here.)

[Create New Task](#)

Show 20 records on each page

« ‹ Page 1 of 1 › »

<input type="checkbox"/> Select All	Label	Tool	Input	Parameters	Date Created	Action
<input type="checkbox"/>	<a href="#">Clone</a> FastTree	FastTreeMP on XSEDE	<a href="#">View (1)</a>	<a href="#">View (27)</a>	4/26/21, 19:57	<a href="#">View Status</a>
<input type="checkbox"/>	<a href="#">Clone</a> RaxML	RAxML-HPC v.8 on XSEDE	<a href="#">View (1)</a>	<a href="#">View (18)</a>	4/26/21, 19:36	<a href="#">View Output</a>
<input type="checkbox"/>	<a href="#">Clone</a> MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a>	<a href="#">View (22)</a>	4/26/21, 19:12	<a href="#">View Output</a>

Move selected to [Tutorial 2](#) [GO](#)

### 30. View the output files.

Create New Task

Show 20 records on each page

« < Page 1 of 1 > »

<input type="checkbox"/>	Select All	Label	Tool	Input	Parameters	Date Created	Action
<input type="checkbox"/>	Clone	FastTree	FastTreeMP on XSEDE	<a href="#">View (1)</a>	<a href="#">View (27)</a>	4/26/21, 19:57	<a href="#">View Output</a>
<input type="checkbox"/>	Clone	RaxML	RAxML-HPC v.8 on XSEDE	<a href="#">View (1)</a>	<a href="#">View (18)</a>	4/26/21, 19:36	<a href="#">View Output</a>
<input type="checkbox"/>	Clone	MAFFT	MAFFT on XSEDE	<a href="#">View (1)</a>	<a href="#">View (22)</a>	4/26/21, 19:12	<a href="#">View Output</a>

Move selected to [Tutorial 2](#) [GO](#)

[Kill Selected](#) [Delete Selected](#)

### 31. View the *fasttree\_result.tree* file

Folders

Total Storage: 1.44 MB

- ↳ Tutorial Example
  - ↳ Data (3)
  - ↳ Tasks (3)
- ↳ Tutorial 2

Output for FastTree

Click on an output file below to review its contents.

<input type="checkbox"/>	Select all	Tool Output	File Name	File Size	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		PROCESS_OUTPUT	STDOUT	166.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			STDERR	1.14 KB	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>		all_results	logfile.txt	35.77 KB	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			stderr.txt	1.14 KB	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			scheduler.conf	76.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			stdout.txt	166.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			start.txt	40.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input checked="" type="checkbox"/>			fasttree_result.tree	2.40 KB	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			done.txt	49.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			_scheduler_stderr.txt	869.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			term.txt	316.00 bytes	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			infile	73.67 KB	<a href="#">View</a>	<a href="#">Download</a>
<input type="checkbox"/>			_JOBINFO.TXT	335.00 bytes	<a href="#">View</a>	<a href="#">Download</a>

[Download Selected](#) [Transfer Selected to CIPRES Share](#)

[Return to Task List](#) [View Current Task](#)

32. Download the *fasttree\_result.tree* file.

The screenshot shows a user interface for viewing task output details. On the left, there's a sidebar titled 'Folders' showing a tree structure with 'Tutorial Example' at the root, containing 'Data (3)' and 'Tasks (3)'. Below that is 'Tutorial 2'. The main area is titled 'View Task Output Details' and contains the following information:

- Save To Current Folder** and **Return** buttons
- Download File** button (highlighted with a red box)
- Tool**: FastTreeMP on XSEDE
- File Name**: fasttree\_result.tree
- File Size**: 2457 Bytes

At the bottom, there's a 'Show/Hide Output Contents' section with a scrollable text area containing phylogenetic tree data in Newick format.

33. Go to iTOL. Upload the file called *fasttree\_result.tree*.

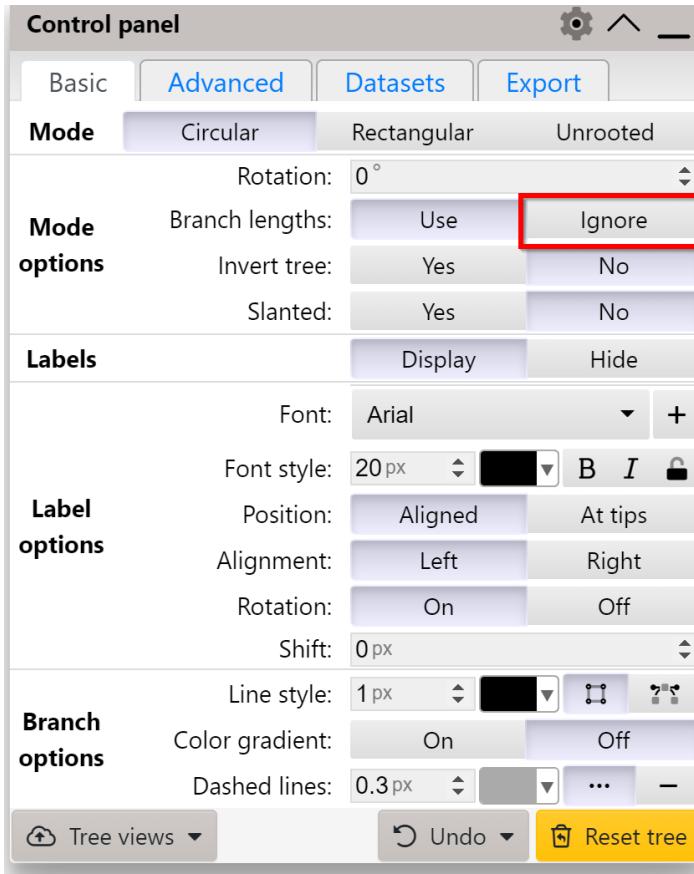
The screenshot shows the 'Upload a new tree' form on the iTOL website. It includes fields for:

- Tree name:** A text input field containing 'FastTree'.
- Tree text:** An empty text area for pasting tree text.
- Tree file:** A file input field showing 'Choose File fasttree\_result.tree'.

A large grey 'Upload' button is at the bottom.

34. It will look a little funky, because one branch is so much longer than the rest.

Choose to ignore branch lengths in the control panel.



You can then view your final tree created using MAFFT and FastTree using CIPRES.

