

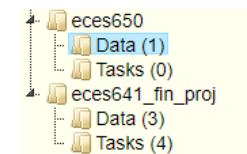
# CIPRES Walkthrough

---

# Uploading Our Data

## Folders

Total Storage: 3.56 MB



## Upload File

Upload your files

Haemophil...e\_16S.fasta

You can select multiple files.

MSIE 9 and below support single uploads only.

You can also enter your data manually below

Label (required)

Data:

# Viewing Our Uploaded Data

## Folders

Total Storage: 3.56 MB

- ecses650
  - Data (1)
  - Tasks (0)
- ecses641\_fin\_proj
  - Data (3)
  - Tasks (4)

## All Data

There is currently 1 data item in this folder.

Upload Data

« < Page 1 of 1 > »

20 records on each page

## User Data

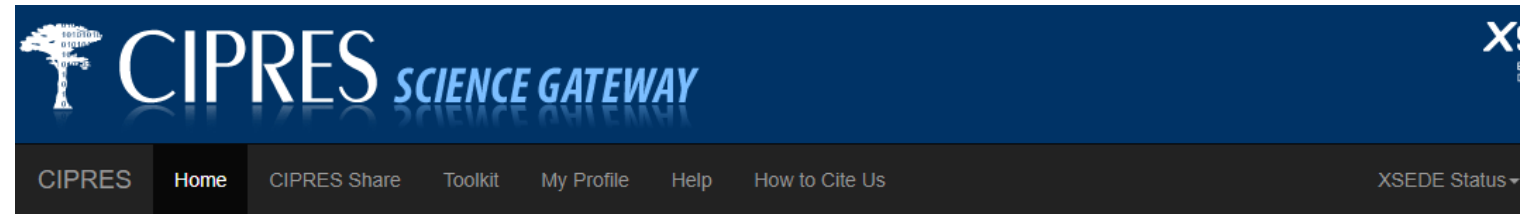
<input type="checkbox"/> Select all	User Data ID	Label	Size	Data Format	Date Created	Parse Status
<input type="checkbox"/>	2774197	Haemophilus_influenzae_16S.fasta	65.13 KB	Unknown	4/21/21, 16:27	✓ Succeeded

Move selected to eces641\_fin\_proj GO

Delete Selected

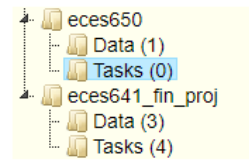
Transfer Selected to CIPRES Share

# Creating a New Task



## Folders

Total Storage: 3.56 MB



There are currently no tasks in this folder.

Create New Task

Continue Editing Current Task

You can: [Report an issue](#) [Request a feature](#) [Subscribe to Portal News](#)

SDSC - UC San Diego, MC 0505 - 9500 Gilman Drive - La Jolla, CA 92093-0505

**SDSC**  
SAN DIEGO SUPERCOMPUTER CENTER

**UC San Diego**

# New Task: Task Summary

CIPRES

Home

CIPRES Share

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status ▾

Logo

Folders

Total Storage: 3.56 MB

ecses650

Data (1)

Tasks (0)

ecses641\_fin\_proj

Data (3)

Tasks (4)

Create new task

Task Summary

Select Data

Select Tool

Set Parameters

You may edit your task using the tabs above.

Current CPU Hr Usage: 9 [Explain this?](#)

Description

mafft

Input

1 Inputs Set

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.

# New Task: Select Data

CIPRES

Home

CIPRES Share

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status

Logout

Folders

Total Storage: 3.56 MB

eces650

Data (1)

Tasks (0)

eces641\_fin\_proj

Data (3)

Tasks (4)

Create new task

Task Summary

Select Data

Select Tool

Set Parameters

You have selected the following data items to be set as input to the current task:

Label	Size	Data Format	Date Created	Parse Status
Haemophilus_influenzae_16S.fasta	65.13 KB	Unknown	4/21/21, 16:27	<input checked="" type="checkbox"/> Succeeded

If you would like to select new input, you may do so below:

You can choose the following data.

Select One	Label	Size	Data Format	Date Created	Parse Status
<input checked="" type="radio"/>	Haemophilus_influenzae_16S.fasta	65.13 KB	Unknown	4/21/21, 16:27	<input checked="" type="checkbox"/> Succeeded

Select Data

Cancel

# New Task: Select Tool

[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-NG on XSEDE \(0.1.5\)](#) ⓘ - Statistical selection of best-fit models of nucleotide and protein substitution, run on XSEDE

[MrBayes Restart on XSEDE \(3.2.x\)](#) ⓘ - Tree Inference Using Bayesian Analysis - run on XSEDE

[MrBayes on XSEDE \(3.2.7a\)](#) ⓘ - Tree Inference Using Bayesian Analysis - run on XSEDE

[Muscle \(3.7\)](#) ⓘ - Create Multiple Alignments from Sequences or Profiles

[NCLconverter \(2.1\)](#) ⓘ - A file format transformation tool

[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-NG on XSEDE \(0.1.5\)](#) ⓘ - Statistical selection of best-fit models of nucleotide and protein substitution, run on XSEDE

[MrBayes Restart on XSEDE \(3.2.x\)](#) ⓘ - Tree Inference Using Bayesian Analysis - run on XSEDE

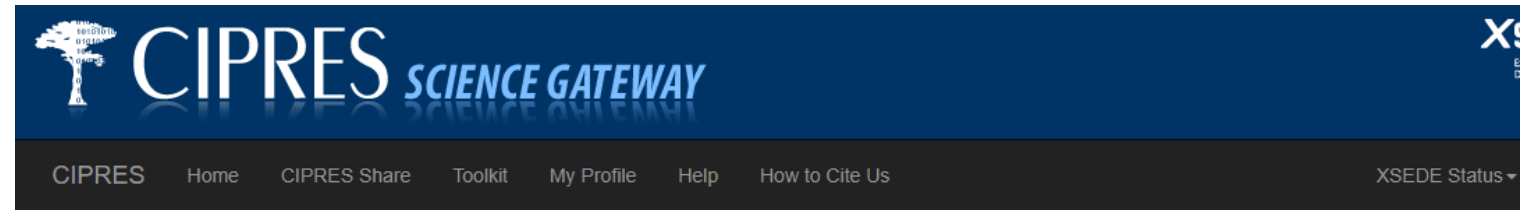
[MrBayes on XSEDE \(3.2.7a\)](#) ⓘ - Tree Inference Using Bayesian Analysis - run on XSEDE

[Muscle \(3.7\)](#) ⓘ - Create Multiple Alignments from Sequences or Profiles

[NCLconverter \(2.1\)](#) ⓘ - A file format transformation tool

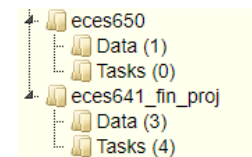
[Noisy on XSEDE \(1.5.12\)](#) ⓘ - Identify homo-plastic characters in multiple sequence alignments - run on XSEDE

New Task:  
Set  
Parameters  
(only if  
necessary)



#### Folders

Total Storage: 3.56 MB



#### Create new task

[Task Summary](#)

[Select Data](#)

[Select Tool](#)

[Set Parameters](#)

MAFFT on XSEDE: Multiple alignment program for amino acid or nucleotide sequences; parallel version ([Kazutaka Katoh](#), [Kei-ichi Kuma](#), [Hiroyuki Toh](#), and [Takashi Miyata](#))

#### [Simple Parameters](#)

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

Which version of MAFFT are you using? \* ☒ V 7.471

I need 60 GB of memory, a previous run ran out of memory ☐

There are unusual characters in the dataset (--anysymbol) \* ☐

Preserve case (--preserve-case) \* ☐

Output format \* ☒ FASTA ☐ ClustalW

#### [Advanced Parameters](#)



# New Task: Save and Run Task

CIPRES

Home

CIPRES Share

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status ▾

Log Out

Folders

Total Storage: 3.56 MB

eces650

Data (1)

Tasks (0)

eces641\_fin\_proj

Data (3)

Tasks (4)

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

Create new task

Task Summary

Select Data

Select Tool

Set Parameters

You may edit your task using the tabs above.

Current CPU Hr Usage: 9 [Explain this?](#)

Description

mafft

Input

1 Inputs Set

Tool

MAFFT on XSEDE

[Click for more info](#)

Input Parameters

46 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.

# View All Tasks

CIPRES

Home

CIPRES Share

Toolkit

My Profile

Help

How to Cite Us

XSEDE St

Folders

Total Storage: 3.90 MB

eces650

Data (1)

Tasks (2)

eces641\_fin\_proj

Data (3)

Tasks (4)

Tasks

Refresh

Current CPU Hrs Usage: 11

Explain this?

CPU Hrs Remaining: 49989

CPU Hrs Expire: 06/30/2021 00:00:00

Running Tasks: 1

There are currently 2 tasks in this tab. (Items 1 - 2 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"/>	Clone	muscle	Muscle	View (1)	View (28)	4/21/21, 16:40	View S
<input type="checkbox"/>	Clone	mafft	MAFFT on XSEDE	View (1)	View (22)	4/21/21, 16:38	View O

Move selected to

eces641\_fin\_proj

GO

Download

Download

### Task Details

Task	mafft
Owner	cjawesome
Group	cjawesome
Date Created	4/21/21, 16:38
Tool	MAFFT on XSEDE
Input	<a href="#">View (1)</a>
Parameters	<a href="#">View (22)</a>
Output	<a href="#">View (3)</a>
Intermediate Results	None
Status	COMPLETED

### Task Messages

Wed Apr 21 16:40:13 PDT 2021 > QUEUE : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Added to run queue.  
Wed Apr 21 16:40:15 PDT 2021 > COMMANDRENDERING : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Command rendered successfully: mafft\_7471.clang\_expance --thread 8 --inputorder --ep 0.0 --auto input.fasta > output.mafft  
Wed Apr 21 16:40:15 PDT 2021 > INPUTSTAGING : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Staging input files to expanse  
Wed Apr 21 16:40:15 PDT 2021 > INPUTSTAGING : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Input files staged successfully to /expanse/projects/ngbt/backend/expanse\_workspace/NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9/  
Wed Apr 21 16:40:16 PDT 2021 > SUBMITTED : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Submitted to expanse as job '2125684'.  
Wed Apr 21 16:41:14 PDT 2021 > LOAD\_RESULTS : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Trying to transfer results.  
Wed Apr 21 16:41:15 PDT 2021 > COMPLETED : SUCCESS : NGBW-JOB-MAFFT\_XSEDE-886AFB027CC54607AD4C02C9D30528D9 : Output files retrieved. Task is finished.

### Task Details

Task	muscle
Owner	cjawesome
Group	cjawesome
Date Created	4/21/21, 16:40
Tool	Muscle
Input	<a href="#">View (1)</a>
Parameters	<a href="#">View (28)</a>
Output	<a href="#">View (5)</a>
Intermediate Results	None
Status	COMPLETED

### Task Messages

Wed Apr 21 16:41:57 PDT 2021 > QUEUE : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Added to run queue.  
Wed Apr 21 16:42:06 PDT 2021 > COMMANDRENDERING : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Command rendered successfully: /projects/ps-ngbt/opt/comet/muscle/3.8.1551/muscle -in infile.fasta -seqtype auto -maxiters 16 -maxmb 30000000 -log logfile.txt -weight1 clustalw -cluster1 upgmb -sueff 0.1 -root1 pseudo -maxtrees 1 -weight2 clustalw -cluster2 upgmb -sueff 0.1 -root2 pseudo -objscore sp -noanchors -out output.fasta  
Wed Apr 21 16:42:06 PDT 2021 > INPUTSTAGING : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Staging input files to comet  
Wed Apr 21 16:42:06 PDT 2021 > INPUTSTAGING : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Input files staged successfully to /projects/ps-ngbt/backend/comet\_workspace/NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E/  
Wed Apr 21 16:42:08 PDT 2021 > SUBMITTED : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Submitted to comet as job '39193889'.  
Wed Apr 21 16:43:34 PDT 2021 > LOAD\_RESULTS : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Trying to transfer results.  
Wed Apr 21 16:43:35 PDT 2021 > COMPLETED : SUCCESS : NGBW-JOB-MUSCLE-A3312860B6AF42A3B4F51AE57977492E : Output files retrieved. Task is finished.

# Task Completion

# Output: MAFFT v. Muscle

J: > School Files > [In progress] ECES650 Stats in Metagenomics > Tutorial 1 Project CIPRES Alignments > mafft.mafft

```
781 >Haemophilus_influenzae_PittGG_uid58593_3
782 -----gaattgaag-----agt
783 ttgatcatggctcagat----tgaacgctggcggcaggcttaacacatgcaagtcgaac
784 ggtagcaggagaaaagcttcttcttctgacgagtgccggacgg-gtgagtaatgcttg
785 ggaatctggccttatgga-gggggataacgacgggaaactgtcgcataataccg----cgta
786 ttatcggg----agatgaaagtgccggactgagaggccgcatgccataggatgagcccaa
787 gtggg-----att
788 aggtagttgggtgggtaaaggcctaccaagcctgc-----gatctctagctggtctga
789 gaggatgaccagccacactggaactgagacacgggtccagact-cctacgggaggcagcag
790 tggggaatattgcgaatggggggaaccctgacgcagccatgccgcgtgaatgaagaagg
791 cttcgggttgtaaagtcttcttgggtattgaggaaggttgatgtttaatagcacatcaa
792 a---ttgacgttaaatacagaagaagcaccggct-aactccgtgccagcagccgcggtaa
793 tacggagggtgcgagcgttaatcggaataactgggcgtaaaggccacgcaggcggttatt
794 taagtgagggtgtaaaagccctgggccttaacctaggaattgcatttcagactgggtaacta
795 gagtacttttagggagggttagaattccacgtgtagcggtgaaat--gcgtagagatgtg
796 aggaataaccgaaggcgaaggcagcccttgggaatgtactgacgctcatgtgcgaagcg
797 tggggagcaaacaggattagataccctggtagtcacgctgtaaacgctgt-----c
798 gatttgggggttgggtttaactctggcgccttagctaacgtgataaatcgaccgctg
799 gggagtacggccg----caaggttaaaactcaaa----tgaattgacggggcccgca
800 caagcggtagcagcatgtggtttaattcgatgcaacgcgaagaaccttactactcttgac
801 atcctaa---gaagagctcagagatgagcttctgcttccgggaacttagagacagggtgc
802 tgcattgctgtgctcag----ctcgtgttgtaaatgtt----gggttaagtcccgcaa
803 -----cgagcgcaacccttatcttctgttgccagcgacttggtcgggaactcaaggag
804 actgccagtataaaactggaggaaggtggggatgacgtcaagtcacatggcccttacga
805 gtagggctacacagctgtacaaatggcgtatacagagggaagcgaagctgcgaggtggag
806 cgaatctcataaagtacg---tctaagtcgggattggagctgcaactcgactccatga
807 agtcggaatcgctagtaatcgcaatcagaatgtcgggtgaa---tacgttccgggcc
808 ttgtacacacgcccgtcacacatgggagtggtgtgaccagaagtagatagcttaaac
809 ttttgaggggcgtttaccacgggtatgattcatgactggggtgaagtcgtaacaaggtaac
810 cgtagggg---aacctgcggttggtacacctcctta-----
```

J: > School Files > [In progress] ECES650 Stats in Metagenomics > Tutorial 1 Project CIPRES Alignments > muscle.fasta

```
218 >Haemophilus_influenzae_PittGG_uid58593_3
219 -----GAATTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCTTAACACA
220 -----TGCAAGTCGAACGGTAGCAGGAG-----AAAGCTT--GCTTTCT-
221 -TGCTGAC-----GAGTGCCGGACGG-GTGAGTAATGCTTGGGAA----
222 -----TCTGGCTTATGG--A-GGGGGATAACGA-----CGGGAAC
223 TGTCGCTAATACCG-----CGTATTATCGGAAGATGAAAGT-GCGGGACTGAGAGGC-
224 -CGCATGCCATAGGATGAGCCCAAGTGGGAT-TAGGTAGT-----TGG
225 TGGGGTAA---AGGCCTACCAAGCCTGCATCTCTAGCTGG-----TCTGAGAGGA
226 TGACCAGCCACACTGGAAGTGA--GACACGGTCCAGACTC-CTACGGGAGGCAGCAGTGG
227 GGAATATTGCGCAATGGGGGAACCTGACG-CAGCCATGCCGCG-----
228 -TGAATGAAGAAG--GCCTTCGGGTGTAAAGTTCTTCGGTATTGAG-GAAGGTGAT
229 GTGTTAATAGCACATCAAAATTGACGTTAAATACAGAAGAAGCACCGGCTAACTCCGTGCC
230 AGCAGCCGCGGTAATACGGAGGGTGCAGCGTTAATCGGAATAACTGGGCGTAAAGGCA
231 CGCAGGCGGTTATTTAAGTGAGGTGTGA----AAGCCCTGGGCTTAACCTAGGAATT--
232 --GCATTT-CAGACTGGGTAAGTAGTACTTTAGGGAGGGGTAGAATCCACGTGTAGC
233 GGTGAAATGCGTAGAGATGTGGAGGAATACCGAAGGCGAAGGCAGCCCTTGGGAATGTA
234 CTGACGCTCATGTGCG-----AAGCGTGGGGAGCAACAGG--ATTAGATACCTT
235 GGTAGTCCACGCTGTAACGCTGTGATTTGGGGTTGGGGTTAACTCTGGCGCCGTA
236 GCTAACGTGATAAATCGACCGCTGGGGAGTACGGCG-----CAAGGTTAAACTCAAA
237 -----TGAATTGACGGG--GCCCGCAC-----AAGCGTGGAGCATGTGGTT----
238 --AATTCGATGCAACGCGAAGAACCTTACCTACTCTTGACATCCTAAGA----AGAGCT
239 CAGAGATGAGCTTGTCCTTCGGGAACCTAGAGACAGGTGCTGCATGGCTGTCGTAG--
240 ---CTCGTGTGTGAAATGTT----GGGTAAAGTCCCGCAACGAGCGCAACCTTATCC
241 TTTGTTGCCAGCG-ACCTGGTCGGGAACCAAGGAGAGTGCAGTGATAAACTGGAGGA
242 AGGTGGGGTAGCGTCAAGTCATCATGGCCCTTACGAGTAGGGCTACACACGTGTACAA
243 TGGCGTATACAGAGGGAAGCGAAGCTGCGAGGTGGAGCGAATCTATAAAGTACGTCTAA
244 GT-CCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGCGAAT
245 CAGAATGTCGCGGTGAATACG-----TTCCCGGGCCTTGTACACACCGCCGTCACACCA
246 TGGGAGTGGGTGTACCAGAAGTAGATAGCTTAACCTTTTGGAGGGCGTTTACCACGGTA
247 TGATTCTGACTGGGGTGAAGTCGTAACAAGGTAAACCGTAGGGG---AACCTGCGGTTGG
248 ATC-ACCTCCTTA-----
```

# References

[Miller, M.A., Pfeiffer, W., and Schwartz, T. \(2010\) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop \(GCE\), 14 Nov. 2010, New Orleans, LA pp 1 - 8.](#)