

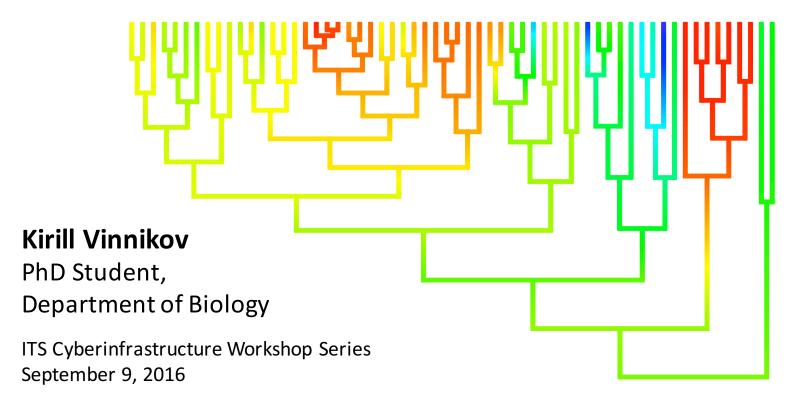








An Introduction to Molecular Phylogenetic Inference



4. PRACTICAL LAB 2

Outline

- Maximum Parsimony test in MEGA 7.0
- Neighbor-joining test in MEGA 7.0
- Choosing the best evolutionary model in jModelTest 2.1.10
- Choosing the best partitioning scheme
- Maximum Likelihood test in RAxML 8.2.9
- Bayesian Analysis in MrBayes 3.2.6

4. PRACTICAL LAB 2

Software requirements

- MEGA 7.0 (www.megasoftware.net)
- jModelTest 2.1.10 (github.com/ddarriba/jmodeltest2)
- PartitionFinder 1.1.1 (www.robertlanfear.com/partitionfinder)
- MrBayes 3.2.6 (mrbayes.sourceforge.net)
- FigTree 1.4.2 (tree.bio.ed.ac.uk/software/figtree)

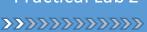
Maximum Parsimony in MEGA 7.0

Import concatenated alignment

- 1. Open File menu and select Convert file format to MEGA
- 2. Open bears.nex
- 3. Save converted file as bears.meg
- 4. Exit the MEGA file editor
- 5. Open bears.meg in **Data** menu

Maximum Parsimony in MEGA 7.0

- 1. Open **Phylogeny** menu and select **Maximum Parsimony** analysis
- 2. Set **bootstrap** test with **100** replicates
- 3. Set **gaps** for complete deletion
- 4. Set **SPR** as tree search method
- 5. Run the analysis
- 6. After completion, root the tree with the outgroup
- 7. Save the tree as image file (png)

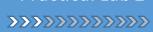




- Does the original tree differ from the bootstrap consensus tree?
- 2. Run the MP analysis with the larger number of bootstrap replications. How did the bootstrap support values change?
- 3. Run the MP analysis using only 1st and 2nd codon positions. How did the tree topology and bootstrap support values change?
- 4. Run the MP analysis with different tree search methods. Did the tree topology and bootstrap support values change?
- 5. Run the MP analysis for two genes separately (open *.fasta files). Does the tree topology differ between the two genes?

Neighbor-Joining test in MEGA 7.0

- 1. Using the same bears dataset, open Phylogeny menu and select **Neighbor Joining** analysis
- 2. Set **bootstrap** test with **100** replicates
- 3. Set gaps for complete deletion
- 4. Set **model** option to **p-distance**
- 5. Run the analysis
- 6. After completion, root the tree with the outgroup
- 7. Save the tree as image file (png)





- Does the original tree differ from the bootstrap consensus tree?
- 2. Run the NJ analysis with the larger number of bootstrap replications. How did the bootstrap support values change?
- 3. Run the NJ analysis using only 1st and 2nd codon positions. How did the tree topology and bootstrap support values change?
- 4. Run the NJ analysis with different substitution models. How did the tree topology and bootstrap support values change?
- 5. How does the tree topology differ between the corresponding NJ and MP trees?

Model Selection with jModelTest

- 1. Start the program **iModelTest.jar**
- Load the alignment for cytb gene: cytb.nex 2.
- Select "Compute Likelihood scores" in the "Analysis" menu 3.
- Select search through 7 substitution schemes and run the analysis 4.
- Select AIC, AICc, BIC or DT in the "Analysis" menu (you can do it by 5. clicking on the menu of each criteria one after another)
- 6. Open the results table in the "Results" menu
- Sort the results by the column with criterion scores (AIC, AICc, BIC, DT) 7.
- 8. Check the parameters of each selected model in the program console window



- 1. Do all four criteria support the same model?
- 2. If not, check how chosen models differ from each other by comparing their parameters (use the table provided on the lecture slide).
- 3. Compare the model chosen by AICc with the model that has the closest AICc score to it. How strong is the difference between the models (check delta)? Which parameters are different between these two models?
- 4. Which model would you choose for the following phylogenetic inference? Explain why.
- 5. Repeat Model Selection analysis for *irbp* gene: irbp.nex. Do *irbp* and *cytb* genes have similar substitution models? If not, explain how they differ.

Search for Partitions with PartitionFinder

- 1. Open examples/nucleotide folder and copy partition finder.cfg file to the folder with your alignment files
- 2. Open partition_finder.cfg file in the text editor (e.g., TextWrangler)
- 3. Change the name of the alignment to bears.phy
- 4. Change models to raxml
- 5. Change the data block section as the following:

```
cytb pos1 = 1-1140\3;
cytb pos2 = 2-1140\3;
cytb pos3 = 3-1140 \ 3;
irbp pos1 = 1141-2420\3;
irbp pos2 = 1142-2420\3;
irbp pos3 = 1143-2420\3;
```

6. Save the file

- 1. Open Terminal (Mac) or Command Prompt (Windows)
- 2. Type "python" (including space!!!)
- 3. Drag and drop PartitionFinder.py file and press space
- 4. Drag and drop the folder with your sequences and partition_finder.cfg file
- 5. Press ENTER
- 6. After analysis completion, you should find the **analysis** folder within the folder containing your sequence files
- 7. Open best_scheme.txt file in your text editor to see the results
- 8. Copy and save the RAxML partition section into a separate file: partitions.txt (the section header should not be included!)

HOMEWORK QUESTIONS

- 1. How many partitions have been revealed in your alignment data?
- 2. Do the partition models correspond with the models that you have obtained in the jModelTest analysis?





Welcome to CIPRES science gateway portal!







Home »

Submissions are working normally

- Codes
- Requirements
- Limitations
- Architecture
- Known Issues
- Usage Statistics
- User Locations
- Survey Results
- Publications

The CIPRES Science Gateway V. 3.3

The CIPRES Science Gateway V. 3.3 is a public resource for inference of large phylogenetic trees. It is designed to provide all researchers with access to NSF XSEDE's large computational resources through a simple browser interface. You can now also access these same capabilities programatically with the CIPRES REST API.

High Performance Parallel Codes for Large Tree Inference and Sequence Alignment on XSEDE: RAXML; MrBayes; BEAST; BEAST2; GARLI; MAFFT; DPPDIV; FastTree, jModelTest2, and Migrate-N. If you need access to PhyloBayes, please inquire.

Serial Codes for Tree Inference:

PAUP* (Inference by Parsimony); Poy (Alignment and Inference);

Serial Codes for Sequence Alignment:

ClustalW; Contralign; MUSCLE; PROBCONS; PROBALIGN

Use the CIPRES Science Gateway



News

CIPRES now offers TreeAnnotator and LogCombiner

July 07, 2016 - 5:35 am

New CIPRES Rest Tool Configuration helper!

February 25, 2016 - 10:55 am

Locations of last 1000 users







Login as a Guest User







CIPRES

Home

Toolkit

Help

How to Cite Us

Missing results?

Send us the job handle, and we may be able to help.

More Information

About Us
Requirements
Known Issues
Usage Statistics
User Locations
Enabled Publications

The CIPRES Science Gateway now offers BEAST2 and PhyloBayes MPI, along with RAxML, MrBayes and other codes.

iPlant users login here:

What is this?

iPlant Collaborative

Empowering A New Plant Biology

First Time Users: Please review the XSEDE Primer and our Fair Use Policy.

CIPRES Login:

Username

Username

Password

Login

Reset

Forgot Password?

Register | Proceed without Registering



Status: All submissions are working normally.

CIPRES Gateway News

CIPRES now provides
TreeAnnotator and LogCombiner

New CIPRES Rest Tool Configuration helper!

Error in Gotoh's alignment algorithm

www.phylo.org









Practical Lab 3





Click on your Guest Folder

Help







CIPRES

Home

Toolkit

t

My Profile

How to Cite Us

XSEDE Status -

Logout

Folders

Total Storage: 0 bytes

☐ Guest Folder

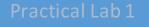


Guest user "Guest-1472866529957" successfully logged in. User "Guest-1472866529957" successfully logged in.

Welcome Guest-1472866529957 to the CIPRES Science Gateway V 3.3. If you would like to view a demo, click here.

Click on a folder in the left panel to manage its contents.





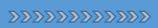


Practical Lab 2

Practical Lab







Upload your DATA files:

bears.phy partitions.txt







CIPRES

Home

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status -

Logout

Folders

Total Storage: 0 bytes

Data (0)

Tasks (0)

There is currently no data in this folder. What would you like to do?

Upload/Enter Data

Create a Task







Upload your DATA files:

bears.phy partitions.txt







CIPRES

Home

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status ▼

Logout

Folders

Total Storage: 0 bytes

Guest Folder

- Data (0)
- Tasks (0)

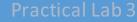
Upload File					
Upload your files	Choose Files no files selected				
	You can select multiple files.				
	MSIE 9 and below support single uploads only.				
You can also enter y	our data manually below				
Label (required)	Label				
Data:	Enter your data				
	Save Cancel				





















CIPRES

Home

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status ▼

Logout

Folders

Total Storage: 0 bytes



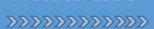


There are currently no tasks in this folder.

Create New Task







Provide the name for your task: RAxML







CIPRES

Home

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status ▼

Logout

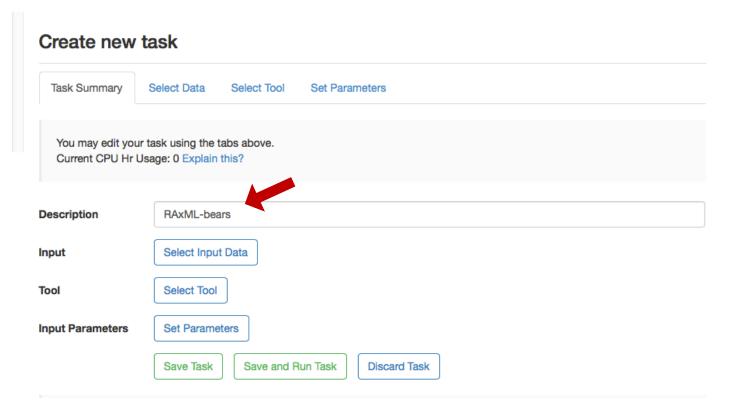
Folders

Total Storage: 0 bytes

Guest Folder

🗅 Data (2)

Tasks (0)



114

Unknown





Select input file: bears.phy





Date Created

9/2/16, 18:41

9/2/16, 18:41



CIPRES

Home

Toolkit

My Profile

Help

Select Data

How to Cite Us

XSEDE Status -

Logout

Folders Create new task Total Storage: 0 bytes Select Data Set Parameters Task Summary Select Tool Guest Folder Data (2) Tasks (0) You can choose the following data. Select One Label Bytes **Data Format** bears.phy 24388 Unknown

partitions.txt

Cancel





Select tool: RAxML-HPC v.8 on XSEDE

BEAST2 on XSEDE (2.3.2)

Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE

BEAST on XSEDE (1.8.0; 1.8.1; 1.8.2; 1.8.3) 📵 - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE

Clearcut (1.0.9) 1 - Fast Implementation of Relaxed Neighbor Joining

ClustalW (1.82) 1 - Create Multiple Alignments from Sequences

Consense (Phylip 3.66) 🕦 - Find A Consensus Tree

DPPDIV on XSEDE (1.0)
 - Estimating species divergence times and lineage-specific substitution rates on a fixed topology run on XSEDE

FastTreeMP on XSEDE (2.1.8) 👔 - Fast (Approximate) Maximum Likelihood tree construction - run on XSEDE

GARLI 2.01 on XSEDE (2.01) 1 - Genetic Algorithm for Rapid Likelihood Inference - run on XSEDE.

GARLI.conf Creator (2.0) 1 - Creates a Garli.conf file for up to five partitions

JModelTest2 on XSEDE (2.1.6) ** - Statistical selection of best-fit models of nucleotide substitution, run on XSEDE

LogCombiner on XSEDE (1.8.3) 1 - Bayesian Evolutionary Analysis by Sampling Trees - run on XSEDE

MAFFT on XSEDE (7.187) (i) - Multiple alignment program for amino acid or nucleotide sequences; parallel version

Migrate-N on XSEDE (3.6.11) 1 - Estimation of Population Sizes and Gene Flow using the Coalescent

MrBayes Restart on XSEDE (3.2.x) 🕡 - Tree Inference Using Bayesian Analysis - run on XSEDE

MrBayes on XSEDE (3.2.6) 🕡 - Tree Inference Using Bayesian Analysis - run on XSEDE

Muscle (3.7) 1 - Create Multiple Alignments from Sequences or Profiles

NCLconverter (2.1) - A file format transformation tool

PAUPRat (Not specified) 1 - Parsimony ratchet searches using PAUP*

POY (4.1.2) - Phylogenetic tree inference using dynamic homologies.

Probalign (1.3) 1 - Multiple sequence alignment using partition function posterior probabilities.

ProbCons (1.12) 1 - Probabilistic Consistency-based Multiple Alignment of Amino/Nucleic Acid Sequences

RAXML-HPC BlackBox (8.2.8) 📵 - Phylogenetic tree inference using maximum likelihood/rapid bootstrapping on XSEDE.

RAXML-HPC2 on XSEDE (8.2.8) 👔 - Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE

RAXML-HPC2 Workflow on XSEDE (8.2.8)
 - Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE



RAXML-HPC v.8 on XSEDE (8.2.8)
- NEW Interface! Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE

Readseq (2.2) 1 - Readseq: biological sequence format interconversion utility







Parameters

Create new task

Task Summary	Select Data	Select Tool	Set Parameters			
	kelihood/				enetic tree in SEDE (Alexand	
Simple Parameters						
Maximum Hours to	Run (click here fo	or help setting thi	s correctly) * 0.25			
Set a name for outp	ut files (-n) * bea	rs_ML				
Enable ML searches						
Outgroup (-o) ()one	or more comma-	separated outgro	oups, see comment f	or syntax)		
Specify the number	of distinct rate ca	ategories (-c) * 2	5			
Disable Rate Hetero	geneity (-V) *					
Supply a tree (Not a	vailable when do	ing rapid bootstr	apping, -x) (-t)	⇒		
Specify a random se	eed value for pars	simony inference	s (-p) * 🔽			
Enter a random seed	d value for parsin	nony inferences	-p "value" gives repr	oducible results	from random starting to	ree) * 12345
Specify an initial rea	rrangement setti	ng (-i) * 🗆				
Specify the distance	from original pru	uning point (-i) *	10			
Constraint (-g)	•					
Binary Backbone (-r) Use a mixed/partitio		partitions but				
Evaluate partitions of						
Estimate individual p	-					
Estimate proportion	of invariable site	s (GTRGAMMA	⊦ I) * ⊜yes ⊙no			
Choose an input file	that excludes th	e range of position	ons specifed in this f	ile (-E)		
Weight characters a	s specifed in this	file (-a)	<u>^</u>			
Disable checking for	•					





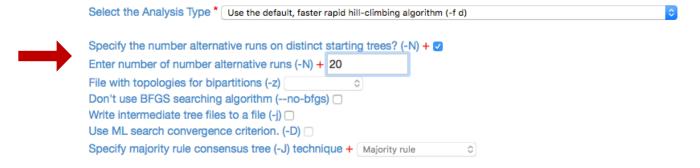
Parameters

Create new task



RAxML-HPC v.8 on XSEDE: NEW Interface! Phylogenetic tree inference using maximum likelihood/rapid bootstrapping run on XSEDE (Alexandros Stamatakis)

Configure the Analysis





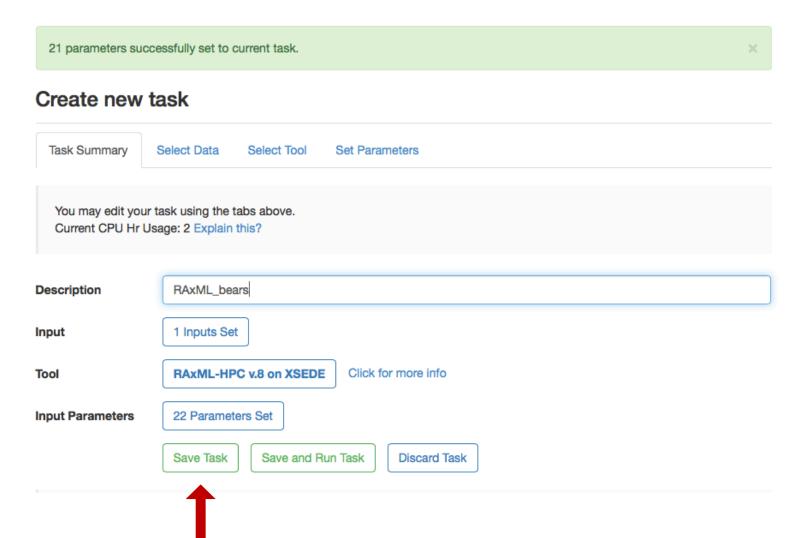
















Run task







CIPRES

Home

Toolkit

My Profile

Help

How to Cite Us

XSEDE Status ▼

C Refresh Tasks

Page 1 of 1

Logout

Folders

Total Storage: 0 bytes

Guest Folder

Data (2)

Tasks (1)

Tasks

Current CPU Hr Usage: 0 Explain this?

There is currently 1 data item in this tab.

Create New Task

records on each page

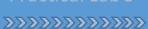
Select All		Label	Tool	Input	Parameters	Date Created	Action
	Clone	RAxML_bears	RAXML-HPC v.8 on	View (1)	View (21)	9/2/16,	Run Task
			XSEDE			19:38	1

Kill and Delete Selected

Move selected to Guest Folder

GO





Proceed to results







CIPRES

Home

Toolkit

My Profile

Help How

How to Cite Us

GO

XSEDE Status -

Page 1 of 1

Logout

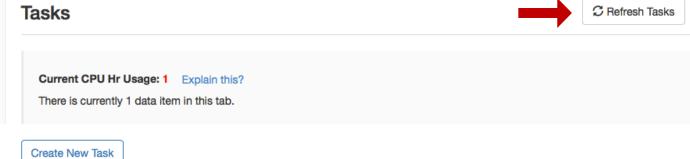
Folders

Total Storage: 0 bytes

Guest Folder

Data (2)

Tasks (1)



Show 20 records on each page

Select All		Label	Tool	Input	Parameters	Date Created	Action
	Clone	RAxML_bears	RAxML-HPC v.8	View (1)	View (21)	9/2/16,	View Output
			on XSEDE			19:38	

Kill and Delete Selected

Move selected to Guest Folder 😊



	RAxML_log.bears_ML.RUN.8	198	View	Download
	RAxML_result.bears_ML.RUN.16	582	View	Download
0	RAxML_log.bears_ML.RUN.2	198	View	Download
0	RAxML_result.bears_ML.RUN.1	582	View	Download
0	RAxML_parsimonyTree.bears_ML.RUN.22	187	View	Download
0	RAxML_log.bears_ML.RUN.12	198	View	Download
0	RAxML_log.bears_ML.RUN.18	198	View	Download
0	done.txt	49	View	Download
0	RAxML_parsimonyTree.bears_ML.RUN.5	187	View	Download
0	RAxML_log.bears_ML.RUN.20	198	View	Download
0	RAxML_log.bears_ML.RUN.10	198	View	Download
	RAxML_bestTree.bears_ML	582	View	Download
0	RAxML_log.bears_ML.RUN.22	198	View	Download
0	RAxML_parsimonyTree.bears_ML.RUN.7	187	View	Download
0	RAxML_parsimonyTree.bears_ML.RUN.18	187	View	Download
0	RAxML_parsimonyTree.bears_ML.RUN.12	187	View	Download
0	RAxML_result.bears_ML.RUN.9	582	View	Download
0	RAxML_result.bears_ML.RUN.14	582	View	Download
0	RAxML_log.bears_ML.RUN.0	198	View	Download
0	RAxML_result.bears_ML.RUN.3	582	View	Download
0	RAxML_parsimonyTree.bears_ML.RUN.20	187	View	Download







Run the RAxML bootstrap analysis

Create new task Task Summary Select Data Select Tool Set Parameters RAxML-HPC v.8 on XSEDE: NEW Interface! 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) * 0.25 Set a name for output files (-n) * bears_boot 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) * 25 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) * 12345 Specify an initial rearrangement setting (-i) * Specify the distance from original pruning point (-i) * 10 Constraint (-g) Binary Backbone (-r) Use a mixed/partitioned model? (-q) partitions.txt 😊 Evaluate partitions only under HKY85 (--HKY85) Estimate individual per-partition branch lengths (-M) * Estimate proportion of invariable sites (GTRGAMMA + I) * yes ono Choose an input file that excludes the range of positions specifed in this file (-E) Weight characters as specifed in this file (-a) Disable checking for sequences with no values (-O) Print output files that can be parsed by Mesquite. (-mesquite)





Run the RAxML bootstrap analysis

Choose a Bootstrapping Type + Non-parametric Boostrapping (-b) Rapid Bootstrapping (-x) No Bootstrapping Enter a random seed value for bootstrapping + 12345 Print branch lengths (-k) Specify bootstrap protocol + Specify an explicit number of bootstraps Let RaxML halt bootstrapping automatically Bootstrap iterations (-N) + 100 Select Bootstopping Criterion: (autoMRE is recommended) + autoMRE Select the criterion for a posteriori bootstopping analysis (-l) + autoMRE File with topologies for a posteriori bootstopping (-z) Save Parameters Reset Cancel

Download the file with the bootstrap trees

View Task Output

Click on an output file below to review its contents.

☐ Select all	Tool Output	File Name	File Size (Bytes)		
	PROCESS_OUTPUT	STDOUT	17177	View	Download
		STDERR	0	View	Download
	all_outputfiles	scheduler.conf	90	View	Download
		stdout.txt	17177	View	Download
0		part.txt	114	View	Download
		infile.txt	24388	View	Download
		_scheduler_stderr.txt	634	View	Download
		done.txt	49	View	Download
		stderr.txt	0	View	Download
		RAxML_info.bears_boot	12690	View	Download
		term.txt	322	View	Download
		start.txt	40	View	Download
		RAxML_bootstrap.bears_boot	20196	View	Download
		_JOBINFO.TXT	343	View	Download

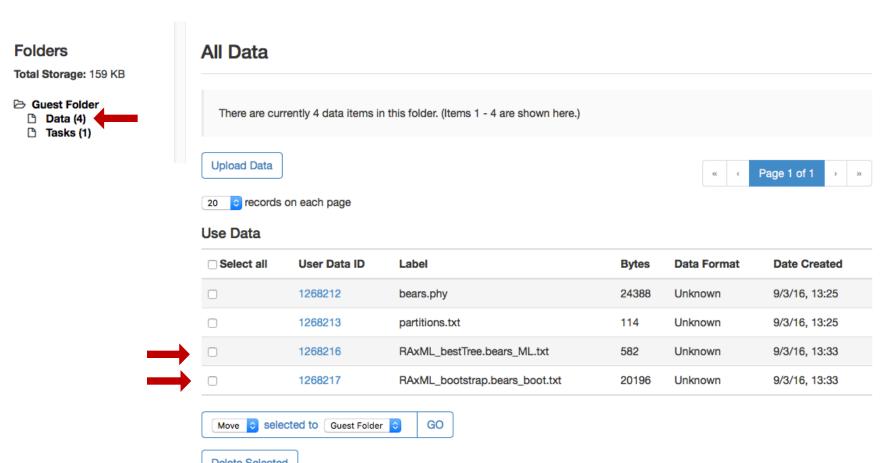






Upload the ML tree and bootstrap files to your data folder









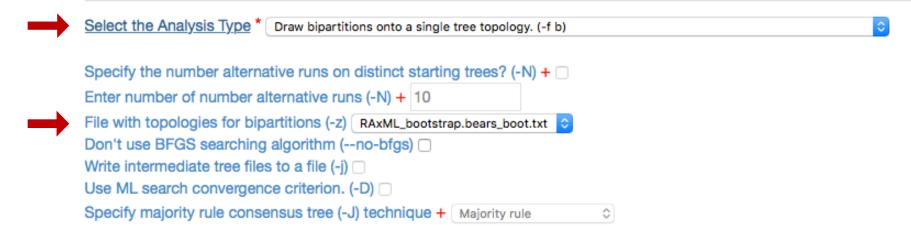
Draw bipartitions on the best ML tree

Create new task

Task Summary	Select Data	Select Tool	Set Parameter	rs			
	kelihood/					c tree inferei (Alexandros	
Simple Parameters							
Maximum Hours to F	Run (click here fo	r help setting th	is correctly) * 0.25	5			
Set a name for output	ut files (-n) * bea	rs_final					
Enable ML searches	under CAT (-F) *						
Outgroup (-o) ()one of	r more comma-s	separated outgr	oups, see comme	nt for syntax)			
Specify the number	of distinct rate ca	ategories (-c) *	25				
Disable Rate Heterog	geneity (-V) *						
Supply a tree (Not av		ing rapid bootst	rapping, -x) (-t) R	AxML_bestTree.b	ears_ML.txt	•	
Specify a random se	ed value for pars	imony inference	es (-p) * 🗸			_	
Enter a random seed	I value for parsin	nony inferences	(-p "value" gives r	eproducible re	sults from rar	ndom starting tree) * 1	2345
Specify an initial rear	rangement settir	ng (-i) * 🗆					
Specify the distance	from original pru	ıning point (-i) *	10				
Constraint (-g)		♦					
Binary Backbone (-r)							
Use a mixed/partition Evaluate partitions o			<u> </u>				
Estimate individual p	-		*				
Estimate proportion	•						
Choose an input file		,	,	is file (-E)		0	
Weight characters as		0		•			
Disable checking for	coguenese with	no voluce (O) C					

Draw bipartitions on the best ML tree

Configure the Analysis

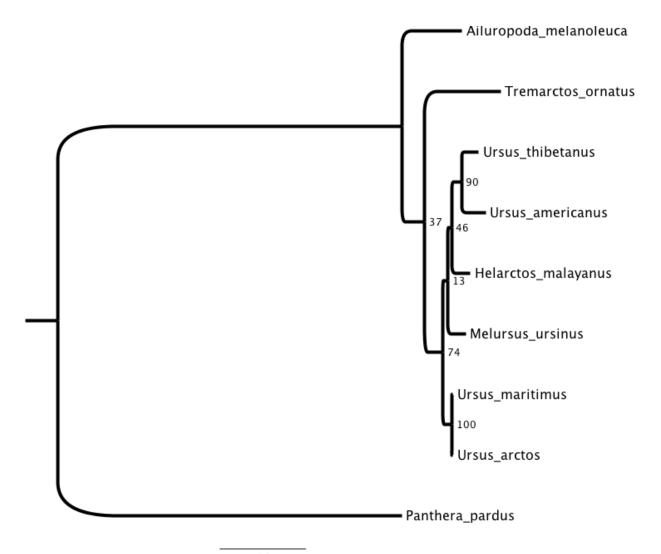


Download the file with the final tree

View Task Output

Click on an output file below to review its contents.									
Select all	Tool Output	File Name	File Size (Bytes)						
	PROCESS_OUTPUT	STDOUT	2588	View	Download				
		STDERR	0	View	Download				
	all_outputfiles	done.txt	49	View	Download				
		part.txt	114	View	Download				
		start.txt	40	View	Download				
		term.txt	317	View	Download				
		RAxML_bipartitionsBranchLabels.bears_final	606	View	Download				
		scheduler.conf	62	View	Download				
		infile.txt	24388	View	Download				
		stdout.txt	2588	View	Download				
		stderr.txt	0	View	Download				
		RAxML_info.bears_final	1475	View	Download				
		_scheduler_stderr.txt	634	View	Download				
		_JOBINFO.TXT	344	View	Download				
		topologies_file.tre	20196	View	Download				
		RAxML_bipartitions.bears_final	592	View	Download				
		tree.tre	582	View	Download				

Visualize the final tree in FigTree







MrBayes v3.2.6 x64

(Bayesian Analysis of Phylogeny)

Distributed under the GNU General Public License

Type "help" or "help <command>" for information on the commands that are available.

> Type "about" for authorship and general information about the program.

MrBayes > help





Bayesian Analysis:

1. Read the alignment file

MrBayes > execute bears.nex

2. Define data partitions

```
MrBayes > charset cytb = 1-1140
```





3. Set the parameters for likelihood model

MrBayes > showmodel

MrBayes > help lset

MrBayes > lset applyto=(1) code=vertmt nst=2 rates=gamma

MrBayes > lset applyto=(2) nst=2 rates=gamma

Bayesian Analysis:

4. Set the priors for phylogenetic model

```
MrBayes > help prset
```

```
MrBayes > unlink statefreq=(all) shape=(all) revmat=(all)
```

```
MrBayes > prset applyto=(1) shapepr = exponential(10.0)
```

Bayesian Analysis:

5. Set the MCMC parameters

```
MrBayes > showmoves
```

MrBayes > help mcmcp

MrBayes > mcmcp ngen=1000000 nruns=2 nchains=2 samplefreq=200 burninfrac=0.2

MrBayes > mcmc

BREAK (10 min)