

Name: Rajashree. U

Contact:rajashree5820@gmail.com

Mobile: +916383331696

Title: Phylogenetic Tree Analysis

Answer the following questions below and upload them in the google form.

1. Investigate the evolutionary relationships among bats, whales, and artiodactyls (even-toed ungulates) groups.
 - a. Choose 3 to 5 species in each group. Mention each species, i. Name, ii. GenBank Accession Number, iii. Sequence Length & iv. Sequence.
 - b. Construct a phylogenetic tree using sequence data for the chosen species.
 - c. Analyze the tree to identify key evolutionary events and relationships in about 350 words.
 - d. Compare and contrast the evolutionary histories of the groups in about 250 words.
 - e. Discuss the implications of your findings for our understanding of mammalian evolution in about 150 words.

PHYLOGENETIC TREE ANALYSIS

Species Details:

The phylogenetic tree analysis is constructed based on the outgroup method. Here it is between mammals (bats, whales, artiodactyls) and non-mammal (peacocks) taken as an outgroup, the analysis is done using a particular gene Cytochrome B (region of the mitochondrial genome) for determining the phylogenetic relationship between organisms due to its sequence variation and to see how there are evolved.

The following gives the details of the species used in this analysis.

Bats:

Rhinolophus ferrumequinum korai mitochondrion, complete genome. NC_016191 /1140bp

>NC_016191.1:14166-15305 *Rhinolophus ferrumequinum korai* CYTB

```
ATGACCAACATTTCGCAAGTCTCACCCACTATTCAAATTATCAACGACTCATTGTCGACCTACCCGCCC
CATCAAGTATCTCTTCTGATGAACTTCGGATCCCTCCTGGGAATCTGCCTAGCTATCCAAATTCTCAC
AGGACTTTTCTAGCAATACACTATACATCAGATACTGCCACAGCCTTCCACTCCGTAACCCACATTTGC
CGAGACGTCAACTACGGTTGAGTCCTGCGCTATCTTCATGCCAACGGAGCCTCTATATTCTTTATCTGCT
TGTTCTACACGTAGGACGAGGAATTTACTACGGCTCCTATACATTCTCAGAAACATGAAACATTGGAAT
CATTCTTCTCTTCGCCGTCATAGCCACAGCATTATAGGCTACGTACTCCCATGAGGCCAAATGTCCTTC
TGAGGAGCAACAGTCATCACAACCTTCTCTCAGCCATTCCATACGTCGGAACAACCTCTAGTTGAATGAG
TCTGAGGCGGATTCTCAGTTGATAAAGCCACACTCACCCGATTCTTCGCTCTGCACTTCCTACTACCTTT
CATCATCGCAGCCATAGTCATAGTCCATCTACTTTTCTCCATGAAACAGGATCAAACAACCCAACTGGA
ATCCCATCAGACGCAGACATAATTCCATTCCACCCTTACTATACCATTAAGATATCCTGGGCCTCGTAC
TAATGCTGATAGCACTACTGTCTTTGGTCCTATTCGCCCCAGACCTATTAGGCGATCCAGACAACCTACAC
TCCAGCCAACCCACTAAACACCCACCCACATTAAACCAGAGTGATACTTTCTATTTGCCTACGCAATC
CTACGTTCAATCCCAAATAAACTCGGCGGAGTTGTAGCCCTAGTCCTATCTATCCTTATCTTAGCTGCCA
TCCCACTACTCCACACATCAAAACAACGCAGCATAGCATTCCGACCCTTAAGCCAATGCCTATTCTGACT
CCTAGTGGCGGATCTCCTCACACTAACCTGAATCGGAGGCCAACCTGTGGAACACCCATTCAATTATCATT
GGACAACCTAGCCTCCATTCTATACTTCCTAATTATCCTCGTCCTAATACCACTTGCAAGCATCGCAGAAA
ACCATCTATTGAAATGAAGA
```

Rhinolophus pumilus mitochondrion, complete genome. NC_005434/1140bp

>NC_005434.1:14171-15310 Rhinolophus pumilus CYTB

ATGACCAACATTCGTAAATCCCACCCATTATTTAAGATTATCAACGACTCATTATCGACTTACCAGCCC
CATCAAGCATCTCCTCCTGATGAAACTTCGGGTCCCTACTAGGAGTCTGCCTAGCCGTACAAATCCTCAC
AGGCCTATTCTAGCCATACACTACACATCAGACACCGCCACTGCCTTCTACTCCGTAACCCATATCTGC
CGAGACGTCAATTACGGCTGAGTCCTACGCTACCTTCATGCCAACGGAGCCTCCATATTCTTTATCTGCC
TATTTCTGCATGTAGGACGGGGAATCTACTATGGCTCCTATACATTCTCTGAAACATGAAATATCGGAAT
TATTCTCCTCTTCGCCGTCATAGCCACAGCATTATAGGCTACGTACTTCCATGAGGCCAAATATCCTTC
TGAGGAGCAACAGTCATCACAACCTCCTCTCAGCCATTCCATACATCGGAACAACCCTAGTAGAATGGG
TCTGAGGAGGATTCTCAGTCGACAAAGCCACACTCACCCGATTCTTCGCCCTACACTTTCTACTCCCCTT
CATTATCTCAGCTATAGTTATAGTCCACCTACTCTTTCTCCACGAAACAGGCTCAAACAACCCAACTGGA
ATTCCATCCGATGTGGATATAATCCATTCCACCCCTACTACACCATTAAGATATCCTAGGCCTCGTGC
TAATACTAATAGCATTACTATCCCTAGTCCTATTGCCCCCGATCTGCTAGGCGACCCAGACAACCTACAC
CCCAGCCAACCCACTTAACACCCACCCCATATTAACAGAGTGATACTTCTATTTGCATATGCAATC
CTACGCTCAATCCCAAACAACTCGGCGGAGTTGTAGCCCTAGTACTGTCTATTCTTGTCTGGCTGTCA
TCCCACTACTCCACACATCAAACAGCGCAGCATGACTTTCCGACCTCTAAGCCAATGCTTATTTTGACT
CCTGGTAGCAGACCTTCTTACACTAACCTGAATTGGAGGCCAACCTGTGGAACACCCATTCAATTATCATC
GGACAACCTAGCCTCTATCCTGTACTTTCTAATTATCCTAGTCCTAATACCACTTGCAAGCATCGCAGAAA
ACCATCTATTAATGAAGA

Rhinolophus monoceros mitochondrion, complete genome. NC_005433 /1140bp

>NC_005433.1:14171-15310 Rhinolophus monoceros CYTB

ATGACCAACATTCGTAAATCCCACCCACTATTTAAGATCATCAACGACTCGTTCATCGACCTACCAGCCC
CATCAAGCATCTCCTCCTGATGAAACTTCGGGTCCCTACTAGGAGTCTGCCTAGCCGTACAAATCCTCAC
AGGCCTATTCTGGCCATACACTACACACCAGACACCGCCACTGCCTTCTACTCCGTAACCCATATCTGC
CGAGACGTCAATTACGGCTGAGTTCTACGCTACCTCCATGCCAACGGAGCCTCCATATTCTTTATCTGCC
TATTTCTGCATGTAGGACGGGGAATCTACTATGGCTCCTATACATTCTCTGAGACATGAAATATCGGAAT
TATTCTCCTCTTCGCCGTTATAGCCACAGCATTATAGGCTACGTACTTCCATGAGGCCAAATATCCTTC
TGAGGAGCAACGGTCATTACAAACCTCCTCTCAGCCATTCCATACGTGGAACAACCCTAGTAGAATGAG
TCTGAGGAGGATTCTCAGTAGACAAAGCCACACTCACCCGATTCTTCGCCCTACACTTTCTACTCCCCTT
CATTATCTCAGCTATAGTTATAGTCCACCTACTCTTCTCCACGAAACAGGCTCAAACAACCCAACTGGA
ATTCCATCCGATGTGGATATAATTCCATTCCACCCCTACTACACCATTAAGACATCCTAGGCCTCGTGC
TAATACTAATAGCACTACTATCTCTAGTCCTATTGCCCCGACCTGCTAGGCGACCCAGACAACCTACAC
CCCAGCTAACCCACTTAACACCCACCCCATATTAACAGAGTGATACTTCTATTTGCATATGCAATC
CTACGCTCAATCCCAAACAACTCGGCGGAGTTGTAGCCCTAGTACTGTCTATTCTTGTCTGGCTGTCA
TCCCACTACTCCACACATCAAACAAACGCAGCATGACTTTCCGACCTCTAAGCCAATGCTTATTTTGACT
CCTGGTAGCAGACCTTCTTACACTAACCTGAATTGGAGGCCAACCTGTTGAACACCCATTCAATTATCATC
GGACAACCTAGCCTCCATCCTATACTTTCTAATTATCCTAGTCCTAATACCACTTGCAAGCATCGCAGAAA
ACCATCTATTAATGAAGA

Whales

Balaenoptera borealis mitochondrion, complete genome.NC_006929 /1140bp

>NC_006929.1:14199-15338 *Balaenoptera borealis* CYTB

```
ATGACCAACATCCGAAAAACACACCCACTAATAAAGATTGTCAACGATGCATTGCTTGATCTCCCCACCC
CATCAAATATCTCCTCATGATGAAATTTGGGCTCCCTACTCGGCCTCTGCTTAATTACACAAATTCTAAC
AGGCCTATTCTAGCAATACACTACACACCAGACACAACAACCGCTTTCTCATCAGTTACACACATTTGC
CGAGACGTAAACTACGGCTGAATTATCCGATACCTACACGCAAACGGAGCCTCCATATTTTTCATCTGCC
TCTACGCCACATGGGACGAGGCCTATACTACGGCTCCTACGCCTTTGAGAAACATGAAATATTGGAGT
TATCCTACTATTACAGTTATAGCCACCGCATTGCTAGGCTACGTCTACCCTGAGGACAAATATCATTT
TGAGGCGCAACCGTCATCACCAACCTCTTATCAGCAATCCCATACATTGGTACTACCCTAGTCGAATGGA
TCTGAGGCGGTTTCTCTGTAGATAAAGCAACACTAACACGCTTTTTTGCCTTCCACTTCATTCTCCCCTT
CATTATTCTAGCACTAGCAATGGTCCACCTCATTTTCTCCATGAAACAGGATCCAACAACCCACAGGT
ATTCCATCCGACATAGACAAAATCCCATTCACCTTACTACACAGTTAAAGACATTCTAGGCGCCCTAC
TACTAATCCTAACCCTACTAATACTAACCCTATTGCGACCCGACCTGCTTGGAGACCCAGACAACTACAC
CCCAGCAAATCCACTCAGTACCCAGCACACATTAAACCAGAATGATATTTCTATTGTCATACGCAATC
CTACGATCAATCCCCAACAAATTAGGCGGAGTCTTAGCCCTATTACTCTCAATCCTAATCCTAGCCCTAA
TCCCAATACTCCACACATCTAAACAACGAAGCATAATATTCCGACCCTTAGCCAATTCCTATTTTGAGT
CCTAGTCGCAGACTTACTAACCCTGACATGAATCGGTGGCCAACCTGTGGAACACCCCTATGTGATCGTA
GGCCAATTTGCATCCATCCTCTATTTCTCCTAATTCTAGTATTAATACCAGCAACTAGCCTTATCGAGA
ACAAACTTATAAAATGAAGA
```

Eubalaena japonica mitochondrion, complete genome. NC_006931

>NC_006931.1:14191-15330 *Eubalaena japonica* CYTB

```
ATGACCAACATCCGAAAAACACACCCACTAATAAAAATTATTAACGACGCCTTCATCGATCTCCCCACTC
CATCAAATATCTCTTCATGATGAAACTTCGGCTCCCTACTTGGCCTTTGTCTAATTATACAAATCCTAAC
AGGCCTATTCTAGCAATACACTACACACCAGACACAACAACCGCCTTCTCATCAATCACACACATCTGT
CGAGACGTAAACTACGGTTGAATTATCCGGTATCTACATGCAAACGGAGCCTCCATATTCTTTATTTGCC
TCTACGCCACATAGGACGCGGCCTATACTACGGCTCCTACGCCTTTCAAGAAACATGAAACATTGGAGT
AATCCTACTATTACAGTCATAGCCACTGCATTGCTAGGCTATGTCTGCCCTGAGGACAAATATCATTC
TGAGGTGCAACCGTCATCACCAATCTCCTATCAGCAATCCCATACATCGGTAACACCTTAGTCGAATGAA
TCTGAGGTGGTTTCTCCGTAGACAAAGCAACACTAACACGCTTCTTGCCTTCCACTTCATCCTTCCCTT
CATCATCCTAGCATTAGCAATCGTCCACCTTCTATTTCTCCACGAAACAGGATCTAACAATCCTACAGGC
ATCCCATCTAACATAGATAAAAATTCCATTCCACCCCTATTACACAATCAAAGATATCCTAGGCGCCCTAC
TACTAATCCTAACCCTACTAATACTGACCCTATTGCGACCTGACCTGCTTGGAGACCCCGATAACTACAC
CCCAGCAAACCCCTCAGCACCCAGCACATATTAACCAGAATGATATTTCTATTGTCATACGCAATC
CTACGATCAATTCCCCAACAAATTAGGCGGAGTCTAGCCCTATTACTCTCAATCCTAATCCTAGCCTTCA
TCCCAATACTCCATACATCCAAACAACGAAGCATAATATTCCGACCCTTCAGCCAATTTCTGTTCTGAGT
ATTAGTCGCAGATCTACTAACCCTAACATGAATTGGCGGCCAACCCGTGGAACACCCCTATGTGATCGTA
GGCCAATTCGCATCCATCCTTTACTTTCTCCTAATTCTAGTTTTAATACCAACGGCTAGTCTTATCGAAA
ATAAACTCATAAAGTGAAGA
```

Eubalaena australis mitochondrion, complete genome. NC_006930

>NC_006930.1:14189-15328 Eubalaena australis CYTB

```
ATGACCAACATCCGAAAAACACACCCACTAATAAAAATTATTAATGACGCCTTCATCGACCTCCCCACCC
CATCAAATATCTCTTCATGATGAAACTTCGGCTCCCTACTTGGCCTTTGTCTAATTATACAAATCCTAAC
AGGCCTATTCTAGCAATACACTACACACCAGACACAACAACCGCCTTCTCATCAATCACACACATCTGT
CGAGACGTAAACTACGGTTGAATTATCCGATATCTACATGCAAACGGAGCCTCCATATTCTTCATTGTC
TCTACGCCCACATAGGACGCGGCCTATACTACGGCTCCTACGCCTTTCAAGAAACATGAAACATTGGAGT
AATCCTACTATTACAGTCATAGCCACTGCGTTCGTGGGCTATGTCTGCCCTGAGGACAAATATCATT
TGAGGTGCAACCGTCATCACCAATCTCCTATCAGCAATCCCATACATCGGTAATACCTTAGTCGAATGAA
TCTGAGGTGGTTTCTCCGTAGACAAAGCAACACTAACACGCTTCTTGCCTTCCACTTCATCCTTCCCTT
CATCATCCTAGCATTAGCAATCGTCCACCTTCTATTTCTCCACGAAACAGGATCTAACAATCCTACAGGC
ATCCCATCTAATATAGATAAAAATTCCATTCCACCCCTATTACACAATCAAGGATATCCTAGGCGCCCTAC
TACTAATCCTAACCCTACTAATACTAACCCTATTTGCACCTGACCTGCTTGGAGACCCCGATAACTACAC
CCCAGCAAACCCCTCAGCACCCACGACATATTAACCAGAATGATATTTCTATTTCGCATACGCAATC
CTACGATCAATTCCTAACAATTAGGCGGAGTCCTAGCCCTATTACTCTCAATCCTAATCCTAGCCTTCA
TCCCAATACTCCATACATCCAAACAACGAAGCATAATATTCCGACCTTCAGCCAATTCCTGTTCTGAGT
ATTAGTCGCAGATCTACTAACCTAACATGAATCGGCGGCCAACCCGTGGAACACCCCTATGTAATCGTA
GGCCAATTCGCATCCATCCTTTACTTTCTCCTAATTCTAGTTTTAATACCAACGGCTAGTCTTATCGAAA
ACAAACTCATAAAATGAAGA
```

Artiodactyles

Ovis aries mitochondrion, complete genome. NC_001941

>NC_001941.1:14159-15298 Ovis aries CYTB

```
ATGATCAACATCCGAAAAACCCACCCACTAATAAAAATTGTAAACAACGCATTTCATTGATCTCCAGCTC
CATCAAATATTTTCATCATGATGAAACTTTGGCTCTCTCCTAGGCATTTGCTTAATTTTACAGATTCTAAC
AGGCCTATTCTAGCAATACACTATACACCTGACACAACAACAGCATTCTCCTCTGTAACCCACATTTGC
CGAGACGTAAACTATGGCTGAATTATCCGATATATACACGCAAACGGGGCATCAATATTTTTTATCTGCC
TATTTATGCATGTAGGACGAGGCCTATACTATGGATCATATACCTTCCTAGAAACATGAAACATCGGAGT
AATCCTCCTATTTGCGACAATAGCCACAGCATTTCATAGGCTATGTTTTACCATGAGGACAAATATCATT
TGAGGAGCAACAGTTATTACCAACCTCCTTTCAGCAATTCATATATTGGCACAAACCTAGTCGAATGAA
TCTGGGGAGGATTCTCAGTAGACAAAGCTACCCTACCCGATTTTTTCGCCTTTCACTTTATTTTCCCATT
CATCATCGCAGCCCTCGCCATAGTTCACCTACTCTTCCTCCACGAAACAGGATCCAACAACCCACAGGA
ATTCCATCGGACACAGATAAAATTCCCTTCCACCCTATTACACCATTAAGACATCCTAGGTGCTATCC
TACTAATCCTCATCCTCATGCTACTAGTACTATTACGCCTGACTTACTCGGAGACCCAGACAACACTACAC
CCCAGCAAACCCACTTAACACTCCCCCTCACATCAAACCTGAATGATACTTCCTATTTGCGTACGCAATC
TTACGATCAATCCCTAATAAACTAGGAGGAGTCCTCGCCCTAATCCTCTCAATCCTAGTCCTAGTAATTA
TACCCCTCCTCCATACATCAAAGCAACGGAGCATAATATTCCGACCAATCAGTCAATGTATATTCTGAAT
```

CCTAGTAGCCGACCTATTAACACTCACATGAATTGGAGGCCAGCCAGTTGAACACCCCTACATCATTATT
GGACAAC TAGCATCTATTATATATTTCTTATCATTCTAGTCATAATACCAGTAGCTAGCATCATCGAAA
ACAACCTCCTAAAATGAAGA

Cervus elaphus songaricus mitochondrion, complete genome. NC_014703

>NC_014703.1:14146-15285 *Cervus elaphus songaricus* CYTB

ATGATCAATACCCGAAAAACCCACCCATTAATAAAAATTGTAAACAACGCATTTATTGACCTCCAGCCC
CATCGAATATTTATCCTGATGAAATTTGGGCTCCTTACTAGGAATTTGTCTAATTCTACAAATTCTCAC
AGGCCTATTCTAGCAATACACTATACATCCGATACAATAACAGCATTTTCCTCTGTCACCCATATCTGT
CGAGACGTCAATTATGGTTGAATTATTCGATACATACACGCAAACGGGGCATCAATATTTTTCATCTGCC
TATTTATACATGTAGGACGAGGCCTATACTACGGGTCATATACTTTTATAGAGACATGAAACATCGGAGT
AATCCTCCTATTTACAGTTATAGCCACAGCATTCGTAGGATATGTCCTACCATGAGGACAAATATCATTC
TGAGGAGCAACGGTCATTACCAACCTTCTCTCAGCAATTCCATACATTGGCACAAACCTAGTCGAATGGA
TCTGAGGAGGCTTTTCAGTAGATAAAGCAACCCTAACCCGATTTTTCGCTTTCCACTTTATTCTCCCATT
TATCATCGCAGCACTCGCTATAGTACACTTACTCTTCCTTCACGAGACAGGATCCAATAACCCAACAGGA
ATCCCATCAGACGCAGACAAAATCCCCTTCATCCTTACTATACCATTAAAGATATCTTAGGCATCTTAC
TTCTAGTACTCTTCCTAATATTACTAGTATTATTCGCACCAGACCTGCTTGGAGACCCAGACAACTATAC
CCCAGCAAATCCACTCAACACACCCCCCTCACATTAAACCTGAATGATATTTCTATTTGCATACGCAATC
CTACGATCGATTCCCAACAACTAGGAGGGGTCTTAGCCCTAGTCTCATCTATCCTAATCTTGATTCTCA
TGCCTCTTCTTCACACGTCCAAACAACGCAGCATGATATTCCGACCATTGAGCCAATGCCTATTCTGAAT
CTTAGTAGCAGACCTACTAACACTCACATGAATTGGAGGACAGCCAGTTGAATACCCCTTTATTATTATT
GGACAAC TAGCATCTGTCTTATACTTTTTCATTATCCTAGTCCTTATACCAATTACCAGCACAATCGAAA
ATAACCTCCTAAAATGAAGA

Bison bison mitochondrion, complete genome. NC_012346

>NC_012346.1:14514-15653 *Bison bison* CYTB

ATGACTAACCTTCGAAAATCCCATCCACTAATAAAAATTGTAAATAACGCATTCATTGACCTTCCAGCTC
CATCAAACATTTATCATGATGAAACTTCGGCTCCCTCCTGGGAATATGCTTAACCTACAAATCCTCAC
AGGCCTATTCTAGCAATACACTACACATCCGATACAACAACAGCATTTTCCTCCGTTGCCCATATCTGC
CGAGACGTGAACTACGGCTGAATCATCCGATACATACACGCAAACGGAGCTTCAATATTCTTTATCTGCT
TATATATGCACGTAGGACGAGGCCTATATTACGGGTCTTATACCTTCCTAGAAACATGAAATATTGGAGT
AATCCTTCTACTTACAGTAATAGCCACAGCATTCATAGGATACGTACTACCATGAGGACAAATATCATTT
TGAGGGGCAACAGTCATTACCAACCTCCTATCAGCAATCCCATACATCGGCACAACTTAGTCGAATGAA
TTTGAGGTGGATTCTCAGTAGACAAAGCAACCCTACCCGATTCTTCGCTTTCCACTTTATCCTTCCATT
TATTATTATAGCAATTGCCATAGTCCACCTATTATTCCTCCACGAAACAGGCTCCAACAATCCAACAGGA
ATCTCCTCAGACATAGACAAAATTCCATTTACCCCTACTATACCATTAAAGACATCCTAGGAGCCTTAT
TACTAATTCTAGCCCTAATACTACTGGTACTATTACACCCGACCTCCTCGGAGATCCAGATAACTACAC
CCCAGCAAATCCACTCAATACACCTCCCCACATCAAACCCGAATGATACTTCTTATTTGCATACGCAATT
TTACGATCAATCCCAATAAACTAGGAGGAGTACTAGCCCTAGCCTTCTCTATCCTAATCCTTGCCCTCA

TTCCCCTACTACACACCTCTAAACAACGAAGCATAATTTTCCGACCACTCAGCCAATGCCTATTCTGAAC
TCTAGTAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTCGAACACCCATATATCATCATC
GGACAAATAGCATCTATTATATACTTCCTTCTCATCCTAGTACTAATACCAACAGCCGGCACAATTGAAA
ACAAACTACTAAAATGAAGA

Outgroup

Polyplectron bicalcaratum(PEACOCK) mitochondrion, complete genome. NC_012900

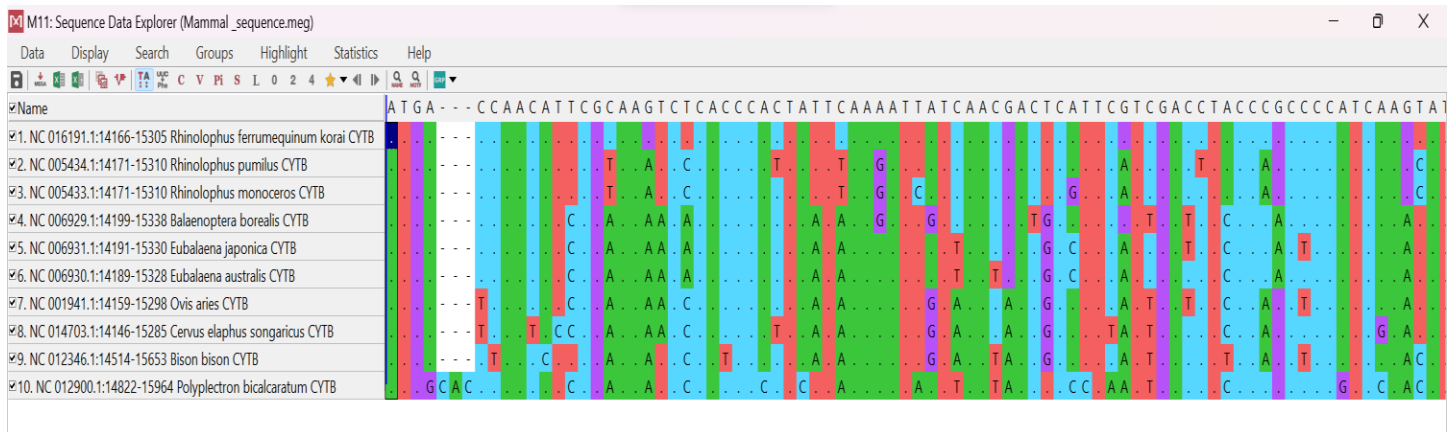
>NC_012900.1:14822-15964 Polyplectron bicalcaratum CYTB

ATGGCACCCCAACATCCGAAAATCCCACCCCCTCTAAAAATAATTAATAACTCCCTAATTGACCTCCCCG
CCCCGTCCAACATCTCCGCCTGATGGAACCTTCGGTTCCTCCTAGCAGTATGTCTTGCCACCCAAATTCT
CACCGGGCTCCTACTAGCCATACACTACACCGCAGACACCTCCCTCGCCTTCTGATCTGTAGCCCACACC
TGCCGAAACGTACAATACGGCTGACTCATCCGTAACCTCCATGCAAACGGCGCCTCATTCTTCTTCATCT
GCATTTTCCTGCACATCGGACGTGGCTTATACTACGGCTCCTACTTATATAAGGAAACCTGAAACACAGG
AGTTATCCTCCTACTCACACTCATAGCAACTGCCTTCGTAGGTTACGTACTCCCATGGGGCCAAATATCG
TTCTGAGGGGGCCACCGTCATACCAACCTATTCTCAGCAATCCCCTACATCGGACAAACCCTAGTGGAAT
GAGCCTGGGGGGGATTCTCAGTTGACAACCCAACCTTACCCGATTCTTCGCCCTACACTTCCTTCTCCC
CTTTGTAATCGCAGGAATCACTATCATTACCTCACATTCCTTCACGAATCAGGTTCAAACAACCCACTA
GGCATTTTCATCCAACCTCTGACAAAATCCCATTCCACCCATACTACTCCCTCAAAGATATCCTAGGCCTAG
CACTCATATTTATCCCATTCCTCGCACTAGCCTTATTTTCCCCCAACTTCCTAGGTGACCCAGAAAACCTT
TACCCCAGCAAATCCACTAGTGACGCCCCCACACATTAAGCCAGAATGATACTTCCTATTTGCCTACGCC
ATCCTACGCTCAATCCCAAACAAACTCGGAGGTGTACTGGCTCTAGCAGCATCAGTACTCATCCTCCTTT
TAATCCCCTTCCTCCACAAATCCAAACAACGAACCATAACATTTGCCCCACTATCCCAAATCCTGTTCTG
ACTTCTAGTAGCCAACCTCCTCGTCCTAACCTGAATTGGGAGCCAACCAGTAGAACACCCTTTCATCATC
ATTGGCCAAGTAGCATCCTTCTCCTACTTCAACATCCTATTAATTCTCTTTCCCATGGCCGGAGCATTAG
AAAACAAAATACTCAACTACTAA

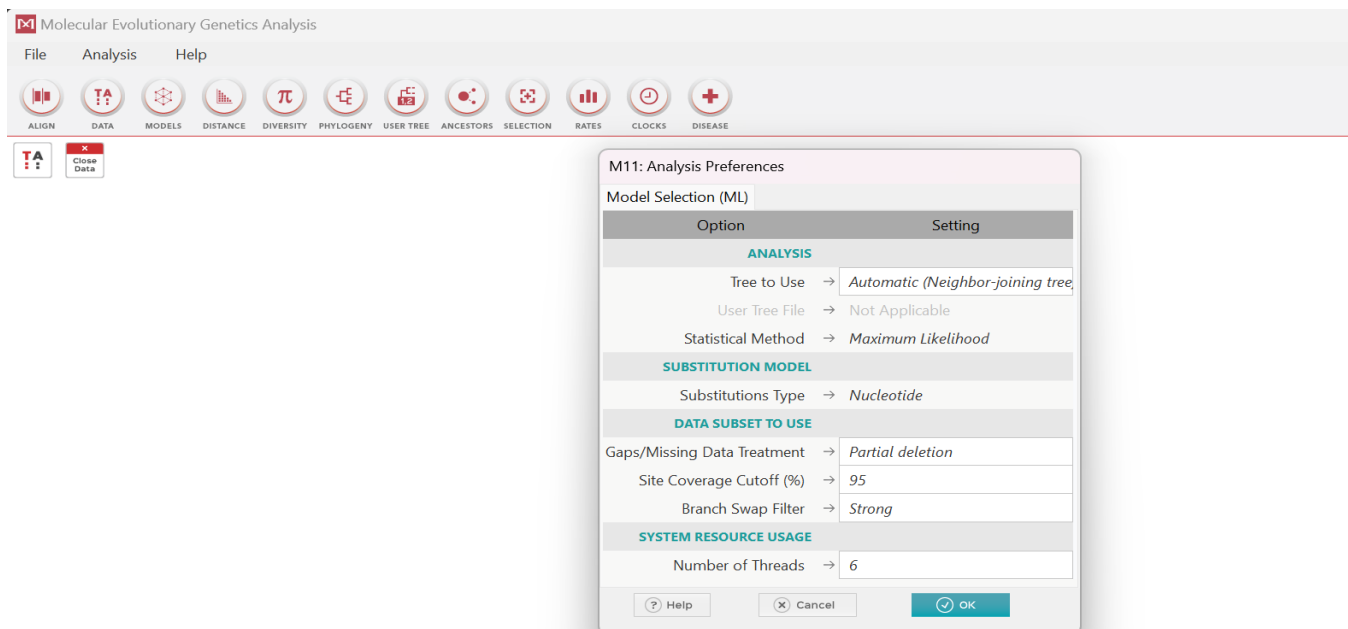
Phylogenetic Tree

The tree is constructed using MEGAX software. The statistical method used was Maximum likelihood the most common one, based on optimality criteria evaluated in terms of probability and propose different models based on the maximum likelihood value the model is selected and the tree is constructed.

Firstly Multiple sequence alignment (MSA) is done with MUSCLE as the data consists of Divergent species it is a more efficient algorithm employed for larger datasets and the results are accurate



Next is evaluating the best-fit model this can be determined by two criteria which are Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). MEGA screens for about 24 different models based on above mentioned criteria and gives scores. The model with the least value is considered to be best fit model.



The result obtained from MEGA is as follows:

Table. Maximum Likelihood fits of 24 different nucleotide substitution models

Model	Parameters		BIC	AICc	lnL	(+I)	(+G)	R	f(A)	f(T)	f(C)
	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)
	r(GT)	r(GC)									
GTR+G+I	27	11575.761		11377.677		-5661.772		0.45	1.69	4.22	0.295
	0.263	0.308	0.134	0.015	0.068	0.056	0.017	0.348	0.004	0.065	0.296
	0.000	0.124	0.007	0.000							
GTR+G	26	11585.148		11394.395		-5671.136		n/a	0.30	4.04	0.295
	0.263	0.308	0.134	0.013	0.073	0.060	0.015	0.338	0.003	0.070	0.288
	0.000	0.133	0.006	0.000							
GTR+I	26	11594.748		11403.996		-5675.936		0.50	n/a	3.54	0.295
	0.263	0.308	0.134	0.022	0.071	0.060	0.025	0.325	0.006	0.068	0.277
	0.000	0.133	0.012	0.000							
TN93+G+I	24	11624.364		11448.276		-5700.085		0.46	1.61	4.40	
	0.295	0.263	0.308	0.134	0.026	0.030	0.056	0.029	0.337	0.013	0.029
	0.287	0.013	0.124	0.026	0.030						
TN93+G	23	11635.346		11466.591		-5710.247		n/a	0.27	4.22	0.295
	0.263	0.308	0.134	0.026	0.031	0.061	0.030	0.326	0.013	0.030	0.278
	0.013	0.135	0.026	0.031							
TN93+I	23	11646.825		11478.070		-5715.987		0.50	n/a	3.47	0.295
	0.263	0.308	0.134	0.031	0.036	0.060	0.034	0.310	0.016	0.034	0.264
	0.016	0.132	0.031	0.036							
HKY+G+I	23	11654.587		11485.833		-5719.868		0.46	1.57	3.62	0.295
	0.263	0.308	0.134	0.028	0.033	0.105	0.032	0.242	0.014	0.032	0.207
	0.014	0.232	0.028	0.033							
HKY+G	22	11656.222		11494.801		-5725.356		n/a	0.27	3.72	0.295
	0.263	0.308	0.134	0.027	0.032	0.106	0.031	0.244	0.014	0.031	0.208
	0.014	0.234	0.027	0.032							
HKY+I	22	11677.010		11515.589		-5735.750		0.51	n/a	2.92	0.295
	0.263	0.308	0.134	0.033	0.039	0.100	0.037	0.231	0.017	0.037	0.196
	0.017	0.221	0.033	0.039							
T92+G+I	21	12016.164		11862.077		-5909.998		0.42	2.21	2.33	0.279
	0.279	0.221	0.221	0.041	0.033	0.155	0.041	0.155	0.033	0.041	0.196
	0.033	0.196	0.041	0.033							
T92+G	20	12016.896		11870.142		-5915.034		n/a	0.39	2.51	0.279
	0.279	0.221	0.221	0.039	0.031	0.159	0.039	0.159	0.031	0.039	0.200
	0.031	0.200	0.039	0.031							
T92+I	20	12019.525		11872.772		-5916.349		0.48	n/a	2.11	0.279
	0.279	0.221	0.221	0.044	0.035	0.151	0.044	0.151	0.035	0.044	0.190
	0.035	0.190	0.044	0.035							

K2+G+I	20	12041.319	11894.566	-5927.246	0.43	2.31	2.40	0.250
	0.250	0.250	0.250	0.037	0.037	0.176	0.037	0.176
	0.037	0.176	0.037	0.037				
K2+G	19	12042.716	11903.297	-5932.615	n/a	0.38	2.57	0.250
	0.250	0.250	0.250	0.035	0.035	0.180	0.035	0.180
	0.035	0.180	0.035	0.035				
K2+I	19	12044.145	11904.726	-5933.330	0.49	n/a	2.17	0.250
	0.250	0.250	0.250	0.039	0.039	0.171	0.039	0.171
	0.039	0.171	0.039	0.039				
GTR	25	12095.638	11912.218	-5931.052	n/a	n/a	1.81	0.295
	0.263	0.308	0.134	0.030	0.116	0.071	0.034	0.241
	0.007	0.156	0.008	0.017				0.004
TN93	22	12238.571	12077.149	-6016.530	n/a	n/a	1.79	0.295
	0.263	0.308	0.134	0.048	0.056	0.066	0.054	0.230
	0.024	0.146	0.048	0.056				0.024
HKY	21	12249.052	12094.964	-6026.442	n/a	n/a	1.77	0.295
	0.263	0.308	0.134	0.047	0.055	0.086	0.053	0.198
	0.024	0.190	0.047	0.055				0.024
JC+I	18	12430.966	12298.881	-6131.410	0.47	n/a	0.50	0.250
	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083
	0.083	0.083	0.083	0.083				
T92	19	12436.883	12297.464	-6129.698	n/a	n/a	1.73	0.279
	0.279	0.221	0.221	0.051	0.040	0.141	0.051	0.141
	0.040	0.178	0.051	0.040				0.040
JC+G+I	19	12439.958	12300.539	-6131.236	0.46	18.41	0.50	0.250
	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083
	0.083	0.083	0.083	0.083				
JC+G	18	12460.415	12328.331	-6146.135	n/a	0.50	0.50	0.250
	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083
	0.083	0.083	0.083	0.083				
K2	18	12479.593	12347.509	-6155.724	n/a	n/a	1.74	0.250
	0.250	0.250	0.250	0.046	0.046	0.159	0.046	0.159
	0.046	0.159	0.046	0.046				
JC	17	12814.798	12690.049	-6327.998	n/a	n/a	0.50	0.250
	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083
	0.083	0.083	0.083	0.083				

NOTE.-- Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the substitution pattern the best. For each model, AICc value (Akaike Information Criterion, corrected), Maximum Likelihood value (lnL), and the number of parameters (including branch lengths) are also presented [1]. Non-uniformity of evolutionary rates among sites may be modeled by using a discrete Gamma distribution (+G) with 5 rate categories and by assuming that a certain fraction of sites are evolutionarily invariable (+I). Whenever applicable, estimates of gamma shape parameter and/or the estimated fraction of invariant sites are shown. Assumed or estimated values of transition/transversion bias (R) are shown for each model, as well. They are followed by nucleotide frequencies (f) and rates of base substitutions (r) for each nucleotide

pair. Relative values of instantaneous r should be considered when evaluating them. For simplicity, sum of r values is made equal to 1 for each model. For estimating ML values, a tree topology was automatically computed. This analysis involved 10 nucleotide sequences. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). There were a total of 1140 positions in the final dataset. Evolutionary analyses were conducted in MEGA11 [2]

Abbreviations: TR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93: Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor./div>

1. Nei M. and Kumar S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

2. Tamura K., Stecher G., and Kumar S. (2021). MEGA 11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution <https://doi.org/10.1093/molbev/msab120>.

Disclaimer: Although utmost care has been taken to ensure the correctness of the caption, the caption text is provided "as is" without any warranty of any kind. Authors advise the user to carefully check the caption before its use for any purpose and report any errors or problems to the authors immediately (www.megasoftware.net). In no event shall the authors and their employers be liable for any damages, including but not limited to special, consequential, or other damages. Authors specifically disclaim all other warranties expressed or implied, including but not limited to the determination of the suitability of this caption text for a specific purpose, use, or application.

Model	Parameters		BIC	AICc	lnL	(+I)	(+G)	R	f(A)	f(T)	f(C)
	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)
	r(GT)	r(GC)									
GTR+G+I 27	11575.761		11377.677		-5661.772	0.45	1.69	4.22	0.295		
	0.263	0.308	0.134	0.015	0.068	0.056	0.017	0.348	0.004	0.065	0.296
	0.000	0.124	0.007	0.000							

From the results, the best-fit model is identified as **GTR + G + I(General Time reversible),(Gamma Distribution) and(Invariant sites)**,with the least value.

Next is constructing a Maximum likelihood phylogenetic tree with GTR+G+I model.

M11: Analysis Preferences

Phylogeny Reconstruction

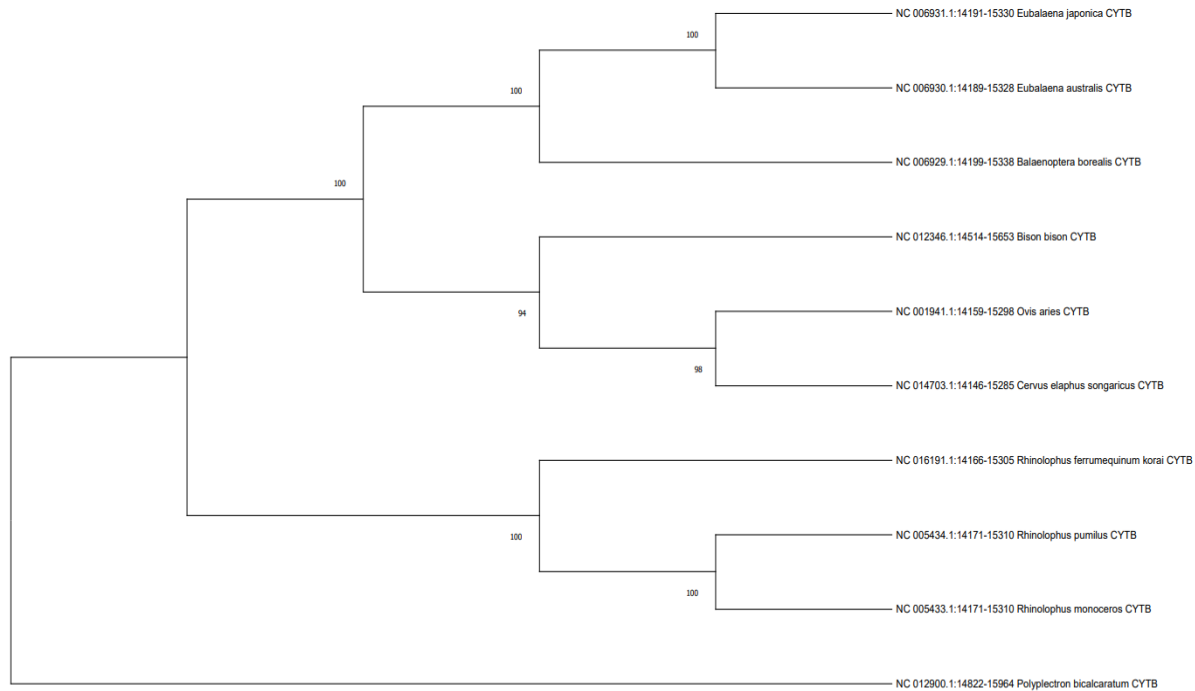
Option	Setting
ANALYSIS	
Statistical Method	→ Maximum Likelihood
PHYLOGENY TEST	
Test of Phylogeny	→ Bootstrap method
No. of Bootstrap Replications	→ 50
SUBSTITUTION MODEL	
Substitutions Type	→ Nucleotide
Model/Method	→ General Time Reversible model
RATES AND PATTERNS	
Rates among Sites	→ Gamma Distributed With Invariant Sites (G+I)
No of Discrete Gamma Categories	→ 5
DATA SUBSET TO USE	
Gaps/Missing Data Treatment	→ Partial deletion
Site Coverage Cutoff (%)	→ 95
TREE INFERENCE OPTIONS	
ML Heuristic Method	→ Nearest-Neighbor-Interchange (NNI)
Initial Tree for ML	→ Make initial tree automatically (Default - NJ/BioNJ)
Initial Tree File	→ Not Applicable
Branch Swap Filter	→ Strong
SYSTEM RESOURCE USAGE	
Number of Threads	→ 6

The above-mentioned image gives details of the parameters used for tree construction. Here the **Bootstrap iteration** was set to **50 values** so that many times based on topology it identifies how many times a particular species occurred in that place and placed it there.

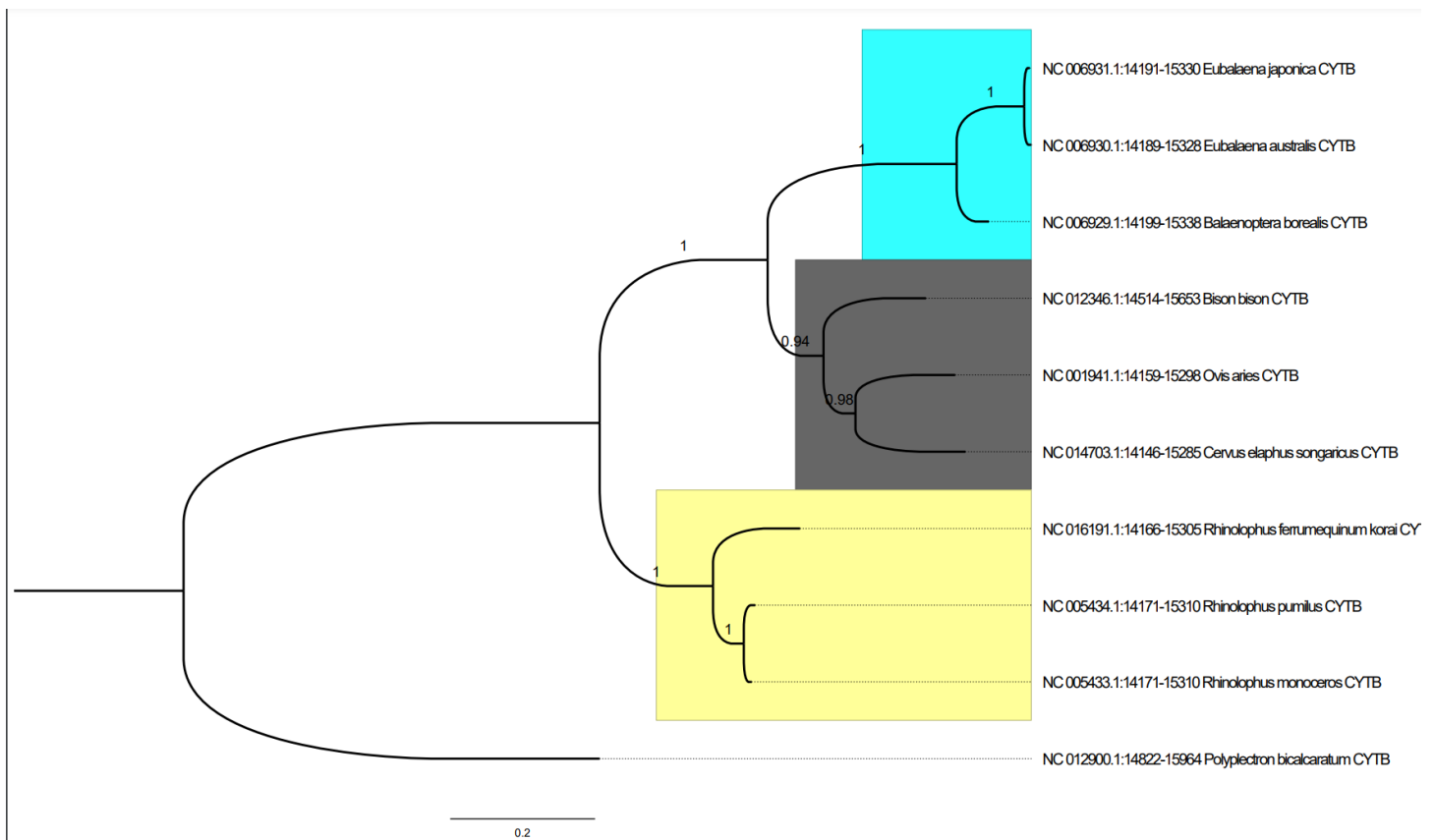
After The results are obtained the tree generated from MEGA can be used as it is or can be **rooted** and modified better with **Fig Tree software**.

The **Newick format** of the tree consists of information like **Branch length** and **Bootstrap value**. Which can be used for further processing.

```
((('NC 006931.1:14191-15330 Eubalaena japonica CYTB':0.007654,'NC 006930.1:14189-15328 Eubalaena australis CYTB':0.009698)[&Bootstrap=1.0]:0.093777,'NC 006929.1:14199-15338 Balaenoptera borealis CYTB':0.044637)[&Bootstrap=1.0,!highlight={3,0.0,#33ffff}]:0.262373,('NC 012346.1:14514-15653 Bison bison CYTB':0.142176,('NC 001941.1:14159-15298 Ovis aries CYTB':0.13845,'NC 014703.1:14146-15285 Cervus elaphus songaricus CYTB':0.152606)[&Bootstrap=0.98]:0.04433)[&Bootstrap=0.94,!highlight={3,0.091307089999999992,#666666}]:0.077604)[&Bootstrap=1.0]:0.233116,('NC 016191.1:14166-15305 Rhinolophus ferrumequinum korai CYTB':0.120885,('NC 005434.1:14171-15310 Rhinolophus pumilus CYTB':0.015934,'NC 005433.1:14171-15310 Rhinolophus monoceros CYTB':0.011068)[&Bootstrap=1.0]:0.042676)[&Bootstrap=1.0,!highlight={3,0.320731430000000004,#ffff99}]:0.157346,'NC 012900.1:14822-15964 Polyplectron bicalcaratum CYTB':1.154995)
```



The image of the tree obtained from MEGA software.



The image of the tree obtained from Fig Tree after rooting and re-orienting the tree.

Evolutionary Relationships and Events in the Phylogenetic Tree:

The tree clearly shows three distinct monophyletic clades – Bats, Whales, and Artiodactyls. Each clade shares unique characteristics due to independent evolution after diverging from a common ancestor. "Bootstrap" values indicate branch support, implying high confidence in the major branching. but as the bootstrap iteration was done only 50 times which is minimum, this tree is of low confidence as 500 to 1000 iterations are considered to give confident results. while evaluating the best-fit model here, we got GTR+G+I (General Time Reversible, Gamma Distribution, Invariant sites) indicating this has both highly variable and highly conserved regions.

The tree shows an early split between bats and the clade containing whales and artiodactyls. This divergence is estimated to have occurred long ago (represented by the longest branch). The grouping of whales with Artiodactyls (including hippos, cows, etc.) is supported by bootstrap values and branch lengths. This clade, suggests a shared terrestrial ancestor before a lineage transitioned to an aquatic lifestyle in whales. Within the whale-artiodactyl clade, whales (represented by Eubalaena, and Balaenoptera) diverged from artiodactyls (represented by Bison, Ovis, and Cervus) around long ago. This suggests a major shift towards aquatic adaptation in whales, distinct from the terrestrial herbivore ZZwhales (dolphins). This transition likely involved adaptations for filter feeding and echolocation in baleen and toothed whales, respectively. Artiodactyls exhibit further diversification within their group, with the bovine lineage (Bison) splitting from the deer lineage (Ovis, Cervus). the bison and sheep cluster together, while the deer diverges slightly earlier. This suggests a relatively more recent common ancestor for bison and sheep compared to deer.

Within artiodactyls, bison diverged first, followed by sheep and deer. Among bats, the horseshoe bats (Rhinolophus species) diverged earlier than the red-whiskered bulbul. These branching events indicate independent evolutionary trajectories within each group.

The bootstrap values indicate the support for each branching point in the tree. Higher values suggest more confidence in the inferred relationship.

The presence of highlighted branches and colours likely signifies specific groups or features of interest in the original tree.

Comparing and Contrasting Evolutionary Histories:

The most striking contrast is the shift in habitat transitions from land to water in whales. Though they are warm blooded like other land mammals differing from sea creatures like sharks, fishes. They involved significant morphological and physiological adaptations like streamlined bodies, echolocation, and flippers. We can also infer Rates of Evolution, The tree implies different evolutionary rates within and between groups. Whales might have undergone rapid adaptations for aquatic life, while artiodactyls show diversification across various terrestrial niches

Artiodactyls show Dietary Specializations as they are diversified with herbivory in various forms (grazers, ruminants), while bats developed diverse feeding strategies like insectivore, frugivory, and echolocation-based hunting.

The mammals here show a lot of Morphological Diversity. Each group shows unique morphologies adapted to their respective niches. Whales have fins for swimming, and artiodactyls have hooves for grazing. Bats exhibit a unique evolutionary trajectory marked by the development of echolocation and flight adaptations. Their early divergence from other mammals suggests the independent evolution of these features. Both whales and bats exhibit echolocation that acts as a common ground between them, but Whales underwent a remarkable transformation from land-based mammals to fully aquatic creatures. This involved changes in limbs, body shape, breathing, and sensory organs. Their close relationship with artiodactyls highlights a shared ancestry, even though their lifestyles are vastly different. In conclusion, analyzing this phylogenetic tree provides valuable insights into the evolutionary history of bats, whales, and artiodactyls. It highlights their unique adaptations, diversification patterns, and their contribution to our understanding of broader mammalian evolution.

Implications for Understanding Mammalian Evolution:

The tree reveals the importance of major evolutionary events, such as the whale-artiodactyl split and the development of flight in bats. It highlights the diverse adaptations within and between mammalian lineages, showcasing the remarkable evolutionary potential of this group. Convergent and Divergent Evolution: The tree highlights both convergent evolution among whales and fish with streamlined bodies and divergent evolution in bats and artiodactyl with diverse feeding strategies.

By understanding these relationships, we can better reconstruct the history of mammals and gain insights into their ecological and morphological diversity. The clade's diversification into various ecological niches showcases the power of natural selection in driving adaptive radiation. While the tree gives relative divergence times, fossil evidence is crucial to pinpoint absolute dates and intermediate ancestral forms. Studying ancestral relationships helps reconstruct past ecosystems and understand how species interactions have shaped present-day biodiversity.