

Evolution of Body Size in Bears



How to Build and Use a Phylogeny

Lesson I

Building a Phylogeny

Objectives for Lesson I:

1. Overview of concepts

1. Observe TA complete a simple walk through for creating a phylogeny.

1. Complete a handout giving specific instructions on how to create a phylogeny.

**Students will use the following example:

- Reconstructing phylogenetic position of the extinct dodo and solitaire, based on Shapiro *et al.* (2002) *Science*

What is a phylogenetic tree?

- A history of evolution
- Often uses characteristics of modern taxa to reconstruct evolutionary history
- Depicts the evolutionary relationships between organisms.

How do we make them?

Phylogenies are built using *characters*. Commonly used characters include:

Morphology:

- Winged/wingless
- Number of limbs
- Type of lung

Genetic:

- Mitochondrial:
 - Evolves rapidly
 - Matrilinear inheritance
 - No recombination
- Nuclear:
 - Evolves more slowly
 - Sexual reproduction
 - Recombination
- RNA
- Neutral vs. coding

How do we make them?

More about *characters*:

- A character that has evolved to a different **state** is an **apomorphy**
- An apomorphy found in multiple taxa is a **synapomorphy**
 - may be **homologous** – inherited from a common ancestor
 - or may be **analogous** – evolved twice or more by convergent evolution
- Taxa are grouped by synapomorphic characters

One method...

Parsimony: the relationship requiring the fewest evolutionary changes is most likely to be correct

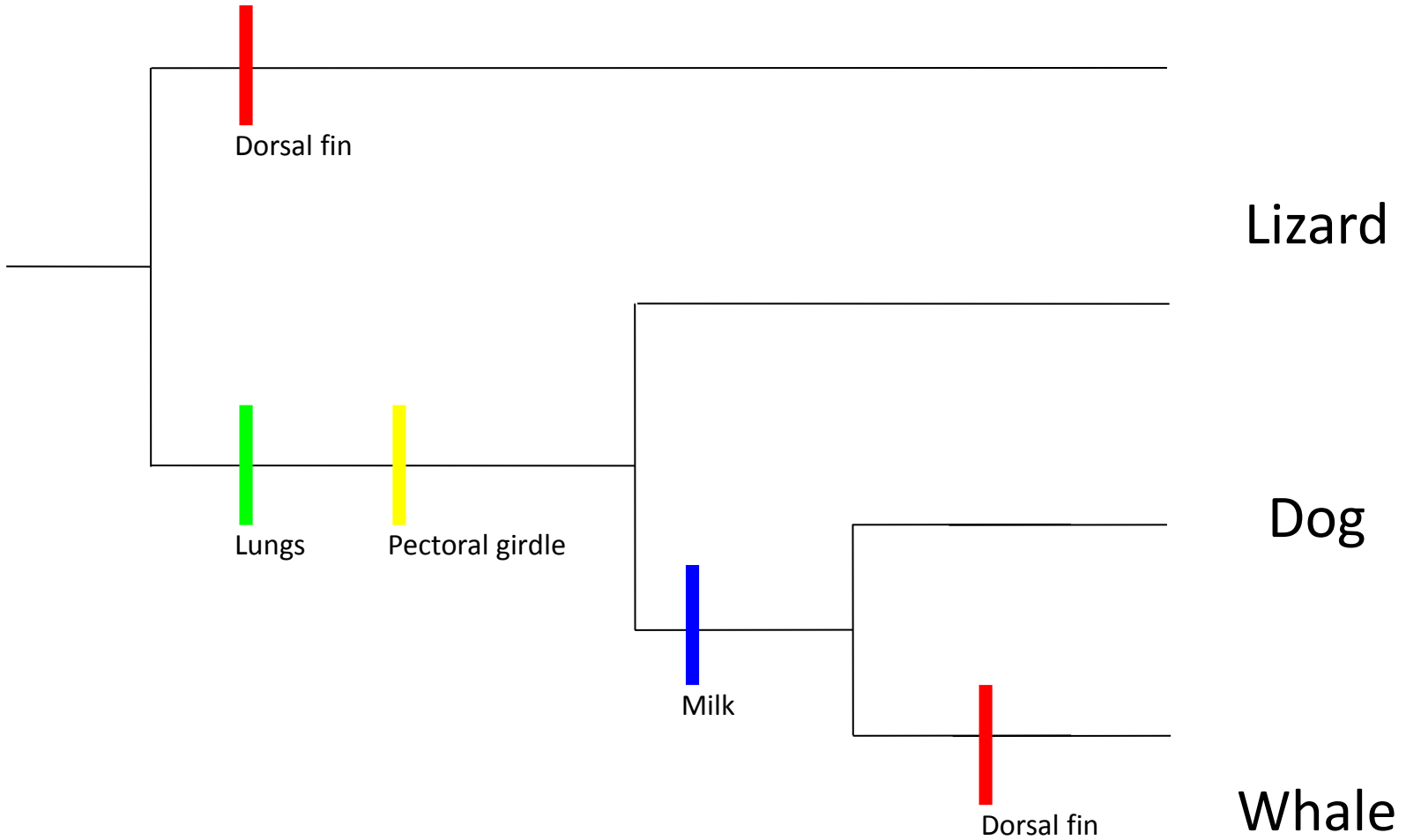
Construct data matrix of apomorphies:

	Shark	Lizard	Dog	Whale
Dorsal fin	1	0	0	1
Pectoral girdle	0	1	1	1
Lungs	0	1	1	1
Milk	0	0	1	1

Count synapomorphies between each taxon:

	Shark	Lizard	Dog	Whale
Shark	-	0	0	1
Lizard	-	-	2	2
Dog	-	-	-	3
Whale	-	-	-	-

Find the tree requiring the fewest evolutionary changes. ^{Shark}



Other methods...

When we have a lot of taxa and characters, the number of potential trees soon reaches millions.

Statistical methods are useful in these cases:

Maximum likelihood: uses an algorithm to evaluate all possible trees and find the one that best explains the observed relationships

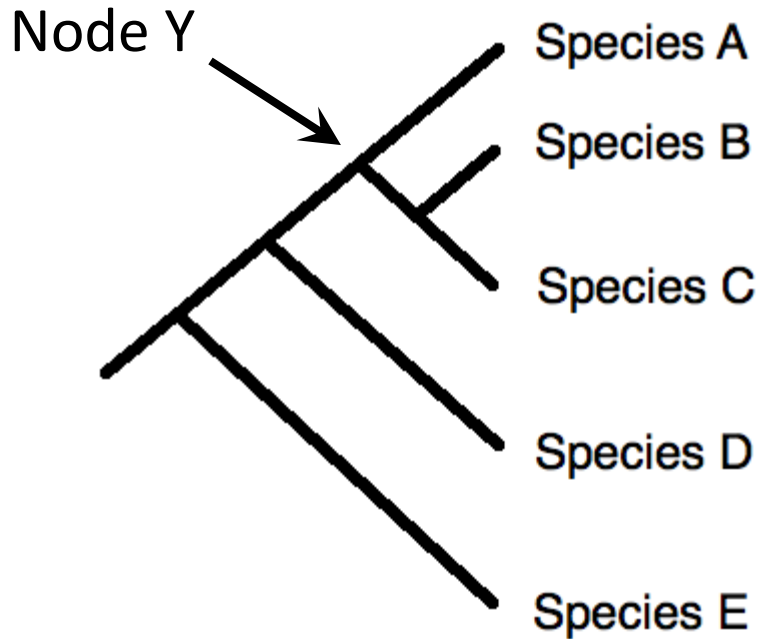
- Incorporates models of how DNA evolves

Other methods...

Other statistical methods include:

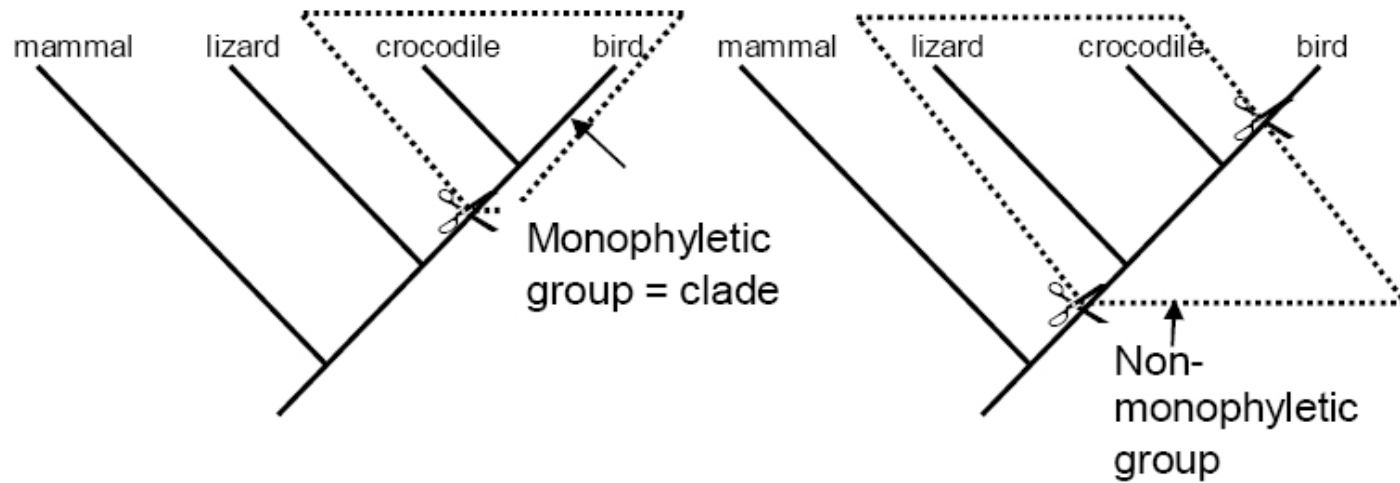
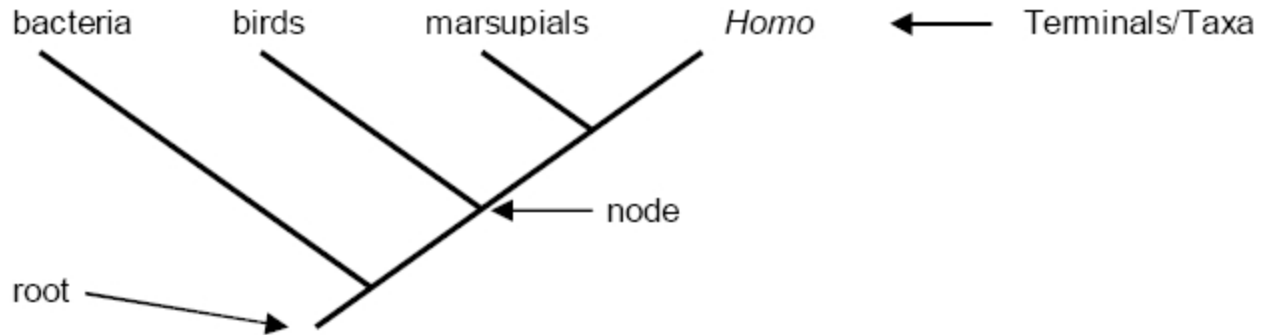
- **Distance:** calculate the amount of evolutionary change between each pair of taxa and then find the tree that best predicts that pattern.
- **Bayesian inference:** uses probability theory: how probable is each possible tree, considering:
 - .Observed data
 - .Models of evolutionary process
 - .Prior beliefs

Terms:



- Species B and C are *sister* taxa
- Species E is an *outgroup* to ABCD
- Node Y is the *ancestor* to A, B & C

Reading a Phylogenetic Tree



home browse help features learning contribute about Search

TREE OF LIFE web project

Explore the Tree of Life

Browse the Site

- [Root of the Tree](#)
- [Popular Pages](#)
- [Sample Pages](#)
- [Recent Additions](#)
- [Random Page](#)
- [Treehouses](#)
- [Images, Movies,...](#)

Search

News

[Darwin 200: the celebration continues...](#)

[read more](#)

Learn about ...

Agaricales
(a group of fungi)

[image info](#)

The Agaricales, or euagarics clade, is a monophyletic group of approximately 8500 mushroom species...

[read more](#)

[more featured pages](#)

The Tree of Life Web Project (ToL) is a collaborative effort of [biologists](#) and [nature enthusiasts](#) from around the world. On more than 10,000 World Wide Web pages, the project provides information about biodiversity, the characteristics of different groups of organisms, and their evolutionary history ([phylogeny](#)).

Each page contains information about a particular group, e.g., [salamanders](#), [segmented worms](#), [phlox flowers](#), [tyrannosaurs](#), [euglenids](#), [Heliconius butterflies](#), [club fungi](#), or the [vampire squid](#). ToL pages are linked one to another hierarchically, in the form of the evolutionary tree of life. Starting with the [root of all Life on Earth](#) and moving out along diverging branches to individual species, the [structure of the ToL project](#) thus illustrates the genetic connections between all living things.

[read more about the Tree of Life Web Project...](#)

"The affinities of all the beings of the same class have sometimes been represented by a great tree... As buds give rise by growth to fresh buds, and these if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever branching and beautiful ramifications."

Charles Darwin, 1859

Reference for relationships of all organisms can be found at Tree of Life Web Project.

Tolweb.org

Building a bear phylogeny:

BMC Evolutionary Biology



Research article

Open Access

Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary

Johannes Krause¹, Tina Unger¹, Aline Noçon¹, Anna-Sapfo Malaspinas²,
Sergios-Orestis Kolokotronis^{3,4}, Mathias Stiller¹, Leopoldo Soibelzon⁵,
Helen Spriggs⁶, Paul H Dear⁶, Adrian W Briggs¹, Sarah CE Bray⁷,
Stephen J O'Brien⁸, Gernot Rabeder⁹, Paul Matheus¹⁰, Alan Cooper⁷,
Montgomery Slatkin², Svante Pääbo¹ and Michael Hofreiter*¹

Genetic Databases:

NCBI Resources How To Sign in to NCBI

GenBank Nucleotide Search

GenBank Submit Genomes WGS HTGs EST/GSS Metagenomes TPA TSA INSDC

GenBank Overview

What is GenBank?

GenBank[®] is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*, 2013 Jan;41(D1):D36-42). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

The complete [release notes](#) for the current version of GenBank are available on the NCBI ftp site. A new release is made every two months. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An example of a GenBank [record](#) may be viewed for a *Saccharomyces cerevisiae* gene.

GenBank Resources

[GenBank Home](#)

[Submission Types](#)

[Submission Tools](#)

[Search GenBank](#)

[Update GenBank Records](#)

<http://www.ncbi.nlm.nih.gov/genbank/>

EMBL-EBI Services Research Training Industry About us

ENA European Nucleotide Archive

ENA Home Search & Browse Submit & Update About ENA Contact FAQ

- ENA Home
- Search & Browse
- Submit & Update
- About ENA
- Contact
- FAQ

NEWS AND ANNOUNCEMENTS

ENA Home »

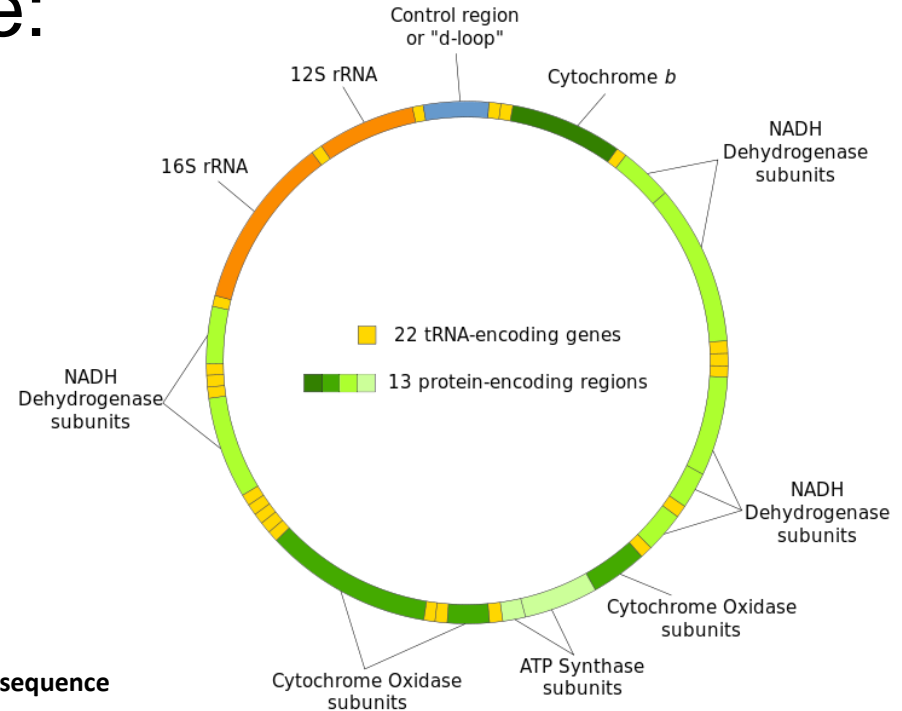
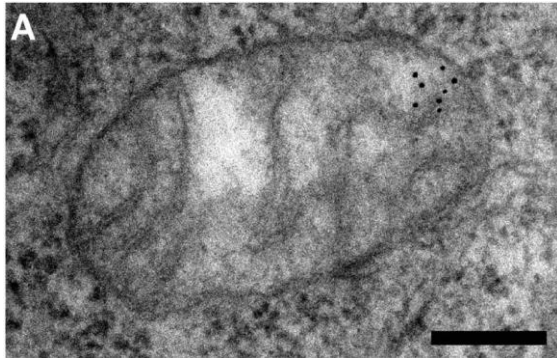
European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation ... [more](#)

Access to ENA data is provided through the browser, through search tools, large scale file download and through the API.

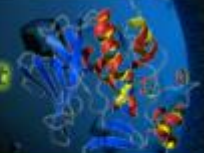
<http://www.ebi.ac.uk/ena/>

Mitochondrial DNA and cytochrome b gene:



***Ailuropoda melanoleuca* (Giant Panda) complete cytochrome b sequence**
FM177761.1

```
TGATCAACATCCGAAAACTCATCCATTAGTTAAAATTATCAACAACCTCATTGACCTTCCAACACCATCAAACATTTCAACATGATGGAACCTTGGGTCCTGTTAGGA
GTGTGTCTGATCTTGCAAACTTAACAGGCTTATTCTAGCCATACACTATACATCAGATACAGCTACAGCCTTTTCATCAGTCGCACACATTTGTCGAGACGTCAACTATGG
TTGATTTATCCGATATATACATGCCAATGGGGCCTCTATATTTTTATCTGCCTATTTATACACGTAGGGCGAGGCTTATACTATGGATCATACCTATTTCCAGAGACATGGA
ATATCGGAATTATCTCCTACTTACAGTTATAGCCACAGCATTATAGGATATGTACTACCTTGAGGACAAAATATCCTTCTGAGGAGCAACCGTCATTACTAACCTACTATCA
GCAATTCCTTACATTGGCACTAATCTAGTGGAGTGAATCTGAGGGGGTTTCTCCGTAGATAAAGCAACACTAACCCGATTTTTGCTTTTCACTTTATCCTCCATTTATCAT
CTCAGCACTAGCAATAGTCCATCTATTATTCCTTACGAAACAGGATCTAATAACCCCTCCGGAATTCATCTGACCCAGACAAAATCCCATTTTACCCCTATCATACAATTA
AAGACATCCTAGGCGTCTATTTCTGTCTCGCCTTAATAACCCCTGGCTTTATTCTCACCAGACCTGTTAGGAGACCTGATAACTATACCCCTGCAAATCCACTAAGTACC
CCGCCACATATTAAGCCTGAATGGTACTTTCTATTTGCTACGCTATCCTGCGATCTATCCCTAATAAACTAGGAGGGGTGCTAGCTCTAATCTTCTCTATTCTAATTCTAAC
TATTATTCCAATTACATACATCAAACAACGAAGCATGATATCCGACCTCTAAGTCAATGCTTATTCTGACTCCTAGTAGCAGACCTACTCACACTAACATGAATTGGAG
GACAGCCAGTAGAACACCCCTCATTATTATTGGGCAATTGGCCTCTATTCTCTACTTTACAATTCTTCTAGTACTTATACCTATCACTAGCATTATTGAGAATAGCCTCTCAA
AATGAAGA
```



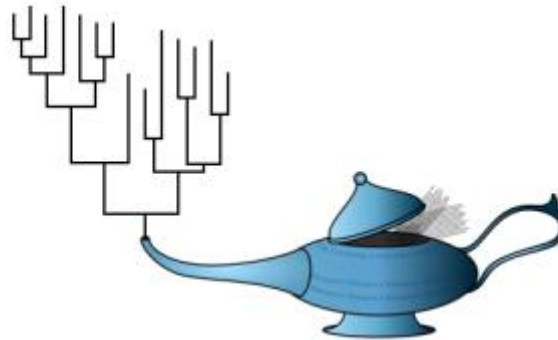
One Click

Advanced

A la Carte

Phylogeny.fr


Robust Phylogenetic Analysis For The Non-Specialist



Phylogeny.fr is a free, simple to use web service dedicated to reconstructing and analysing phylogenetic relationships between molecular sequences.

Phylogeny.fr runs and connects various bioinformatics programs to reconstruct a **robust phylogenetic tree** from a set of sequences.

If you use this site, please cite:

 Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.-F., Guindon S., Lefort V., Lescot M., Claverie J.-M., Gascuel O. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist* Nucleic Acids Research. 2008 Jul 1; 36 (Web Server Issue):W465-9. Epub 2008 Apr 19. (PubMed)

Web search www.phylogeny.fr.
In the pull down “Phylogeny Analysis” pick “One Click”

Méthodes et Algorithmes pour la Bio-informatique LIRMM

Information Genomique et Structurale

Home Phylogeny Analysis Blast Explorer Online Programs Your Workspace Documentation Downloads Contacts

PhyML 3.0 aLRT (doc + aLRT)

1. Overview 2. Data & Settings 3. Results

Datatype: auto-select protein DNA/RNA

Upload your alignment (FASTA, Phylip, Clustal, EMBL or NEXUS format) from a file:

Or paste it here (load example of alignment)

Maximum dataset size: (number of taxa) * (number of taxa) * (sequence size) = 80000000.

Settings

Statistical tests for branch support:

Approximate Likelihood-Ratio Test (aLRT):

Bootstrapping procedure: Number of bootstraps:


Substitution model:

[▶ Advanced Settings...](#)


To receive the results by e-mail, enter your address(es):

http://www.phylogeny.fr/version2 CGI/one_task.cgi?task_type=phyml

Page should be seen after “one click”



Methodes
et
Algorithmes
pour la
Bio-informatique
LIRMM



Information
Genomique et
Structurale

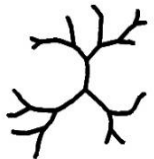
Home
Phylogeny Analysis
Blast Explorer
Online Programs
Your Workspace
Documentation
Downloads
Contacts

PhyML 3.0 aLRT (doc + aLRT)

[e-mail results](#)

1. Overview
2. Data & Settings
3. Results


Phylogeny: PhyML




PhyML has been launched! Maximum likelihood computation in progress, please wait...

Please cite:


1. Dereeper A., Audic S., Claverie J.M., Blanc G. *BLAST-EXPLORER helps you building datasets for phylogenetic analysis*. BMC Evol Biol. 2010 Jan 12;10:8. (PubMed)
2. Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.F., Guindon S., Lefort V., Lescot M., Claverie J.M., Gascuel O. *Phylogeny.fr: robust phylogenetic analysis for the non-specialist*. Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W465-9. Epub 2008 Apr 19. (PubMed) * joint first authors
3. Guindon S., Gascuel O. *A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood*. Syst Biol. 2003, Oct;52(5):696-704. (PubMed)
4. Anisimova M., Gascuel O. *Approximate likelihood ratio test for branches: A fast, accurate and powerful alternative*. Syst Biol. 2004, Aug;53(4):539-52. (PubMed)




Marseille - Nice genopole®



CNRS
Centre
National
de la
Recherche
Scientifique
département des bioinformatiques



Réseau National des Génomiques © 2003



Institut de Biologie de la Méditerranée

Loading screen should appear as sequences are analysed.

"One Click" Mode

Alignment MUSCLE → Curation Gblocks → Phylogeny PhyML → Tree Rendering TreeDyn

1. Overview 2. Data & Settings 3. Alignment 4. Curation 5. Phylogeny 6. Tree Rendering

Tree Rendering results

ENA_FM177761_FM177761.1_Ailuropoda_melanoleuca_complete_mitochon
 ENA_FM177760_FM177760.1_Ursus_spelaeus_complete_mitochondrial_ge
 gi_19343502_ref_NC_003427.1_Ursus_arctos_mitochondrion_complete
 gi_19343516_ref_NC_003428.1_Ursus_maritimus_mitochondrion_comple
 gi_19343488_ref_NC_003426.1_Ursus_americanus_mitochondrion_compl
 ENA_FM177759_FM177759.1_Ursus_thibetanus_complete_mitochondrial
 ENA_FM177765_FM177765.1_Helarctos_malayanus_complete_mitochondri
 ENA_FM177763_FM177763.1_Melursus_ursinus_complete_mitochondrial
 ENA_FM177762_FM177762.1_Arctodus_simus_complete_mitochondrial_ge
 ENA_FM177764_FM177764.1_Tremarctos_ornatus_complete_mitochondria
 gi_5834857_ref_NC_001325.1_Phoca_vitulina_mitochondrion_complete

0.3

Figure 1: Phylogenetic tree (the branch length is proportional to the number of substitutions per site).

Dynamic Tree Edition

Color leaf using color

branch and assign the group name

Add annotations using color

Reset to original tree

Reroot (outgroup)

Swap subtrees

Tree manipulation :

Reroot using mid-point rooting

Flip subtree

Change leaf name

A tree will generate in the program, click on "Tree in Newick Format" under input. Copy the code that generates.

1. Overview 2. Data & Settings 3. Alignment 4. Curation 5. Phylogeny 6. Tree Rendering

Tree Rendering results

0.2

Figure 1: Phylogenetic tree (the branch length is proportional to the number of substitutions per site).

Dynamic Tree Edition

Color leaf using color: blue branch and assign the group name: Add annotations using color: red

Reset to original tree Reroot using mid-point rooting Reroot (outgroup) Flip subtree Swap subtrees Change leaf name

Tree manipulation

Display:

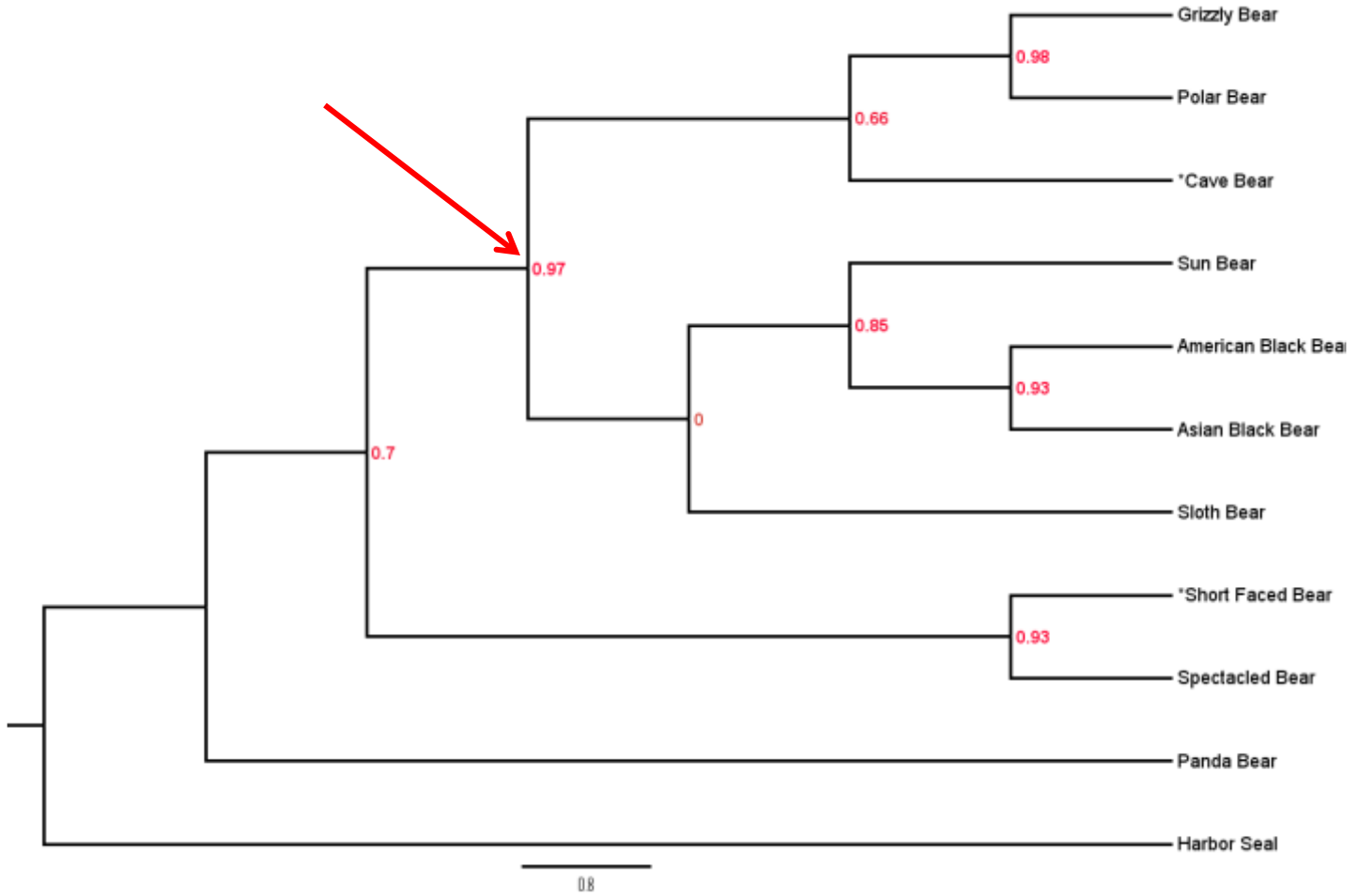
Branch annotation: Branch support values Branch length values Use color: red

Legend at position: 25 120 Update Use color: dimgray

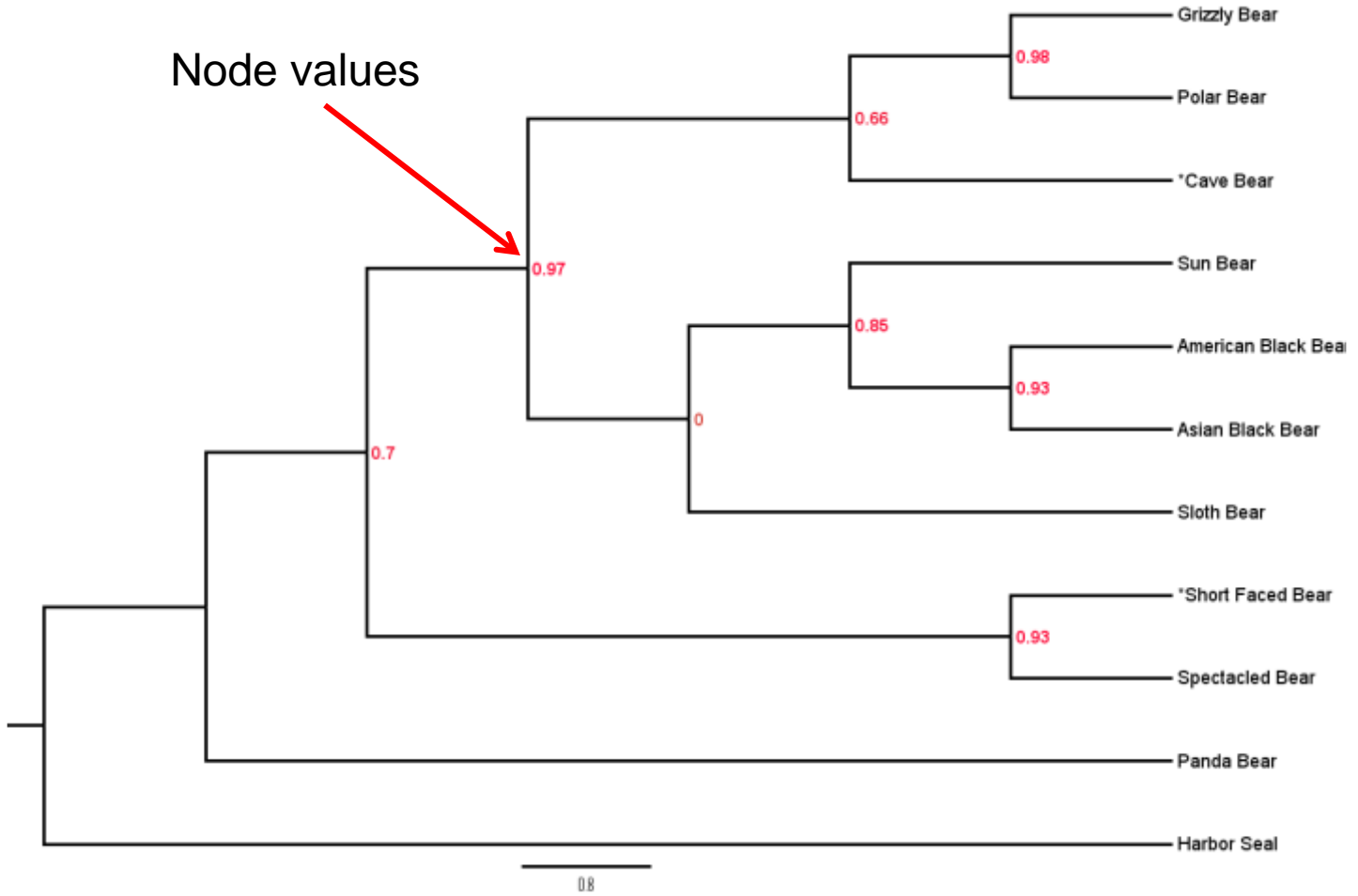
Ignore branch length

Leaves font: Times normal 10 Update

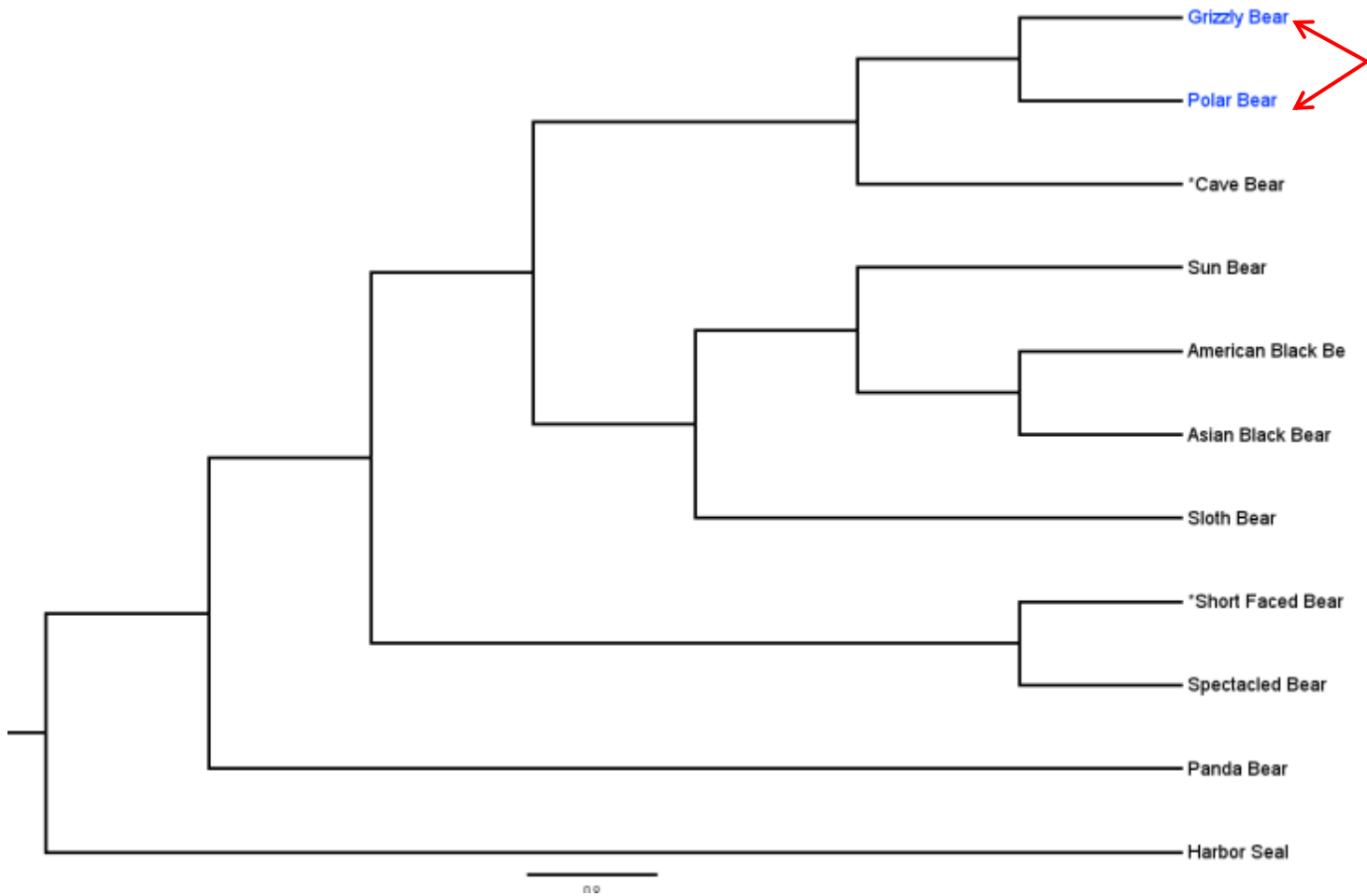
Edit the tips using the “Change leaf name” tool.



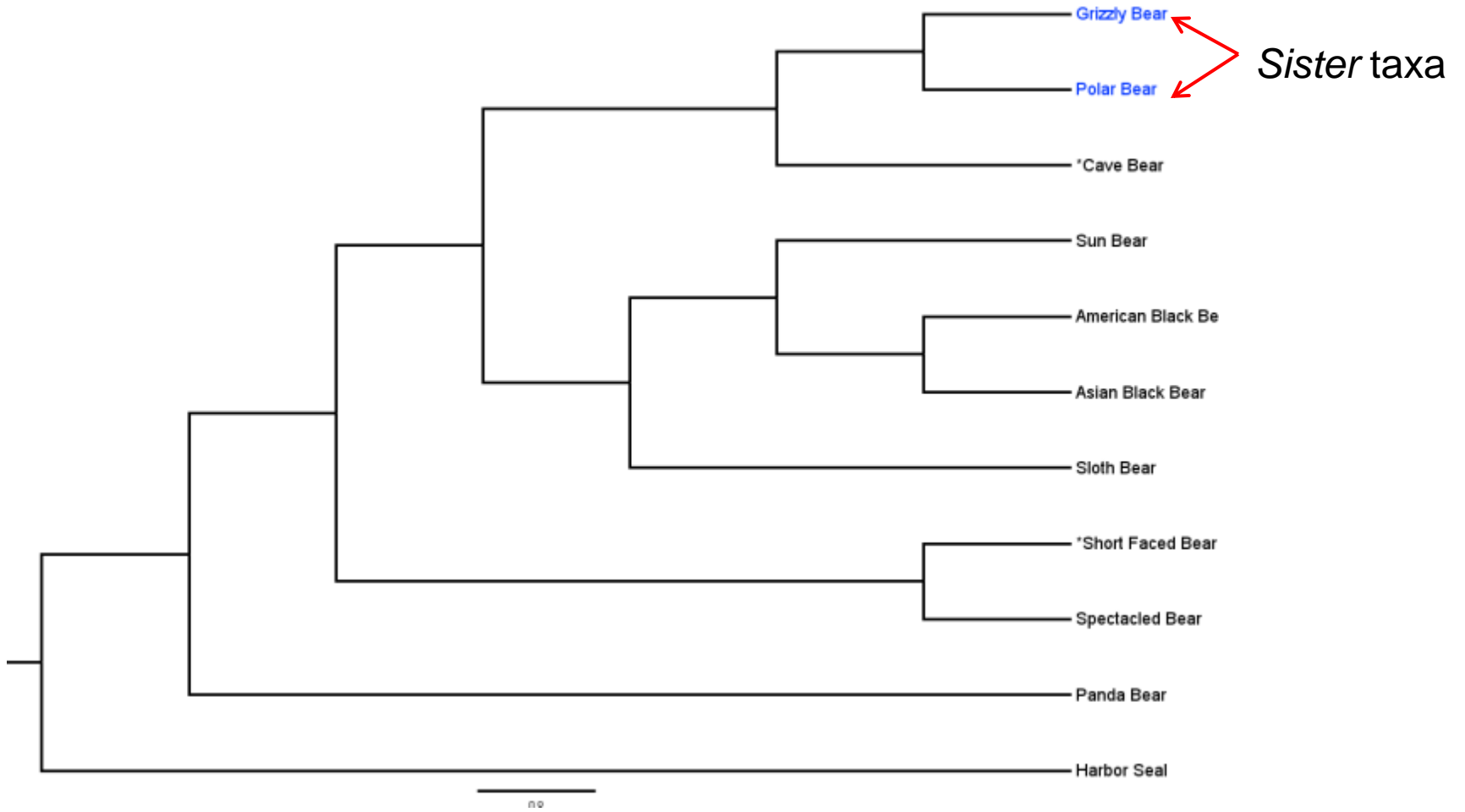
* Extinct species



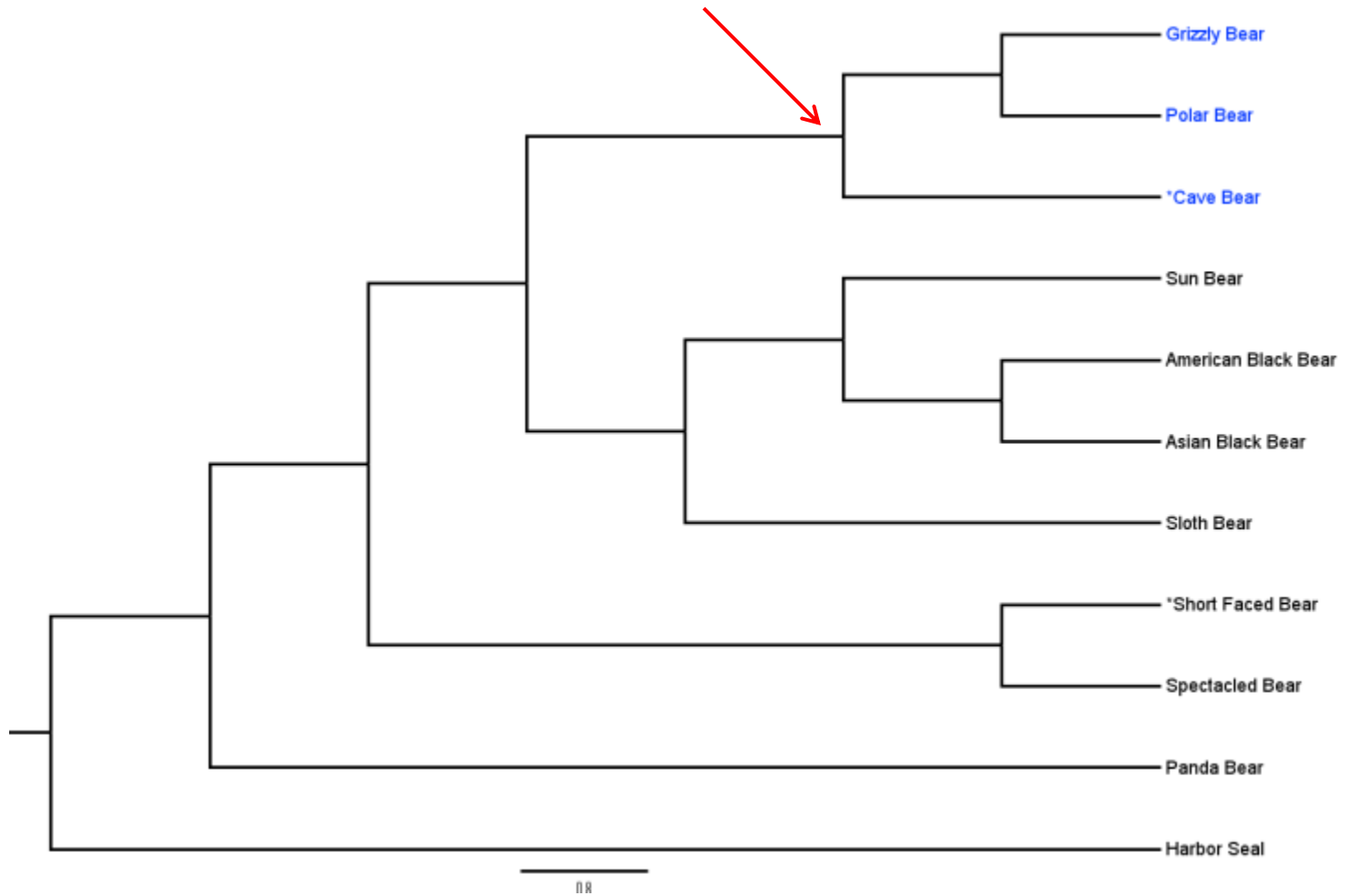
* Extinct species



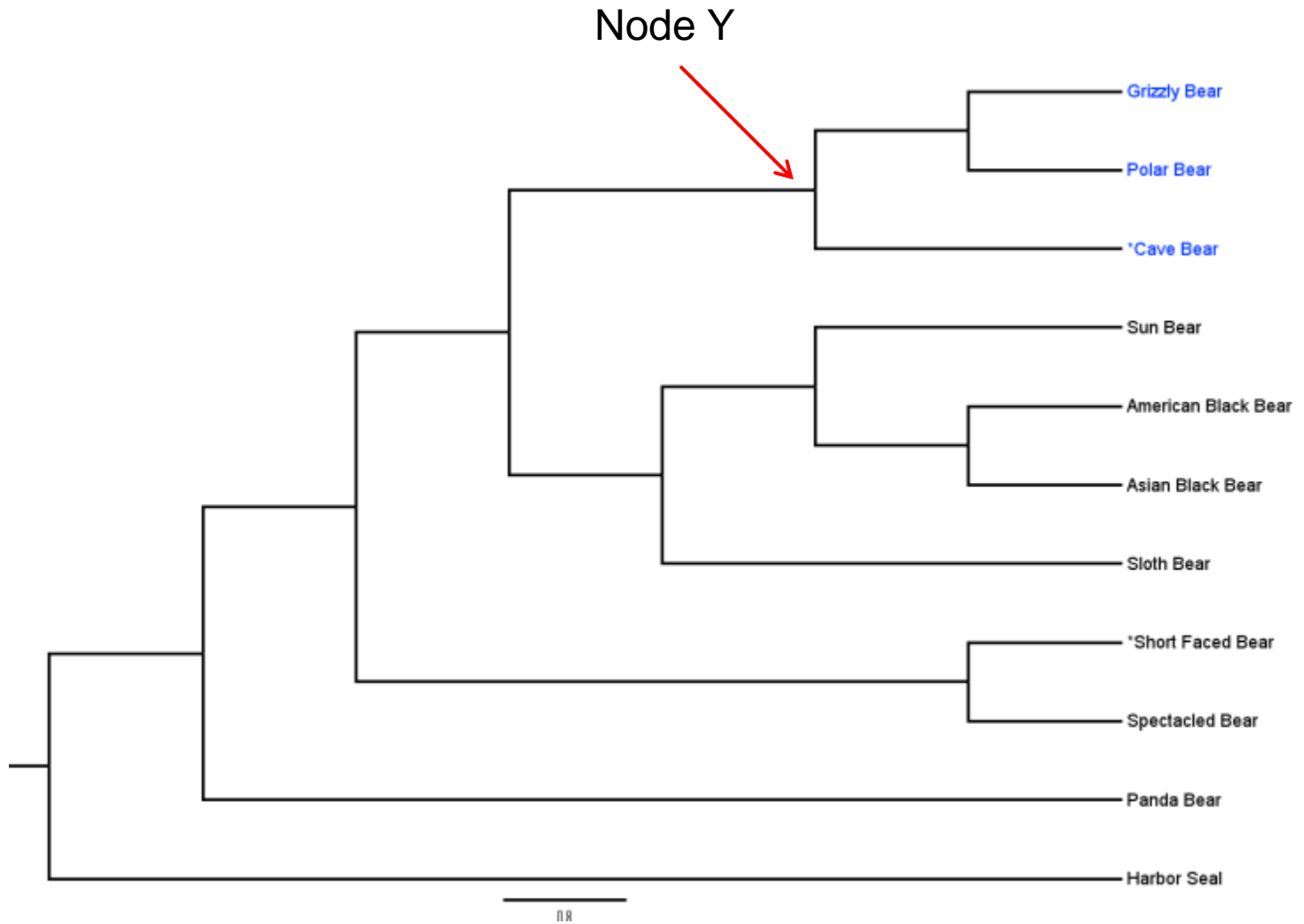
* Extinct species



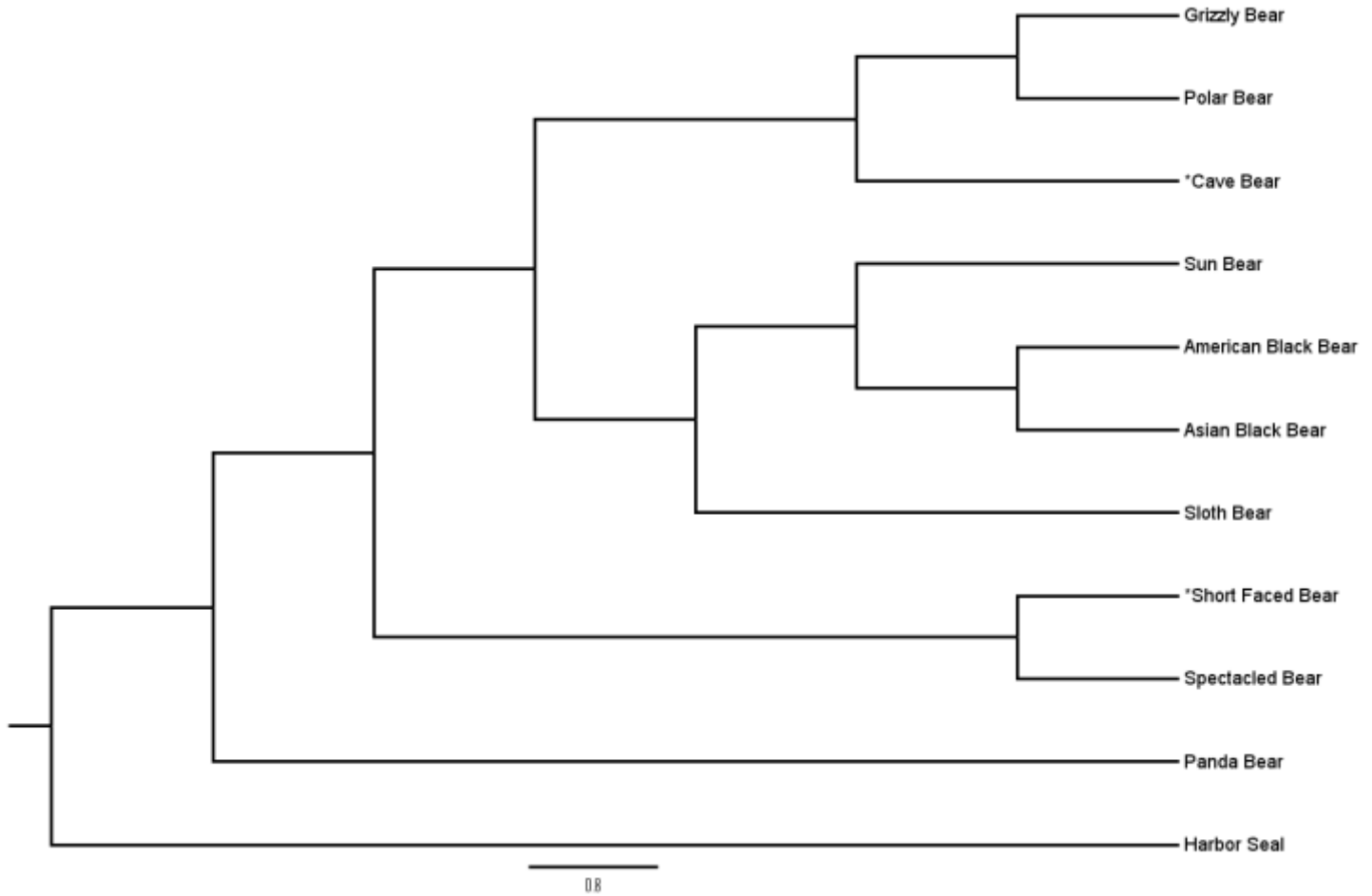
* Extinct species



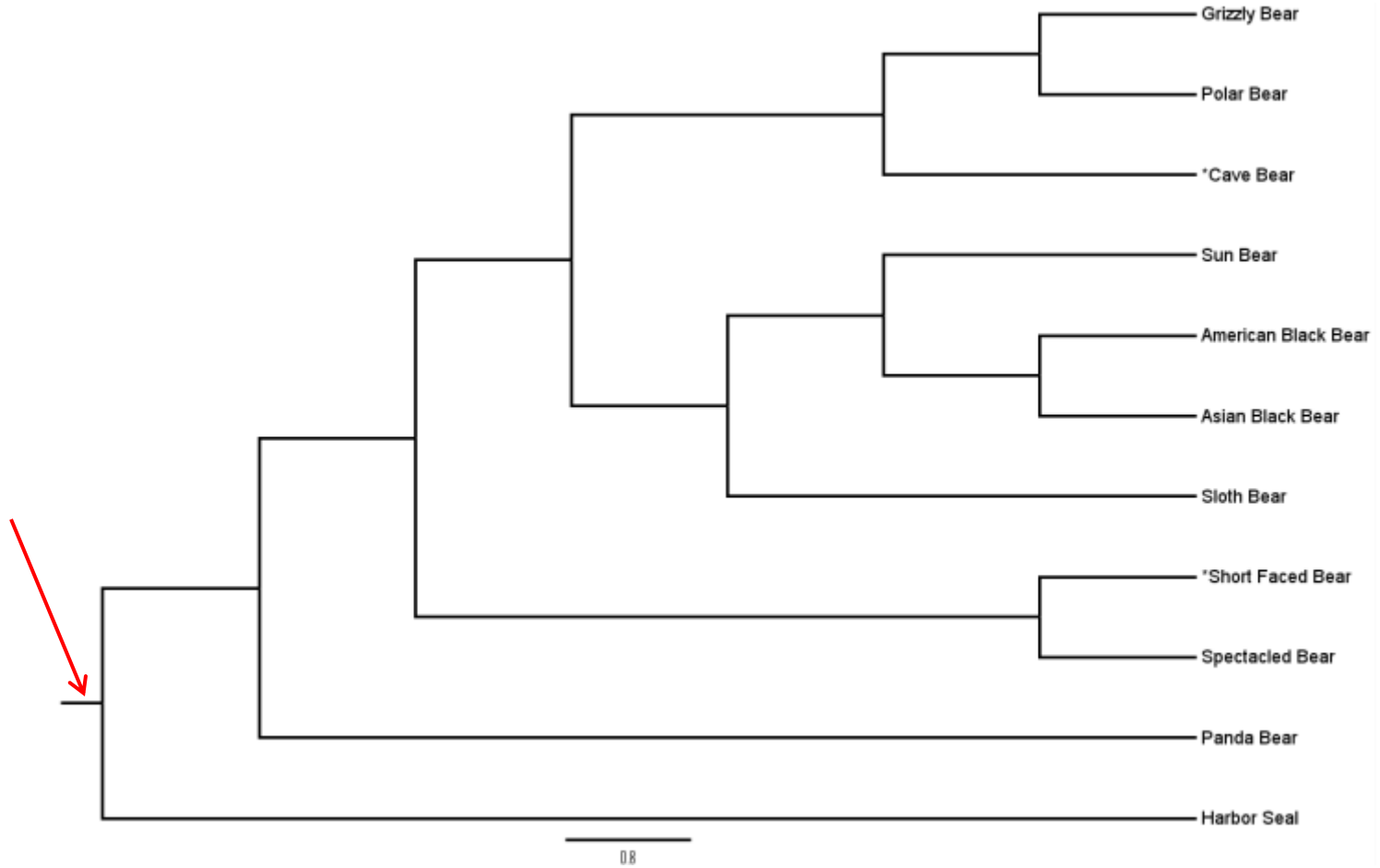
- Node Y is *ancestor* to grizzly bear, polar bear and cave bear



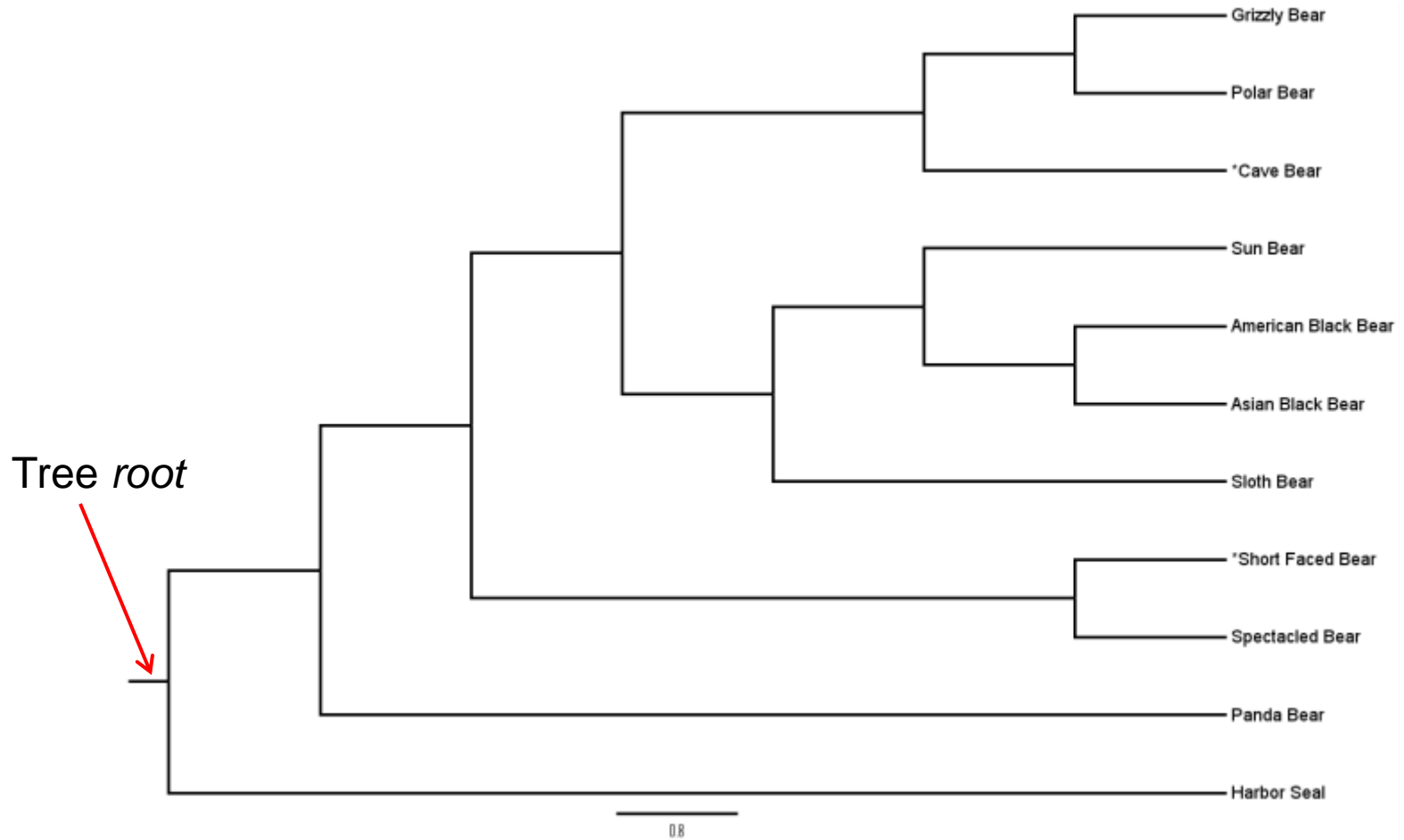
- Node Y is _____ to grizzly bear, polar bear and cave bear



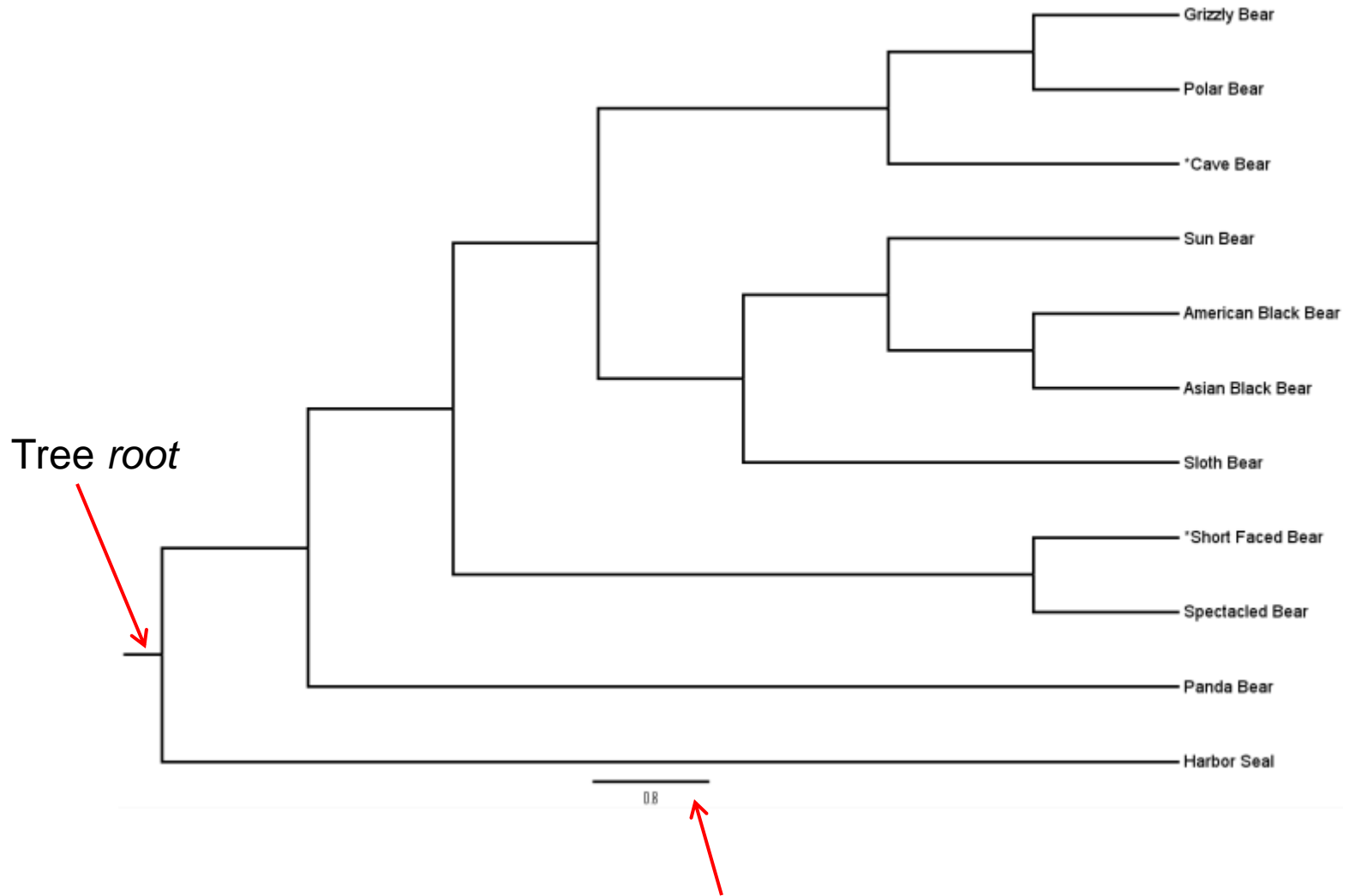
* Extinct species



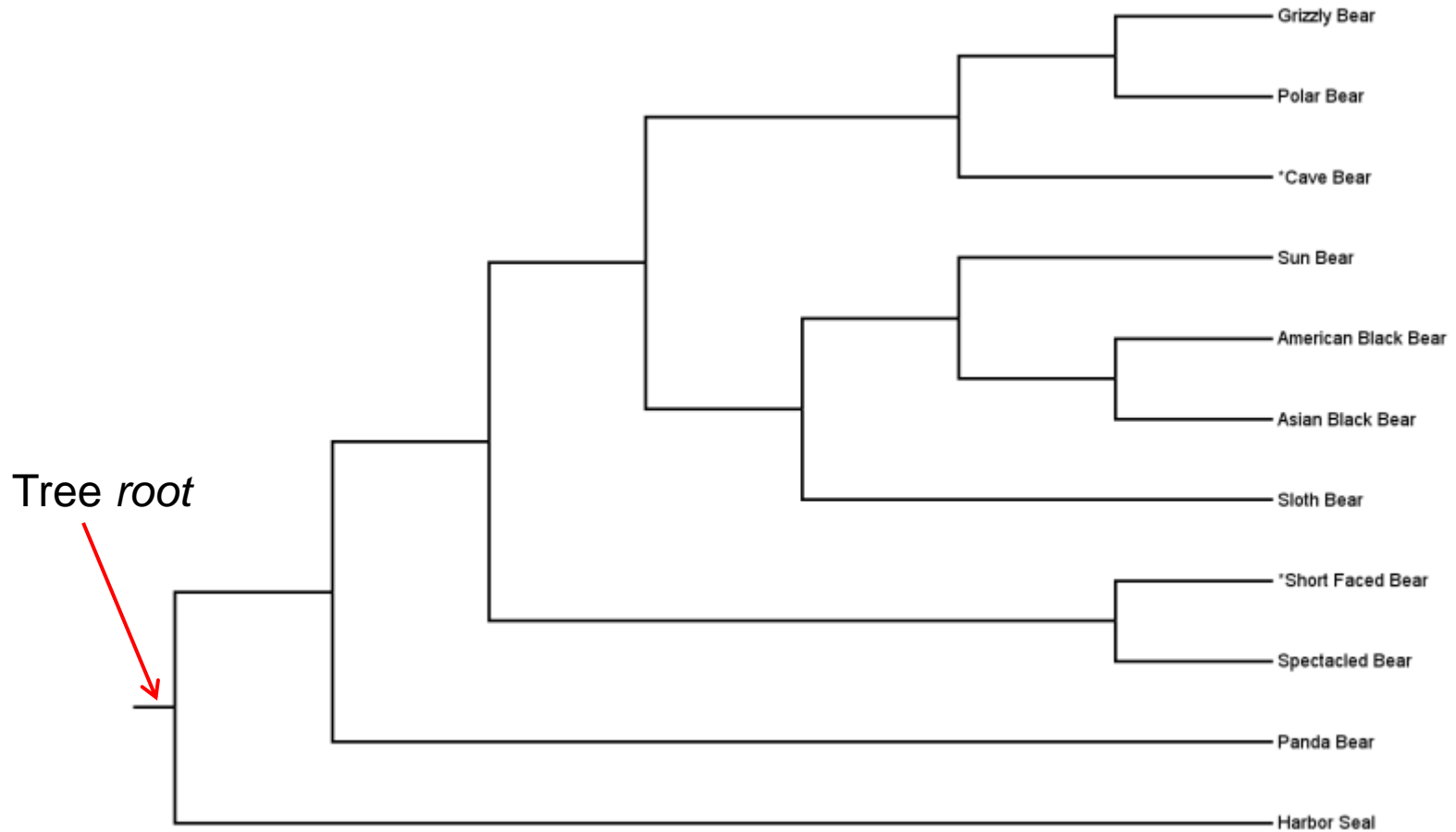
* Extinct species



* Extinct species



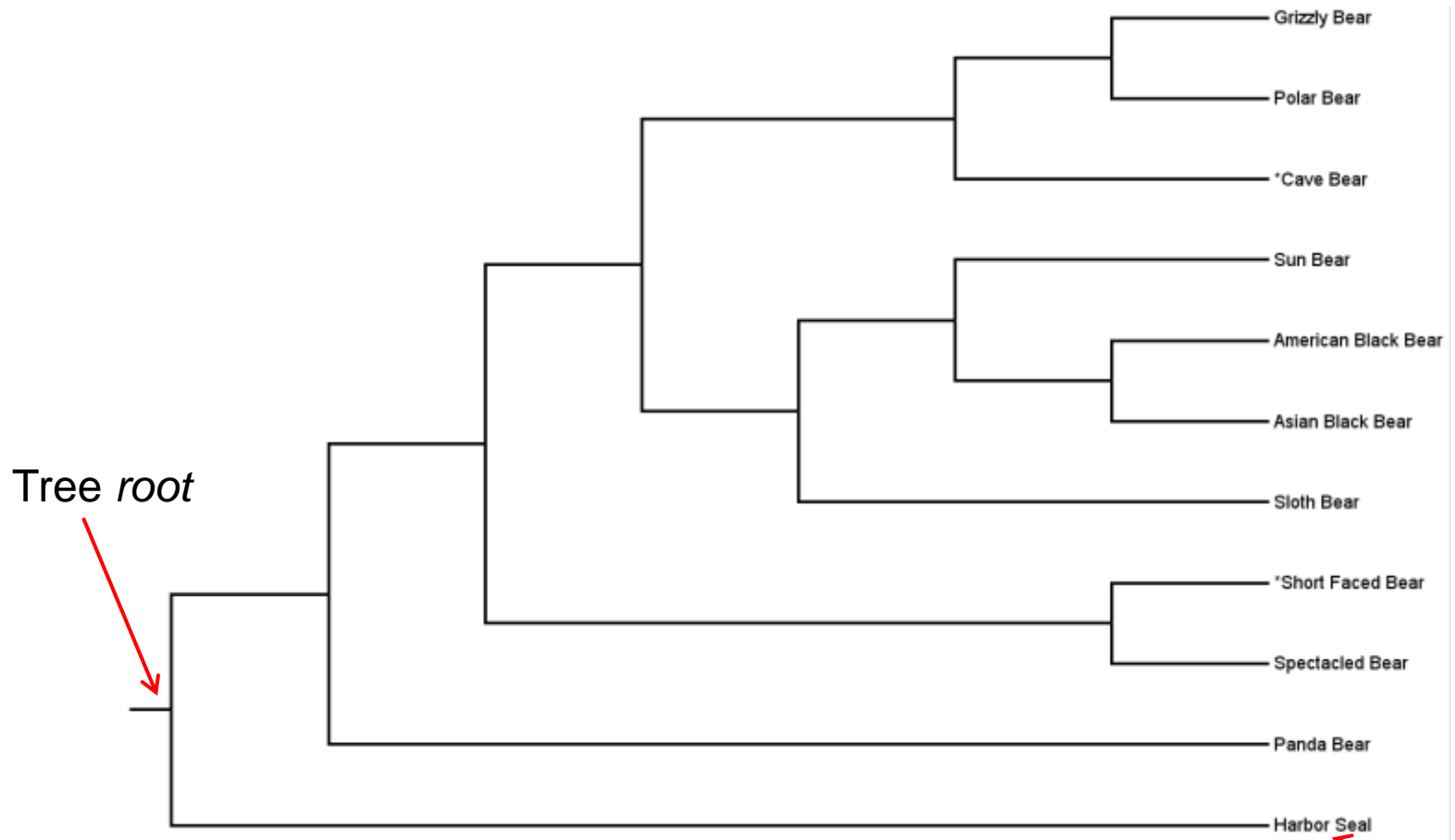
* Extinct species



08

Scale Bar

* Extinct species



0.8

Scale Bar

* Extinct species



Knowing how to create and read a phylogeny, now you will create a tree using mitochondrial DNA sequences of bird taxa including the dodo bird.