IDEAS Module 2 Spring 2018:

Introduction to Phylogenetic Comparative Methods

Outline

- Intro to Phylogenies

- Statistical thinking and the problem of shared evolutionary history

Outline

- **Intro to Phylogenies**

- Statistical thinking and the problem of shared evolutionary history

How are these species related?

The Linnaean Classification System

Least specific	Kingdom
	Phylum
	Class
	Order
	Family
	Genus
Most specific	*species*

Family	Genus	species
Emydidae	*Graptemys*	*nigrinoda*
Emydidae	*Graptemys*	*flavimaculata*
Emydidae	*Terrapene*	*ornata*
Emydidae	*Pseudemys*	*concinna*
Geoemydidae	*Rhinoclemmys*	*nigrinoda*

Family	Genus	species
Emydidae	*Graptemys*	*nigrinoda*
Emydidae	*Graptemys*	*flavimaculata*
Emydidae	*Graptemys*	*geographica*
Emydidae	*Graptemys*	*ouachitensis*
Emydidae	*Graptemys*	*versa*

How are these species related?

How are these species related?

Lepidochelys olivacea

Carreta carreta

Dermochelys coriacea

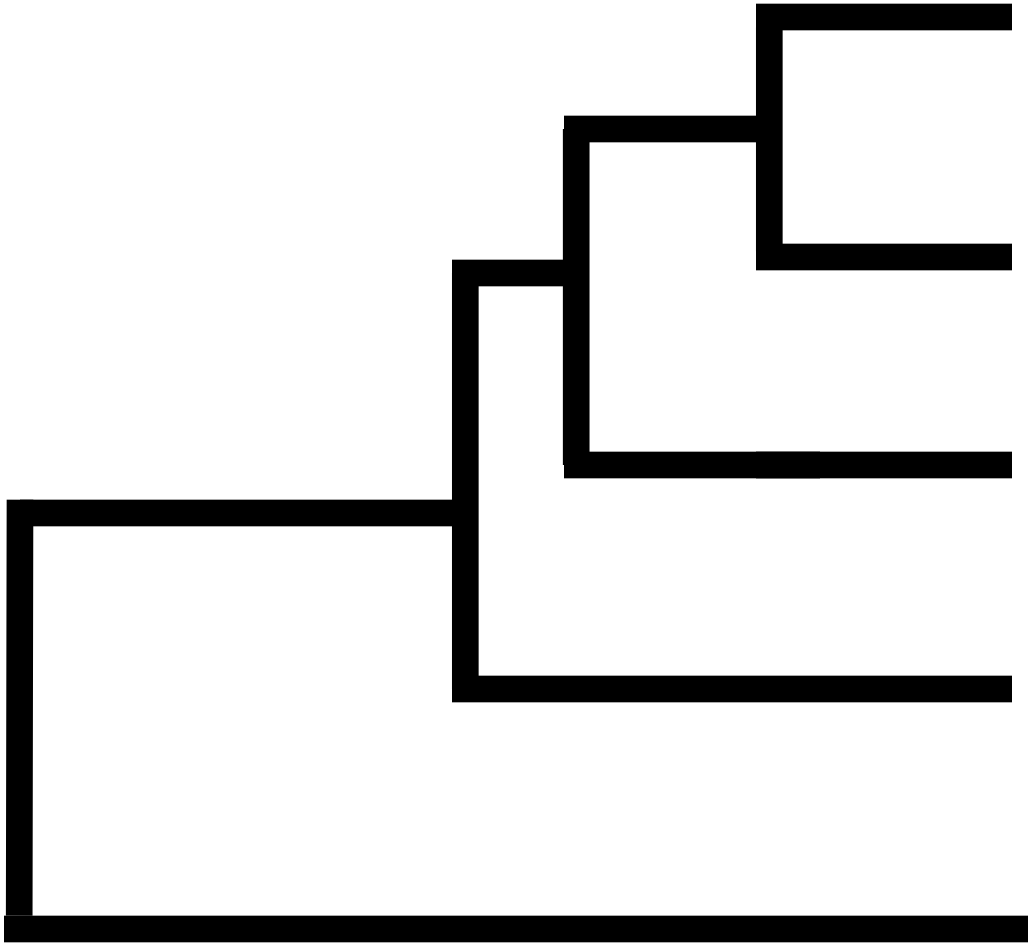
Eratemochelys imbricata

Chelonia mydas

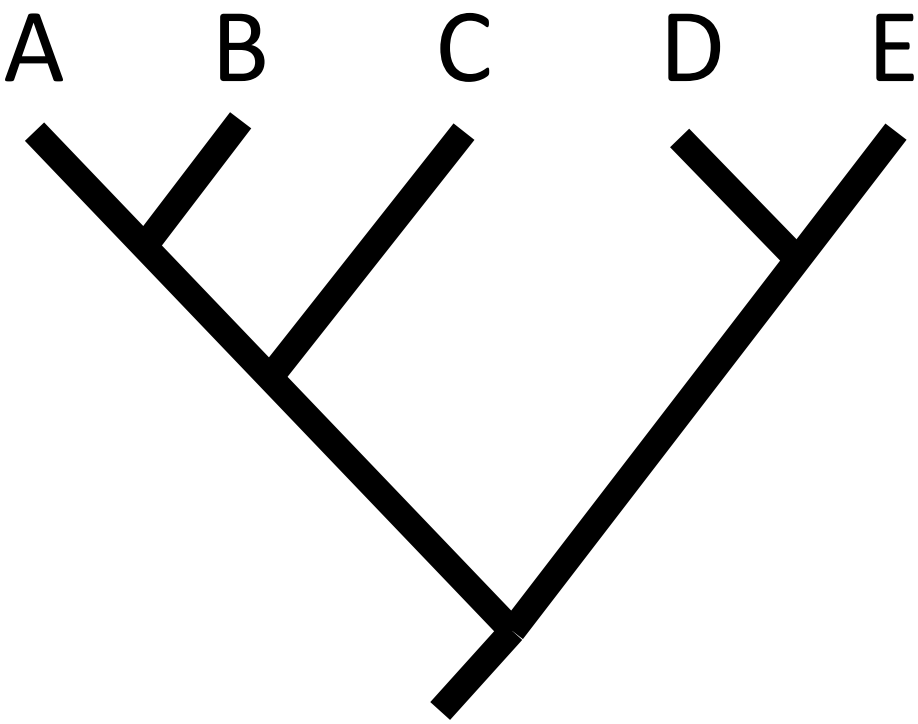
Family: Cheloniidae

Family: Dermochelyidae

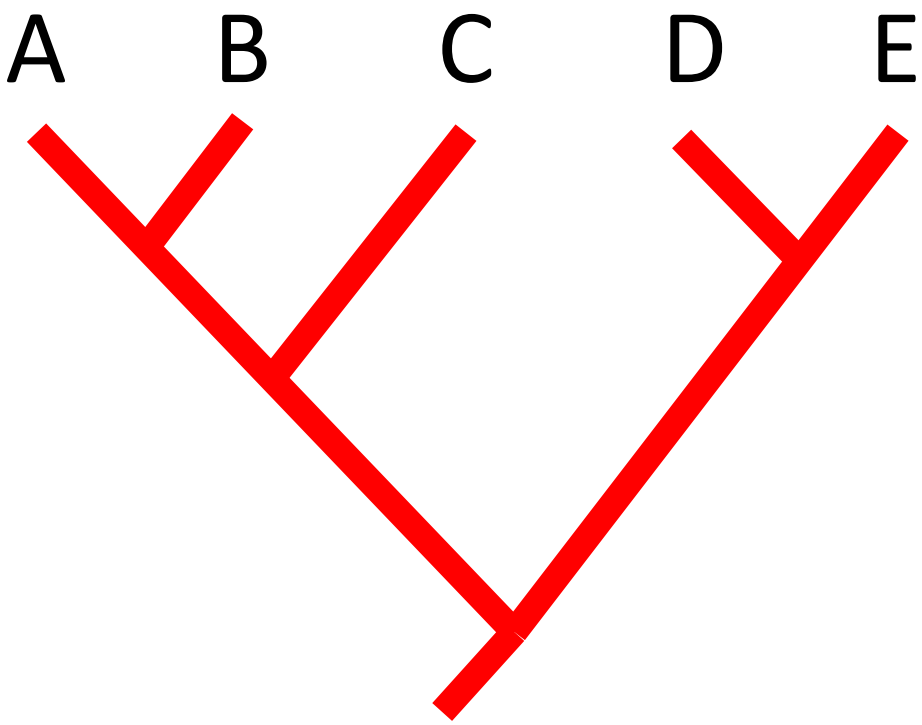
A phylogeny summarizes the evolutionary relationships of a set of species



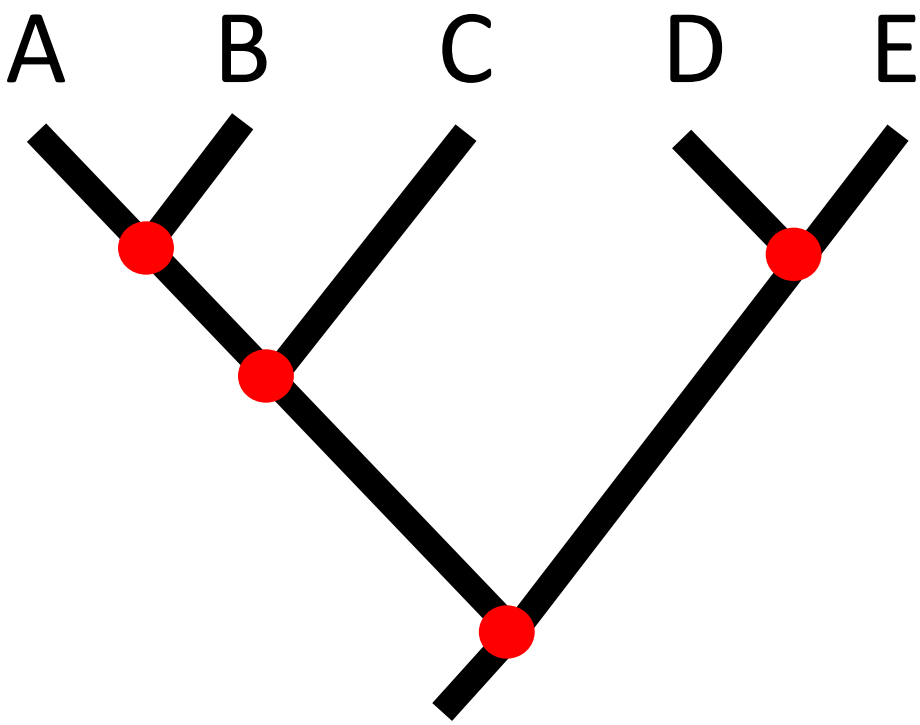
Basic Phylogenetic Terminology



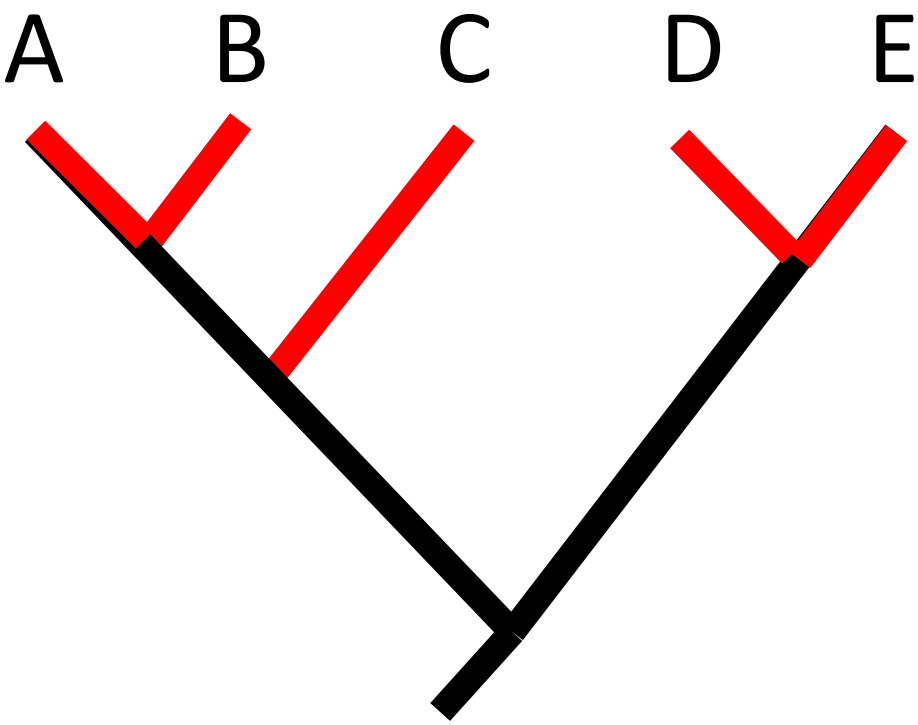
Branches



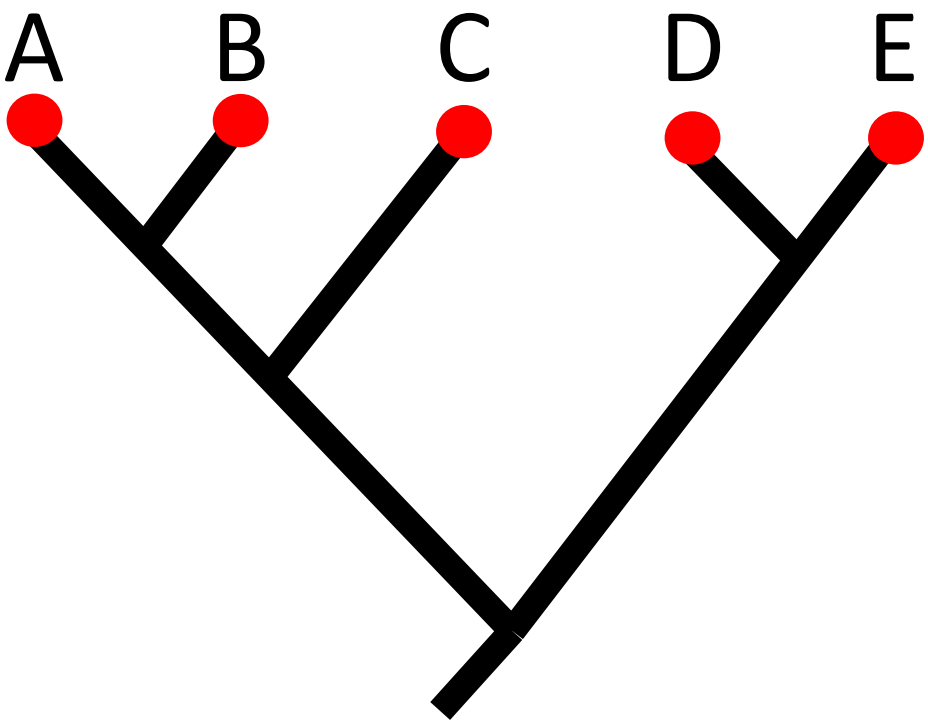
Nodes



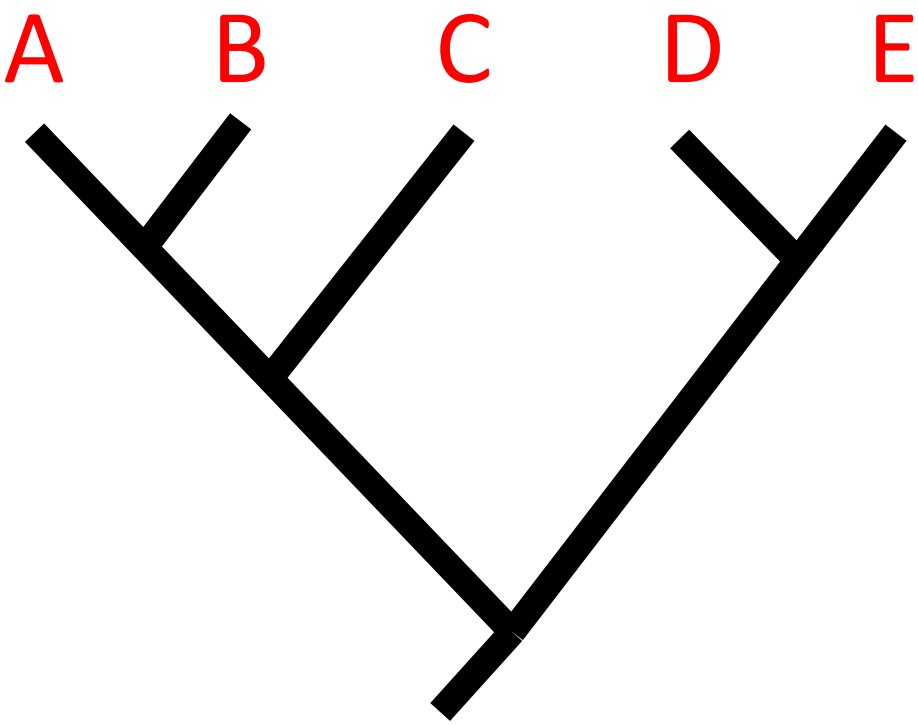
Terminal Branches



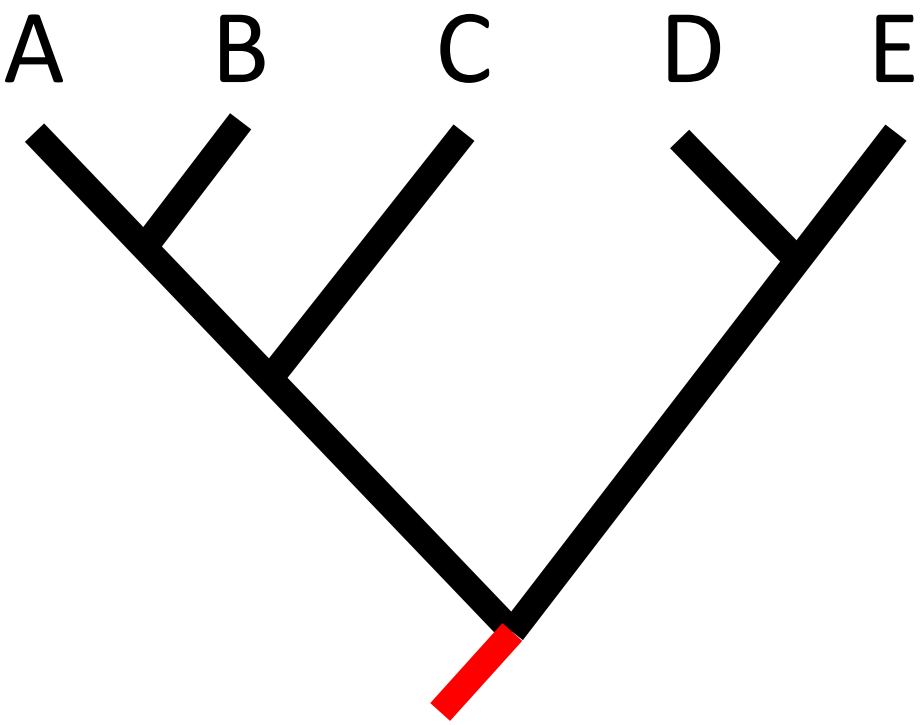
Tips



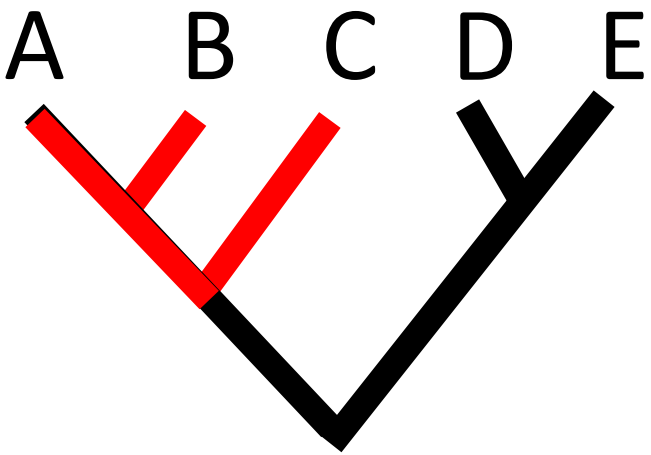
Tip Labels

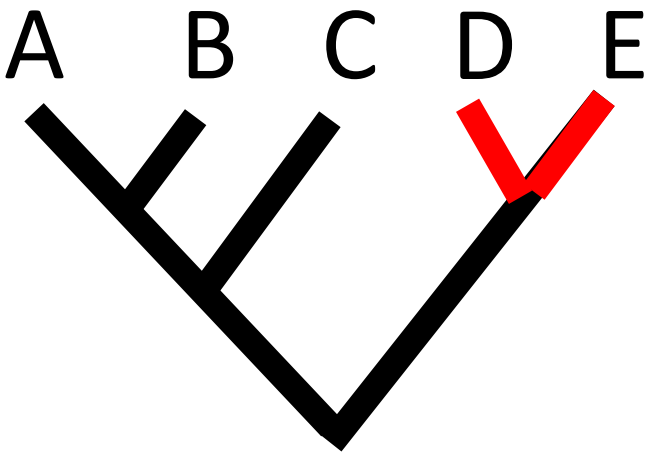


Root

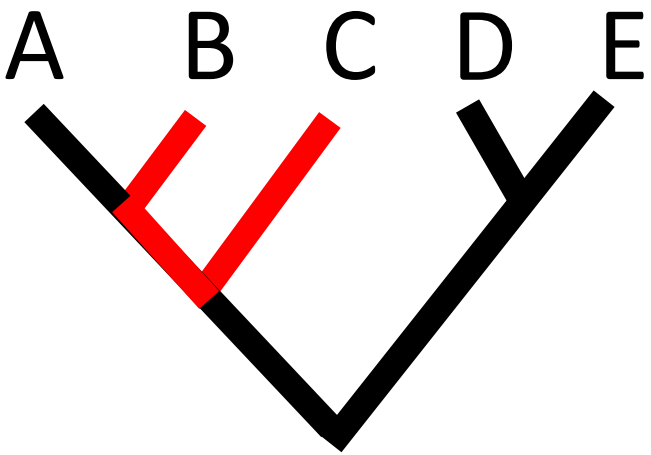


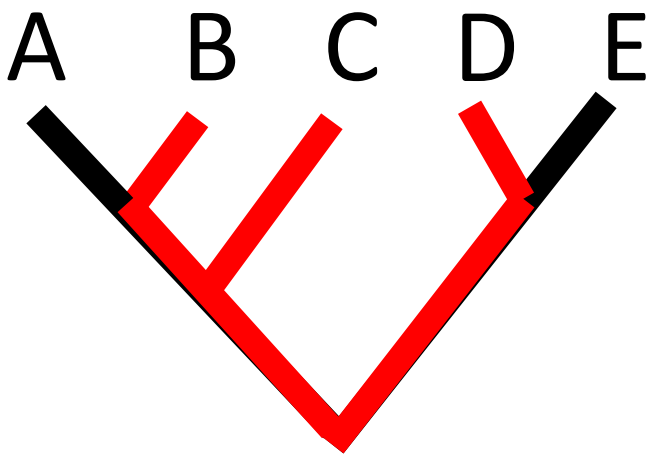
monophyletic groups (clades)



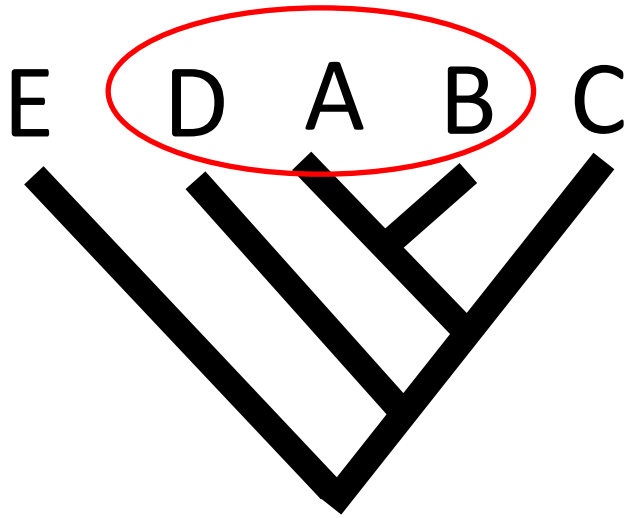


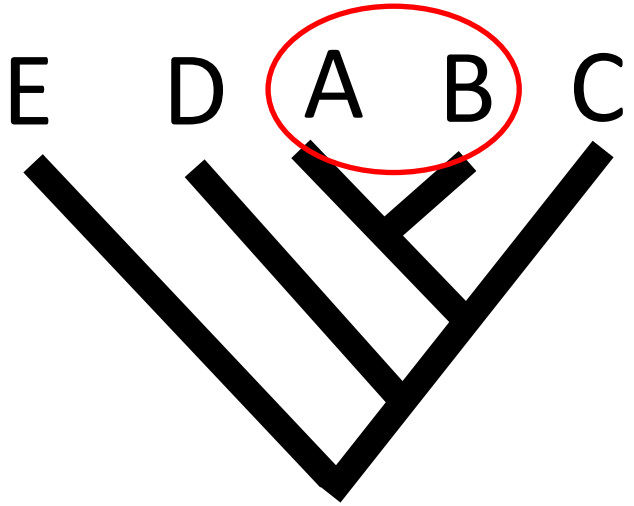
paraphyletic groups (non-monophyletic)

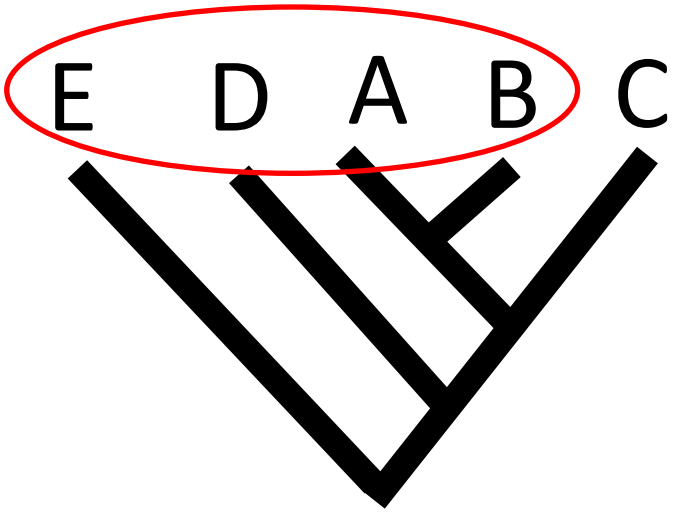


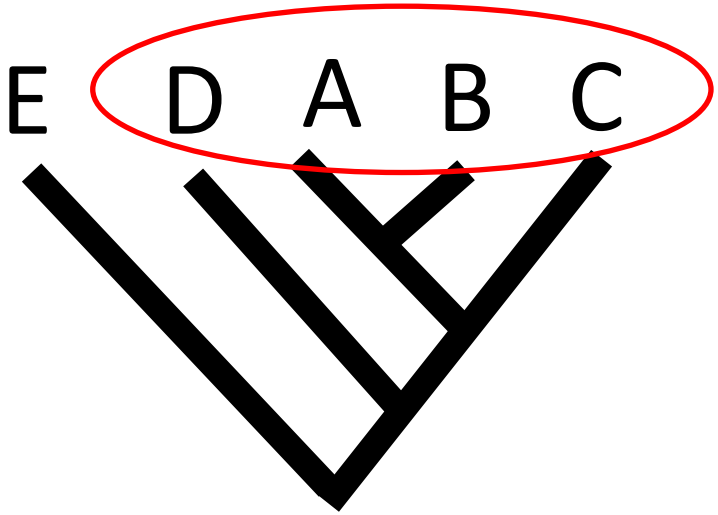


Which are monophyletic?

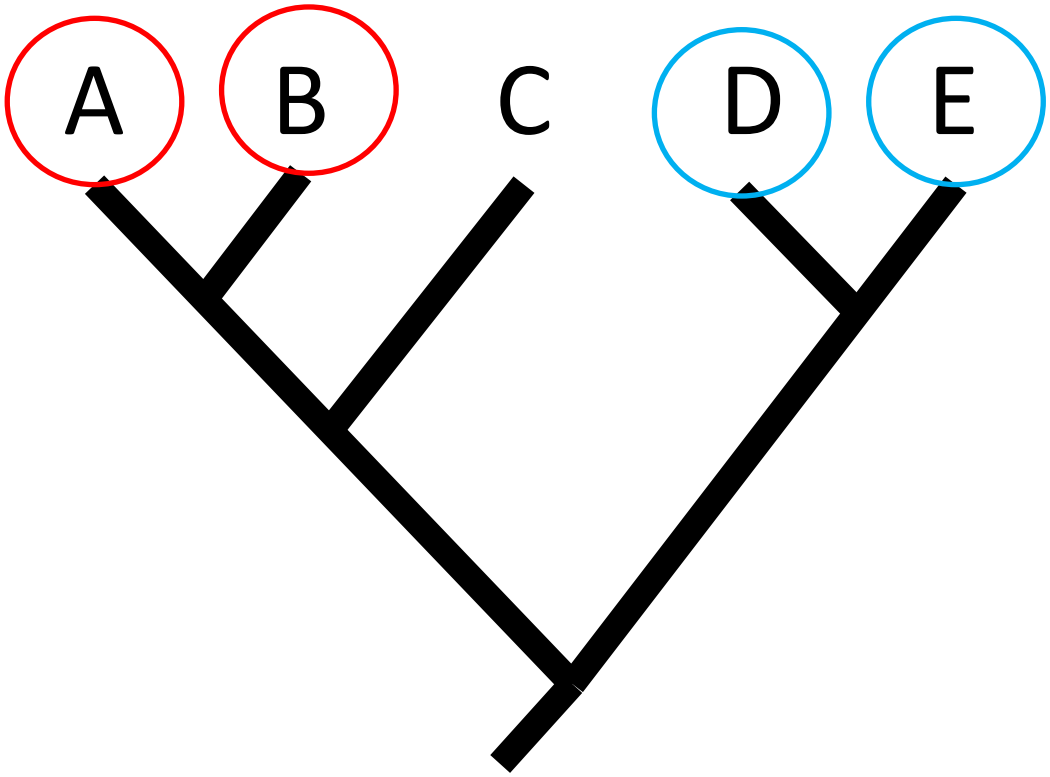




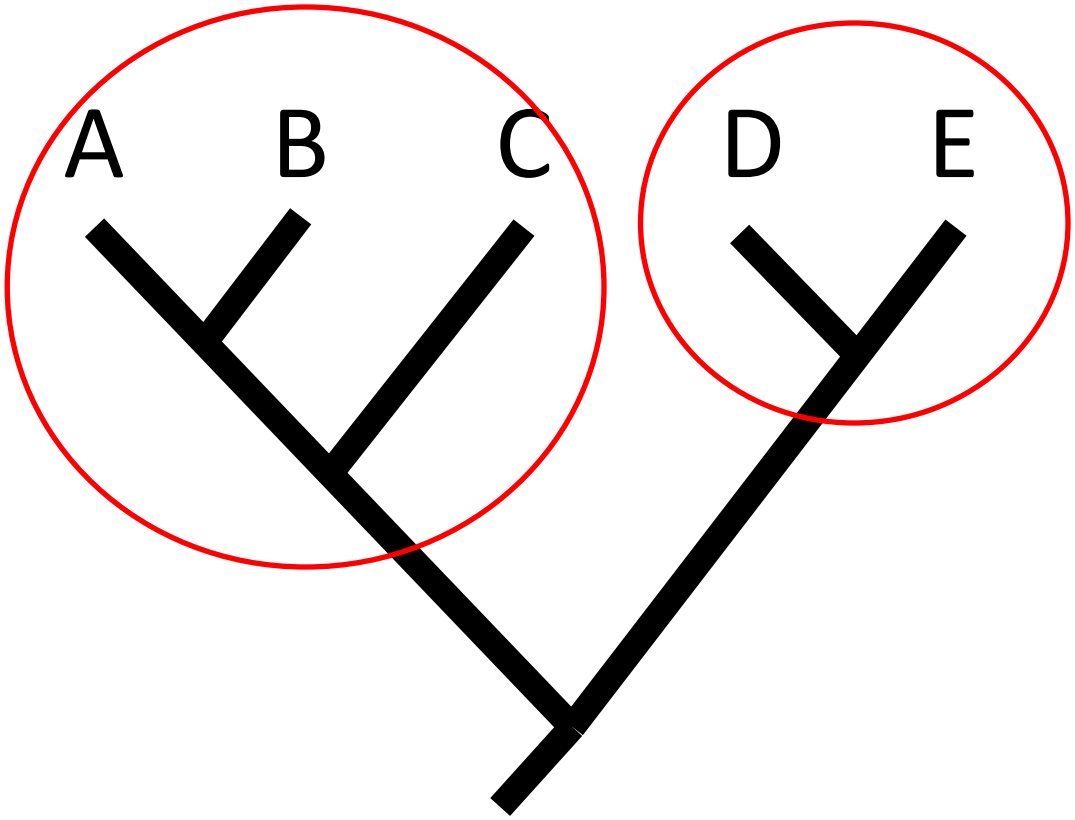




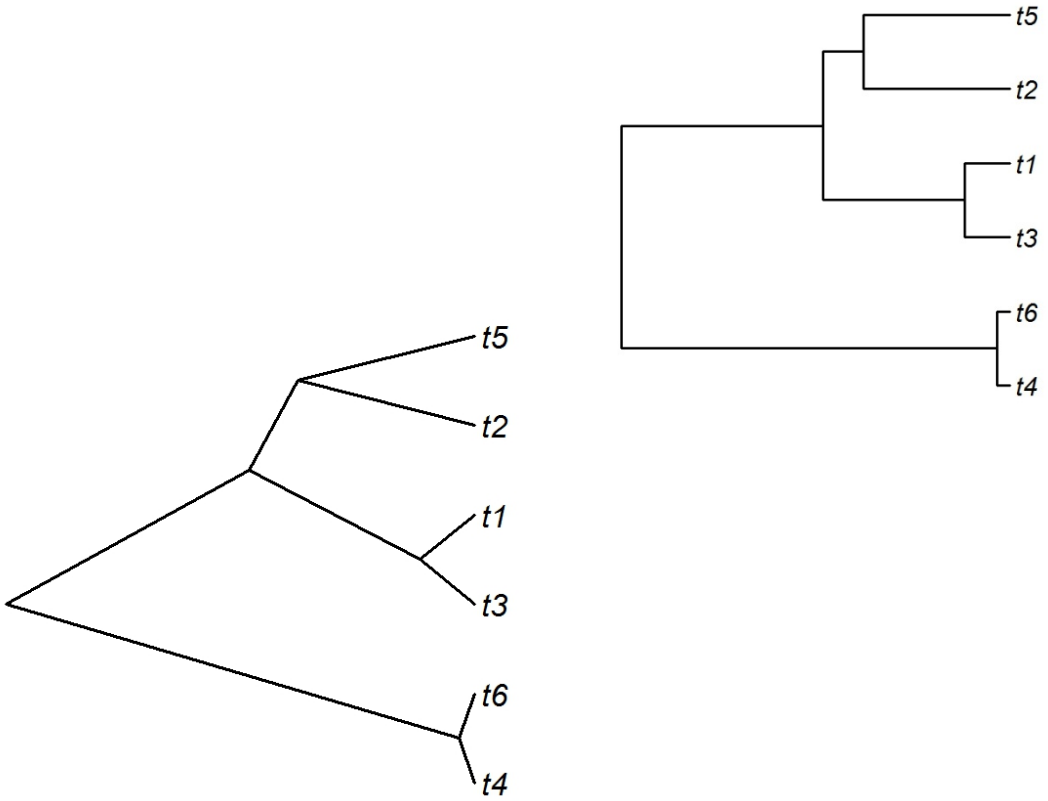
Sister Species



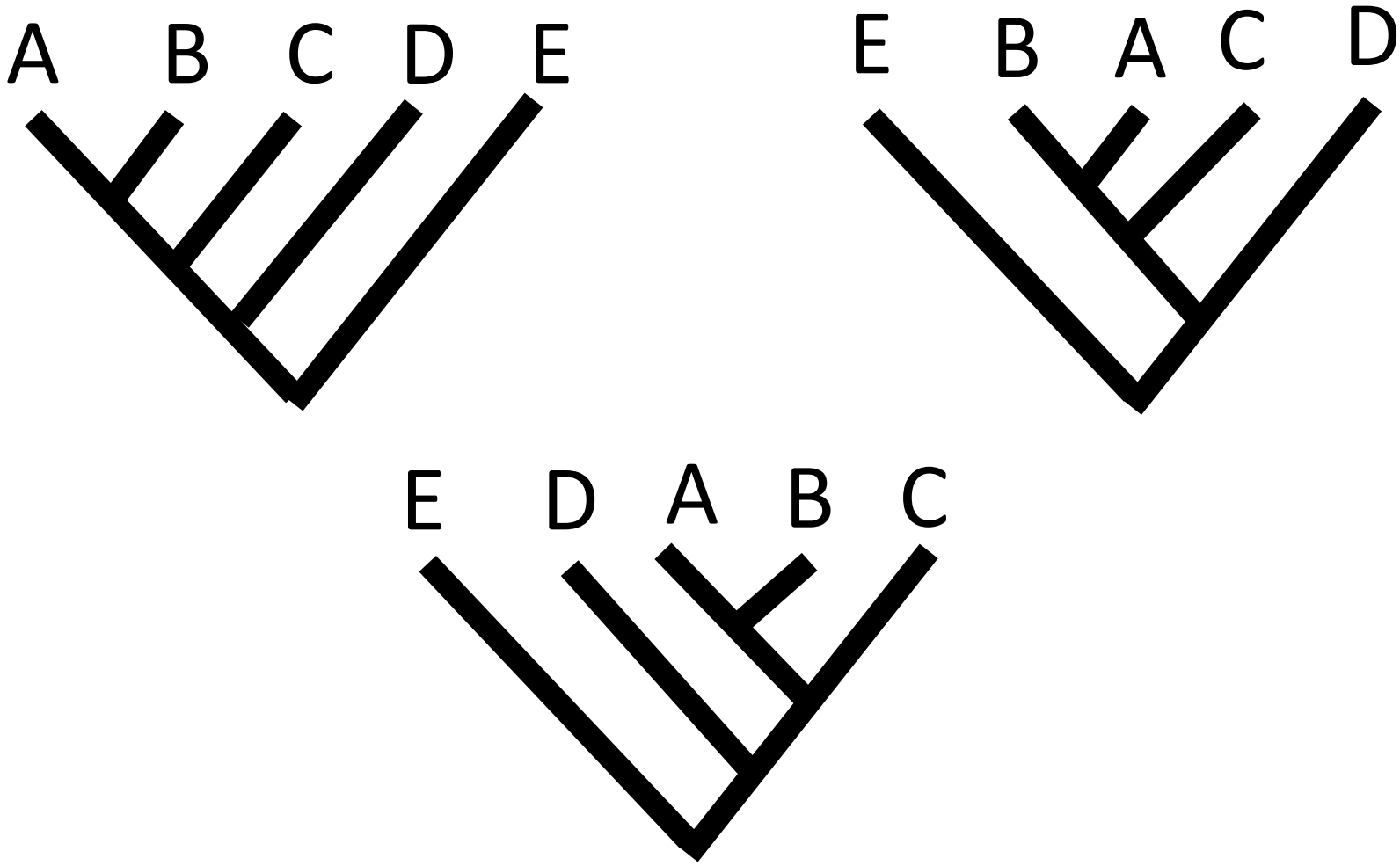
Sister Clades



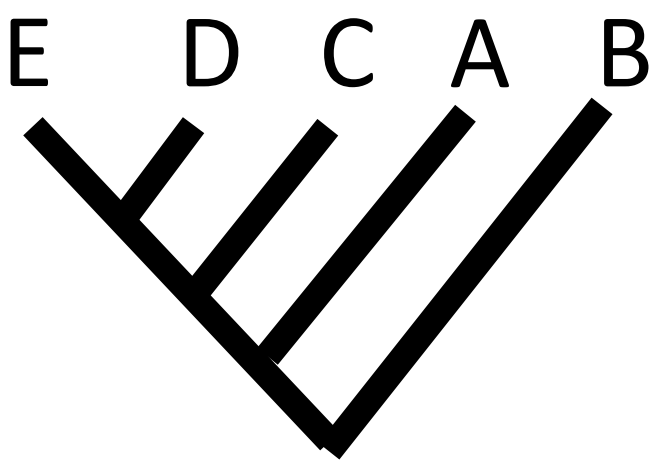
These cladograms contain the same information

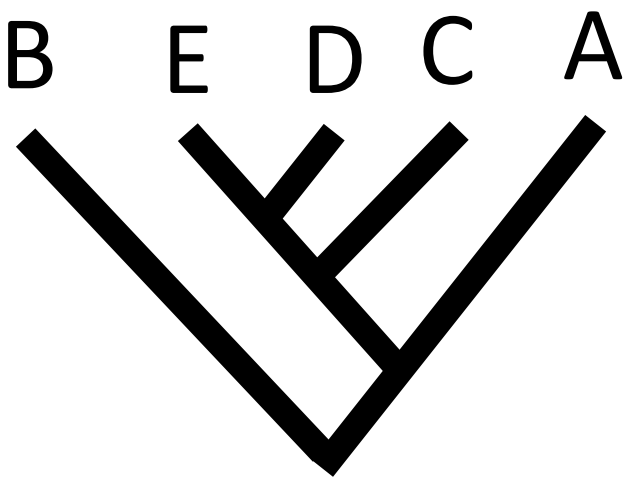


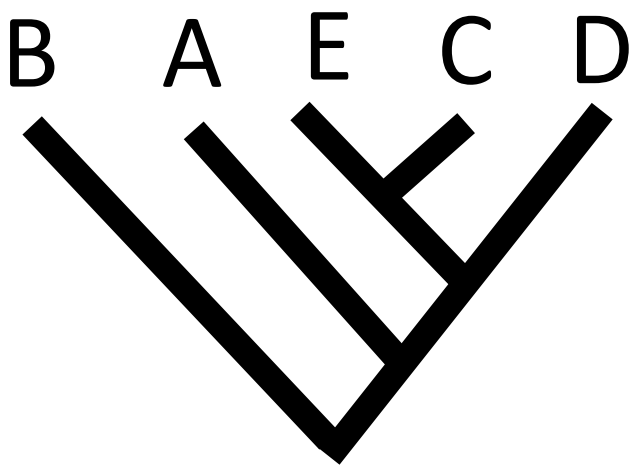
These cladograms are identical

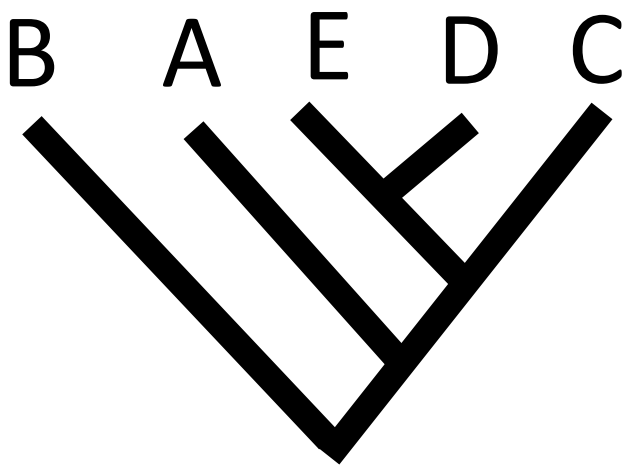


Which of these trees is not the same?

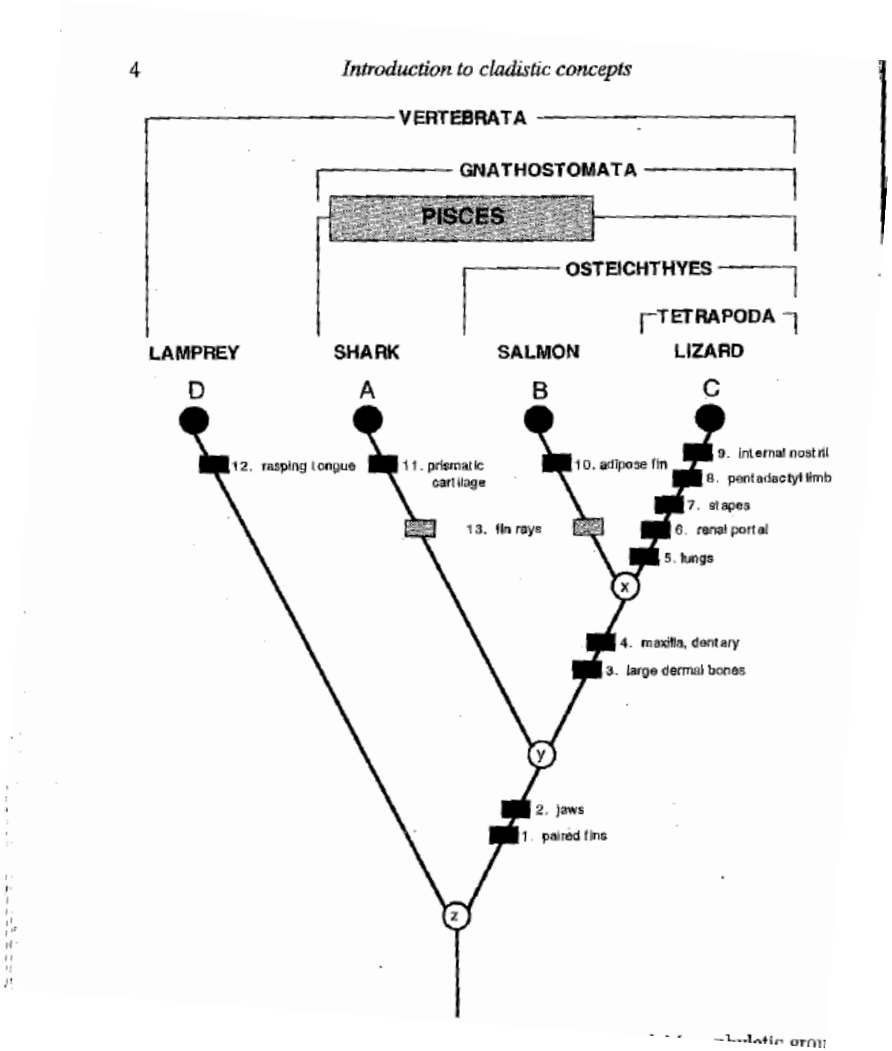




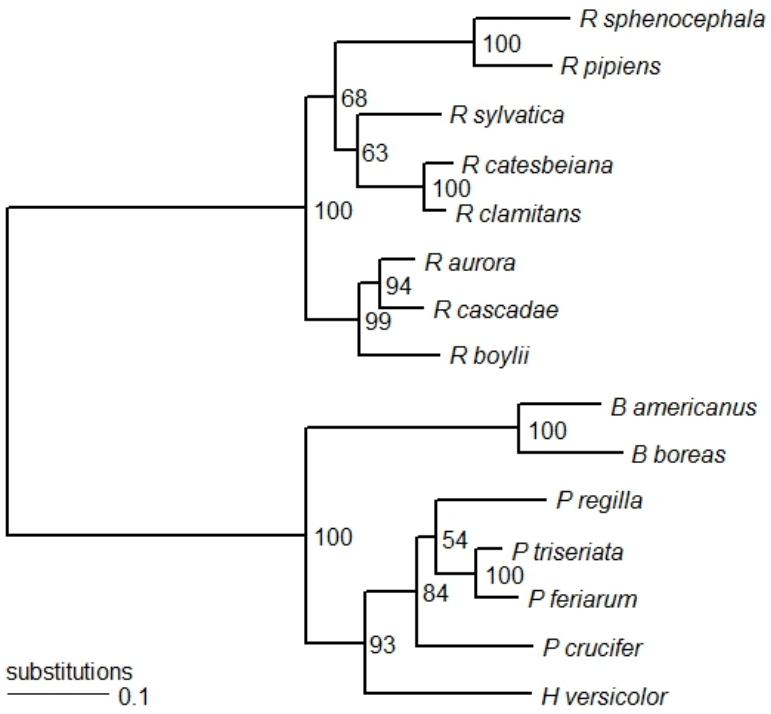


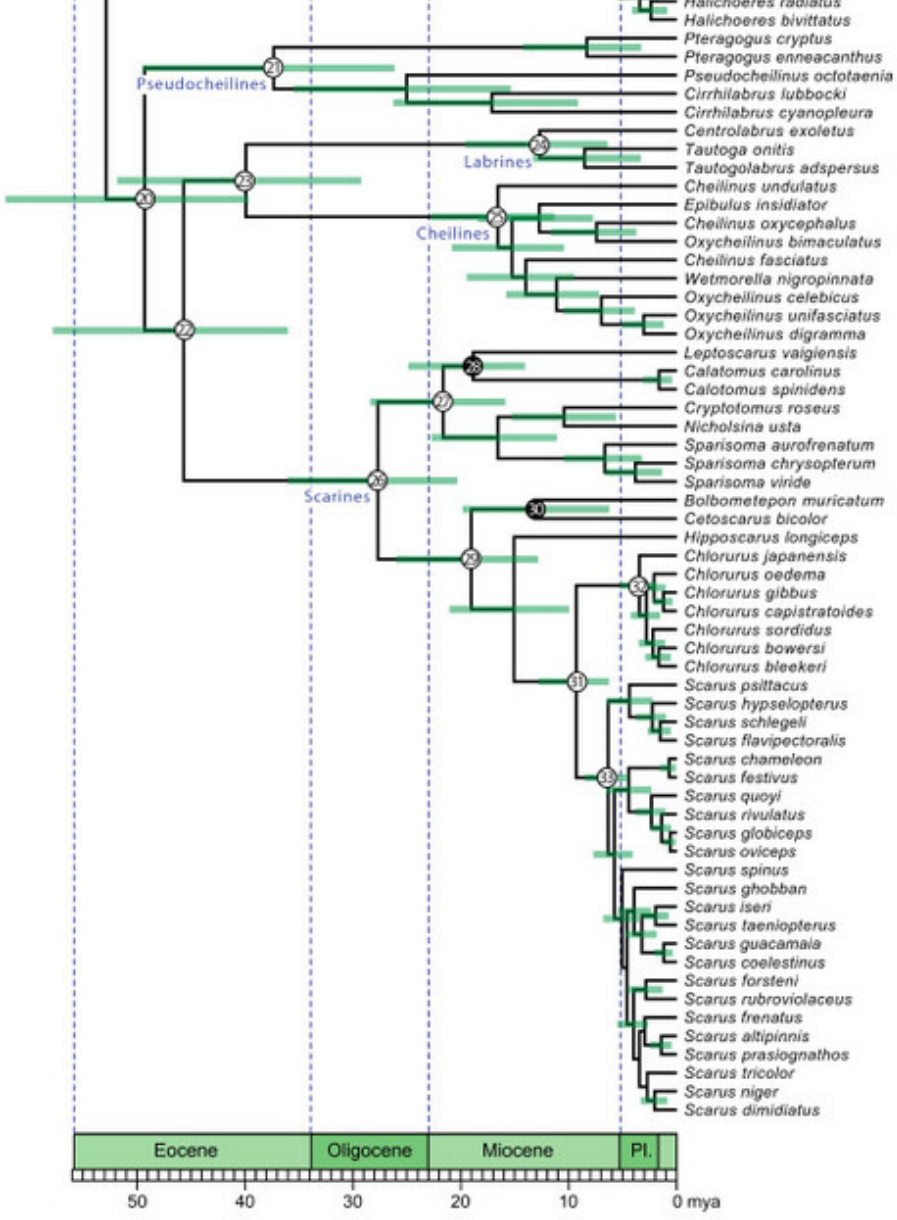


A cladogram


4
Introduction to cladistic concepts
VERTEBRATA
GNATHOSTOMATA
PISCES
OSTEICHTHYES
TETRAPODA
LAMPREY
SHARK
SALMON
LIZARD
D
A
B
C
12. rasping tongue
11. prismatic cartilage
10. adipose fin
9. internal nostril
8. pentadactyl limb
7. stapes
6. renal portal
13. fin rays
5. lungs
x
4. maxilla, dentary
3. large dermal bones
y
2. jaws
1. paired fins
z

A phylogram


R sphenocephala
100
R pipiens
68
R sylvatica
63
R catesbeiana
100
R clamitans
100
R aurora
94
R cascadae
99
R boylii
B americanus
100
B boreas
P regilla
100
54
P triseriata
100
84
P feriarum
93
P crucifer
H versicolor
substitutions
0.1



A chronogram

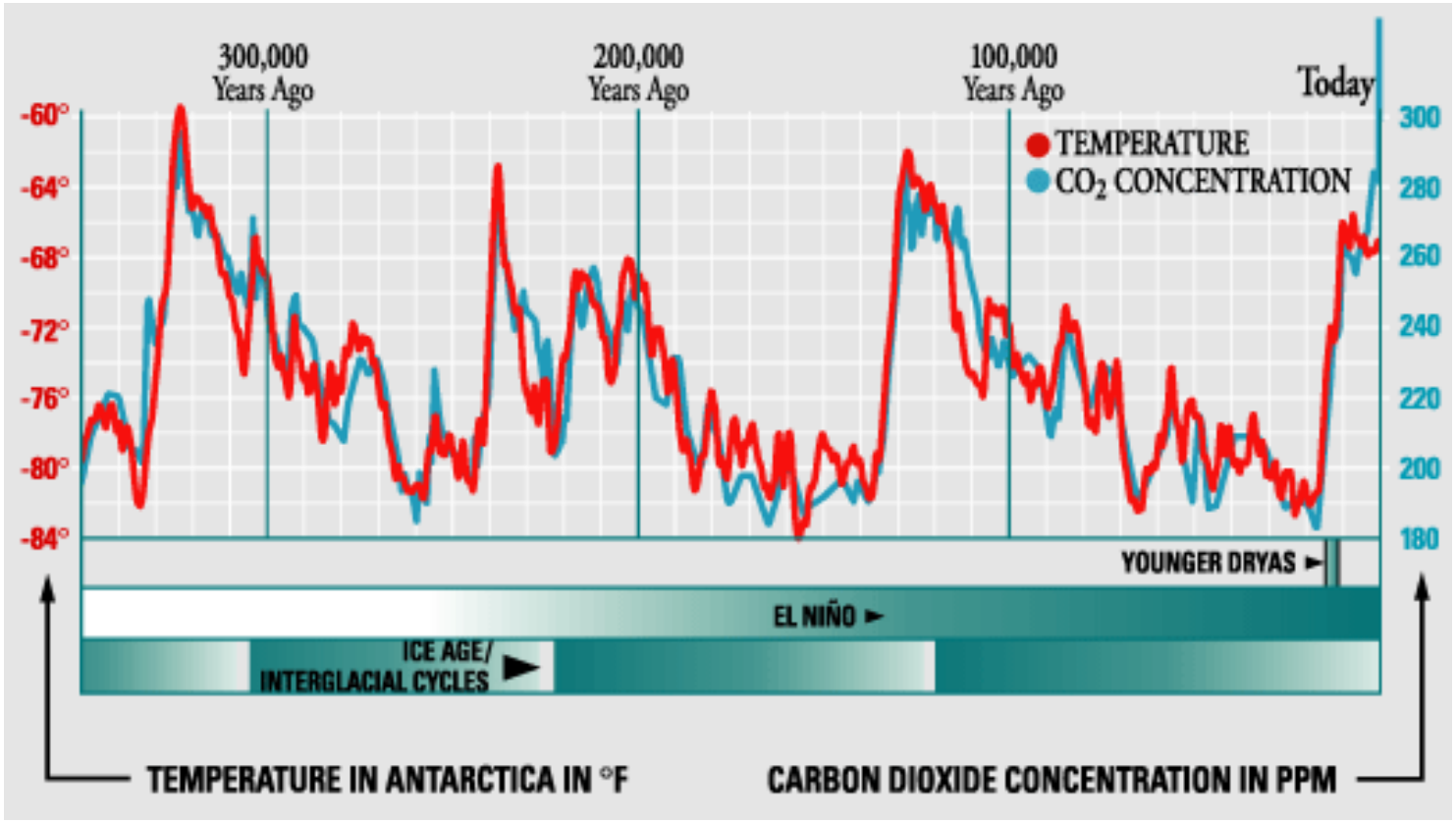
Outline

- Intro to Phylogenies

- **Statistical thinking and the problem of shared evolutionary history**

We will often see patterns in chance occurrences.

Is a pattern that we observe likely to have arisen by chance?



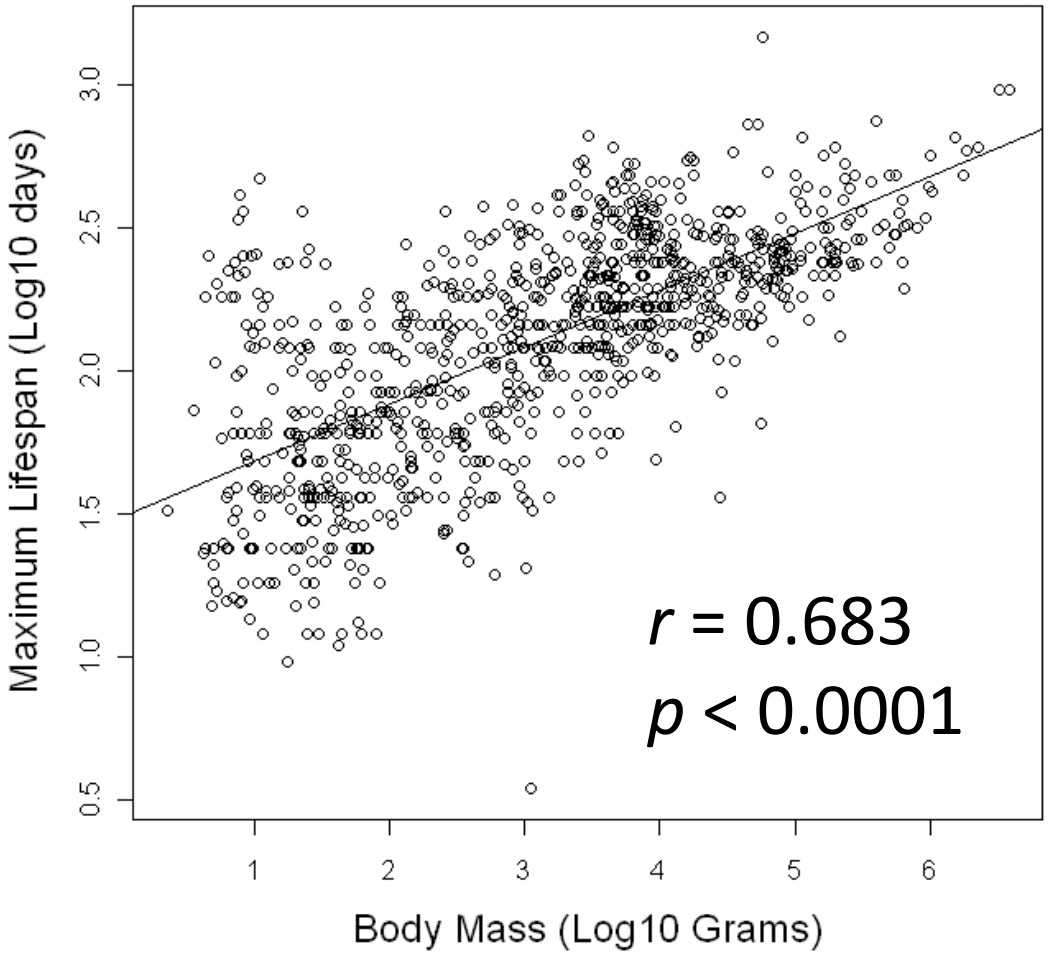
Is a pattern that we observe likely to have arisen by chance?

	Attempted suicide		Did not attempt suicide	
	Mean	SD	Mean	SD
Alcohol and drug use history				
Age at onset of regular drinking (years)	16.0	2.3	18.0	4.8
Age at onset of alcohol dependence (years)	20.6	6.0	23.1	8.0
Maximum number of drinks in 24 hours	34.2	21.8	24.9	19.3
Number of DSM-III-R alcohol dependence criteria endorsed	7.3	1.9	5.9	2.2
Number of withdrawal symptoms	4.4	3.2	2.3	2.9
Number of alcohol-related incidents of violence	2.9	1.5	2.1	1.5
Number of alcohol-related physical problems	0.5	0.7	0.3	0.6
Number of substances dependent on	1.3	1.2	0.8	1.1

- H_0: The pattern that we observe is due to chance alone.

- H_1: The pattern observed is not due to chance alone.

Regression Analysis

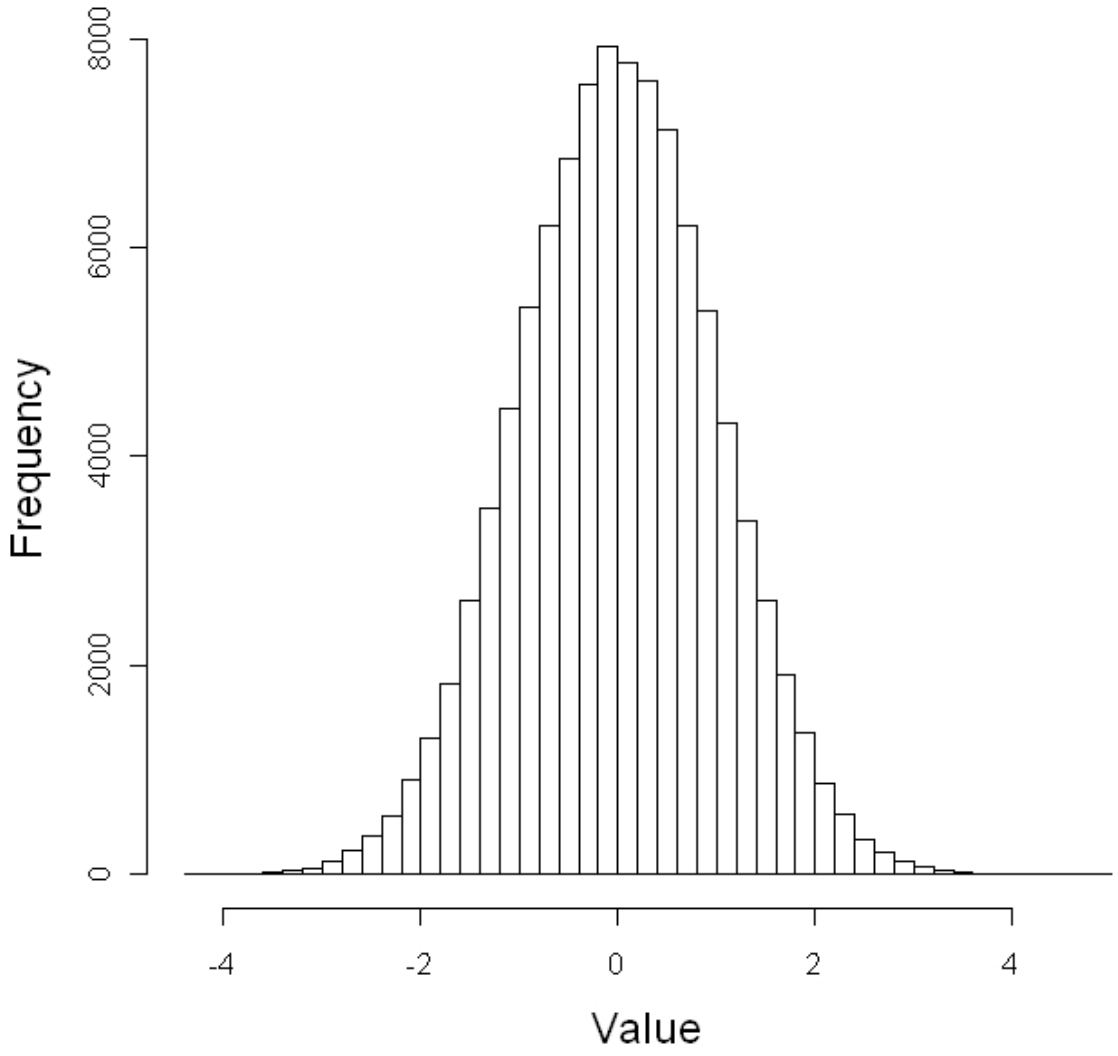


Regression Analysis

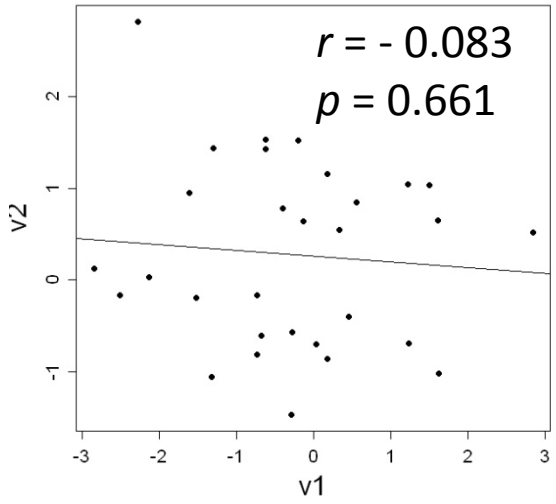
- *r* : the strength of the observed correlation

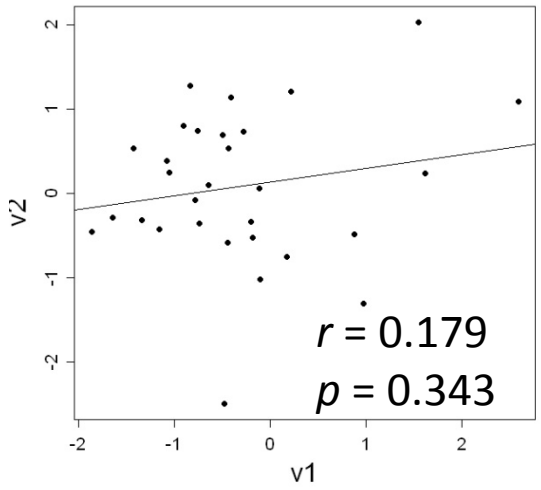
- *p*-value: the chance of a correlation at least as strong as the one observed occurring by chance

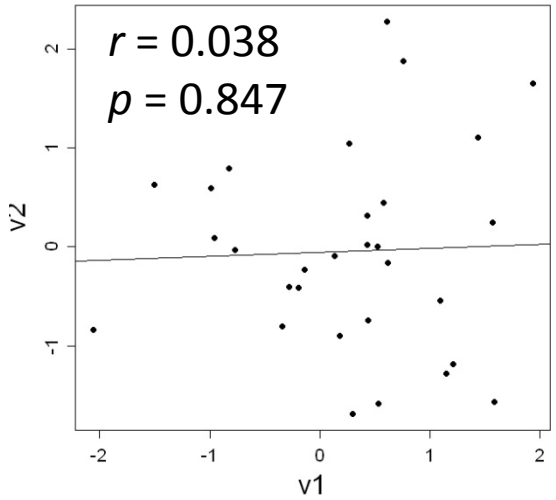
100,000 draws from random normal distribution

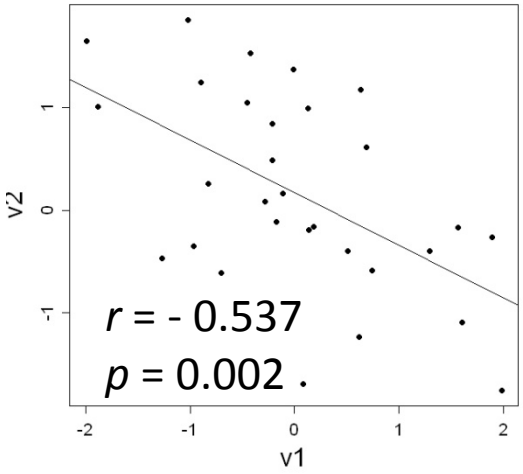


A series of random correlations

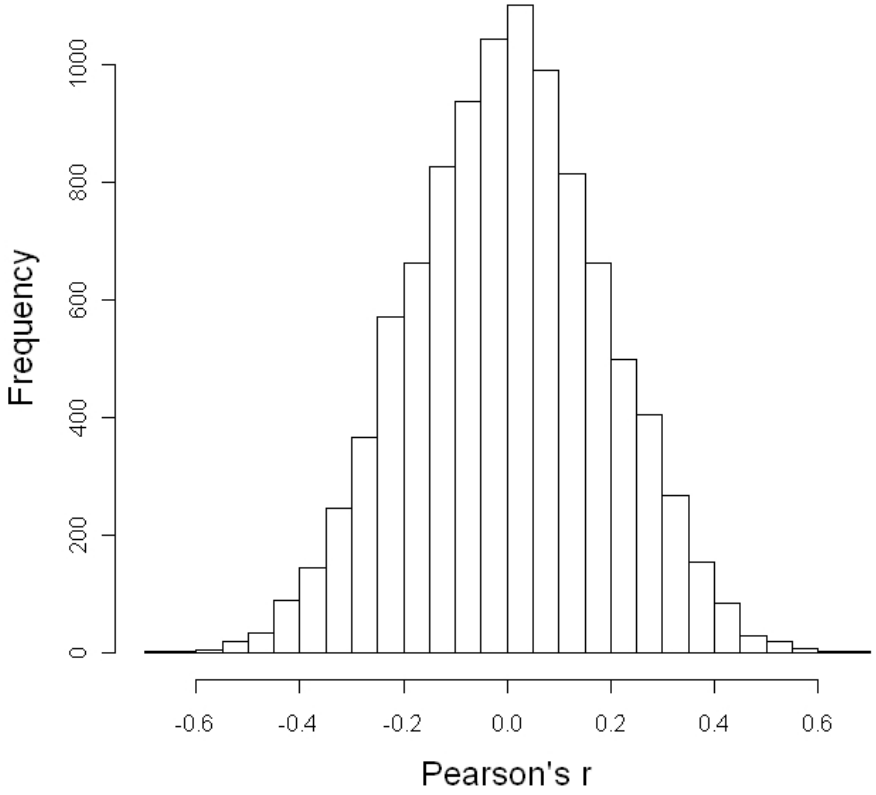








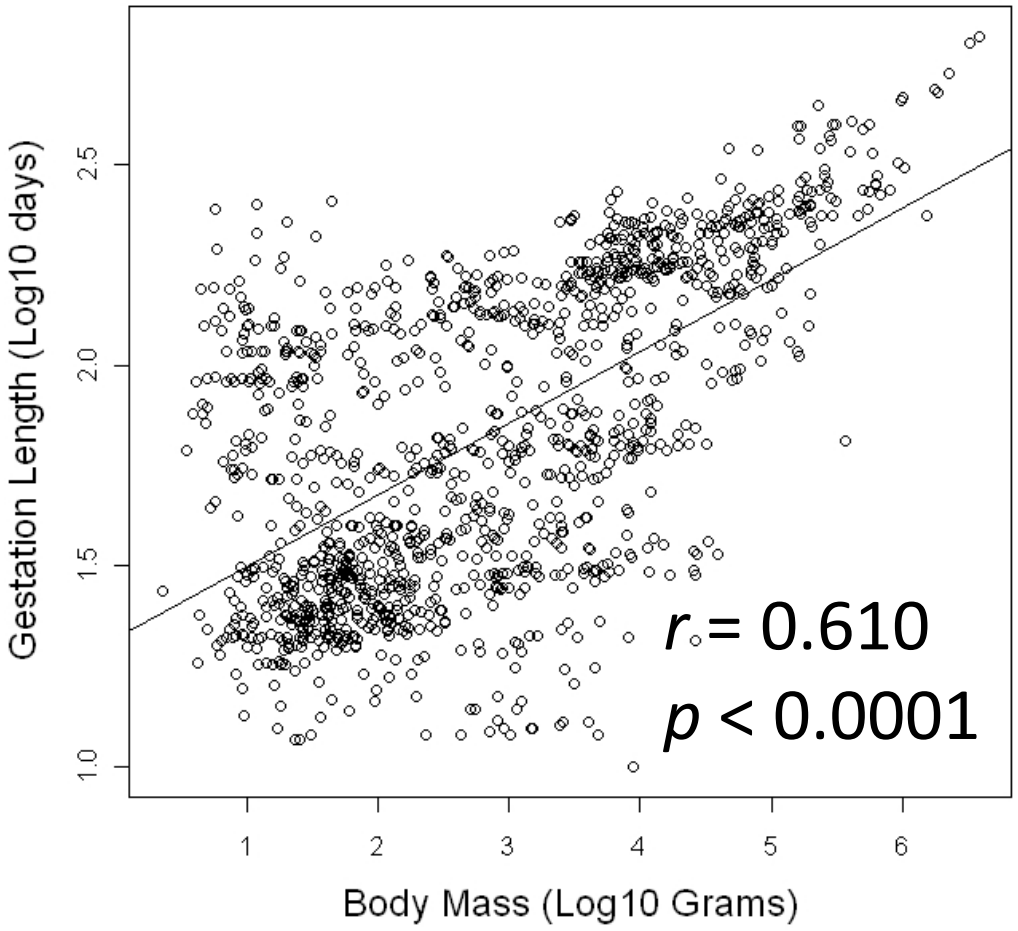
10,000 random correlations: A null distribution



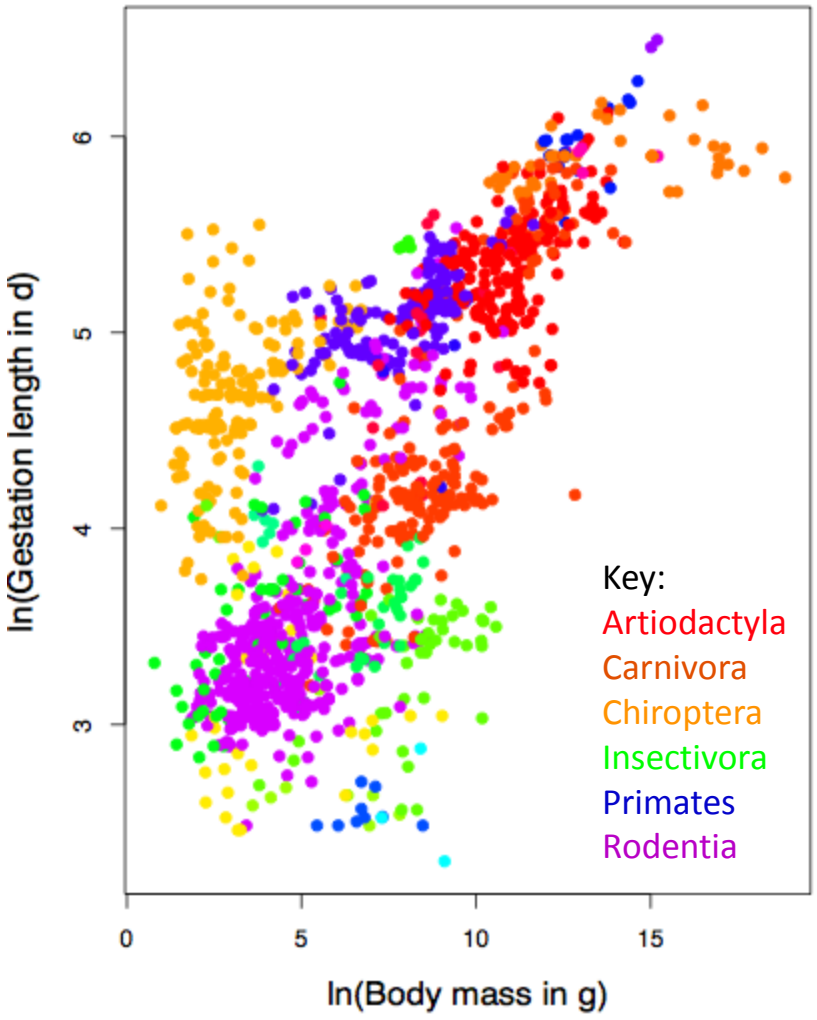
- H_0: The correlation that we observe is due to chance alone.

- H_1: The observed correlations is not due to chance alone.

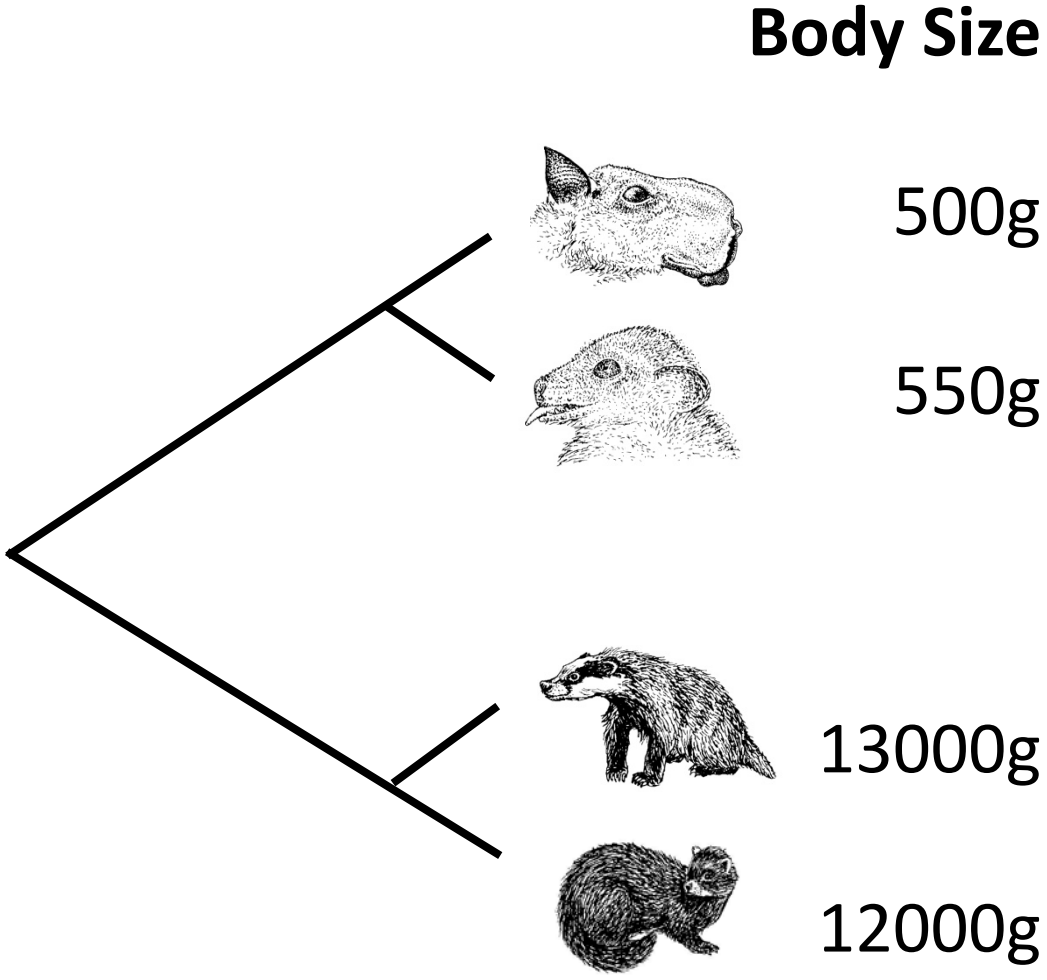
Regression Analysis



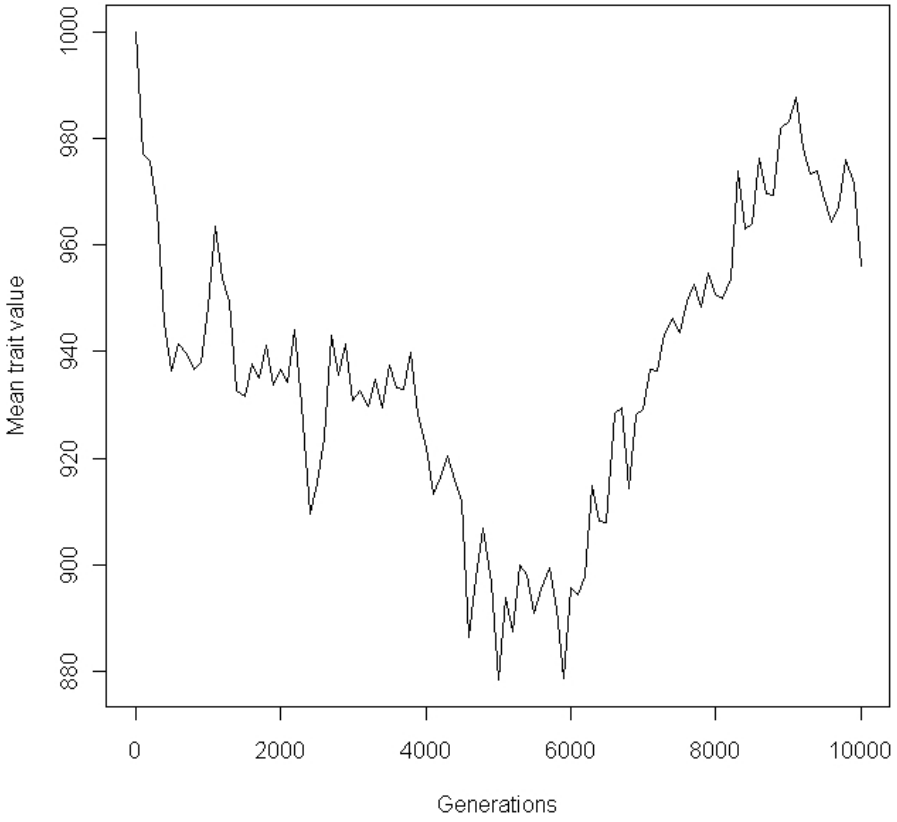
Closely related species have similar traits



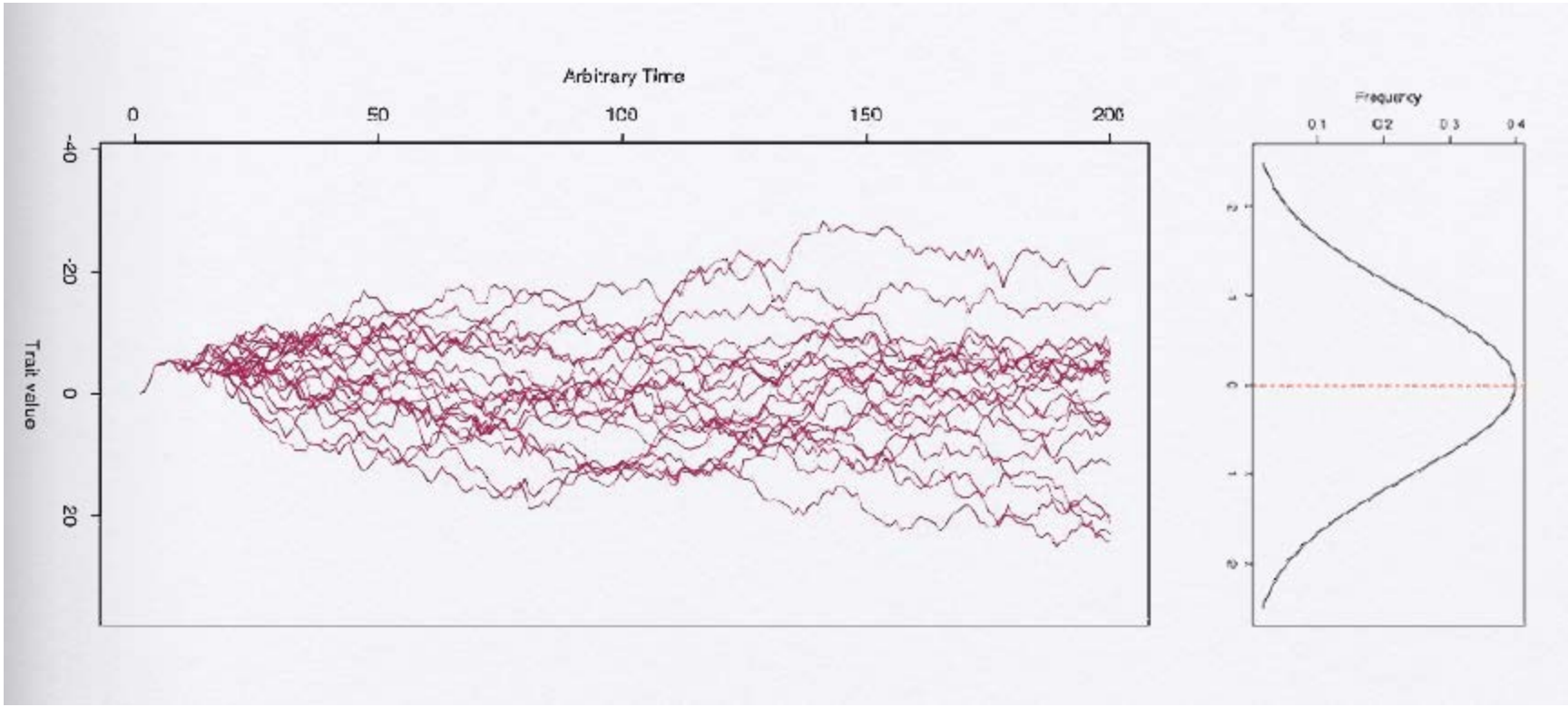
Closely related species have similar traits

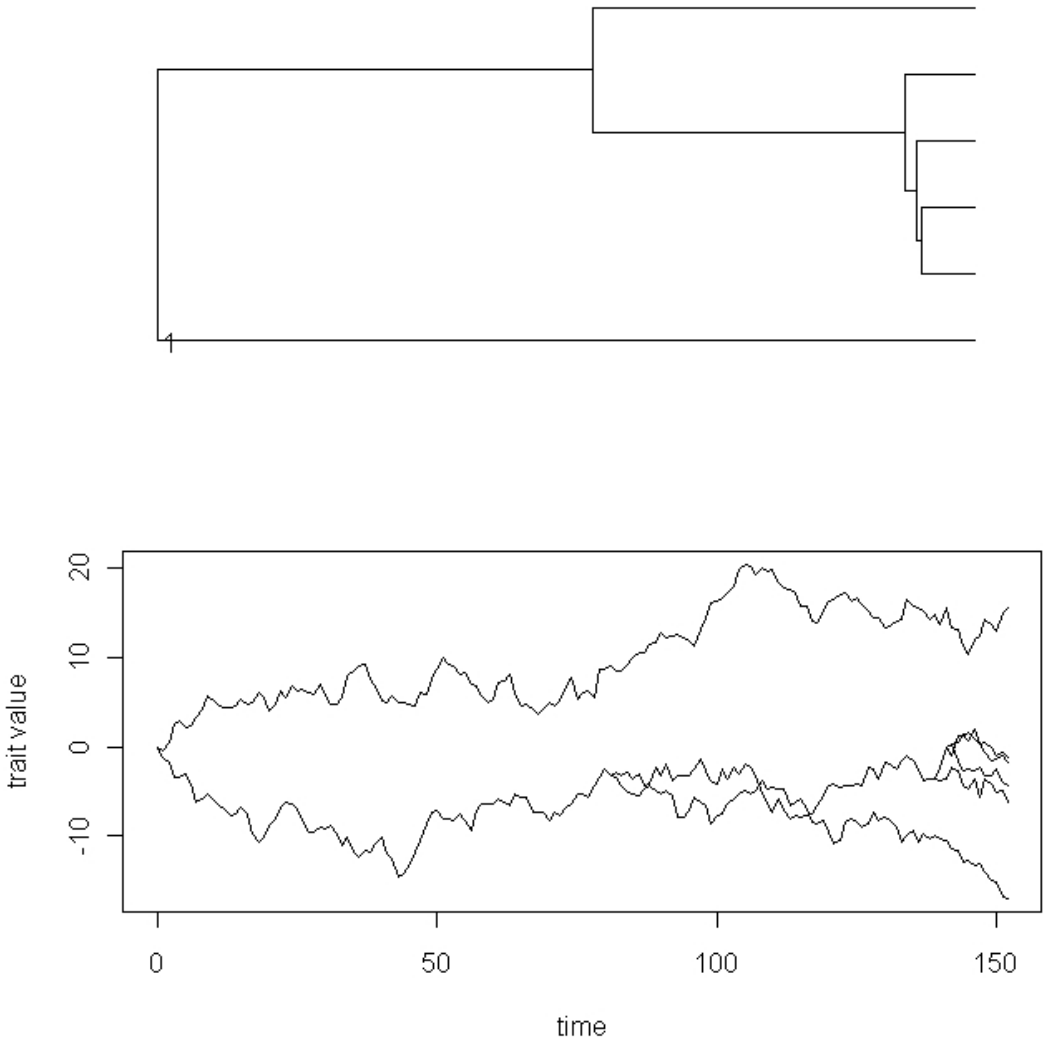


Brownian Motion Model of Trait Evolution

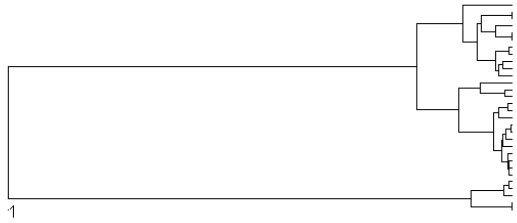


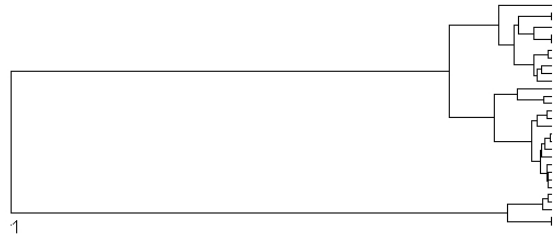
Generates normal character distributions

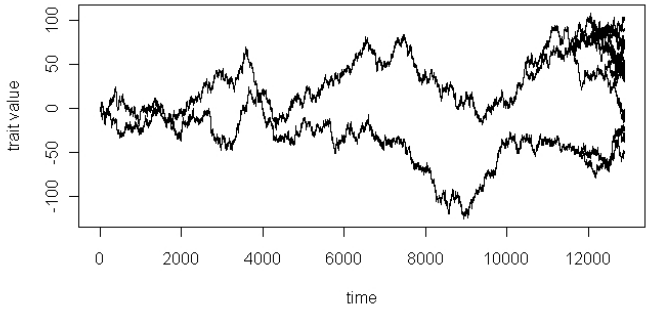


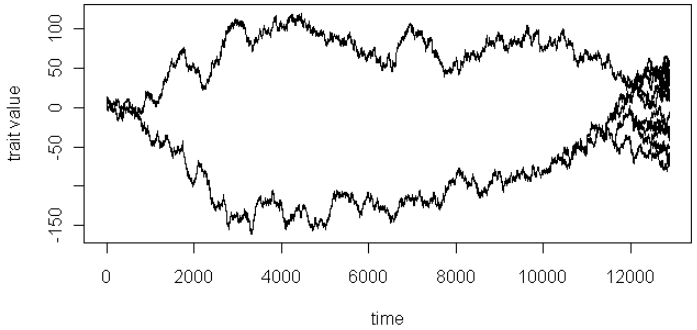

1
trait value
20
10
0
-10
0
50
100
150
time

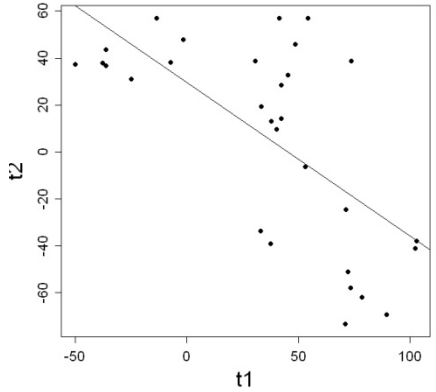
Correlation between randomly evolved characters



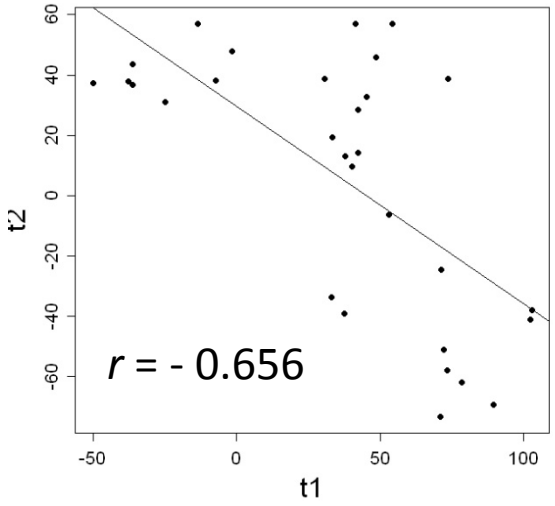


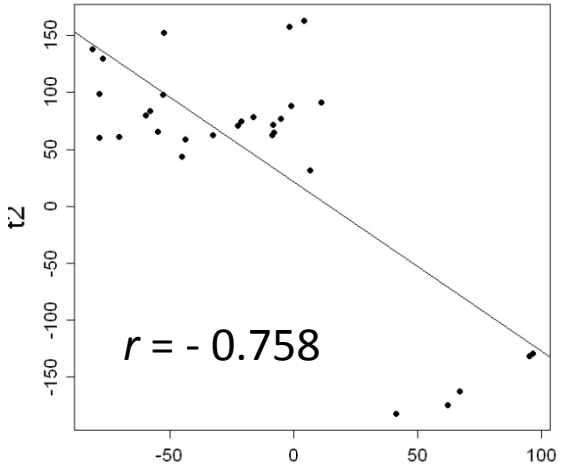


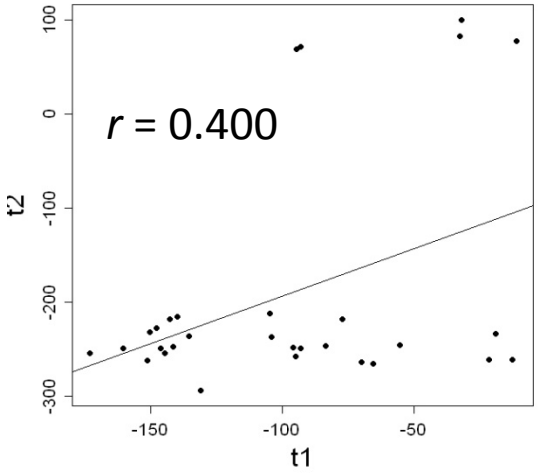


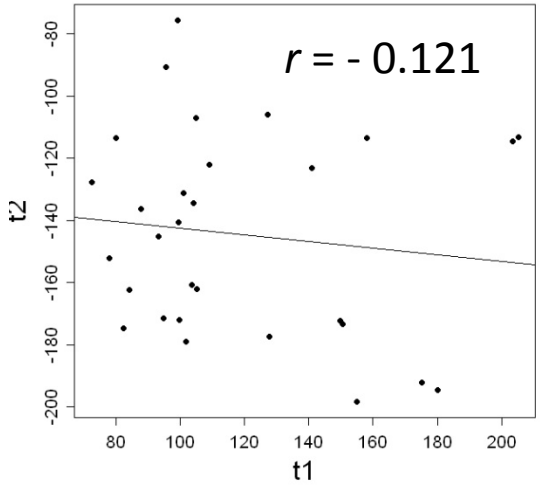


A series of correlations

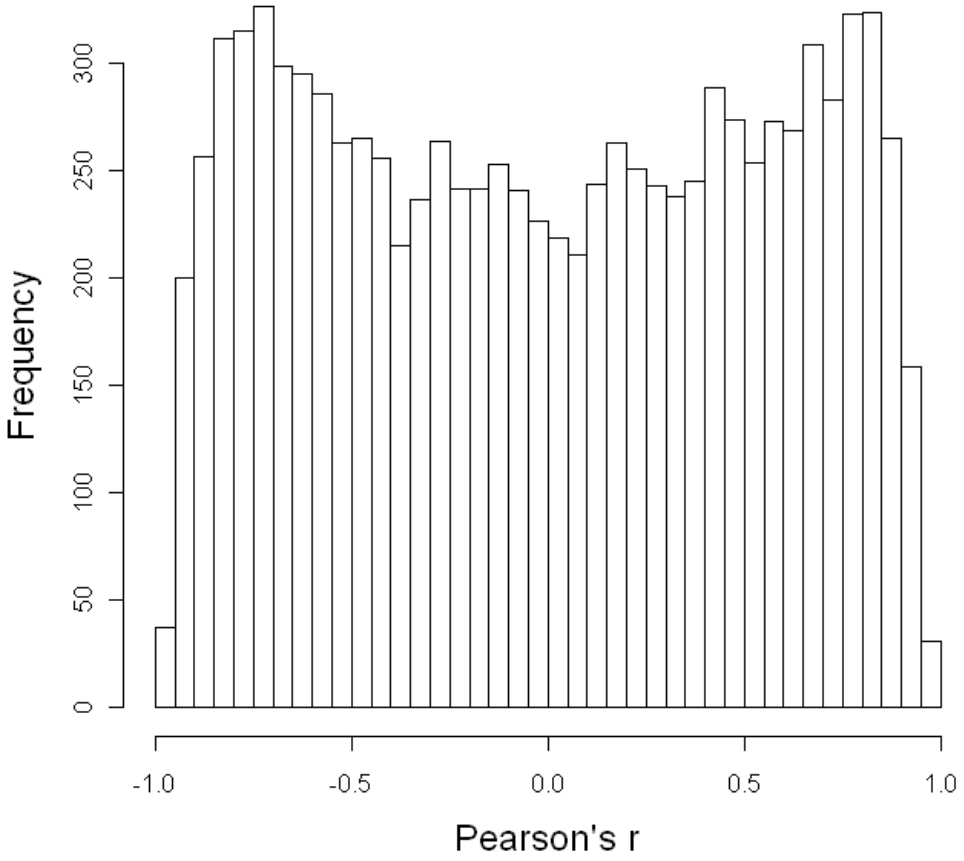




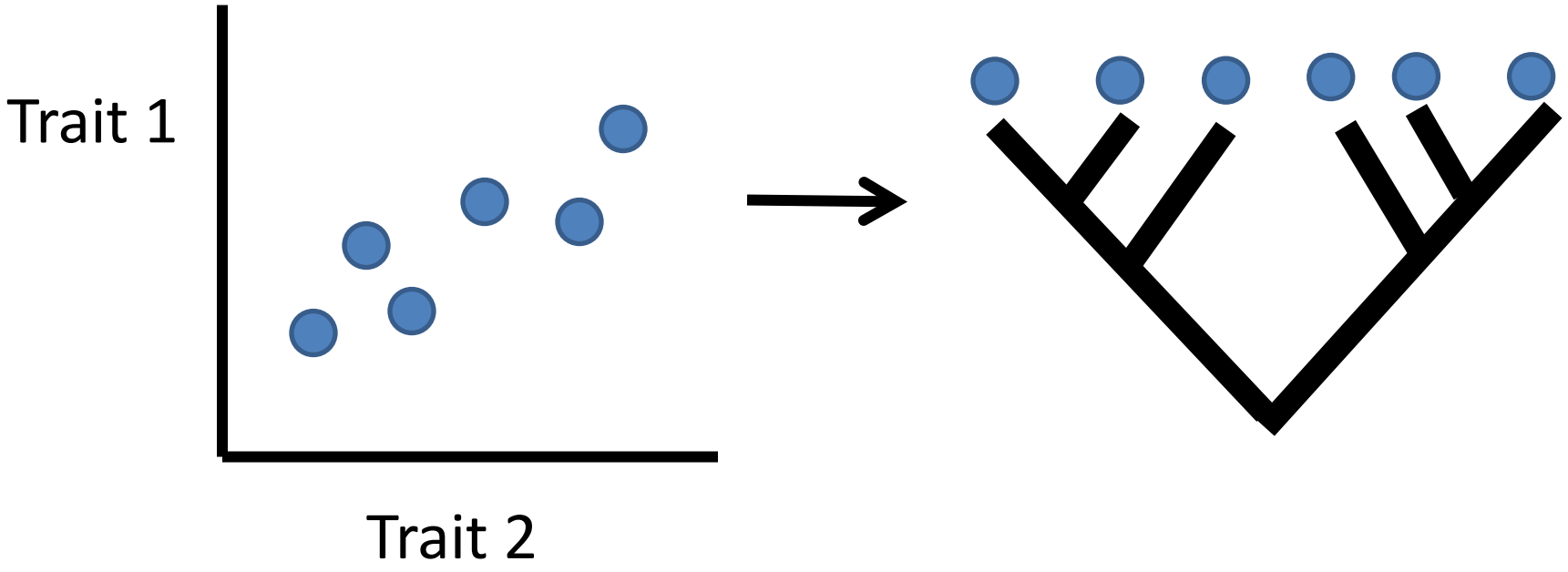




Correlations among 10,000 pairs of randomly evolved traits



Summary of the problem



Summary of the problem