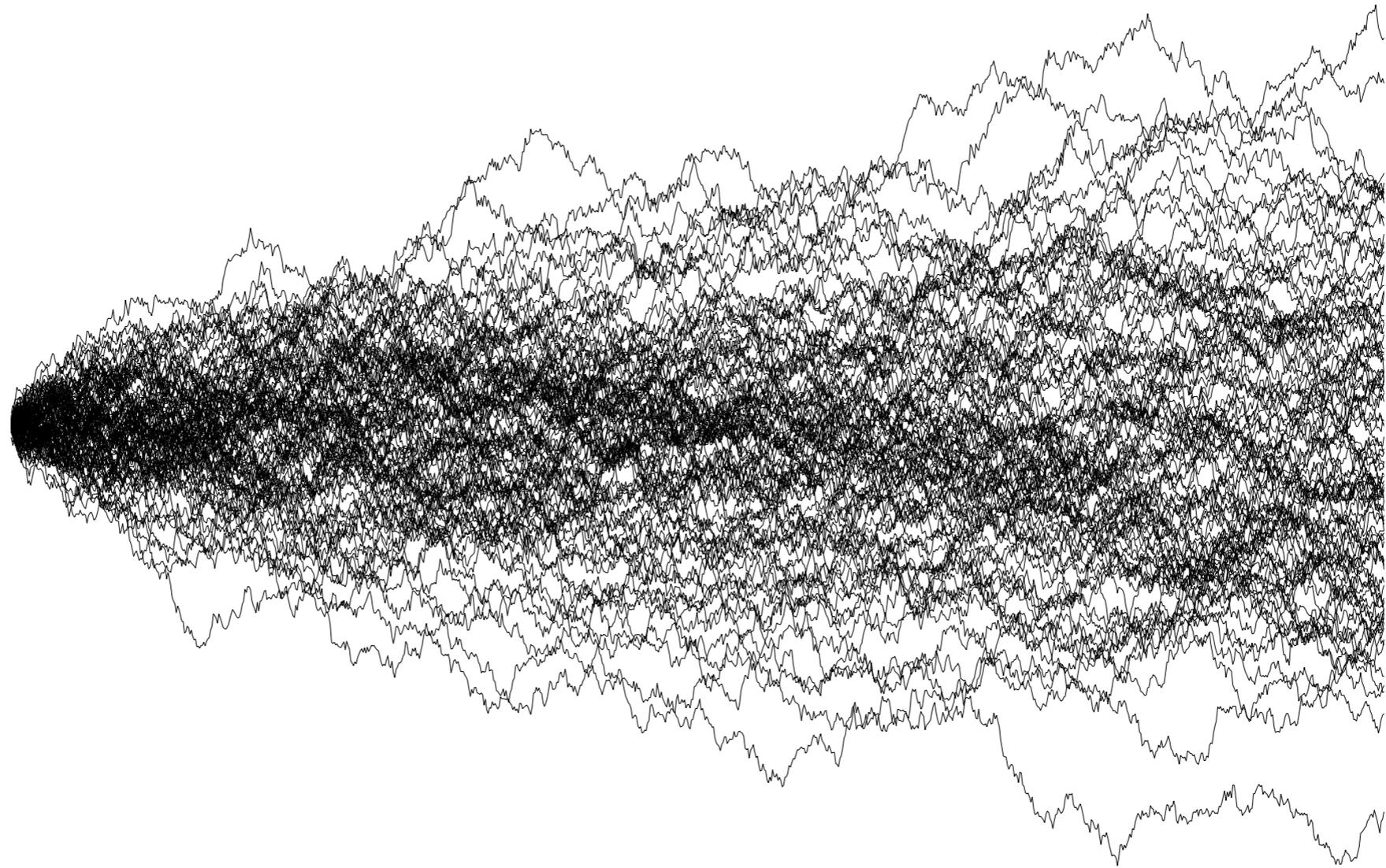


Brownian Motion



Brownian Motion

- A model for the evolution of continuously-valued characters
- States change continuously through time
- After some time, expected character states follow a normal distribution

Outline - BM

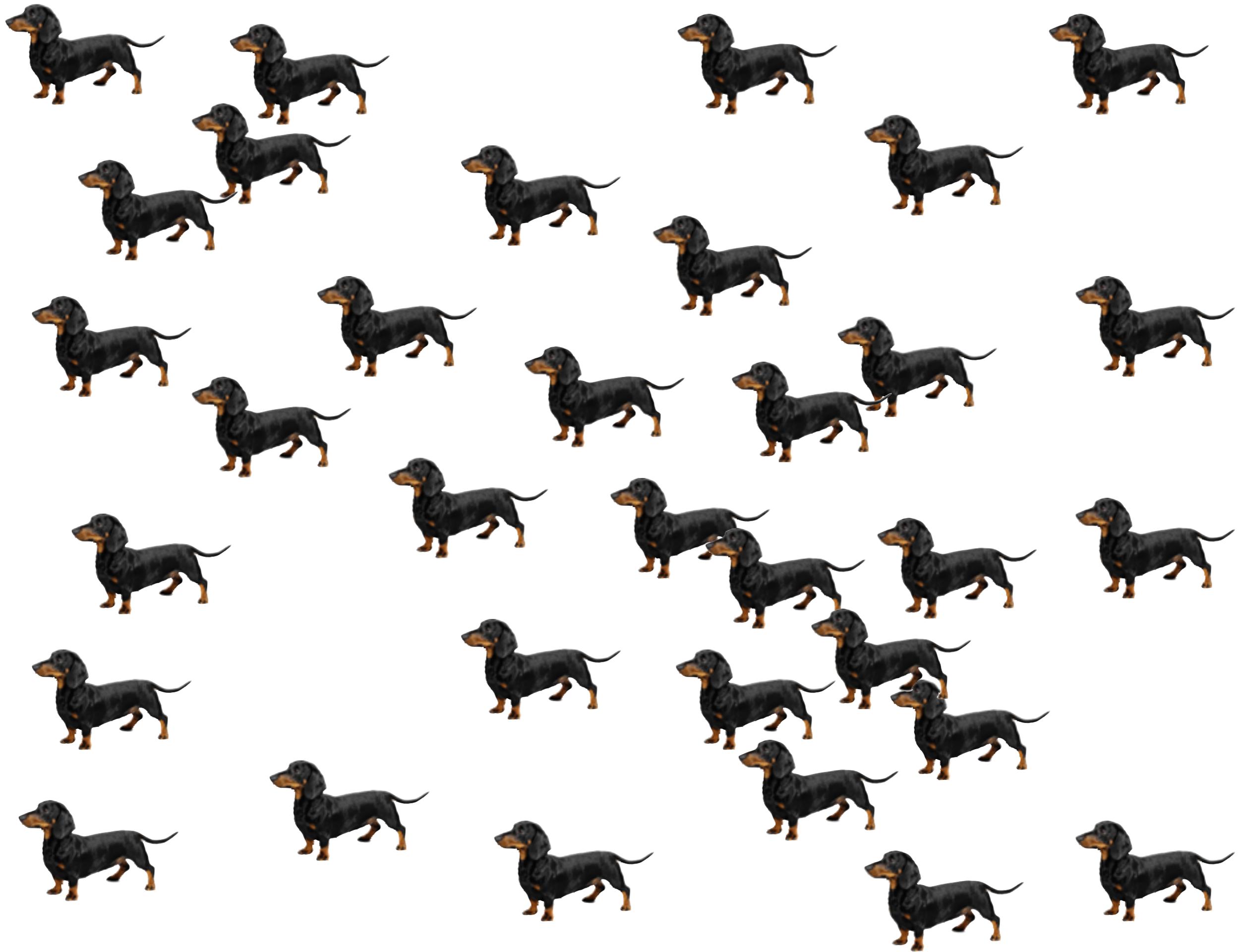
- What is Brownian motion?
- When might characters evolve in a Brownian-like way?
- Simulating Brownian motion on trees

Outline - BM

- What is Brownian motion?
- When might characters evolve in a Brownian-like way?
- Simulating Brownian motion on trees

Brownian Motion: The Model

- A continuous-time stochastic process
- Describes a “random walk” of evolution for continuously-valued characters
- Sometimes called the Wiener process



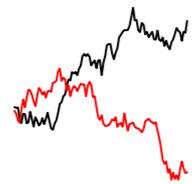
Three Facts Describe Brownian Motion

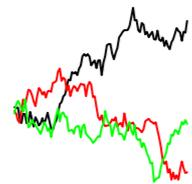
- Let $W(t)$ be the value of the character at time t . Then:
 - $E[W(t)] = W(0)$
 - Successive steps are independent
 - $W(t) \sim N(W(0), \sigma^2 t)$

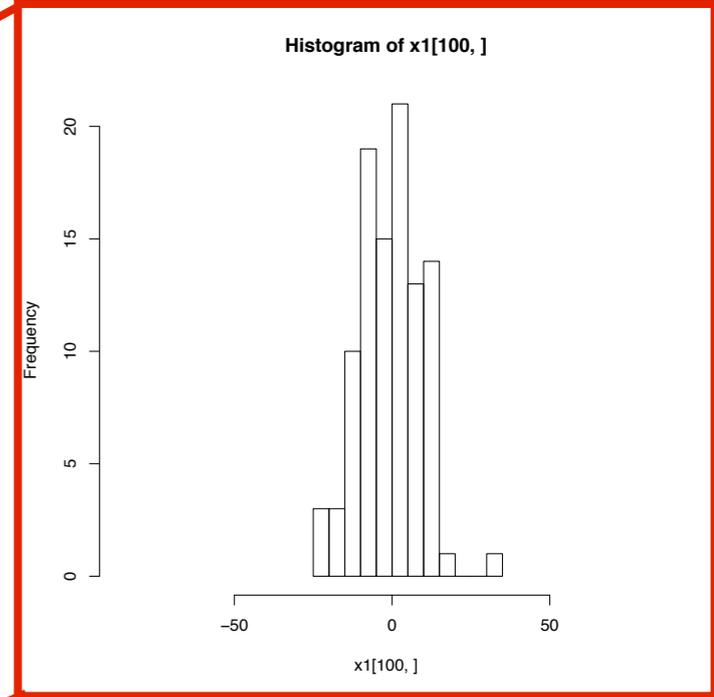
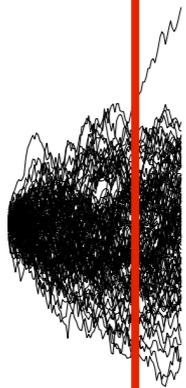
Parameters of BM

- Brownian motion models have two parameters:
 - Θ , the starting value; $W(0) = \Theta$
 - σ^2 , the rate parameter

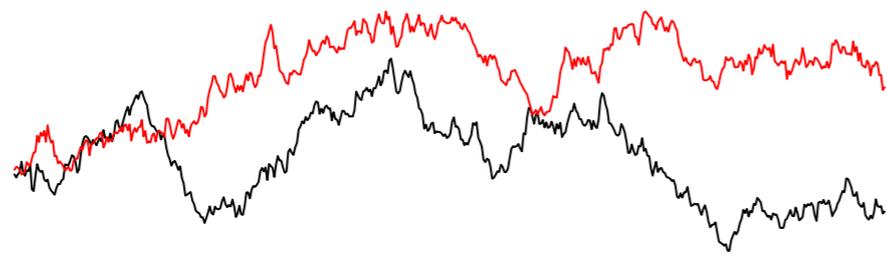
A small, jagged, black scribble or mark on a white background. The mark consists of several sharp, irregular peaks and valleys, resembling a stylized signature or a random noise pattern. It is located in the lower-left quadrant of the image.

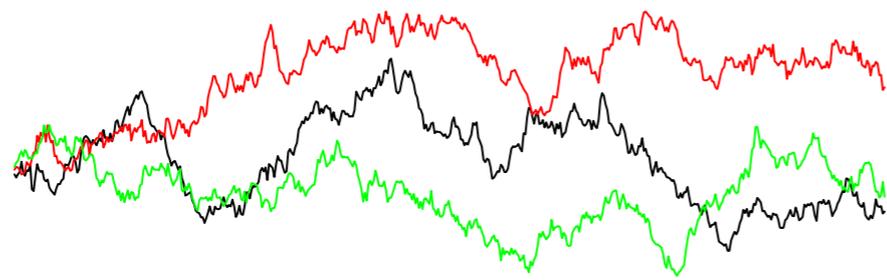


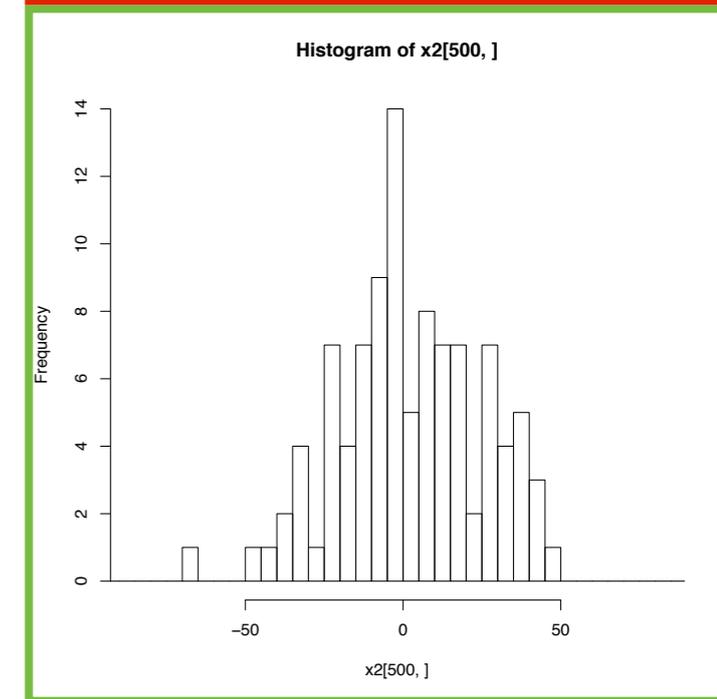
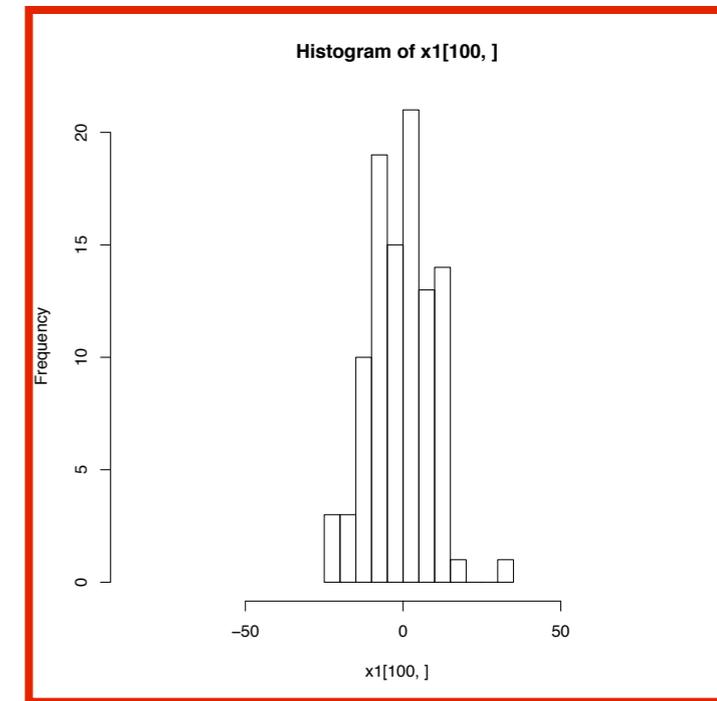
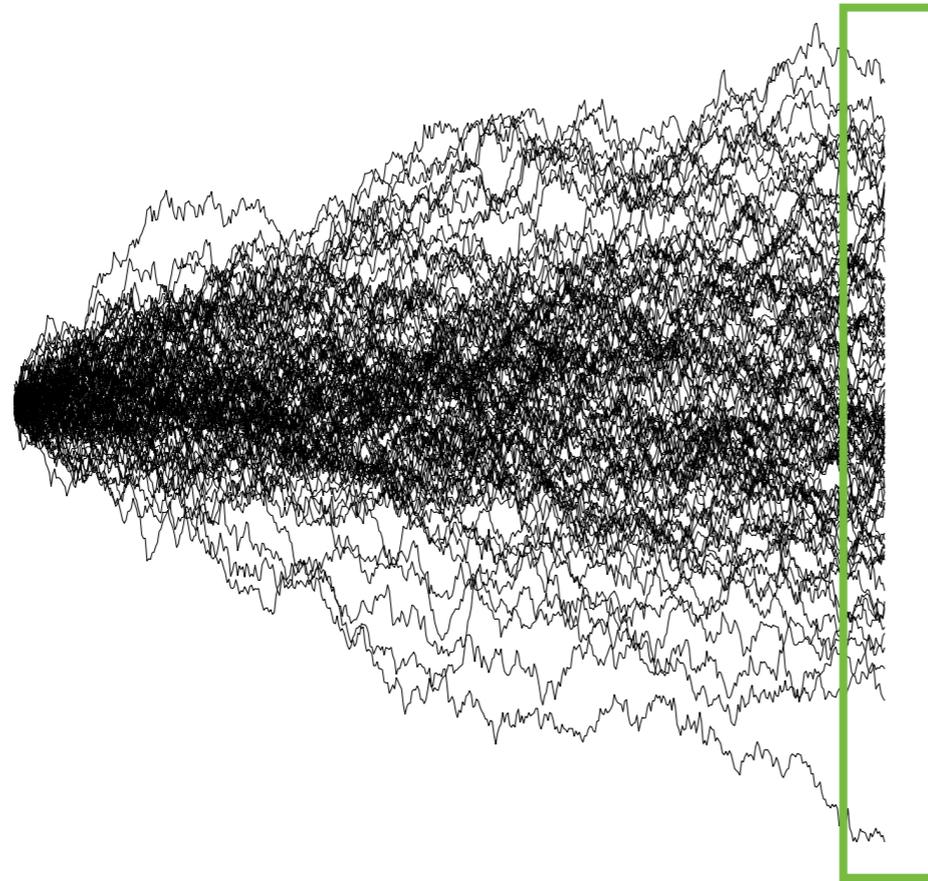


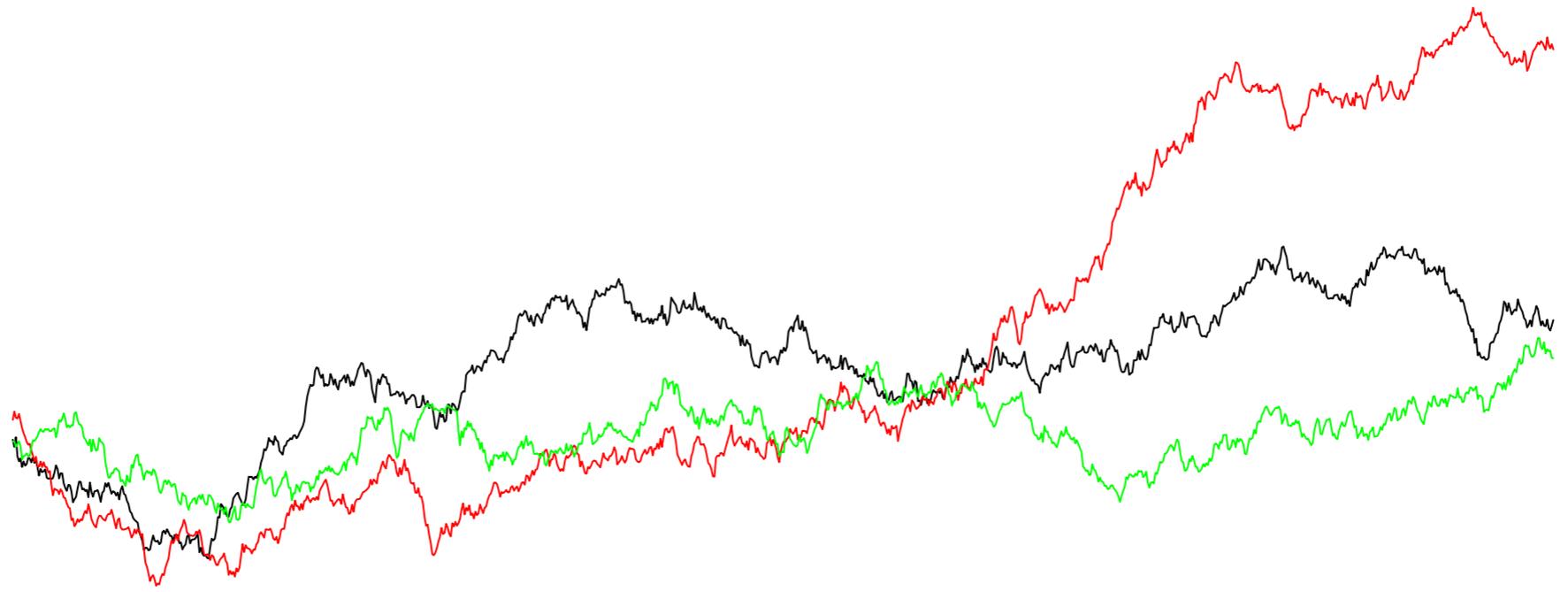


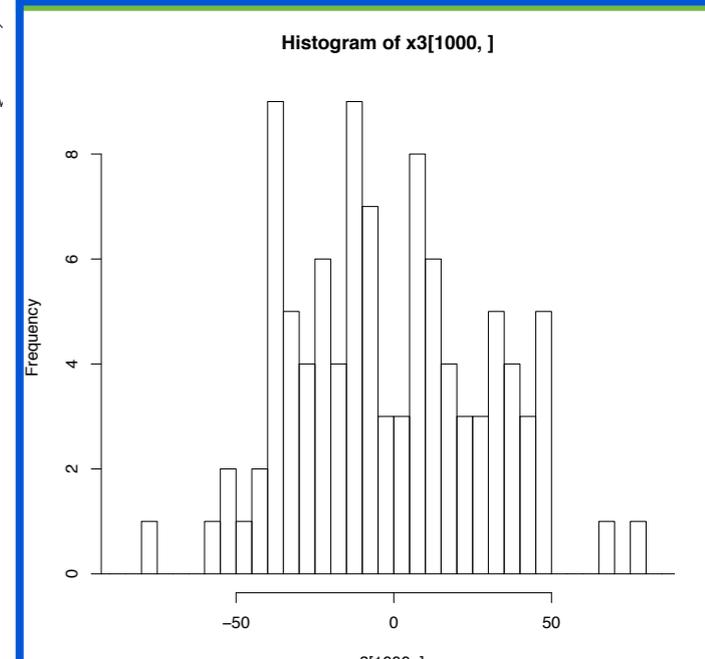
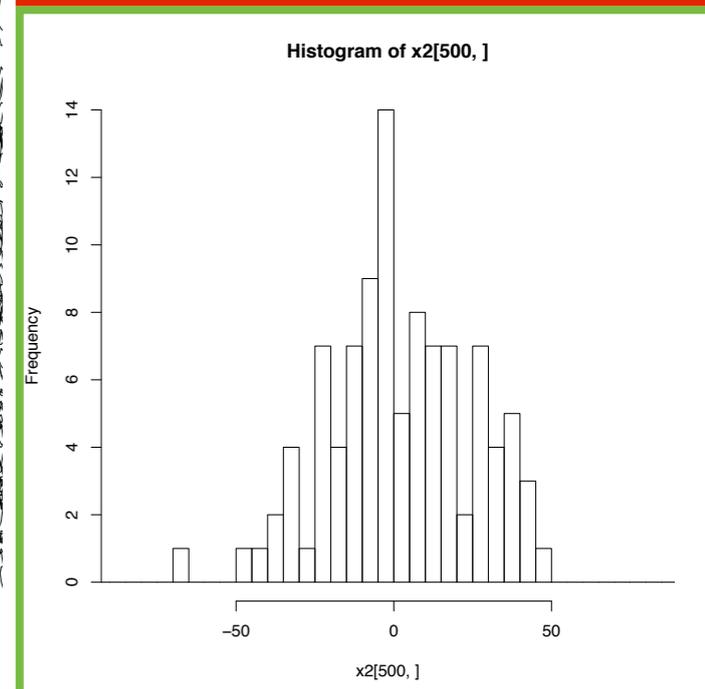
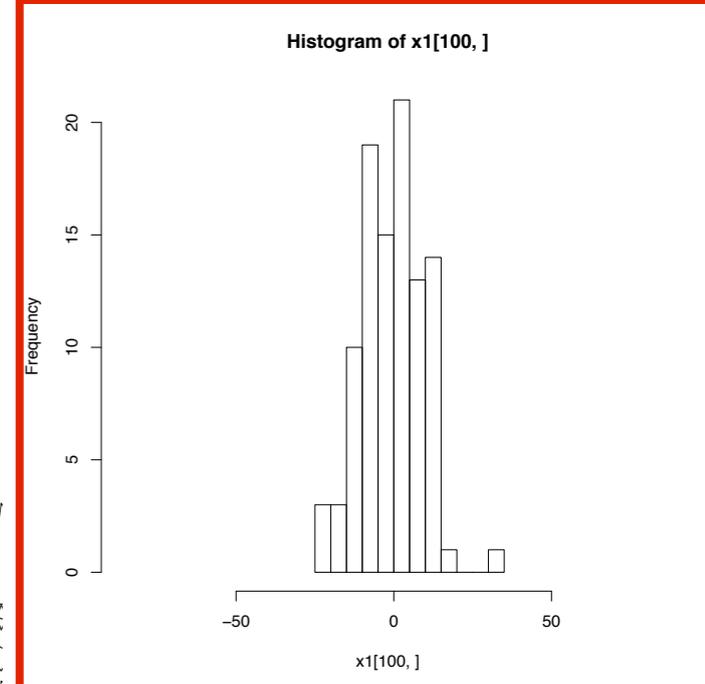
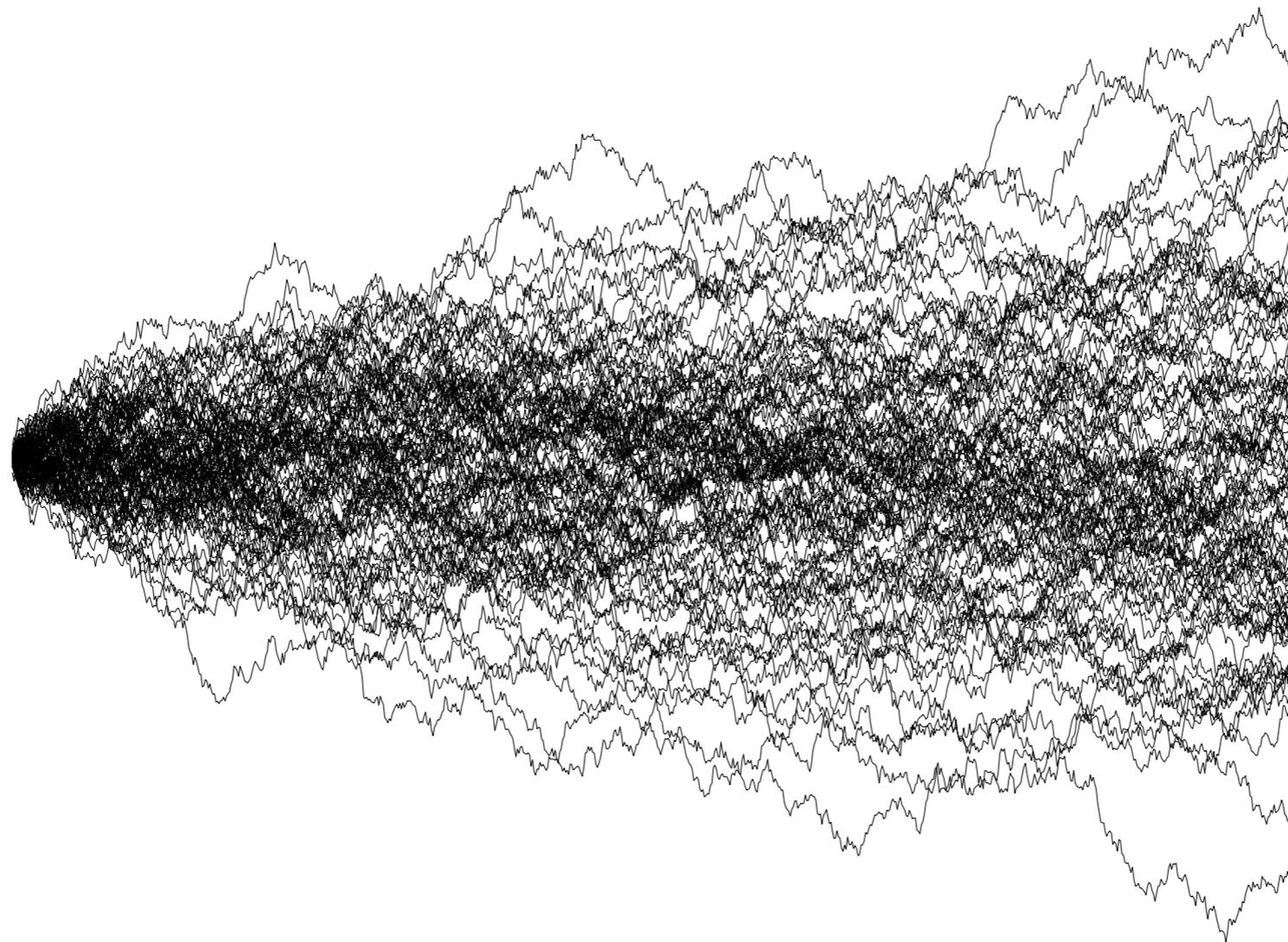




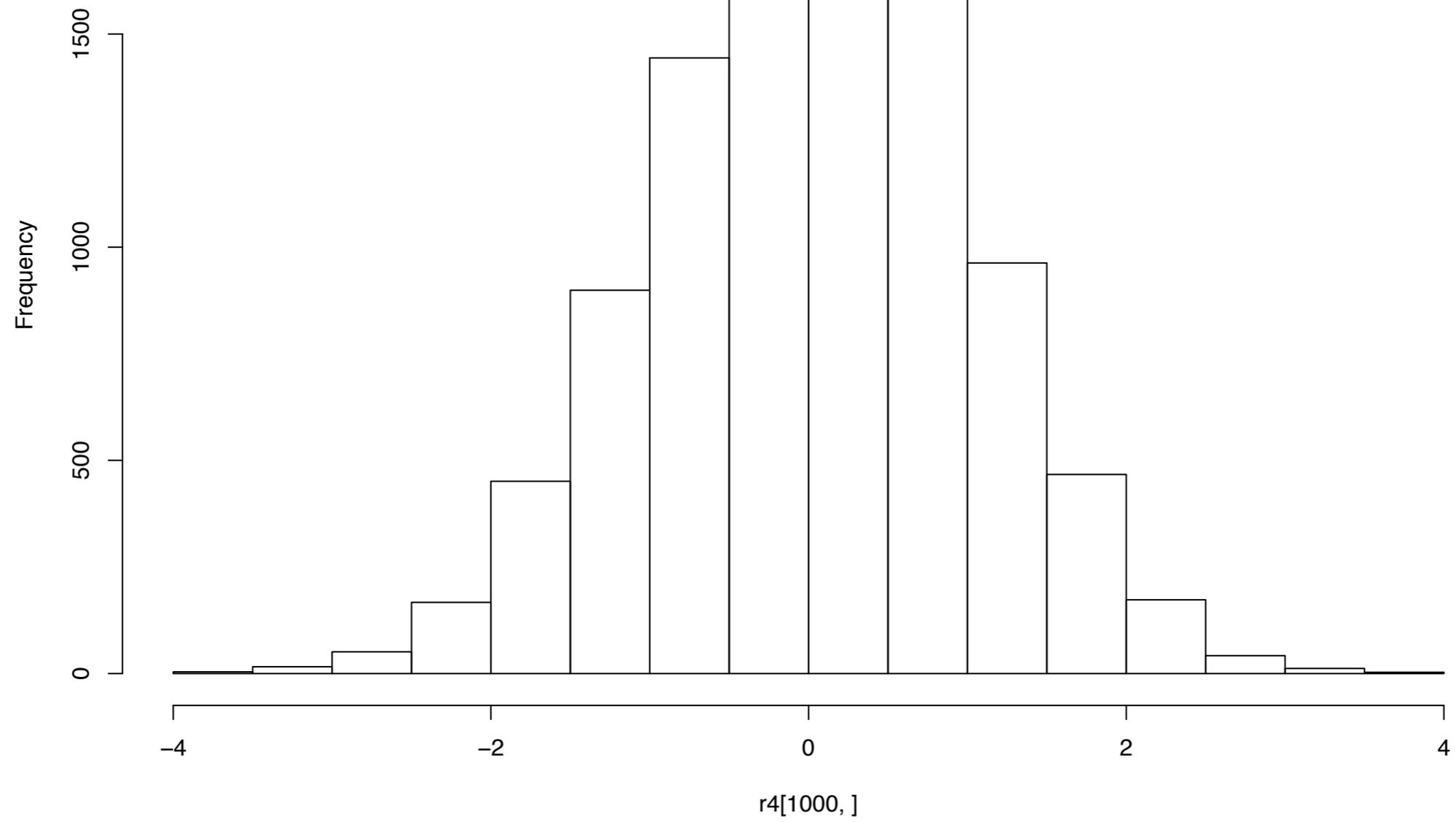


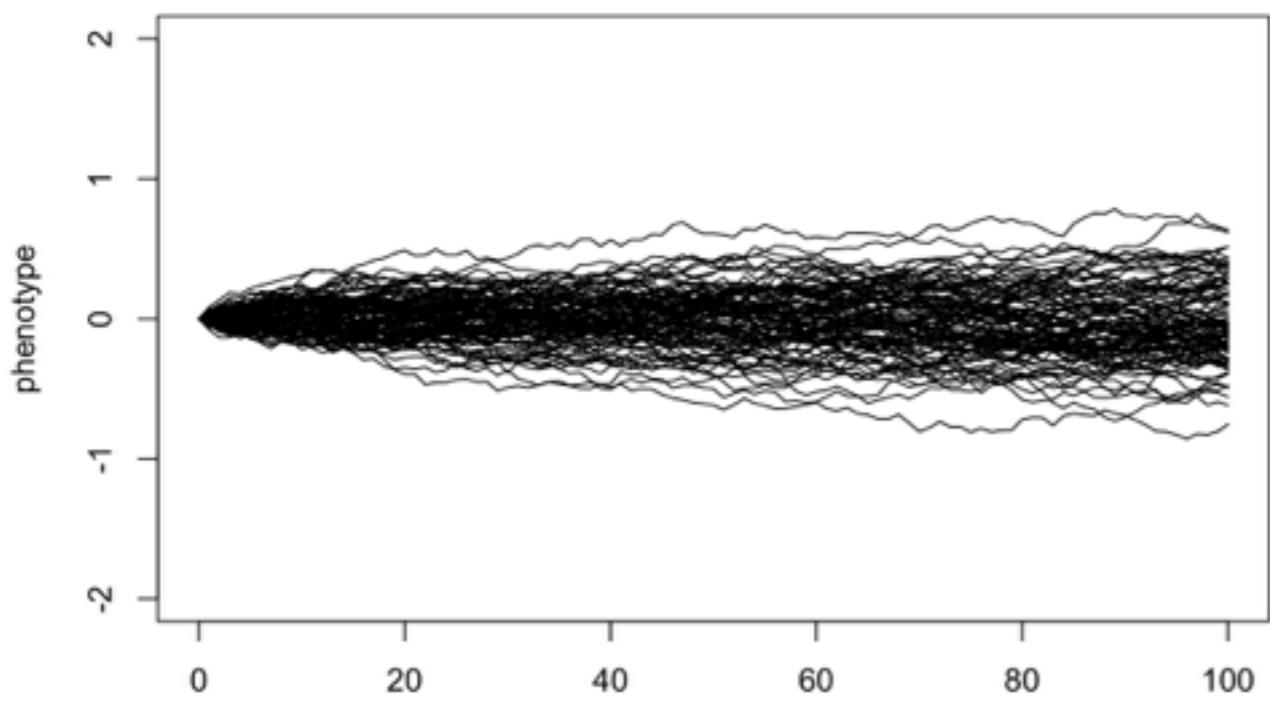




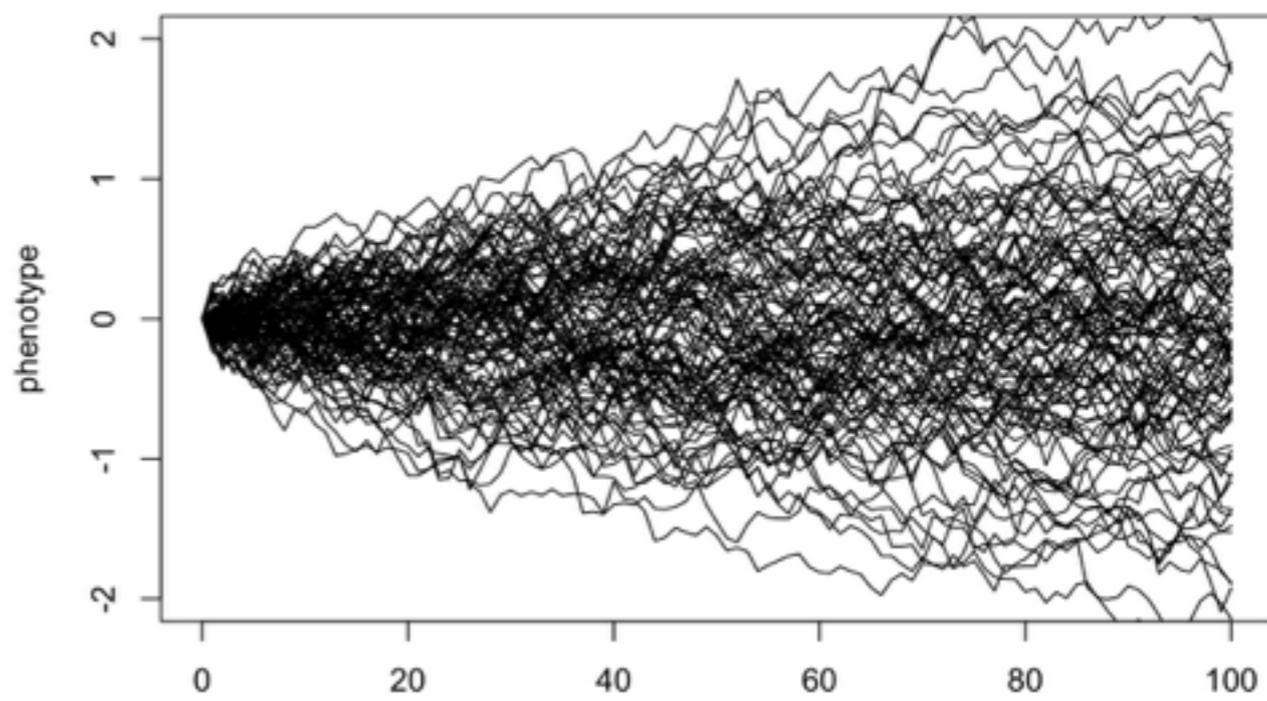


Histogram of r4[1000,]



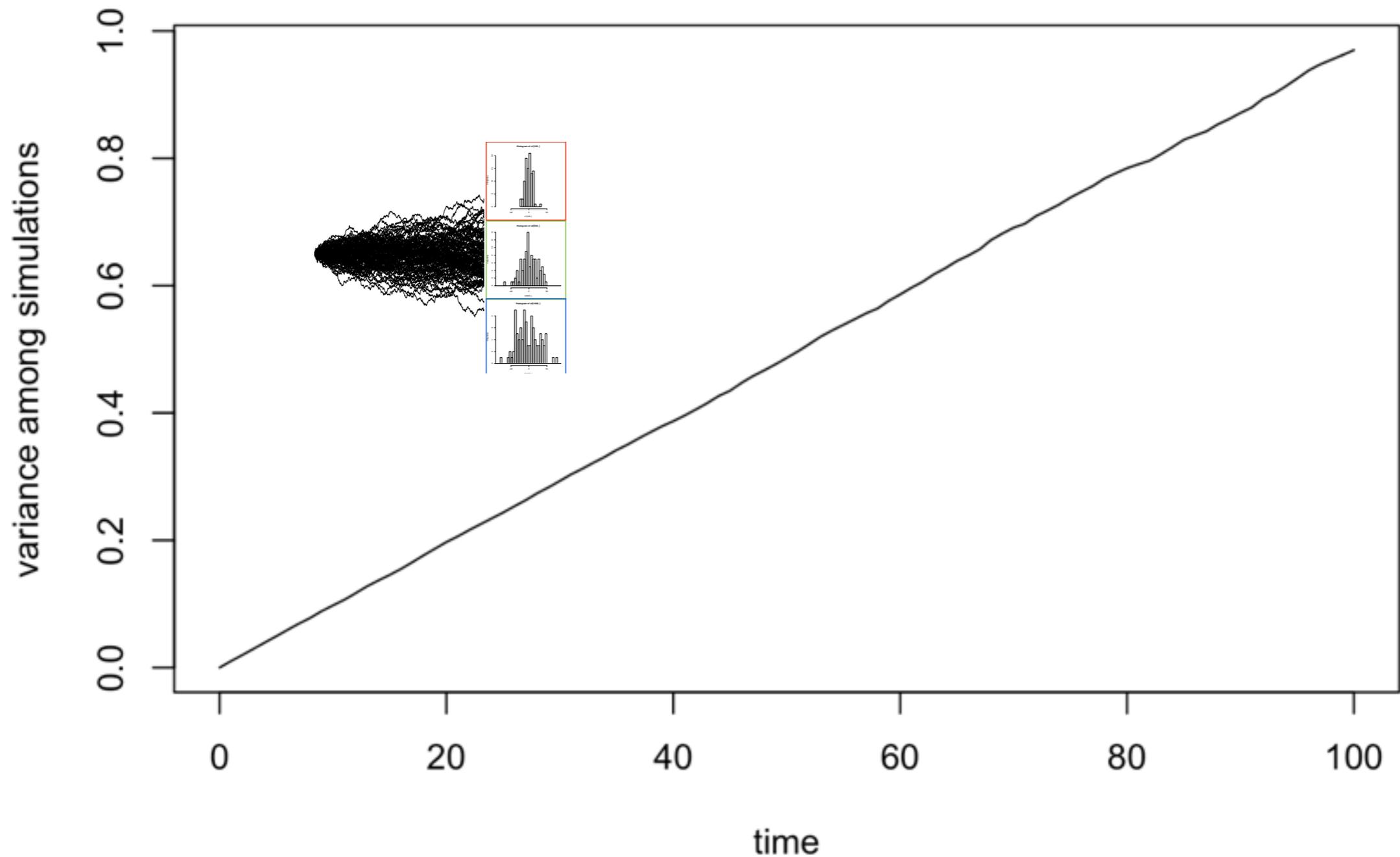


$$\sigma^2 = 0.01$$



$$\sigma^2 = 0.1$$

variance increases with time



$$\text{Var}(X) = \sigma^2 * \textit{time}$$

A physical model for BM



Why Normal?

- BM can be used to describe motion that results from the combination of a large number of independent weak forces
- Adding many small independent variables result in normal distributions, no matter the original distribution (Central limit theorem)

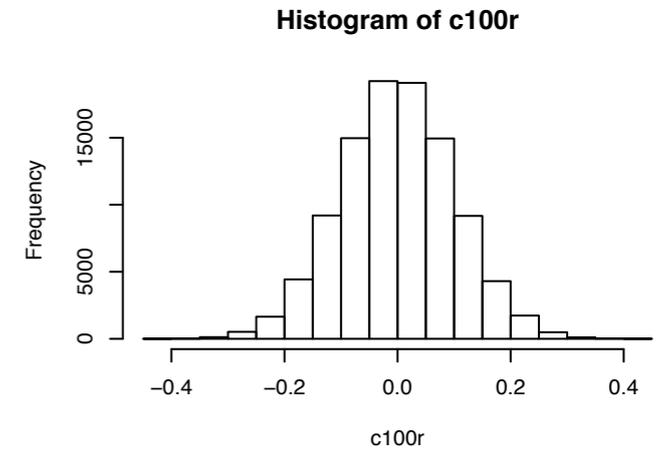
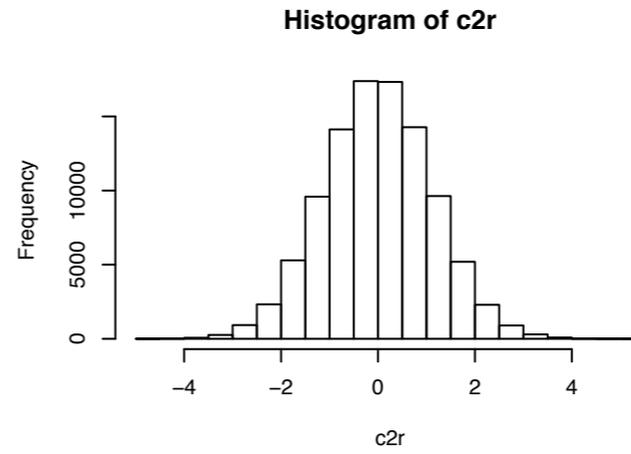
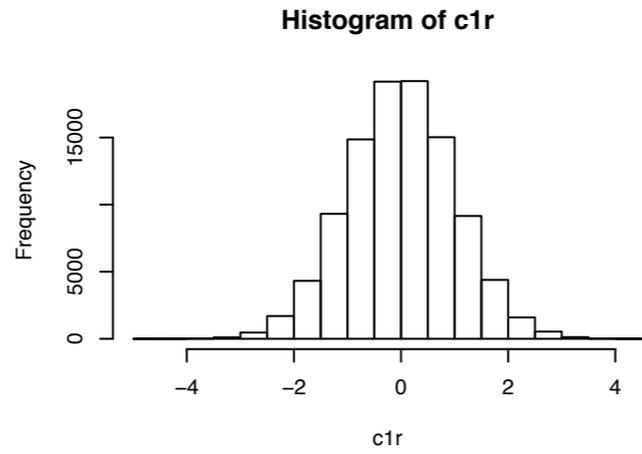
Mean of

$n=1$

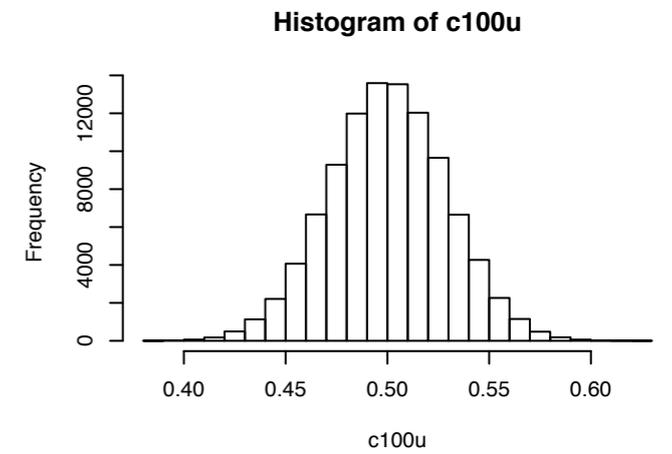
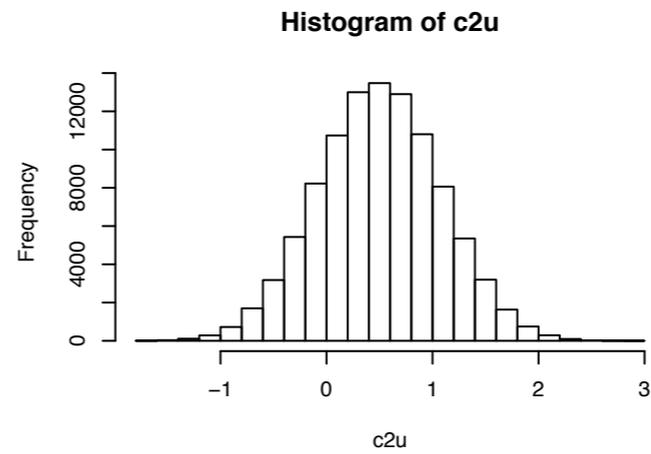
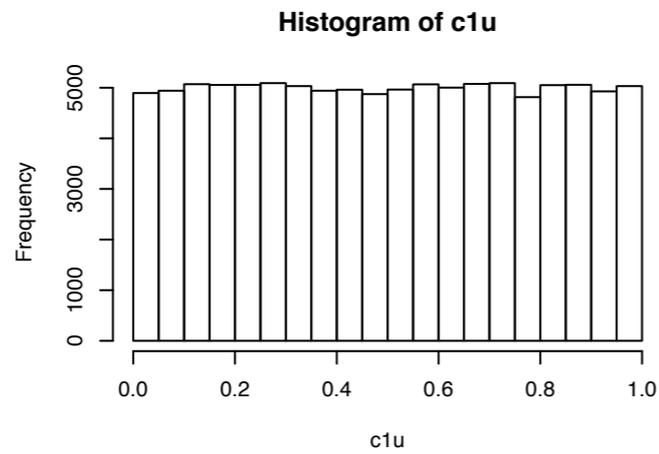
$n=2$

$n=100$

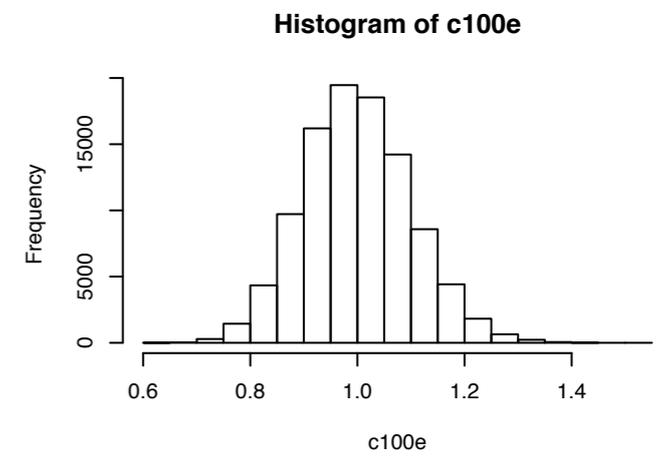
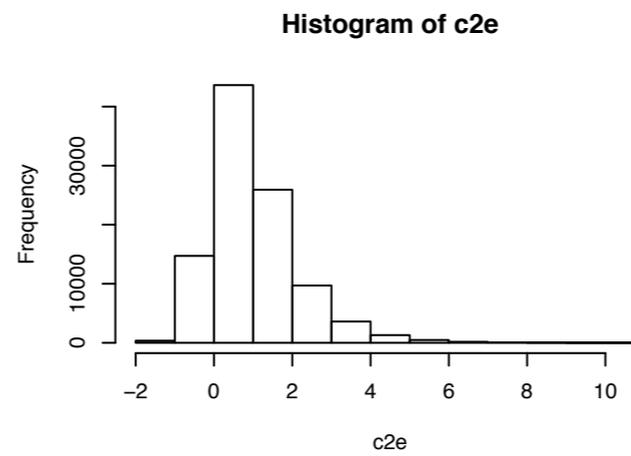
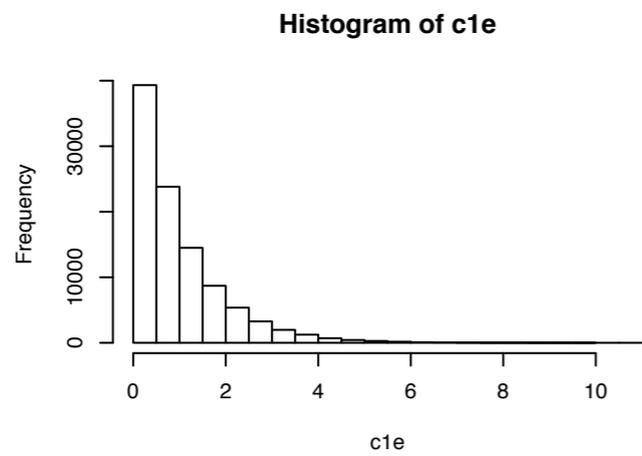
Normal



Uniform



Exponential

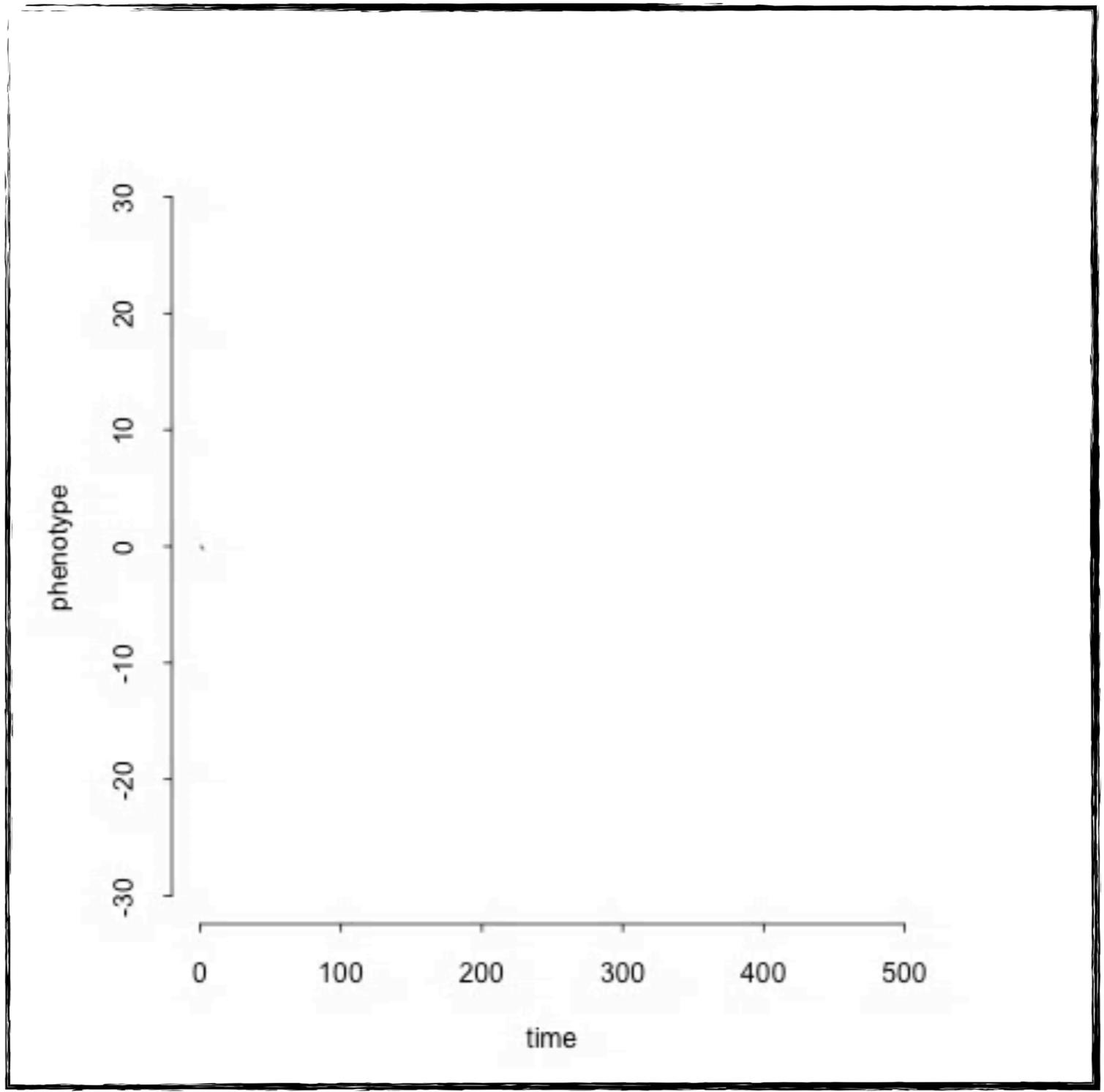


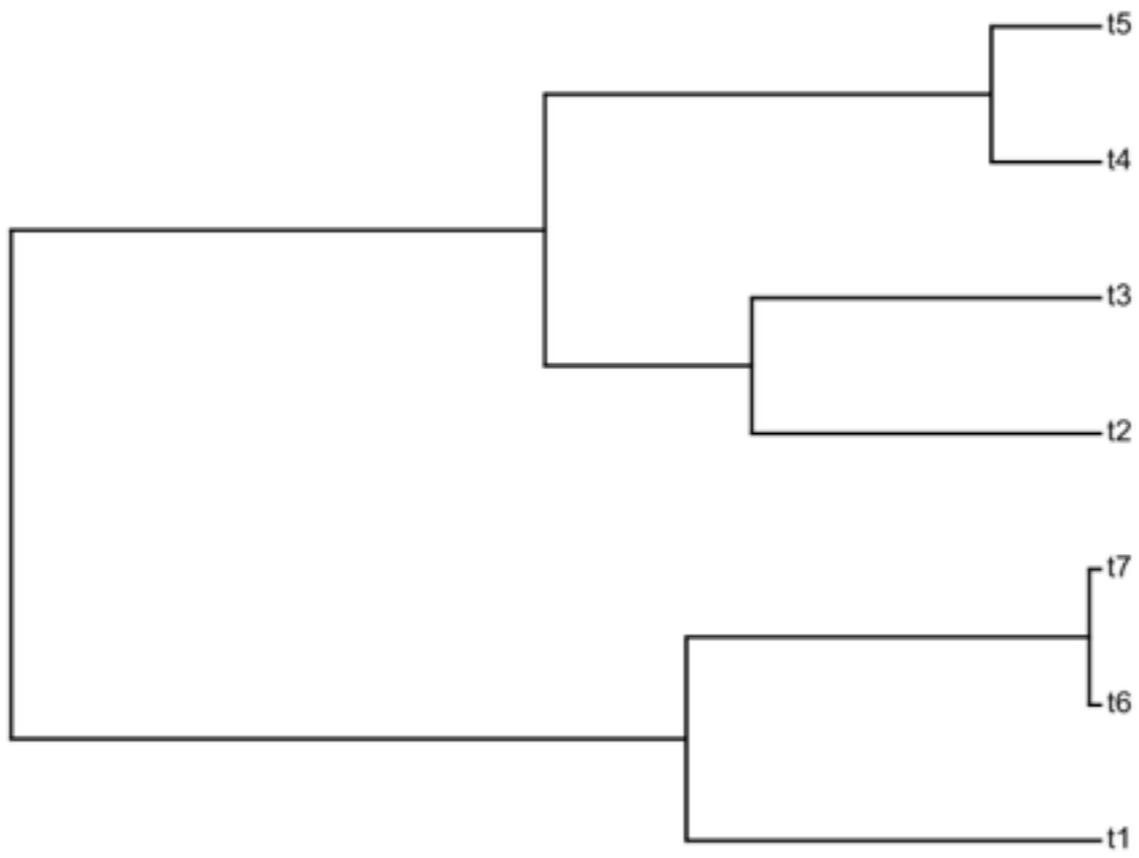
Evolution might approximate BM...

- Genetic drift
- Random punctuated change
- Selection that is weak relative to the time interval considered
- Selection that changes randomly through time

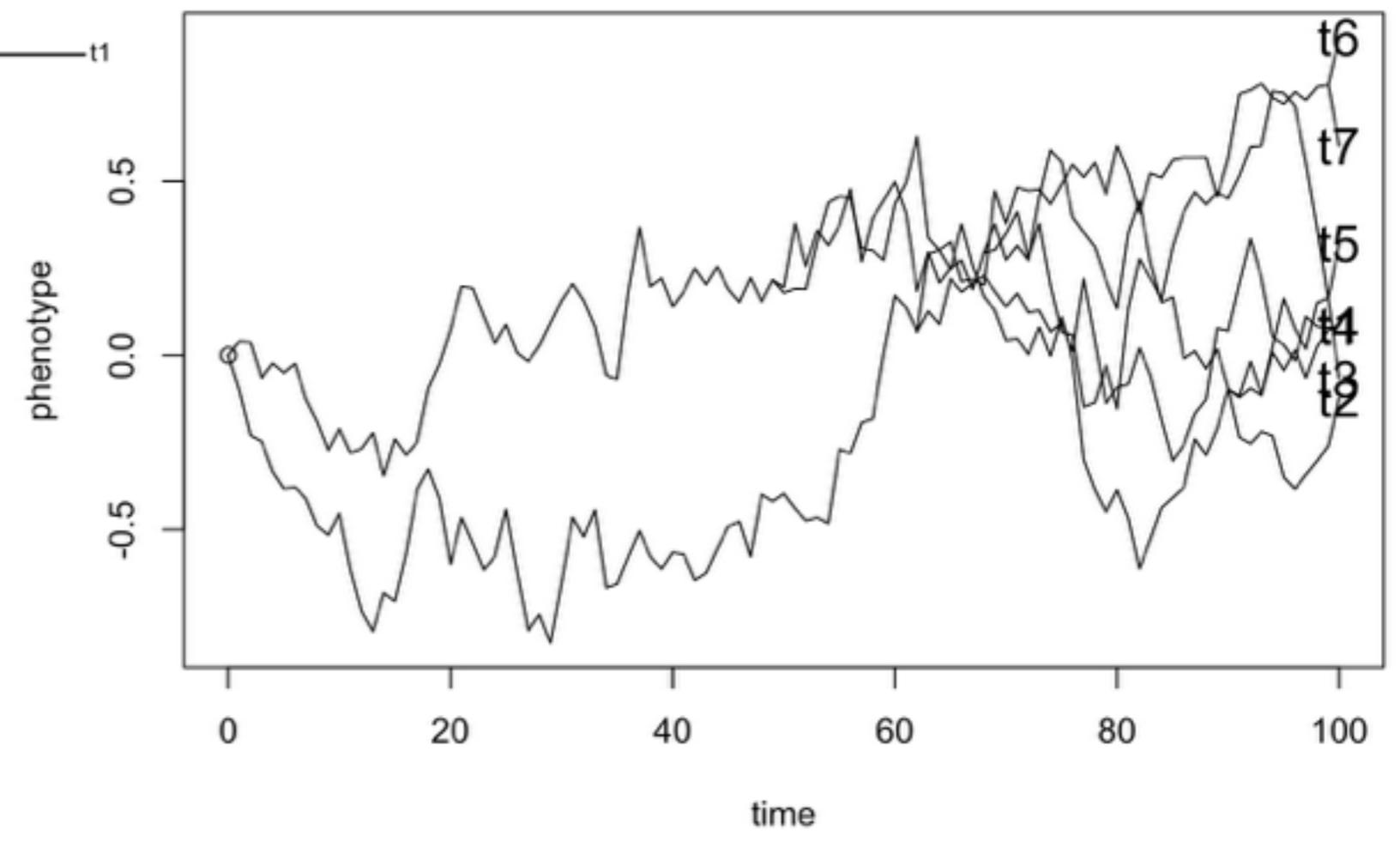
Outline - BM

- What is Brownian motion?
- When might characters evolve in a Brownian-like way?
- Simulating Brownian motion on trees





<http://tinyurl.com/macro-2016-p3>

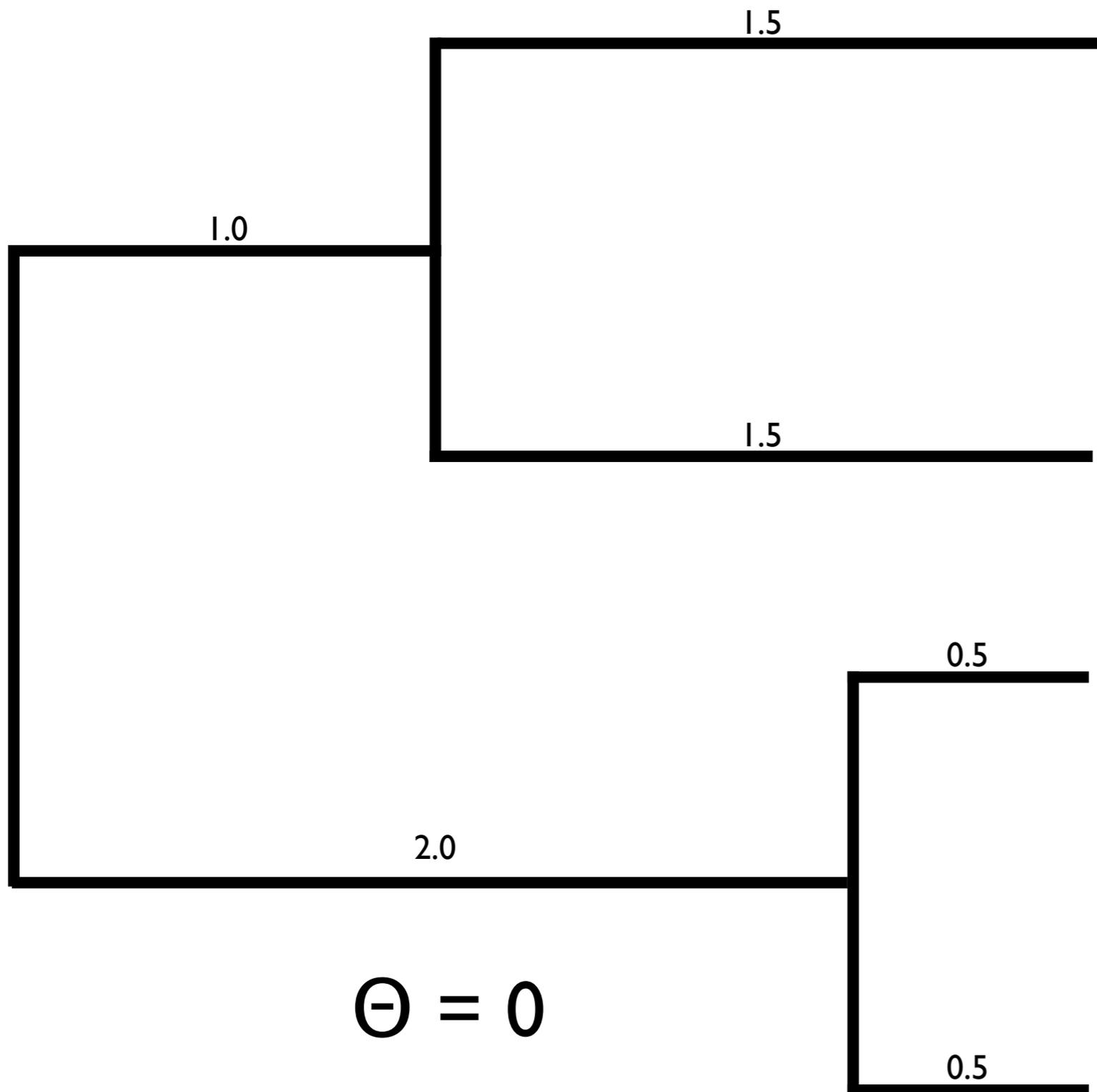


Explain how the phylogeny influences patterns of phenotypic evolution.

Would you expect more or less diversity in phenotype if the species were evolving completely independently?

Simulating BM

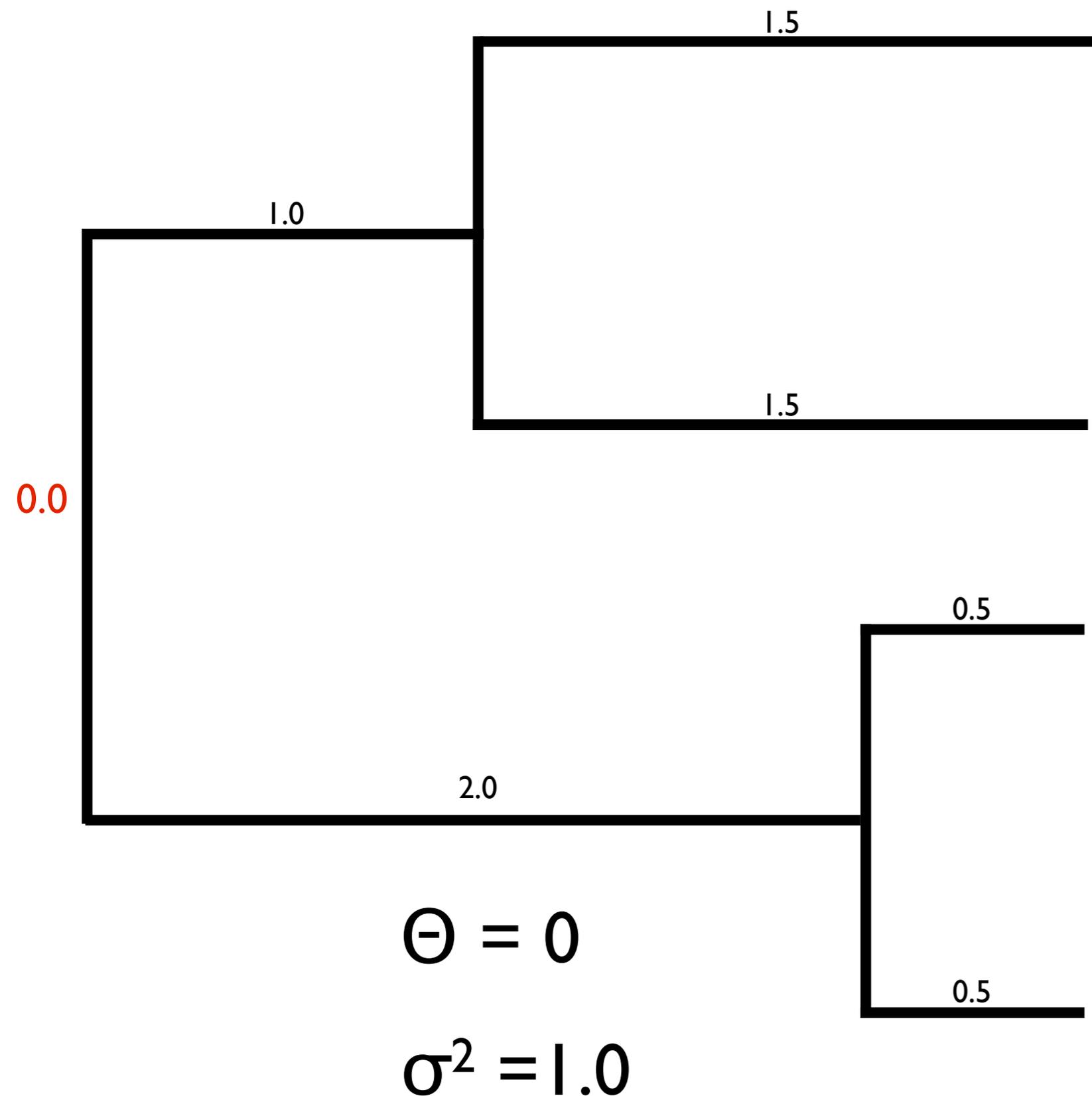
- Simulating Brownian motion involves drawing values from normal distributions
- Variance of the distribution depends on σ^2 and t
- Values along adjacent branches are added from the root to the tips of the tree



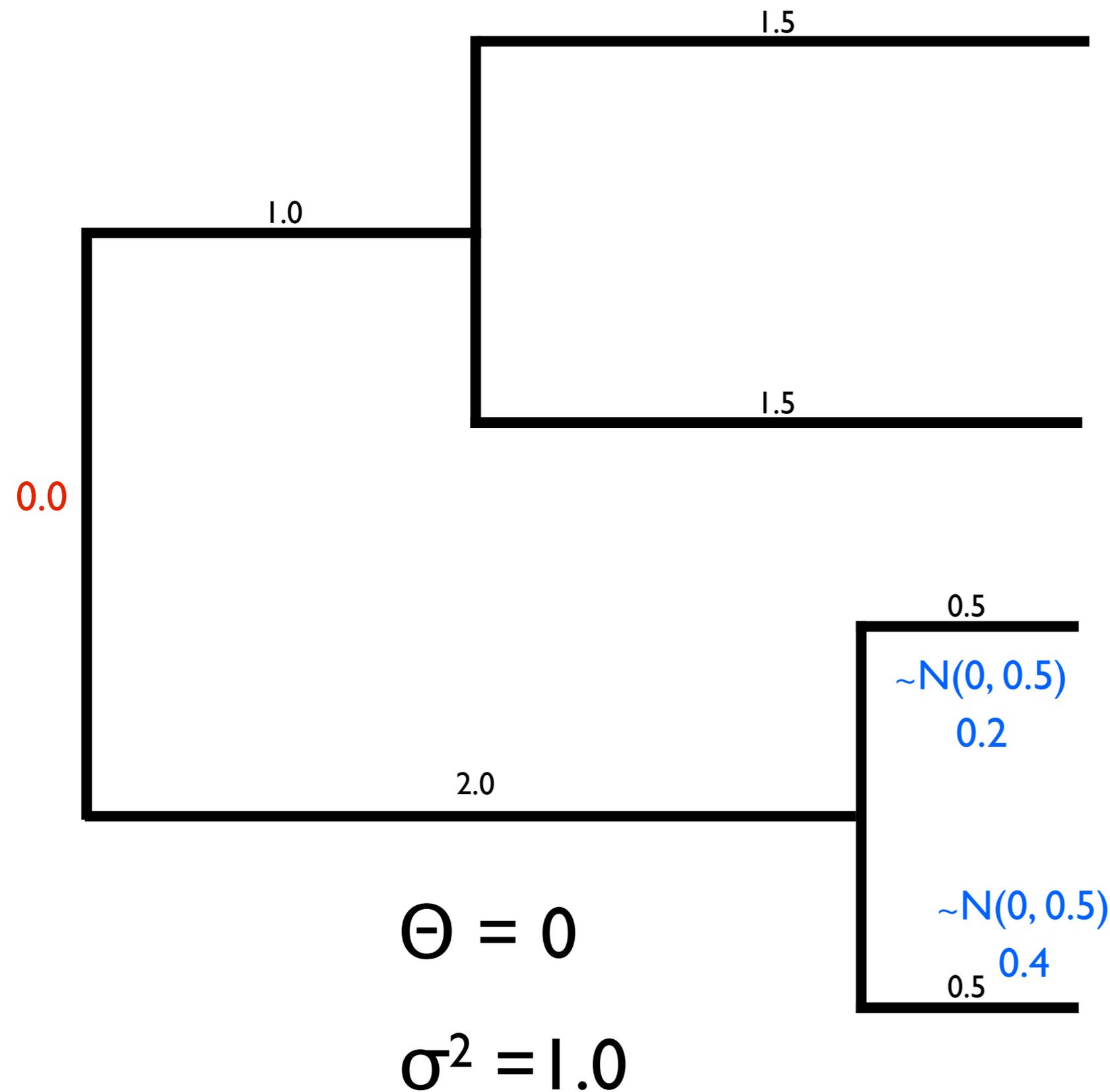
$$\Theta = 0$$

$$\sigma^2 = 1.0$$

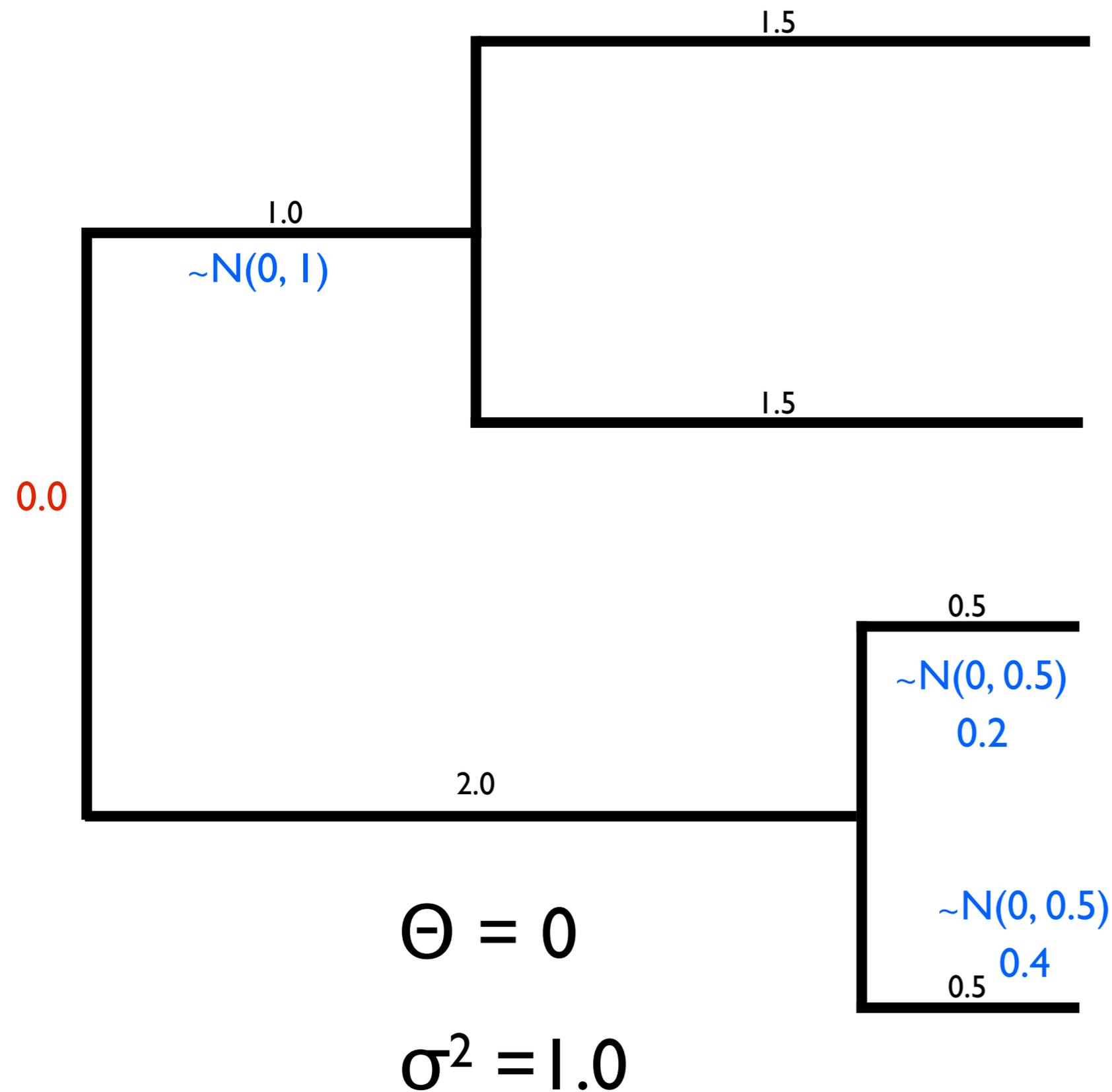
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



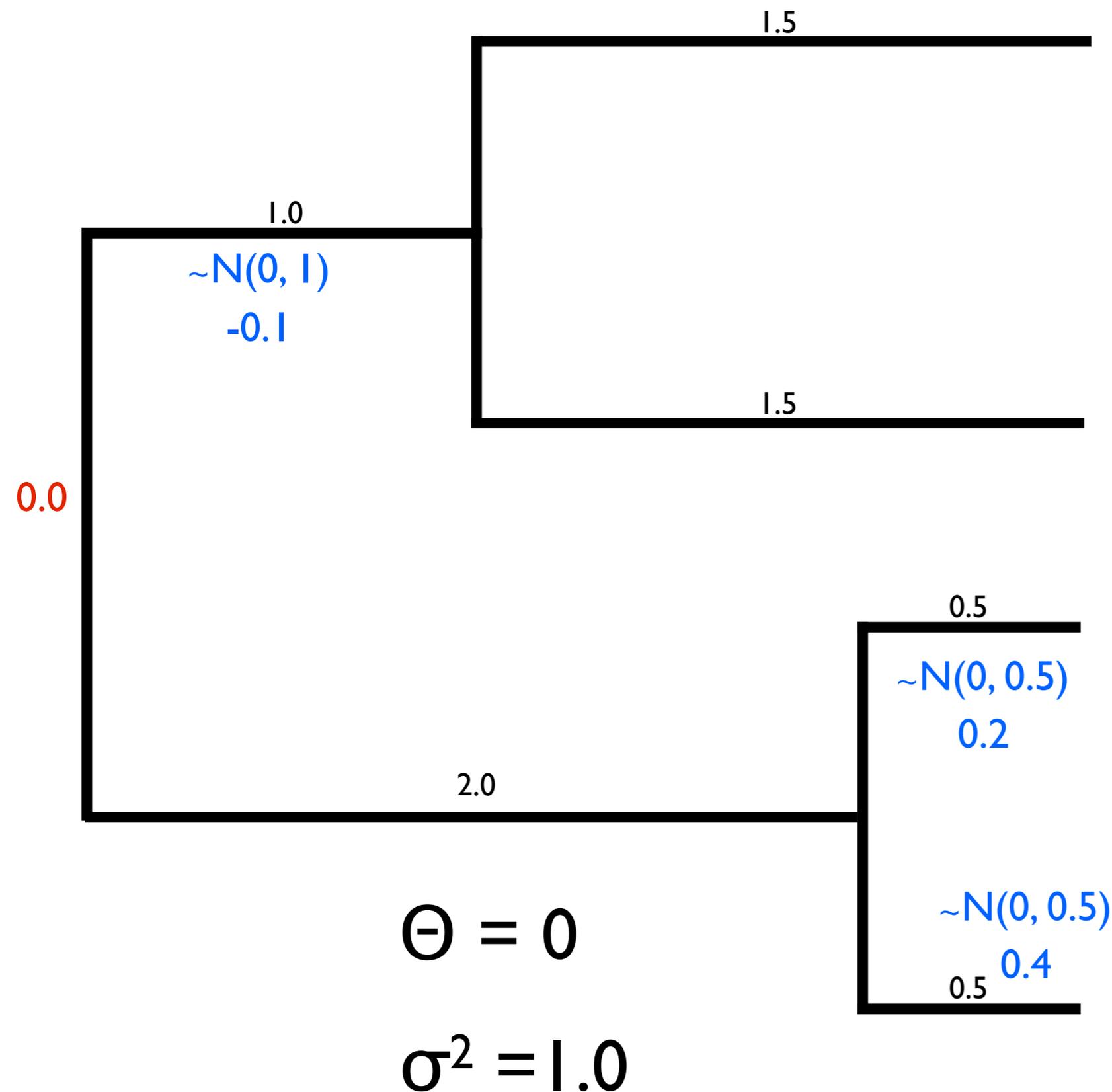
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



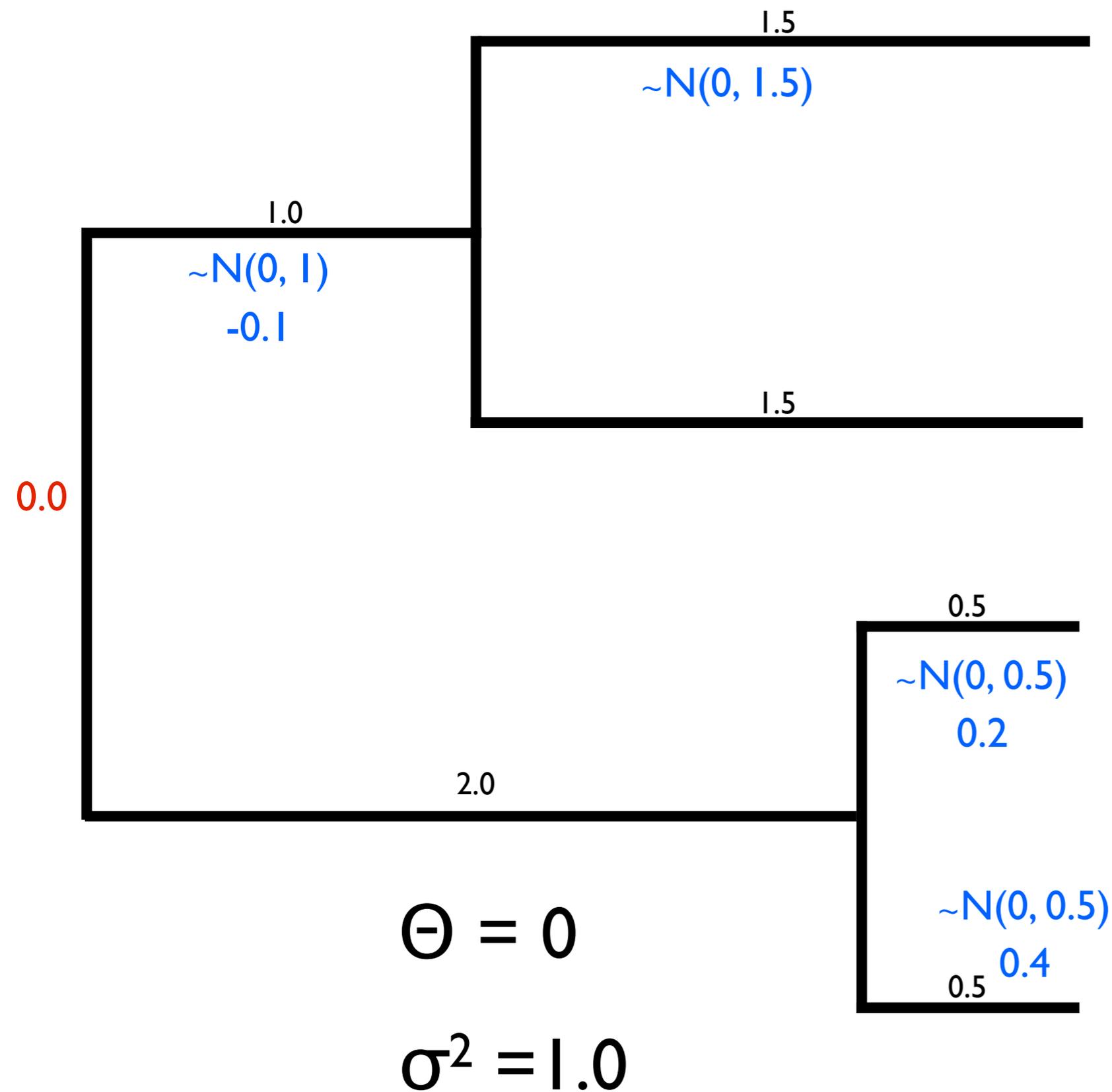
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



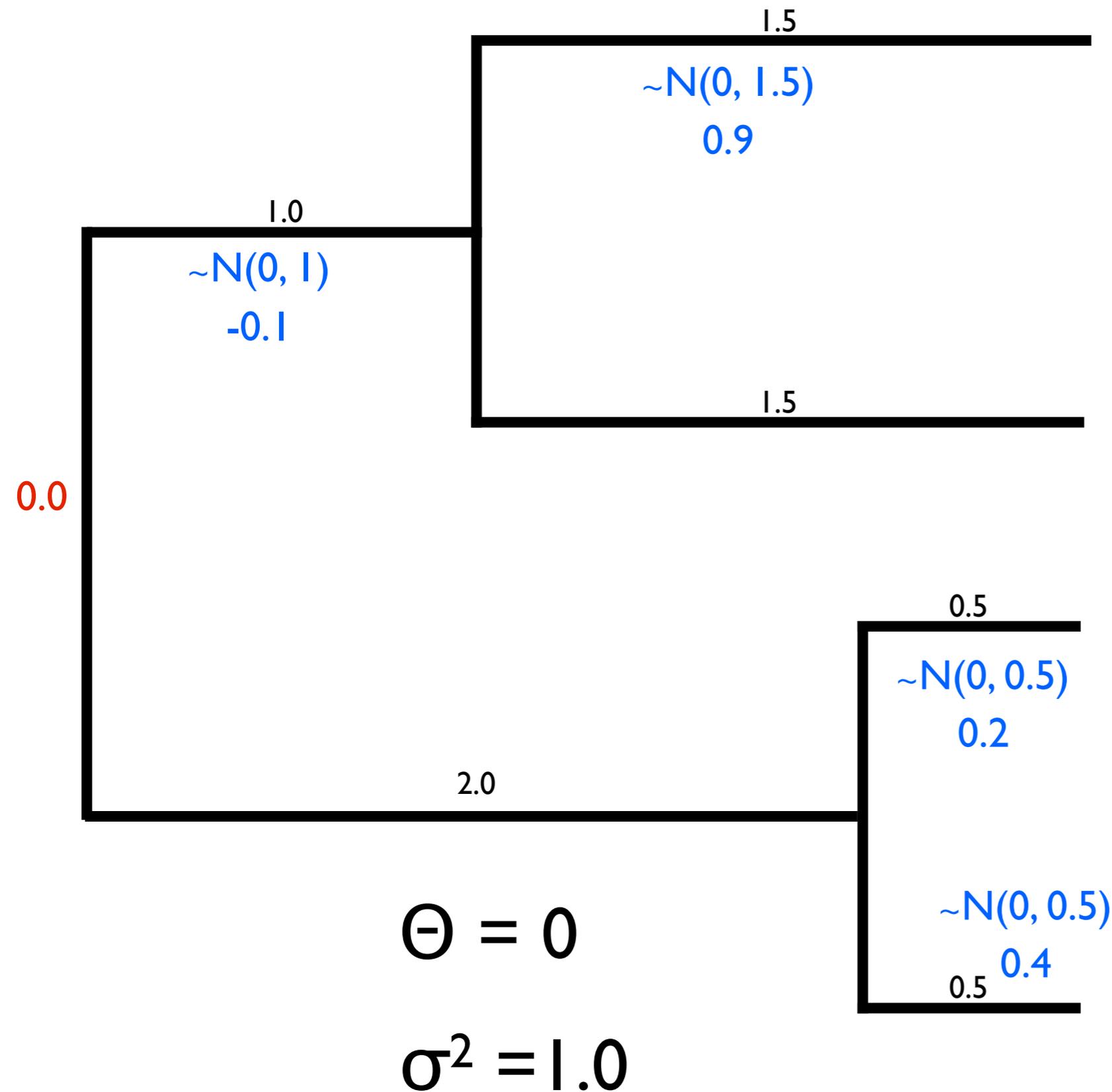
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



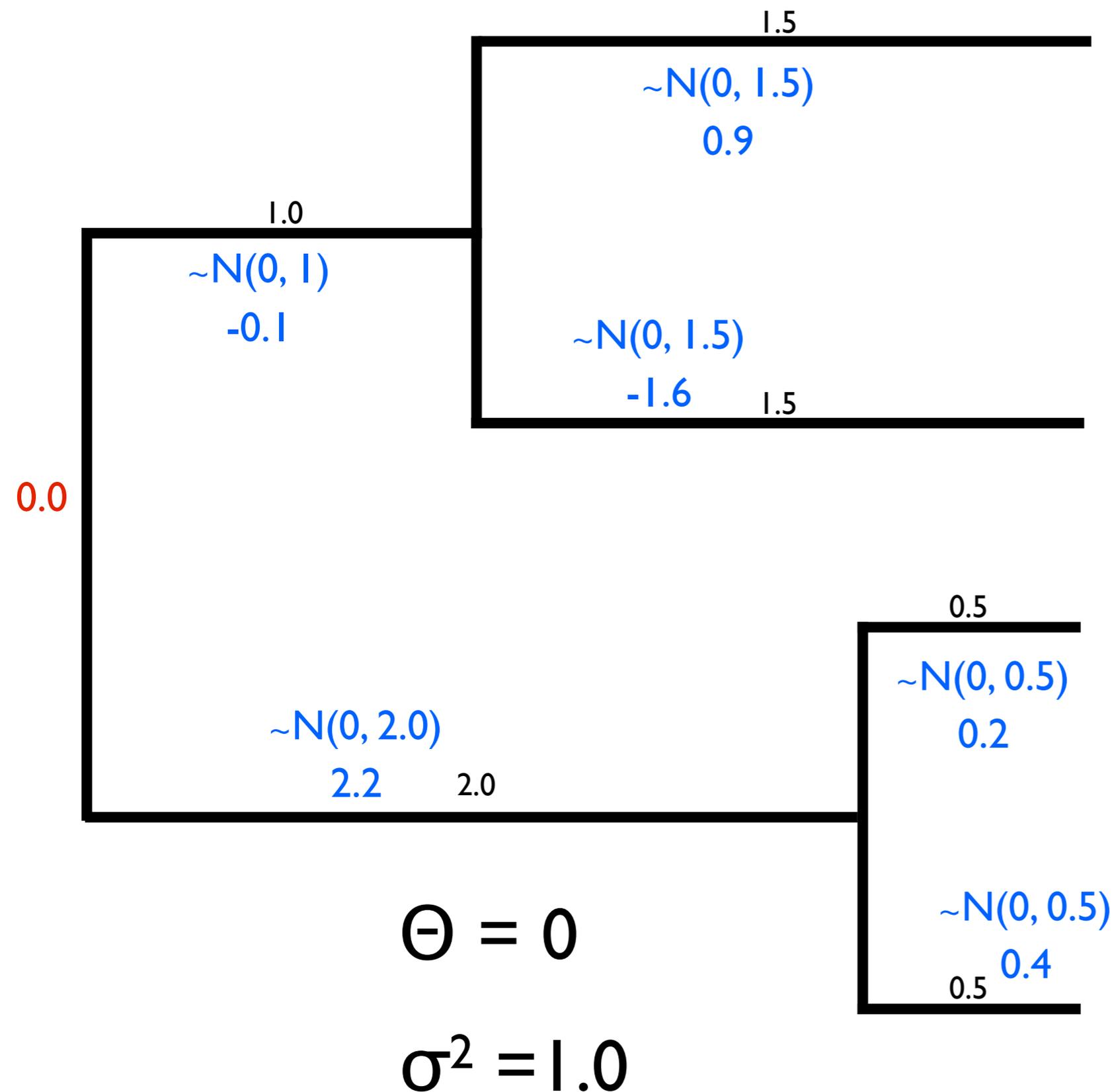
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



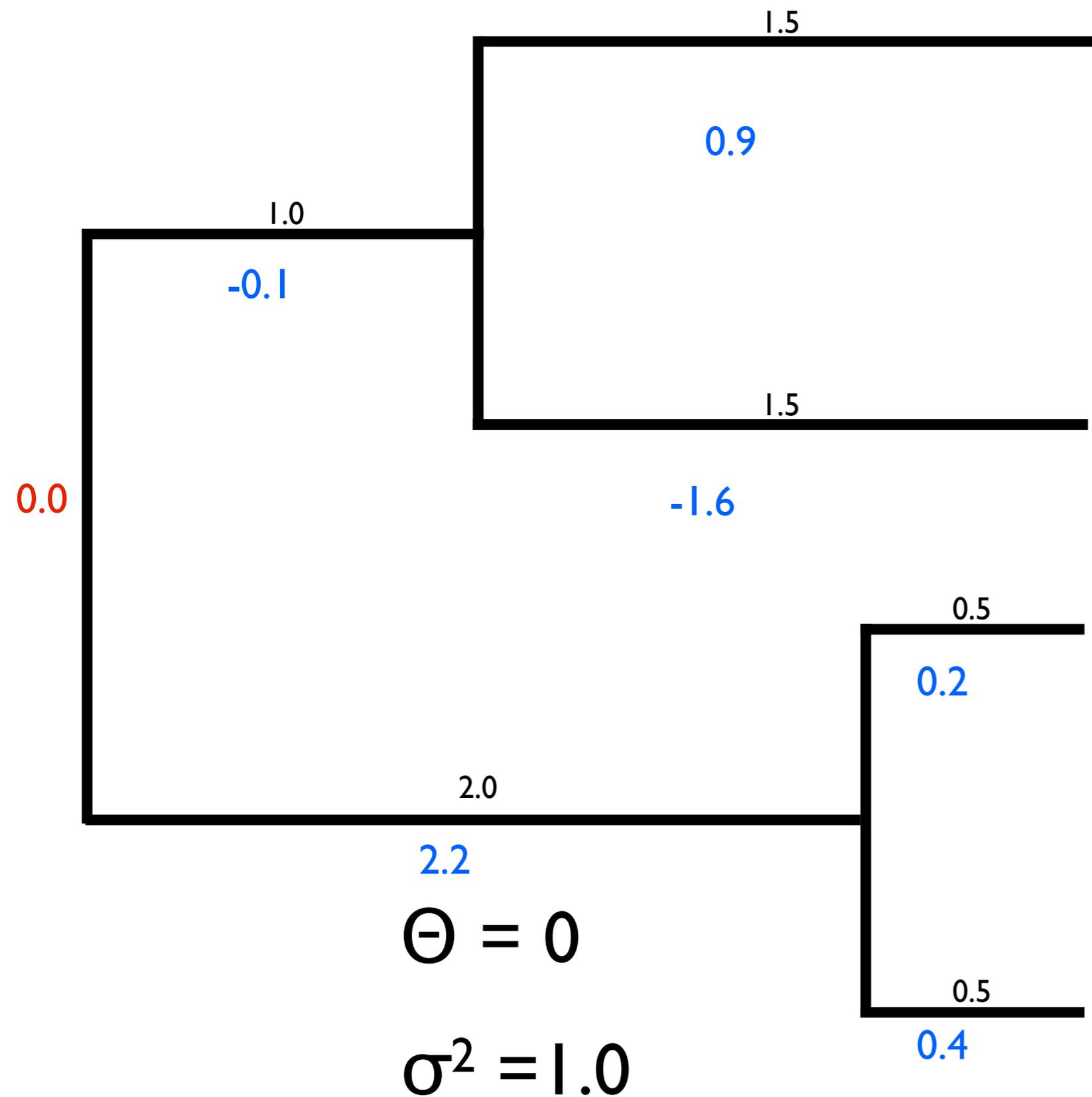
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



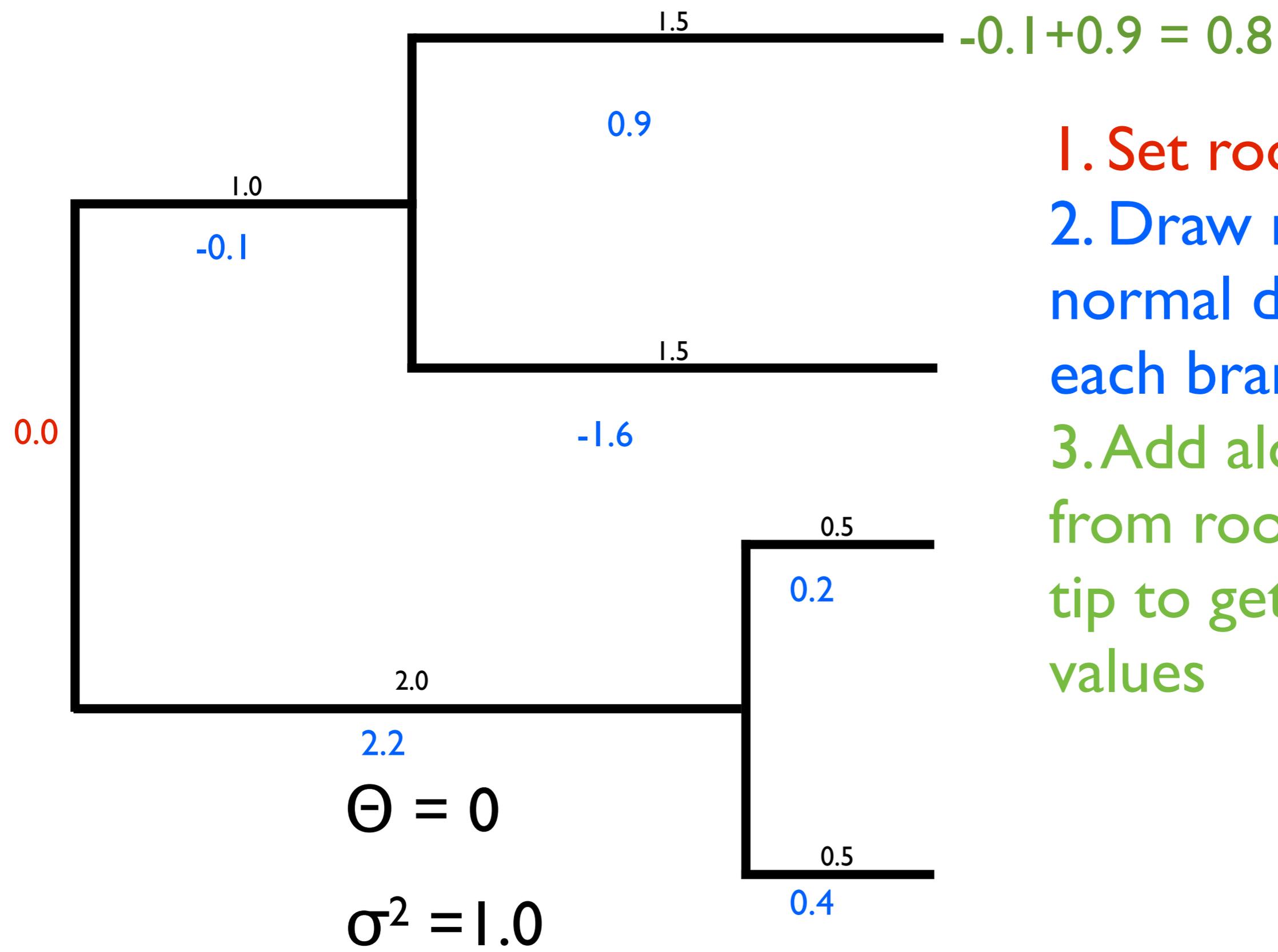
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



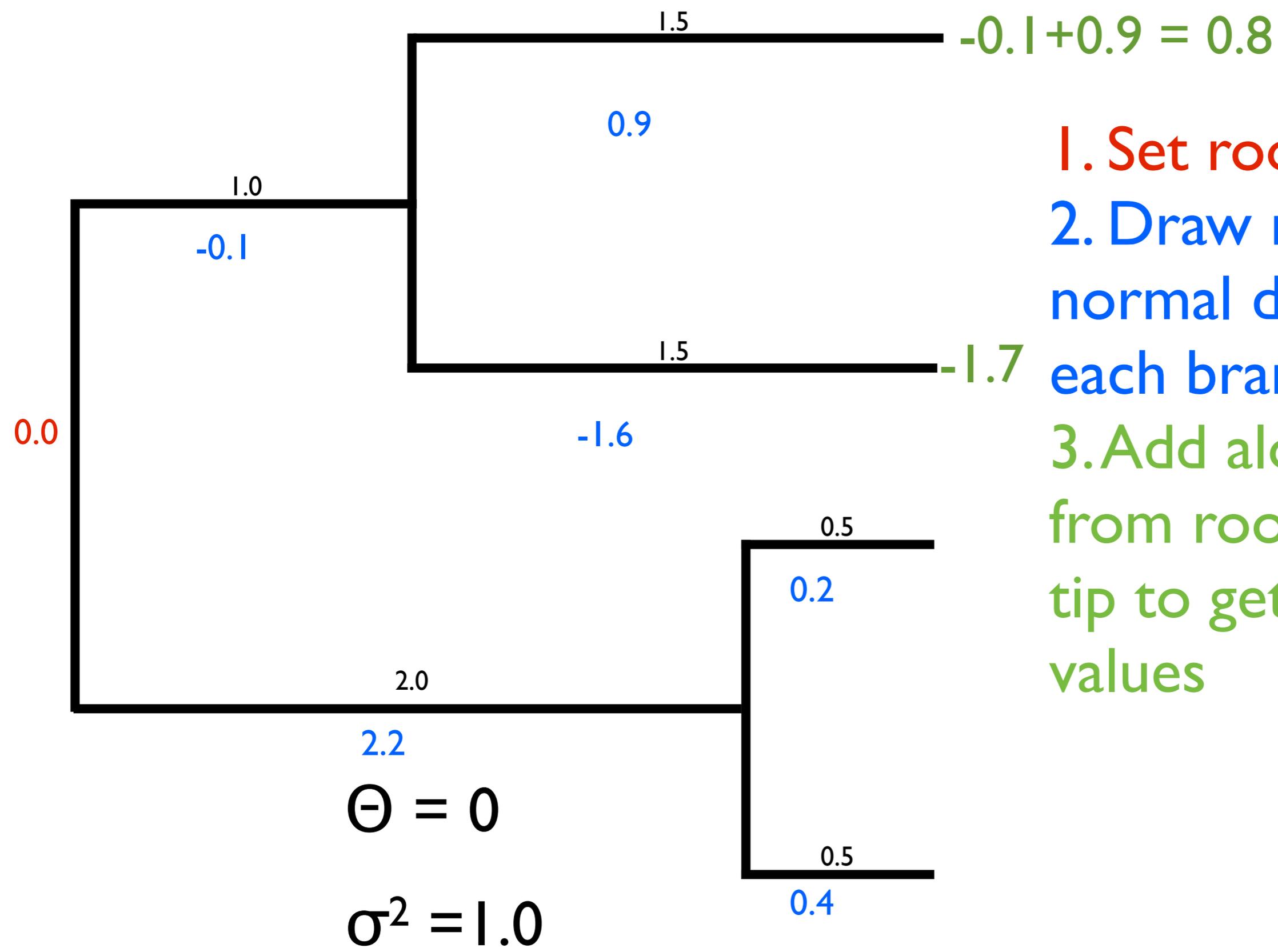
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



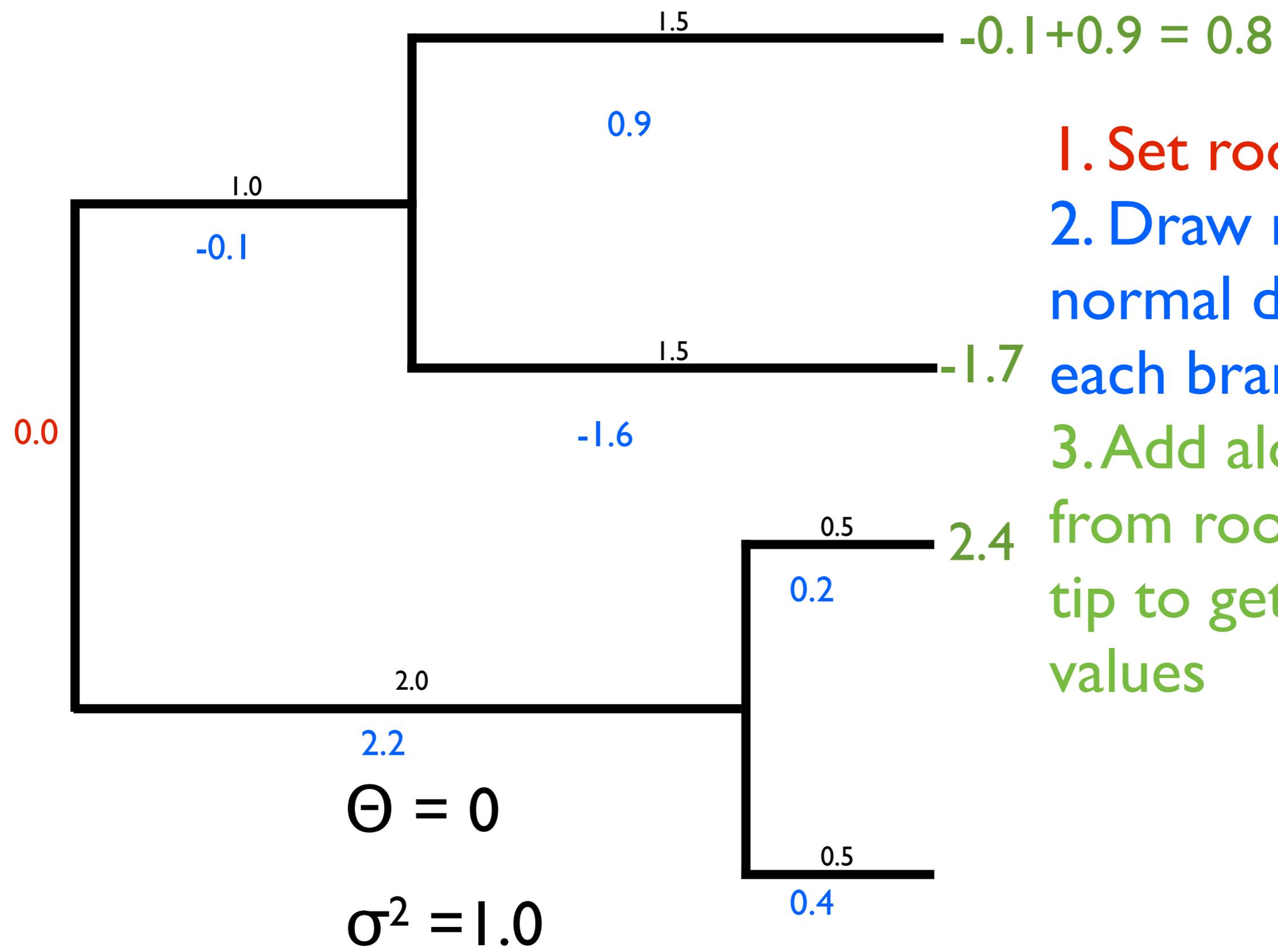
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



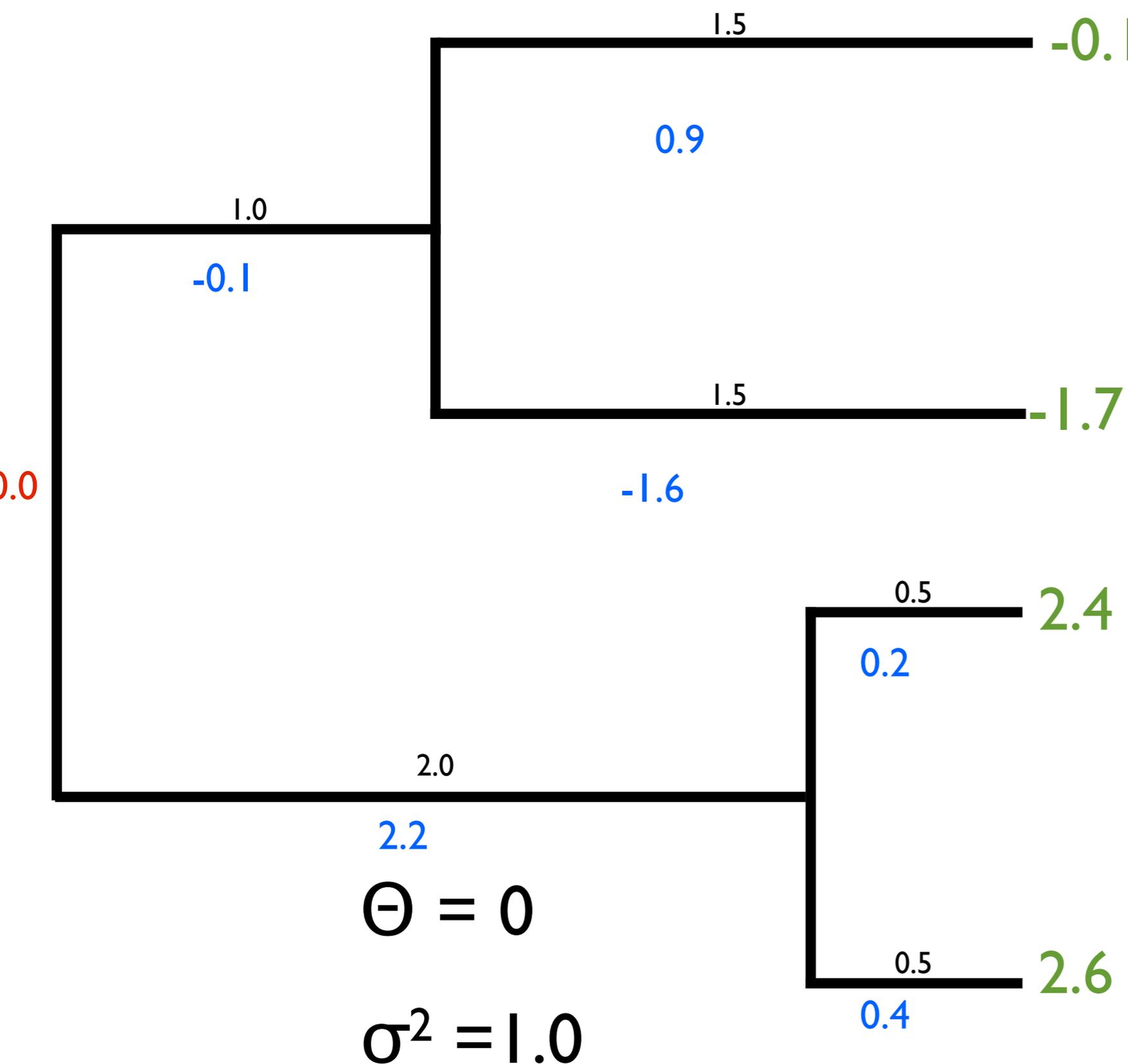
1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

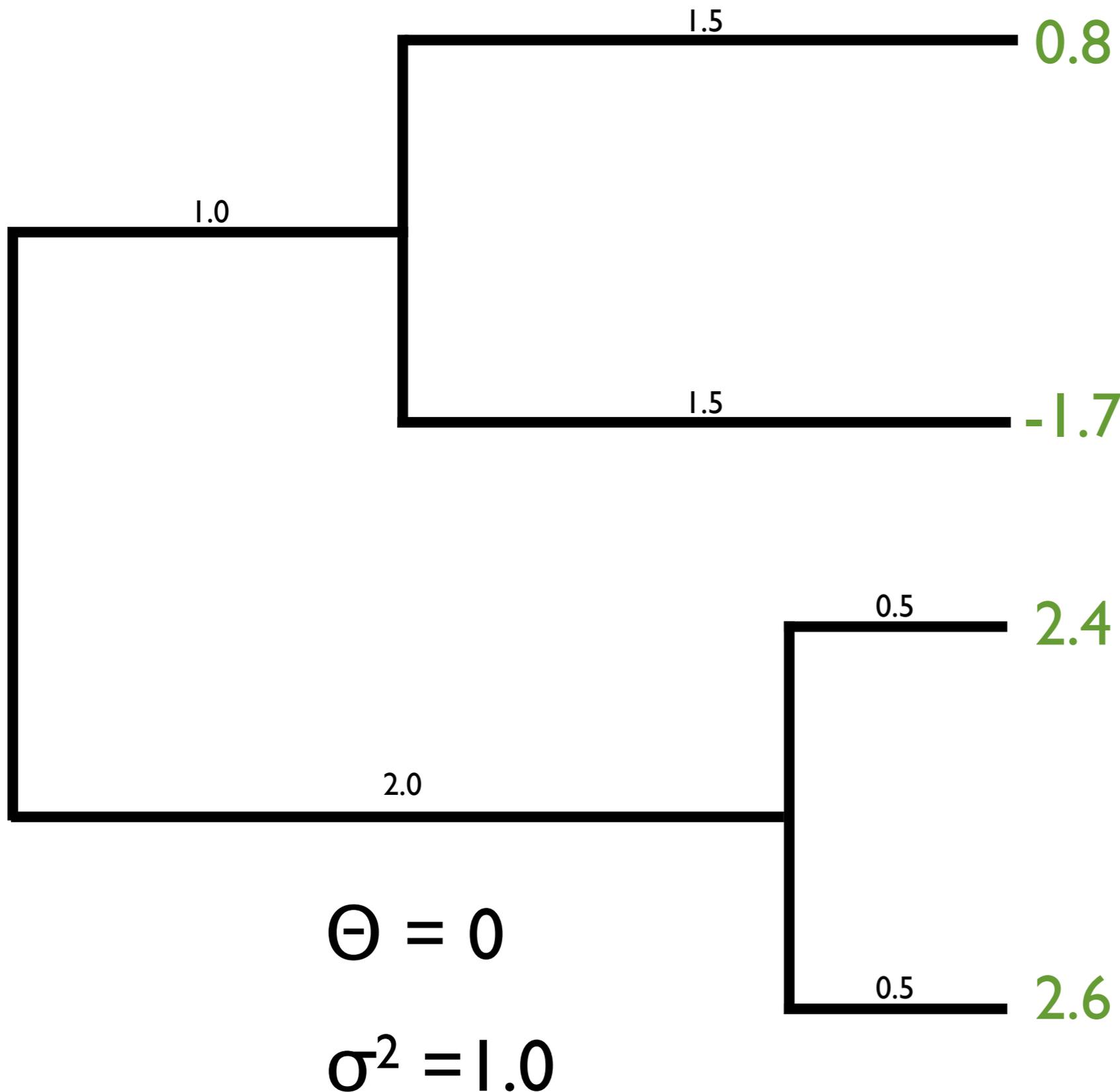


1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values



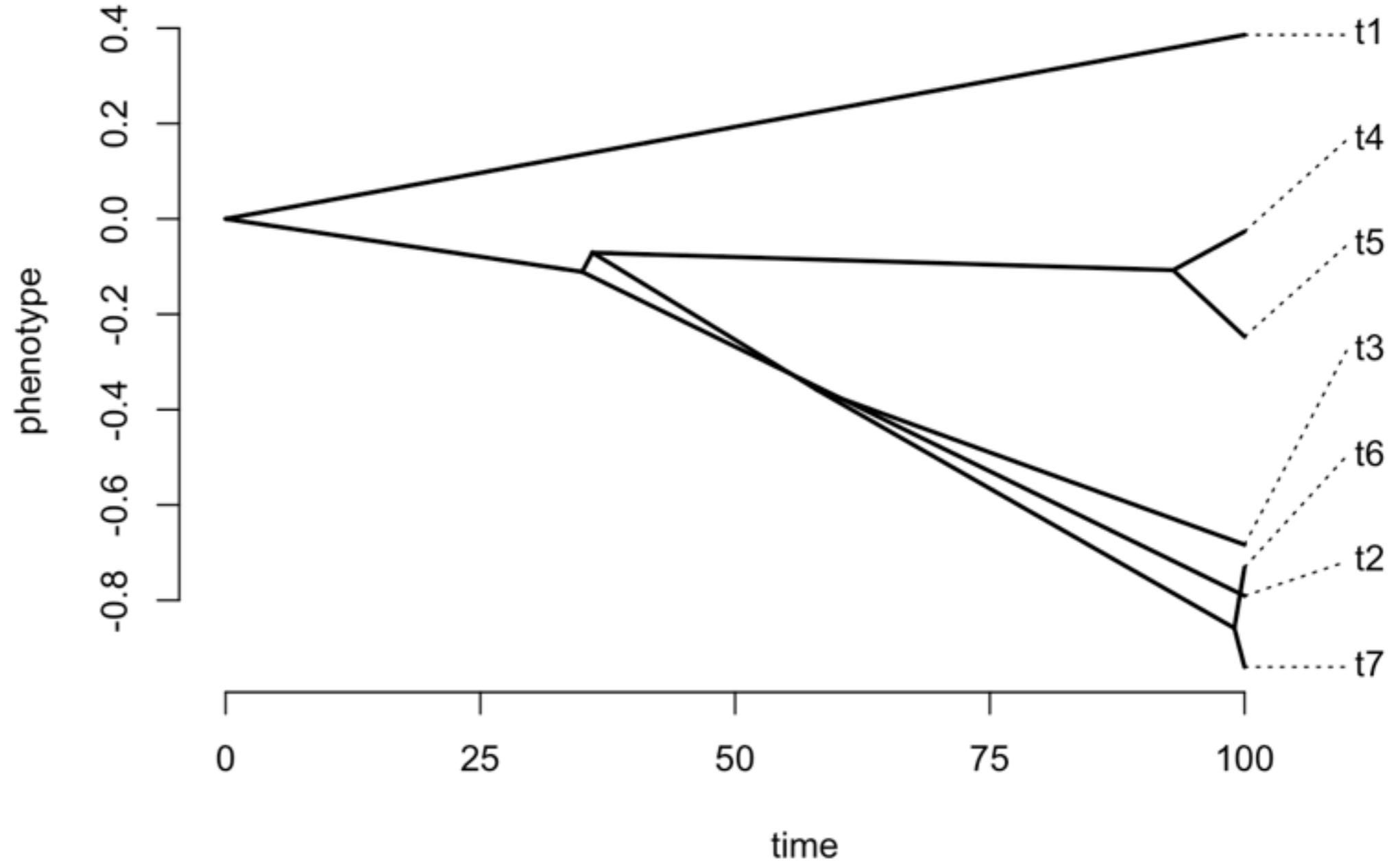
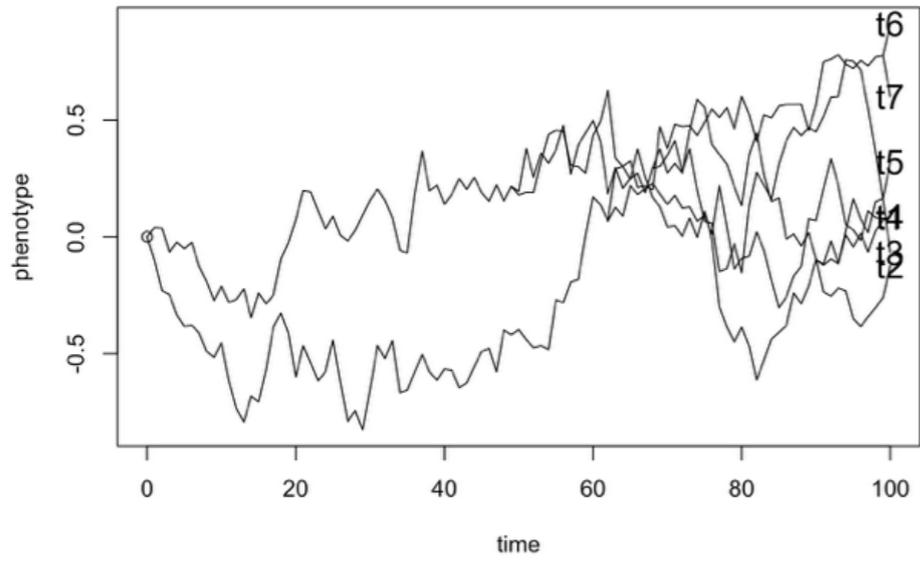
$-0.1 + 0.9 = 0.8$

1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

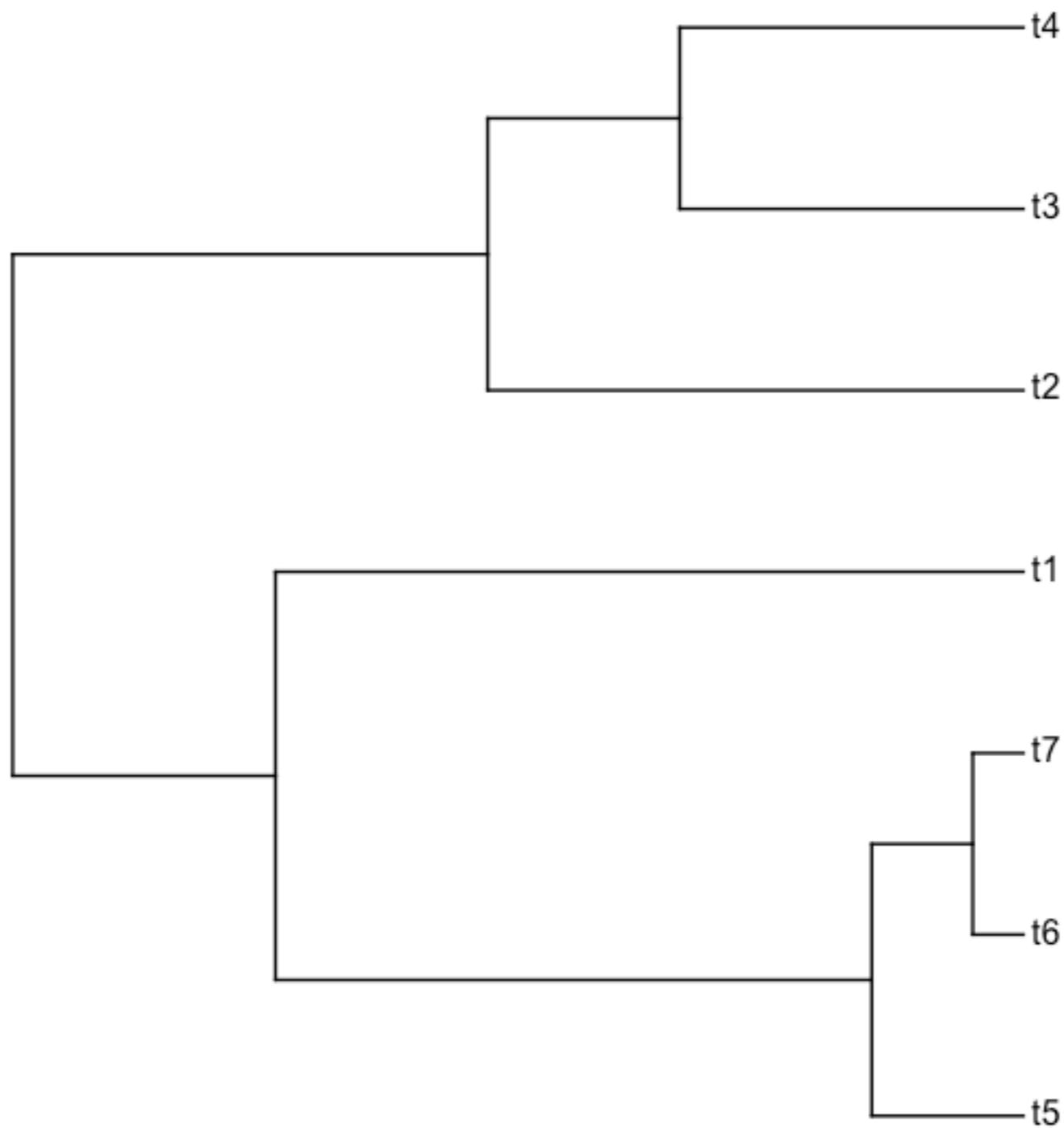


1. Set root state
2. Draw random normal deviate for each branch
3. Add along path from root to each tip to get tip values

“shortcut” simulations



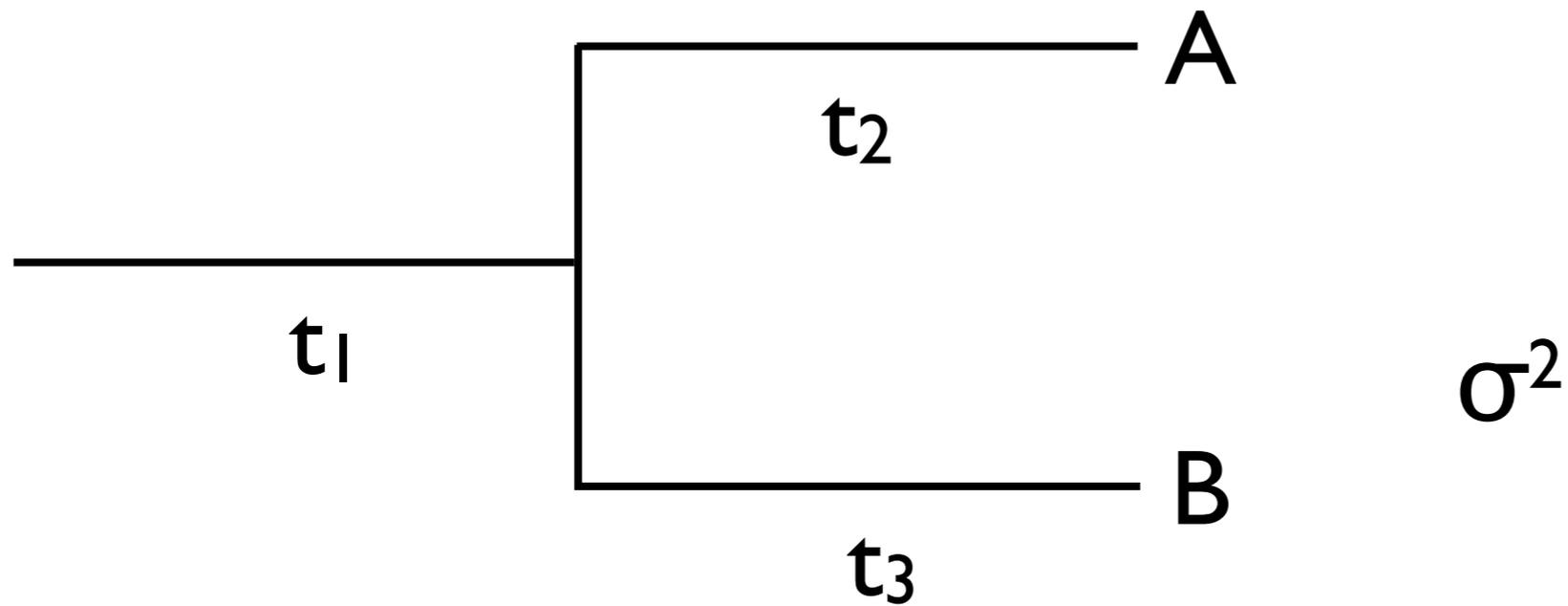
demo of BM simulations



Properties of BM on trees

- Variance increases with both σ^2 and t
- Expected (mean) value of any tip is always Θ
- Closely related species tend to be similar (they covary)

How do they covary?

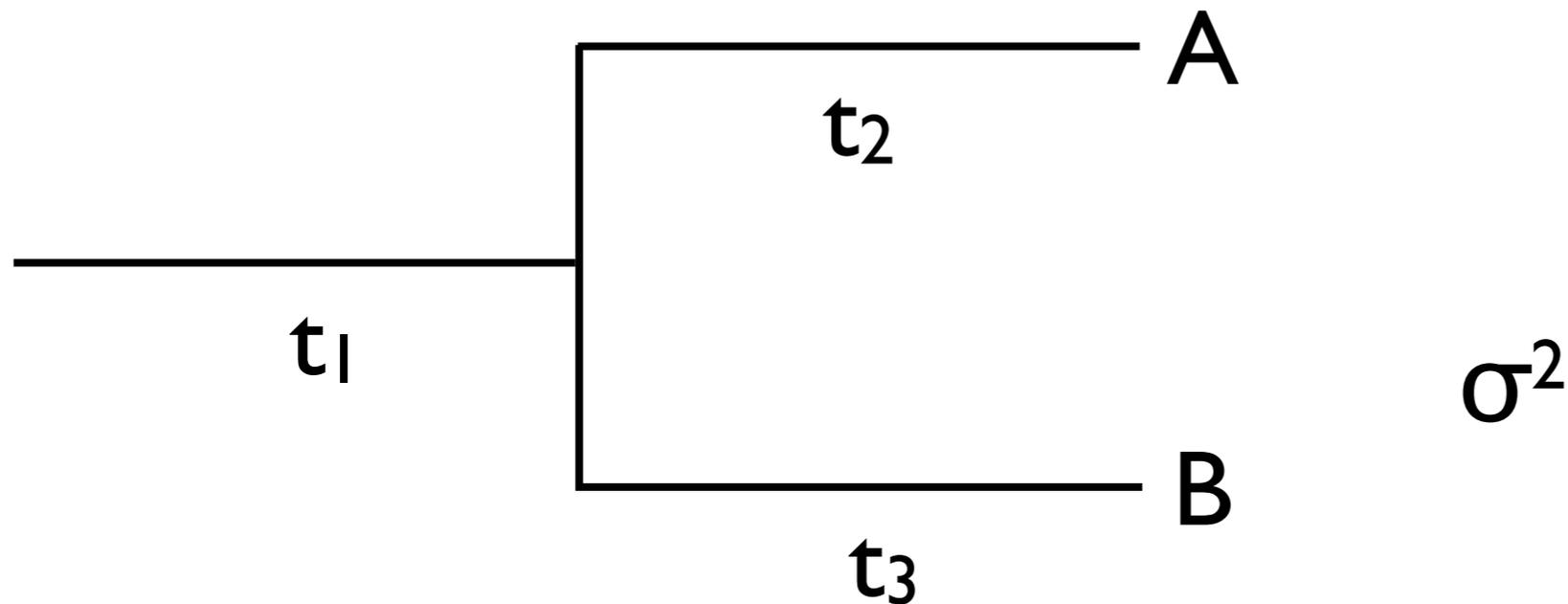


$\text{var}(A)$

$\text{cov}(A,B)$

$\text{var}(B)$

How do they covary?

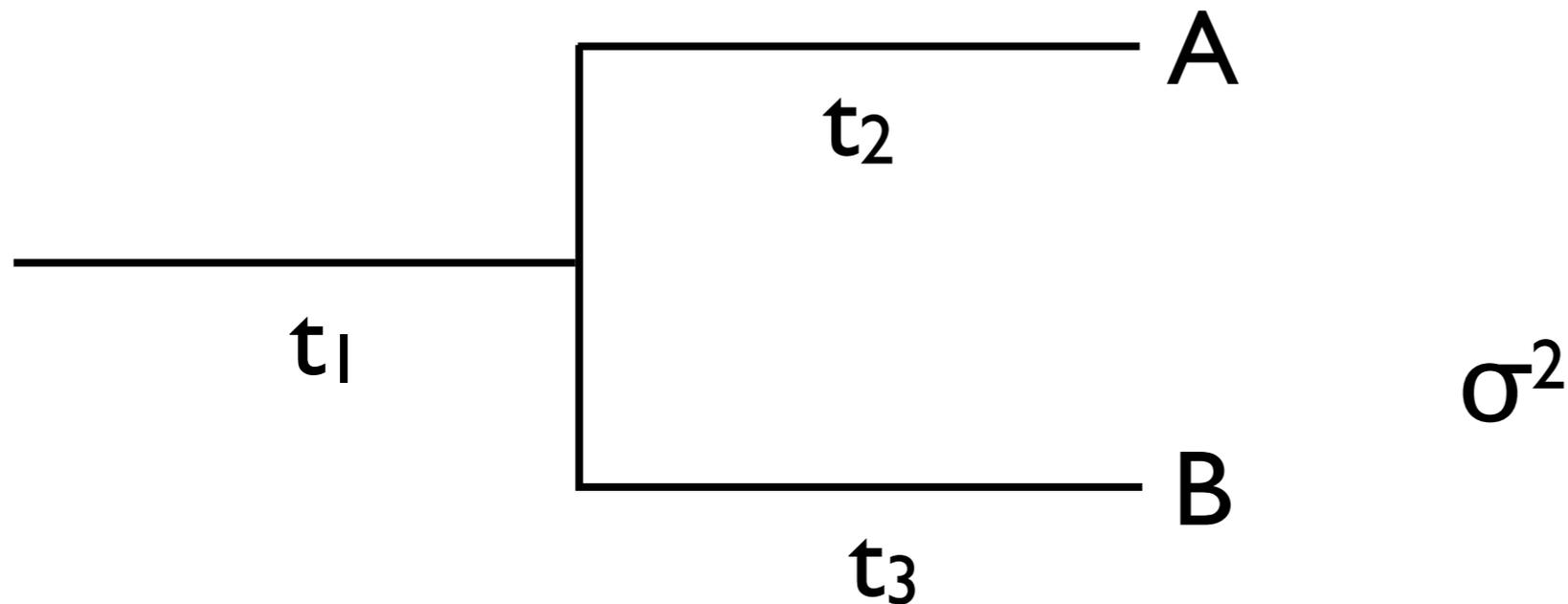


$$\text{var}(A) = \sigma^2(t_1 + t_2)$$

$$\text{cov}(A, B)$$

$$\text{var}(B)$$

How do they covary?

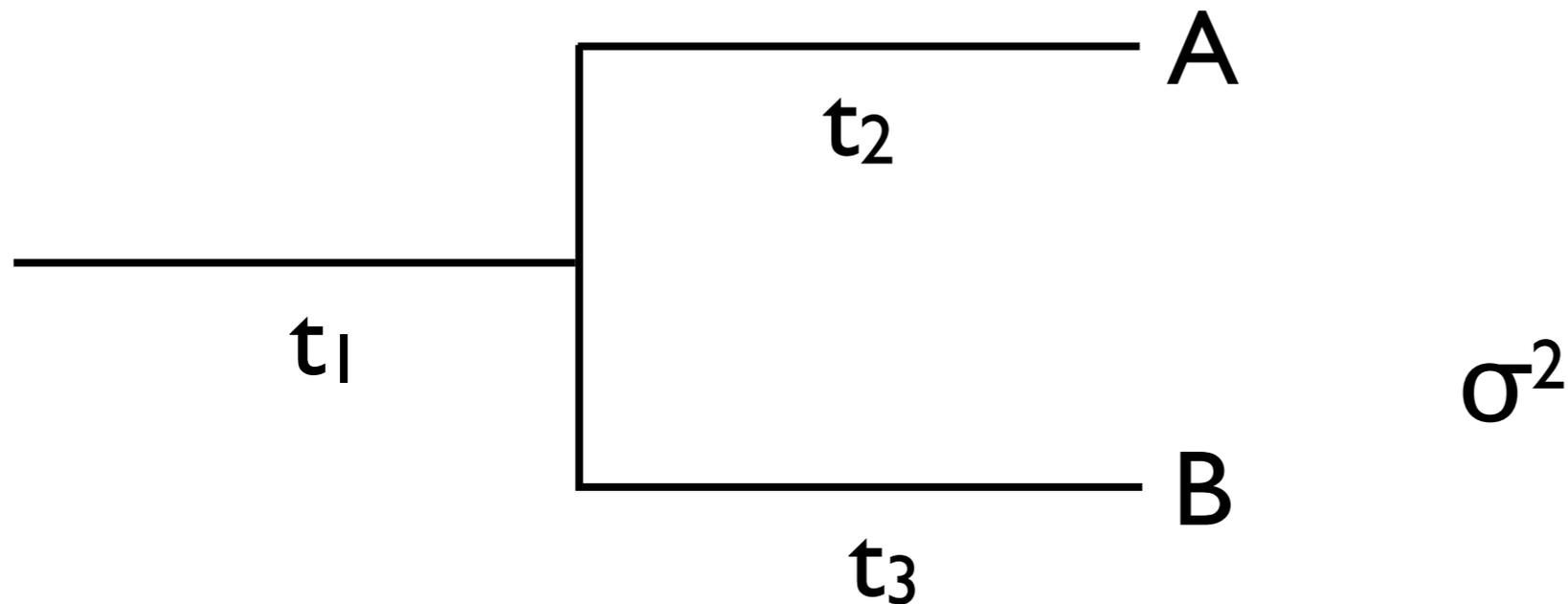


$$\text{var}(A) = \sigma^2(t_1 + t_2)$$

$$\text{cov}(A, B)$$

$$\text{var}(B) = \sigma^2(t_1 + t_3)$$

How do they covary?

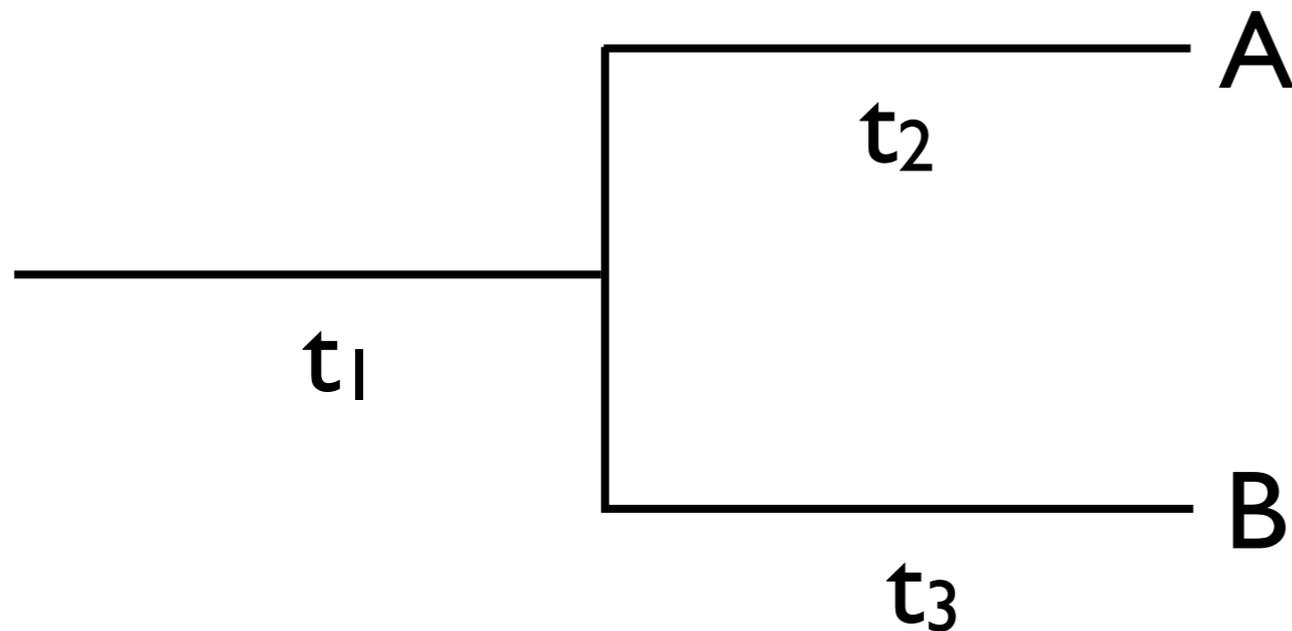


$$\text{var}(A) = \sigma^2(t_1 + t_2)$$

$$\text{cov}(A, B) = \sigma^2(t_1)$$

$$\text{var}(B) = \sigma^2(t_1 + t_3)$$

How do they covary?



**variance-covariance
matrix**

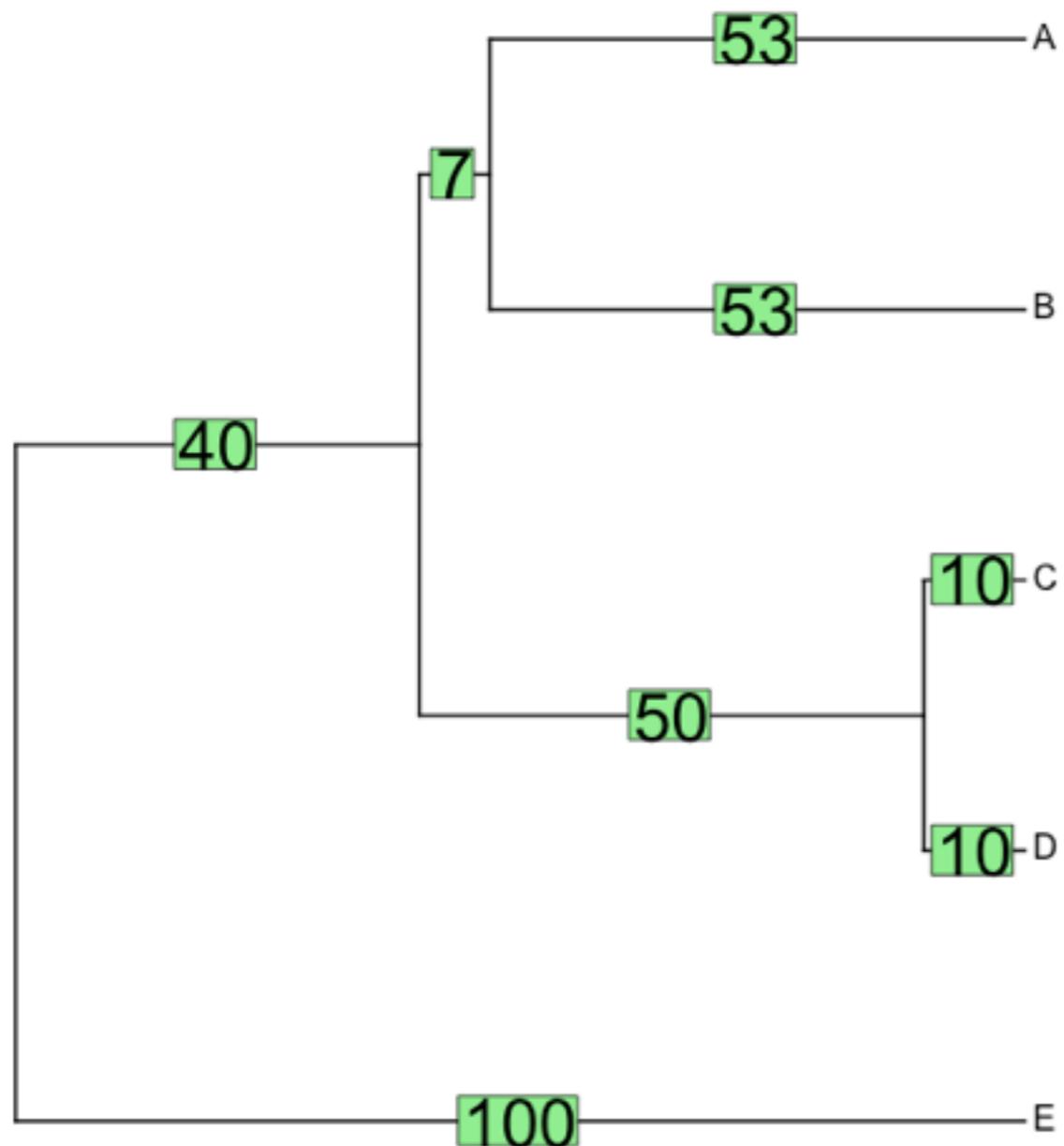
$$\sigma^2 \begin{bmatrix} t_1 + t_2 & t_1 \\ t_1 & t_1 + t_3 \end{bmatrix}$$

$$\text{var}(A) = \sigma^2(t_1 + t_2)$$

$$\text{cov}(A, B) = \sigma^2(t_1)$$

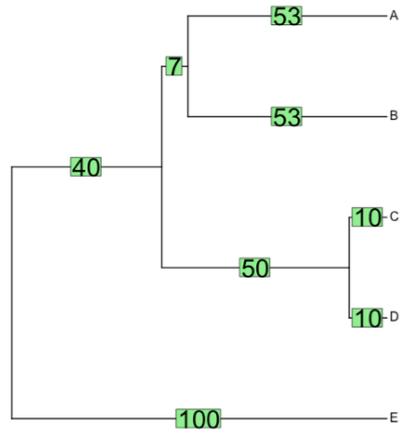
$$\text{var}(B) = \sigma^2(t_1 + t_3)$$

Under Brownian motion model, tip values on a tree are distributed as multivariate normal with expected value = θ and variance/covariance = time of common ancestry \times Brownian rate, σ^2

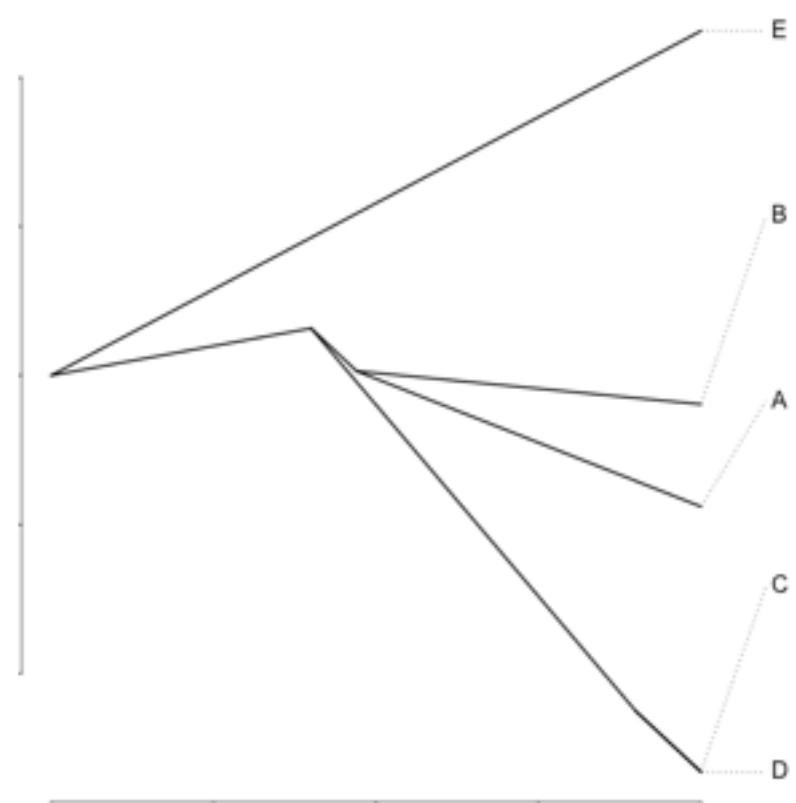
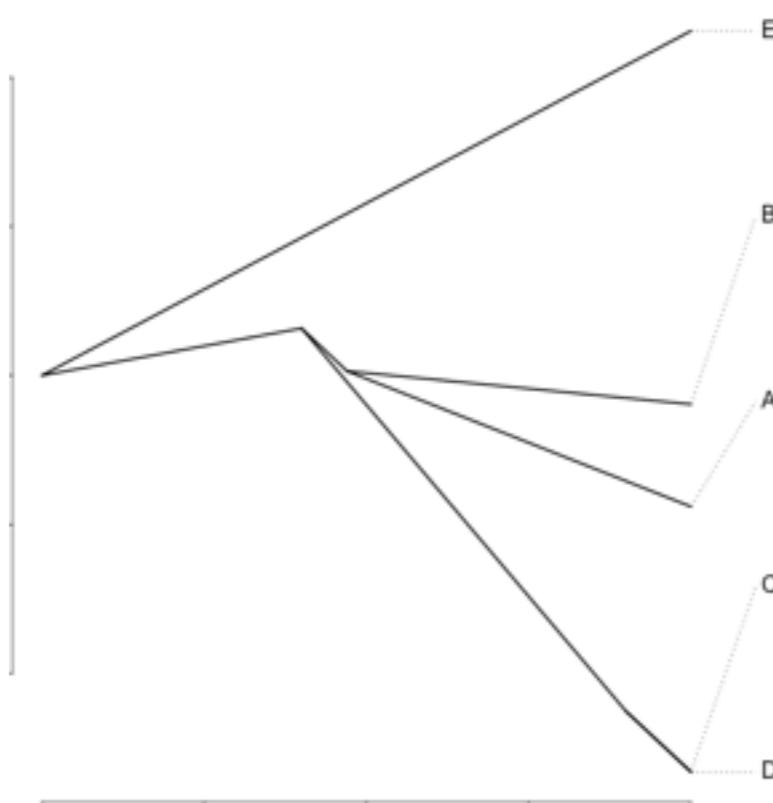
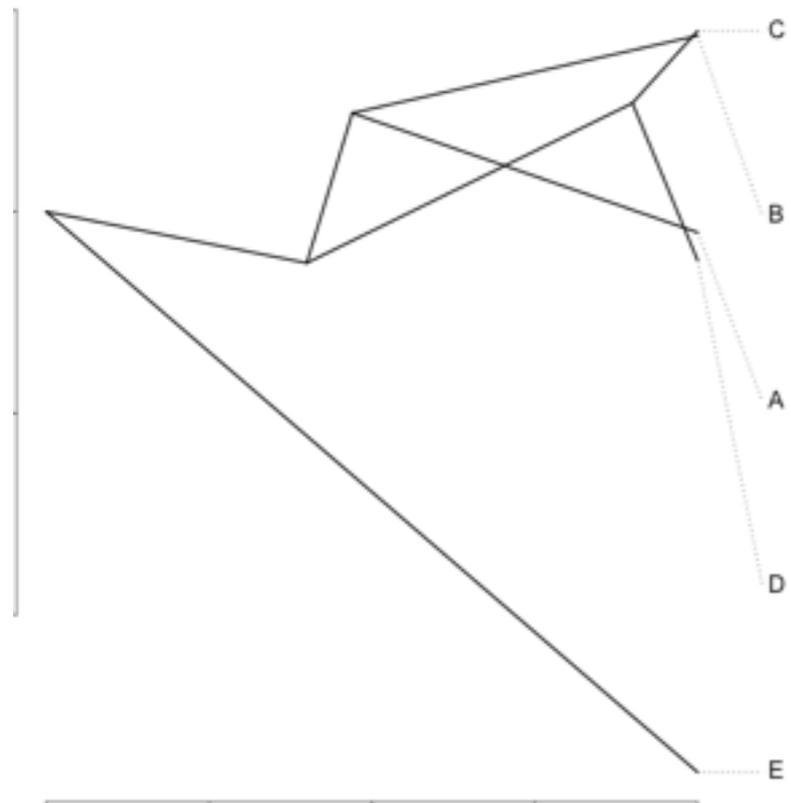


$$V = \sigma^2 \begin{bmatrix} A & B & C & D & E \\ 100 & 47 & 40 & 40 & 0 \\ 47 & 100 & 40 & 40 & 0 \\ 40 & 40 & 100 & 90 & 0 \\ 40 & 40 & 90 & 100 & 0 \\ 0 & 0 & 0 & 0 & 100 \end{bmatrix}$$

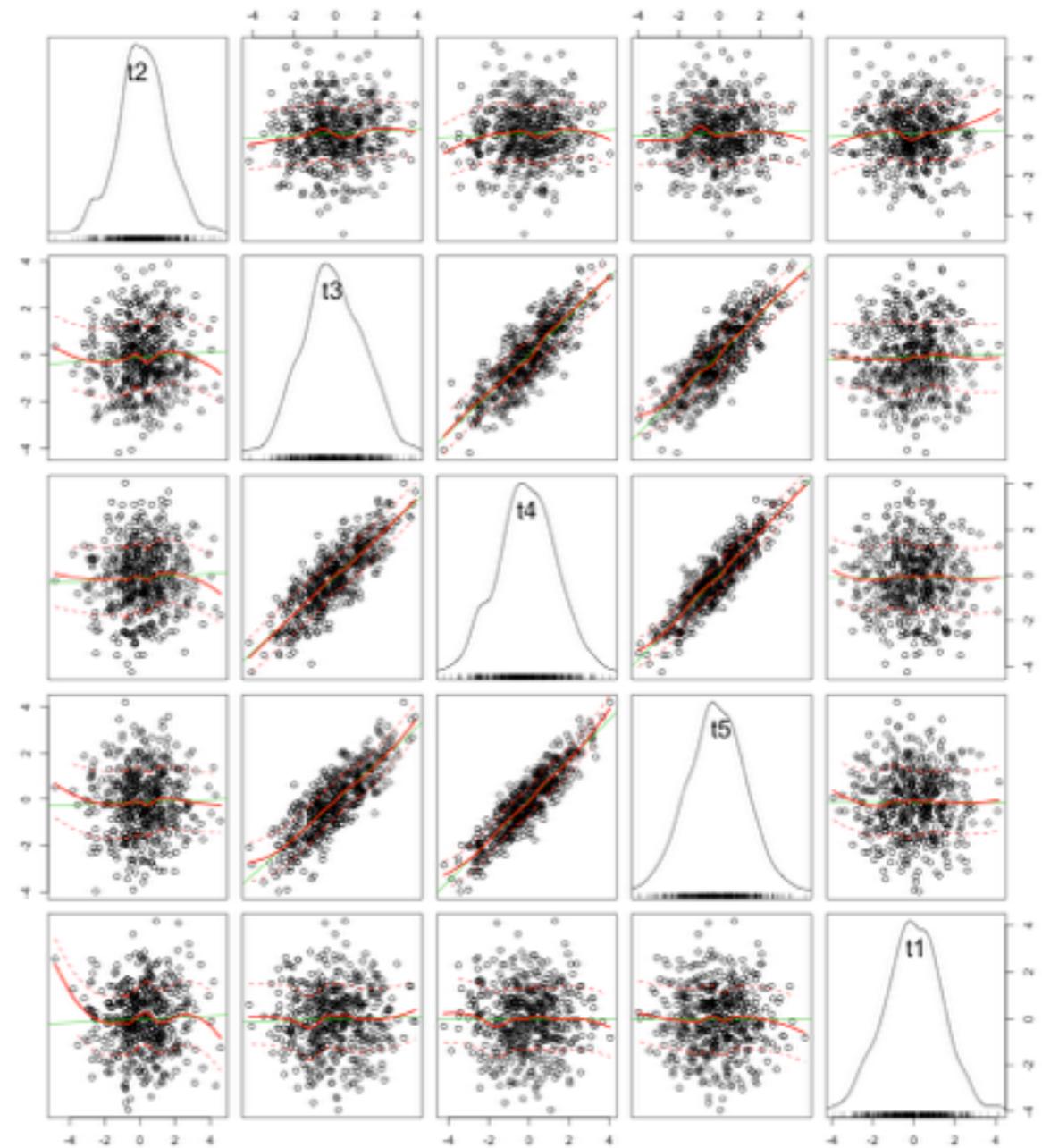
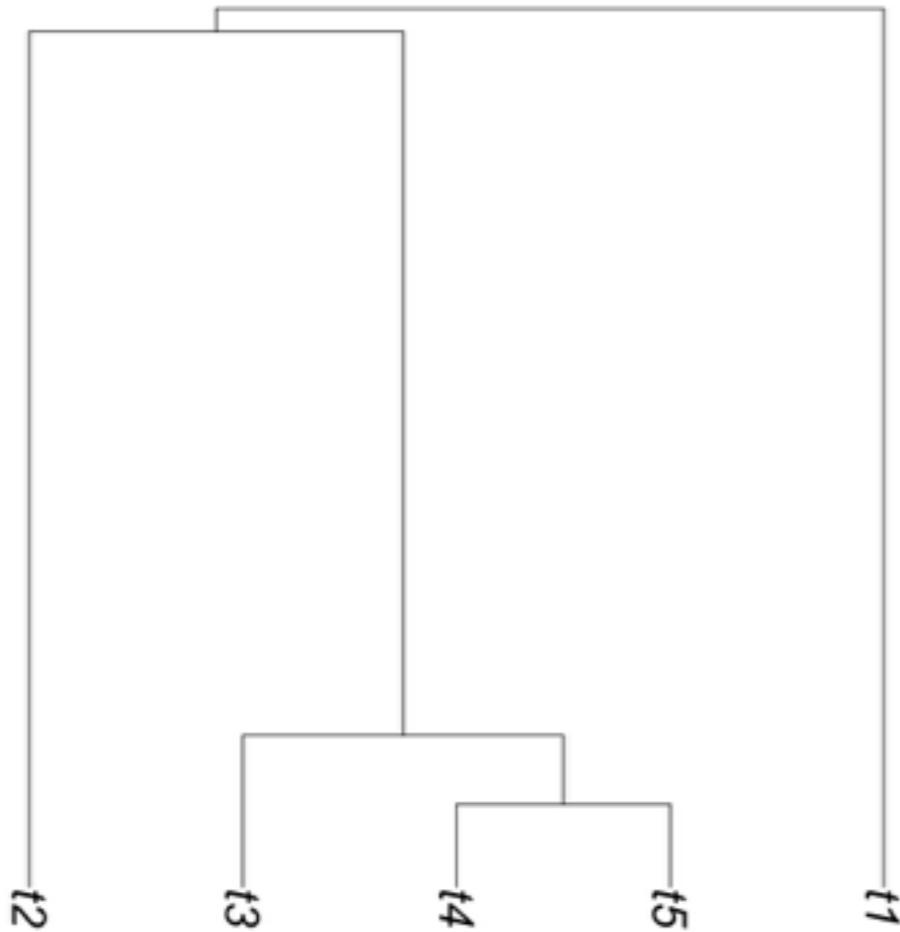
We can draw from this distribution to quickly simulate evolution under this model



$$V = \sigma^2 \begin{bmatrix} 100 & 47 & 40 & 40 & 0 \\ 47 & 100 & 40 & 40 & 0 \\ 40 & 40 & 100 & 90 & 0 \\ 40 & 40 & 90 & 100 & 0 \\ 0 & 0 & 0 & 0 & 100 \end{bmatrix}$$

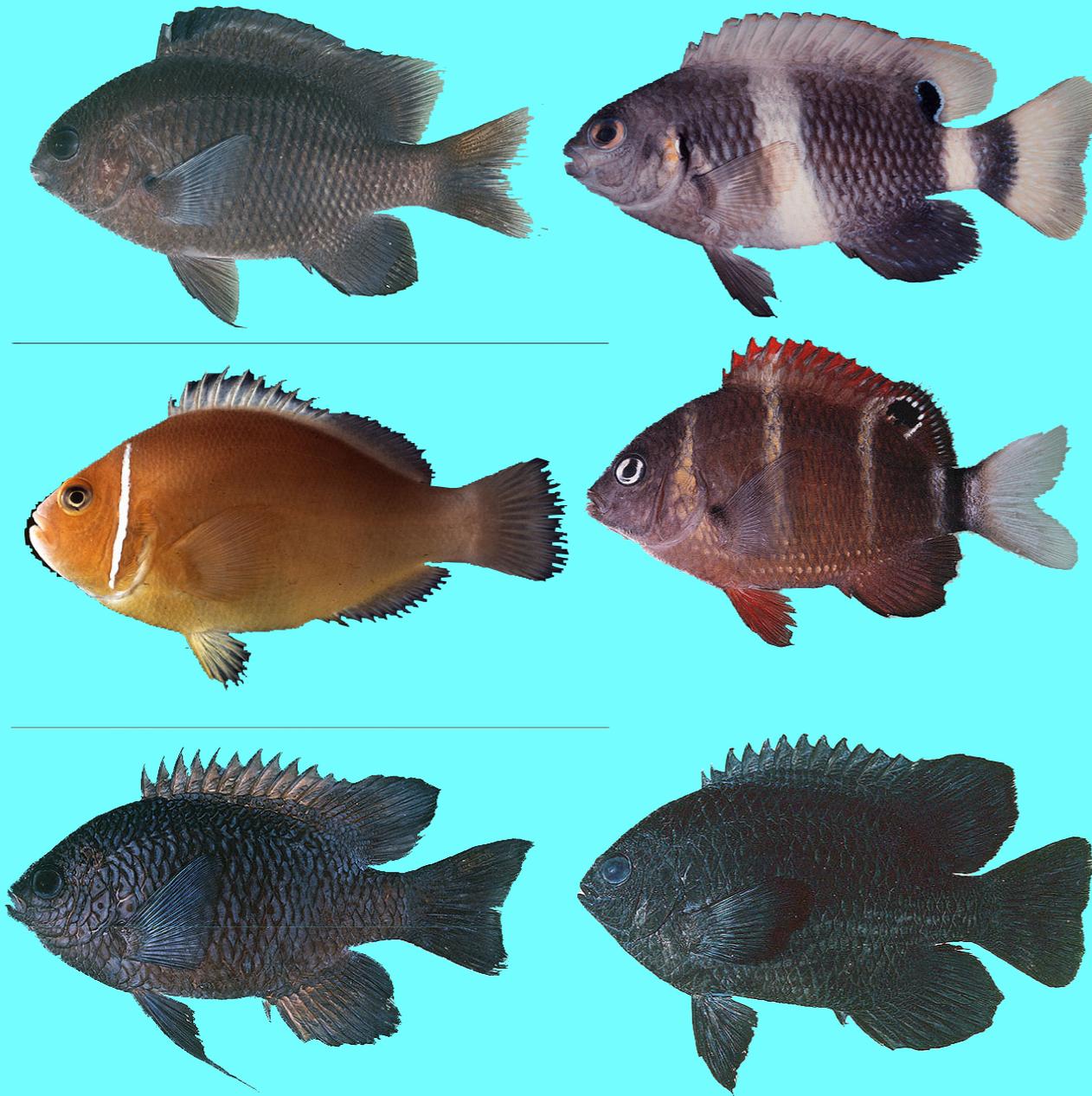


shared ancestry creates
covariation across tips



using BM in comparative
studies

Studying rates of morphological evolution



Family Pomacentridae
~320 species

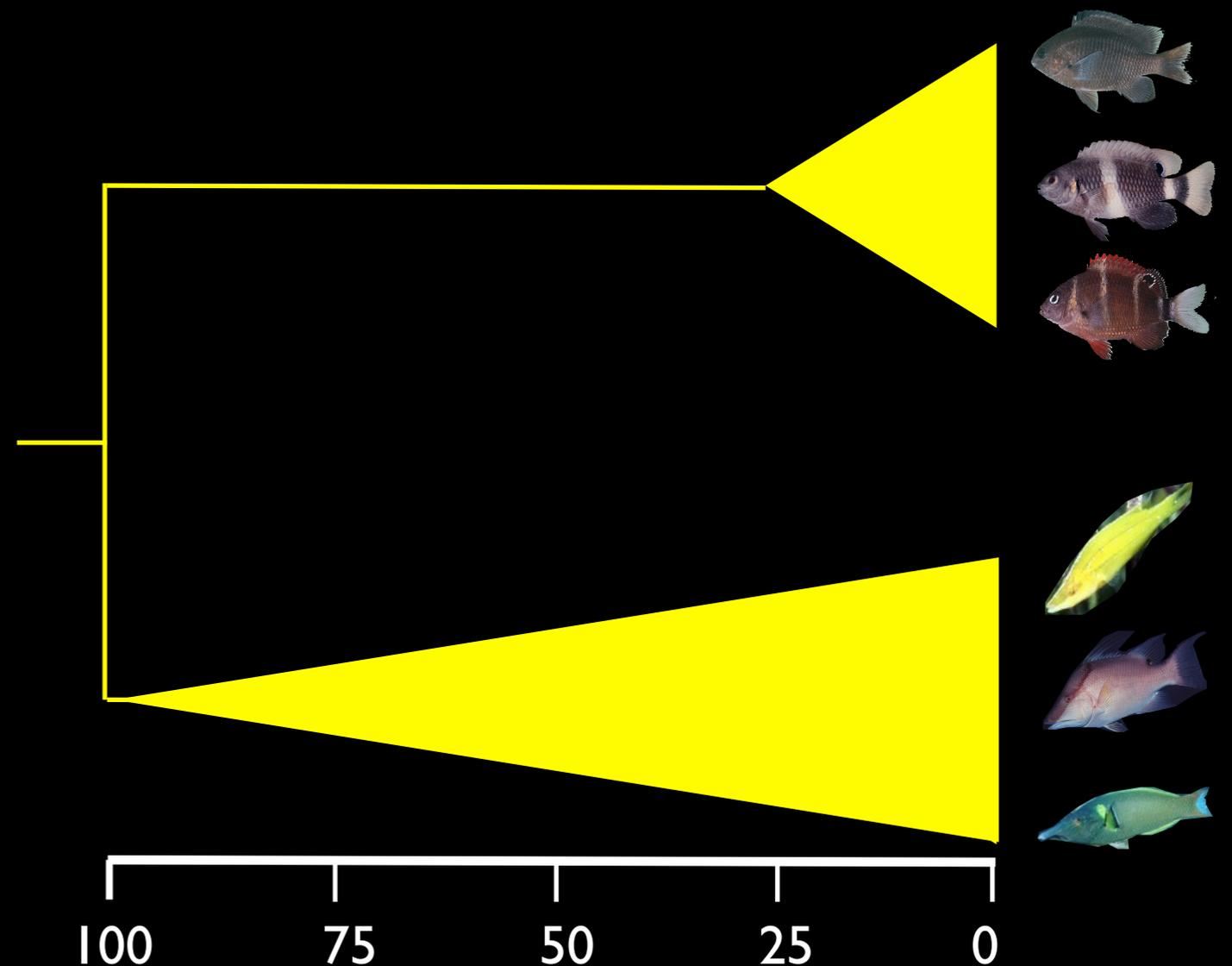


Family Labridae
~500 species

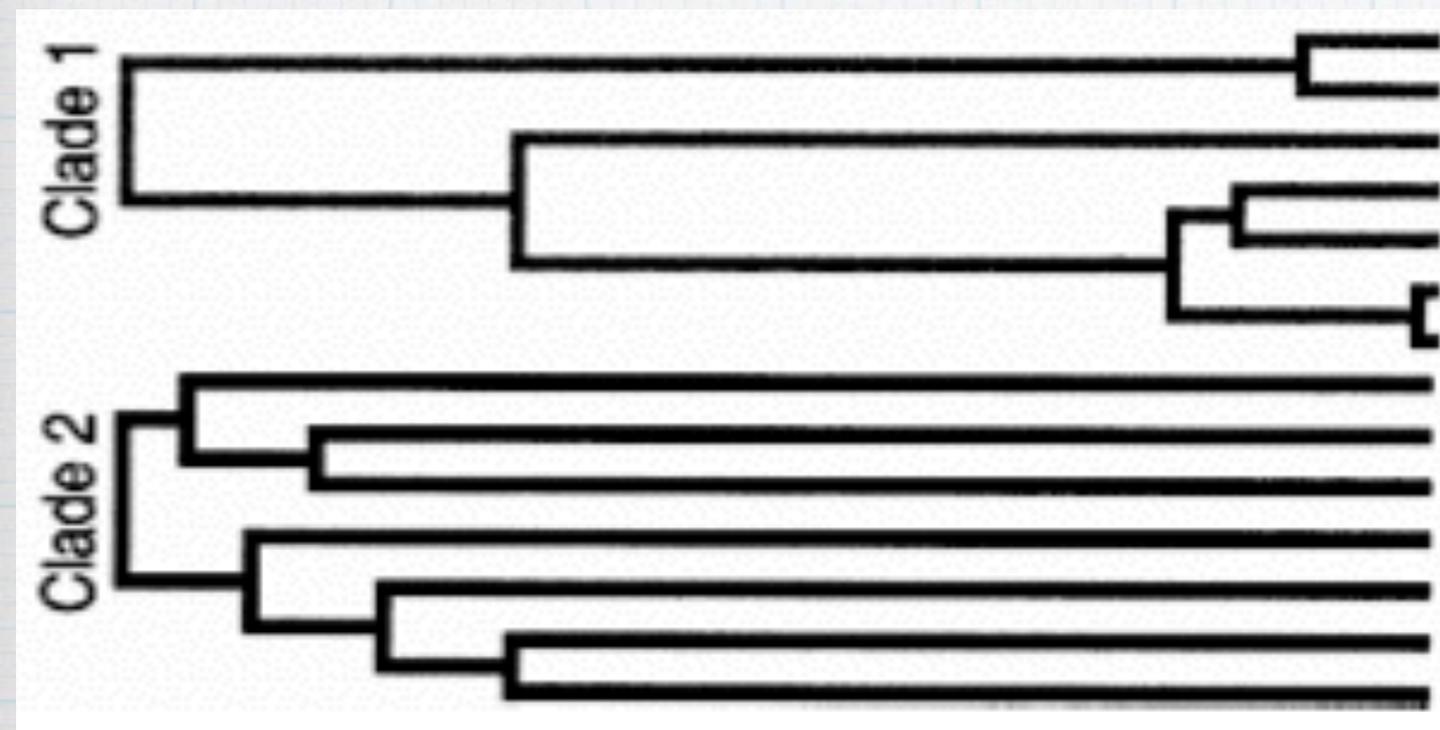
Has morphology evolved more quickly in wrasses?

Perhaps....

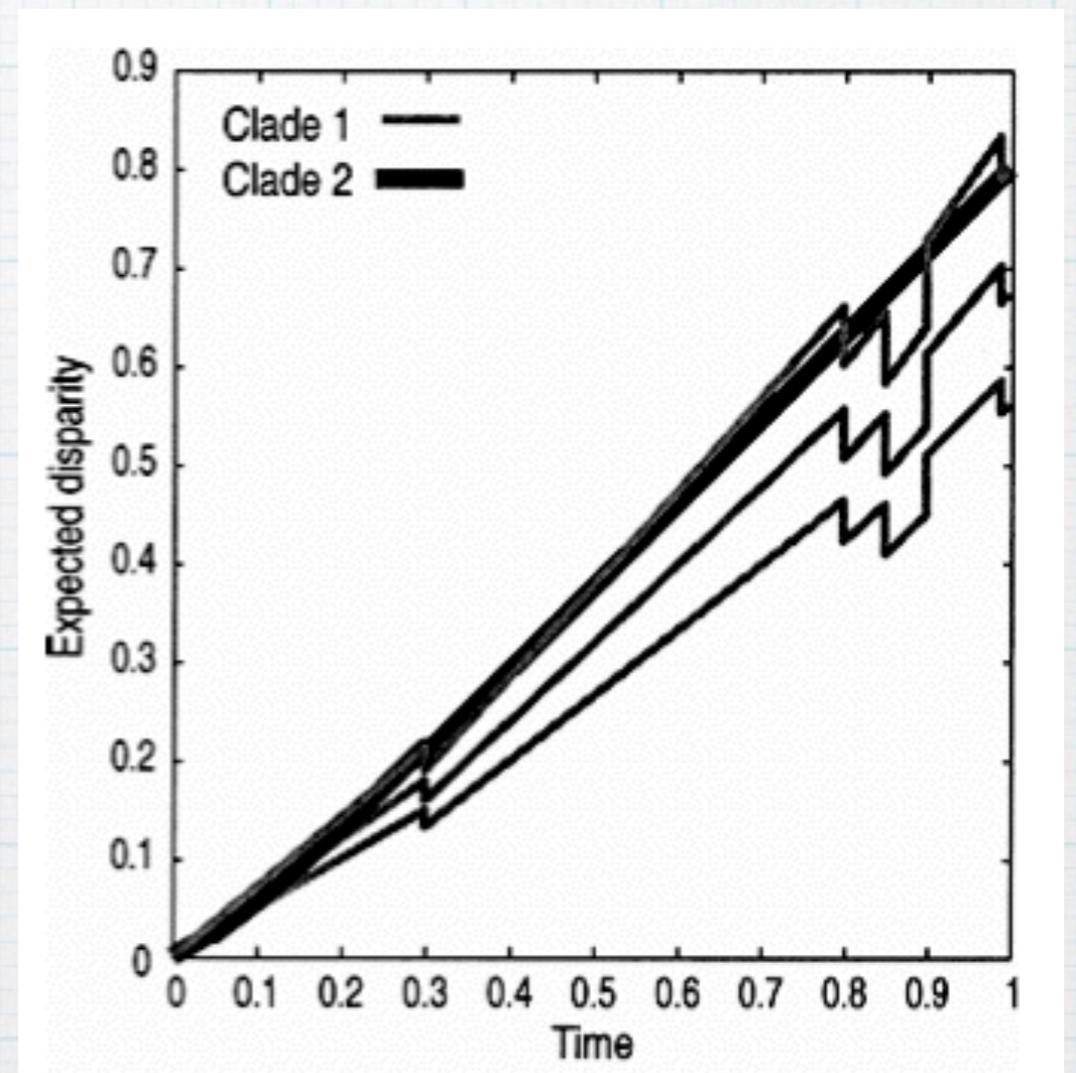
- time and rate of morphological evolution will influence patterns of diversification
- older clades should have more diversity
- how can we tease out influences of time and rate to identify clades that are exceptional in their morphological diversity?



Comparing rates of morphological evolution



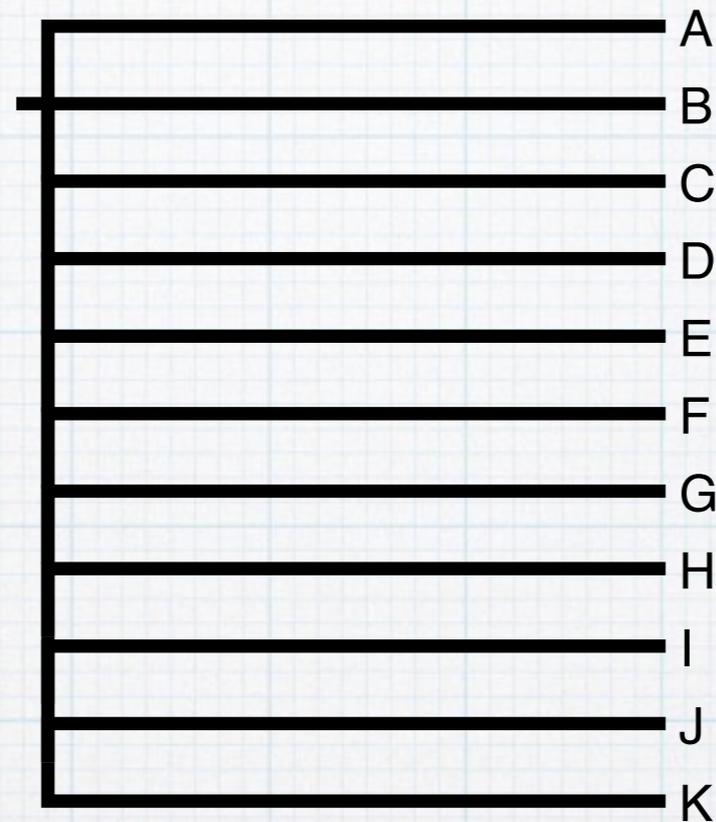
which clade should be more diverse?



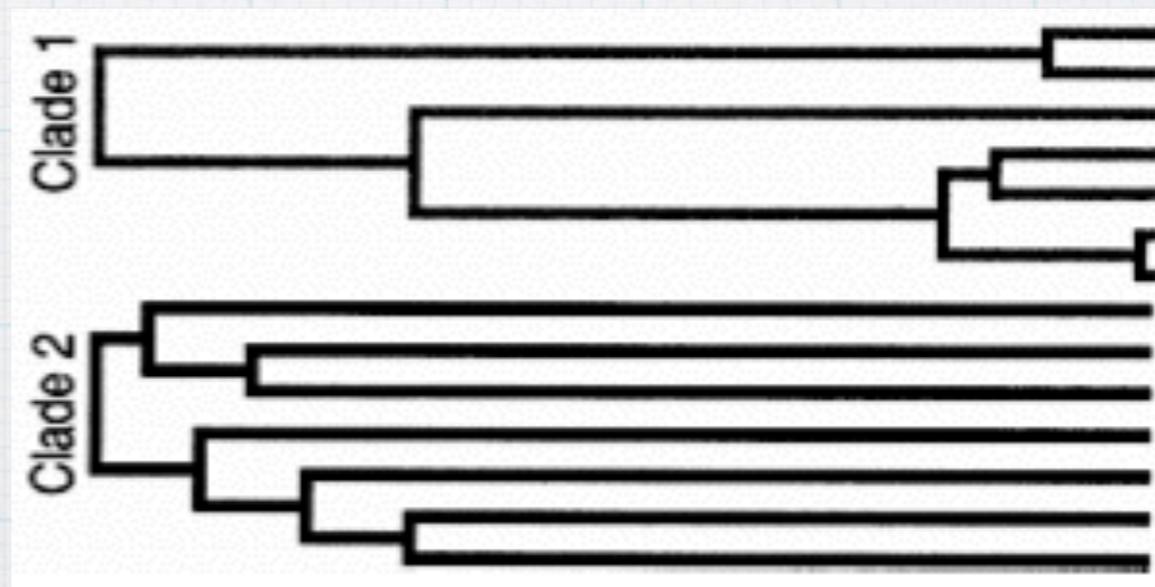
* change along branches is independent since MRCA

* change along common branches is shared

* clades with deep splitting events should have more diversity



expect maximal morphological evolution



expect phylogenetically-constrained morphological evolution

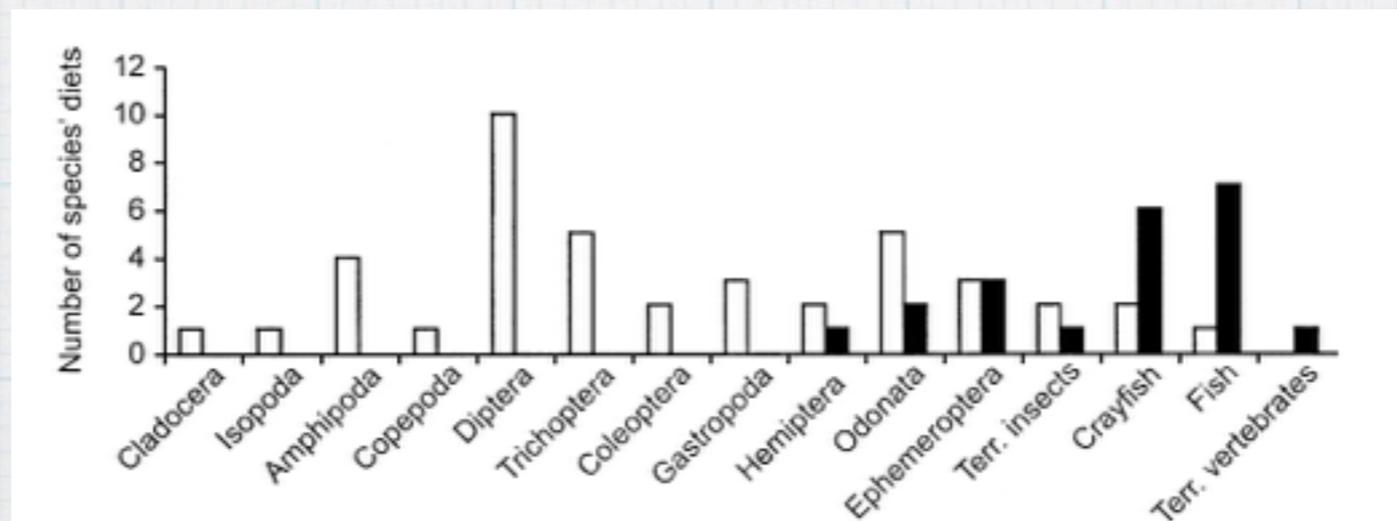
Rates of Morphological Evolution in Centrarchids



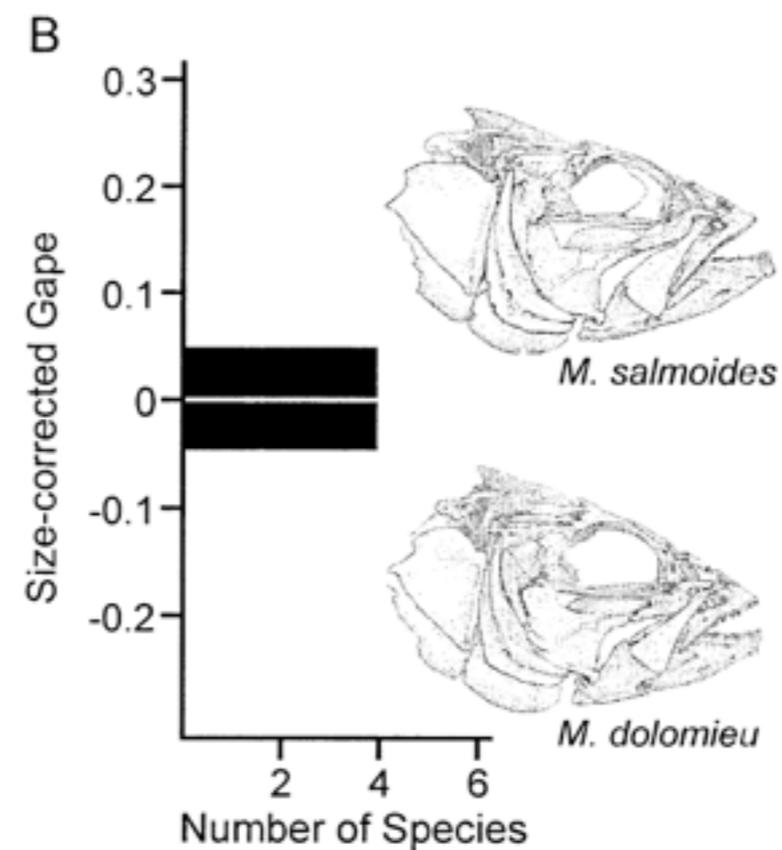
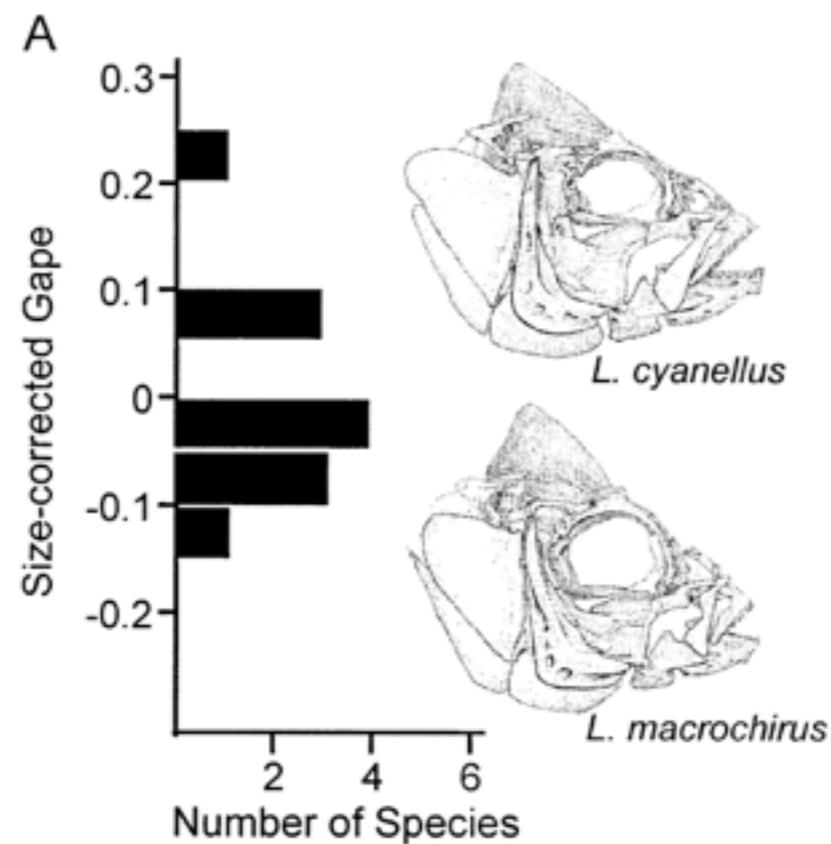
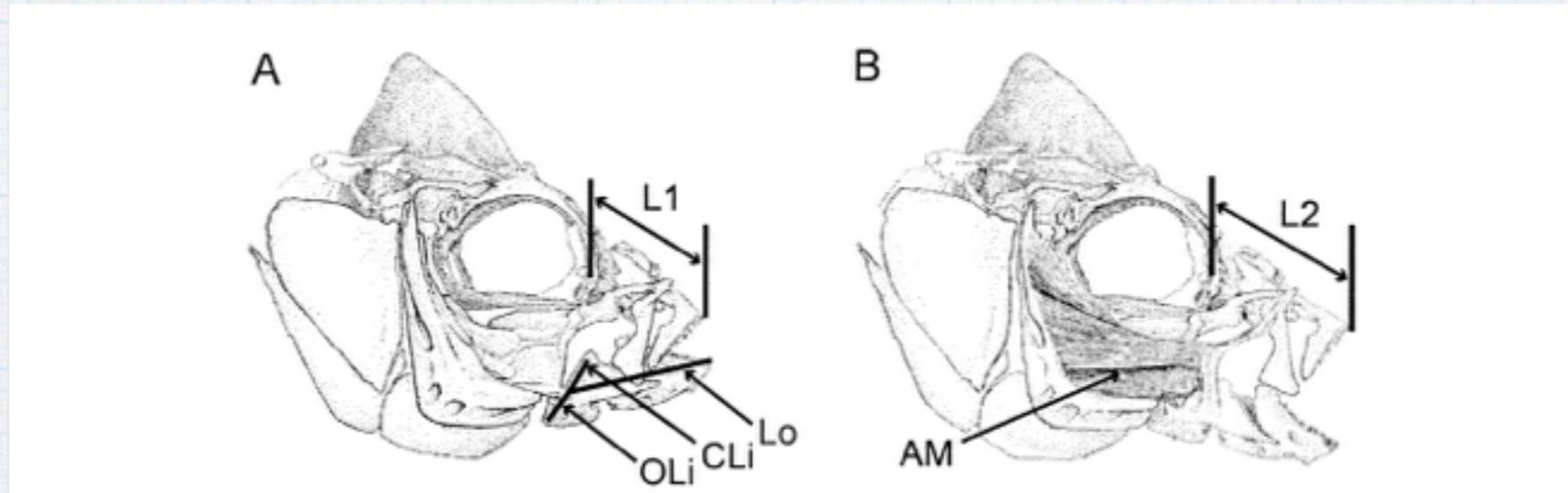
Lepomis (sunfishes): high dietary diversity

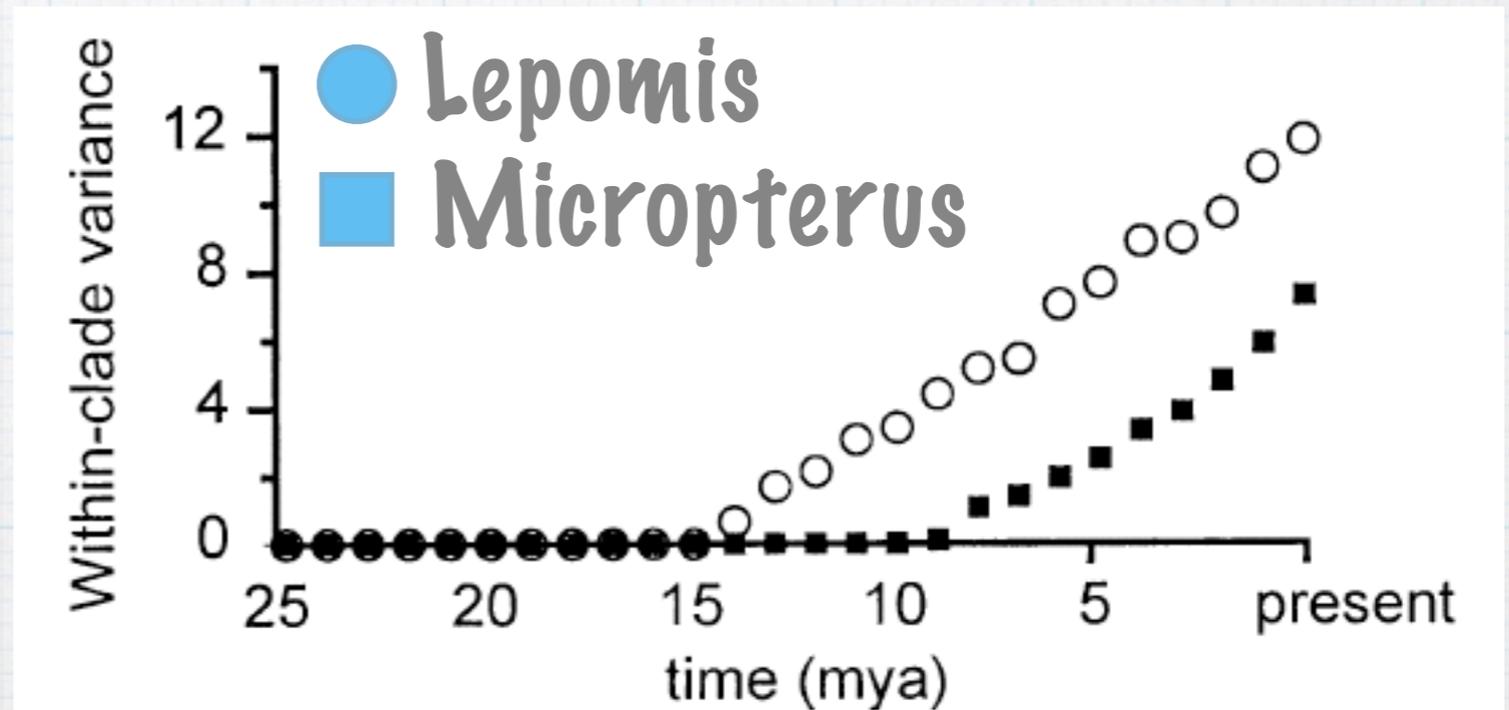
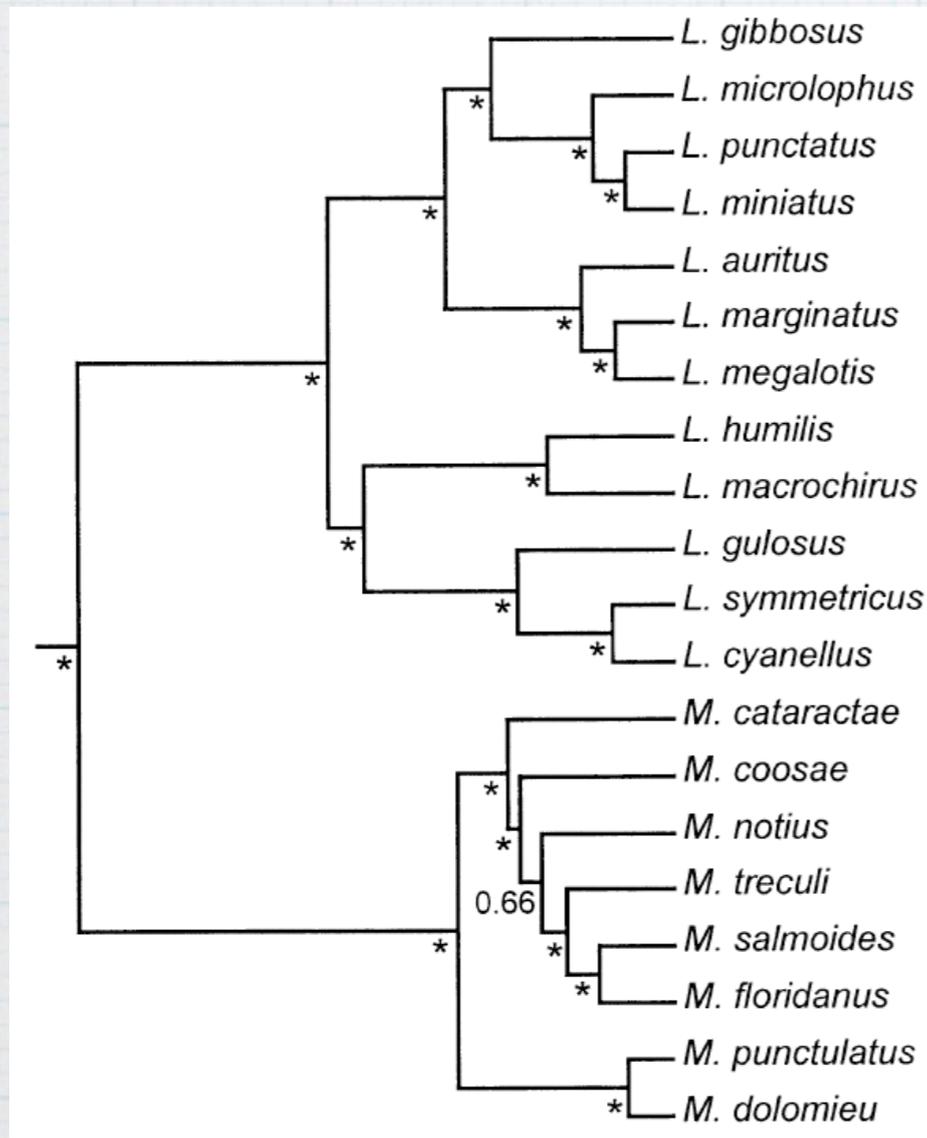


Micropterus (blackbasses): low dietary diversity



Morphological diversity higher in Lepomis

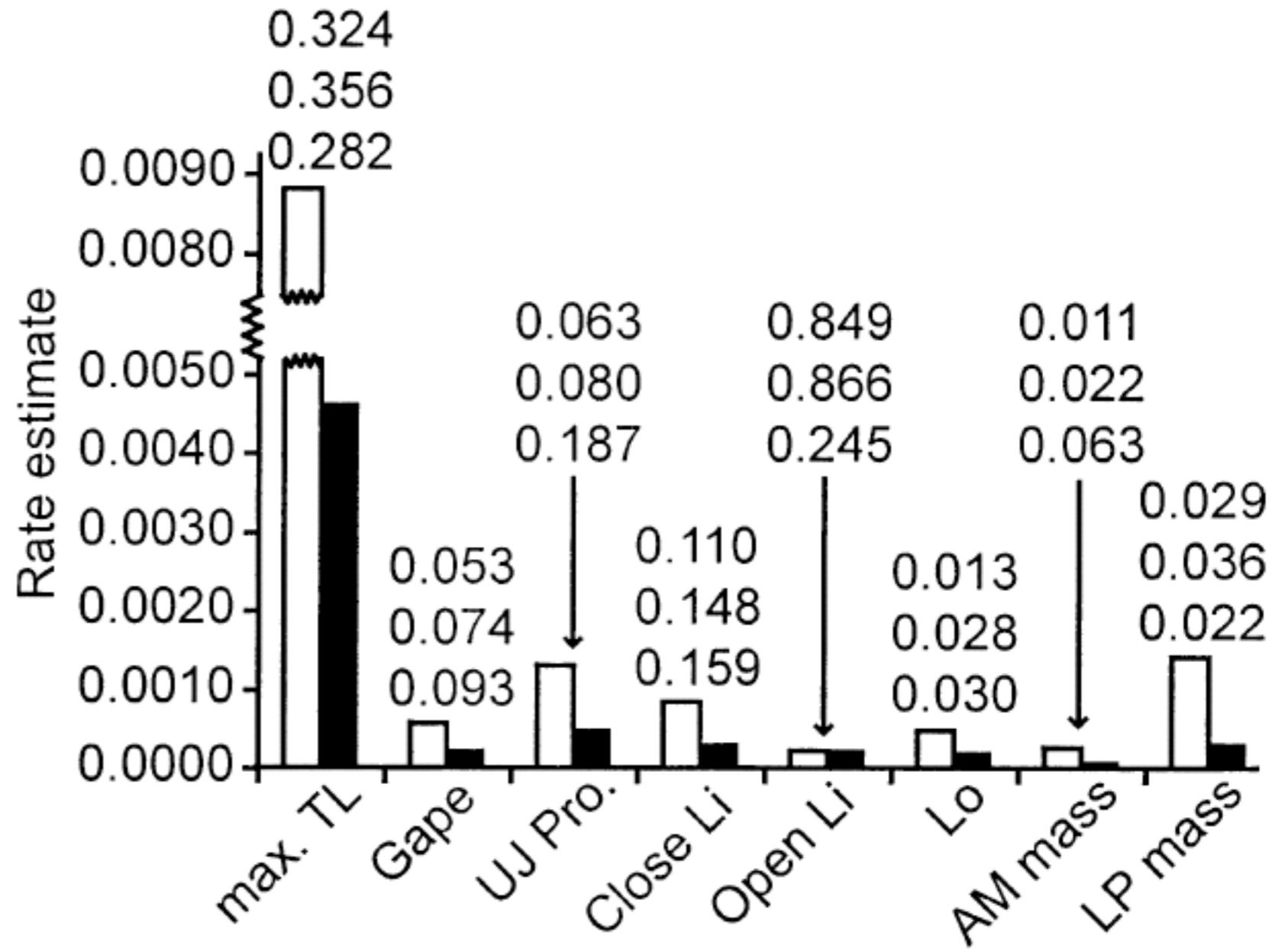




However, we expect this. Is diversity higher in *Lepomis* once we control for clade age?

Yes!

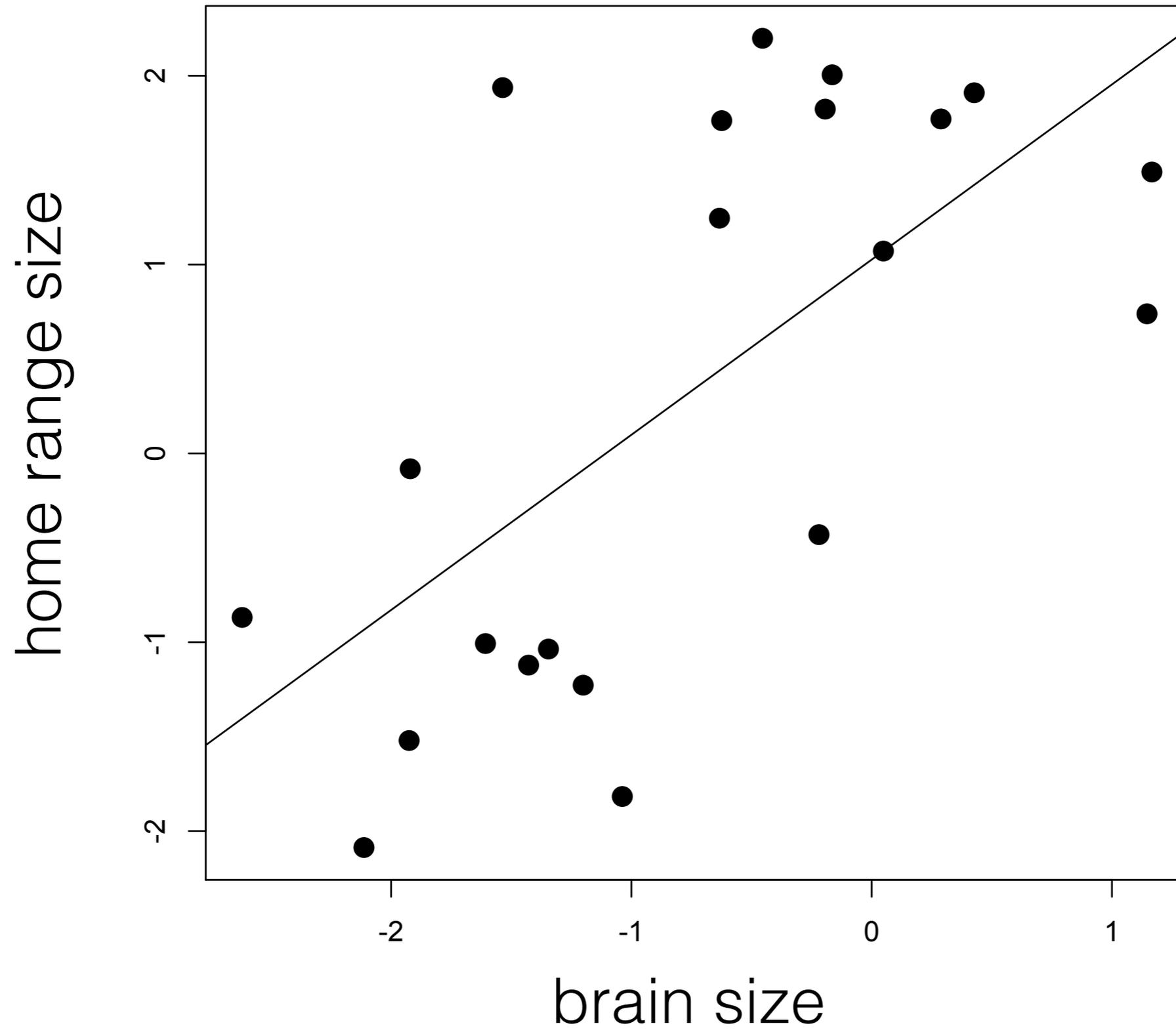
* rates of evolution up to 9X greater in Lepomis



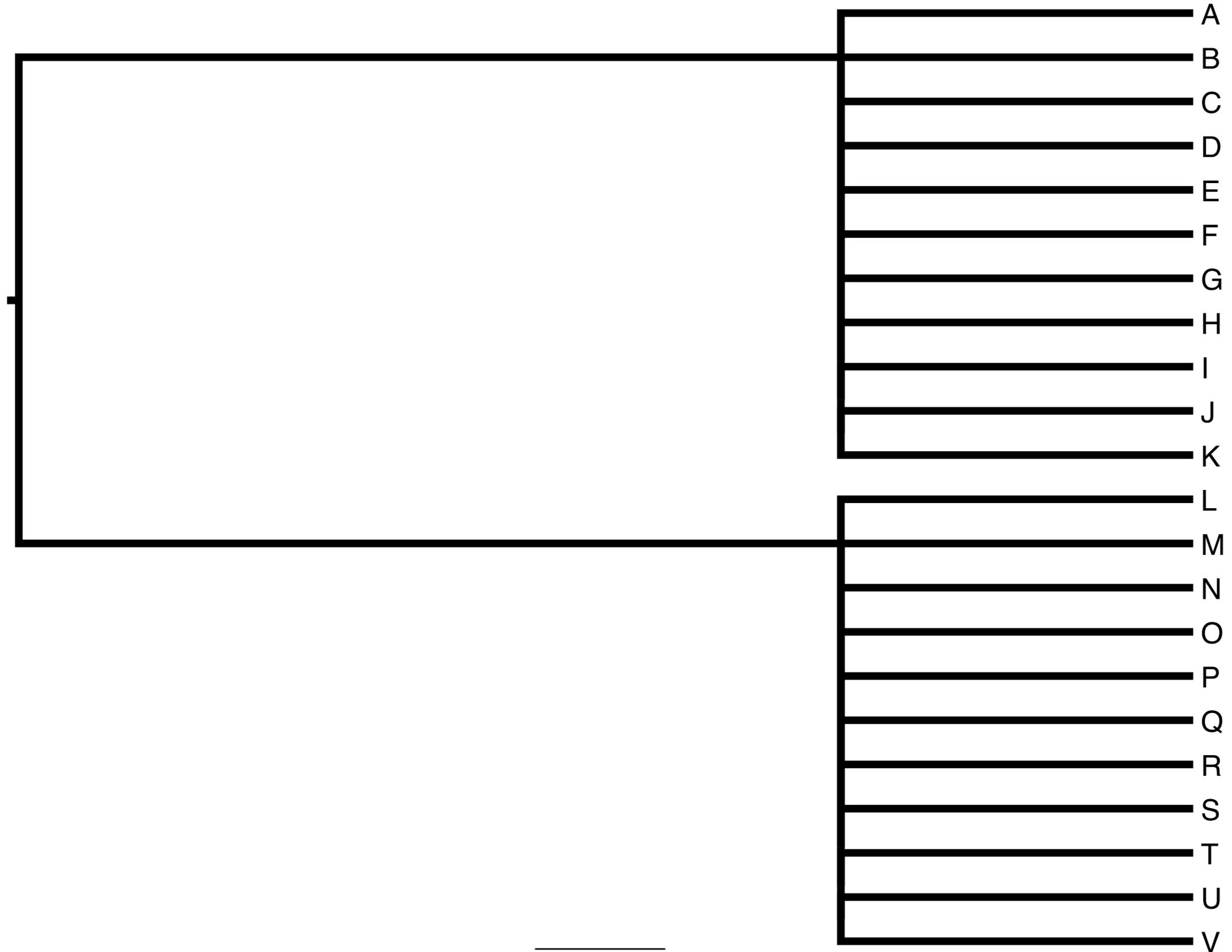
Phylogenetic non-independence

- any set of species in an analysis will share an evolutionary history
- we expect closely related species to be more similar to each other than more distantly related species
- treating species as independent in statistical analyses can be misleading

do species with bigger brains have larger home ranges?



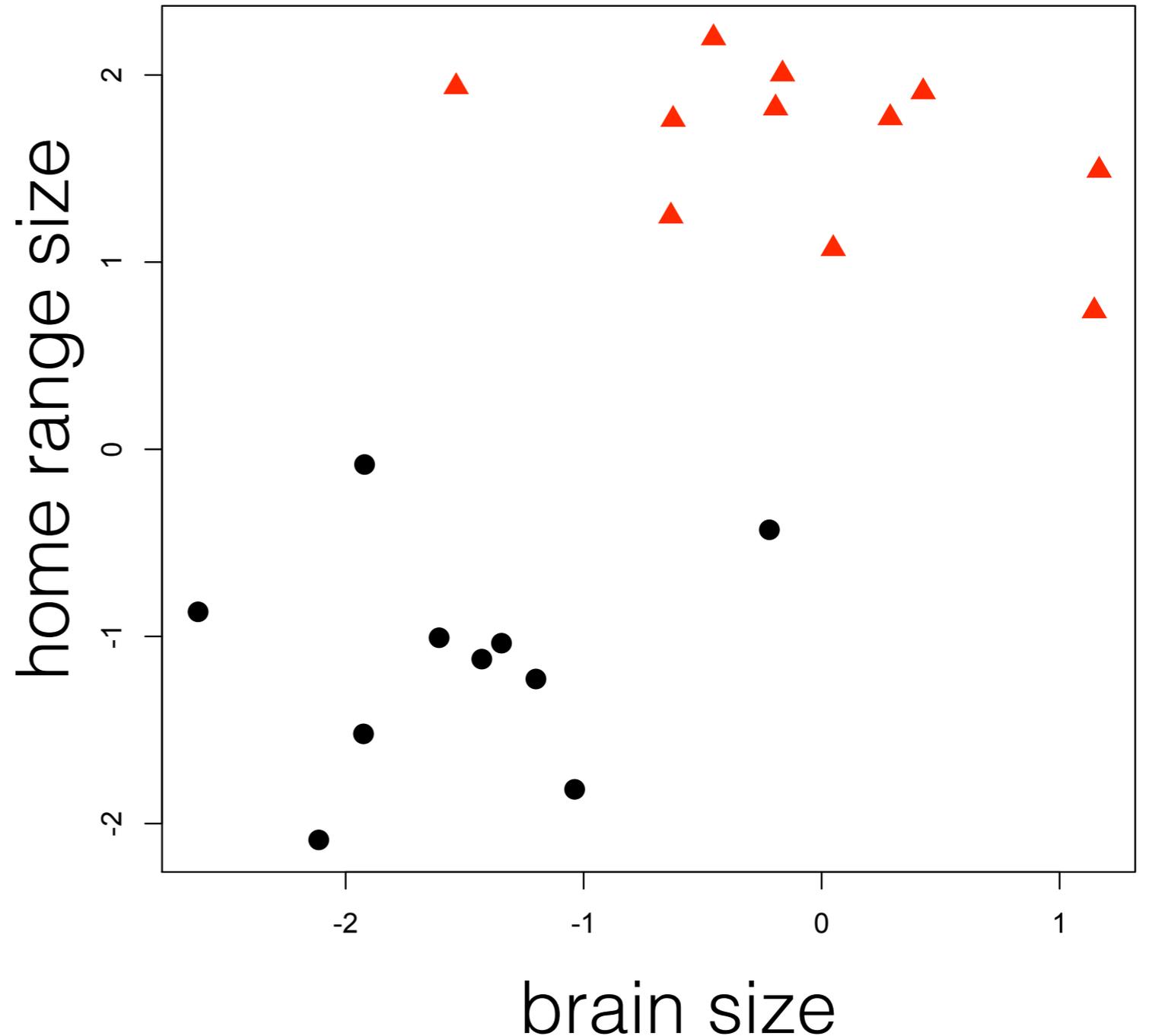
what if the phylogeny looks like this?



0.11

Need to account for phylogeny

- effectively have closer to two independent points, instead of 20
- are brain size and range size related?



Independent contrasts is a procedure for accounting for phylogenetic relatedness when making interspecific trait comparisons

- input: N trait measures (continuous value), tree
- output: $N-1$ **contrasts**
- contrasts can be used in statistical analysis (correlation, regression, ANOVA, etc)

Motivation

- species pairs share a common ancestor



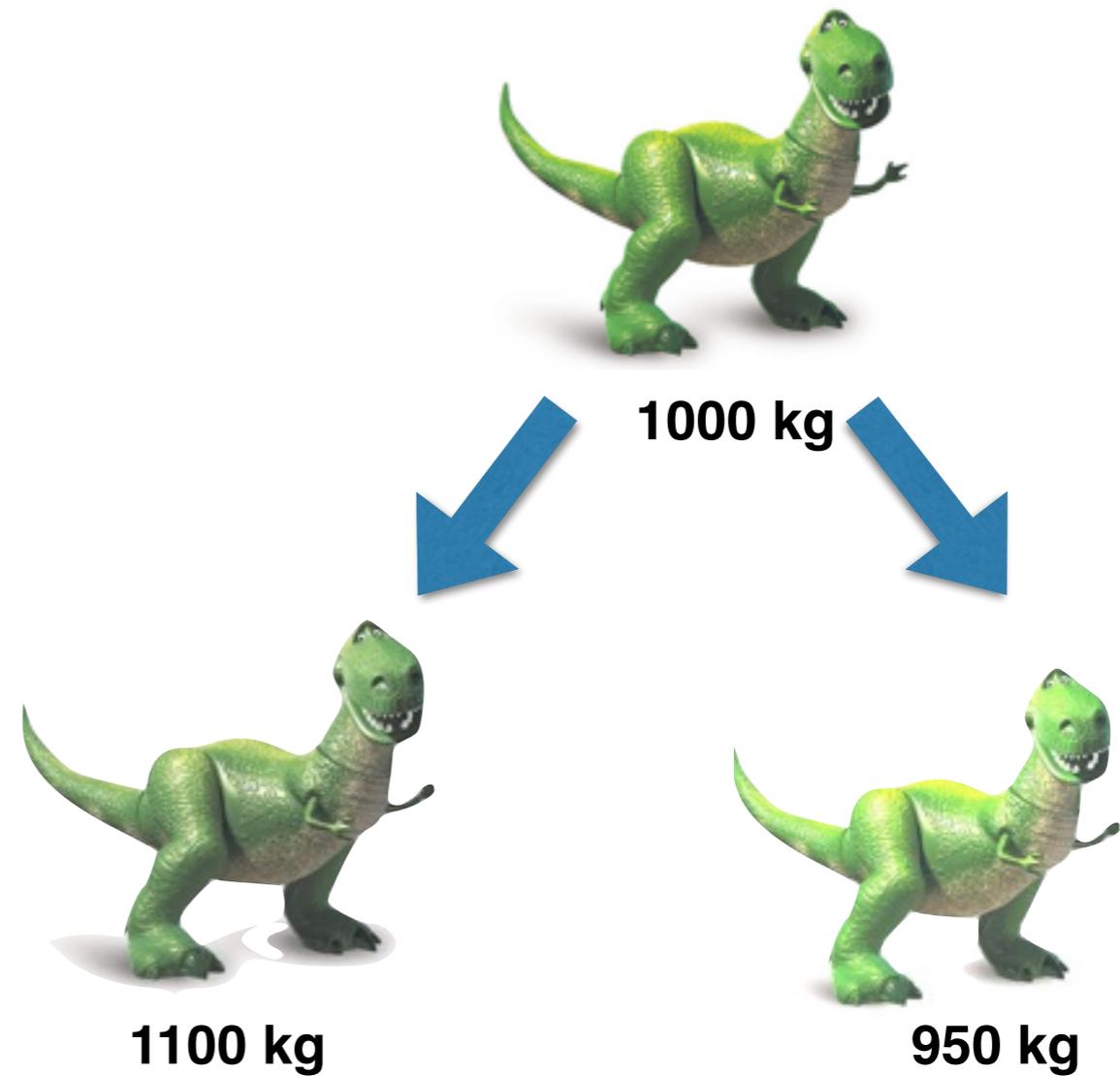
1100 kg



950 kg

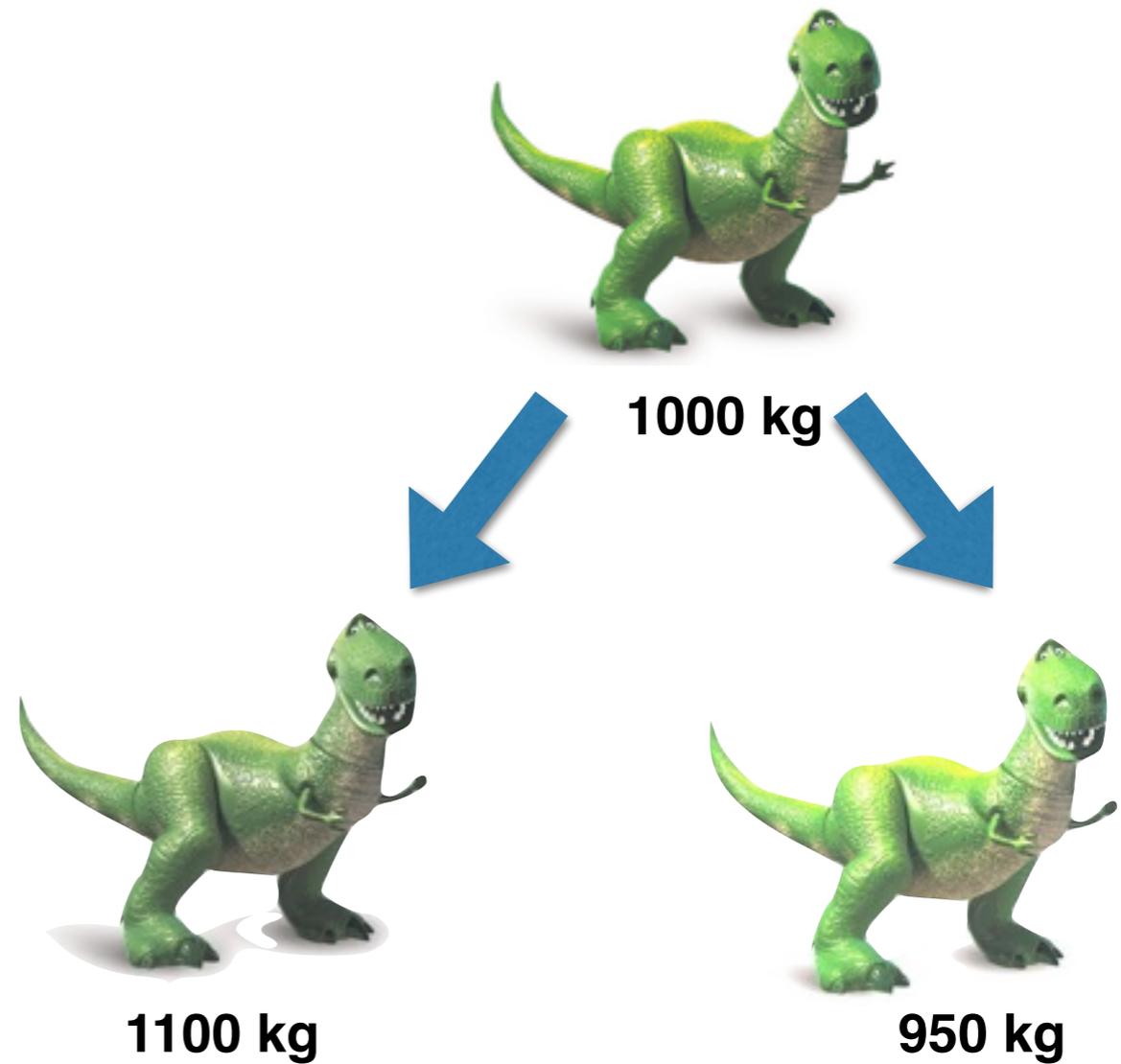
Motivation

- species pairs share trait value at ancestor



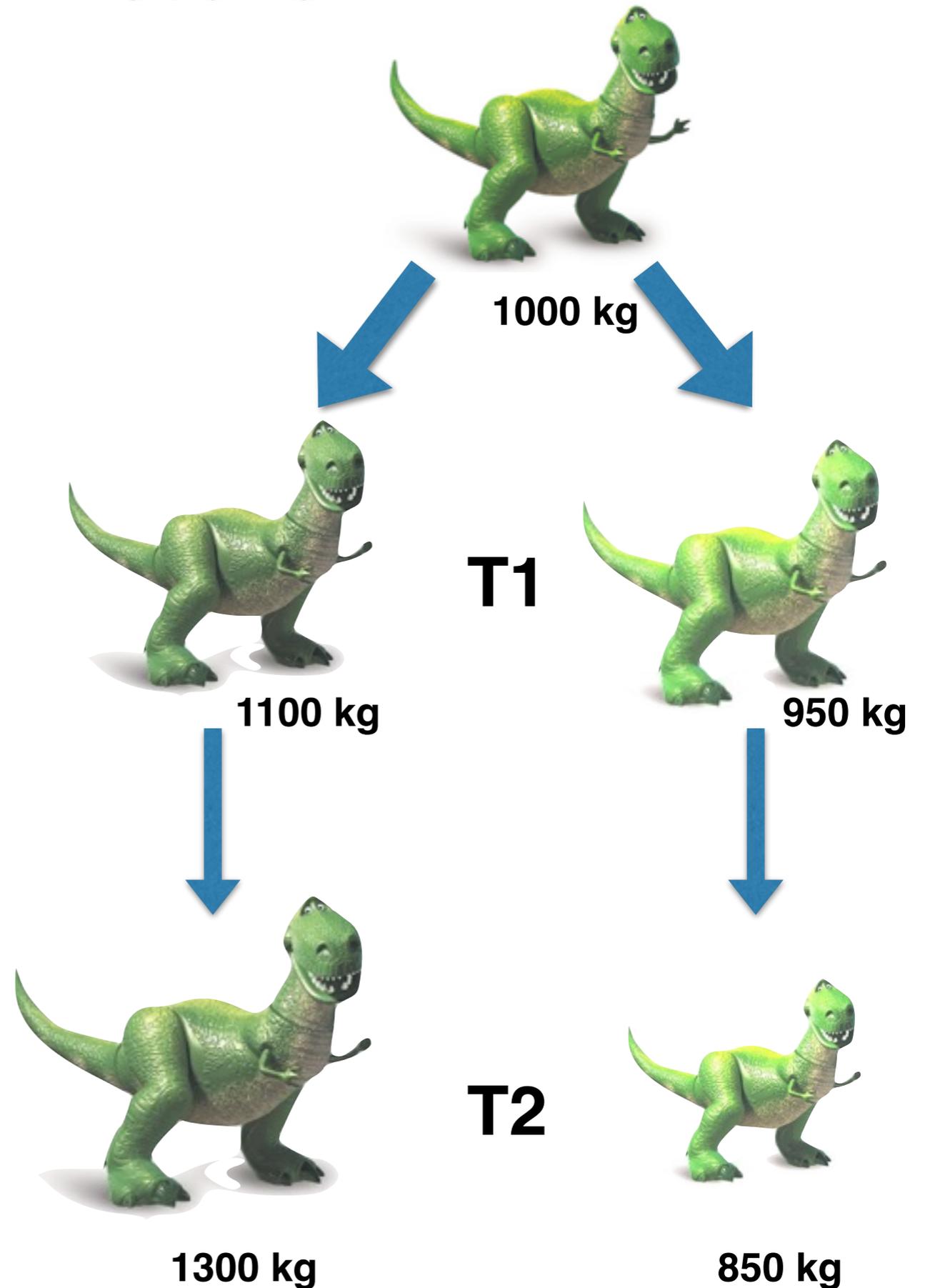
Motivation

- species pairs share trait value at ancestor
- with time since divergence, *difference* should increase



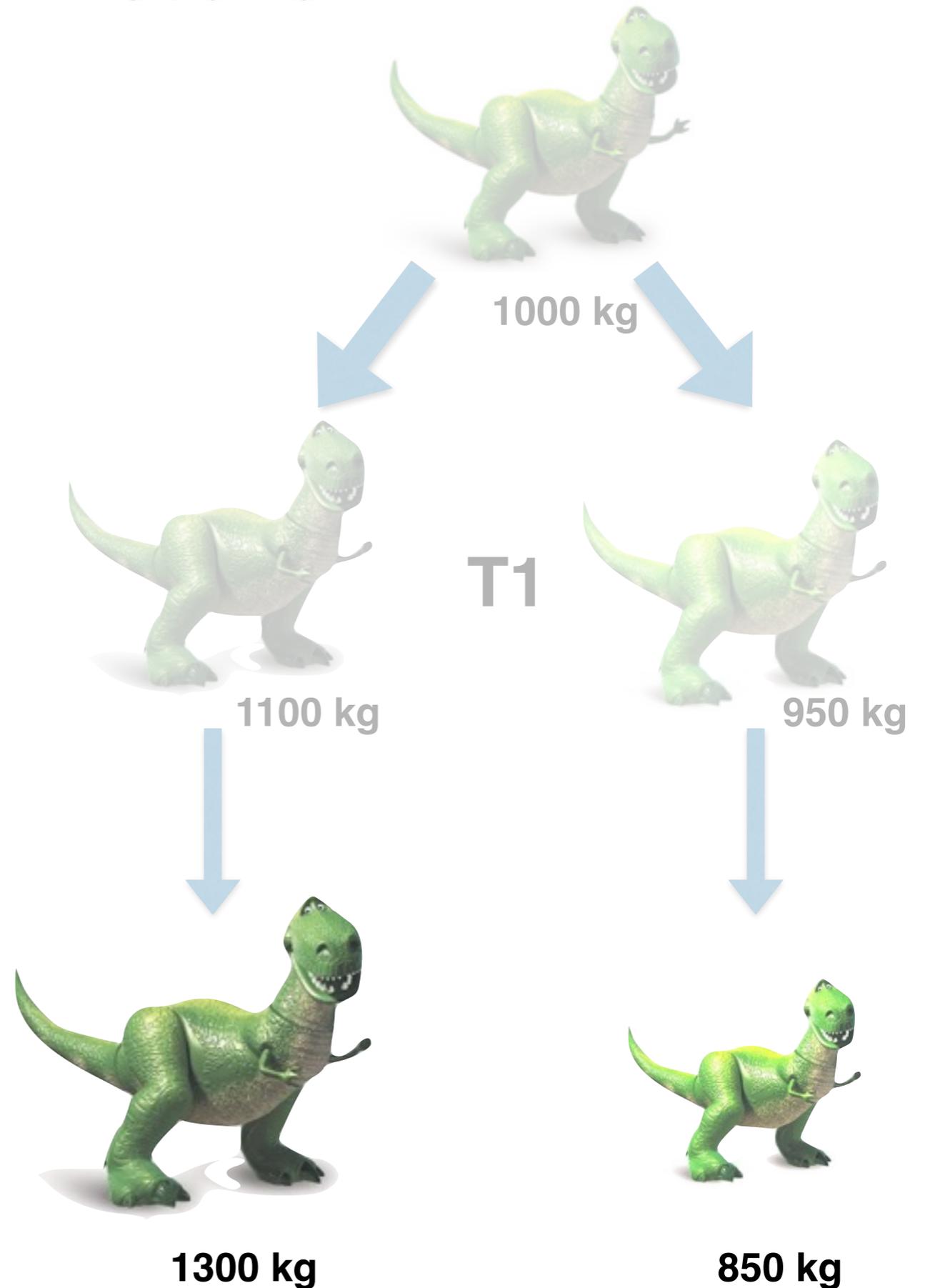
Motivation

- species pairs share trait value at ancestor
- with time since divergence, *difference* should increase



Motivation

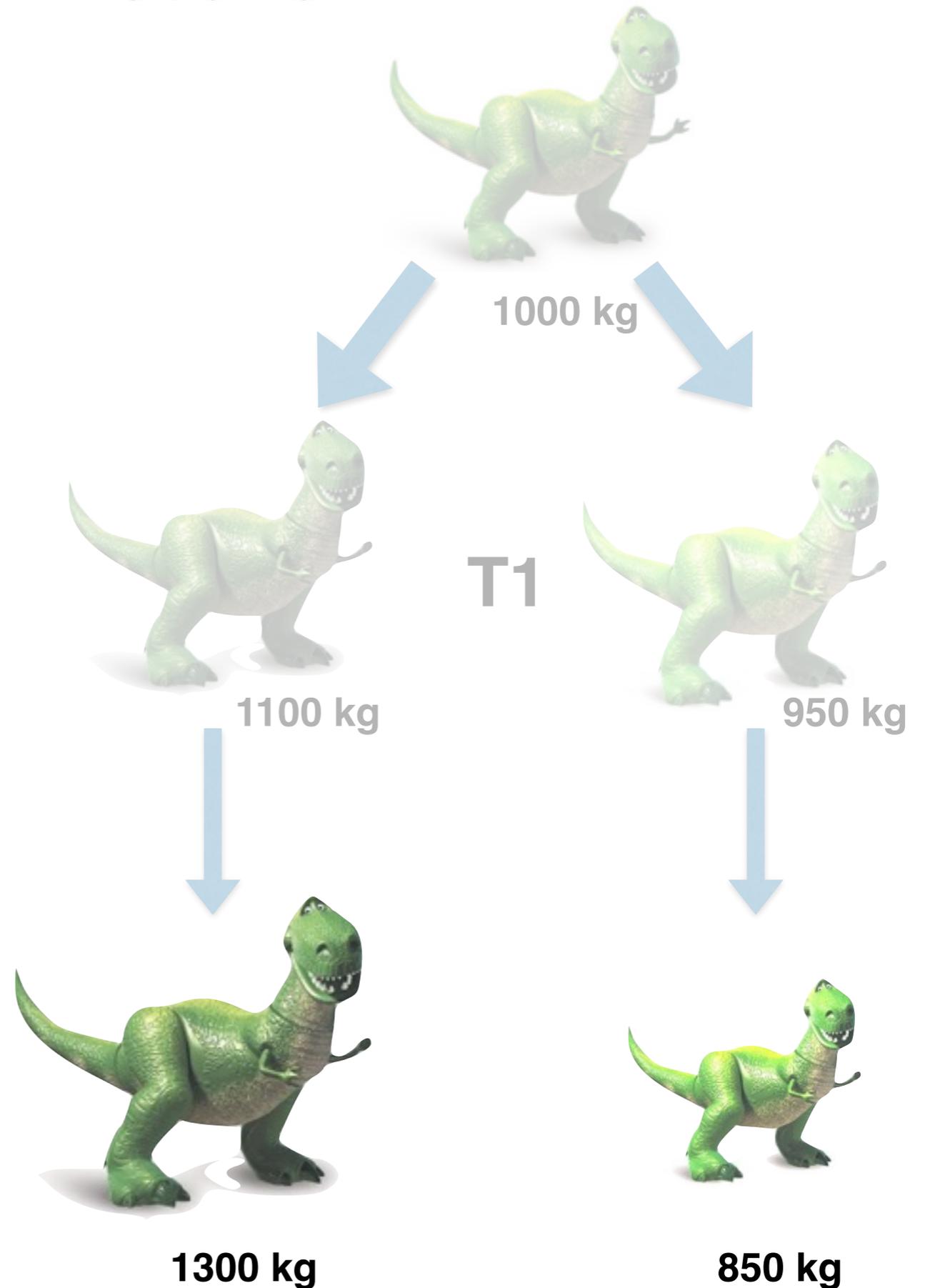
- a contrast tells you how much difference has evolved since shared common ancestor



1300 kg - 850 kg

Motivation

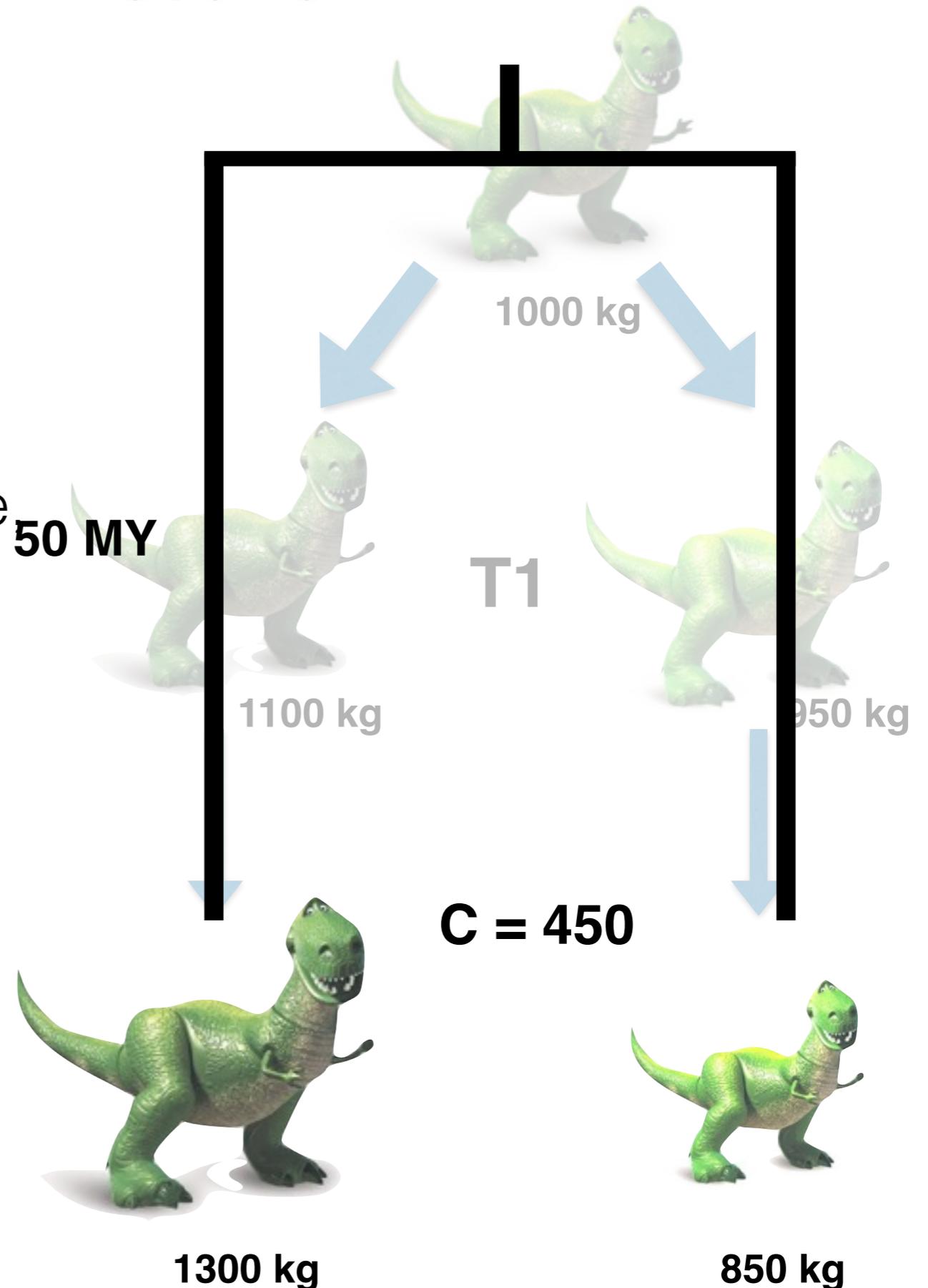
- a contrast tells you how much difference has evolved since shared common ancestor



1300 kg - 850 kg

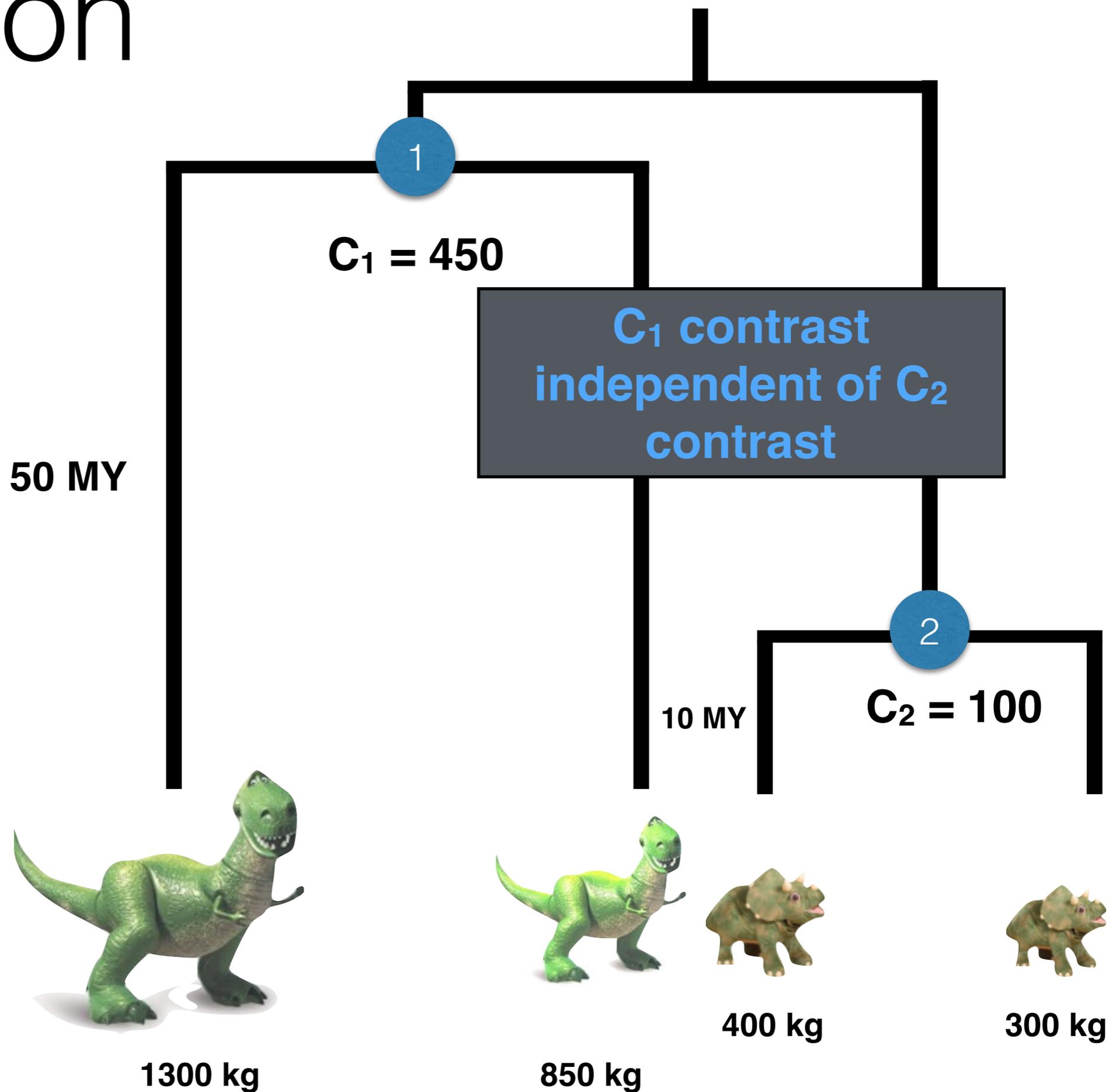
Motivation

- species pairs share trait value at ancestor
- with time since divergence *difference* should increase
- a contrast tells you how much difference has evolved since common ancestor
- (with branch lengths, we can learn about **rate**)



Motivation

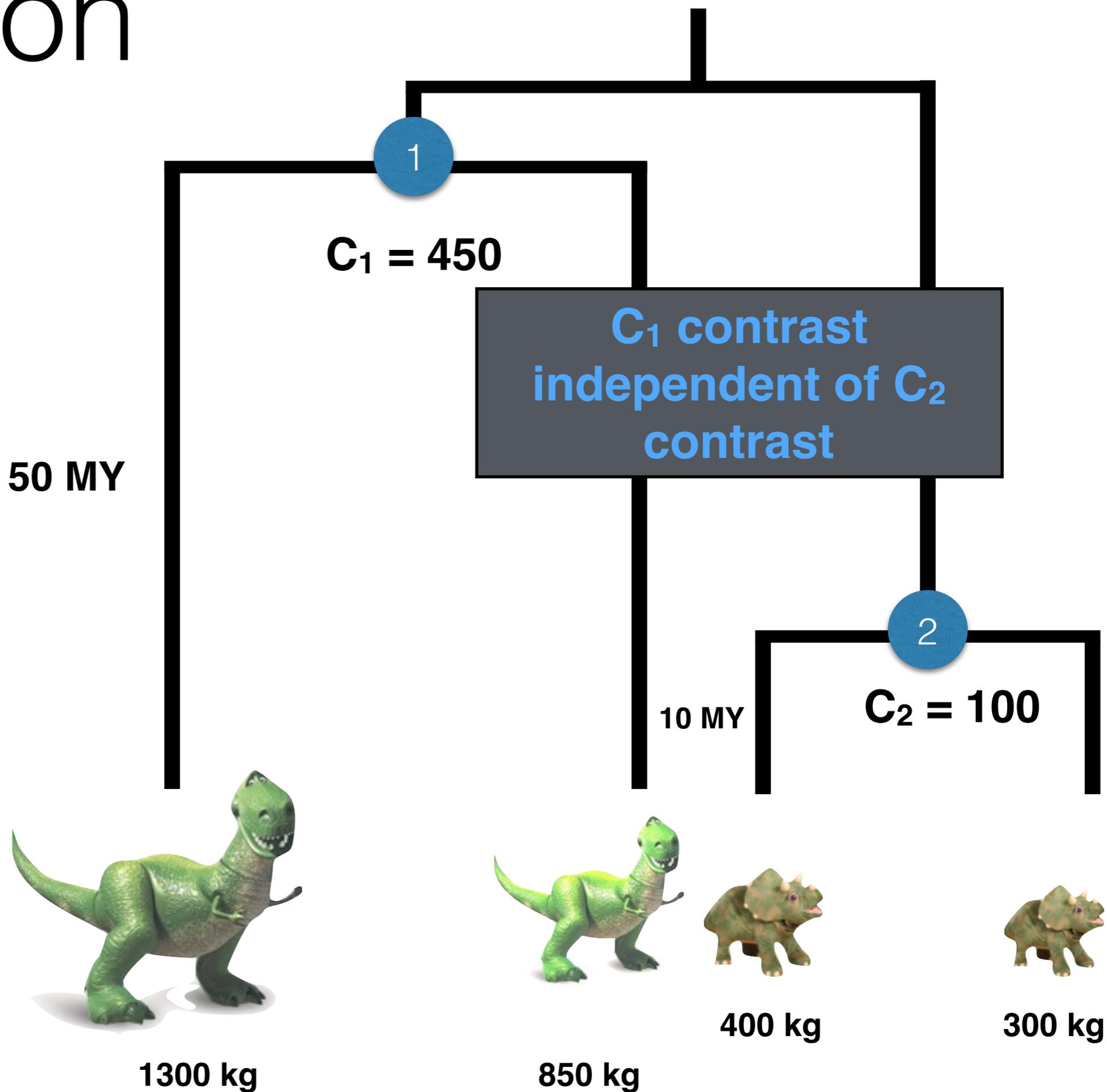
- a contrast tells you how much difference has evolved
- contrasts from other subclades are *independent*



Motivation

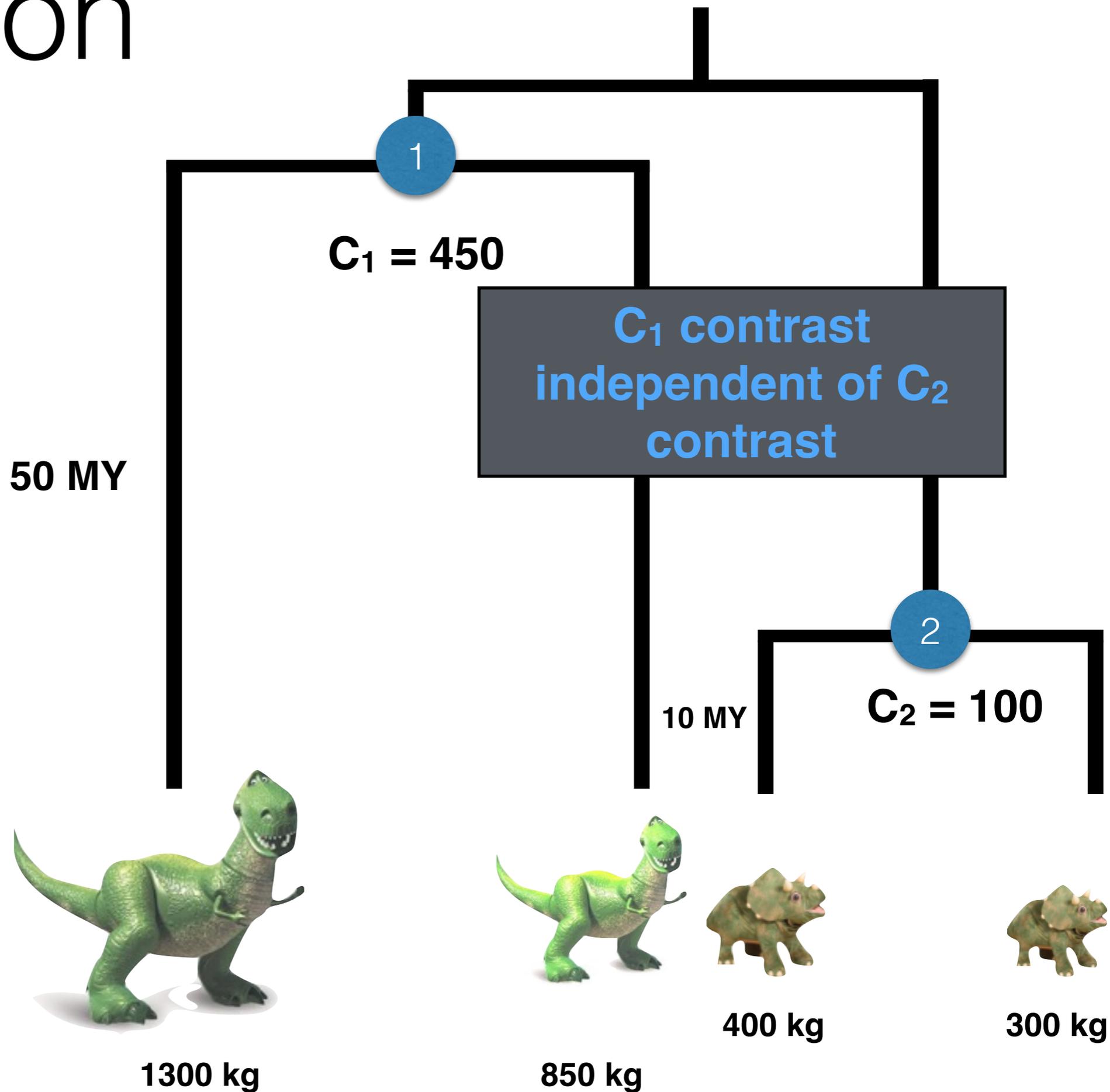
- a contrast tells you how much difference has evolved
- contrasts from other subclades are *independent*

are you surprised that C_1 is greater than C_2 ?



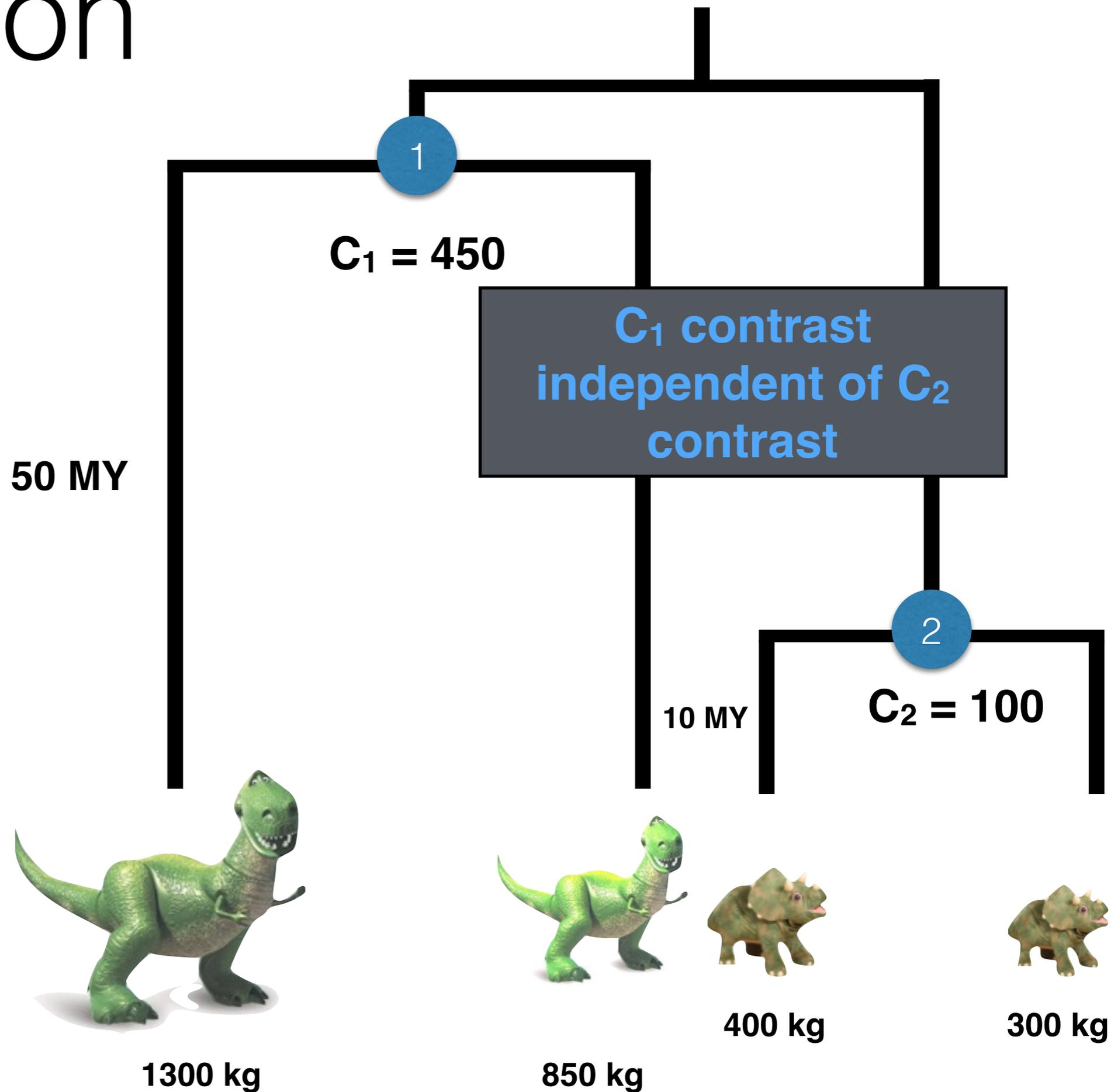
Motivation

- a contrast tells you how much difference has evolved
- contrasts from other subclades are *independent*
- we need to account for branch lengths to compare contrasts

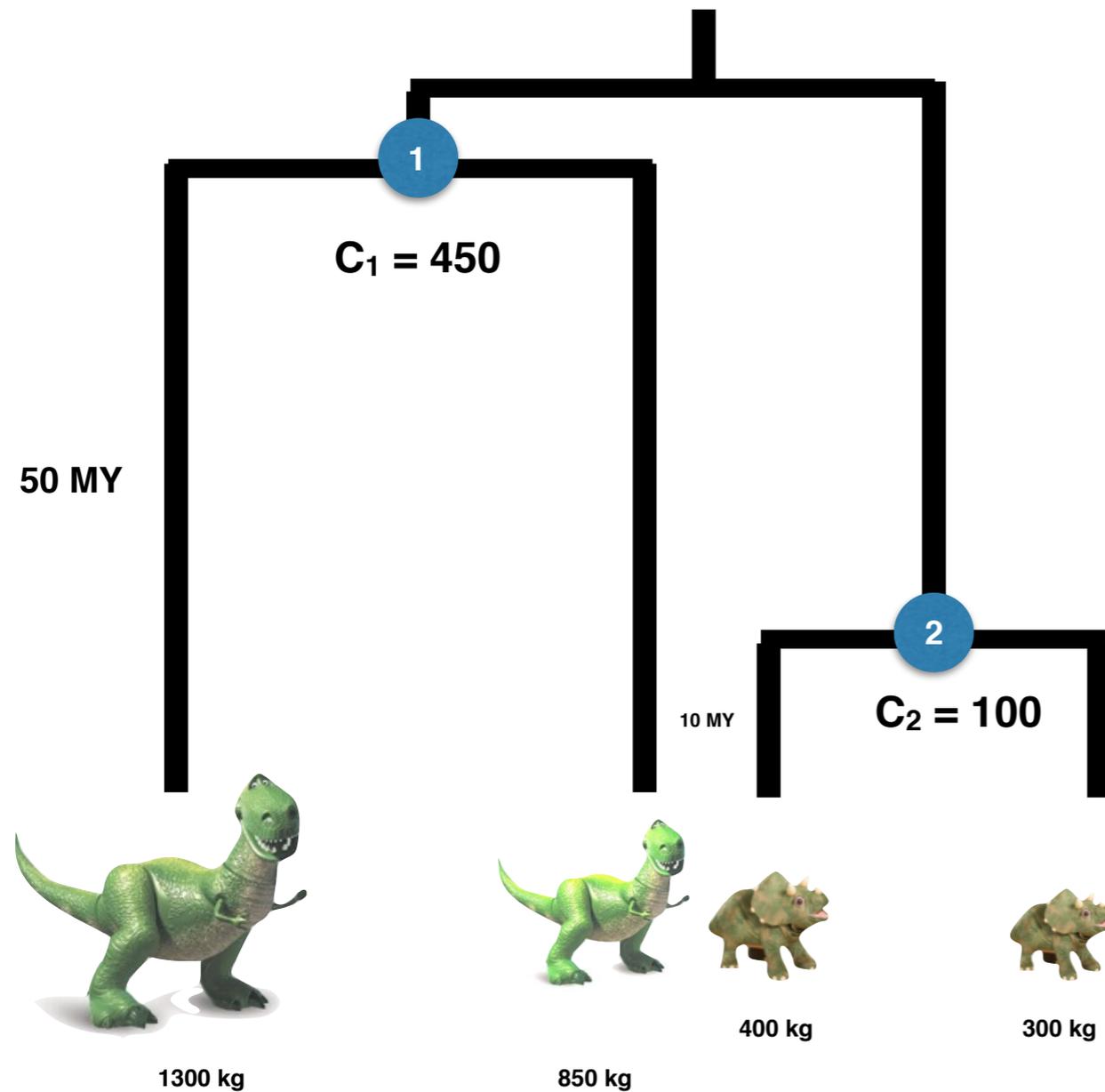


Motivation

- a contrast tells you how much difference has evolved
- contrasts from other subclades are *independent*
- we need to account for branch lengths to compare contrasts
- once we do this, contrasts provide independent estimates of the evolutionary rate

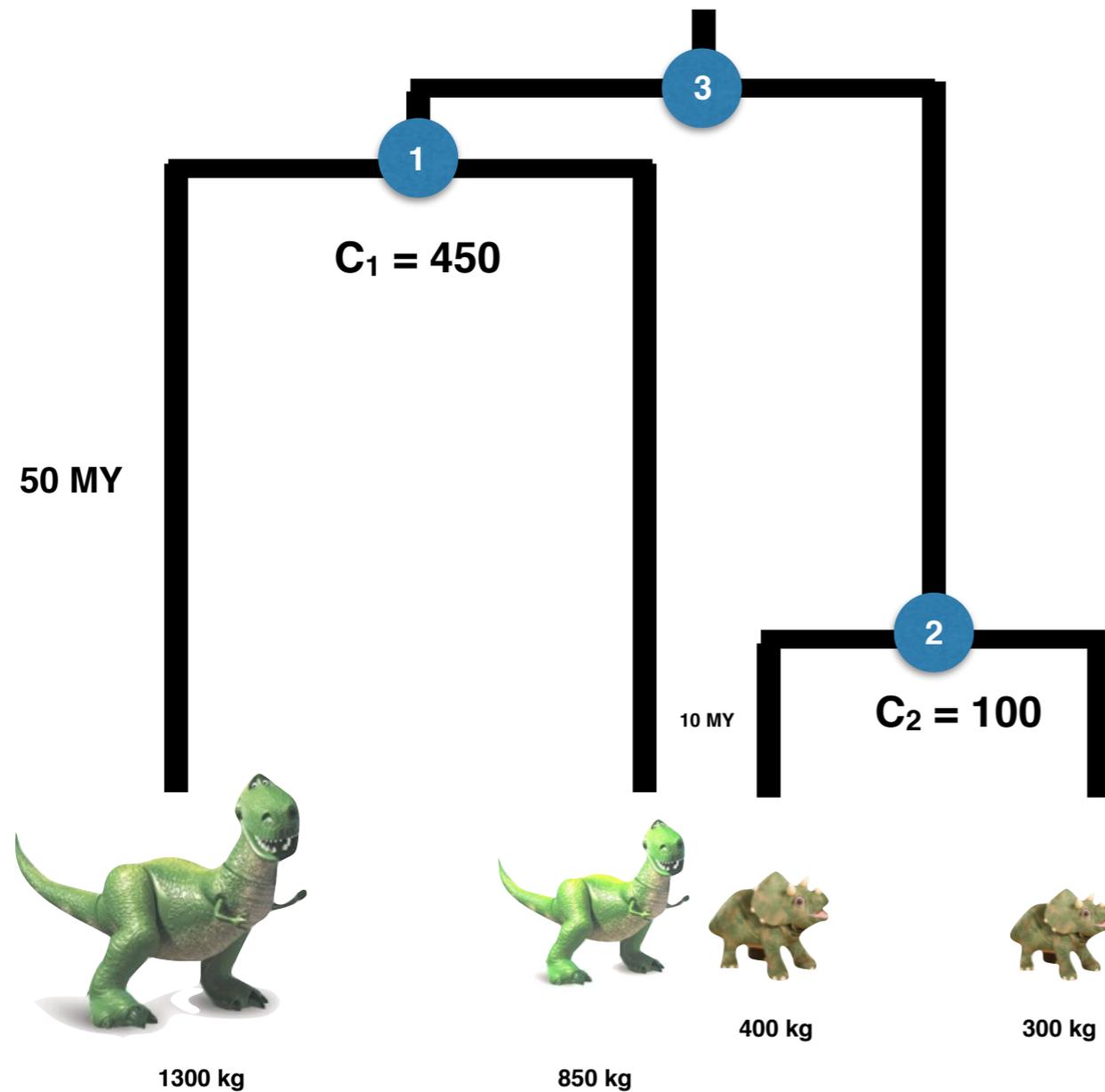


Any phylogenetic tree with N tips provides for N-1 independent contrasts



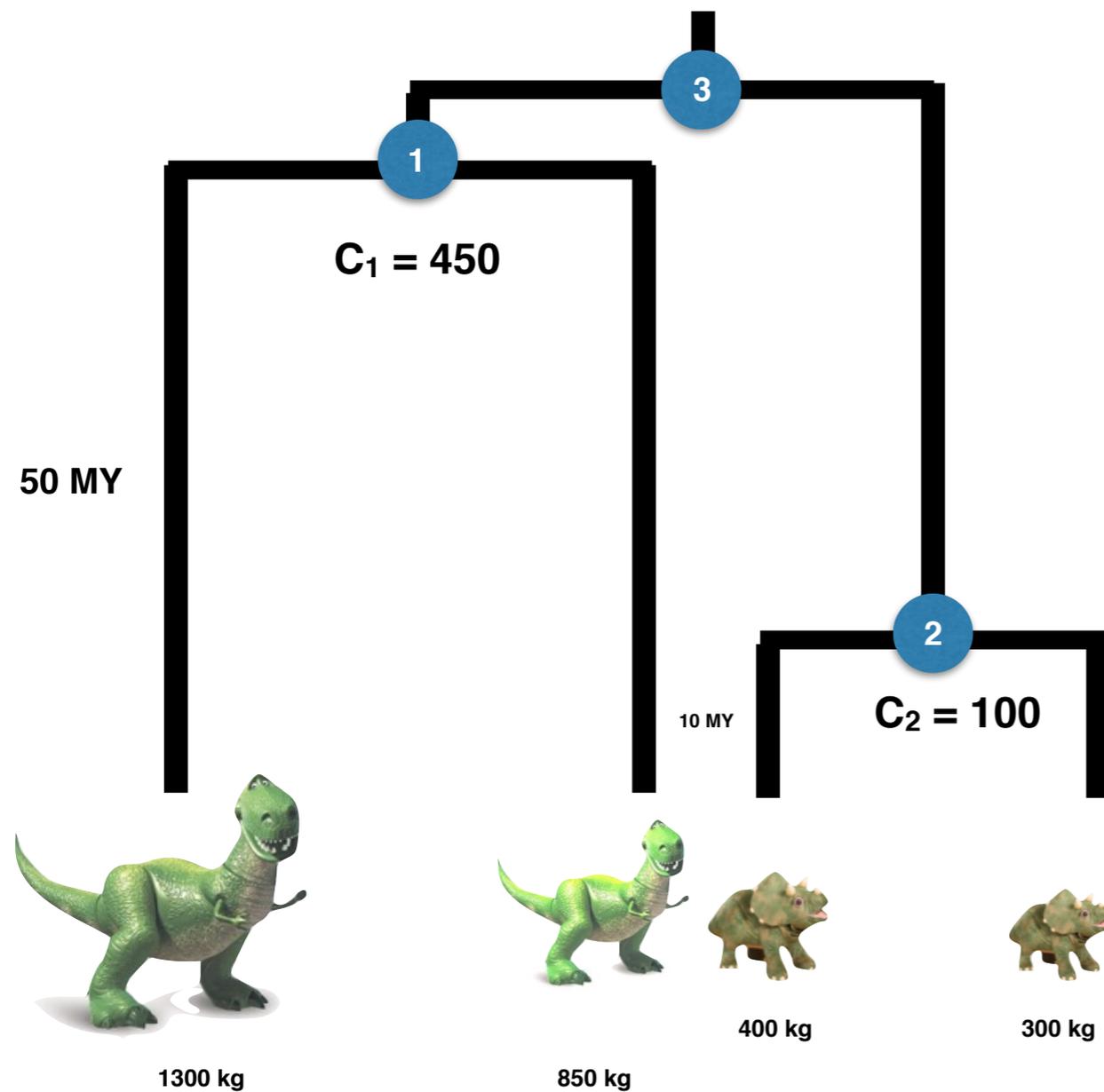
where is the third contrast?

Any phylogenetic tree with N tips provides for N-1 independent contrasts



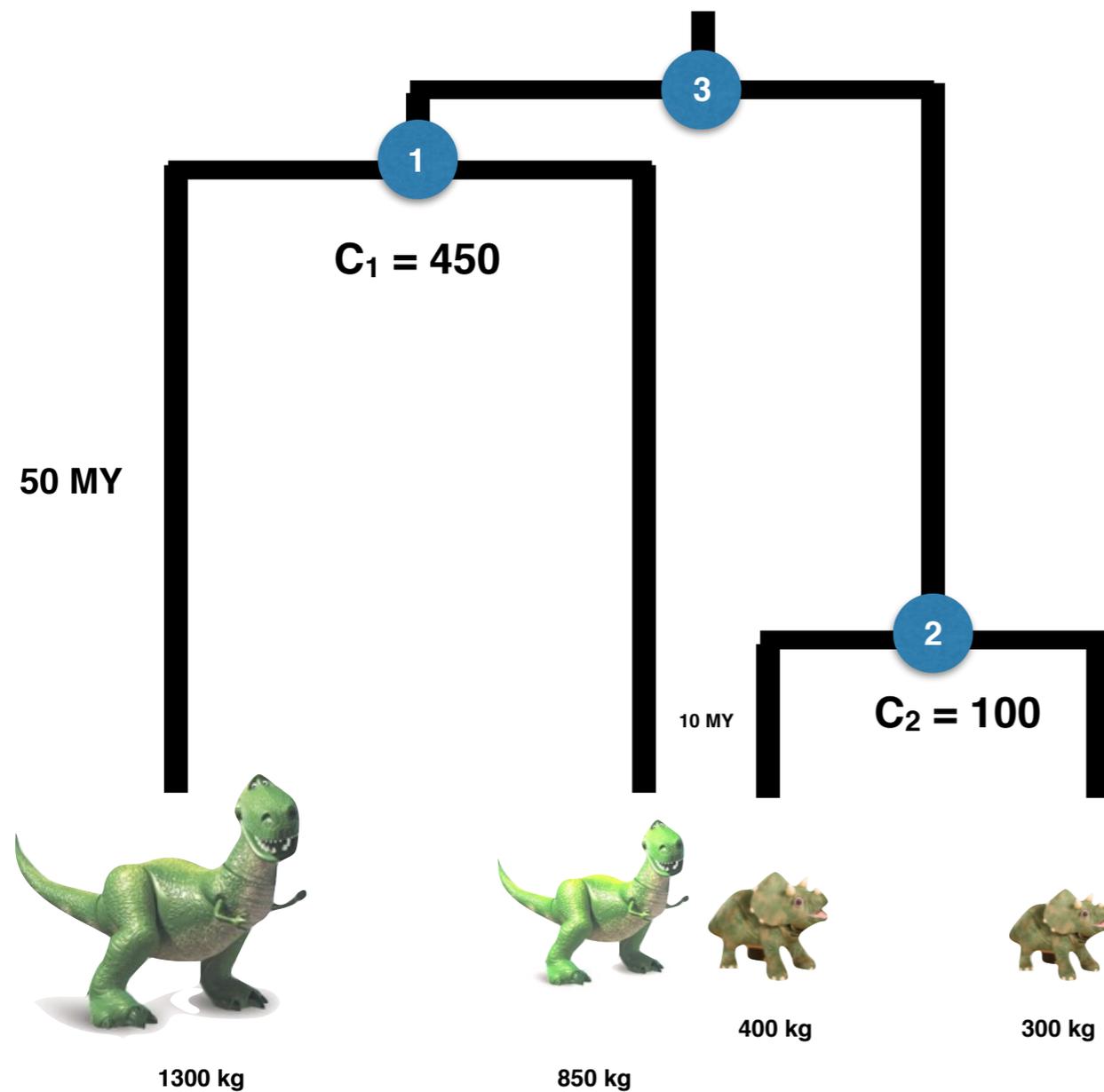
what is unusual about this contrast?

Any phylogenetic tree with N tips provides for N-1 independent contrasts



C_3 is contrast in values between (unobserved) species value of node 1 and (unobserved) species value of node 2

Any phylogenetic tree with N tips provides for N-1 independent contrasts

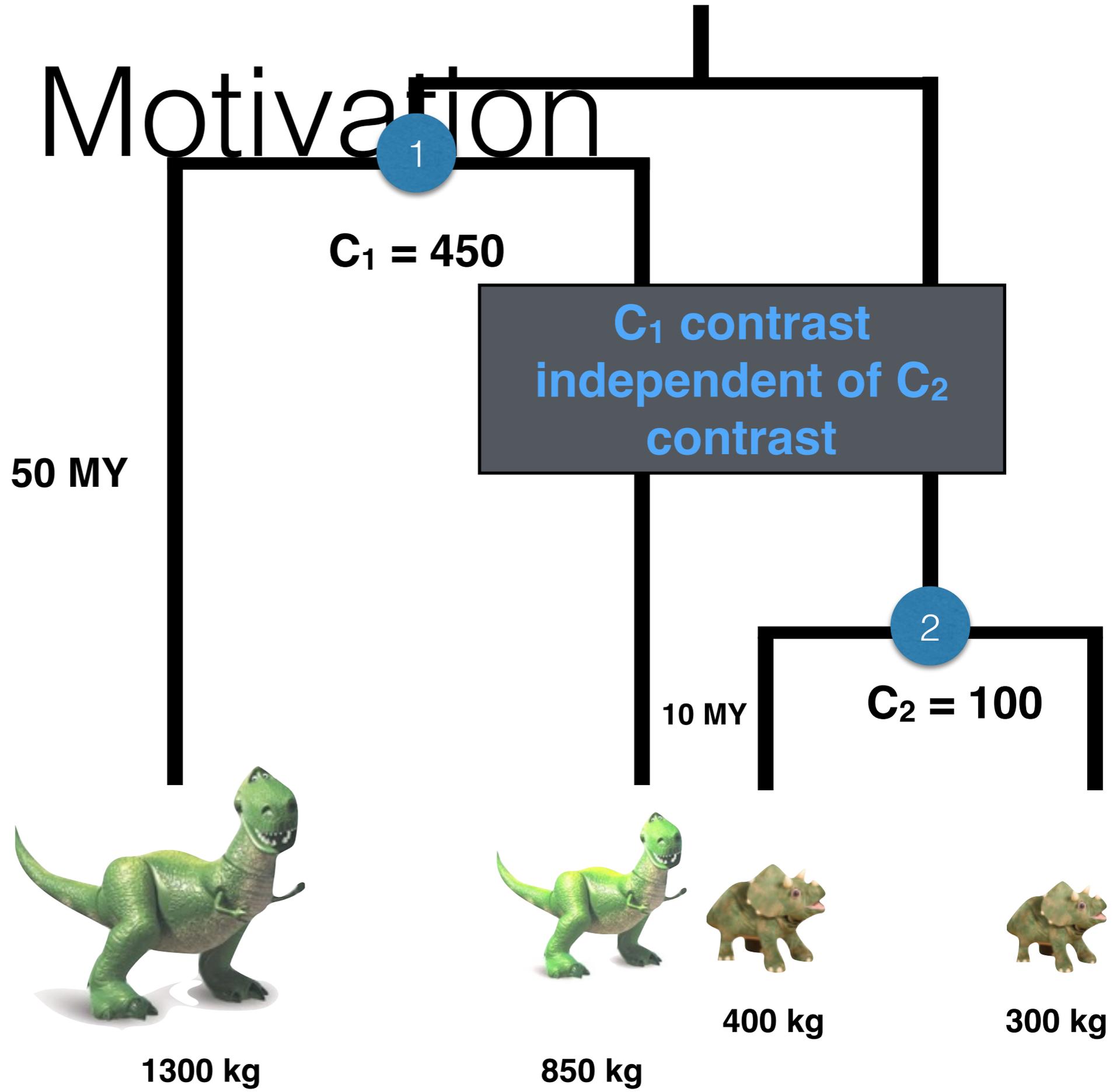


C_3 is contrast in values between (unobserved) species value of node 1 and (unobserved) species value of node 2

unobserved values can be estimated from their respective daughter species

Motivation

- a contrast tells you how much difference has evolved
- each contrast provides an *independent* measures of evolutionary rate
- contrast in node 1 depends *only* on evolution in node 1 daughter species



Phylogenetic non-independence

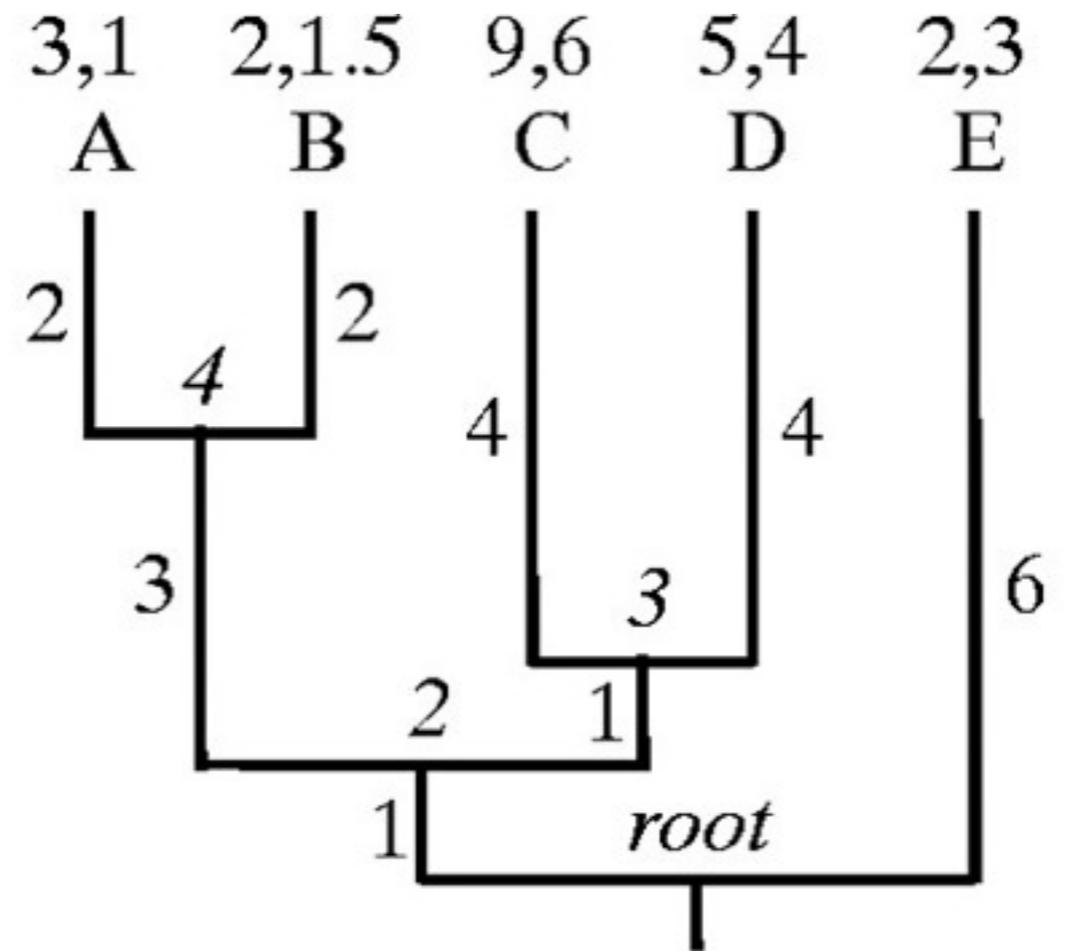
- any set of species in a comparative analysis will share an evolutionary history
- we expect closely related species to be more similar to each other than more distantly related species
- treating species as independent in statistical analyses can be misleading

the contrasts algorithm: overview

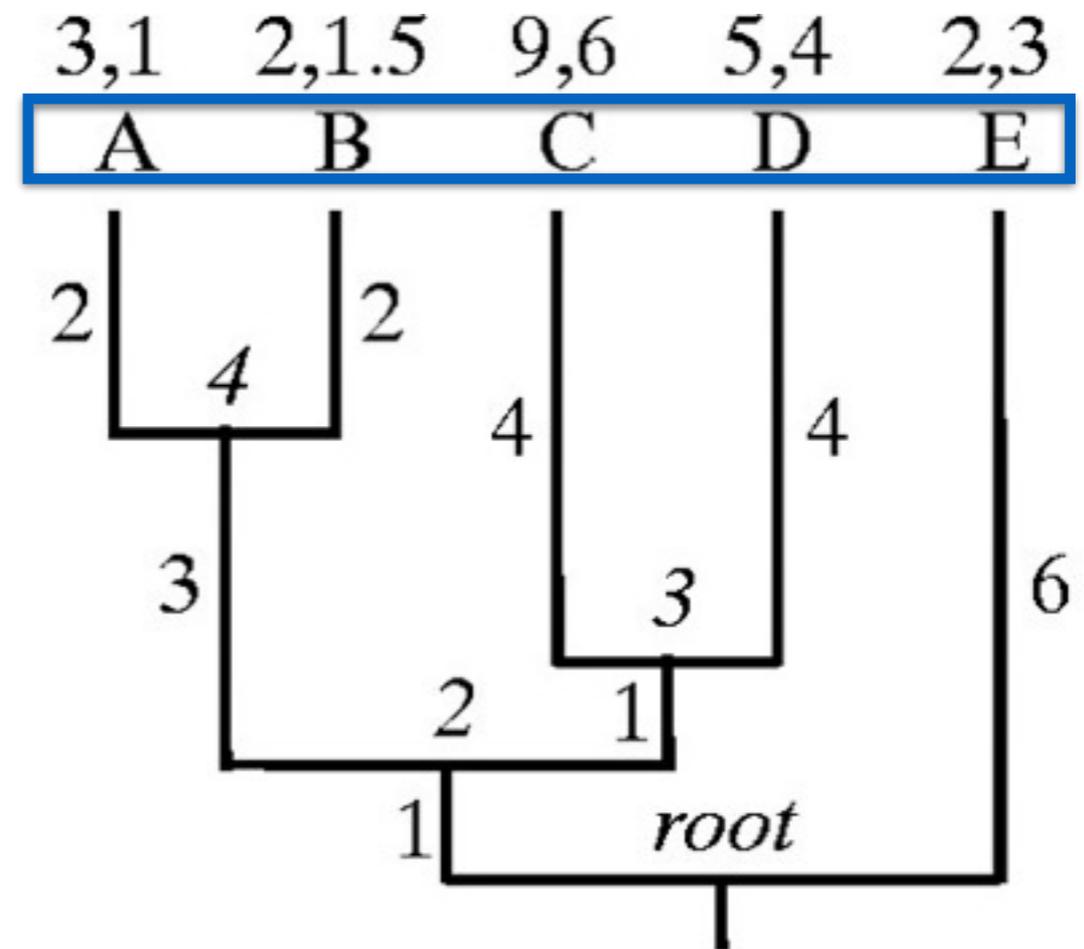
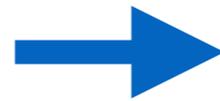
For each internal node....

1. Calculate Raw Contrast
2. Calculate Standardized Contrast
3. Get Expected Trait Value at that node
4. Get Length Correction for Subtending Branch

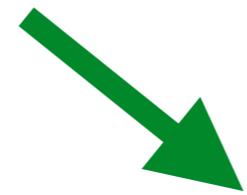
Worked Example



5 species (A-E)



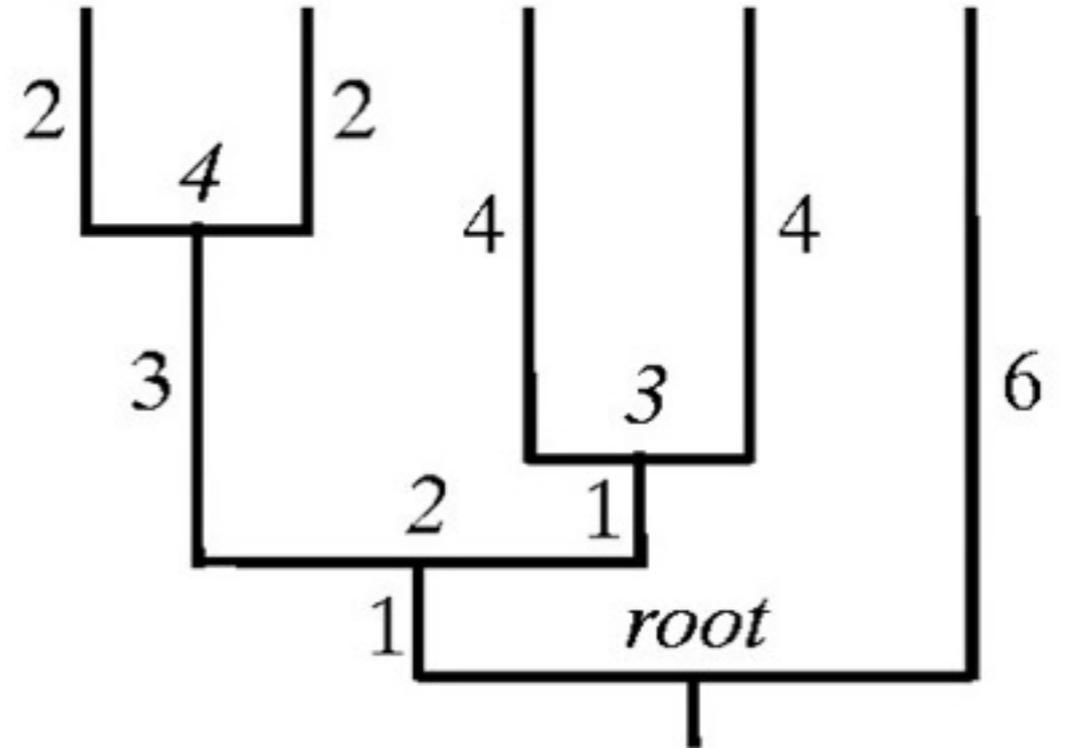
trait values (trait 1, trait 2)



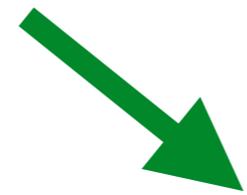
5 species (A-E)



3,1	2,1.5	9,6	5,4	2,3
A	B	C	D	E



trait values (trait 1, trait 2)

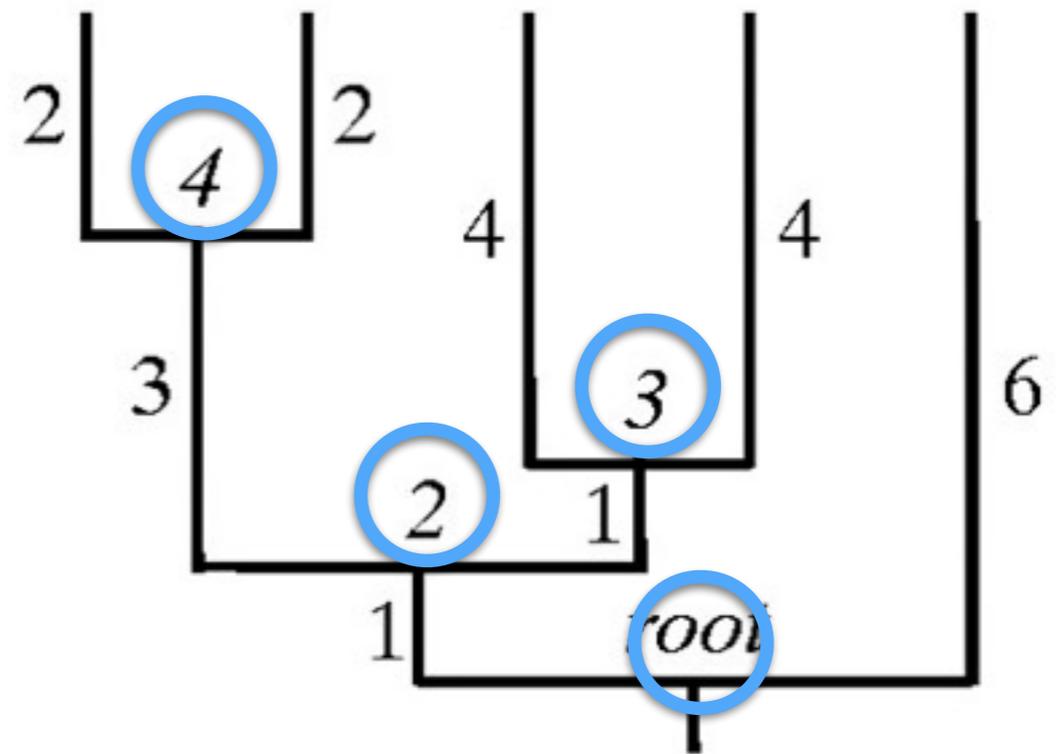


5 species (A-E)



3.1	2.1.5	9.6	5.4	2.3
A	B	C	D	E

4 internal nodes



step 1 get raw contrast

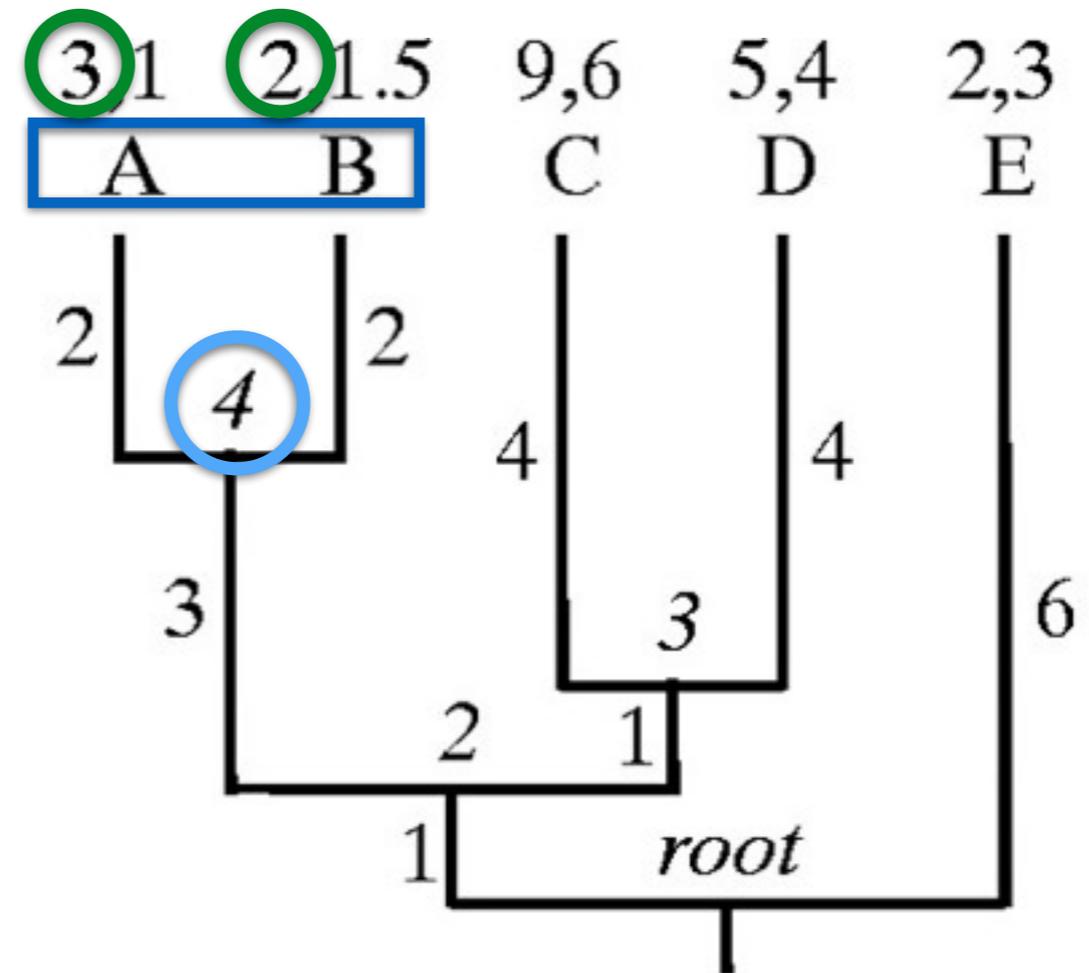
trait values

species

internal nodes

branch lengths

trait 1 raw contrast of node 4: $3 - 2 = 1$



step 2

standardize contrast

$$\text{contrast}_S = \frac{\text{contrast}_R}{\sqrt{\sum \text{branchlengths}}}$$

$$1 / (2 + 2)^{0.5} = 0.5$$

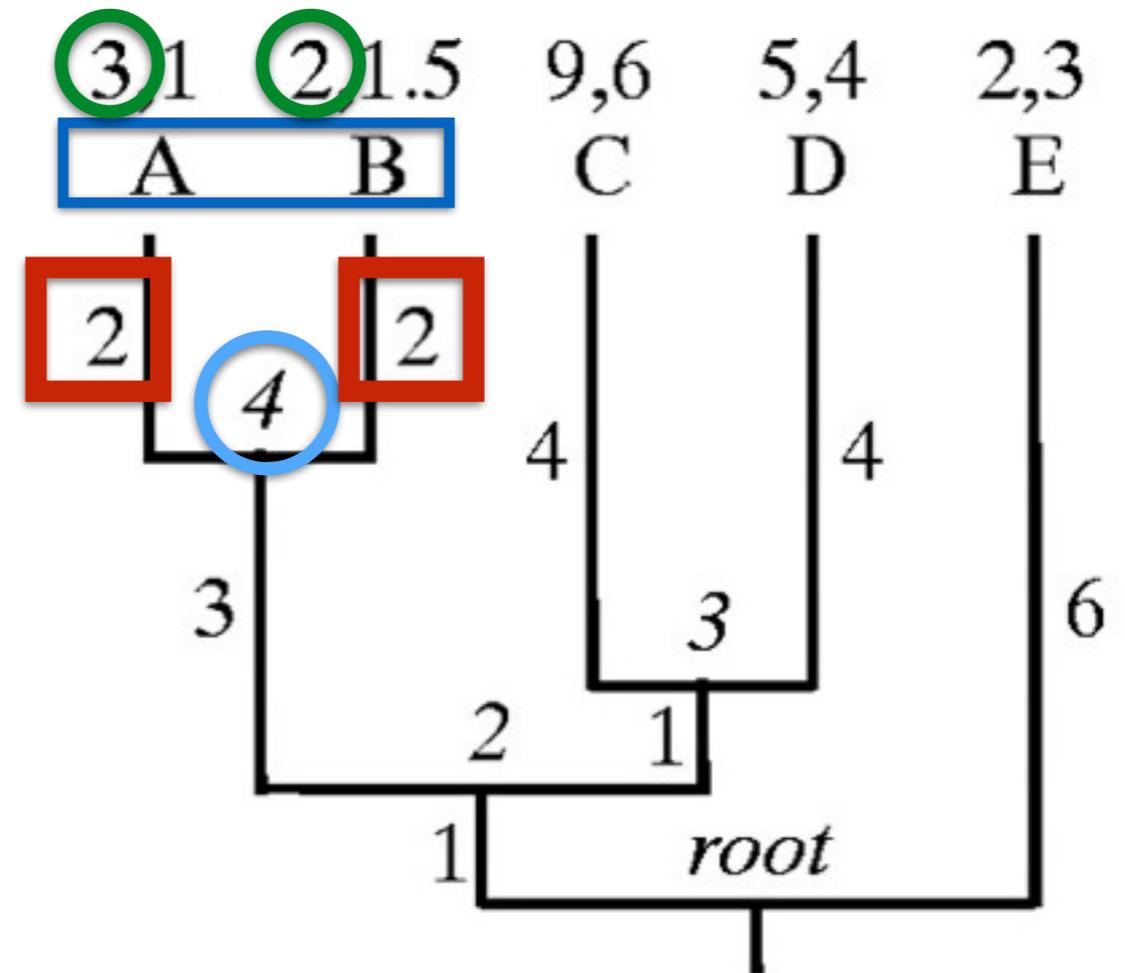
we standardize by amount of branch length since common ancestor to make contrasts comparable to other contrasts

trait values

species

internal nodes

branch lengths



step 2

standardize contrast

$$\text{contrast}_S = \frac{\text{contrast}_R}{\sqrt{\sum \text{branchlengths}}}$$

get the standardized contrast for node 3

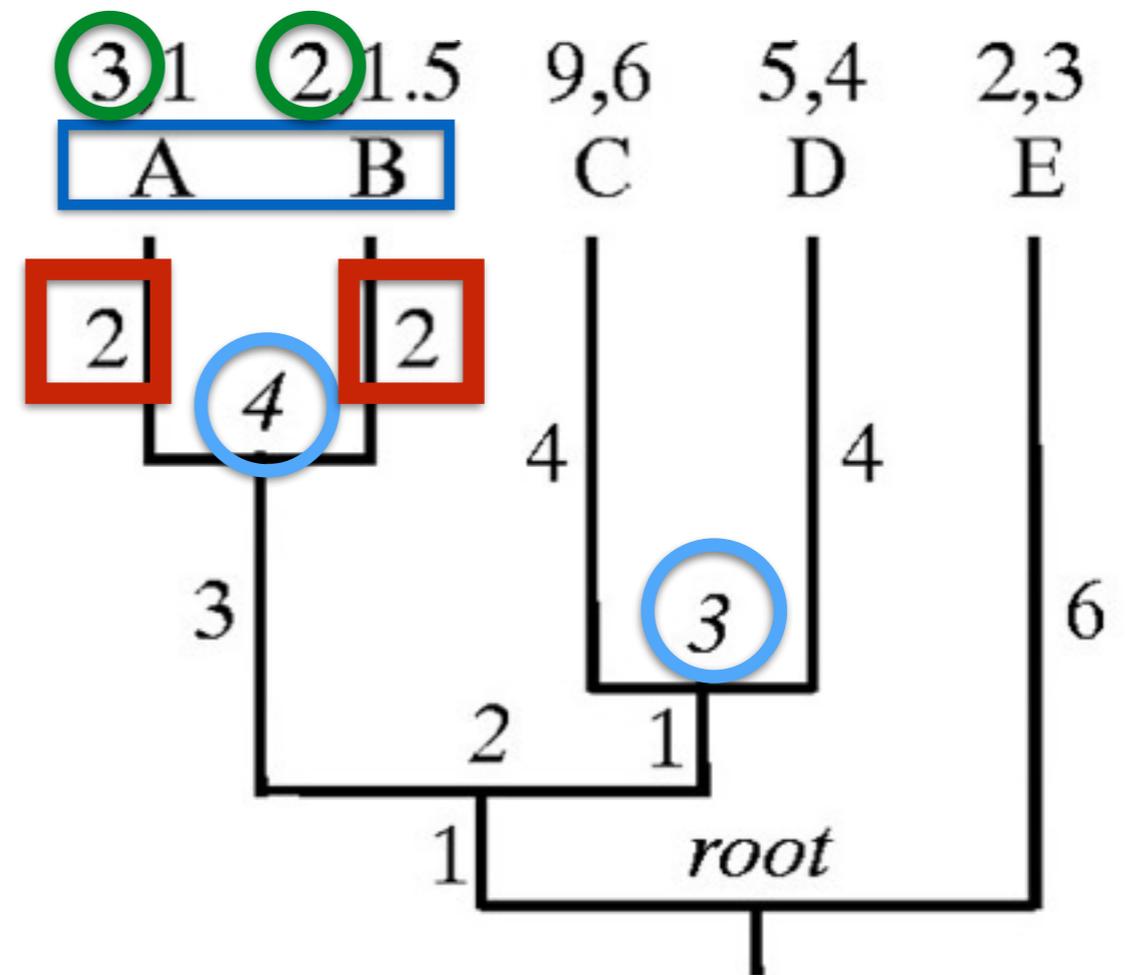
we standardize by amount of branch length since common ancestor to make contrasts comparable to other contrasts

trait values

species

internal nodes

branch lengths



step 2

standardize contrast

$$\text{contrast}_S = \frac{\text{contrast}_R}{\sqrt{\sum \text{branchlengths}}}$$

get the standardized contrast for node 2

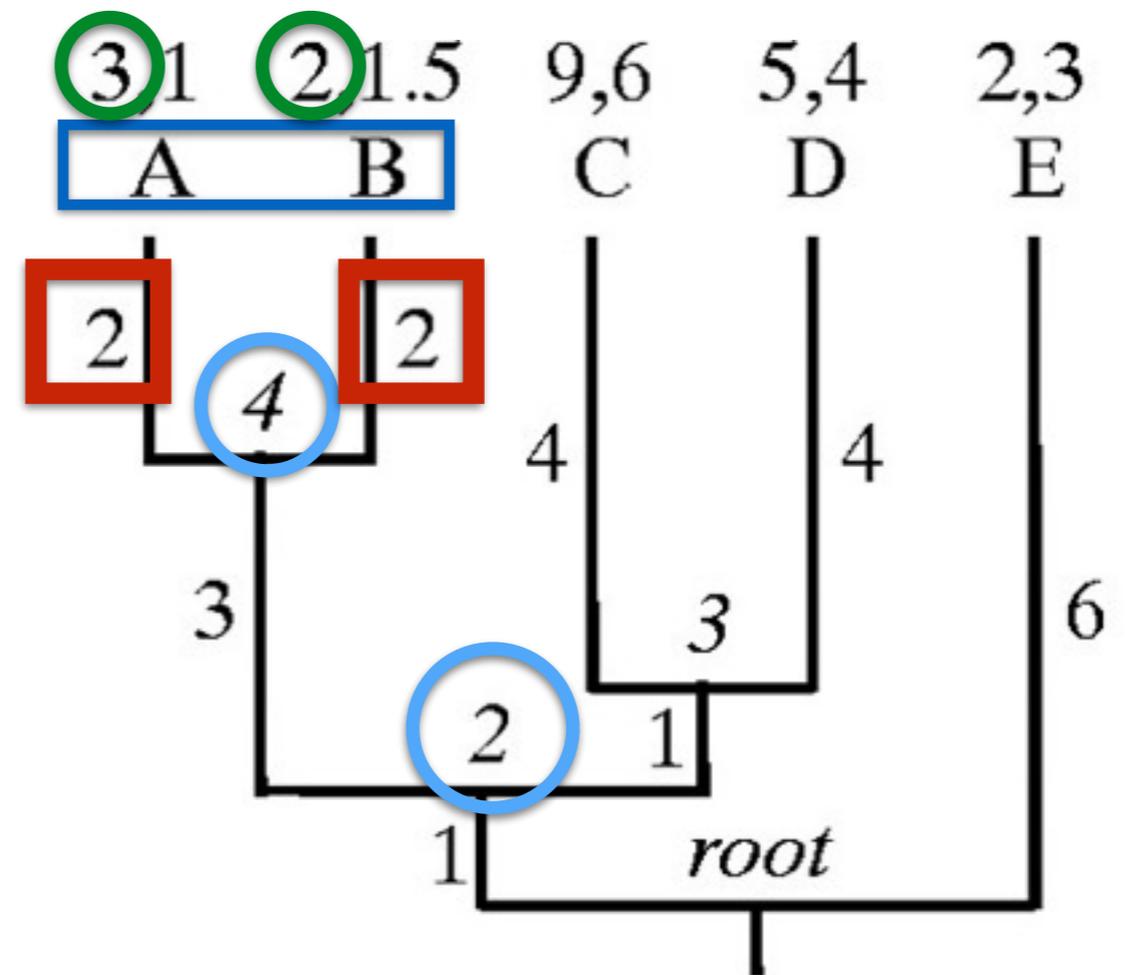
we standardize by amount of branch length since common ancestor to make contrasts comparable to other contrasts

trait values

species

internal nodes

branch lengths



step 2

standardize contrast

$$\text{contrast}_S = \frac{\text{contrast}_R}{\sqrt{\sum \text{branchlengths}}}$$

get the standardized contrast for node 2

(we need to know the trait values for internal nodes)

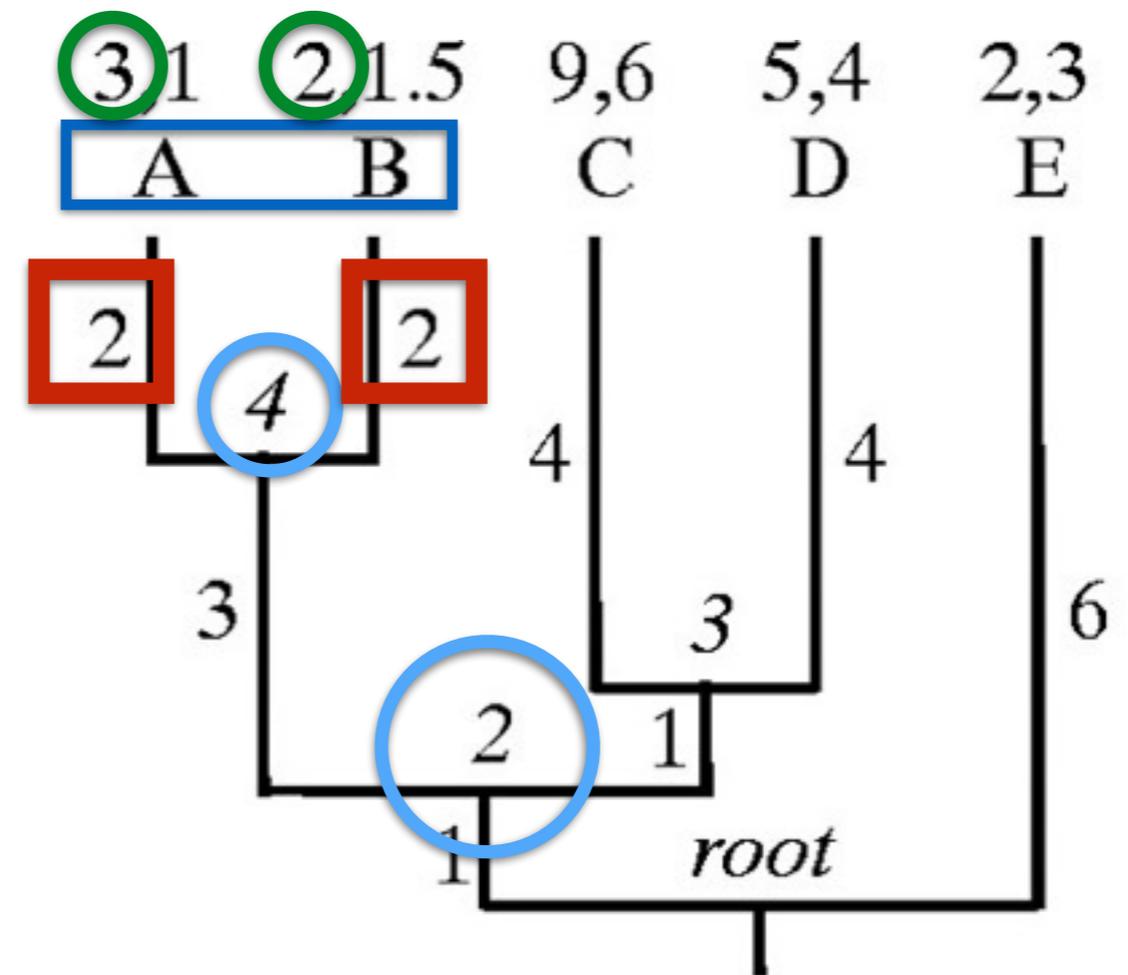
we standardize by amount of branch length since common ancestor to make contrasts comparable to other contrasts

trait values

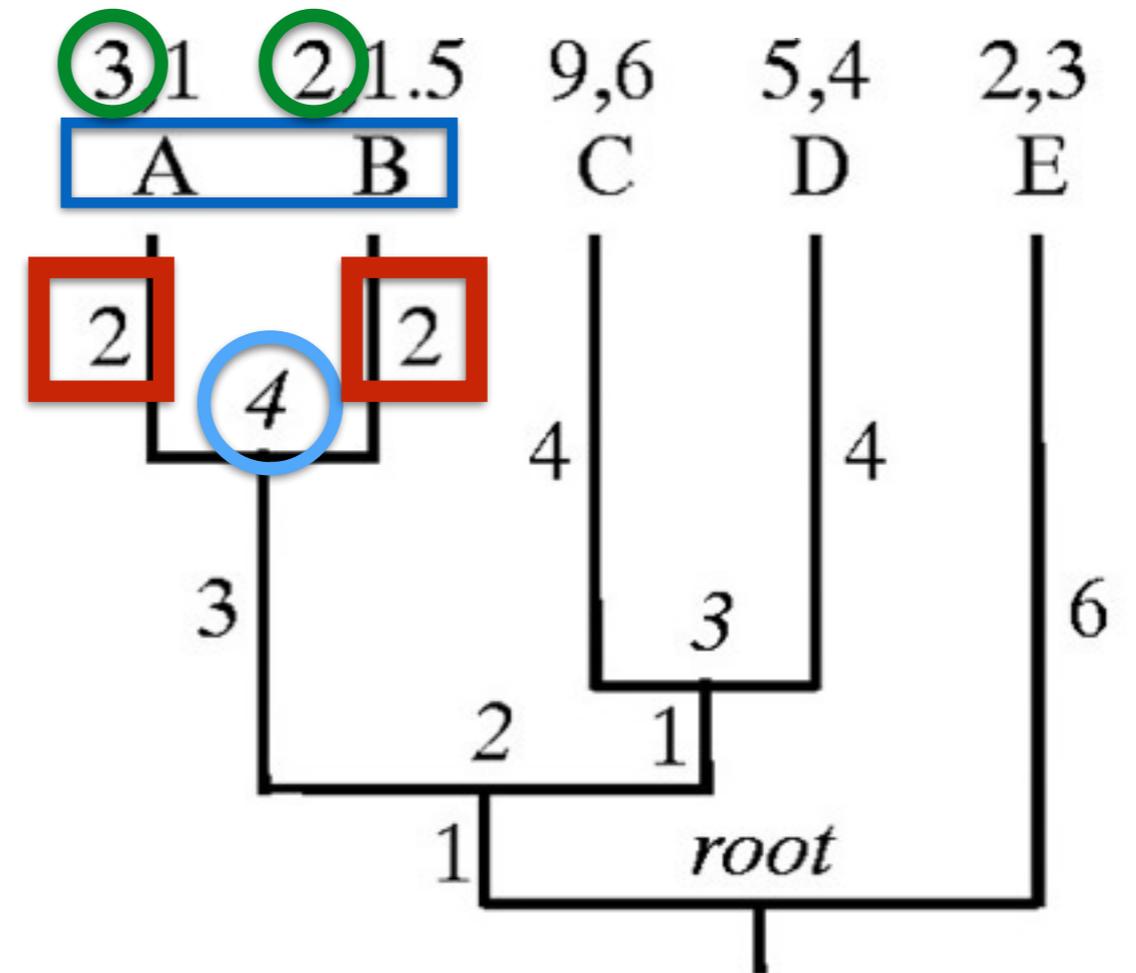
species

internal nodes

branch lengths



Step 3: get internal node value of trait

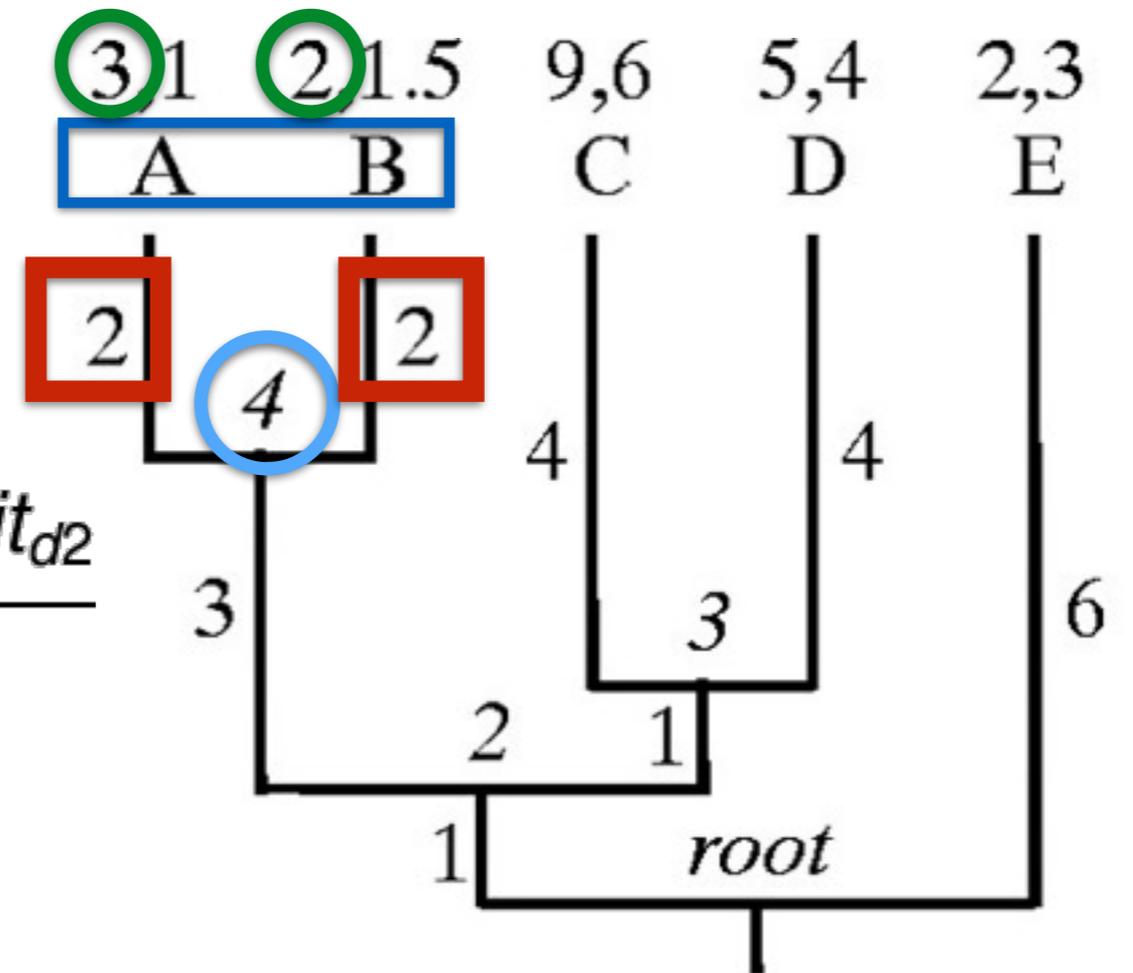


*trait value of the internal node is the average of the trait values in the daughter species **weighted** by branch length*

Step 3: get internal node value of trait

$$trait_{internal} = \frac{\frac{1}{branch_1} * trait_{d1} + \frac{1}{branch_2} * trait_{d2}}{\frac{1}{branch_1} + \frac{1}{branch_2}}$$

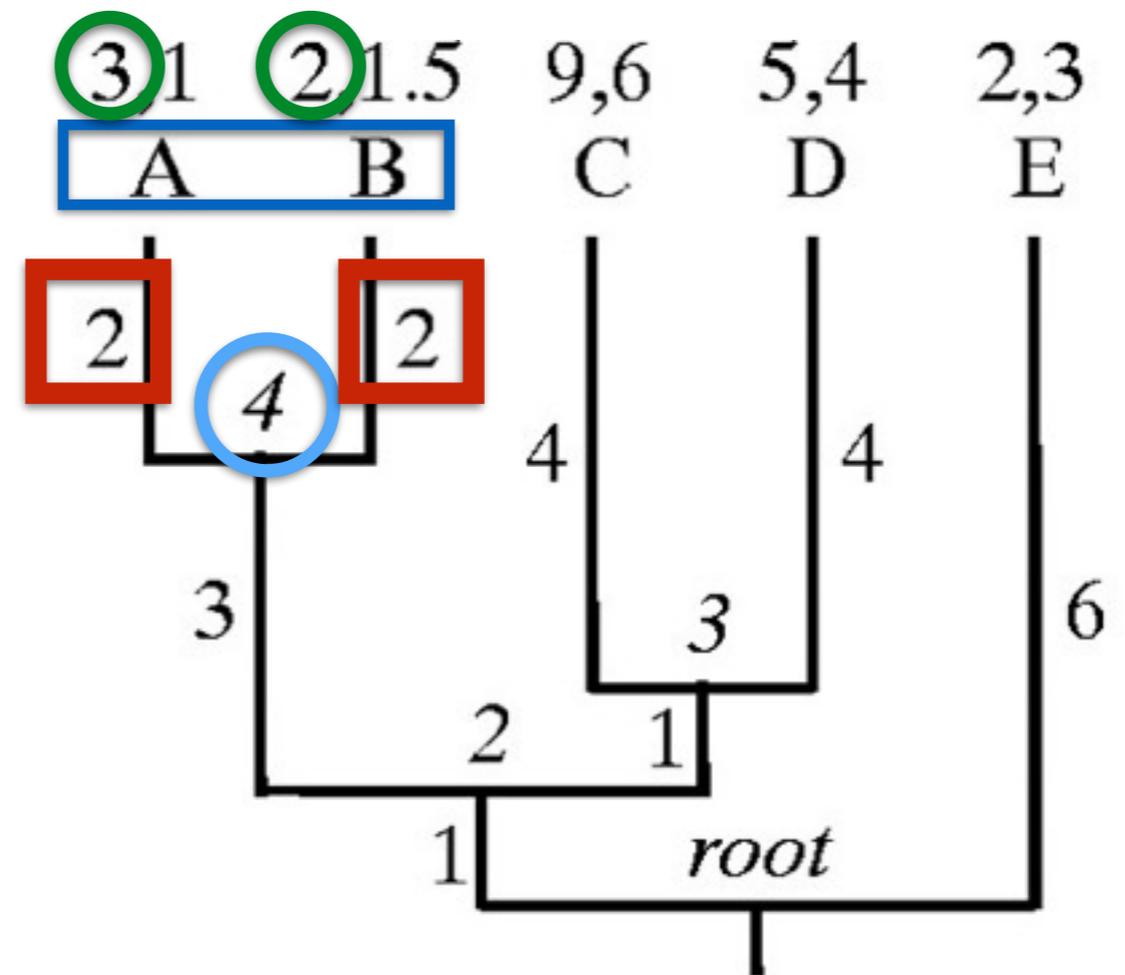
*trait value of the internal node is the average of the trait values in the daughter species **weighted** by branch length*



Step 3: get internal node value of trait

$$\begin{aligned} \text{trait}_{\text{internal}} &= \frac{\frac{1}{\text{branch}_1} * \text{trait}_{d1} + \frac{1}{\text{branch}_2} * \text{trait}_{d2}}{\frac{1}{\text{branch}_1} + \frac{1}{\text{branch}_2}} \\ &= \\ &= \mathbf{(3/2 + 2/2)/(1/2 + 1/2)} \\ &= \\ &= \mathbf{2.5} \end{aligned}$$

*trait value of the internal node is the average of the trait values in the daughter species **weighted** by branch length*

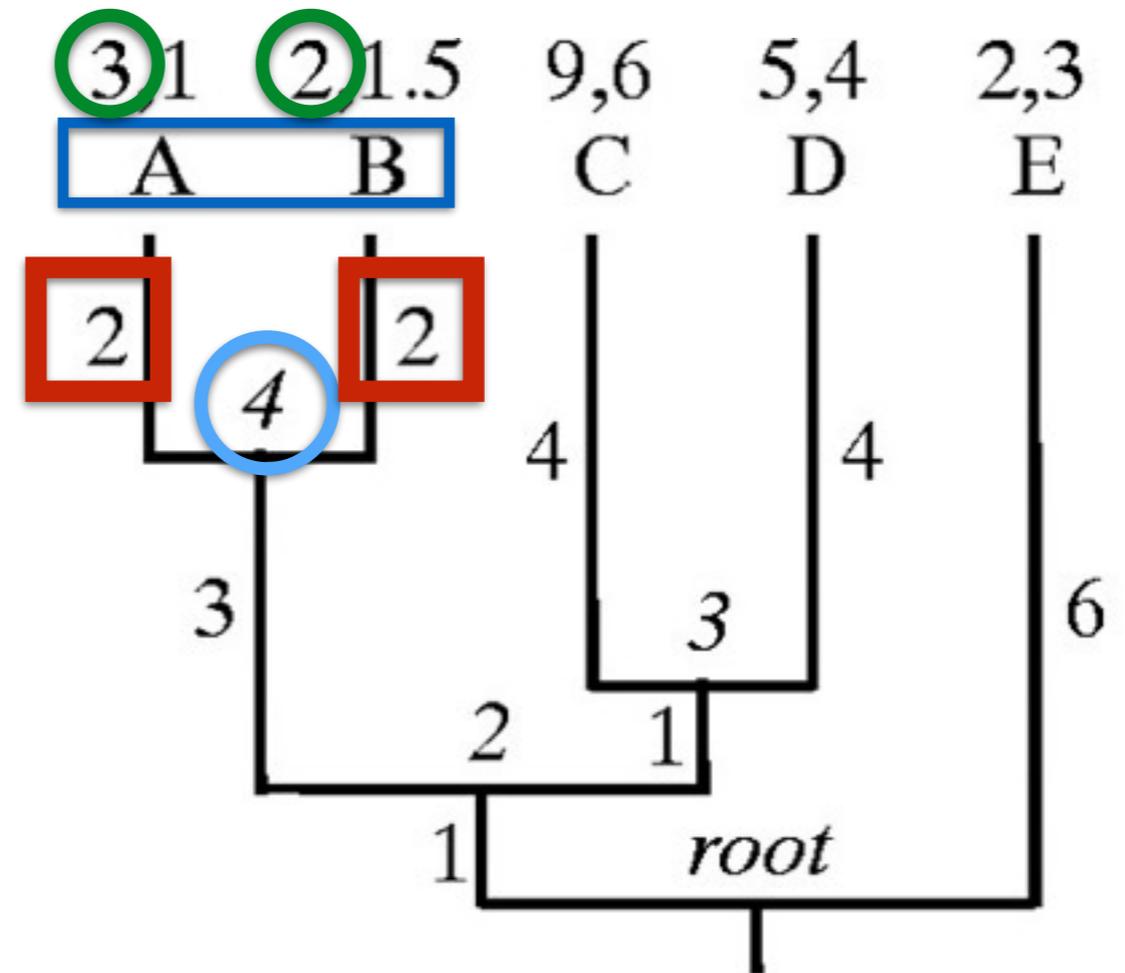


Step 3: get internal node value of trait

$$trait_{internal} = \frac{\frac{1}{branch_1} * trait_{d1} + \frac{1}{branch_2} * trait_{d2}}{\frac{1}{branch_1} + \frac{1}{branch_2}}$$

get internal value of node 3

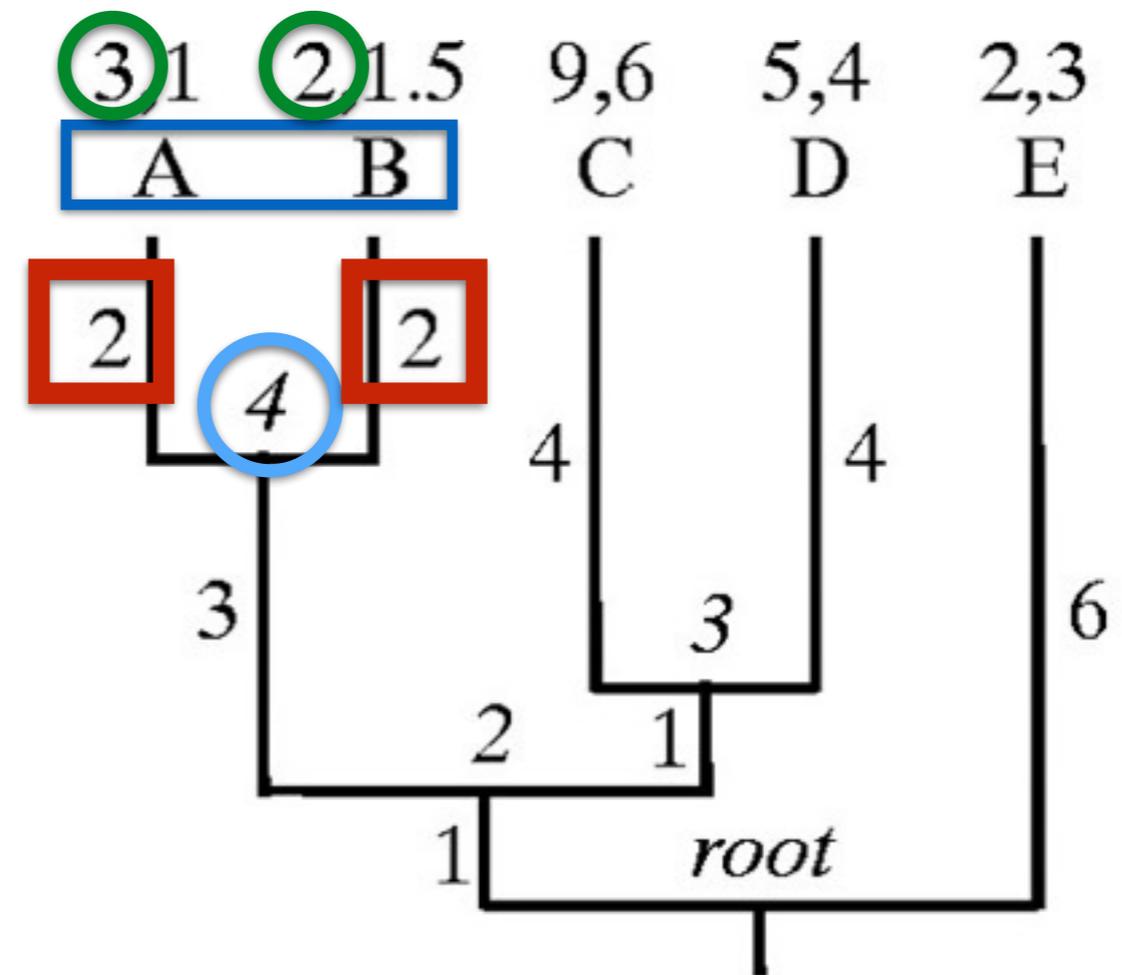
*trait value of the internal node is the average of the trait values in the daughter species **weighted** by branch length*



Step 4: calculate error of branch length leading to interior contrast

$$\text{brancherror} = \frac{\text{branch}_1 * \text{branch}_2}{\text{branch}_1 + \text{branch}_2}$$

error in trait estimation of internal nodes can be expressed as a burst of evolution at that node

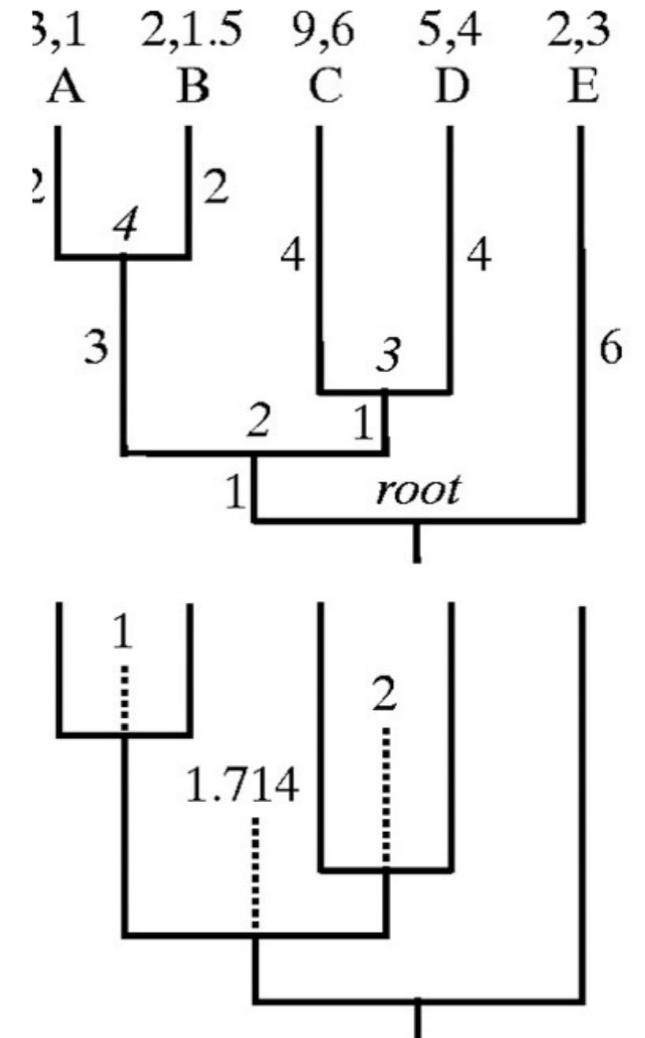


Step 4: calculate error of branch length leading to interior contrast

$$\begin{aligned} \text{brancherror} &= \frac{\text{branch}_1 * \text{branch}_2}{\text{branch}_1 + \text{branch}_2} \\ &= \frac{(2 * 2) / 4}{(2 + 2)} \\ &= 1 \end{aligned}$$

error in trait estimation of internal nodes can be expressed as a burst of evolution at that node

this new length is used for all subsequent calculations using that branch



Exercise: work in pairs, calculate all standardized contrasts for traits 1 and 2

step 2 (standardize)

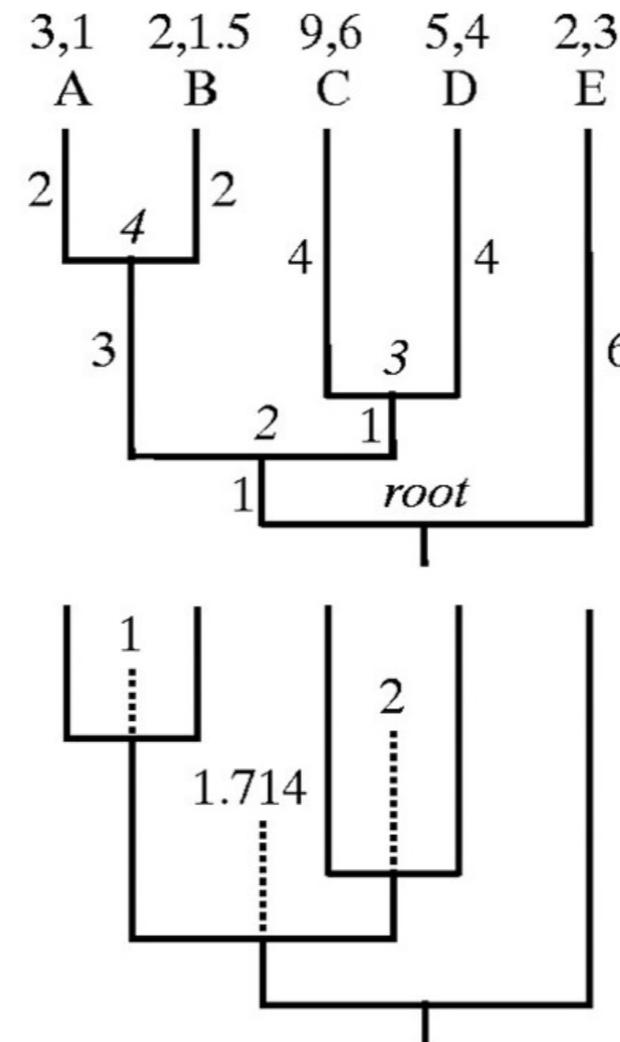
$$\text{contrast}_S = \frac{\text{contrast}_R}{\sqrt{\sum \text{branchlengths}}}$$

step 3 (internal value)

$$\text{trait}_{\text{internal}} = \frac{\frac{1}{\text{branch}_1} * \text{trait}_{d1} + \frac{1}{\text{branch}_2} * \text{trait}_{d2}}{\frac{1}{\text{branch}_1} + \frac{1}{\text{branch}_2}}$$

step 4 (error)

$$\text{brancherror} = \frac{\text{branch}_1 * \text{branch}_2}{\text{branch}_1 + \text{branch}_2}$$

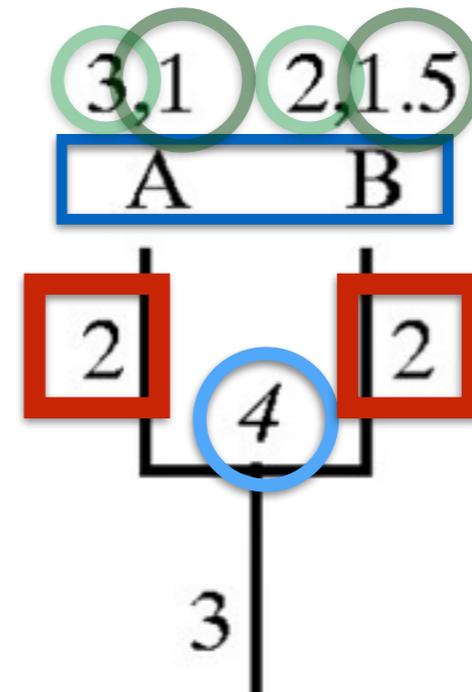


Phylogenetically independent contrasts calculations for trait 1

Node name	Contrast name	Contrast values	Raw contrast	Standard deviation	Standardized contrast	(Corrected) length of each branch	
						Left	Right
4	A-B	3-2	1.0	2.0	0.5	2.0	2.0
3	C-D	9-5	4.0	2.828	1.414	4.0	4.0
2	4-3	2.5-7	-4.5	2.646	-1.701	[4.0]	[3.0]
root	2-E	5.071-2	3.071	2.952	1.040	[2.714]	6.0

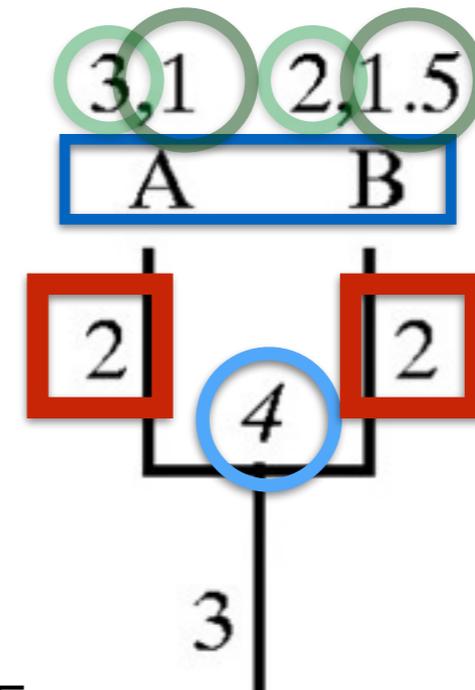
covariation between original traits and trait contrasts does not change

- regression analysis on contrasts allows for tests of correlated evolution
- to preserve covariation, contrasts on traits must be calculated in same order



covariation between original traits and trait contrasts does not change

- regression analysis on contrasts allows for tests of correlated evolution
- to preserve covariation, contrasts on traits must be calculated in same order

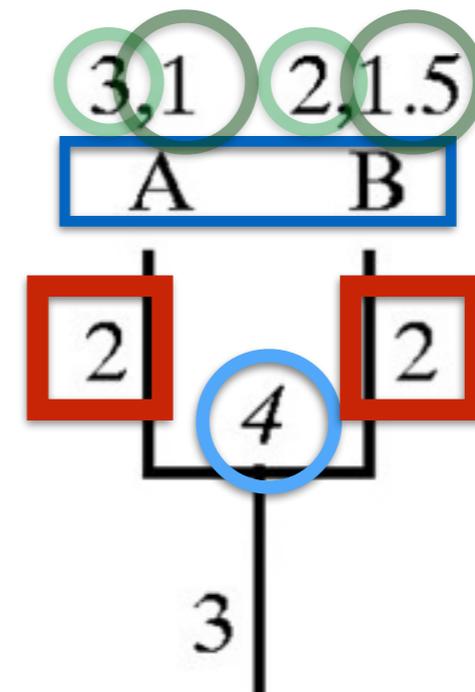


$$\begin{aligned}C1 &= 3 - 2 = 1 \\C2 &= 1 - 1.5 = -0.5 \\C1 &= 2 - 3 = -1 \\C2 &= 1.5 - 1 = 0.5\end{aligned}$$

yes

covariation between original traits and trait contrasts does not change

- regression analysis on contrasts allows for tests of correlated evolution
- to preserve covariation, contrasts on traits must be calculated in same order



$$C1 = 3 - 2 = 1$$
$$C2 = 1 - 1.5 = -0.5$$

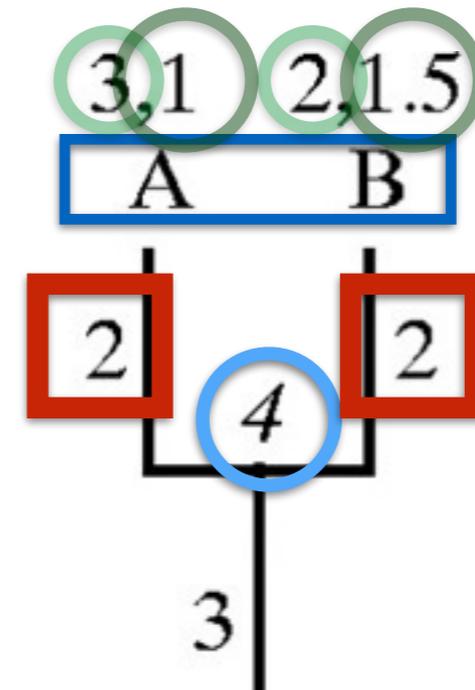
$$C1 = 2 - 3 = -1$$
$$C2 = 1.5 - 1 = 0.5$$

$$C1 = 3 - 2 = 1$$
$$C2 = 1.5 - 1 = 0.5$$

yes

covariation between original traits and trait contrasts does not change

- regression analysis on contrasts allows for tests of correlated evolution
- to preserve covariation, contrasts on traits must be calculated in same order



$$C1 = 3 - 2 = 1$$
$$C2 = 1 - 1.5 = -0.5$$

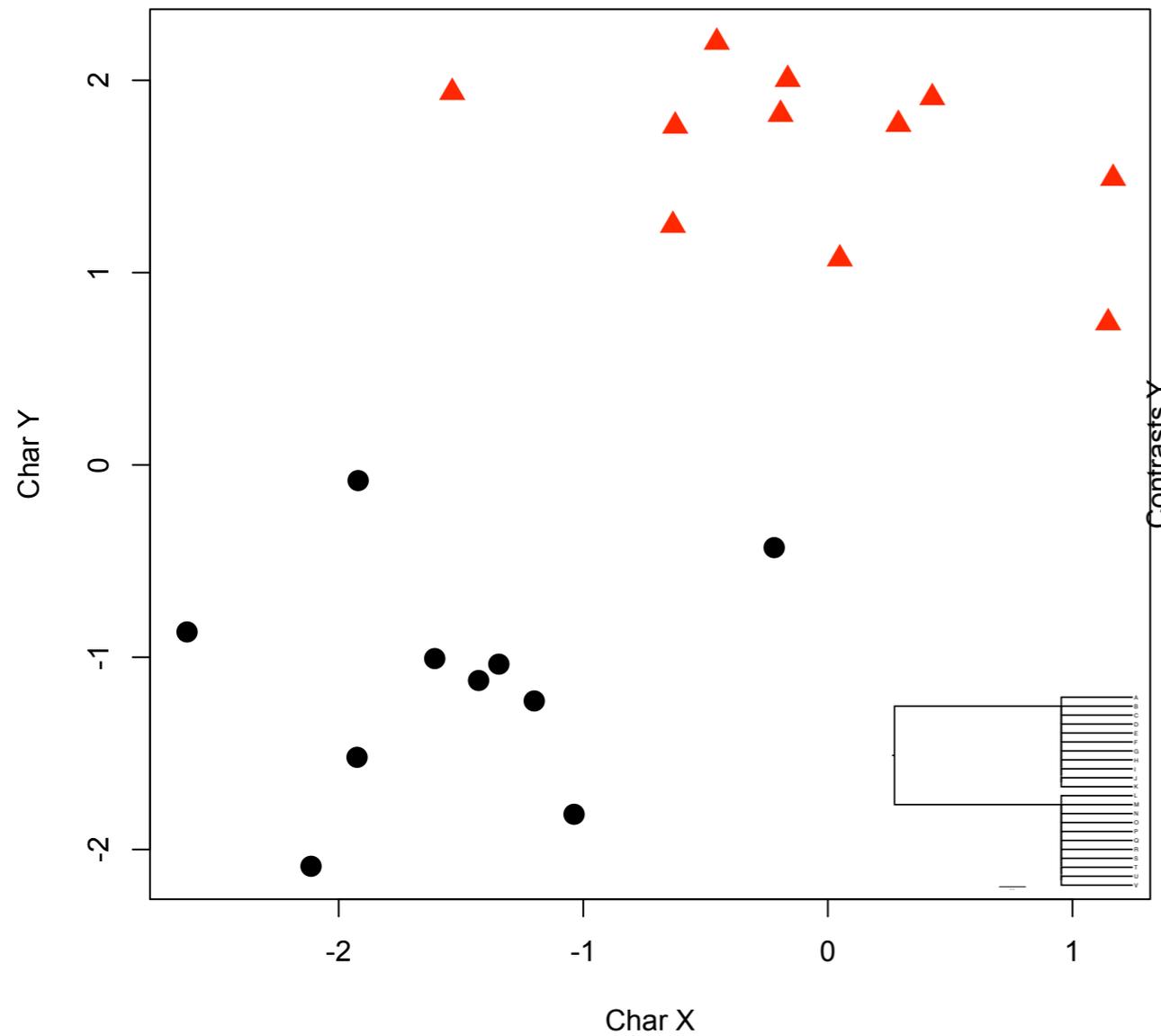
$$C1 = 2 - 3 = -1$$
$$C2 = 1.5 - 1 = 0.5$$

yes

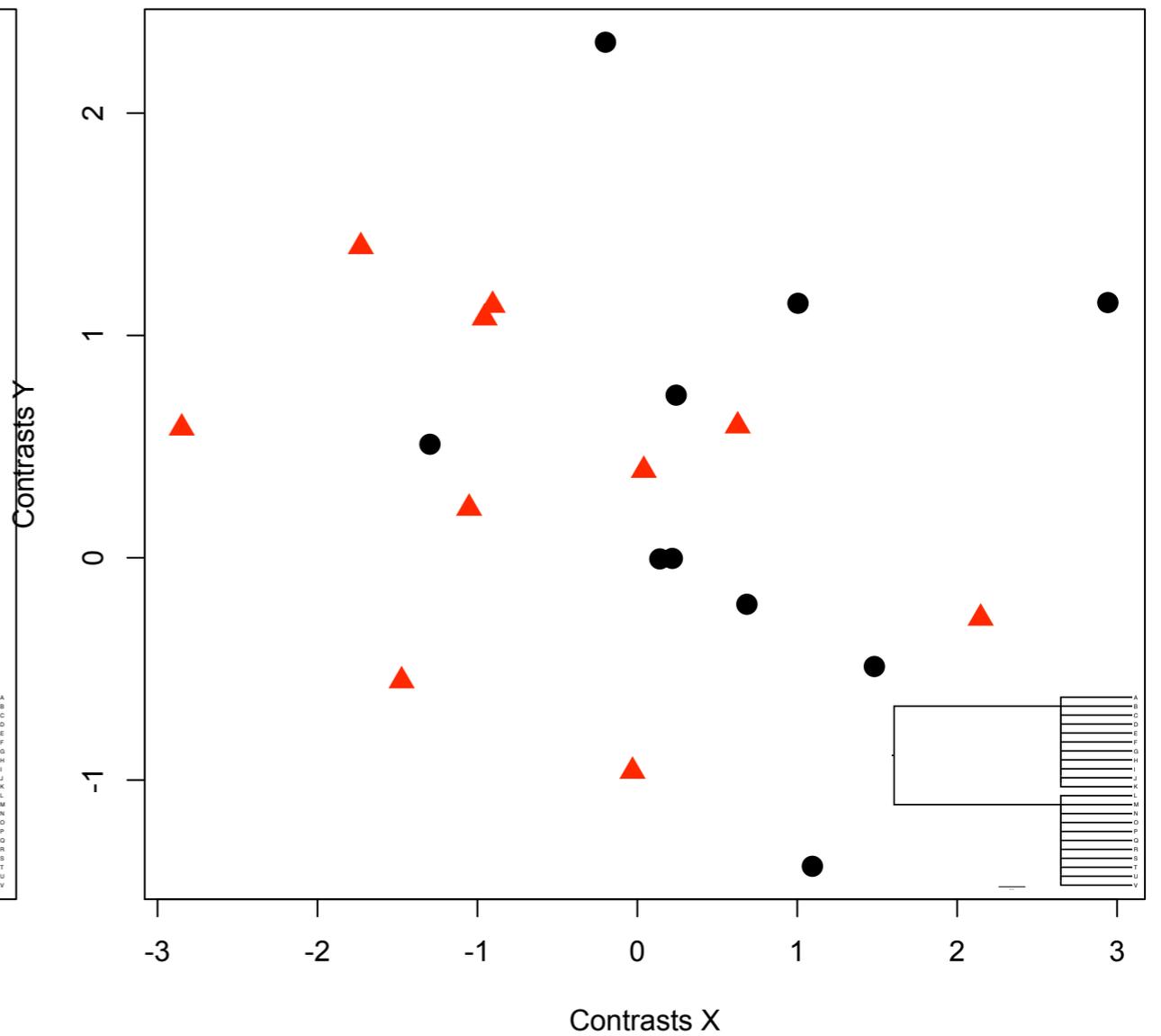
~~$$C1 = 3 - 2 = 1$$
$$C2 = 1.5 - 1 = 0.5$$~~

no

contrasts can then be used in multivariate analyses



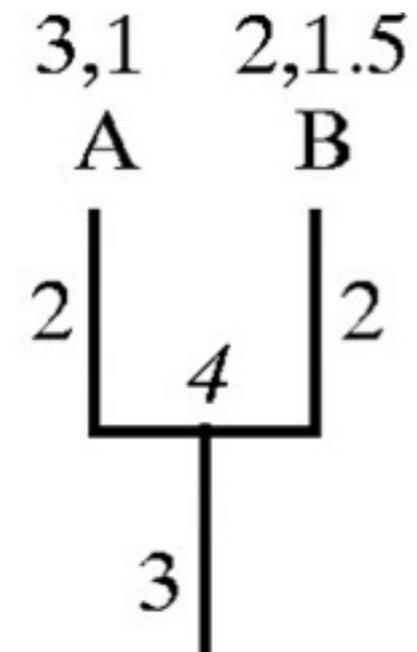
RAW data



contrast data

one last constraint on contrasts....

- regression models must be forced through origin
- biological reason: at node, expected divergence between traits is 0
- failure to force regression through origin means that at node, you expect some divergence in trait Y when divergence in trait X is 0



What are contrasts?

- Each standardized contrast is telling us something about the RATE of evolution

$$S_{ij} = \frac{X_i - X_j}{v_i + v_j} \quad \begin{array}{l} \text{“amount of change”} \\ \text{“time”} \end{array}$$

What are contrasts?

- The contrasts have a close relationship with σ^2 , the rate parameter from BM
- The sum of the squared contrasts divided by n gives the ML estimate of σ^2

Using and Interpreting Contrasts

- Independent contrasts should be thought of as vectors
- They summarize information about the **amount and direction** of evolution at each node in the tree
- Standardized contrasts provide information about the **rate** of evolution

Using and Interpreting Contrasts

$$\hat{\sigma}^2 = \frac{\sum S_i}{n-1}$$

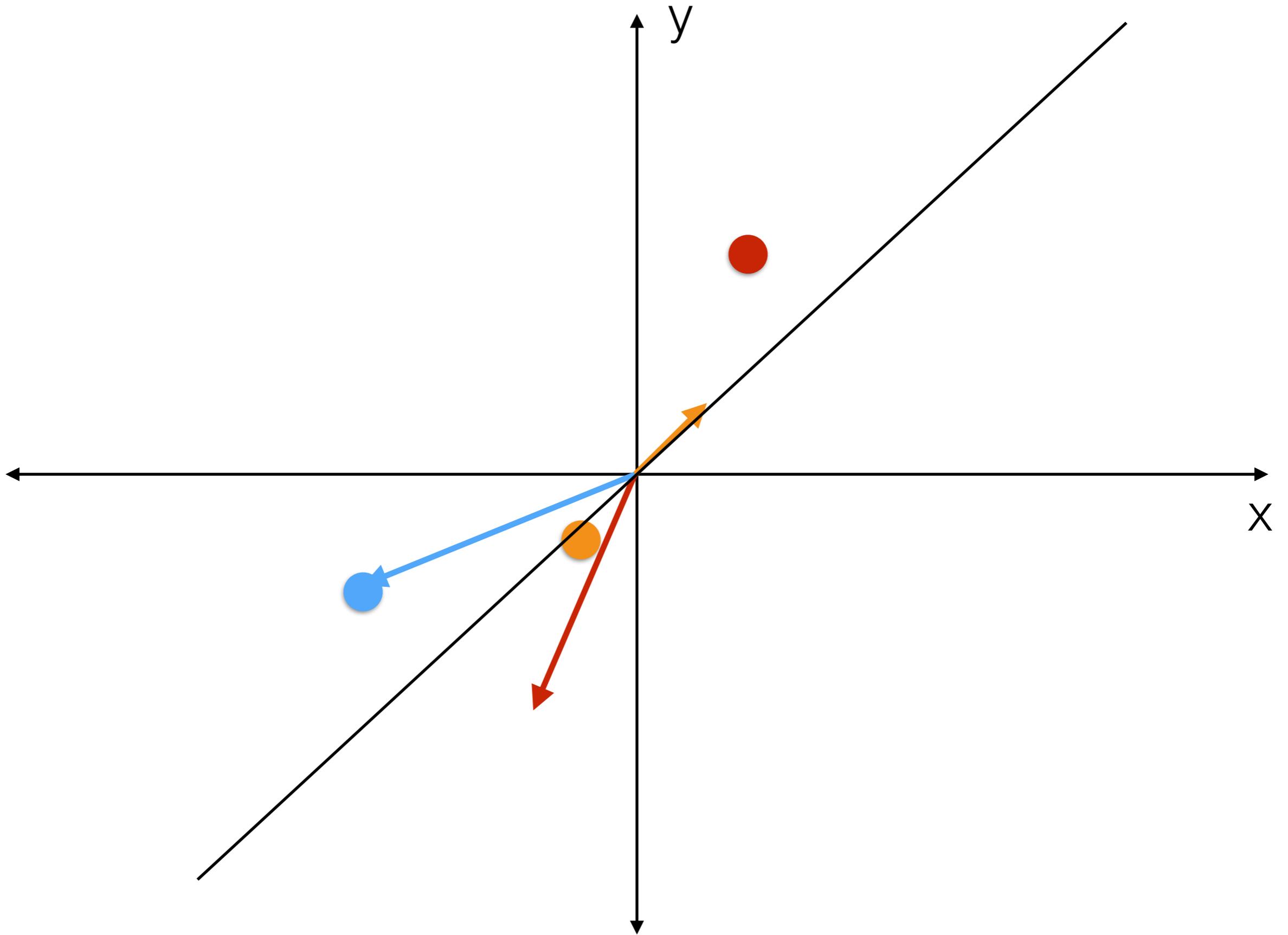
- This provides an unbiased estimate of evolutionary rate
- The expected value of this estimate is equal to the actual rate parameter
- The maximum likelihood estimate of the rate parameter is biased

Character correlations

- Most common use for ICs: testing for character correlations
- Are two characters evolving in a correlated fashion?

Character correlations

- Calculate independent contrasts for two characters, x and y
- Carry out a regression analysis of y on x with **no intercept** (force regression line through the origin)
- $P < 0.05$, then reject the null hypothesis of no evolutionary correlation



Character correlations

- Why force contrasts through the origin?
- Because, for each contrast, the direction of subtraction is arbitrary; the signs of all the contrasts could be reversed
- Regression through the origin treats the contrasts as vectors

Wrasse Feeding Diversity

Crustacivores



Planktivores



Molluscivores



Piscivores

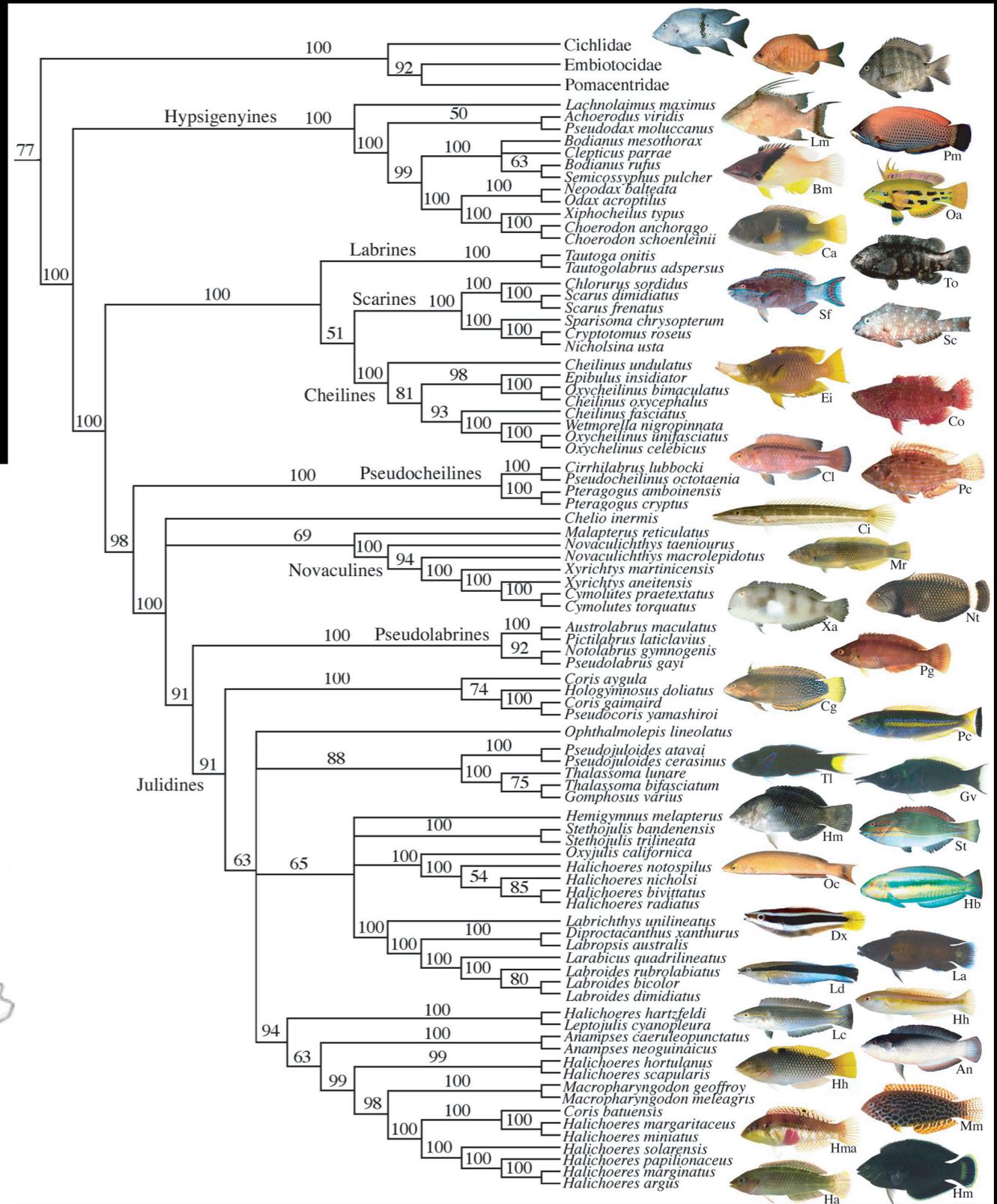
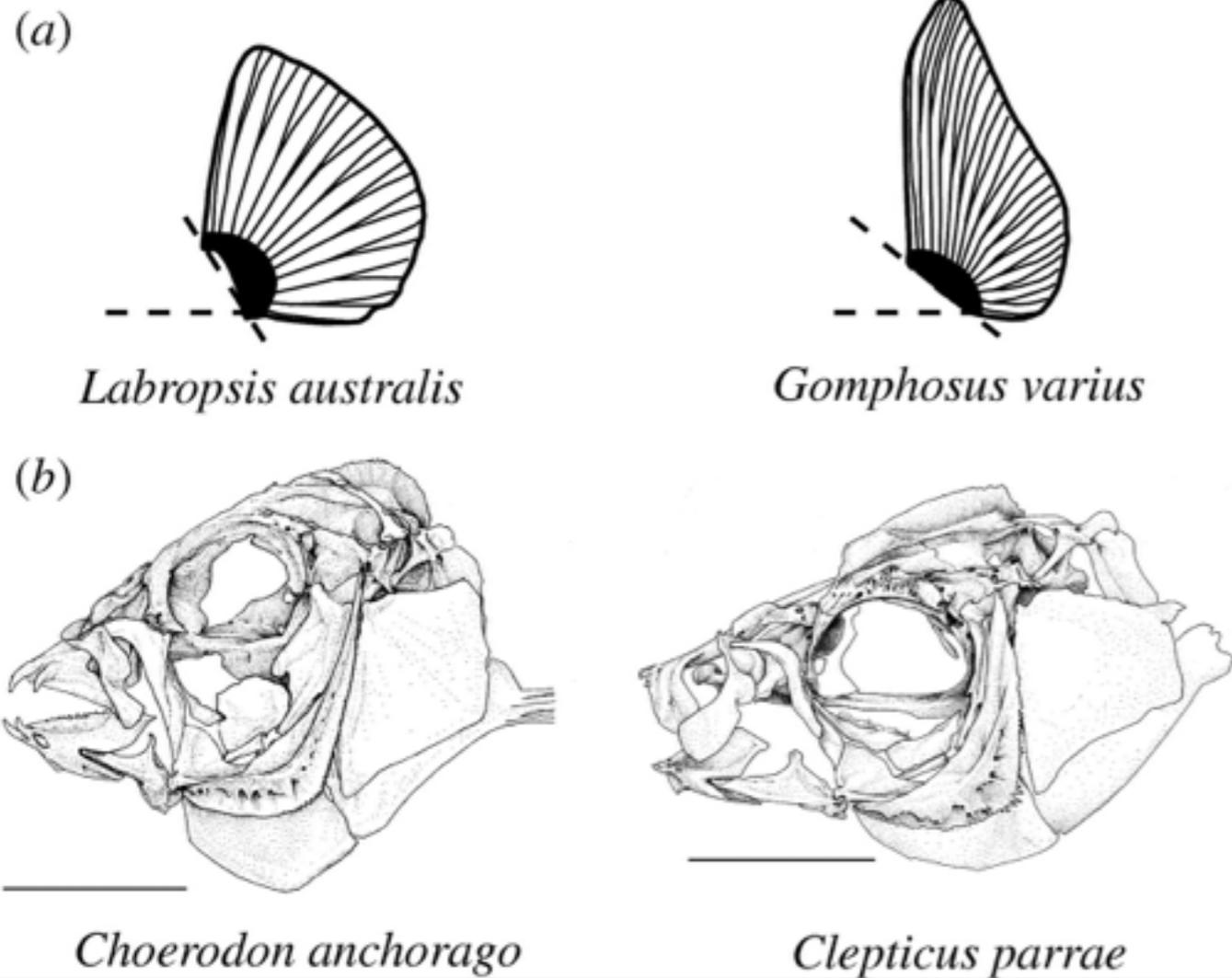


What is the functional basis of diversity?

Herbivores

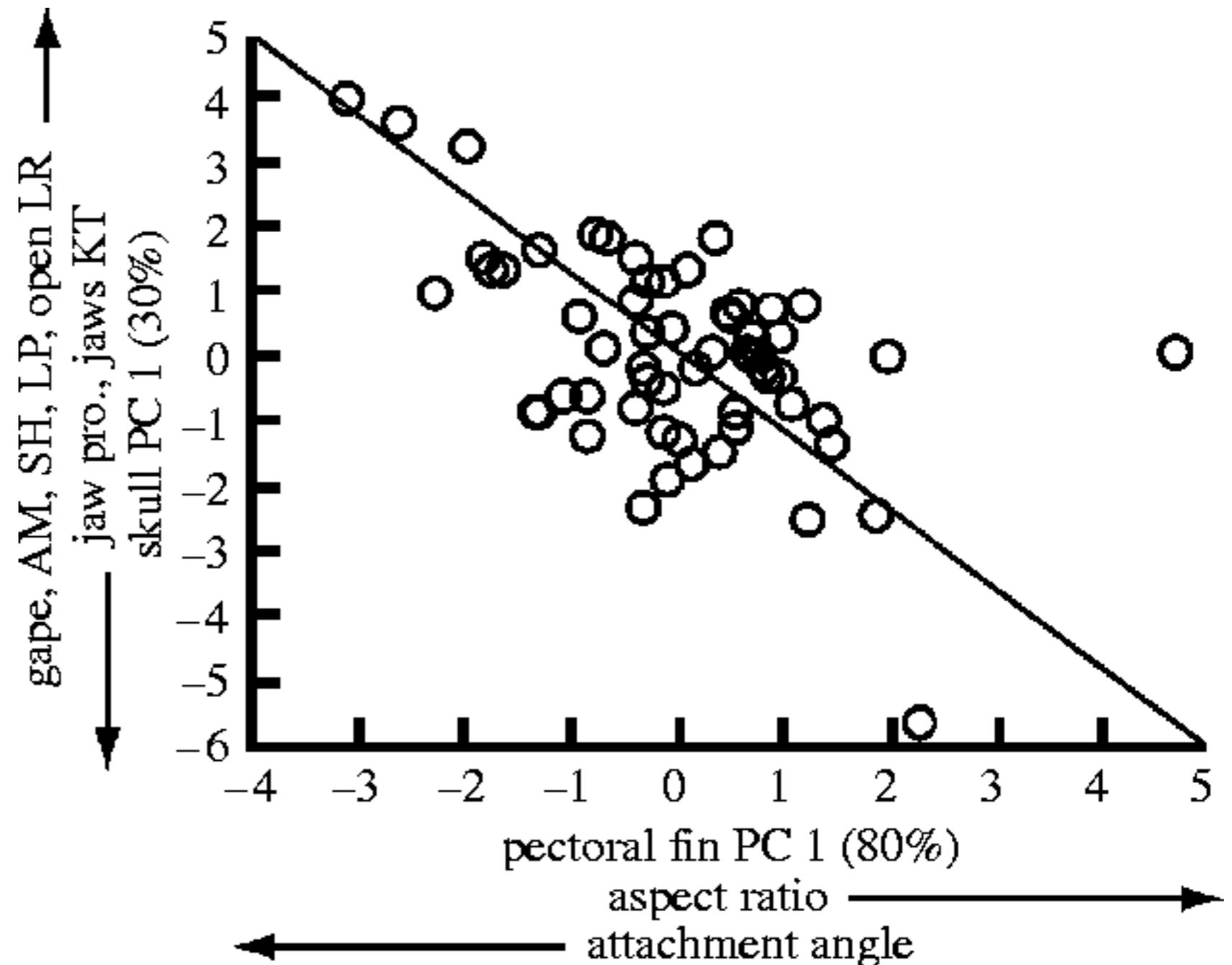


Have skull and pectoral fins evolved together?



Integrated feeding and locomotor evolution in wrasses

- changes in the skull to increase jaw speed are associated with pectoral fin shape changes that increase swimming speed performance;
- increases in the skull's force output are associated with changes in the pectoral fin to produce greater thrust at slow speeds.



do bats in complex habitats have bigger brains?

184 K. Safi and D. K. N. Dechmann *Brain adaptation to habitat complexity in bats*

