Other Models for Continuous Characters

Other Models for Phenotypic Evolution

- Brownian Motion (BM)
- Early Burst (EB)
- Ornstein-Uhlenbeck (OU)

Brownian Motion (BM)

- Brownian motion model with a constant rate of evolution
- Two parameters: starting value (Θ) and rate (σ^2)

Ornstein-Uhlenbeck (OU)

$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

Ornstein-Uhlenbeck (OU)

$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

brownian motion

Ornstein-Uhlenbeck (OU)

$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

change towards optimum

brownian motion

Ornstein-Uhlenbeck (OU)

$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

optimal value

Ornstein-Uhlenbeck (OU)

$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

pull towards “optimum”

Ornstein-Uhlenbeck (OU)

$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

strength of selection is proportional to distance of trait from optimal value

Ornstein-Uhlenbeck (OU)

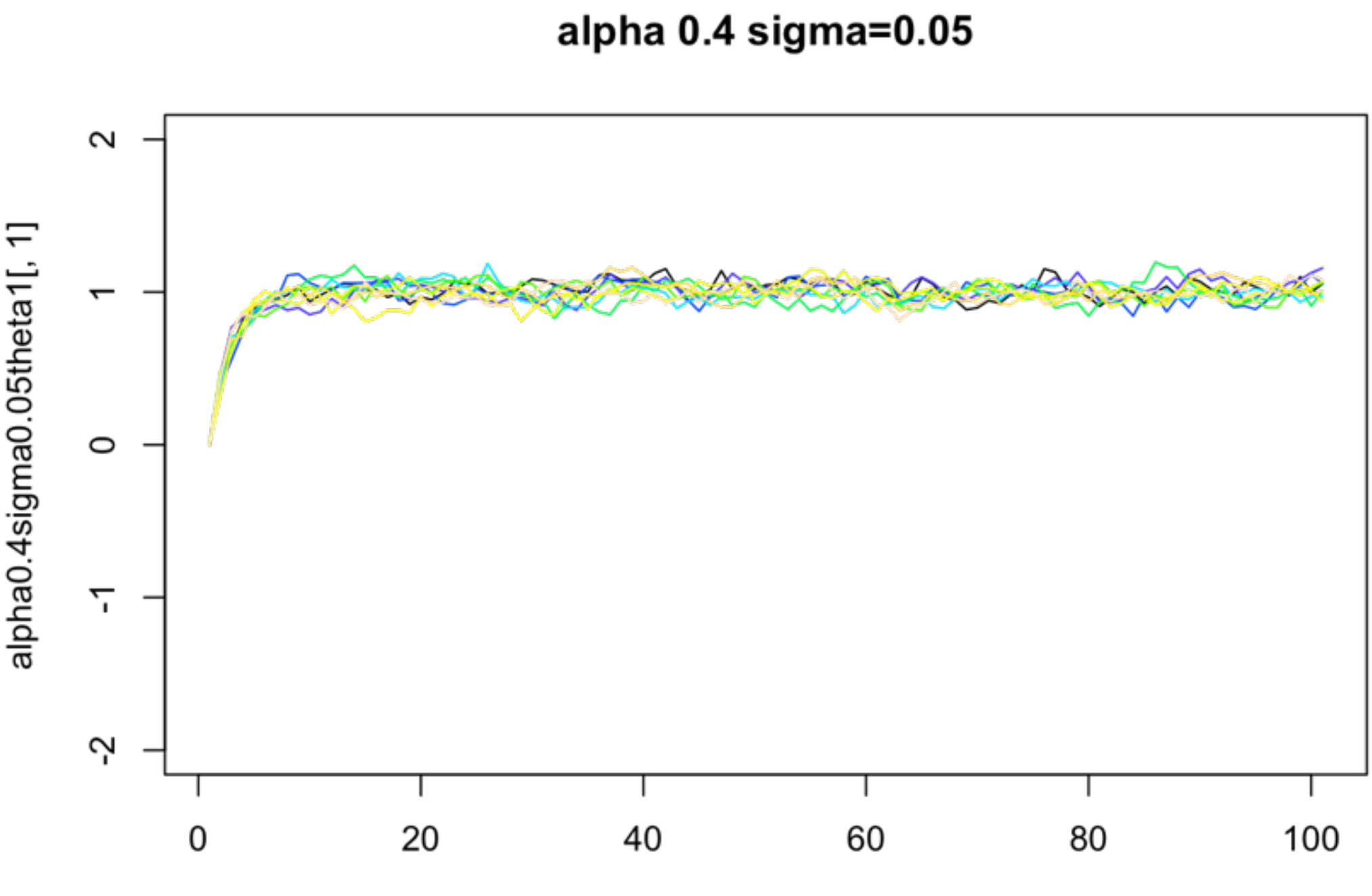
$$dX_{(t)} = \mathbf{0}[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

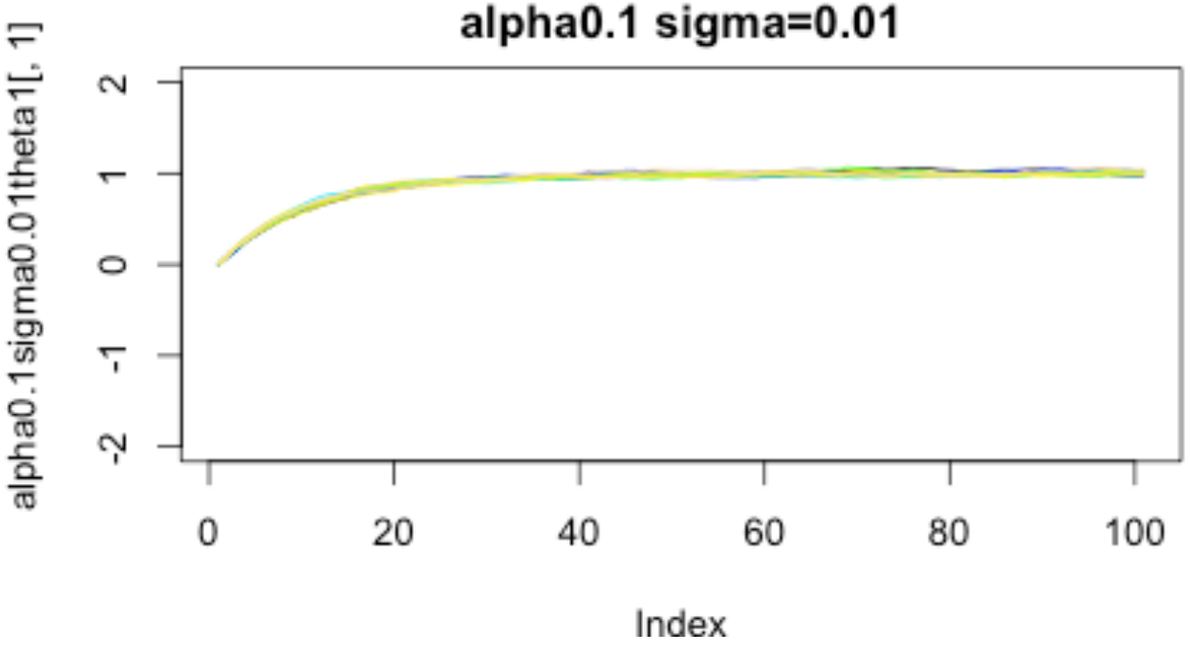
when alpha is 0, OU becomes BM

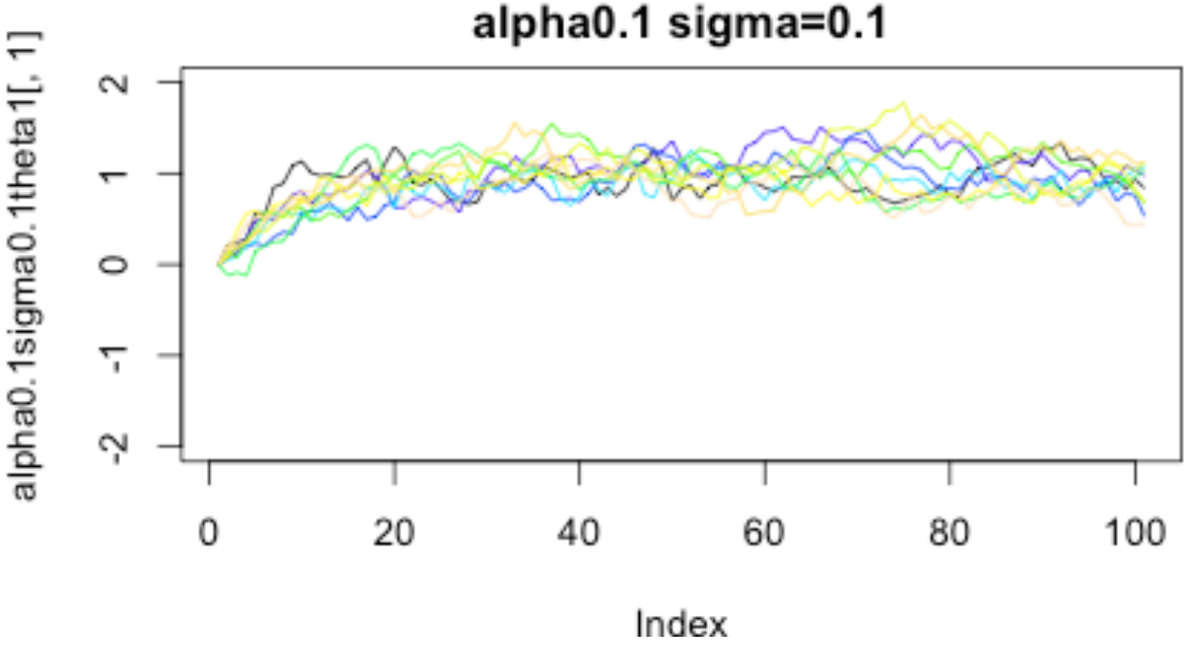
Ornstein-Uhlenbeck (OU)

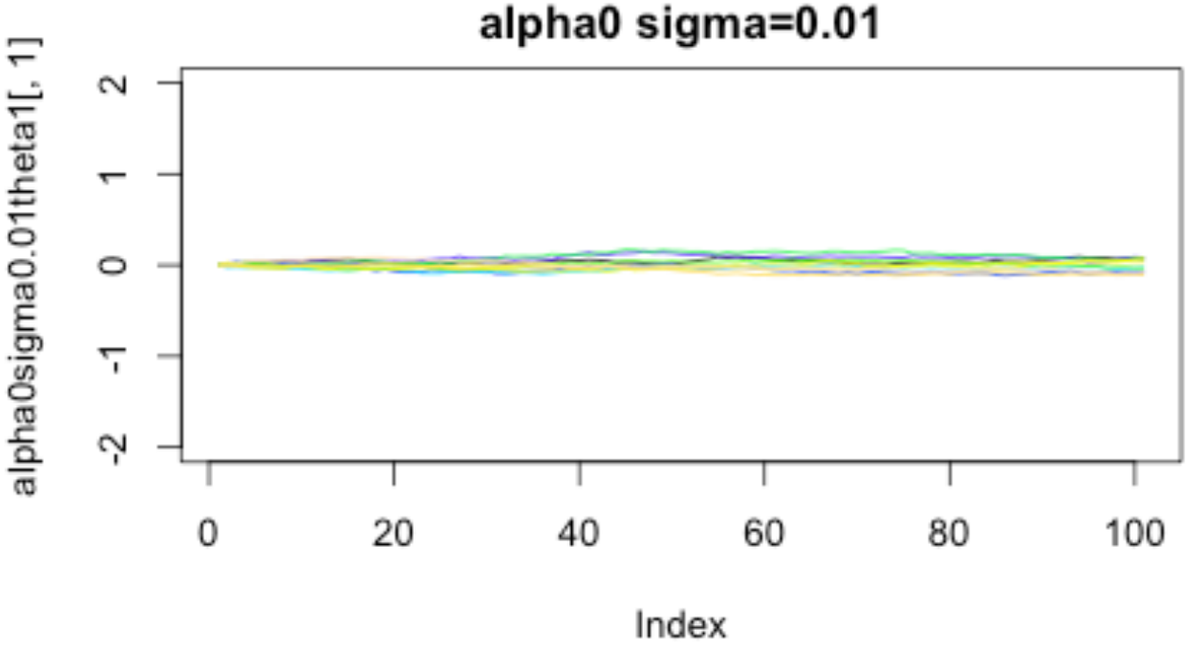
$$dX_{(t)} = \alpha[\Theta - X_{(t)}]dt + \sigma dB_{(t)}$$

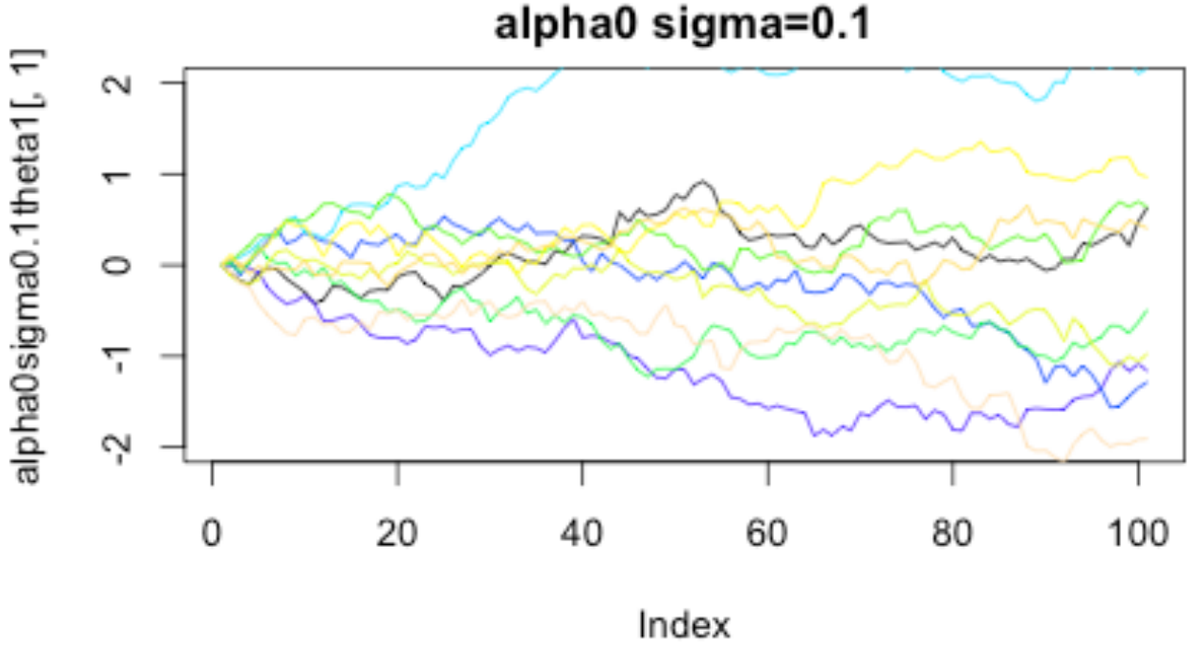
OU evolution




alpha0.1 sigma=0.01
alpha0.1sigma0.01theta1[, 1]
-2
-1
0
1
2
0
20
40
60
80
100
Index


alpha0.1 sigma=0.1
alpha0.1sigma0.1theta1[, 1]
-2
-1
0
1
2
0
20
40
60
80
100
Index

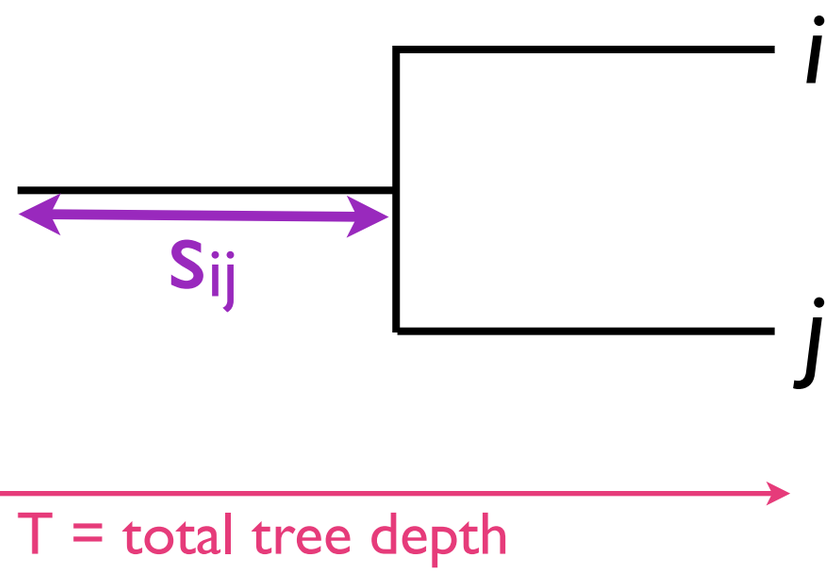

alpha0 sigma=0.01
alpha0sigma0.01theta1[, 1]
-2
-1
0
1
2
0
20
40
60
80
100
Index


alpha0 sigma=0.1
alpha0sigma0.1theta1[, 1]
-2
-1
0
1
2
0
20
40
60
80
100
Index

Ornstein-Uhlenbeck Model (OU)

- Evolution has a tendency to move towards some medial value
- "Brownian motion with a spring"
- Three parameters: starting value (Θ), rate (σ^2), and constraint parameter (α)

$$V_{ij} = \frac{\sigma^2}{\alpha} e^{-2\alpha(T-s_{ij})}(1 - e^{-2\alpha s_{ij}})$$

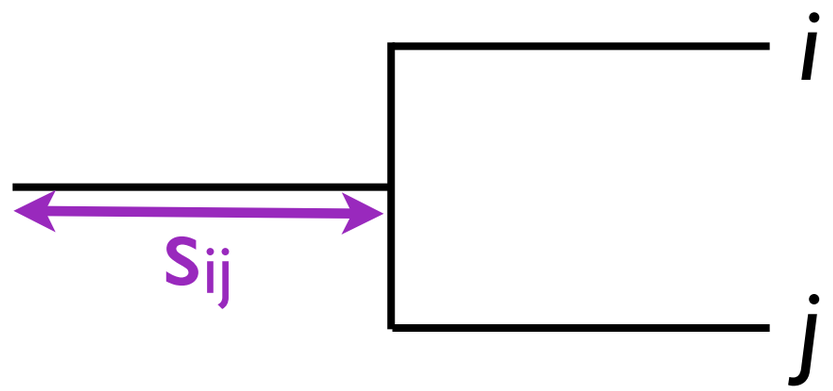


Early Burst Model (EB)

- Rate of evolution slows through time
- Highest rate at the root of the tree
- Three parameters: starting value (Θ), starting rate (σ^2_o), and rate change (r)

$$r(t) = \sigma_0^2 \, e^{rt}$$

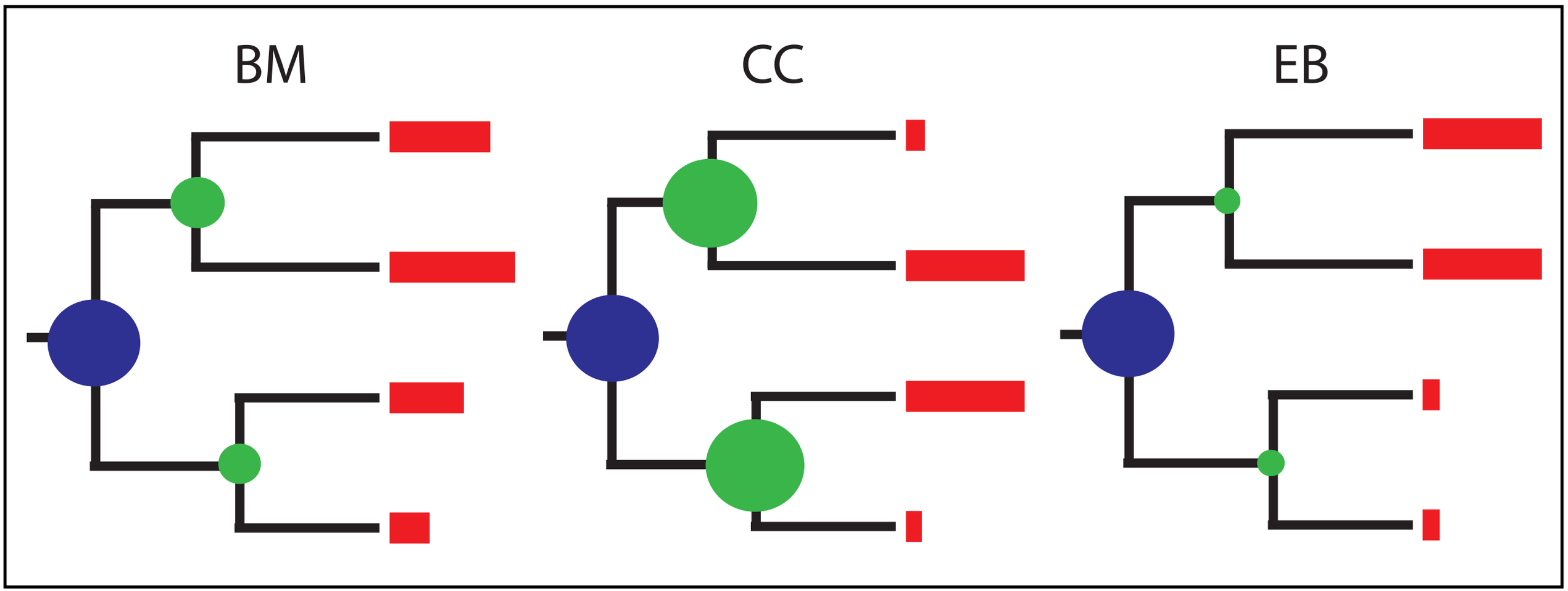
$$V_{ij} = \int_0^{s_{ij}} \sigma_0^2 \, e^{rt} \, dt = \sigma_0^2 \, \frac{e^{r s_{ij}}}{r}$$



Why these three?

- BM is assumed by almost all phylogenetic comparative methods
- EB corresponds to one idea of adaptive radiation
- OU may capture the importance of constraints on evolution

How do we tell these models apart?


BM
CC
EB

Brownian motion = drift

OU = stabilizing selection

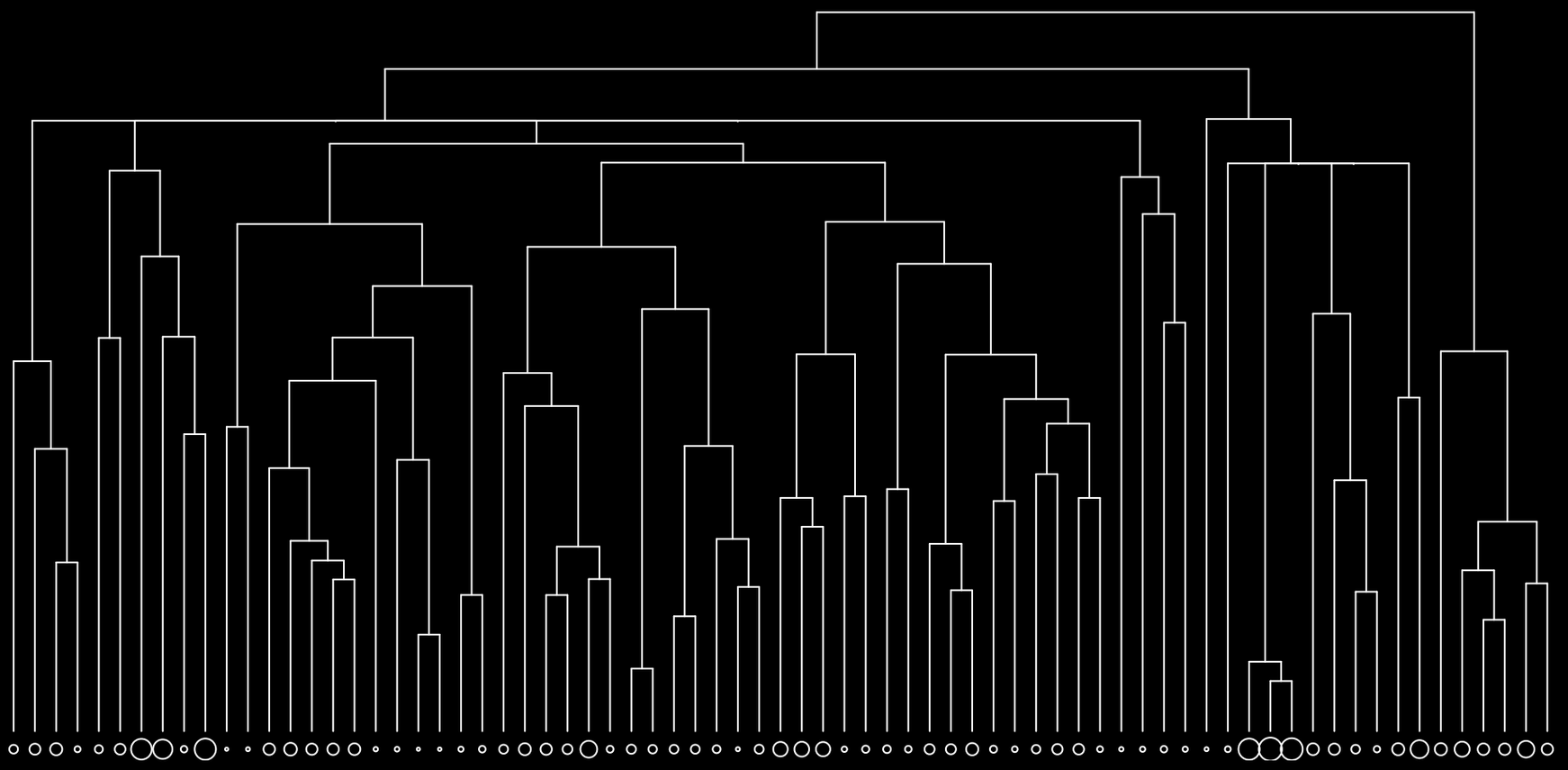
Brownian motion = drift or many other processes

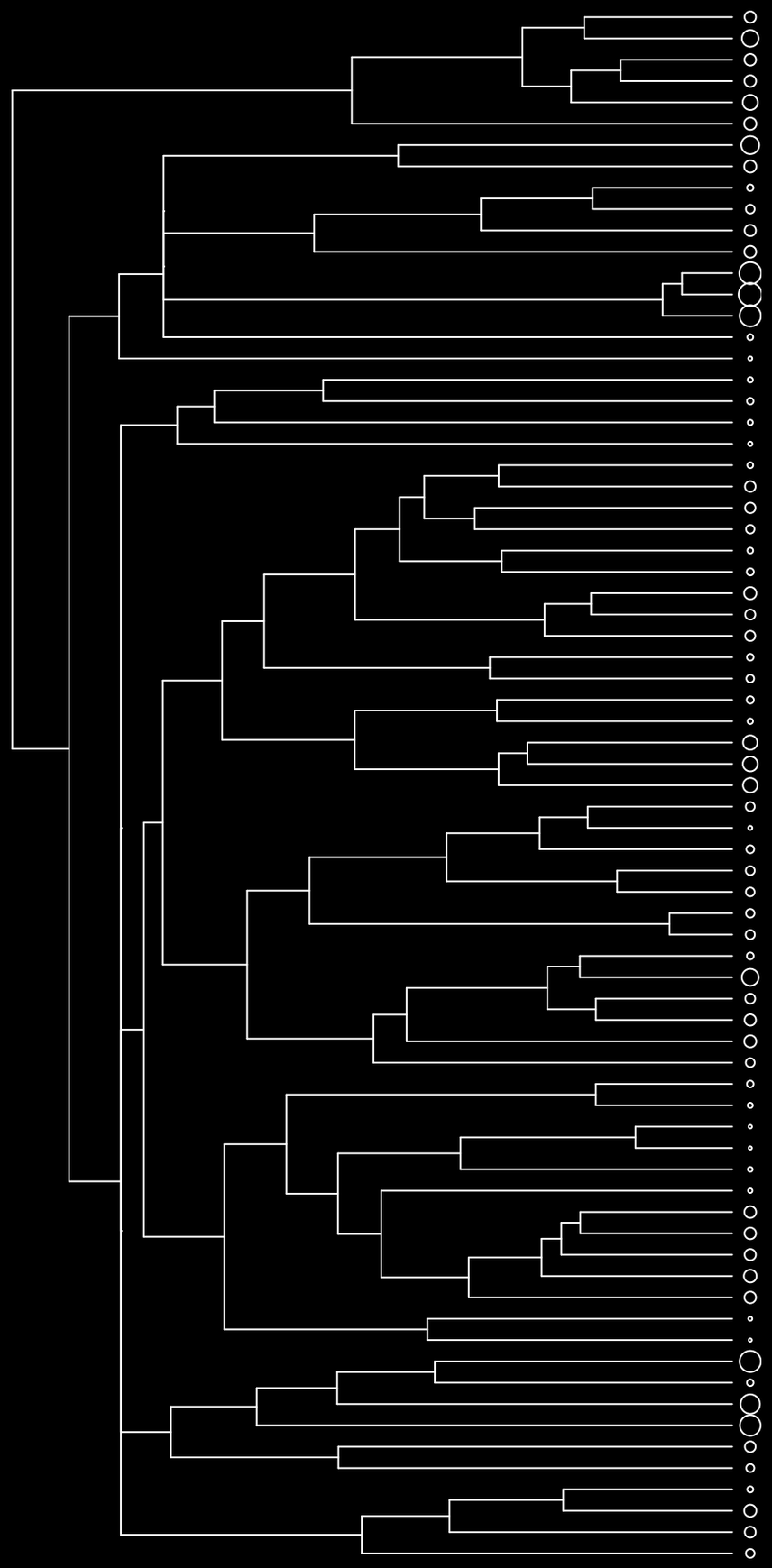
OU = stabilizing selection or many other processes

Example: *Anolis* lizards

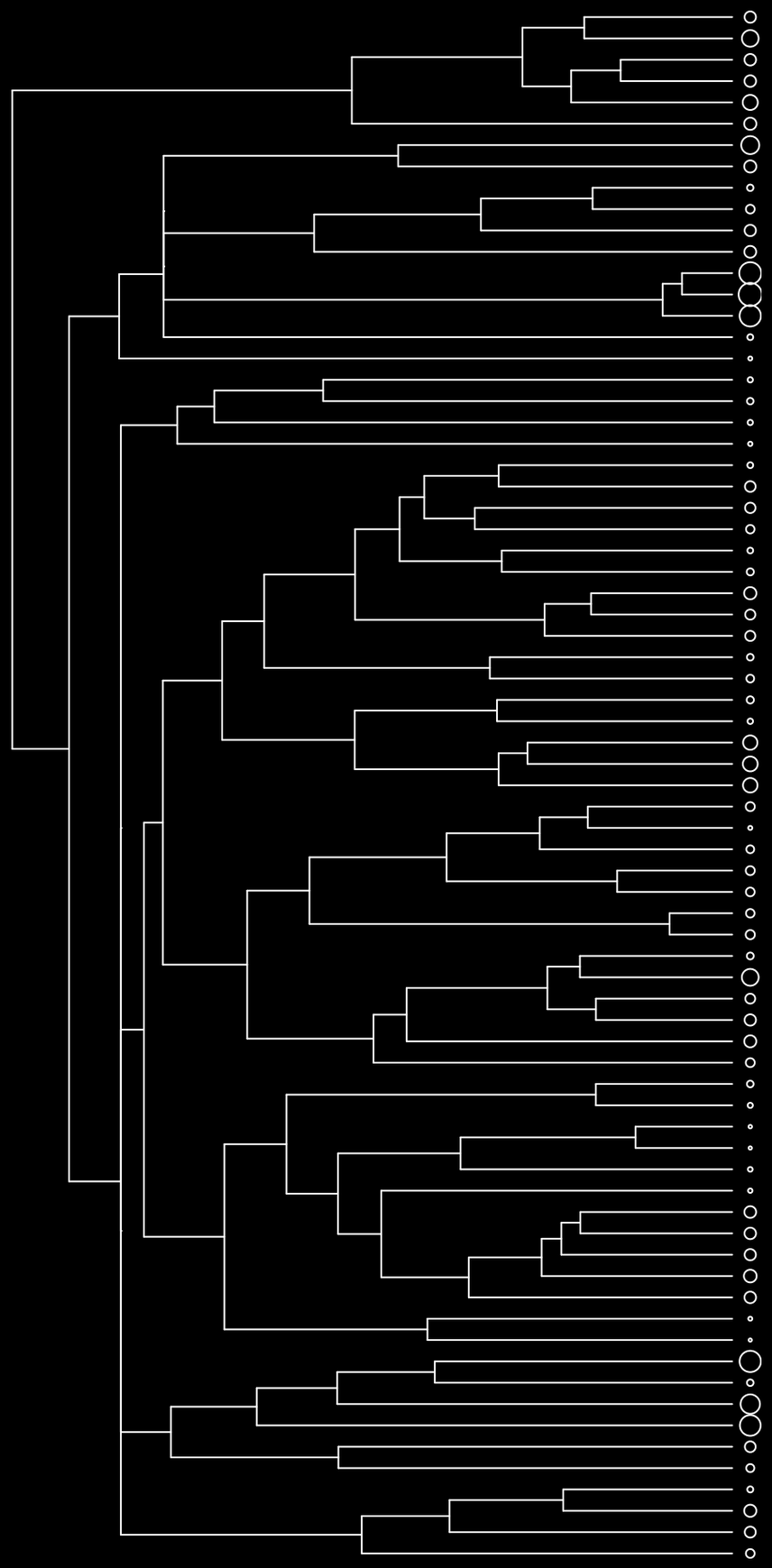
- Lizards on Caribbean islands
- Phylogenetic and body size data for 73 species (out of ~140 total)

Anolis baleatus

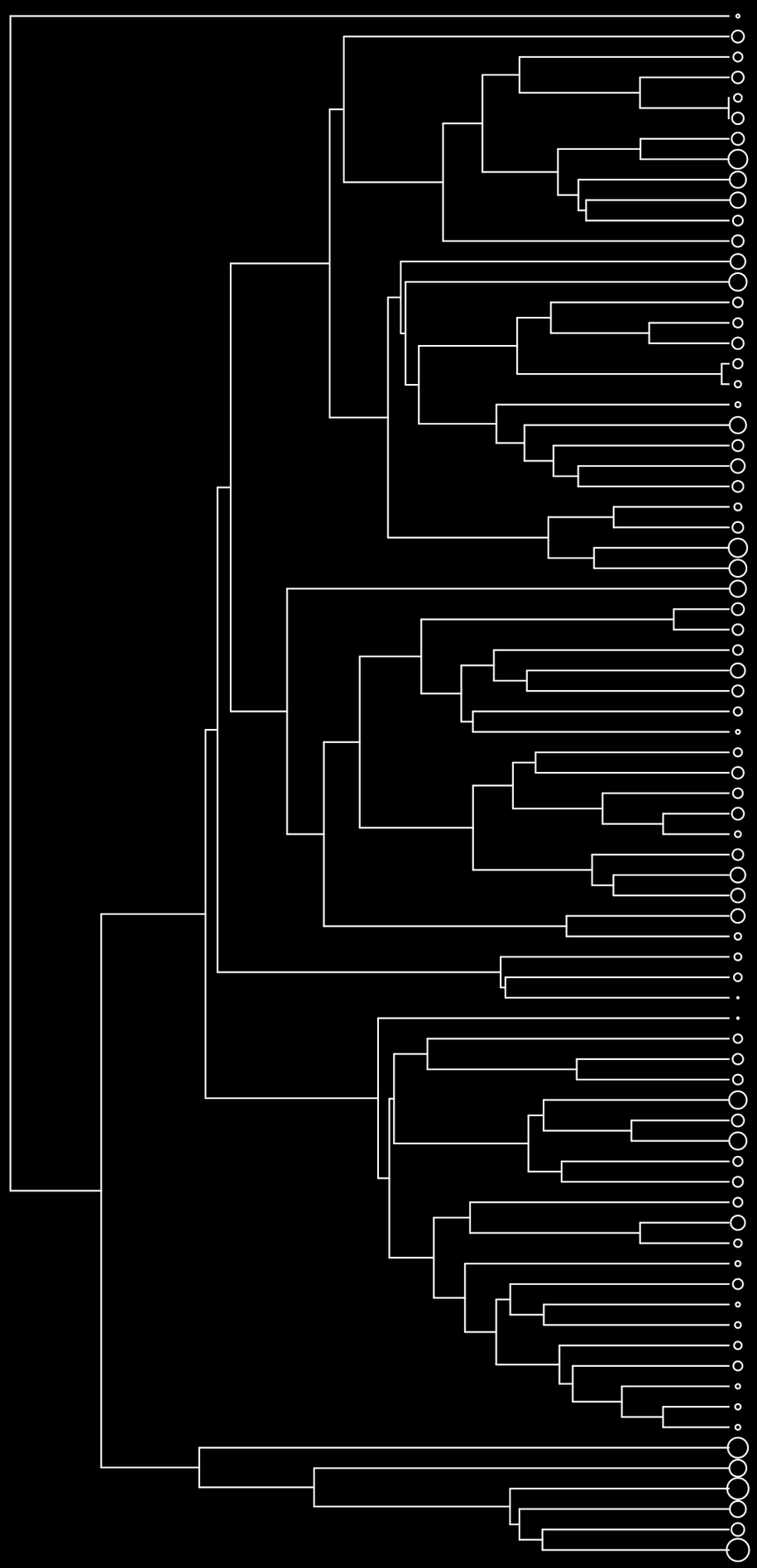




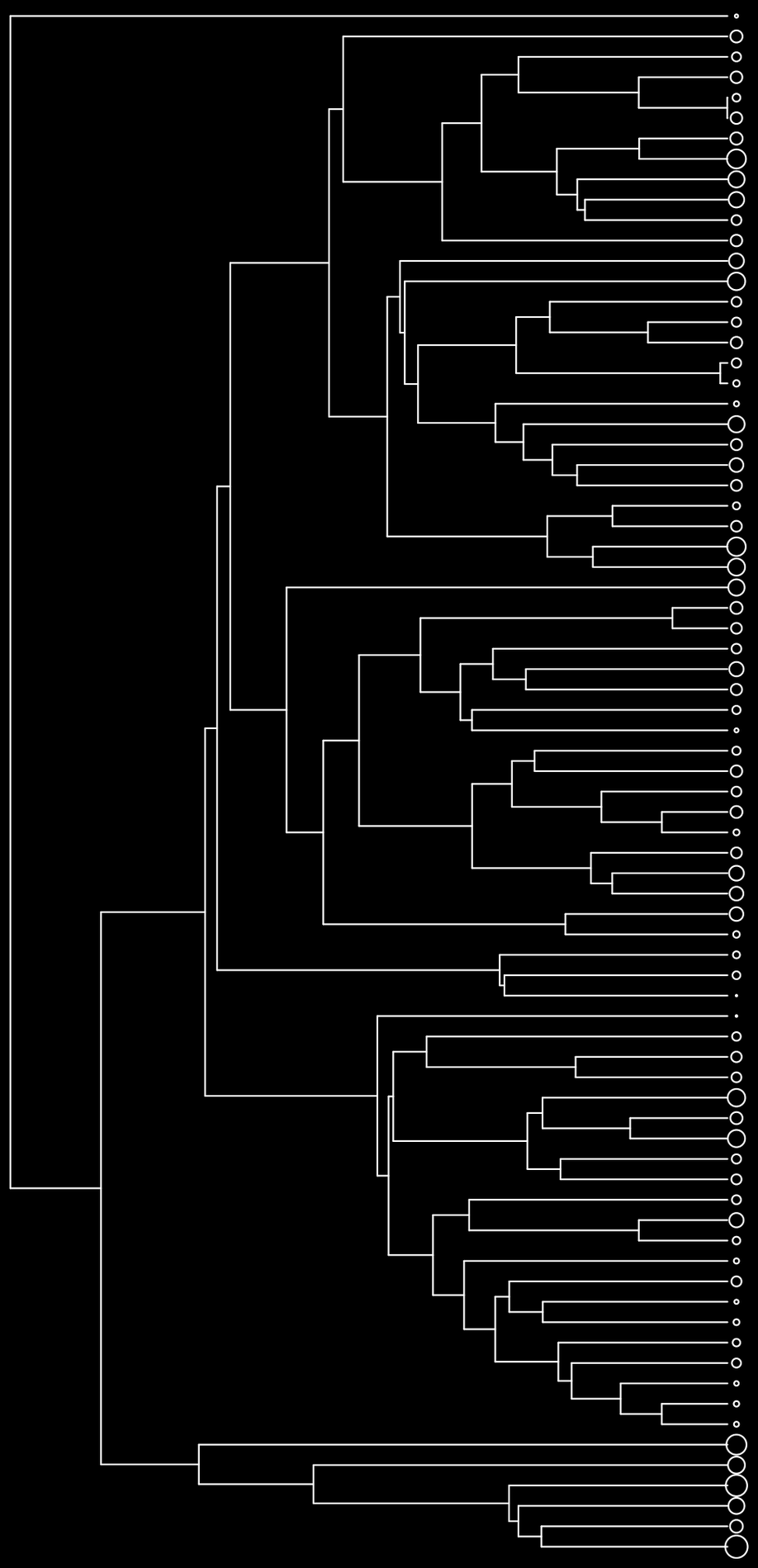
Model	Parameter estimates	lnL	Akaike weight
BM			
EB			
OU			



Model	Parameter estimates	lnL	Akaike weight
BM	$\sigma^2 = 0.004$	-18.2	0.58
EB	$\sigma^2 = 0.006$ $r = -0.01$	-18.1	0.2
OU	$\sigma^2 = 0.004$ $\alpha = 0$	-18.2	0.22



Cichlids in Lake Tanganyika



Cichlids in Lake Tanganyika

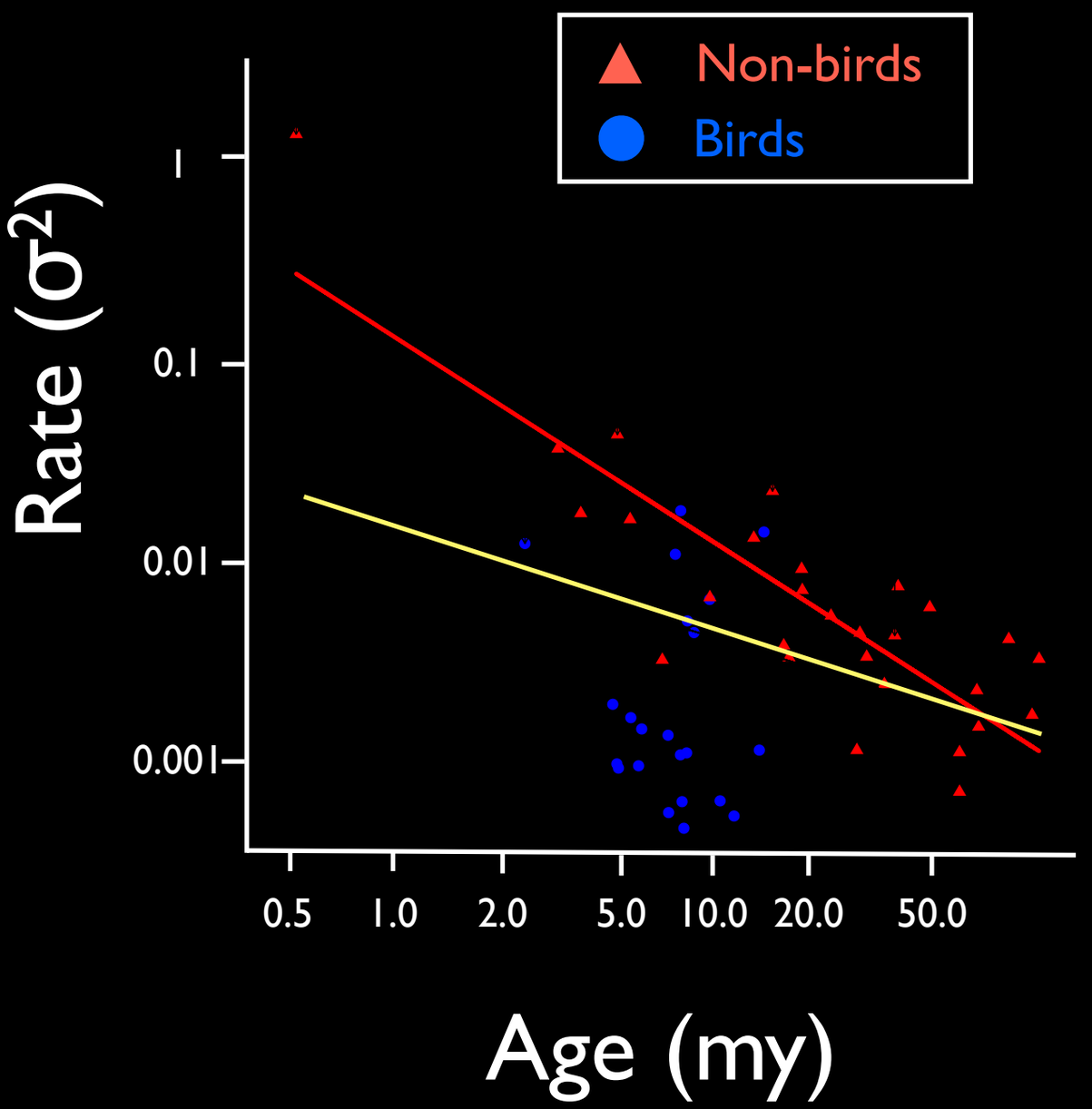
Model	Parameter estimates	lnL	Akaike weight
BM	$\sigma^2 = 0.02$	-62.3	0
EB	$\sigma^2 = 0.02$ $r = 0$	-62.3	0
OU	$\sigma^2 = \ldots$ $\alpha = \ldots$	-33.3	1

Squamates
Birds
Fish
Insects
Mammals
Amphibians
Body size
BM
CC
EB
NA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49
Body shape

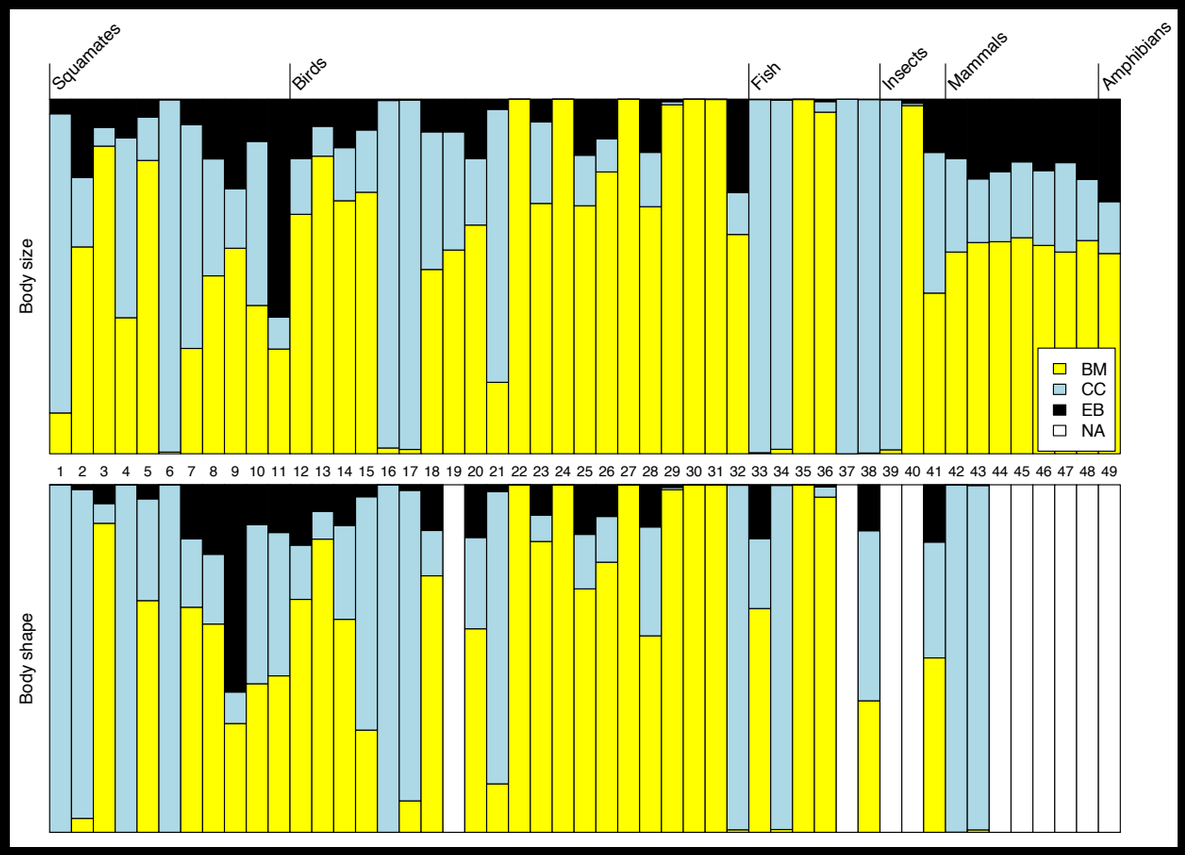
Table 1. Number of clades and subclades showing support for each of the three models (BM, CC, EB) for body size and body shape. We count both the number of clades with the highest AICc values for a particular model ("maximum w") and those with weights greater than 0.95 ("w > 0.95").

Clades	Data set	n	Criterion	BM	CC	EB
All full clades	Body size	49	Maximum w	35	13	1
			w > 0.95	9	8	0
	Body shape	39	Maximum w	24	14	1
			w > 0.95	8	8	1
All subclades	Body size	284	Maximum w	200	74	10
			w > 0.95	0	22	0
	Body shape	205	Maximum w	99	101	5
			w > 0.95	0	41	0

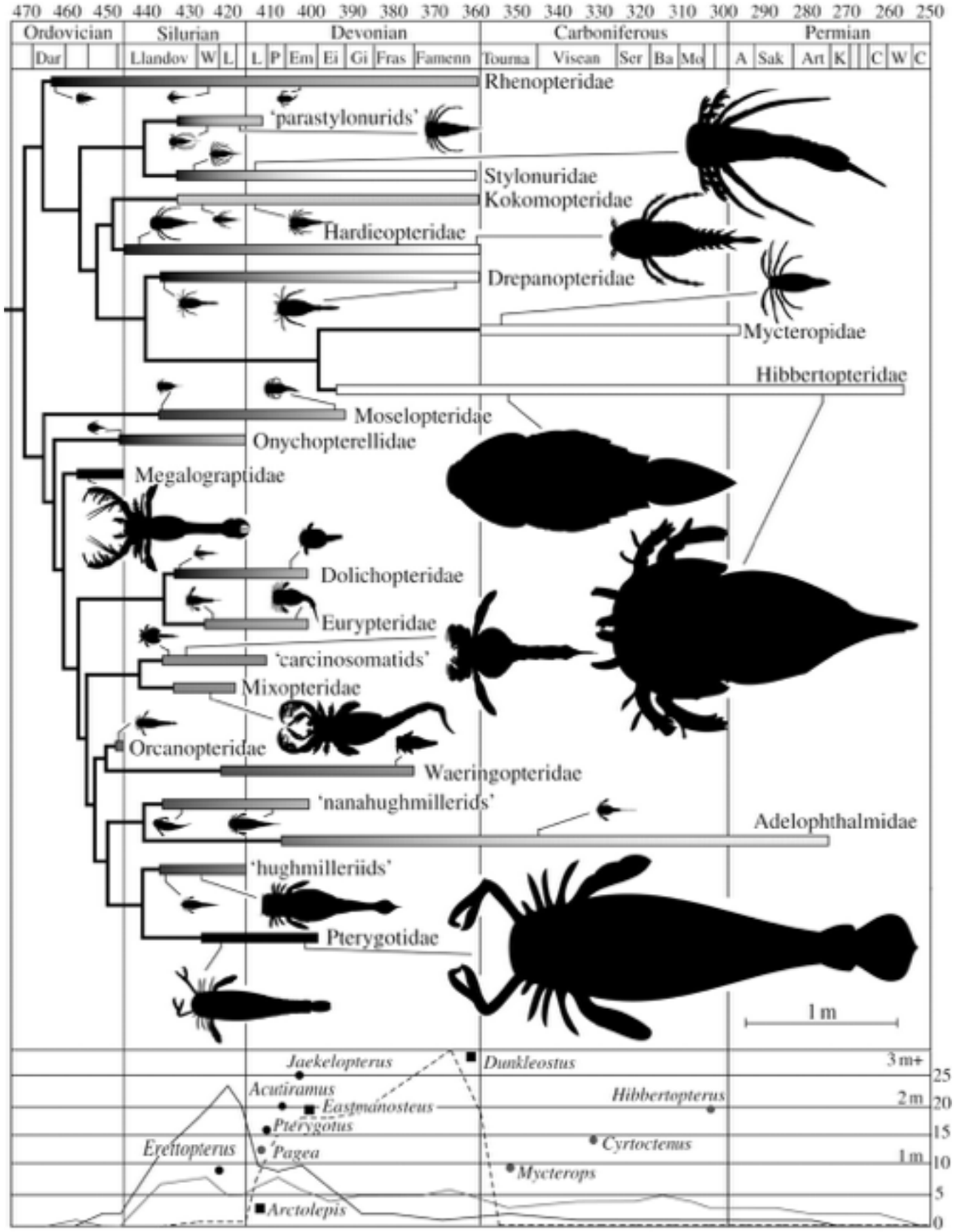
Brownian "Rates" Scale with Time



- "Adaptive radiation" pattern very rare in this data set
- Constraints dominate over long time periods
- Brownian motion is sometimes a poor fit to real data

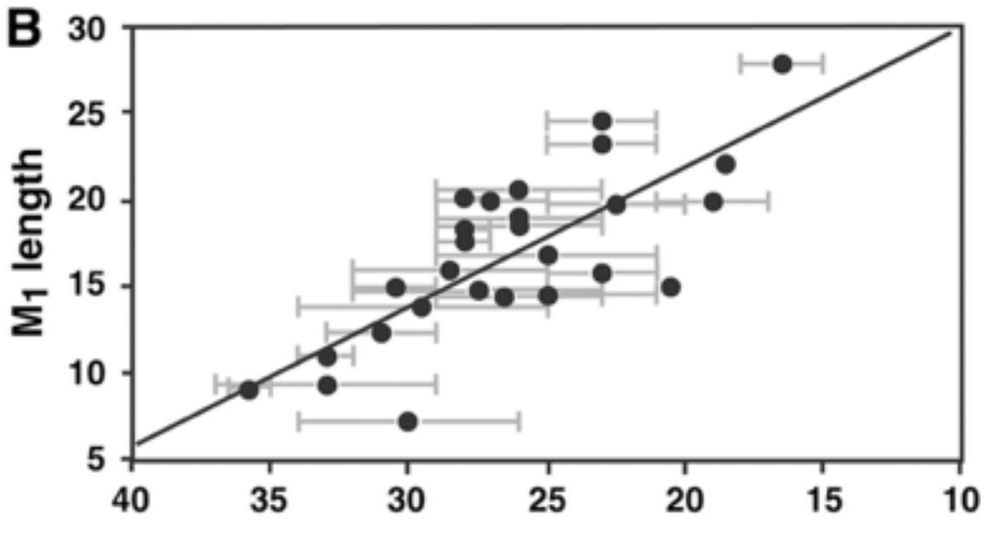


more models…



Lamsdell, J. C., and S. J. Braddy. 2010. Cope's Rule and Romer's theory: patterns of diversity and gigantism in eurypterids and Palaeozoic vertebrates. Biol Letters **6**:265-269.

BM with trend



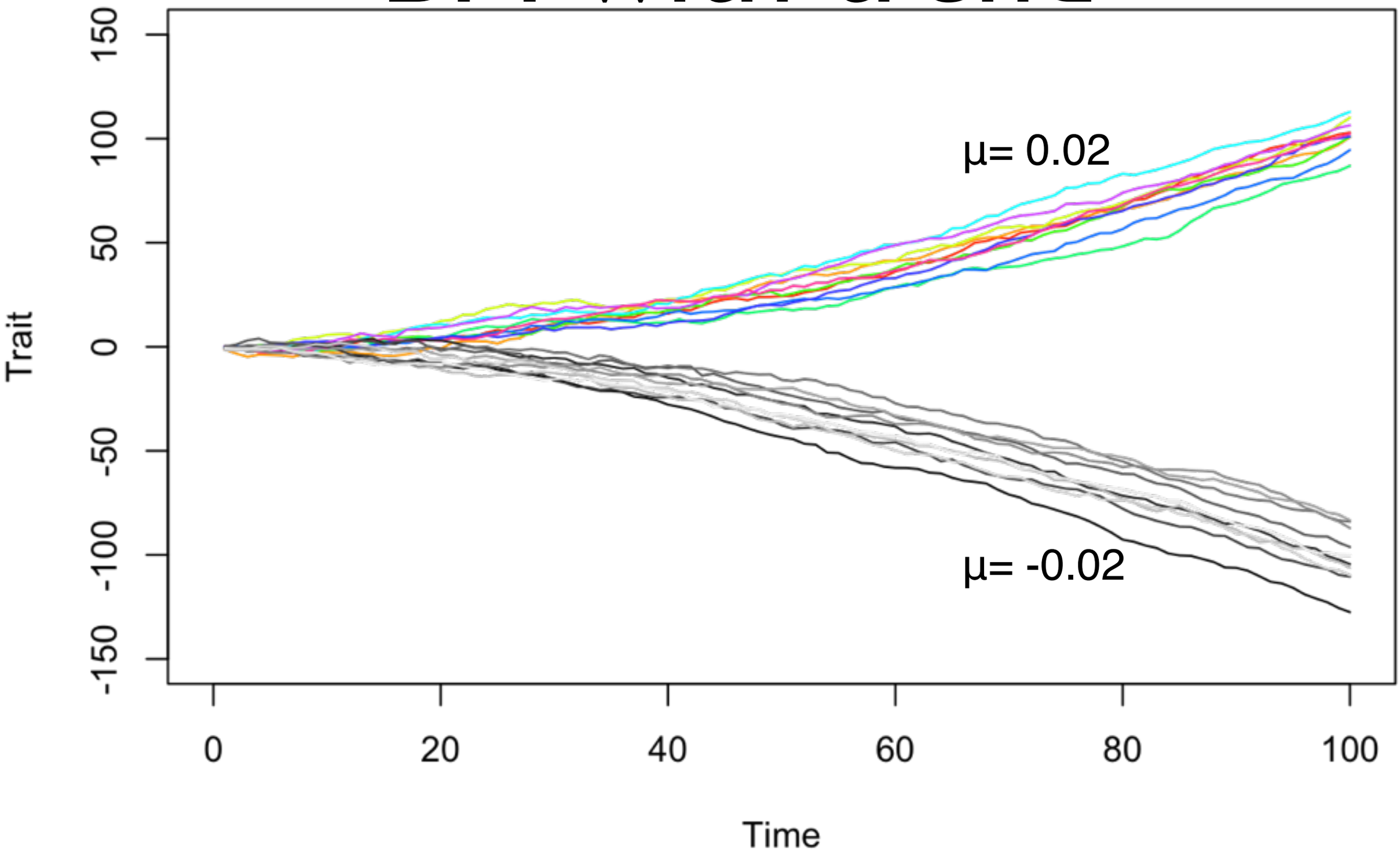
$$dX_{(t)} = \sigma dB_t$$

rate

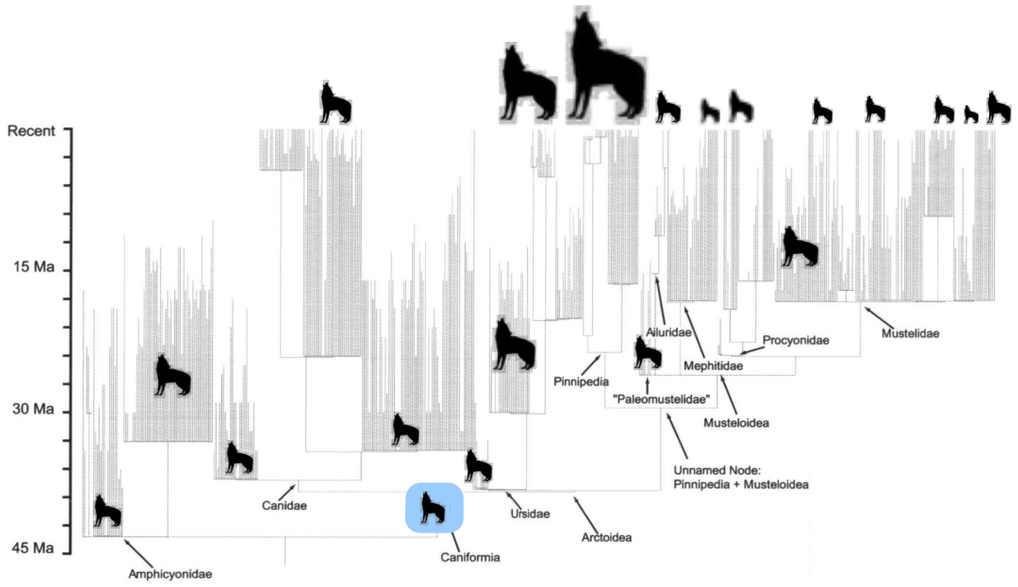
normal distribution where mean = t * μ

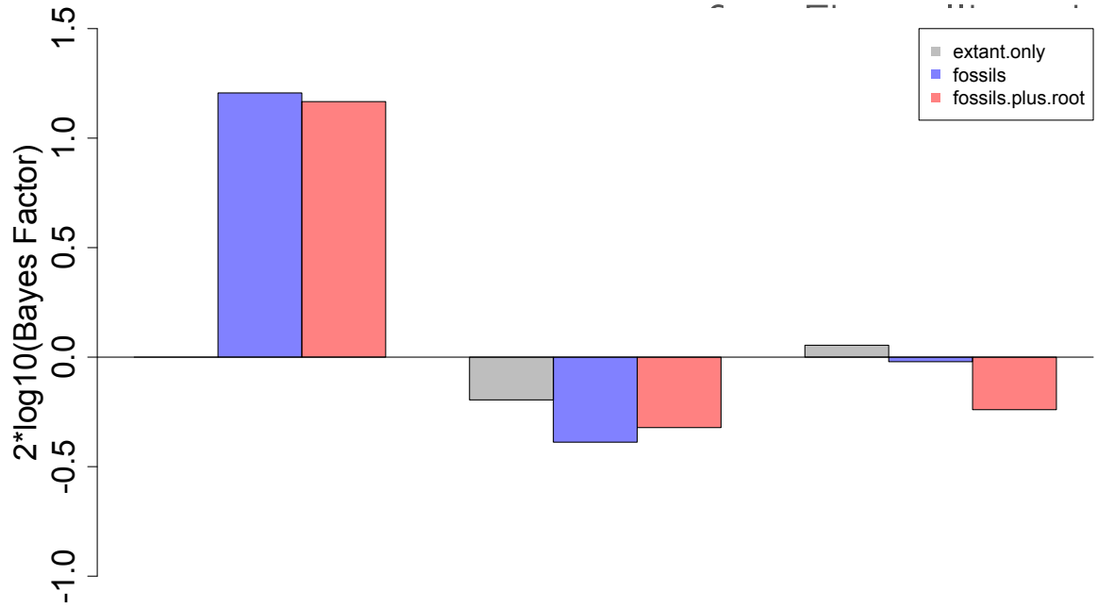
trait increases when μ > 0, decreases when μ < 0

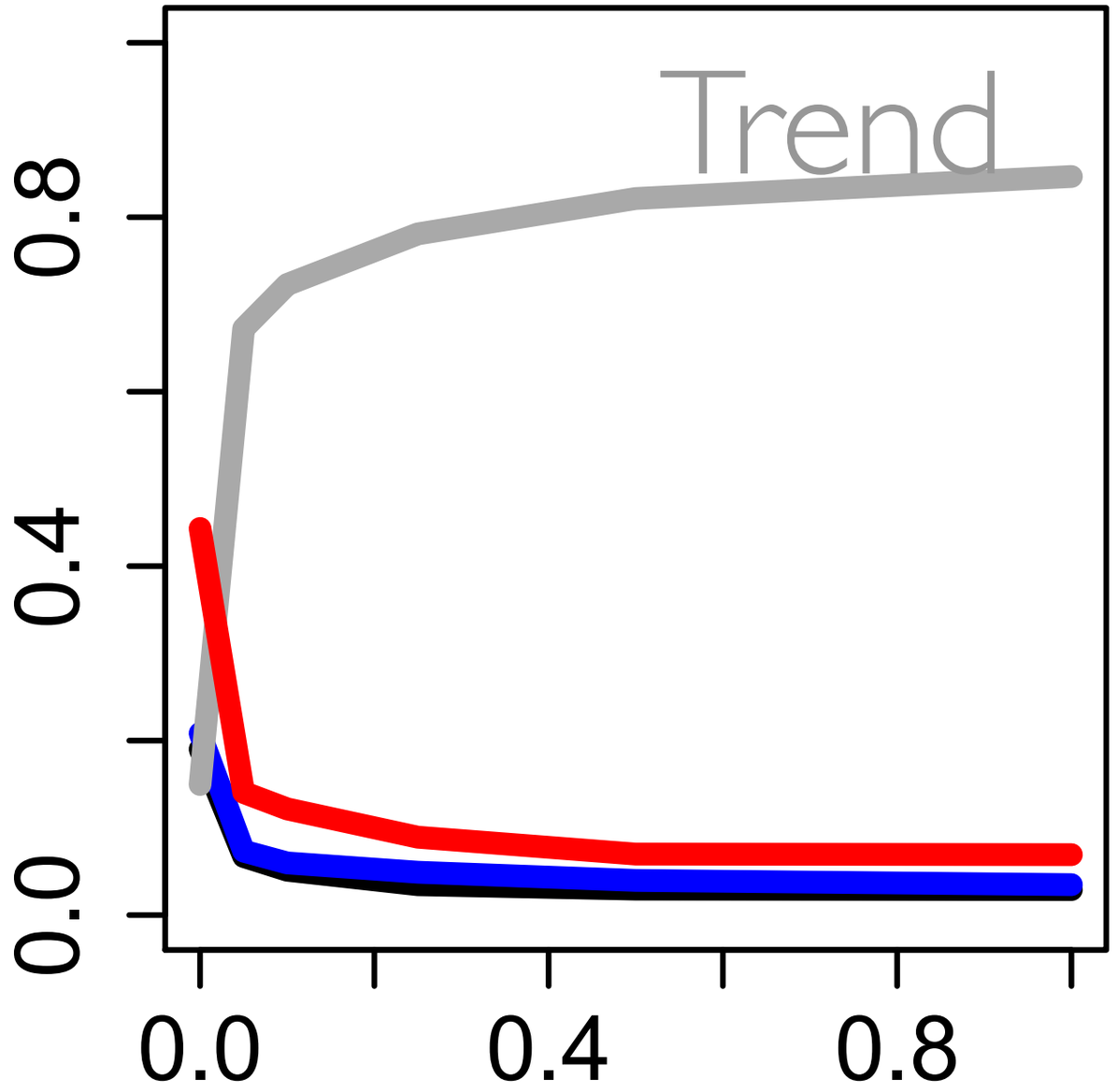
BM with trend



usually need fossils to detect trends

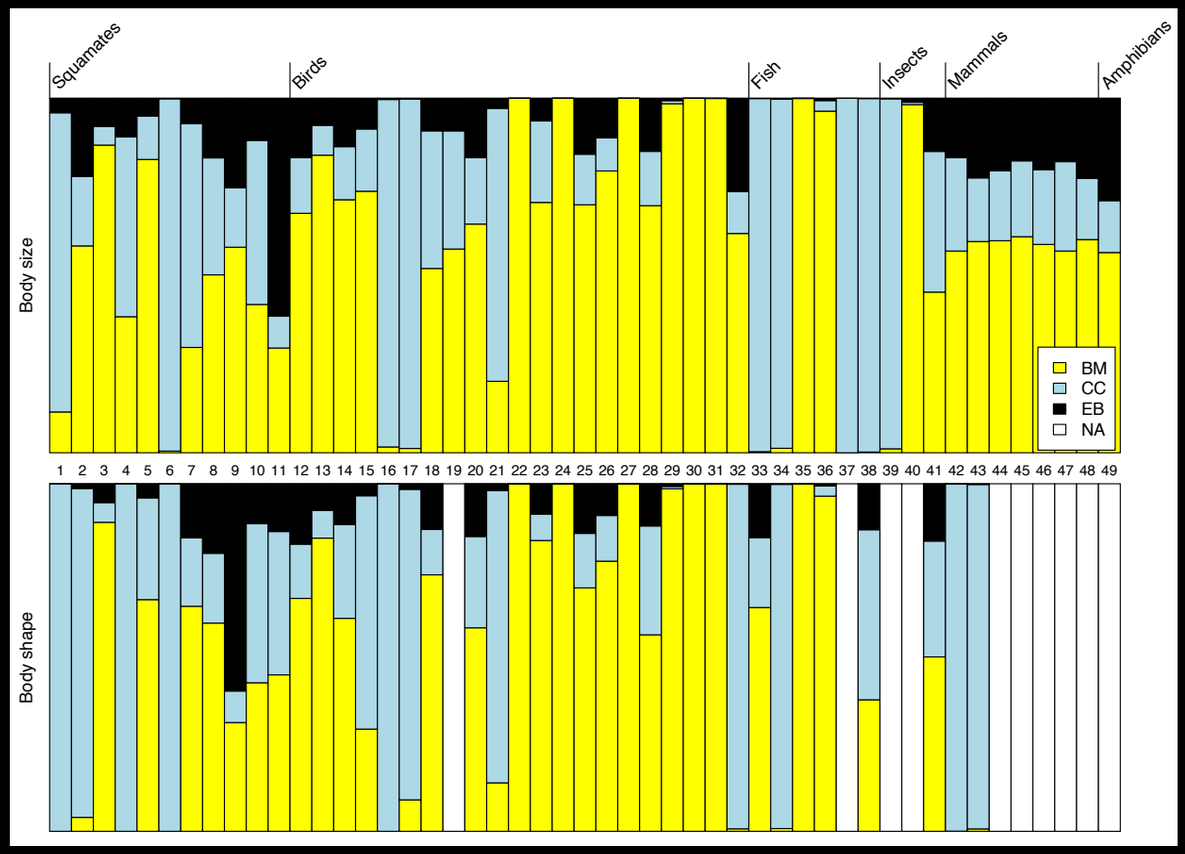






Model Adequacy

- we commonly select from among a pool of models
- model adequacy asks if any of them are good



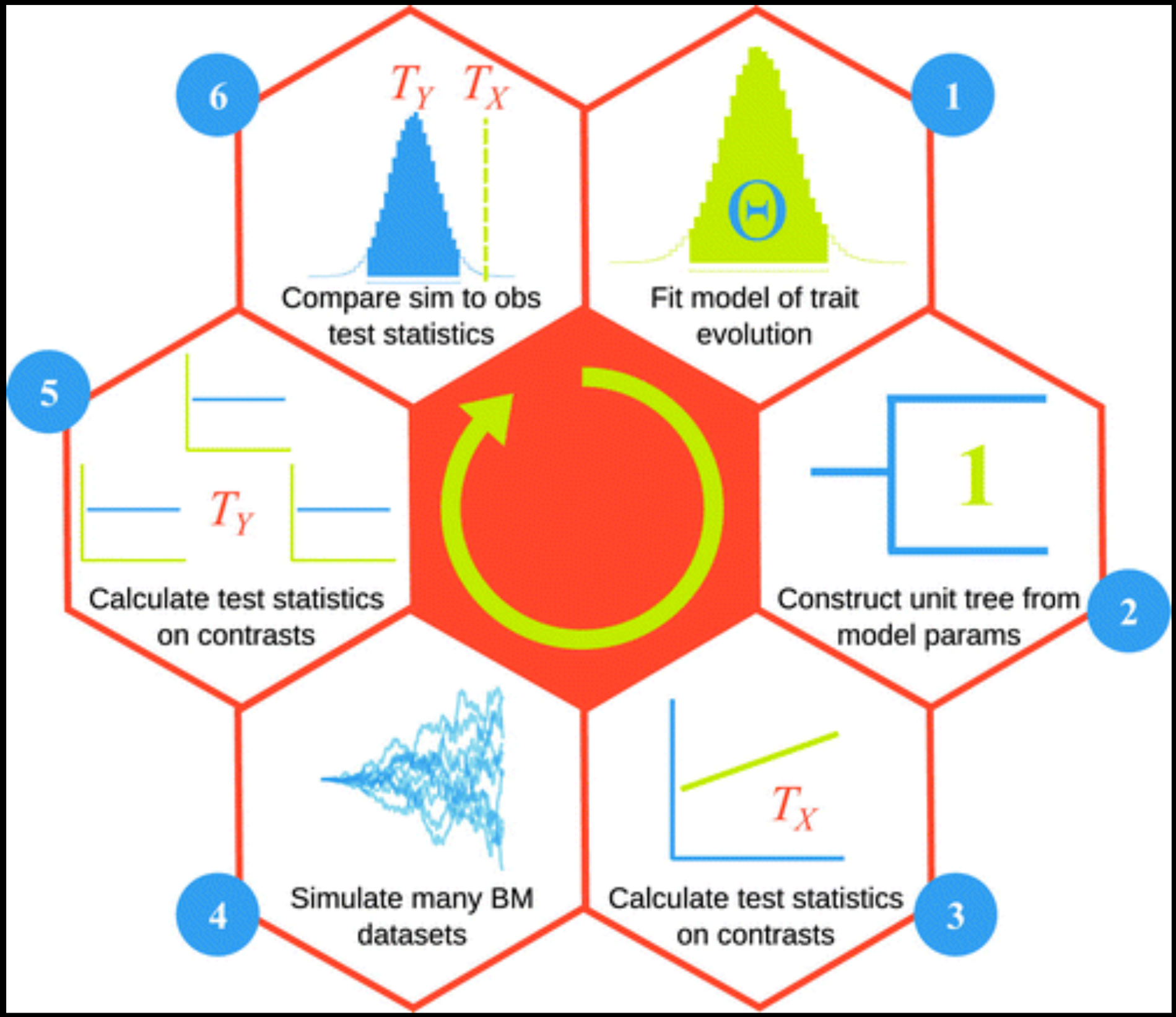
Assess inadequacy with Arbutus

- https://github.com/mwpennell/arbutus
- library(devtools)
- install_github("mwpennell/arbutus")

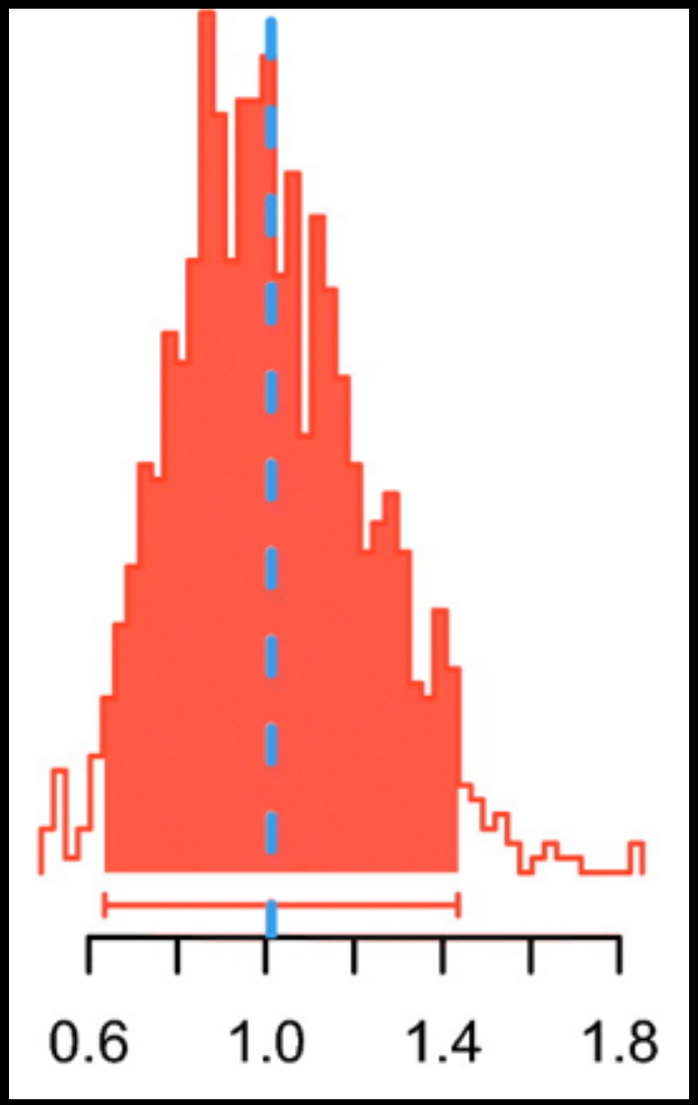
Arbutus

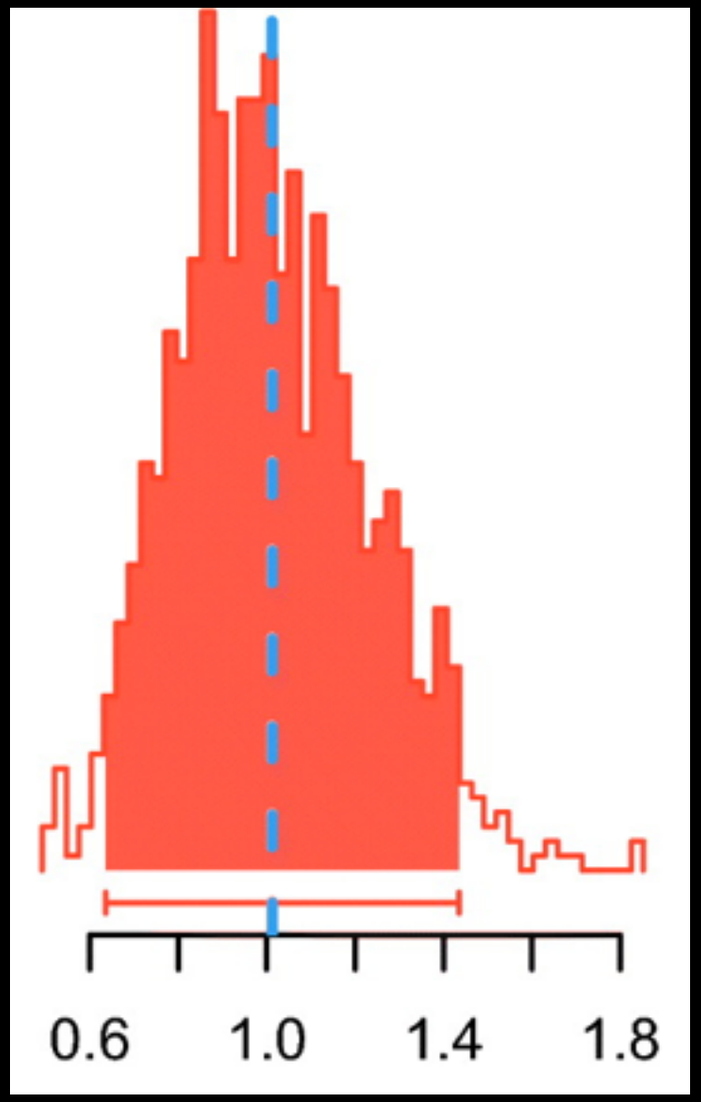
- asks if best fit model parameters predict summary statistics from the data
- most of these summary stats are based on contrasts
  - mean of squared contrasts
  - CV of absolute contrasts
  - slope of abs contrasts against expected variance

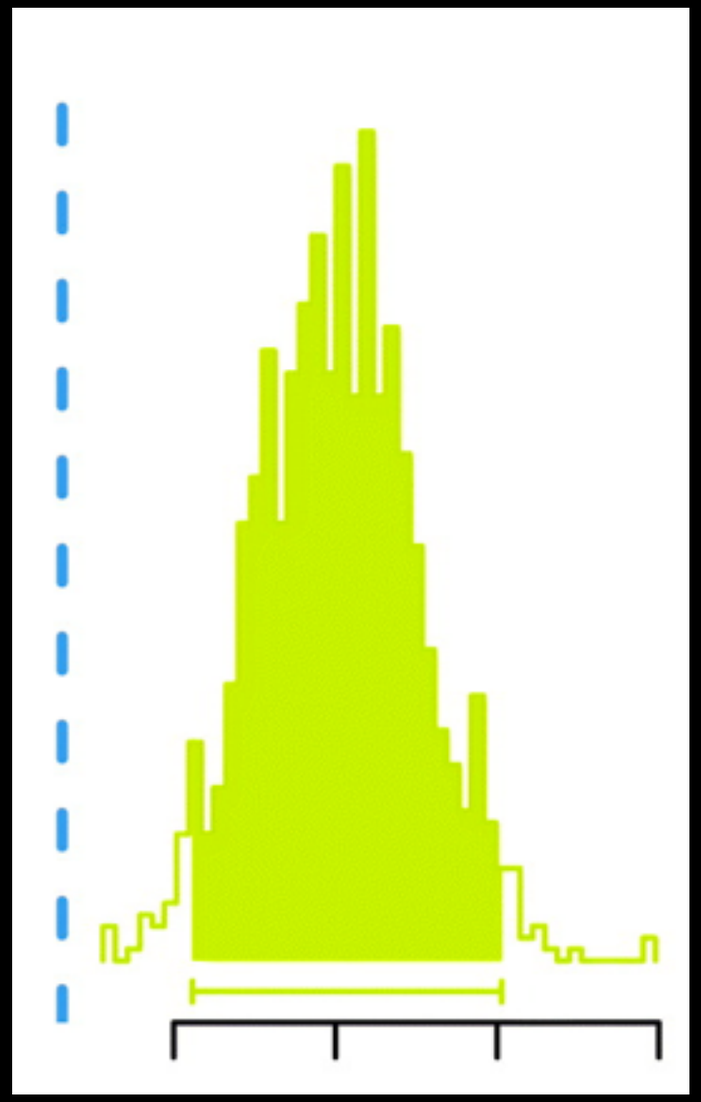
etc…

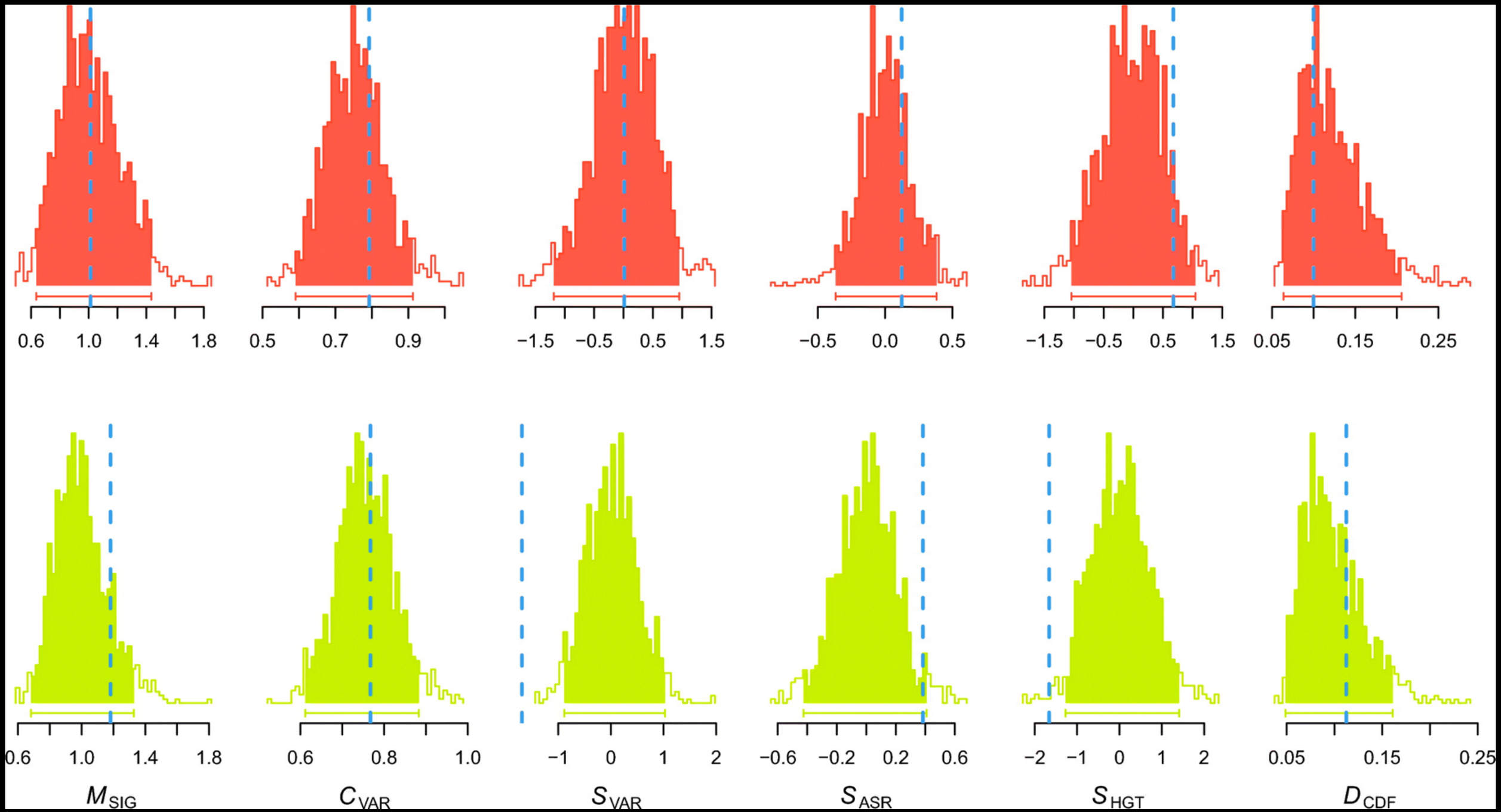


Model Adequacy and the Macroevolution of Angiosperm Functional Traits, Pennell et al 2015


0.6
1.0
1.4
1.8


0.6
1.0
1.4
1.8




0.6 1.0 1.4 1.8
0.5 0.7 0.9
−1.5 −0.5 0.5 1.5
−0.5 0.0 0.5
−1.5 −0.5 0.5 1.5
0.05 0.15 0.25
0.6 1.0 1.4 1.8
0.6 0.8 1.0
−1 0 1 2
−0.6 −0.2 0.2 0.6
−2 −1 0 1 2
0.05 0.15 0.25
M_{SIG}
C_{VAR}
S_{VAR}
S_{ASR}
S_{HGT}
D_{CDF}

What about “phylogenetic signal”?

“Phylogenetic signal”

A pattern where closely related species on a phylogenetic tree have trait values that are more similar than expected by chance.

1. We expect phylogenetic signal under a wide range of evolutionary models.

— Brownian motion
— OU with small alpha
— multi-peak OU
— early burst

2. Phylogenetic signal is a pattern, not a process

3. Phylogenetic signal is NOT a constraint

In fact, unconstrained models (like BM) create lots of phylogenetic signal, while constrained models (like OU) can result in very little phylogenetic signal

Measuring Phylogenetic Signal

- Blomberg's K statistic
- Pagel's lambda

Measuring Phylogenetic Signal

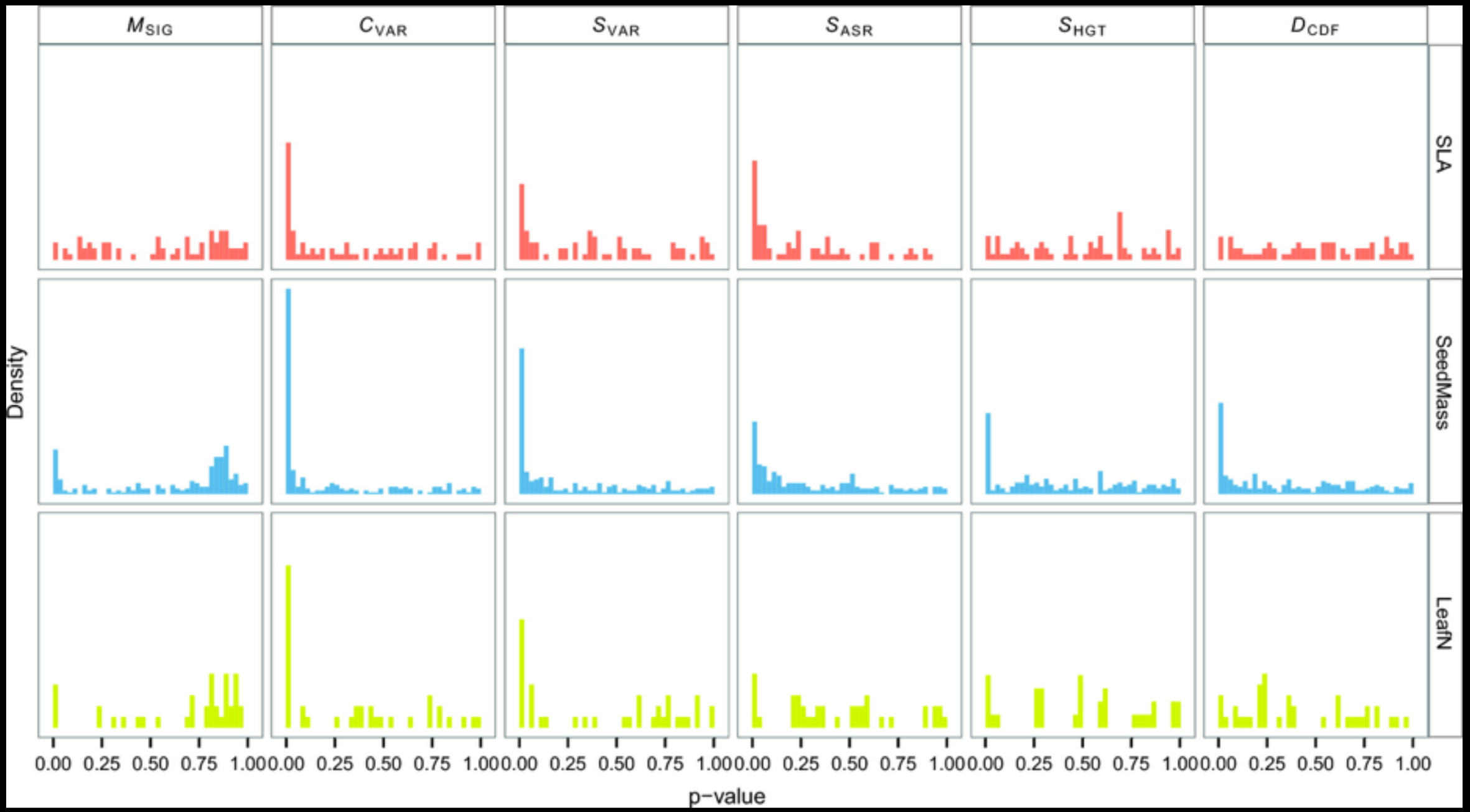
- Blomberg's K statistic
- measure of partitioning of variance (compared to BM)
    - K >1 variance among
    - K < 1 variance within

Measuring Phylogenetic Signal

- Blomberg's K statistic (comparison to BM)
  - measure of partitioning of variance (compared to BM)
    - K >1 variance among
    - K < 1 variance within
- **Pagel's lambda (branch length transformation)**
  - **similarity of species correlations compared to expected under BM**
  - **lambda = 0: no correlation;**
  - **lambda =1: correlation same as Brownian**

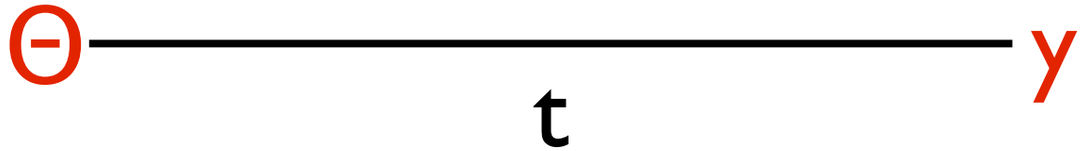
Table 1: Description of test statistics used to assess model adequacy

Test statistic	Description
M_{SIG}	The mean of the squared contrasts. This is equivalent to the restricted maximum likelihood estimator of the Brownian motion rate parameter σ^2 (Garland et al. 1992; Rohlf 2001). M_{SIG} is a metric of overall rate. Violations detected by M_{SIG} indicate whether the overall rate of trait evolution is over- or underestimated.
C_{VAR}	The coefficient of variation (standard deviation/mean) of the absolute value of the contrasts. If C_{VAR} calculated from the observed contrasts is greater than that calculated from the simulated contrasts, it suggests that we are not properly accounting for rate heterogeneity across the phylogeny. If C_{VAR} from the observed is smaller, it suggests that contrasts are even more than the model assumes. We use the coefficient of variation rather than the variance because the mean and variance of contrasts can be highly correlated.
S_{VAR}	The slope of a linear model fitted to the absolute value of the contrasts against their expected variances (following Garland et al. 1992). Each (standardized) contrast has an expected variance proportional to the sum of the branch lengths connecting the node at which it is computed to its daughter lineages (Felsenstein 1985). Under a model of Brownian motion, we expect no relationship between the contrasts and their variances. We use it to test whether contrasts are larger or smaller than we expect based on their branch lengths. If, for example, more evolution occurred per unit time on short branches than long branches, we would observe a negative slope. If S_{VAR} calculated from the observed data deviates substantially from the expectations, a likely explanation is branch length error in the phylogenetic tree.
S_{ASR}	The slope of a linear model fitted to the absolute value of the contrasts against the ancestral state inferred at the corresponding node. We estimated the ancestral state using the least squares method suggested by Felsenstein (1985) for the calculation of contrasts. (We note that this is not technically an ancestral state reconstruction [see Felsenstein 1985]; it is more properly thought of as a weighted average value for each node.) We used this statistic to evaluate whether there is variation in rates relative to the trait value. For example, do larger organisms evolve proportionally faster than smaller ones?
S_{HGT}	The slope of a linear model fitted to the absolute value of the contrasts against node depth (after Purvis and Rambaut 1995). This is used to capture variation relative to time. It is alternatively known as the "node-height test" and has been used to detect early bursts of trait evolution during adaptive radiations (for uses and modifications of this test see Freckleton and Harvey 2006; Slater and Pennell 2014).
D_{CDF}	The D statistic obtained from a Kolmolgorov-Smirnov test from comparing the distribution of contrasts to that of a normal distribution with mean 0 and standard deviation equal to the root of the mean of squared contrasts (the expected distribution of the contrasts under Brownian motion; see Felsenstein 1985; Rohlf 2001). We chose this to capture deviations from normality. For example, if traits evolved via a "jump-diffusion"-type process (Landis et al. 2013) in which there were occasional bursts of rapid phenotypic evolution (Pennell et al. 2013), the tip data would no longer be multivariate normal owing to a few contrasts throughout the tree being much larger than the rest (i.e., the distribution of contrasts would have heavy tails).


M_{SIG}
C_{VAR}
S_{VAR}
S_{ASR}
S_{HGT}
D_{CDF}
SLA
SeedMass
LeafN
Density
0.00 0.25 0.50 0.75 1.00
p−value

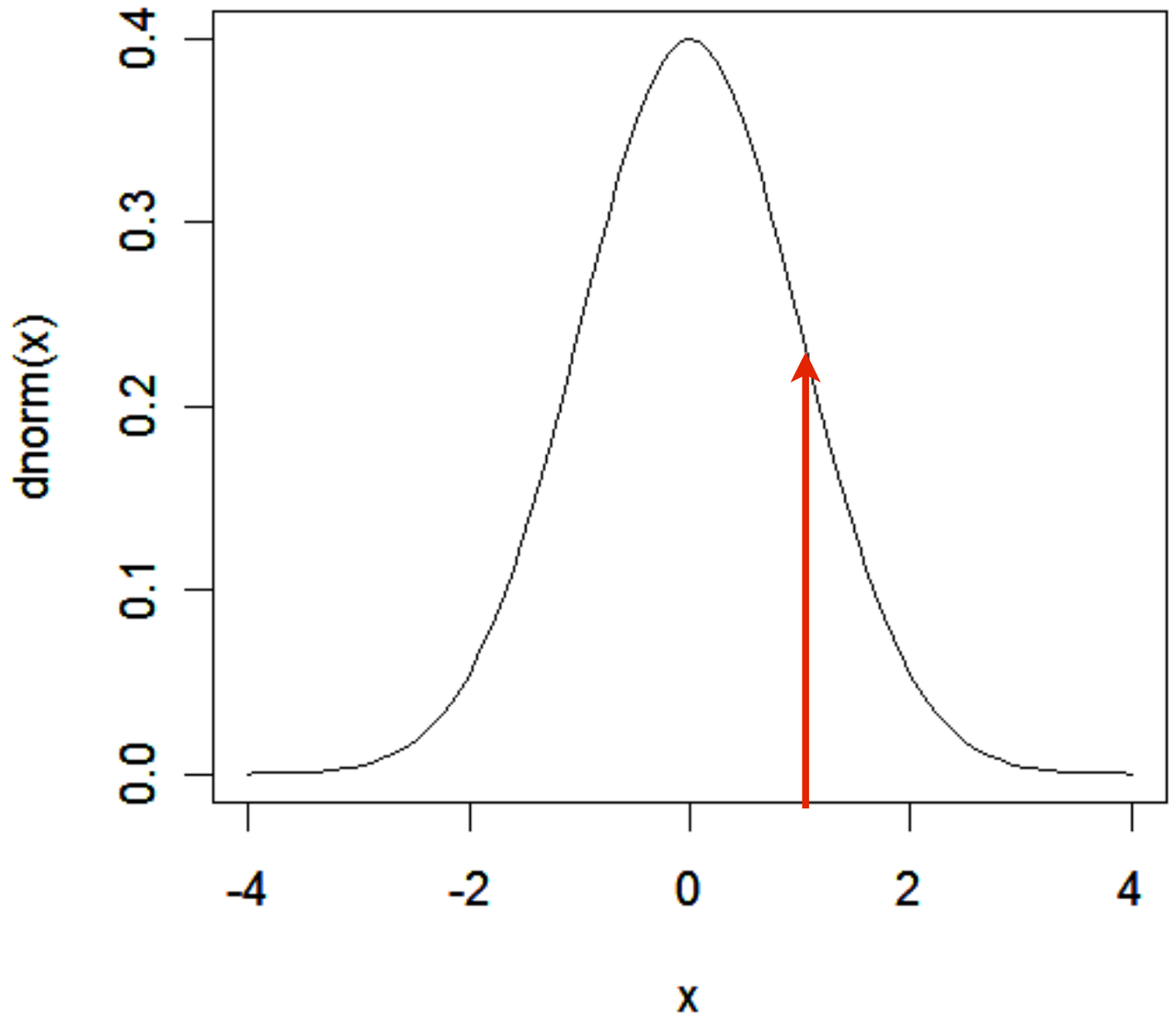
Likelihood for a single character

Brownian motion

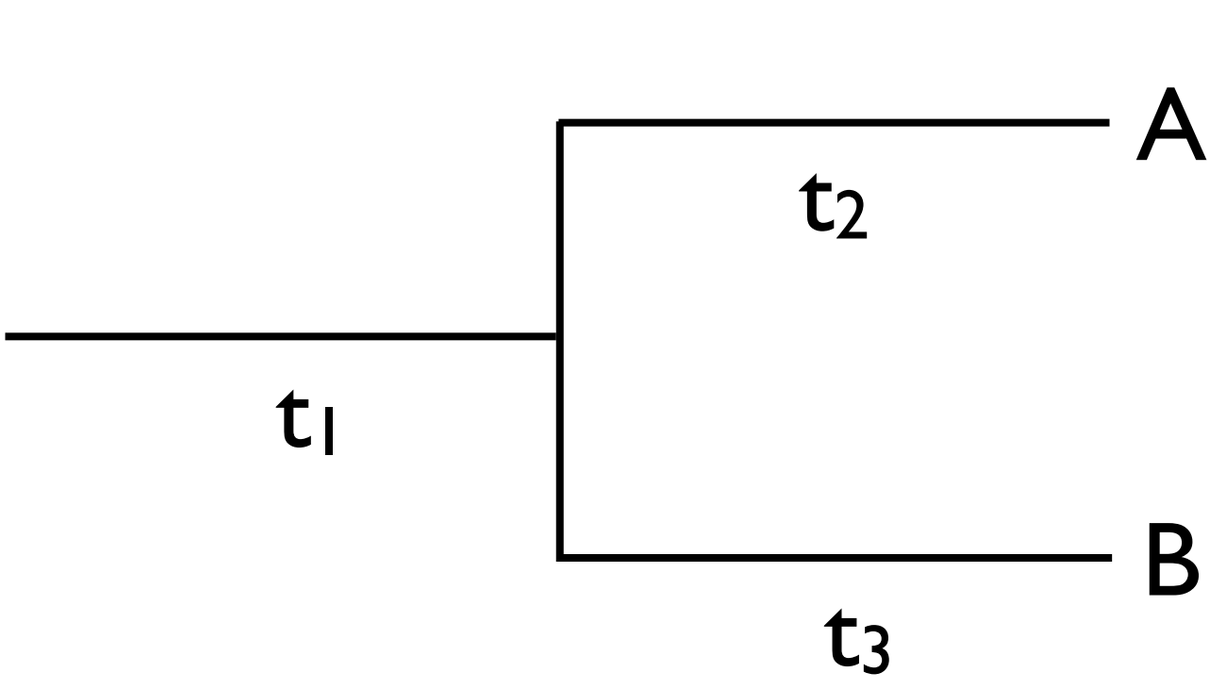


$$y \sim N(0, \sigma^2 * t)$$

$$\Pr[y = x] = \frac{1}{\sigma\sqrt{2\pi t}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2 t}\right)$$


dnorm(x)
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
x

Multivariate Normal



variance-covariance matrix

σ^2

t_1+t_2	t_1
t_1	t_1+t_3

C

$var(A) = \sigma^2(t_1+t_2)$ $cov(A,B) = \sigma^2(t_1)$

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

Two dimensions (x, y) correspond to tree with n=2

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - \frac{2\rho xy}{(\sigma_x\sigma_y)}\right)\right)$$

More dimensions gets more complicated

Easy to do with computers