

Phylogenetic Comparative Methods

Laura Salter Kubatko
Departments of Statistics and
Evolution, Ecology, and Organismal Biology
The Ohio State University
lkubatko@stat.ohio-state.edu

May 27, 2010

Contrasts

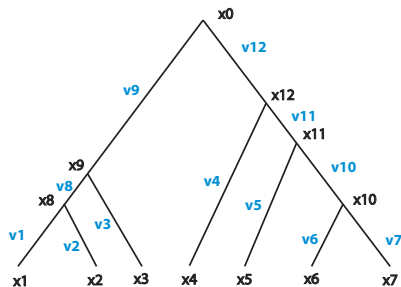
- ▶ Recall our discussion of **contrasts** – differences in trait values at tips
- ▶ It would be most desirable to have a set of these contrasts that were independent.
- ▶ Felsenstein (1985) showed how to construct these.
- ▶ Go back to our example from last time to see how we might do this.

Phylogenetic Independent Contrasts

- ▶ For any pair of sister taxa, consider the difference in their trait values.

▶ Example: $x_{1i} - x_{2i}$

$$\begin{aligned} \text{var}(x_{1i} - x_{2i}) &= \text{var}(x_{1i}) + \text{var}(x_{2i}) - 2\text{cov}(x_{1i}, x_{2i}) \\ &= \sigma_i^2(v_1 + v_8 + v_9) + \sigma_i^2(v_2 + v_8 + v_9) - 2\sigma_i^2(v_8 + v_9) \\ &= \sigma_i^2(v_1 + v_2) \end{aligned}$$



Phylogenetic Independent Contrasts

- ▶ Now we have one contrast – let's separate it from the remainder of tree by replacing this clade with a node whose trait value is a weighted average of x_{1i} and x_{2i}
- ▶ So we want to define $x'_8 = fx_{1i} + (1 - f)x_{2i}$ – how should we pick f ?
- ▶ Pick the value that gives independence of x'_8 and $x_{1i} - x_{2i}$ – find f to give $cov(x_{1i} - x_{2i}, fx_{1i} + (1 - f)x_{2i}) = 0$
- ▶ After a little work, we find $f = \frac{v_2}{v_1 + v_2}$



Phylogenetic Independent Contrasts

- ▶ Now we have two important results:
 - ▶ $cov(x'_8, x_{ji}) = cov(x_{1i}, x_{ji}) = cov(x_{2i}, x_{ji})$
 - ▶ $var(x'_8) = \sigma_i^2 \left(\frac{v_1 v_2}{v_1 + v_2} + v_8 + v_9 \right)$
 - ▶ Note the added branch length
- ▶ Continuing in this manner gives a set of phylogenetic independent contrasts – [R package demo](#)

Shortcomings of Brownian Motion Model

- ▶ Problem with BM model: biological realism
- ▶ Recall that
 - ▶ expected value of the trait at each tip is x_0 for all tips
 - ▶ Variance increases as a function of time
- ▶ As an example, we'll consider the paper:
Butler, M.A. and A.A. King. 2004. Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *American Naturalist* 164(6): 683-695.

Shortcomings of Brownian Motion Model

► Examples of some BM processes

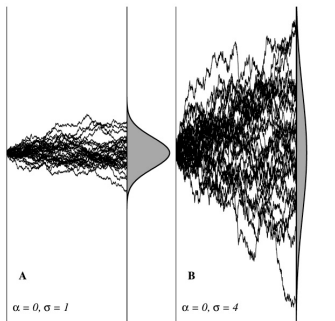
From Am Nat 164(6):683-695.

© 2004 by The University of Chicago.

For permission to reuse, contact journalpermissions@press.uchicago.edu.



CHICAGO JOURNALS



Shortcomings of Brownian Motion Model

- ▶ BM model omits two important evolutionary concepts
 - ▶ Selection: we expect traits that are more fit (=better adapted to the environment) to be maintained in the population, while traits that are less fit to die out
 - ▶ Fitness varies over
 - Time - e.g., ice age vs. present time
 - Lineages (branches in the tree) - e.g., speciation might happen due to available habitat
- ▶ Still want to include the process of genetic drift: random fluctuation in which traits are maintained or die out

The OU Model

- ▶ Goal: develop a phylogenetic model to incorporate both selection and drift
- ▶ OU Process: $dX_t = \alpha(\theta - X_t) + \sigma dW_t$
- ▶ Deterministic term: $\alpha(\theta - X_t)$
 - ▶ Used to model selection
 - ▶ α = strength of selection
 - ▶ θ = optimum trait value
 - ▶ Force of selection is proportional to the distance of current trait value from the optimum ($\theta - X_t$)

The OU Model

- ▶ Goal: develop a phylogenetic model to incorporate both selection and drift
- ▶ OU Process: $dX_t = \alpha(\theta - X_t) + \sigma dW_t$
- ▶ Diffusion term: σdW_t
 - ▶ Used to model random genetic drift
 - ▶ σ = intensity of drift

The OU Model

► Examples of some OU processes

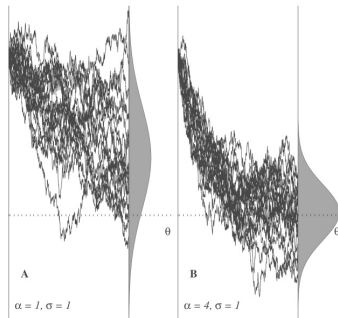
From Am Nat 164(6):683-695.

© 2004 by The University of Chicago.

For permission to reuse, contact journalpermissions@press.uchicago.edu.



CHICAGO JOURNALS



The OU Model

- ▶ Notes:
 - ▶ If $\alpha=0$, the model reduces to BM
 - ▶ Often called the Hansen model in phylogenetics
 - ▶ Assume the molecular clock (distance from tip to root is the same for all species)
- ▶ To apply the model, we need:
 - ▶ Data on the distribution of quantitative characters across species
 - ▶ A phylogeny with branch lengths relating the species of interest
 - ▶ Hypotheses regarding selective forces operating on each branch

The OU Model

▶ A two-taxon example

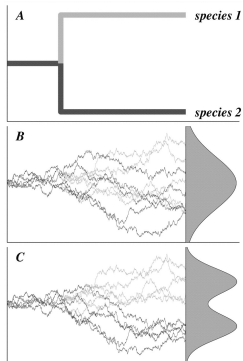
From Am Nat 164(6):683-695.

© 2004 by The University of Chicago.

For permission to reuse, contact journalpermissions@press.uchicago.edu.



CHICAGO JOURNALS



The OU Model

- ▶ Extensions beyond two taxa:
 - ▶ For an n taxon tree, there will be $2n - 2$ branches
 - ▶ Assume that there are a smaller number of selective optima, and assign these to branches.

- ▶ We'll look at an example: Anolis lizard evolution

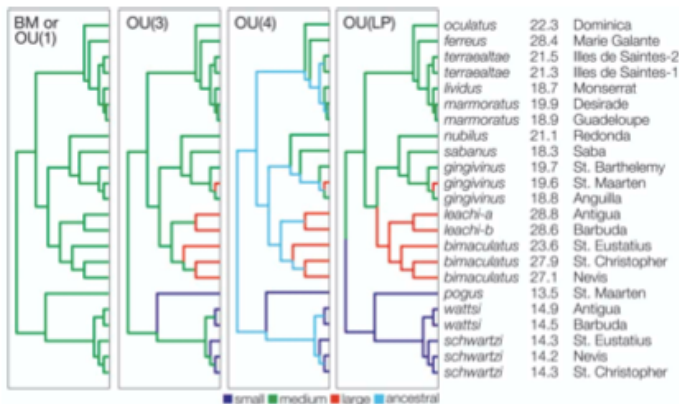
An Example: Anolis lizard evolution (Butler and King (2004))



- ▶ We'll look at variation in body size on the Lesser Antilles Islands
- ▶ Islands with two species: one large and one small
- ▶ Islands with one species: intermediate body size
- ▶ Hypothesis of interest is character displacement: did two initially intermediate-sized species come into contact and then evolve in opposite directions?

An Example: Anolis lizard evolution (Butler and King (2004))

- Models to be compared:



An Example: Anolis lizard evolution (Butler and King (2004))

► Results:

The American Naturalist, 164(6):683-695 - Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution.

5/21/09 12:45 AM

[Go Back](#) | Image of typeset table: **1****Table 1:** Performance of alternative models for body size evolution in the character displacement study

	BM	OU(1)	OU(3)	OU(4)	OU(LP)
$-2\log\mathcal{L}$	-34.66	-34.66	-40.21	-47.22	-49.69
AIC	-30.66	-26.66	-28.21	-33.22	-37.69
SIC	-28.39	-22.12	-21.40	-25.27	-30.88
LR		0	5.55	12.56	15.03
<i>P</i> value		1	.24	.028	.0046

Note: For each model, the likelihood values ($-2\log\mathcal{L}$), Akaike Information Criterion (AIC), likelihood ratio test value (LR), and associated *P* values are given. The LR tests were conducted for each model versus Brownian motion. Columns contain alternative evolutionary models. Model abbreviations are as follows: BM = Brownian motion, OU = Ornstein-Uhlenbeck with one (1), three (3), or four (4) optima or based on linear parsimony (LP) reconstruction of the colonization events (see text). For the OU(1) model, the best-fit

An Example: Anolis lizard evolution (Butler and King (2004))

► Results:

The American Naturalist, 164(6):683-695 - Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution.

5/21/09 12:46 AM

[Go Back](#) | Image of typeset table: **1****Table 2:** Parameters estimated for the five models comparing character displacement with alternative hypotheses

	BM	OU(1)	OU(3)	OU(4)	OU(LP)
α		0	.32	14.67	2.49
σ	.21	.21	.20	.47	.22
θ_0	2.95	2.95	3.99	... ^a	.86
θ_{small}		... ^a	-1.40	2.58	2.75
θ_{medium}			.18	3.11	3.24
θ_{large}			2.71	3.30	3.56
$\theta_{\text{ancestral}}$				2.83	

Note: The evolving character is log body size (mm). Model abbreviations are given in table 1. The model parameters (in rows) are as follows: α = strength of selection in OU models; σ =

Comparative Methods

- ▶ Some issues
 - ▶ Known phylogeny - never really true
 - ▶ Numerical methods to get MLEs
 - ▶ Model selection and hypothesis testing
 - ▶ Amount/type of data - sufficient to fit models?

- ▶ Concerns about performance of typical measures (e.g., BIC) and asymptotic properties - see Ane, 2008. *The Annals of Applied Statistics* 2(3): 1078-1102.