

Phylogenetic comparative methods

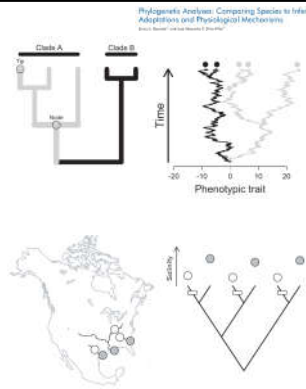
All three analyses indicate that theropod cranial form is significantly correlated with phylogeny. When subjected to the permutation test in MorphoJ, all four data sets are found to have significant phylogenetic signal (24-taxon data set: tree length = 0.1689822, $P < 0.001$; 26-taxon data set: tree length = 0.24750721, $P < 0.0001$; 31-taxon data set: tree length = 0.17546221, $P < 0.001$; 36-taxon data set: tree length = 0.26278840, $P < 0.0001$). MPVR reveals that form metrics (PC scores) exhibit significant correlations with phylogeny in all four data sets (all sample sizes, both with and without oviraptorosaurs) (Table S7). Finally, Blomberg's methods show that PC1 of form has a significant phylogenetic signal (Table S8) in all data sets, whereas PC2 of form usually is significantly correlated with phylogeny. We also note that MPVR and Blomberg's methods show that function is strongly correlated with phylogeny (see Tables S7, S8).

Table S8: Significance and strength of phylogenetic signal using Blomberg's K . K -statistic indicating strength of phylogenetic signal, observed variance of the PIC ($PIC_{var,obs}$), mean variance of PIC from the permutation ($PIC_{var,rand}$), Z -score of observed vs. random variance of PIC and p -value of observed vs. random variance of PIC (Kembel *et al.*, 2009) are shown for each variable on the various trees. Significance codes: 0.05, *, 0.01, **, 0.001, ***.

Dataset	K	$PIC_{var,obs}$	$PIC_{var,rand}$	Z-score	p-value
36-taxa					
PC1 _{form}	0.664	2.24×10^{-3}	9.37×10^{-4}	-2.74	$1.00 \times 10^{-41***}$
PC2 _{form}	0.468	1.47×10^{-3}	2.71×10^{-4}	-2.08	$4.10 \times 10^{-37**}$
PC3 _{form}	0.297	7.48×10^{-3}	1.01×10^{-3}	-1.00	0.125
PC4 _{form}	0.165	9.76×10^{-3}	8.54×10^{-3}	0.545	0.766
PC5 _{form}	0.302	4.65×10^{-3}	7.16×10^{-3}	-1.41	$3.56 \times 10^{-2*}$

Independent Contrasts

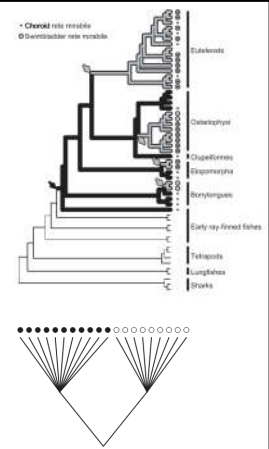
- Any question that involves comparisons among multiple species should be done in a phylogenetic context.
- For most traits, expectation is that phenotypic differences will correlate with phylogenetic differences.



"Nothing in biology makes sense except in the light of evolution"
T. Dobzhansky, 1973

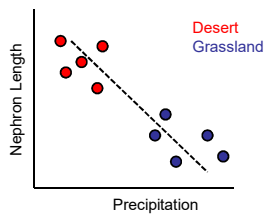
Independent Contrasts

- Evolution of the rete mirabile associated with swimbladder. Occurred 4 times independently, sample size is thus 4.
- In this other example, a trait changed once. Sample size is 1 regardless of how many species sampled.



Independent Contrasts

- Two habitats: desert and grassland
- Within each habitat you sample 5 species, measure a response variable related to precipitation (nephron length).
- Hypothesis:** nephron length evolves in response to selection pressures in desert environments.



Independent Contrasts

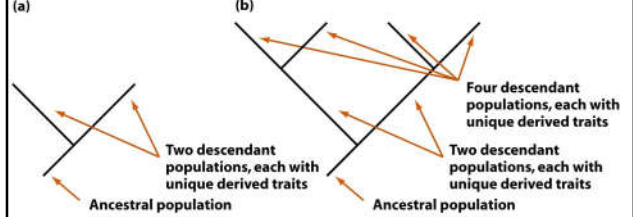
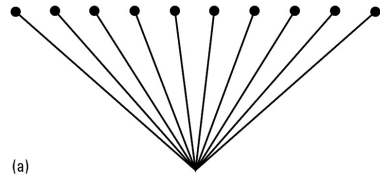


Figure 4-2. Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

- Hypothesis assumes statistical independence for each species. For each species, desert conditions selected for longer nephrons independently (and vice versa).
- We know species inherit traits from common ancestors (phylogenetic constraints).
- Assuming we know the "true" phylogenetic history of your 10 species, how do we test for or correct for this?

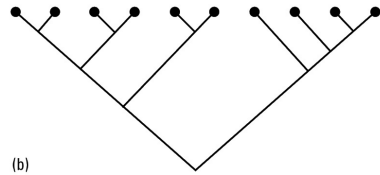
Independent Contrasts

- 10 species, all derived from one common ancestor. No phylogenetic information.

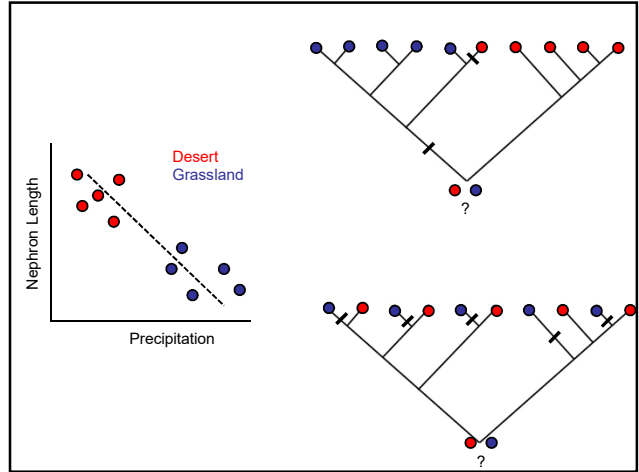


(a)

- Phylogeny known. Is trait of interest evolved independently?



(b)



PHYLOGENIES AND THE COMPARATIVE METHOD

JOSEPH Felsenstein
Department of Genetics SK-30, University of Washington, Seattle, Washington 98195
Vol. 125, No. 1 The American Naturalist January 1985

COMPARATIVE METHOD

FIG. 8.—An example of a phylogeny, assumed known, from which we can define independent contrasts between taxa. This tree is highly symmetric, so that $v_1 = v_2 = v_3 = v_4 = v_5 = v_6 = v_7 = v_8 = v_9 = v_{10} = v_{11} = v_{12}$, and $v_{13} = v_{14}$.

TABLE 1
THE FOUR CONTRASTS EXTRACTED FROM THE PHYLOGENY SHOWN IN FIGURE 8, EACH WITH ITS VARIANCE, AS COMPUTED USING STEPS 1-4 IN THE TEXT

CONTRAST	VARIANCE
$X_1 - X_2$	$v_1 + v_2$
$X_3 - X_4$	$v_3 + v_4$
$X_5 - X_6$	$v_5 + v_6$
$X_7 - X_8$	$v_7 + v_8$

where

$$X_9 = \frac{v_1 X_1 + v_2 X_2}{v_1 + v_2}$$

$$v_9 = \frac{v_1 v_2}{v_1 + v_2}$$

$$X_{10} = \frac{v_3 X_3 + v_4 X_4}{v_3 + v_4}$$

$$v_{10} = \frac{v_3 v_4}{v_3 + v_4}$$

$$X_{11} = \frac{v_5 X_5 + v_6 X_6}{v_5 + v_6}$$

$$v_{11} = \frac{v_5 v_6}{v_5 + v_6}$$

$$X_{12} = \frac{v_7 X_7 + v_8 X_8}{v_7 + v_8}$$

$$v_{12} = \frac{v_7 v_8}{v_7 + v_8}$$

FIG. 9.—A less symmetrical phylogeny. The independent contrasts for this phylogeny are given in table 1.

Packages

- Analyses of Phylogenetics and Evolution
 - Reading/writing phylogenetic trees
 - Manipulating phylogenetic
 - Reading/writing DNA sequences
 - Computing distances from DNA sequences
 - Phylogenetic comparative methods
 - Estimation of ancestral characters
 - Analysis of diversification
- <http://ape.mpl.ird.fr/>
- Geiger
- Picante
- Ade4
- phytools

Independent Contrasts

- Function `pic` (*ape* package)
 - `pic(x, phylo)`
 - X is the trait of interest
 - Phylo is a phylogeny
 - The names in x must match the names in the phylogeny
- Output is a series of independent contrasts of x as outlined in Felsenstein(1985).

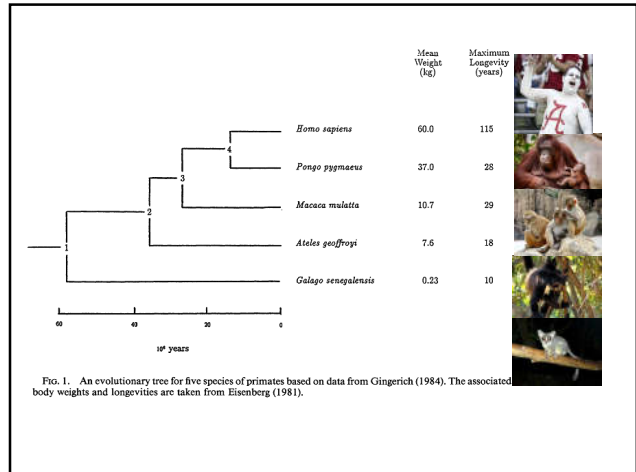
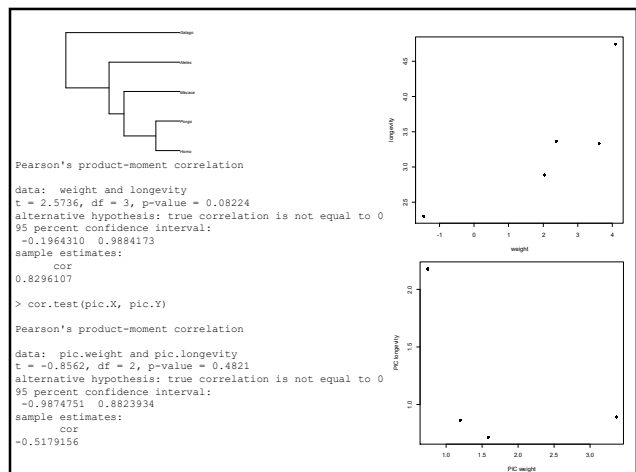
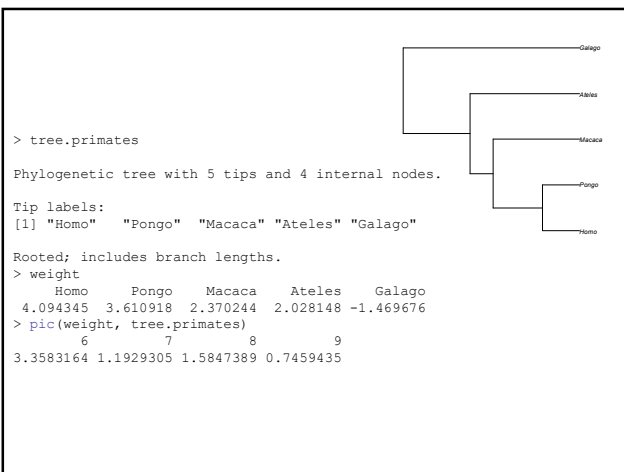


FIG. 1. An evolutionary tree for five species of primates based on data from Gingerich (1984). The associated body weights and longevities are taken from Eisenberg (1981).



Phylogenetic Signal

- Given a phylogenetic tree (topology and branch lengths), do closely related species tend to be more similar?
- If so, variance in independent contrasts will be low
- Permuting the tips of the trees and recalculating variance provides a test of significance:

$$K = \frac{\text{observed MSE}_{ij}}{\text{expected MSE}_{ij}}$$

- $K < 1$ – no phylogenetic signal
- $K > 1$ – close relatives more similar than expected at random

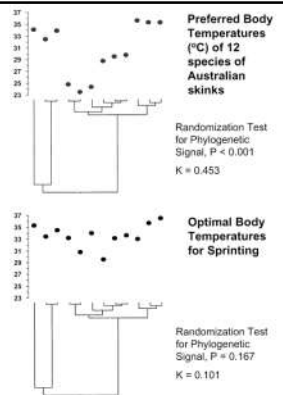


FIG. 4. Example in which significant phylogenetic signal is detected for one trait but not for another. (A) Preferred body temperature of some Australian skinks ($P < 0.001$, $K = 0.453$). (B) Optimal body temperature for sprinting ($P = 0.167$, $K = 0.101$). Data and tree are from Huey and Bennett (1987) and Garland et al. (1991), respectively.

TESTING FOR PHYLOGENETIC SIGNAL IN COMPARATIVE DATA: BEHAVIORAL TRAITS ARE MORE LABILE
 SIMON P. BLANDIN¹, THOMAS GALANTO, JR.^{2,3} and ANDREW R. TRAY^{1*}
¹Department of Biology, University of California, Riverside, California 92521
²Department of Biology, University of Illinois, Urbana, Illinois 61801
³Department of Biology, University of Illinois, Urbana, Illinois 61801

Phylogenetic signal

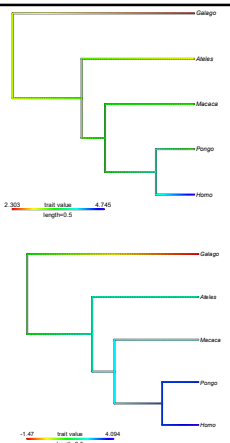
- Function `phylosignal` (picante package)
 - `Phylosignal(x, phylo, reps, checkdata=TRUE)`
 - X is the trait of interest
 - Phylo is a phylogeny
 - Reps= number of permutations
 - Checkdata=TRUE will check that the phylogeny names match the trait names, otherwise the function assumes the order is the same
- Returns measure of K (Bloomberg et al. 2003), observed variation in PIC, randomized variation in PIC and probability.

Phylogenetic Signal

- Phylogenetic signal in longevity
- Phylogenetic signal in weight

K PIC.variance_obs PIC.variance_rnd.mean PIC.variance.P PIC.variance.S
 1 0.7045102 2.27495 2.354906 0.5624376

K PIC.variance_obs PIC.variance_rnd.mean PIC.variance.P PIC.variance.S
 1 1.911768 5.2564 13.86626 0.00693007 -1.338417

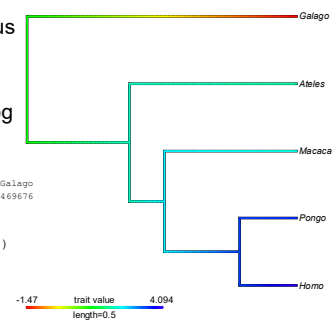


Mapping traits onto a phylogeny

- Function `contMap` (phytools package) will color a phylogeny by a continuous trait.
- Note that these are the log transformed values:

```
> weight
  Homo Pongo Macaca Ateles Galago
  4.094345 3.610918 2.370244 2.028148 -1.469676
```

```
contMap(tree.primates, weight)
```



Phylogenetic Signal

- Pagels λ
 - Scale the internal branches of the tree

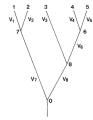
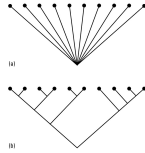


Fig. 3—A less symmetrical phylogeny. The independent contrasts for this phylogeny are given in table 1.

- $\lambda = 0$, "star tree" and no signal
- $\lambda = 1$, full branch lengths best fit



Inferring the historical patterns of biological evolution

Mark Pagel
School of Animal and Molecular Sciences, University of Reading, Whiteknights, Reading RG6 2AL, UK
MARTIN.PAGEL@Rdg.AC.UK | 28 OCTOBER 1999 | www.nature.com

Pagel, M. 1999 Inferring the historical patterns of biological evolution. *Nature* **401**, 877-884.

Phylogenetic Signal

- Function `fitContinuous` (geiger package)
 - `fitContinuous(tree.primates, longevity, model="lambda")`
- Requires tree, trait of interest and specify the model as "lambda"
- Be careful that names in phylogeny match names associated with trait
- Can pass multiple traits (or a whole matrix) but each one is evaluated separately
- Returns
 - Parameters of fitting model to the data
 - Log likelihood
 - AIC
 - Model complexity
 - λ

Longevity

```
> fitContinuous(tree.primates, longevity,
model="lambda")
GEIGER-fitted comparative model of continuous data
fitted 'lambda' model parameters:
  lambda = 0.000000
  sigma = 0.6504503
  s0 = 3.327478
model summary:
log-likelihood = -6.019467
AIC = 18.038933
AICc = 42.038933
free parameters = 3
Convergence diagnostics:
  optimization iterations = 100
  failed iterations = 0
  frequency of best fit = 0.79
object summary:
  'lik' -- likelihood function
  'bnd' -- bounds for likelihood search
  'res' -- optimization iteration summary
  'opt' -- maximum likelihood parameter
> lambdaFit$opt
$lambda
[1] 7.34599e-209
$sigma
[1] 0.6504503
$s0
[1] 3.327478
$lnL
[1] -6.019467
$method
[1] "L-BFGS-B"
$K
[1] 3
$aic
[1] 18.03893
$aicc
[1] 42.03893
```

Weight

```
> fitContinuous(tree.primates, weight,
model="lambda")
GEIGER-fitted comparative model of continuous data
fitted 'lambda' model parameters:
  lambda = 1.000000
  sigma = 3.153841
  s0 = 1.183727
model summary:
log-likelihood = -9.194399
AIC = 24.388797
AICc = 48.388797
free parameters = 3
Convergence diagnostics:
  optimization iterations = 100
  failed iterations = 0
  frequency of best fit = 0.13
object summary:
  'lik' -- likelihood function
  'bnd' -- bounds for likelihood search
  'res' -- optimization iteration summary
  'opt' -- maximum likelihood parameter
> lambdaFit$opt
$lambda
[1] 1
$sigma
[1] 3.153841
$s0
[1] 1.183727
$lnL
[1] -9.194399
$method
[1] "L-BFGS-B"
$K
[1] 3
$aic
[1] 24.3888
$aicc
[1] 48.3888
```

Models of Evolution

- Phylogenetic signal indicates closely related species will be more similar than expected at random.
- How do traits change?

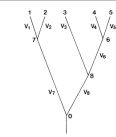
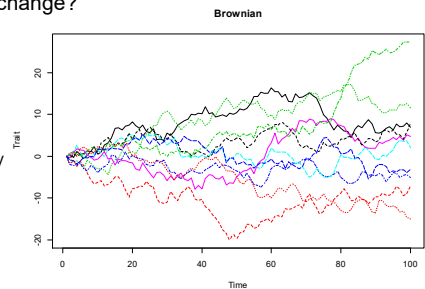


Fig. 3—A less symmetrical phylogeny. The independent contrasts for this phylogeny are given in table 1.

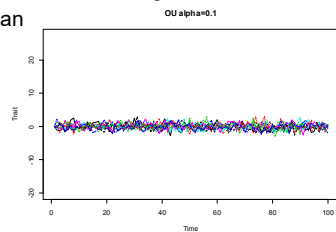
Brownian motion: traits change randomly.

Differences among species should be tightly correlated with phylogenetic distance (time since divergence)



Models of Evolution

- Ornstein-Uhlenbeck model is Brownian motion with an attraction or tendency to return to the mean.
- In evolutionary terms, this would represent random variation around but not away from an optima.
- Parameter alpha represents the strength of attraction.
 - Alpha=0 is Brownian



Models of Evolution

- Function `fitContinuous` will evaluate multiple models of evolution for a trait of series of traits (model="lambda" etc.)
 - Default is Brownian Motion (leave out model=)
 - "lambda" – Pagel's lambda
 - "ou" - Ornstein-Uhlenbeck
 - "white" – white noise (all species from random normal distribution with no phylogenetic signal)
 - Others...see help file
- Similar model comparison logic we used for AIC model selection can be used to evaluate these models – **which is the best predictor of the observed data?**

For longevity

Brownian, lambda, ou, white noise delta AIC:

```
[1] 2.162422 2.000013 1.999927 0.000000
```

For weight

Brownian, lambda, ou, white noise delta AIC:

```
[1] 0.00000 2.00000 2.00000 2.49594
```

What is the conclusion?

Is this consistent with the other analyses?

Assignment

- Reading
 - FYI: Felsenstein, J. 1985 Phylogenies and the Comparative Method. *The American Naturalist* 125, 1–15.
 - Blomberg, S. P., Garland Jr, T. & Ives, A. R. 2003 Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57, 717–745.
 - Collar, D. C., O'Meara, B. C., Wainright, P. C. & Near, T. J. 2009 Piscivory limits diversification of feeding morphology in Centrarchid fishes. *Evolution* 63, 1557–1573.

Assignment

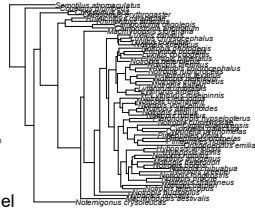
- Starter script and phylogenetic tree:
 - Loads tree based on mtDNA data (652 bp COI) for 63 species

- Trait data: egg diameter for 63 species (minnow_data.csv)

- Coburn, M. M. 1986 Egg diameter variation in eastern North American minnows (Pisces: Cyprinidae): correlation with vertebral number, Habitat, and spawning behavior. Ohio Journal of Science 86, 110-120.

- Is there a phylogenetic signal? Which model is the best fit?

- Does egg size differ among habitats?



Fish Dataset

Egg diameter, vertebral counts, habitat, and spawning behavior for some eastern North American cyprinids.

Name	Female sample size	Egg diameter (mm)	Source	Average mean diameter (mm)	Mean vertebral count	Source	Habitat	Spawning behavior
<i>Campostoma anomalum</i>	1	1.3-1.5	Becker 1983	1.7	41.5	Present Study	U-M**	3***
<i>Campostoma oligolepis</i>	*	2.0	Beck 1958		*		U-M	3
<i>Coxius plumbeus</i>	*	1.29-1.67	Brown & Haunmer 1970	1.48	40.5	Scott & Grossman 1973	L	2
<i>Eriymba buccata</i>	3	1.6	Becker 1983					
<i>Eriymba buccata</i>	107	0.75	Hoyt 1971	0.78	34.7	Present Study	L-U	1
<i>Eriymba buccata</i>	20	0.78	Wallace 1973					
<i>Eriymba buccata</i>	6	0.82	Present Study					

species	egg	habitat	GB
<i>Campostoma anomalum</i>	1.7	U-M	JN024821:
<i>Campostoma oligolepis</i>	1.4	U-M	JN024834:
<i>Coxius plumbeus</i>	1.48	L	EU524531:
<i>Cyprinella galactura</i>	1.6	U-U	JN025219:
<i>Cyprinella lutrensis</i>	0.9	L-U	EU751769:
<i>Cyprinella pyrrhomelas</i>	1.39	L-U	JN025246:
<i>Eriymba buccata</i>	0.78	L-U	JN027477:
<i>Hybopsis amblops</i>	0.96	L-U	JN026798:
<i>Hybopsis amnis</i>	0.8	L	JN026800:
<i>Hybopsis dorsalis</i>	0.91	L-U	JN027526:
<i>Luxilus chrysocephalus</i>	1.24	L-U	EU524764:
<i>Luxilus coccogenis</i>	1.41	U-M	JN027401:
<i>Luxilus cornutus</i>	1.25	U-M	EU524781:
<i>Luxilus cornutus</i>	1.37	U-M	JN027112:
<i>Lythrurus airdens</i>	0.92	L-U	JN027118:

