Phylogenetic Comparative Methods: I

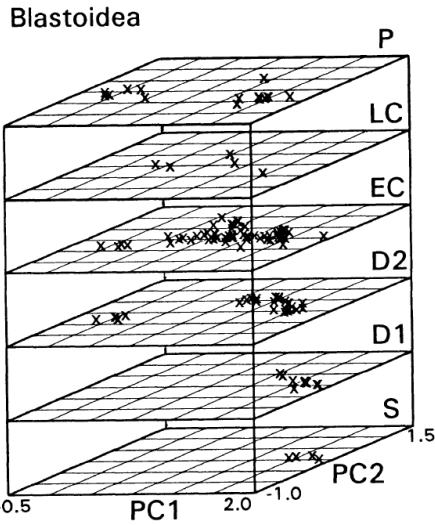# Macroevolutionary patterns of phenotypic diversity through time

Blastoidea

FIGURE 2. Temporal pattern of morphospace occupation in Blastoidea. Note overall increase in range of morphospace occupied. Early Carboniferous taxonomic di-
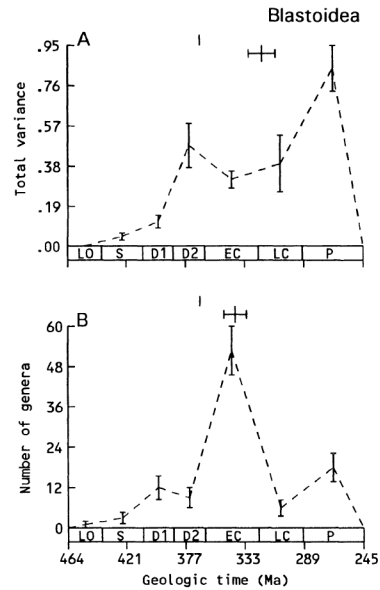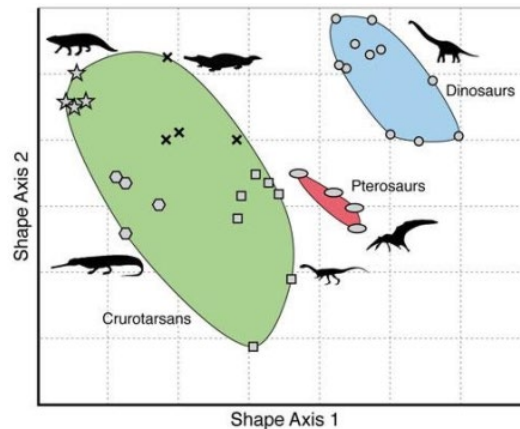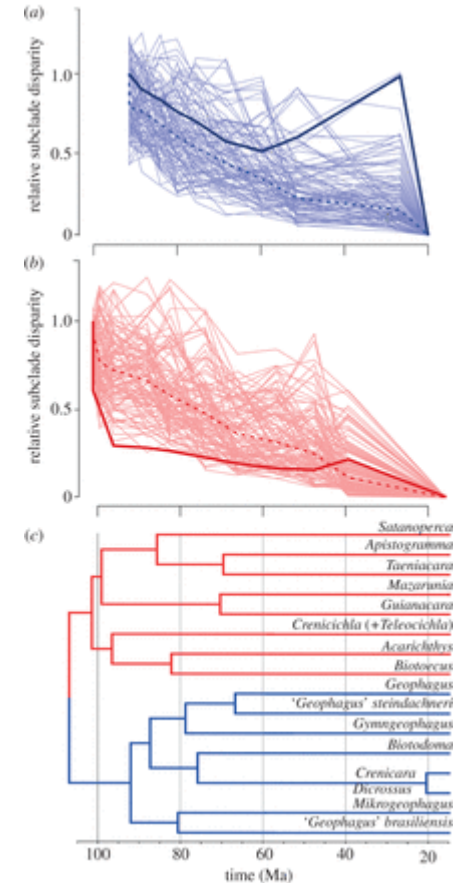
Foote (1993)

Blastoidea

FIGURE 3. Comparison of morphological and taxonomic diversity in Blastoidea. Morphological diversity in this
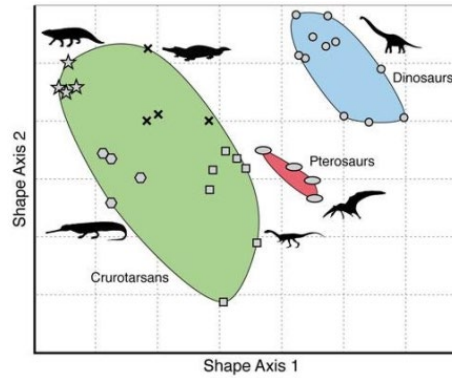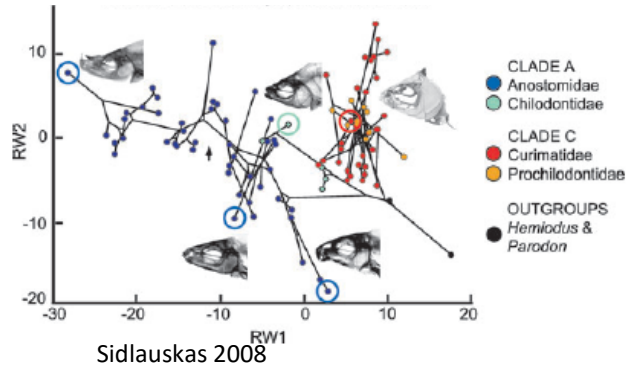

Brussatte et al. (2008)

Arbour and López-Fernández (2013)

Accumulation of Phenotypic Diversity

Morphological diversity

Sidlauskas 2008

Brussatte et al. 2008

Tempo: the pace of evolution

Mode: the manner/direction of evolution
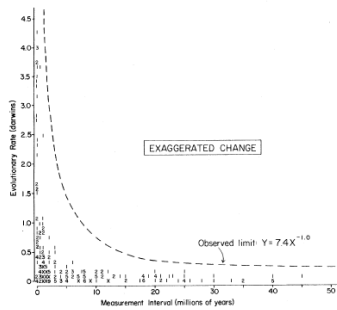
Gingerich 1983

Uyeda et al. 2011
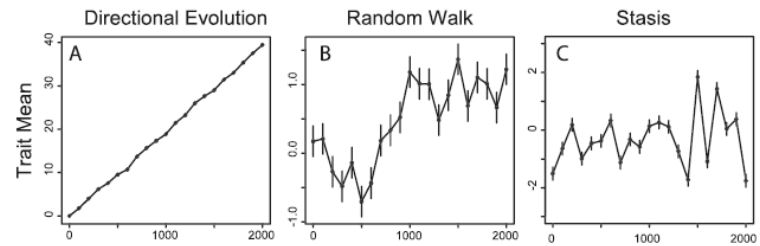
Hunt 2012

•How do we characterize patterns, and hypothesize processes?

Comparative Evolutionary Biology

One way: examine trait correlations across species

Hemmingsen (1960)

Primate data

Harvey & Clutton-Brock (1981)
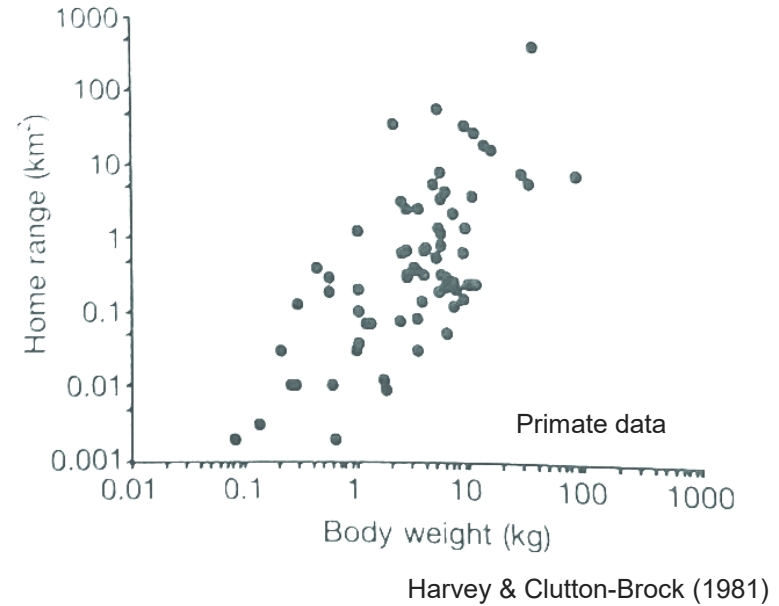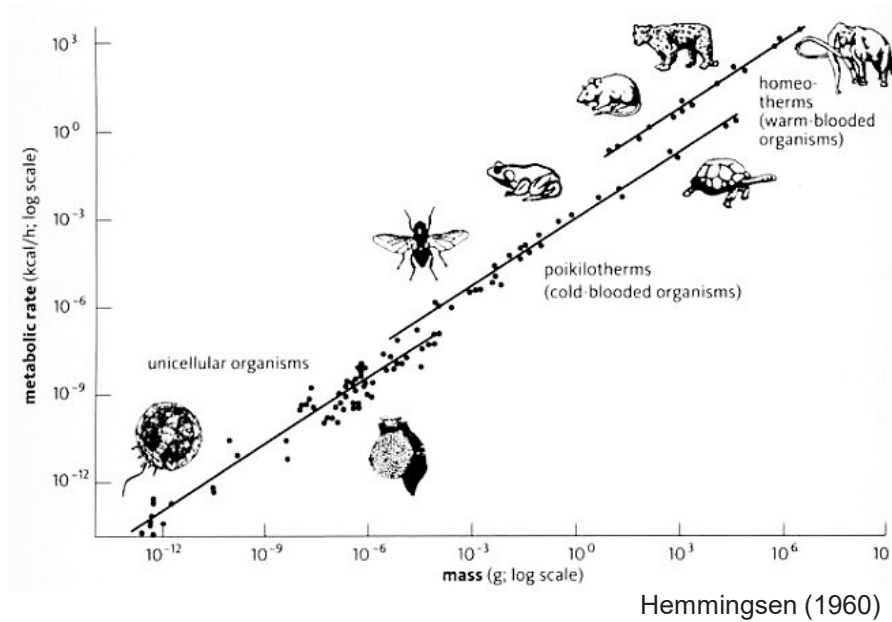
Can provide evidence of adaptation
 & coevolution

Unfortunately such comparisons often lead us astray

Comparative Evolutionary Biology

Species-level analyses (or analyses of higher-taxonomic groups) ignore evolutionary history

An impressive pattern, but…
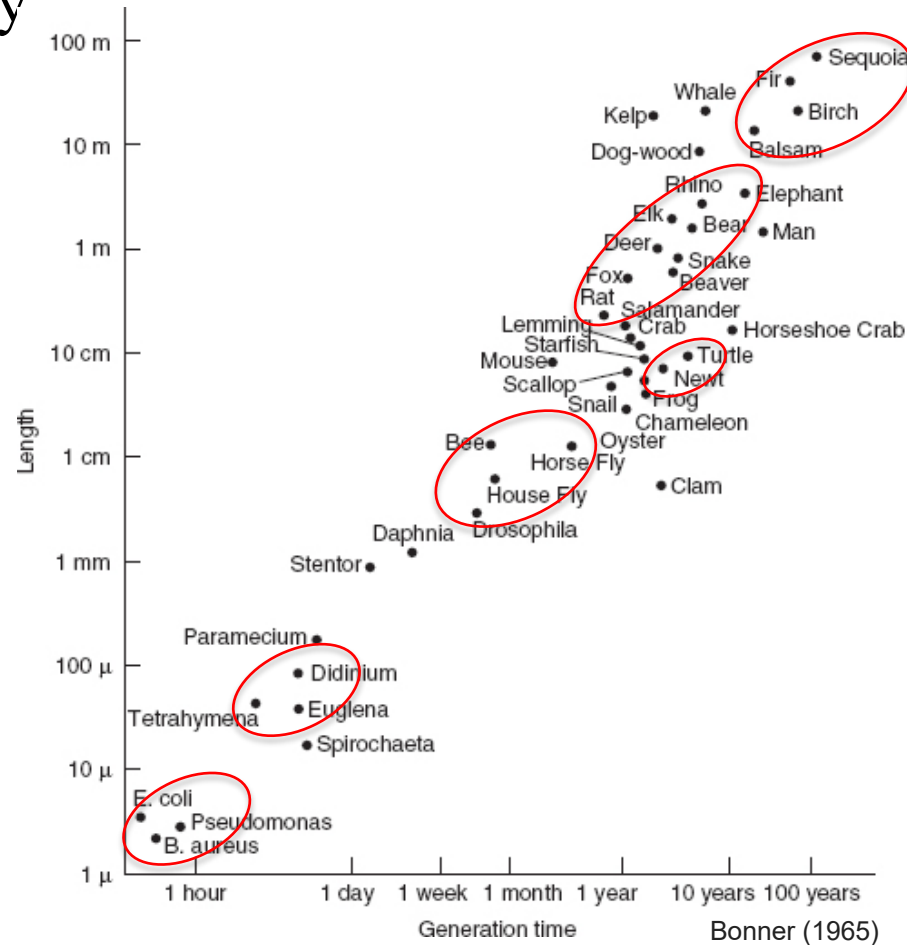
Taxa in circles are part of larger evolutionary lineages

Bonner (1965)

How many 'independent' data points do we have here???
This can make an enormous difference in inference and interpretation!

Say we observe the following pattern

Species	Behavior	Coloration
1	Solitary	Cryptic
2	Solitary	Cryptic
3	Solitary	Cryptic
4	Solitary	Cryptic
5	Solitary	Cryptic
6	Solitary	Cryptic
7	Solitary	Cryptic
8	Solitary	Cryptic
9	Gregarious	Aposematic
10	Gregarious	Aposematic
11	Gregarious	Aposematic
12	Gregarious	Aposematic
13	Gregarious	Aposematic
14	Gregarious	Aposematic
15	Gregarious	Aposematic
16	Gregarious	Aposematic

	Solitary	Gregarious
Cryptic	8	0
Aposematic	0	8

$X^2 = 12.25$; $P = 0.0004$

Pretty impressive association!

Inspired by data of
Sillen-Tullberg (1988)

Differing Evolutionary Interpretations

But what if evolutionarily we have:

So

So

So

So

So

So

So

So

Gr

Gr

Gr

Gr

Gr

Gr

Gr

Gr

Apo

Gr
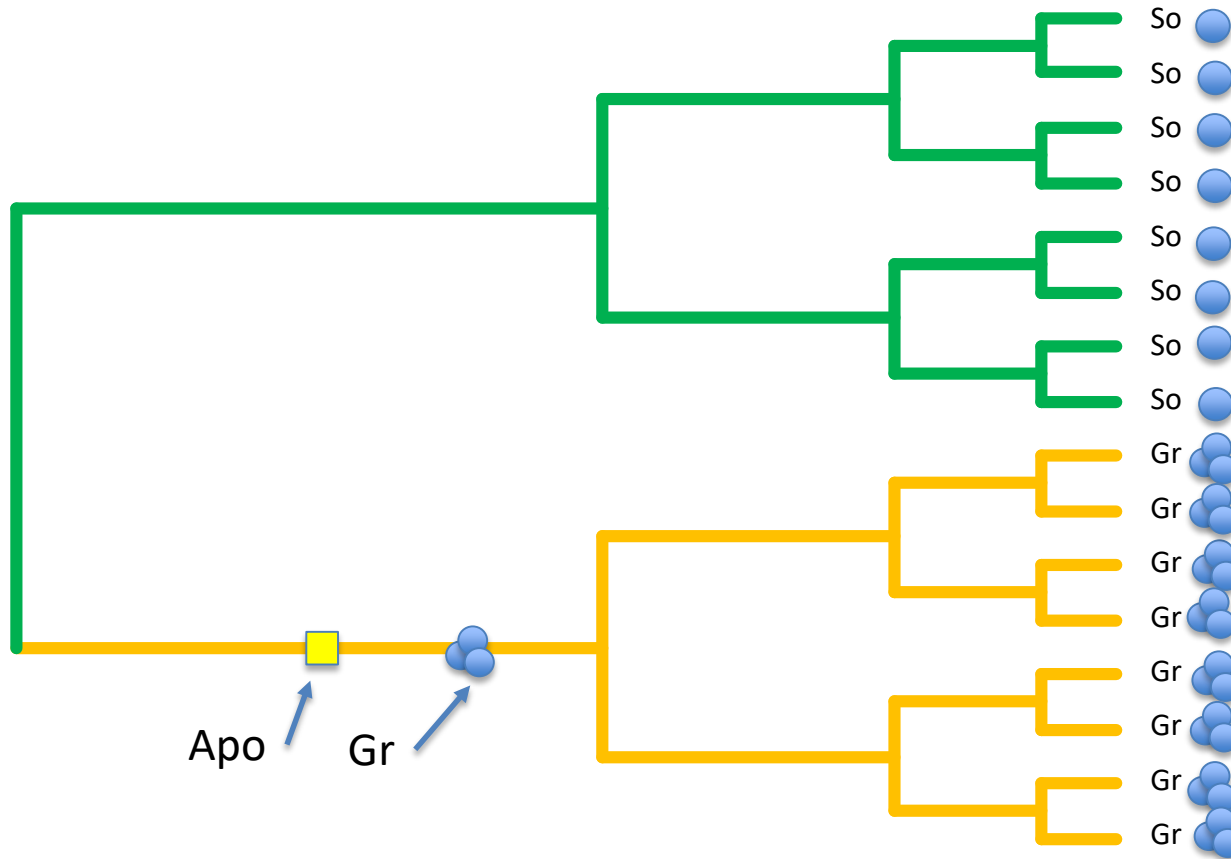
ONE evolutionary change can explain pattern

Not so impressive any longer!

Differing Evolutionary Interpretations

Or what if evolutionarily we have:
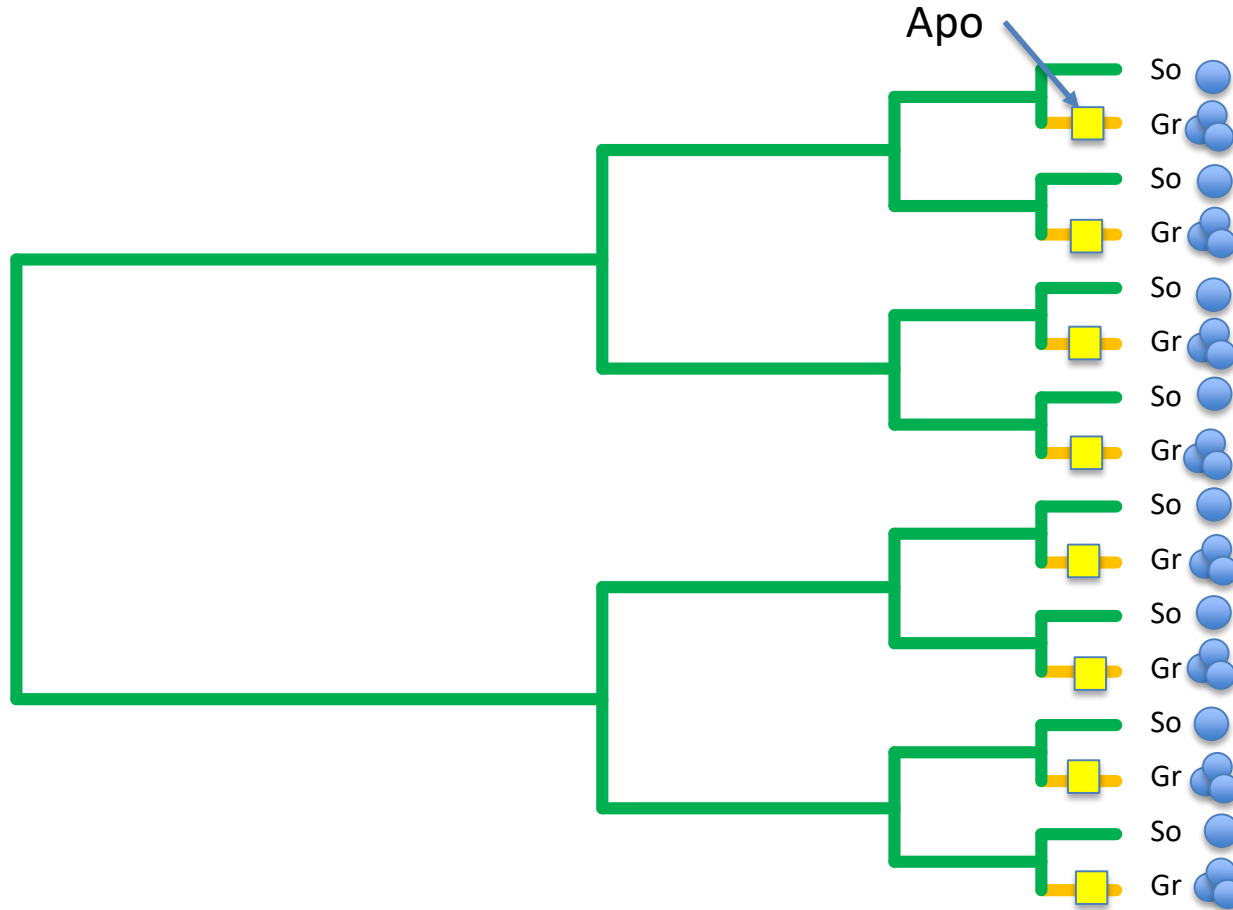
This is pretty impressive!

Eight correlated evolutionary changes of both traits!

Or what if we observe this pattern:

$r = 0.915; P > 0.0001$
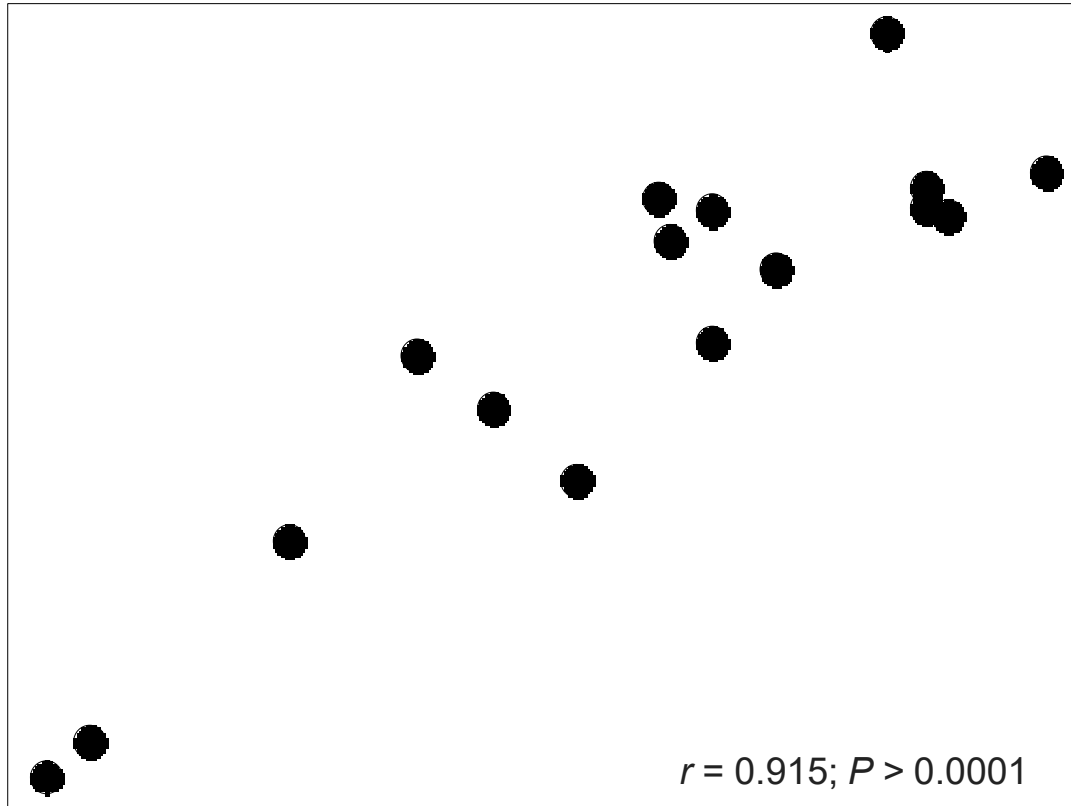
An impressive association!

But what if evolutionarily we have:

Not so impressive, closely related species are similar!

Traits seem to change with phylogeny

Differing Evolutionary Interpretations

Or what if evolutionarily we have:

Still pretty impressive!
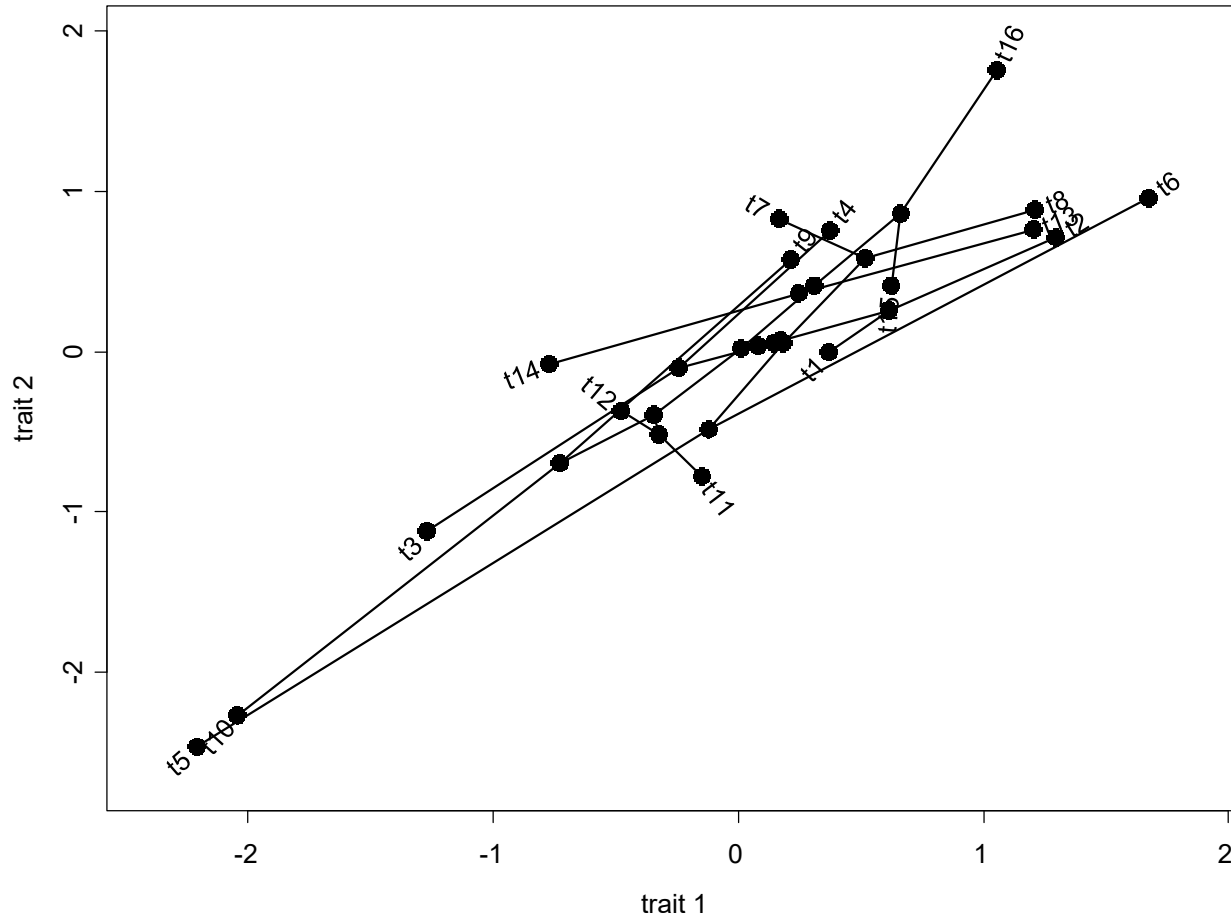
Lots of correlated evolutionary change in both traits

The point:

-Taxa are not independent

-Ignore evolutionary history AT ONE'S PERIL!
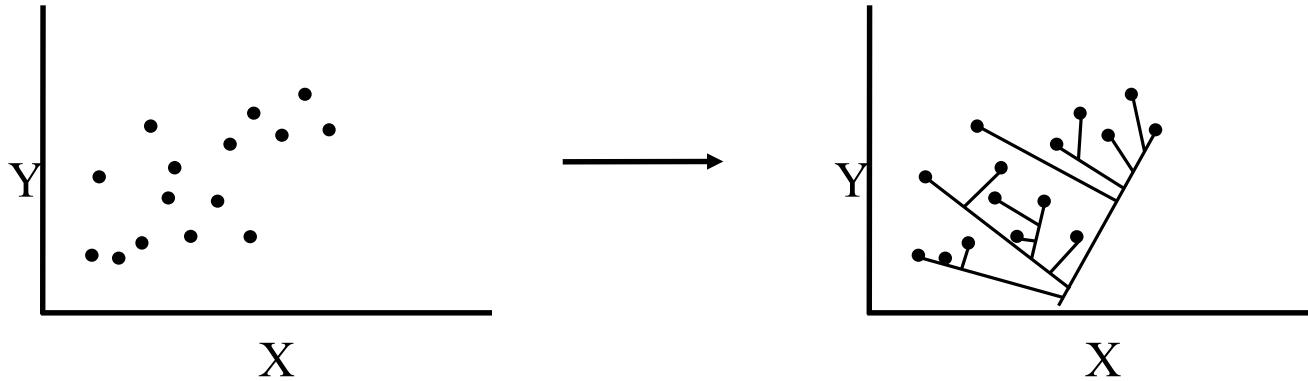
Phylogenetic comparative methods _condition_ the data on the phylogeny to account for lack of independence during the analysis

Phylogenetic Comparative Methods

An Incomplete History

2000s: Maturation phase
Synthesis: PIC, PGLS, Phylo-transform
Complex model comparison (BM1, BMM, OU1, OMM)
Bayesian methods
Parameter-shift methods (e.g., MEDUSA, BAMM)
Discrete diversification associations (BiSSE family)

1985: The Breakthrough
Phylogenetic Independent
Contrasts (PIC)

2010s: Multivariate approaches
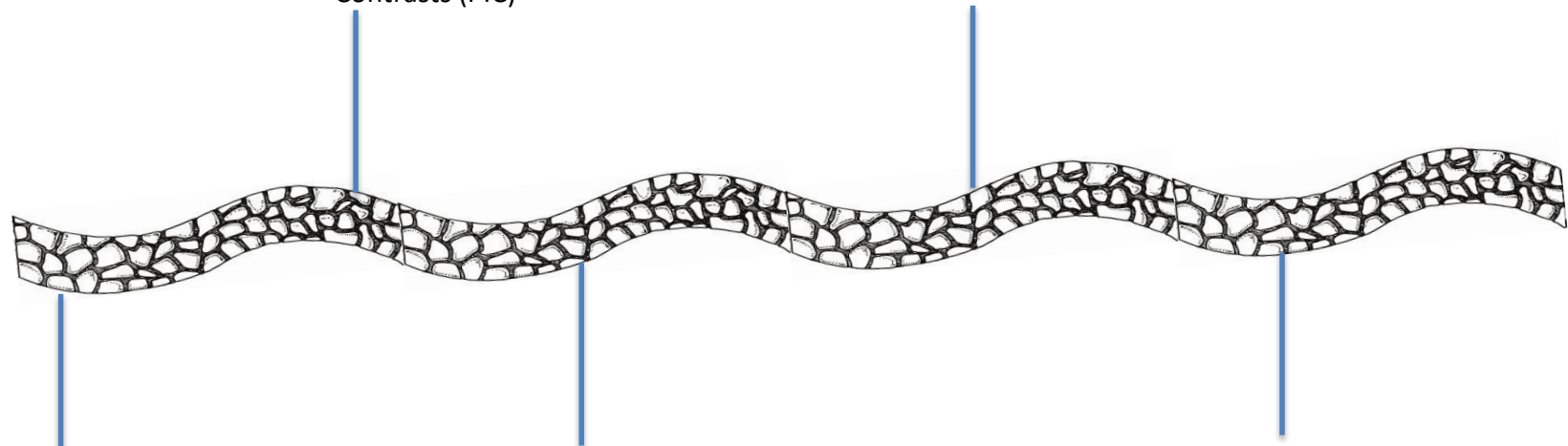
70s – early 80s: early attempts
Nested ANOVA
Phylogenetic autocorrelation
Discrete change correlations

80s – 90s: 'niche expansion'
PGLS
Phylogenetic signal (λ, K)
Phylogenetic ANOVA
Evolutionary models (BM1, OU1, ACDC, λ)
Diversity plots (LTT & DTT)
Diversification rates
Discrete trait change models

TODAY: PCMs: A diverse toolkit for evaluating evolutionary hypotheses

Testing association of discrete traits without phylogeny is problematic

Species	Behavior		Coloration	
1	Solitary	🔵	Cryptic	🟩
2	Solitary	🔵	Cryptic	🟩
...				
15	Gregarious	🔵🔵🔵	Aposematic	🟨
16	Gregarious	🔵🔵🔵	Aposematic	🟨

Tally of co-occurrence of states ignores number of *INDEPENDENT EVOLUTIONARY TRANSITIONS*

	Solitary	Gregarious
Cryptic	8	0
Aposematic	0	8

$X^2 = 12.25$; $P = 0.0004$
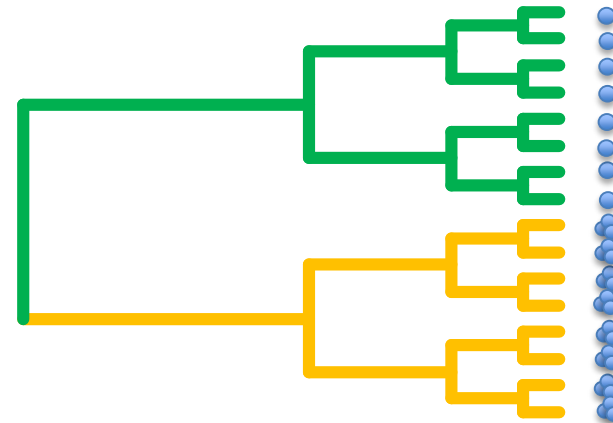
But what if association of traits not independent of phylogeny?

Discrete Data: Counting Evolutionary Events

Identify the number of evolutionary *changes* in each trait

-Associate changes with one another

Procedure:

-Map traits to phylogeny

-Estimate ancestral states*

-Tally transitions in each trait

-X^2 test for significance

Apo Gr
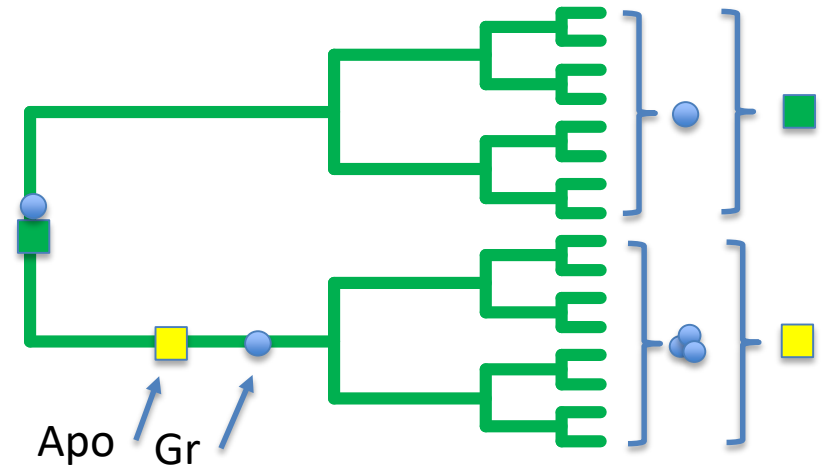
NOTE: method counts transitions, but not direction of change

	Solitary	Gregarious
Cryptic	1	0
Aposematic	0	1

$X^2 = 0.00$; $P = 1.00$

Ridley (1983)

Discrete Data: Counting Evolutionary Events

Another example:

	Solitary	Gregarious
Cryptic	8	0
Aposematic	0	8

$X^2 = 12.25$; $P = 0.0004$

Problem: counting evolutionary events (ala Ridley 1983) does not consider the distribution of changes in one trait *relative* to changes in the other

CCT: Evaluates whether trait change is concentrated in specific regions of the phylogeny
  -This would happen if changes in one character facilitate changes in another

This allows testing of *directional* hypotheses:
    "Does aposematism facilitate social (gregarious) behavior?"

    -Motivation: if one state (e.g., aposematism) is common, then we expect gregariousness to evolve in its presence more often just by chance. Thus we must account for it.

Maddison (1990)

H_0: Gains/losses in trait 1 are randomly distributed across phylogeny
H_1: Changes in trait 1 depend on changes in trait 2

Procedure:
  -Map traits to phylogeny & estimate ancestral states
  -Identify locations of changes in each trait
  -Calculate:

$$Pr = {}^A\!/_B$$

A: $\#(G,L_{\text{obs.trt1}}|$ distribution of trait 2, $\tau)$
B: $\#(G,L_{\text{obs.trt1}}| \tau)$

-The number of ways gains/losses could occur on phylogeny (B), versus the number of ways G/L occur
       while accounting for distribution of trait 2 (A)
-Calculations are a bit involved, but straightforward
-Use probability rules for A & B (**the <u>OR</u> rule**: for details see Maddison 1990)!

τ: the phylogeny

Discrete Data: Concentrated Changes

Is signaling behavior associated with ecology in lizards?

Large display repertoire Closed habitat

Large display repertoire Arboreality

Black: trait present

| Closed habitats | + via perceptual evolutionary inertia or niche diversity | Signal diversity |

Home range size	- via relaxed degradation constraints	
Arboreality	+ via greater visual acuity	Signal complexity
Diet of mobile prey	+ via resource defensibility	
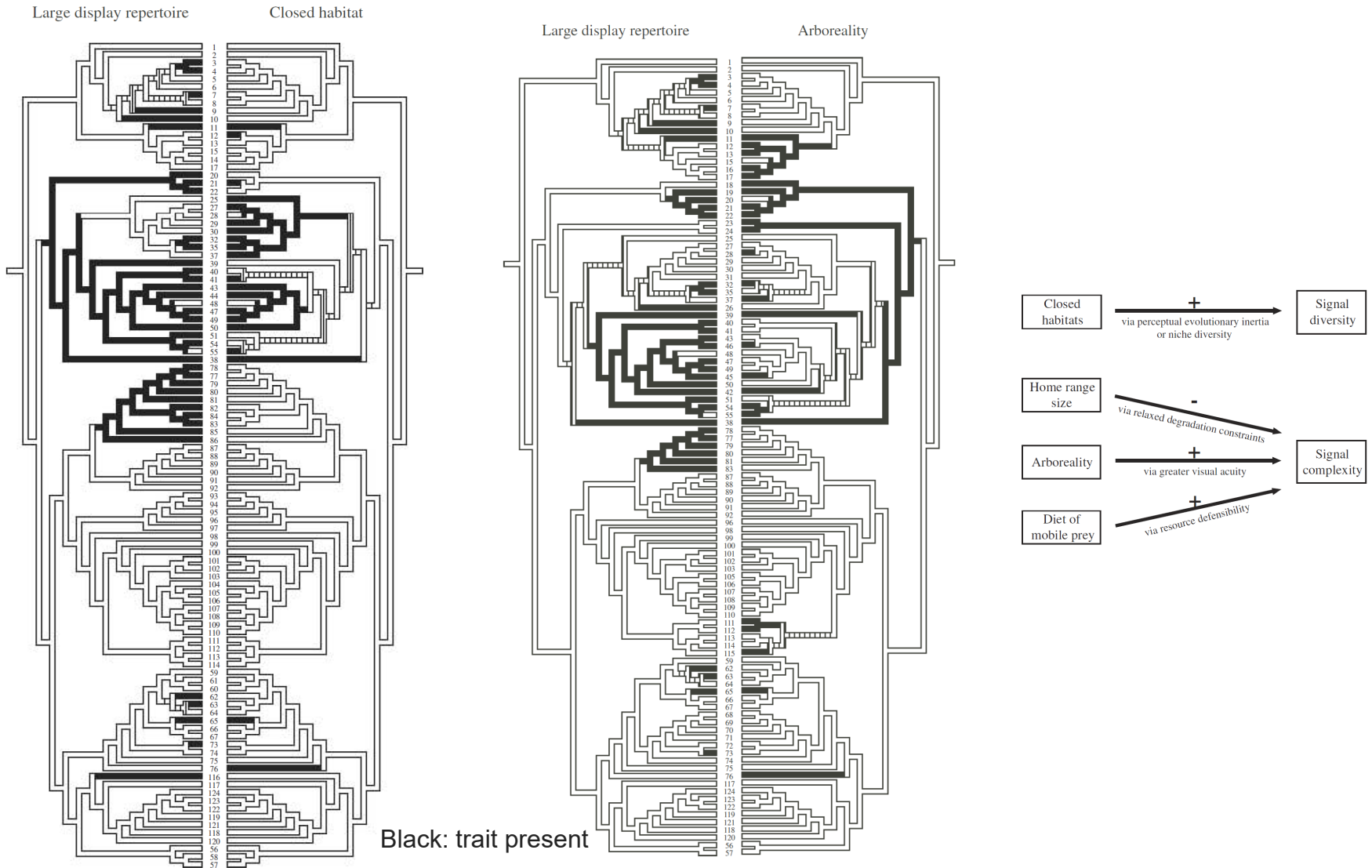
Ecology facilitated evolution of increased signal complexity

Ord et al. (2012)

The problem can be modeled statistically

   -Estimate probability transition matrix between states ($\mathbf{P}$)

$$\mathbf{P}(t) = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

Looks like P matrices in phylogenetics!
(and in fact is derived from that formulation)

   -p_{ij} is probability of transition i→j at time t

   -Rows sum to 1 (so $p_{00} = (1-p_{01})$)

More useful to consider transition *RATES:* $\mathbf{P}(t) = e^{\mathbf{Q}(t)}$

$$\mathbf{Q}(t) = \begin{bmatrix} (1 - q_{01})dt & q_{01}dt \\ q_{10}dt & (1 - q_{10})dt \end{bmatrix}$$

$\mathbf{Q}$ is estimated using maximum likelihood (or Bayesian methods)

$$\mathcal{L}(\mathbf{Q}) = \Pr(\boldsymbol{X}|\mathbf{Q}, \boldsymbol{\tau})$$

Pagel (1994)

Correlated change via expanded **Q** matrix

$$
\begin{array}{c}
\text{X,Y} \\
\begin{array}{cccc}
 & \text{0,0} & \text{0,1} & \text{1,0} & \text{1,1}
\end{array} \\
\begin{array}{c}
\text{0,0} \\
\text{0,1} \\
\text{1,0} \\
\text{1,1}
\end{array}
\begin{bmatrix}
1-(q_{12}+q_{13}) & q_{12} & q_{13} & 0 \\
q_{21} & 1-(q_{21}+q_{24}) & 0 & q_{24} \\
q_{31} & 0 & 1-(q_{31}+q_{34}) & q_{34} \\
0 & q_{42} & q_{43} & 1-(q_{42}+q_{43})
\end{bmatrix}
\end{array}
$$

Q describes transitions in both traits across the phylogeny

Hypotheses of trait evolution evaluated by comparing different q-values

Pagel (1994)

Hypothesis testing via model comparison:

1: Jointly estimate $\mathbf{Q}$ & $\mathcal{L}(\mathbf{Q})$. This is the unconstrained model

2: Define 'constrained' model based on biology

-These define dependent state changes: $\mathbf{Q_d}$

-e.g., G/L rates identical ($q_{13} = q_{31}$);

| test	hypothesis	description	statistic
omnibus χ^2 (4 d.f.)	$L(I) < L(D)$	correlated evolution	$L(I)$ versus $L(D)$
contingent change (1 d.f.)	$q_{12} \neq q_{34}$	change in Y depends upon state of X	$L(D_7)$ versus $L(D_8)$
	$q_{21} \neq q_{43}$	change in Y depends upon state of X	$L(D_7)$ versus $L(D_8)$
	$q_{13} \neq q_{24}$	change in X depends upon state of Y	$L(D_7)$ versus $L(D_8)$
	$q_{31} \neq q_{42}$	change in X depends upon state of Y	$L(D_7)$ versus $L(D_8)$
temporal order or relative rate (1 d.f.)	$q_{12} \neq q_{13}$	order of acquisition of X versus Y	$L(D_7)$ versus $L(D_8)$
	$q_{42} \neq q_{43}$	order of loss of X versus Y	$L(D_7)$ versus $L(D_8)$
alternative models	$q_{ij} = 0$	q_{ij} transitions excluded	$L(D_7)$ versus $L(D_8)$
branch transformation	$\kappa \neq 1$	transformation of branch lengths	$L(D_n)$ versus $L(D_{n+1})$

3: Jointly estimate $\mathbf{Q_d}$ & $\mathcal{L}(\mathbf{Q_d})$

4: Model comparison with LRT: $X^2 = -2\log(\mathcal{L}(\mathbf{Q})/\mathcal{L}(\mathbf{Q_d}))$ *

*Bayesian methods may also be used

Discrete Data: Transition Model

Does sex-determination co-evolve with chromosome changes in turtles?
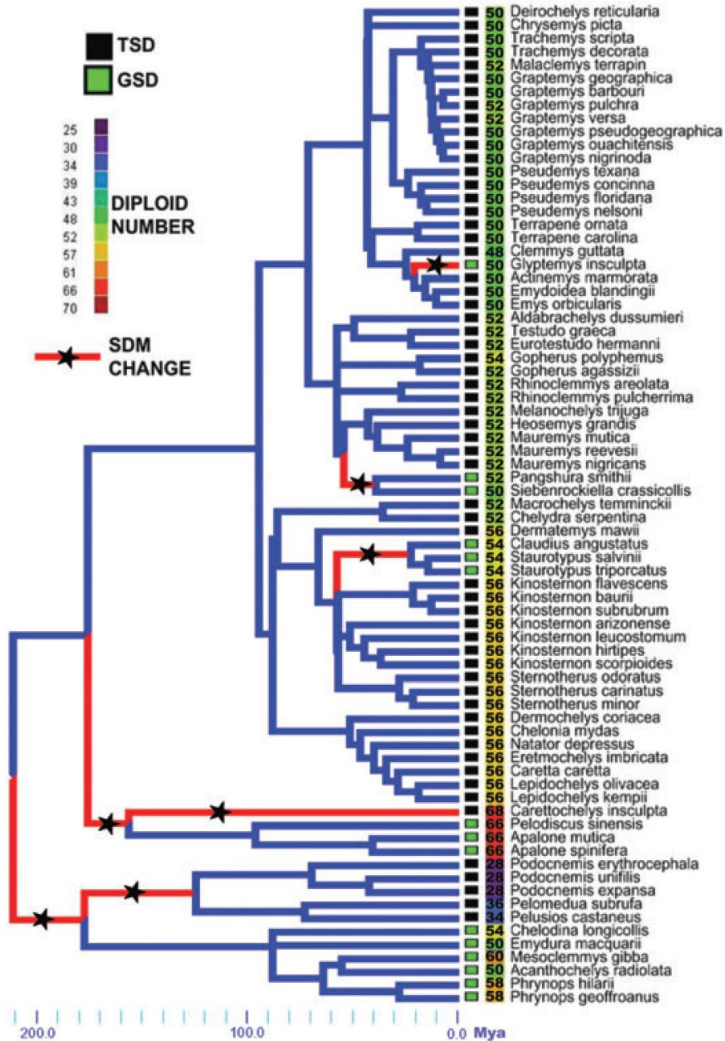
Did changes in SDM precede
Chrom or vice-versa?

SDM: TSD/GSD
Chrom: >52/ <52*
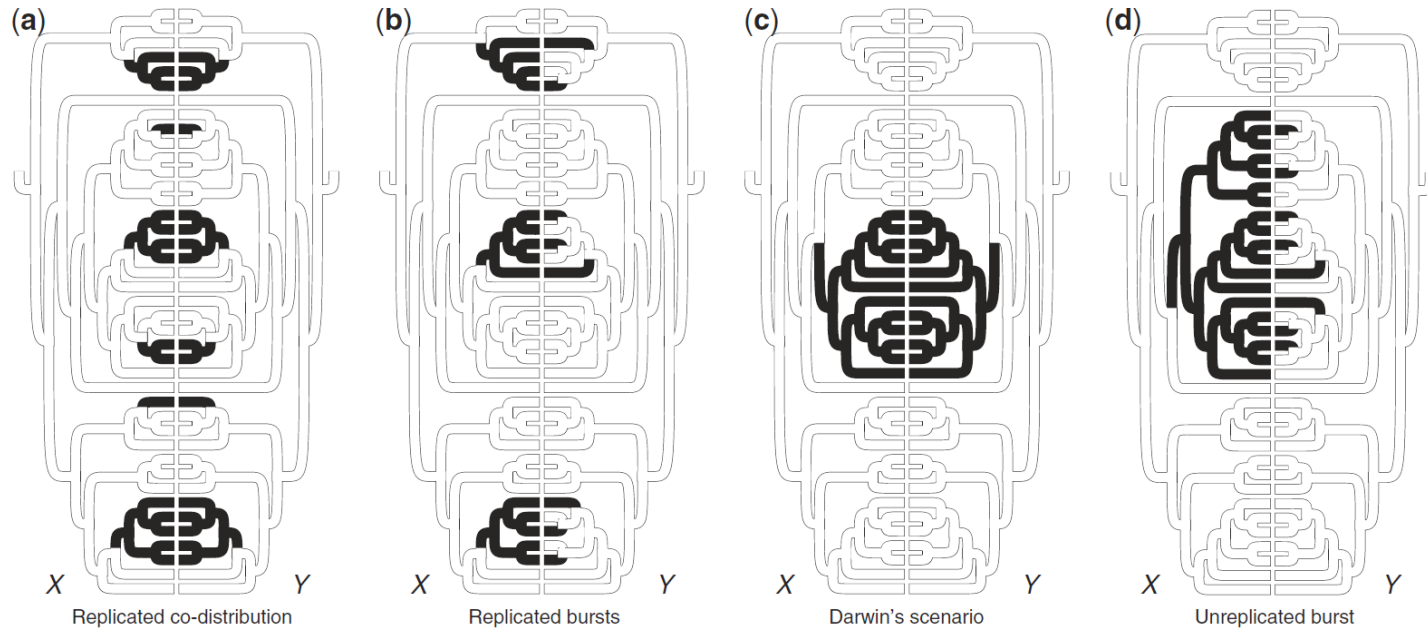
Q_d: equal transition rates for
SDM/Chrom
$$q_{12}=q_{13}; \; q_{42}=q_{43}$$

-Changes in traits were associated
-No support for one preceding the other

Valenzuela and Adams (2011)

Discrete Data: The Problem of replication

Discrete trait methods sensitive to number of transitions on phylogeny*

(a) Replicated co-distribution
(b) Replicated bursts
(c) Darwin's scenario
(d) Unreplicated burst

C&D lead to inference problems

CAREFUL in interpretation!

Always plot traits/transitions (don't just examine significance levels)

*We will see same issue with BiSSE models

Maddison & FitzJohn (2015)

Similar issues arise when evaluating association of continuous traits
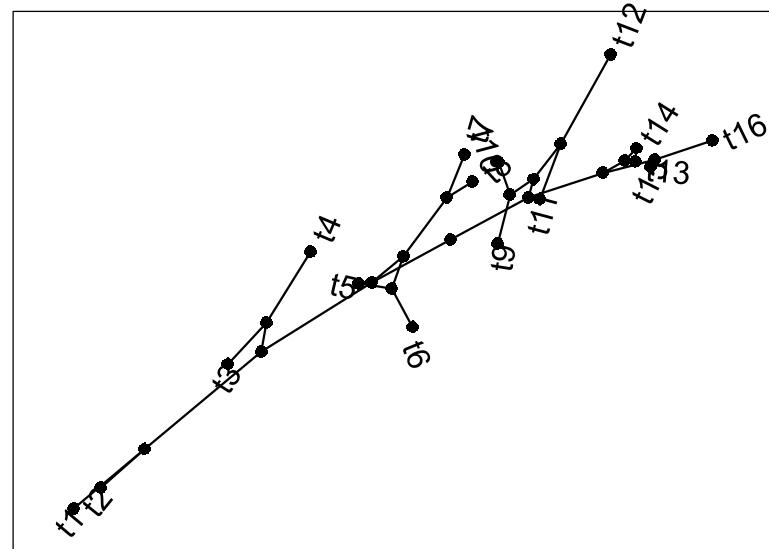

$r = 0.915; P > 0.0001$

Say we have this correlation,

Here is the same pattern with the phylogeny superimposed

Clearly, closely related taxa are similar.
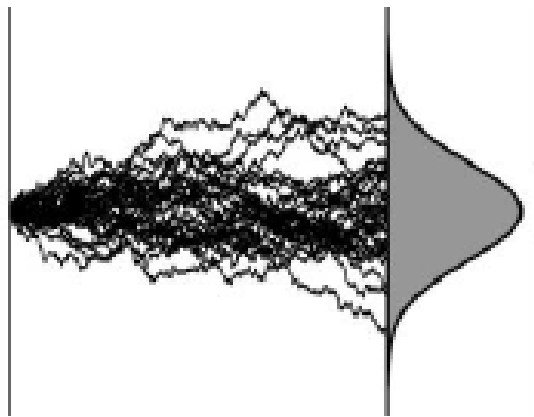
How can we account for this?

Accounting for non-independence due to phylogeny requires a *null model* of evolutionary change

Brownian motion is the null model:

Trait changes are independent from time step to time step

Outcome: no change in μ, but $\sigma^2_y \uparrow \propto$ time

Side-note: this is the continuous-trait model equivalent of the Markov process we discussed earlier

Given this *model*, one can calculate contrast scores between taxa that are independent of one another relative to the phylogeny

Felsenstein (1985)

Phylogenetically Independent Contrasts (PIC)

Algorithm:
- Calculate contrasts for sister taxa
- Estimate ancestral value
- Continue recursively down phylogeny

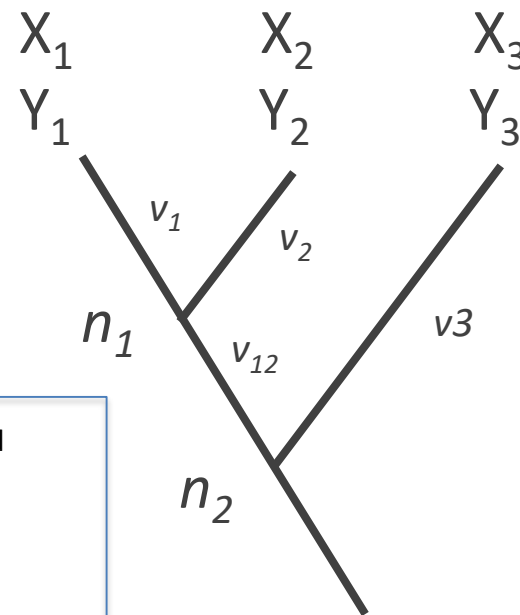**Contrast scores**

$$Y_{pic} = \frac{Y_i - Y_j}{\sqrt{v_i + v_j}}$$

Trait difference scaled by time

Internal nodes: weighted average

$$Y_{n1} = \frac{1/v_1 \, Y_1 + 1/v_2 \, Y_2}{1/v_1 + 1/v_2}$$

*NOTE: Internal branches adjusted FOR THESE CALCUATIONS as:

$$v_{ij}^* = v_{ij} + \left(\frac{1}{1/v_i + 1/v_j} \right)$$

X_1 X_2 X_3
Y_1 Y_2 Y_3

v_1 v_2

n_1 v_{12} $v3$

n_2

What is this???
The pruning algorithm!!!!

NOTE: For polytomies 'resolve' them using zero-length branches (neat tidbit: doesn't matter for which taxon!)
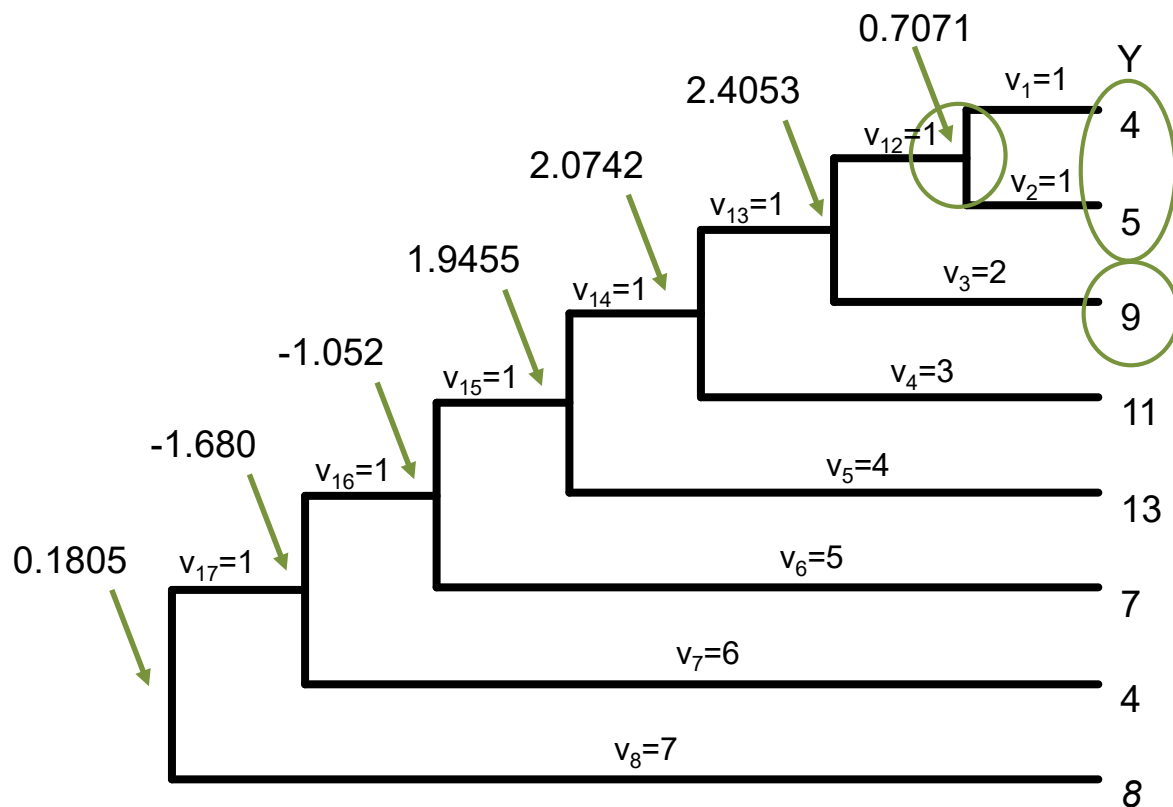
Felsenstein (1985)

Continuous Data: Independent Contrasts

$$c_1 = \frac{5 - 4}{\sqrt{1 + 1}}$$

$$c_2 = \frac{9 - 4.5}{\sqrt{1.5 + 2}}$$

Why? **The pruning algorithm!**

(need ancestral estimate, which is based on weighted average and rescaled branch lengths)

Internal nodes: weighted average

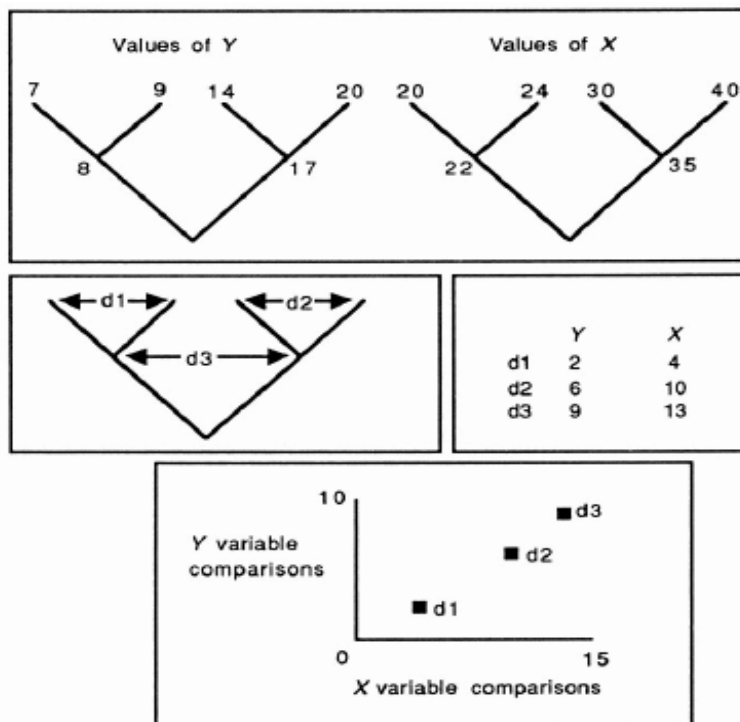$$Y_{n1} = \frac{{}^{1}\!/_{v_1}\, Y_1 + {}^{1}\!/_{v_2}\, Y_2}{{}^{1}\!/_{v_1} + {}^{1}\!/_{v_2}}$$

*NOTE: Internal branches adjusted FOR THESE CALCUATIONS as:

$$v_{ij}^{*} = v_{ij} + \left(\frac{1}{1/v_i + 1/v_j} \right)$$

Felsenstein (1985)

What are the contrasts?
 -Amount of evolutionary change since common ancestor

 Represents the independently-evolved trait differences between taxa

  -There are *N-1* contrasts for *N taxa*

Statistical analysis (regression) performed on contrasts rather than tips

data $\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}_{pic}^{t}\mathbf{X}_{pic}\right)^{-1}\mathbf{X}_{pic}^{t}\mathbf{Y}_{pic}$ (note: through origin because order of tips in contrast irrelevant)

Felsenstein (1985)
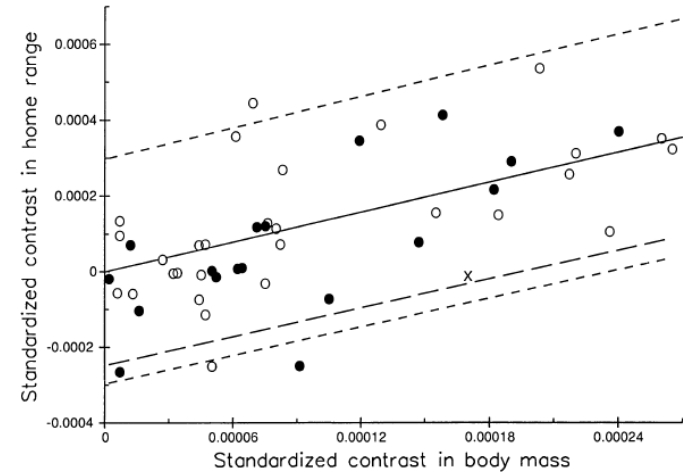
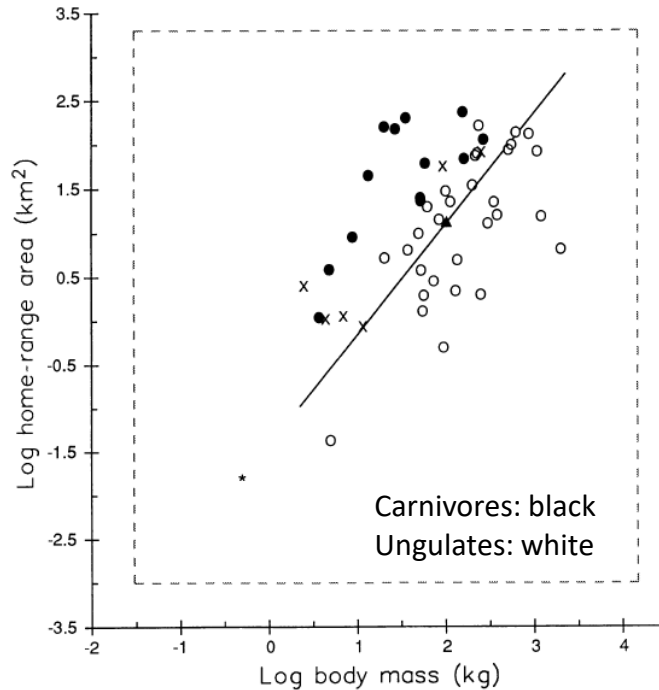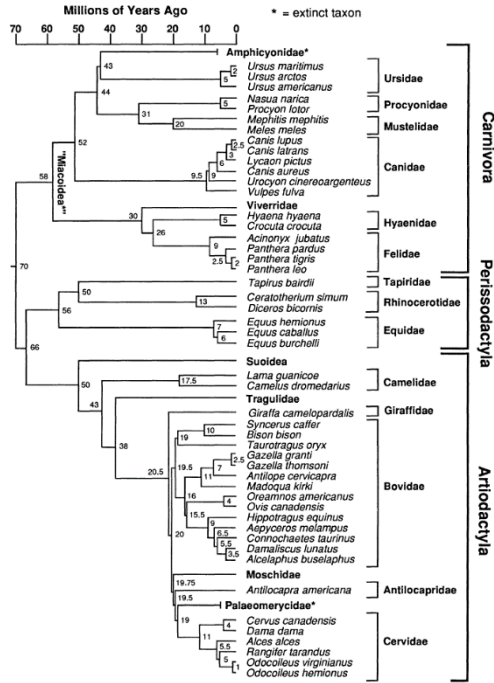# Continuous Data: Independent Contrasts

How are leaf traits associated across taxa (prior non-PCM analyses showed correlations)?
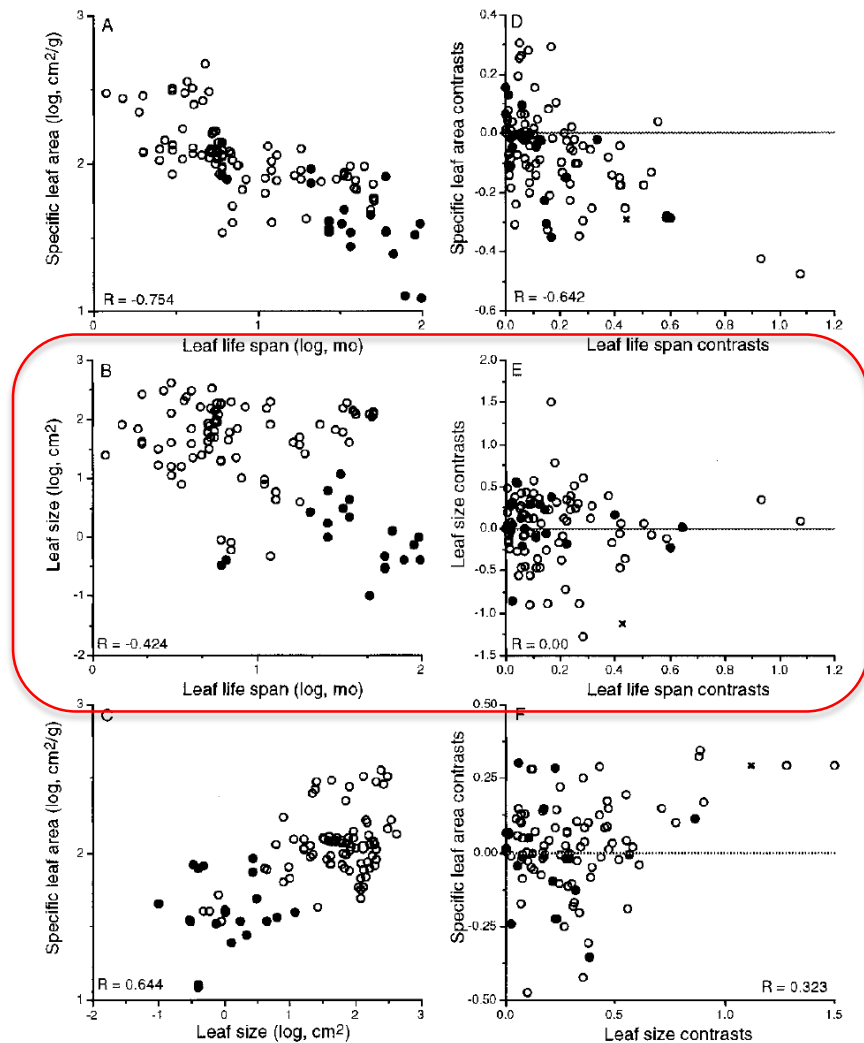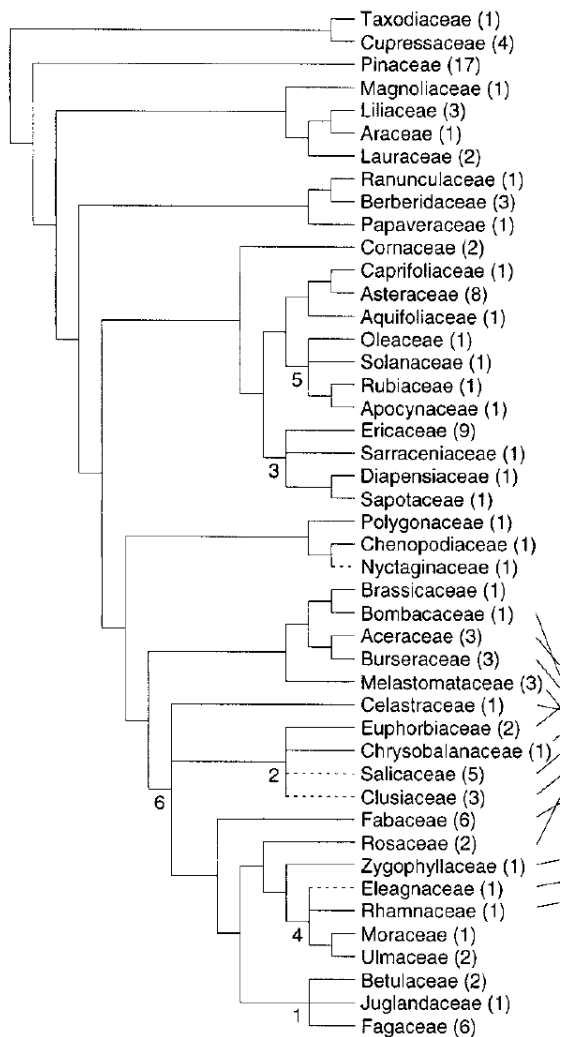
Non-phylogenetic association due to angiosperm-gymnosperm differences

Ackerly and Reich (1999)

Continuous Data: Independent Contrasts

Is there an evolutionary relationship between body size and range in mammals?

Carnivores: black
Ungulates: white

ANCOVA: Significant BM-HR slope
carnivore HR > ungulate HR

ANCOVA: Significant BM-HR regression
no HR differences

Garland et al. (1993)

Conclusions
  -PICs account for non-independence due to phylogeny
  -Clever use of pruning algorithm*
  -Can work with multifurcations (polytomies)

Unanswered questions
  -Is it general?
  -How to examine ANOVA models? Factorial models?
  -What is the relationship between PIC and general statistical theory?

-Stay tuned!

*For a fascinating geometric interpretation and matrix-algebra
way of obtaining PICs see Rohlf (2001)