PHYLOGENETICALLY INDEPENDENT CONTRASTS

LIZA MARIA GARCIA

4/8/2009

PHYLOGENETICALLY INDEPENDENT CONTRASTS

Definition

PIC is a comparative method for testing hypotheses of how organisms are adapted to their environment. The method of independent contrasts uses phylogenetic information, to transform interspecific data into values that, in principle, are independent and identically distributed, and hence can be analyzed with standard statistical methods like regression.

"**PIC** (Felsenstein, 1985) **is the method best justified, best understood and most versatile of existing phylogenetically based statistical methods. So far, it must be viewed as the method of choice for comparative analyses of continuous-valued characters and is widely applied in ecological, evolutionary morphology and physiology studies.**"

(Garland et al., 1999).

OUTLINE

❖ **History**

    ❀ **Problems with previous approaches**

❖ **Methodology and assumptions**

❖ **Why and when to use Phylogenetically Independent Contrasts (PIC)**

❖ **PIC limitations**

❖ **Conclusions**

HISTORY

- ❖ Use of regression approaches to asses relationship between variables

  - @ **Insect chromosome numbers and social behavior** (Sherman, 1979)

  - @ **Population density and body weight** (Damuth, 1981)

  - @ **Brain weight and body weight** (Martin, 1981)

- ❖ **Problem: such approach ignores differences in evolution rates and phylogeny**

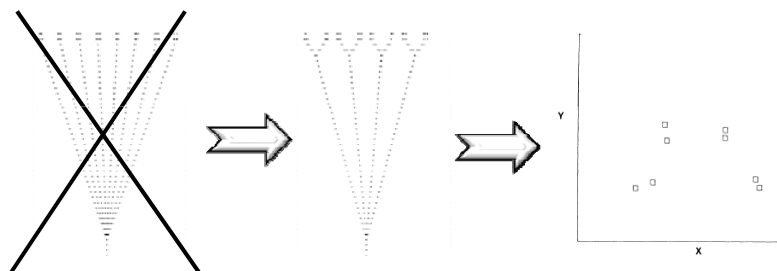 - @ **Samples considered independent from each other**

Felsenstein, 1985

HISTORY (continued)
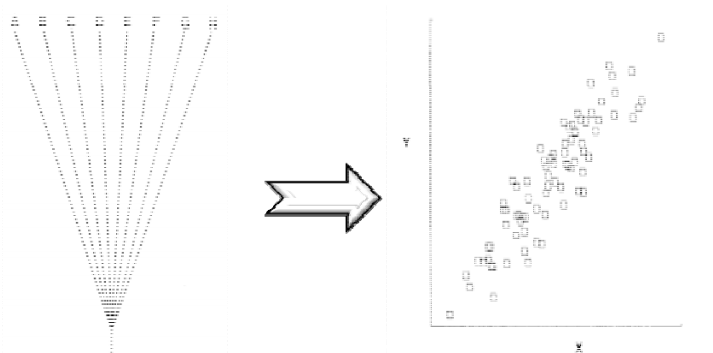
Assuming the 8 species resulted from a single explosive adaptive radiation.
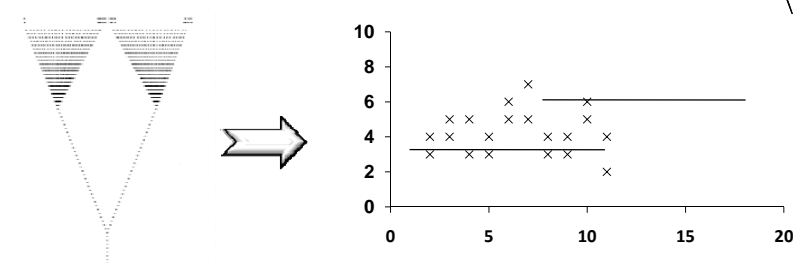
Felsenstein, 1985

Problems with previous approaches

What if the actual phylogeny tree was not this one …

But this other one …?

This would be the resulting scatter plot considering the related species

Felsenstein, 1985

"worst case" scenario

- ◉ **Phylogeny for 40 species in which prove to be 2 groups, each of 20 close relatives**

- ◉ **Sampling equal number of individuals from each branch does not solve the problem**

Felsenstein, 1985

The exception

❖ **There is one case in which the problem does not arise**

  ◉ When characters respond instantaneously to natural selection in the current environment

  ◉ Phylogenetic inertia is essentially absent

  ◉ In that case the phenotype can be correlated with the environment

  ◉ If comparative studies are done without accounting for phylogeny, the absence of phylogenetic inertia should be presumed

Phylogenetic inertia
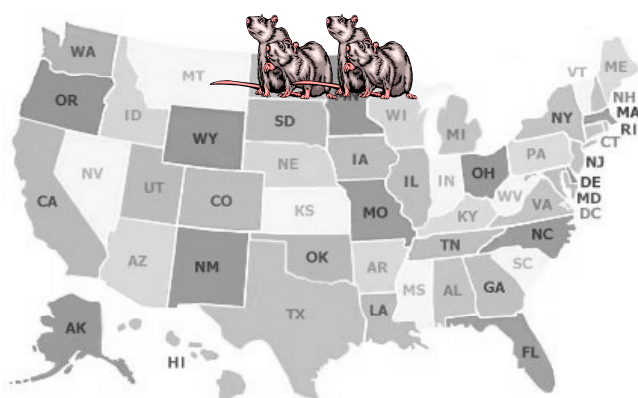
How to correct for phylogeny?

Before the development of PIC several approaches were attempted:

❖ **Nested analysis of variance** (Clutton-Brock and Harvey, 1977)

❖ **Analyzing if the same relationships were held within different families** (Baker and Parker, 1979)

❖ **Reconstruct phylogeny using parsimony** (Gittleman, 1981)

What if we do not account for phylogeny?

Statistical consequences

❖ **Inflation of Type I error rates when testing hypothesis due to overestimation of degrees of freedom**

❖ **Lowered power to detect significant relationships**

❖ **Inefficient estimates of evolutionary parameters**

The Method

If we know the phylogeny and have a model of evolutionary change, it should be possible in principle to correct for the nonindependece of taxa
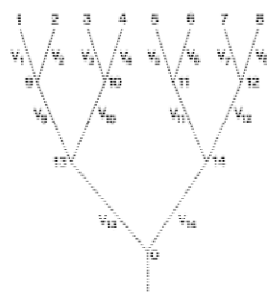
- X_i =phenotype in X specie X_1 - X_2 and X_3 –X_4
- v = Units of time
- X_1 --X_2 has expectation zero and variance S^2_x
- S^2_y = Y variance after v units of time.
- Contrasts obtained by dividing each contrast by its standard deviation
- Seven independent contrast in X at seven independent contrast in Y
- The X contrast will be independent of each other but not of the Y contrast

Felsenstein, 1985

PIC ASSUMPTIONS

The three main assumptions of independent contrast are:

❖ Correct topology

❖ Branch-lengths measured in units of expected variance of character evolution

❖ Brownian motion (BM) model of character evolution

PIC yield the nominal type I error rates

Diaz-Uriarte and Garland, 1996

PIC Applications

Most of the phylogenetically based "comparative methods" have focused on testing for character correlations, with continuous traits

- ❖ **body size**
- ❖ **limb proportions**
- ❖ **metabolic rate**
- ❖ **home range area**

Garland, 1993

PIC Applications (continued)

Biologists often wish to compare mean phenotypes among groups of organisms defined on the basis of phylogenetic affinity or on behavioral, ecological, or physiological criteria.
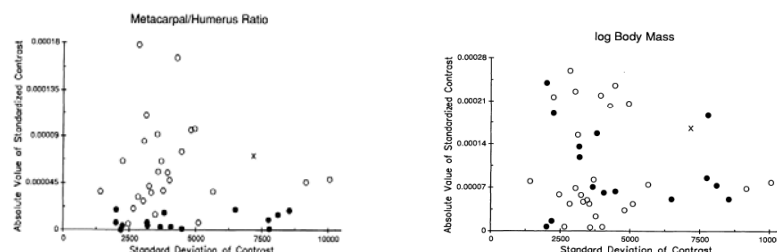
Garland, 1993

Comparing rates of phenotypic evolution among clades

❖ **Garland (1992), used PIC, to hypothesized phylogenetic relationships and divergence times between carnivora and ungulates**
❖ **Five phenotypic characters:**
- **average adult body mass**
- **maximal sprint running speed**
- **metatarsal-femur ratio**
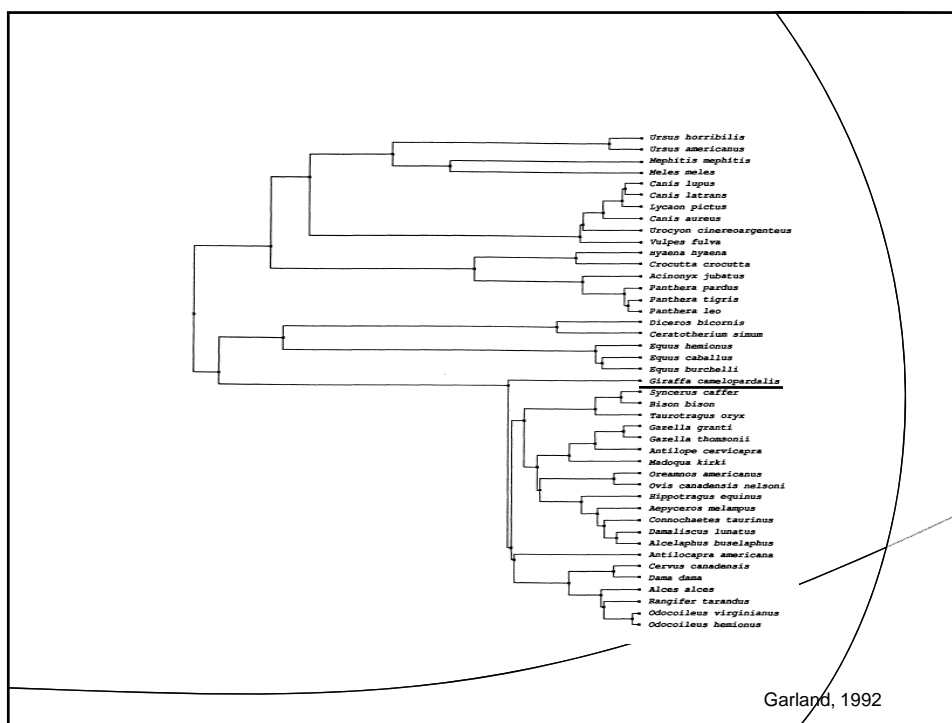- **metacarpal-humerus ratio**
- **home-range**

Garland, 1992

Phenotype plots

- **Comparison between carnivora and ungulate**
- **X represented the basal contrast between carnivore and ungulate**
- **A simple statistical comparison of means of the absolute values of the standardized independent contrast was done**

Garland, 1992

Garland, 1992

Biological significance

❖ **The five traits analyzed do not diverge on the same time scales and evolve different in carnivora and ungulata.**

❖ **Branches lengths in unit of time have intuitive appeal for yielding rate comparisons.**

❖ **Divergence seems not to occur without limit, evolution slows down over longer time periods within both carnivora and ungulata.**
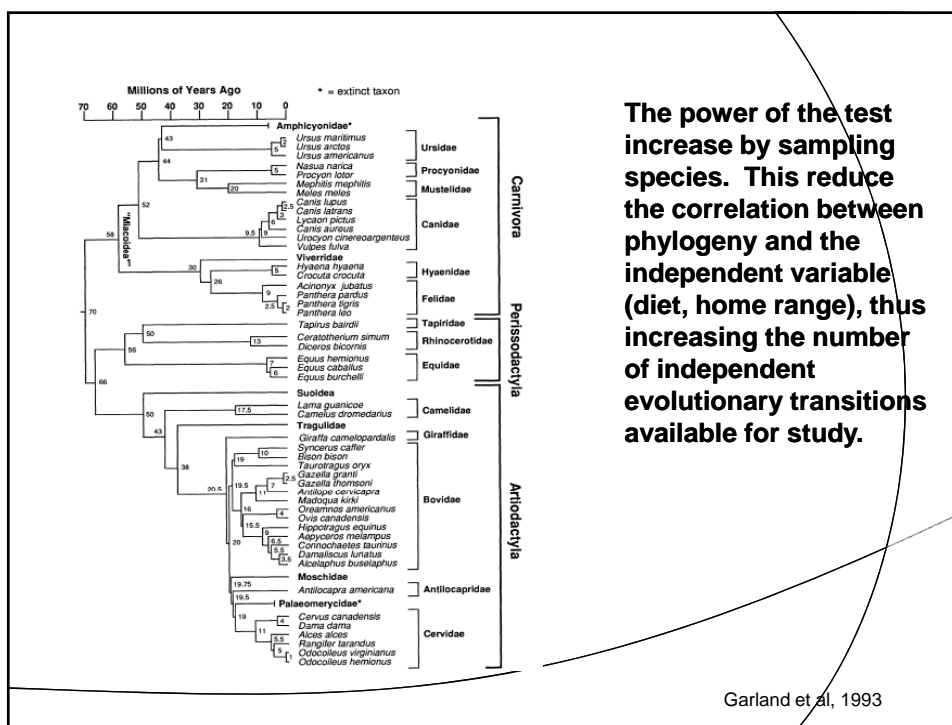
Garland, 1992

Other example

Garland et al., 1993, proposed a computer program that perform simulations under a variety of evolutionary models.

They analyzed the data using ANCOVA and applied phylogenetic autocorrelation between carnivora and ungulata, to show how PIC can be used to test for clade differences.

Garland et al, 1993

The power of the test increase by sampling species. This reduce the correlation between phylogeny and the independent variable (diet, home range), thus increasing the number of independent evolutionary transitions available for study.

Garland et al, 1993

Translocation success

❖ **Wolf et al., 1998 tested the hypothesis of a phylogenetic relationship between avian and mammalians species**

❖ **used three variables associated with translocation success**

- **the habitat quality of the release location**

- **the number of individuals released**

- **the range of the release area relative to the historical distribution of the species.**

Wolf et al., 1998

Translocation success

Avian phylogeny   **Mammalian phylogeny**
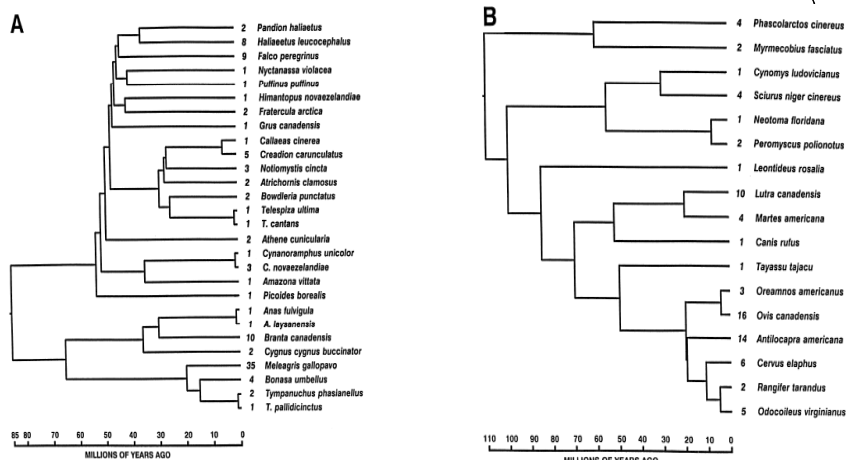
Wolf et al., 1998

Factors affecting translocation success
(Example of PIC application)

| Variable | Nonphylogenetic | | | | Regression of Phylogenetically Independent Contrasts[e] | |
| | Logistic Regression[b] | | Ordinary Regression[c] | | | |
	B	p	B	p	B	p
Taxonomic Class	1.181	0.0231	0.413[d]	0.0262	0.028[d]	0.9735
Diet		0.0003		0.0414		0.1135
Carnivorous	2.070		0.275[d]		−0.558[d]	
Omnivorous	2.008		0.551[d]		0.539[d]	
Migratory	−0.532[d]	0.2936	−0.332	0.0746	−0.140[d]	0.5360
Status	1.967	0.0001	0.513	0.0055	0.093[d]	0.4651
Habitat Quality	0.817	0.0098	0.209	0.0500	0.176	0.0570
# Animals ($\log_{10}$)	2.512	0.0001	1.018	< 0.0001	0.190	0.0592
Range	3.847	0.0119	2.269	< 0.0001	0.400	0.0033
Range × # Animals	−1.415[d]	0.1023	−0.907	0.0003	0.415	0.0296
Constant	−7.777		−0.413		—	
Model Statistics						
X^2	74.637 (9df)					
p value	< 0.0001					
r^2			0.3365		0.1599[e]	
F			9.638		3.617	
p value			< 0.0001		0.0004	

Translocation success is nor related to Taxonomic Class !

Wolf et al., 1998

Conventional statistic analyses can be used to suggest important predictor variables in an interspecific comparative data set.

Phylogenetically based procedures are theoretically more reliable in terms of hypothesis testing and parameter estimation.
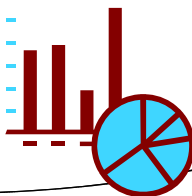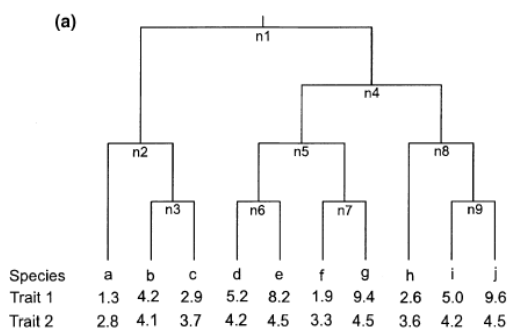
Wolf et al., 1998

PIC Limitations

❖ **Before using PIC we should consider the shape of the relationship between traits**

❖ **Non linear relationship between traits**

   ✓ **Although transformations can be made (log, etc) to ensure that Brownian assumptions are satisfied**

   ✓ **A reduction in statistical power is observed when data are nonlinear**
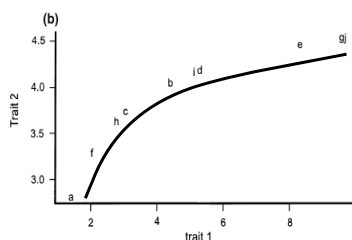
Quader et al., 2004

Example of problems with nonlinear relationships

Species	a	b	c	d	e	f	g	h	i	j
Trait 1	1.3	4.2	2.9	5.2	8.2	1.9	9.4	2.6	5.0	9.6
Trait 2	2.8	4.1	3.7	4.2	4.5	3.3	4.5	3.6	4.2	4.5

Each branch segment has a length of either 1 or 2 units.   The internal nodes are labeled n1–n9.
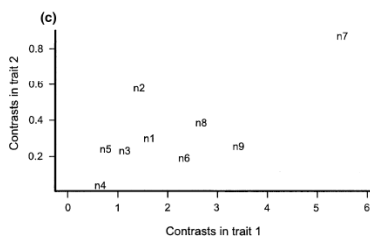
Quader et al., 2004

Plots of nonlinear relationships

- Nonlinear relationship between trait 1 and trait 2
- Note that not particular clade is responsible for nonlinearity

- The tight relationship above has degenerated into a loose scatter of point (right)
- The points are labeled by the nodes they represent in the tree

Quader et al., 2004

Conclusions

❖ **Because species usually will not represent independent data points in the statistical sense, conventional parametric and non- parametric methods are inappropriate for hypothesis testing with interspecific data**
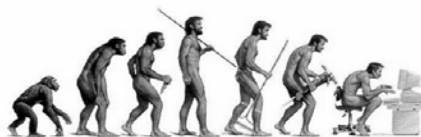
Conclusions

❖ **PIC is useful for comparisons of evolutionary rates of different clades**

❖ **PIC is useful in analyses of comparative data sets used in conservative projects**

❖ **PIC is commonly used for exploring cross-taxon relationship between traits**

❖ **Nonlinear relationships may require data transformation prior to analysis**

In summary ...

"Phylogenies are fundamental to comparative biology; there is no doing it without taking them into account."

Felsenstein, 1985