"PRINCIPLES OF PHYLOGENETICS: ECOLOGY AND EVOLUTION"
Integrative Biology 200 Spring 2014
University of California, Berkeley D.D. Ackerly

March 7, 2014

Readings

Hansen, T. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341-1351.

Ackerly, D. D. 2009. Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. Proceedings of the National Academy of Sciences USA 106:19699-19706.

Other strongly recommended readings!

Blomberg, S. P., and T. Garland, Jr. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. J. Evol. Biol. 15:899-910.

Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. Evolution 57:717-745. (Focus on pp. 719-723 on methods for measuring phylogenetic signal)

Harmon, L. J., J. B. Losos, T. J. Davies, et al.. 2010. EARLY BURSTS OF BODY SIZE AND SHAPE EVOLUTION ARE RARE IN COMPARATIVE DATA Evolution 64:2385-2396.

I. Phylogenetic signal

Pattern and Process

Pattern is not process. No matter how many times we hear and repeat this phrase, the temptation to infer the action of specific processes from pattern alone is inevitable. In comparative biology, it has been all to common to equate stasis (pattern) with 'inertia' or 'constraint' (process?) and evolutionary change (pattern) with the action of selection (process). It is essential at the outset to see these dichotomies of pattern and process as orthogonal, setting up a 2x2 table:

	Pattern of trait evolution	
Evolutionary process	Stasis	Change
Adaptive	Stabilizing selection Fluctuating directional selection	Directional or disruptive selection
Non-adaptive	Lack of genetic variation (= constraint?) Antagonistic correlations among traits under selection Swamping by gene flow	Mutation Genetic drift Genetic correlations with traits under selection

In this case, I equate the term 'adaptive' with patterns resulting from the action of selection. It is not a statement about the functional significance of a particular trait, nor is it the same as stating that a particular trait represents an 'adaptation'. More on that in the next lecture.

It is critical to remember, as well, that these processes are neither exclusive nor independent of each other. For example, the action of drift and selection in small populations may simultaneously contribute to changes in gene frequency, and it may be difficult or impossible in an individual instance to separate their contributions. Replication in space (populations) or time (across generations) is critical, as selection is the only process we currently know of in evolutionary biology that can cause repeated, heritable change in the same direction in response to environmental conditions. Note also that selection leads to the loss of genetic variation. Thus, strong selection on a trait may lead to fixation of genes responsible for variation, such that future maintenance of the trait is due to the absence of sufficient variation (even if the selective context shifts).

As discussed in the Blomberg and Garland paper, there is a long history of usage of terms such as phylogenetic inertia or constraint, and other more recent terms such as phylogenetic effect and conservatism. To a greater or lesser degree, all of these terms imply the action of some underlying process. The term *phylogenetic signal* is useful as the implication is clear that we are only speaking about the pattern. I also like phylogenetic conservatism, though it is now taking on its own intellectual baggage and may suggest particular hypotheses about process (e.g., see Losos 2008 Ecology Letters).

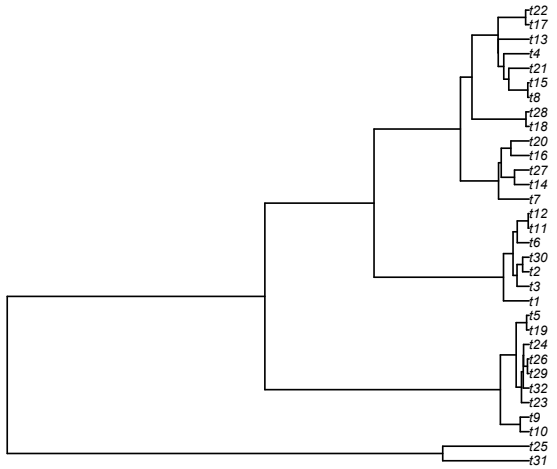
Quantifying phylogenetic signal for continuous traits
Several statistical have been developed to quantify phylogenetic signal, and test for significance relative to a null model of no signal. Freckleton et al. (200x) discuss the λ parameter, which is derived from Pagel's model fitting methods. I have worked more with Blomberg et al. (2003) K statistic. λ and K have the same general interpretation:

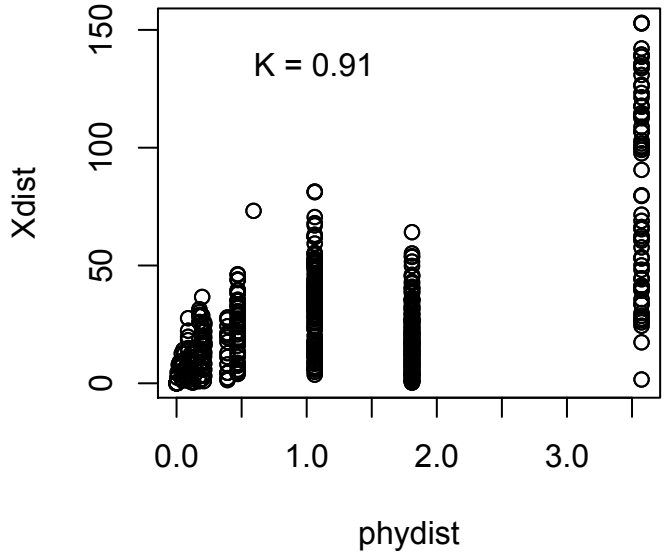
λ, $K < 1$	less signal than expected under Brownian motion
λ, $K = 1$	expectation under B.M.
$K > 1$	more signal than expected under B.M.

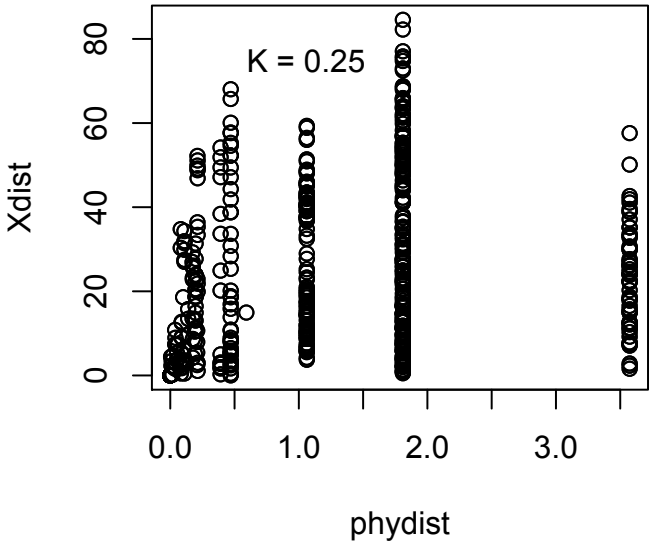
Quantitatively, one of the main differences is that λ has a maximum around 1.1 or 1.2, so it has little resolution to describe patterns with a high degree of signal (Freckleton et al. don't report values > 1). K on the other hand has a theoretical maximum of infinity and can provide some insight into differing degrees of signal greater than expected under Brownian motion.

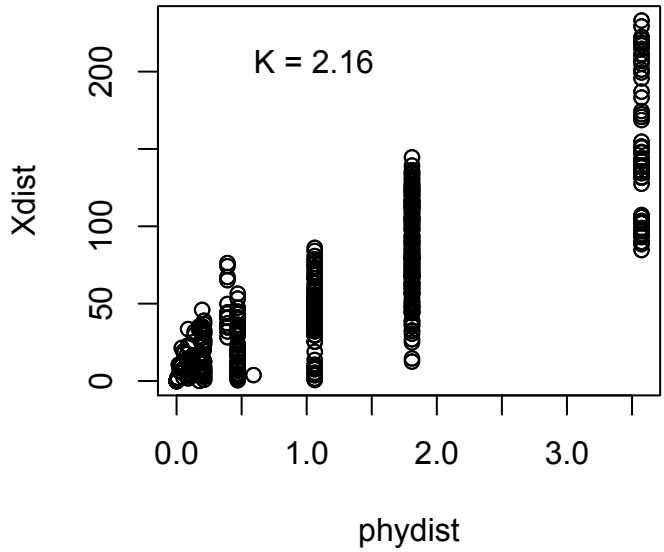
Fortunately, there is a simple graphical interpretation of these statistics, based on a plot of the difference in trait values between species pairs (all pairwise comparisons, not just sister taxa!) versus the phylogenetic distance between the taxa (Figure on next page). Remember from Brownian motion, the expectation of the squared difference between trait values (the variance) goes up linearly with the phylogenetic distance. It turns out that this translates approximately into a filled half triangle in a plot of the absolute difference vs. phylogenetic distance for all species pairs (assuming the tree is ultrametric). Concave curvature of the upper surface of this plot means that close relatives are less similar than expected, compared to distant relatives ($K < 1$). Convex curvature, or an absence of points in the lower right, means that close relatives are more similar than expected ($K>1$). The details of the plot structure will depend on tree topology, but in general examining these plots is quite informative. Significance tests for the correlations in such plots must be tested using Mantel tests for matrix correlations (another topic!).

THE TREE







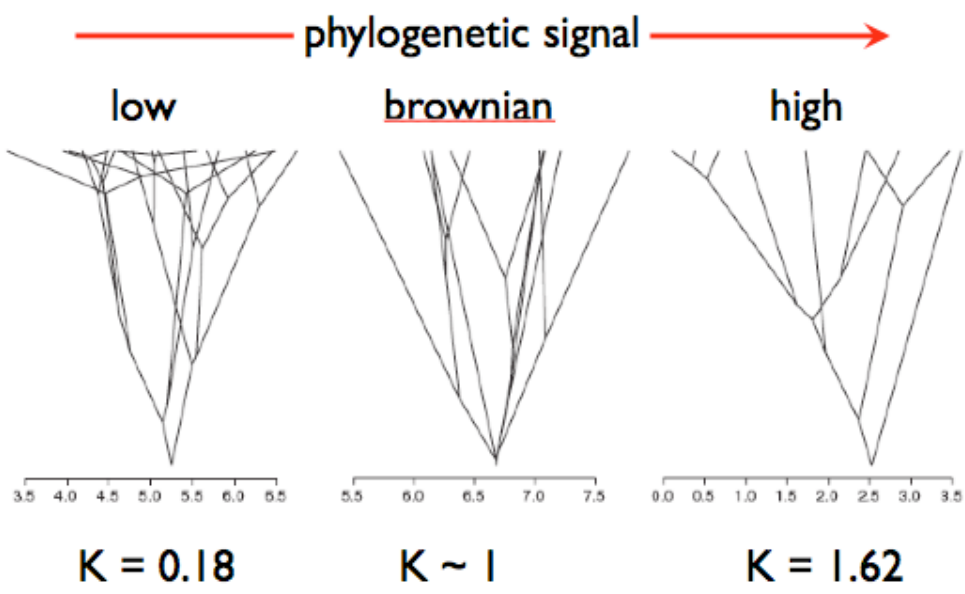


Another approach to asking about phylogenetic signal and patterns is to ask whether the trait data fit an alternative model of trait evolution, and whether the parameters of this model then shed light on the processes influencing the trait and its evolutionary history. One of the very important alternative models is to introduce stabilizing selection, i.e. a model in which there is a 'pull' towards some central optimal value (which may fluctuate along different lineages), and Brownian motion reflects the random processes and/or the excursions in the action of selection around this underlying optimum. The stabilizing selection model is known as an Ornstein-Uhlenbeck stochastic process, so you will see references to OU models (and there is an R library called 'ouch' for OU models).

Another graphical approach is to draw a 'traitgram', a phylogeny where the tips are arranged left to right based on species trait values, and internal nodes show the maximum likelihood reconstructions. This can get a bit messy, as the branches cross over each other. Lots of crossovers shows that a trait has exhibited many reversals of trait evolution during its history. If those crossovers occur particularly among the most recent divergence events, then it's evidence of rapid evolution and a departure from Brownian motion, as the trait is not continuing to diverge at constant rates. Figure below is from Ackerly (2009).

As you can see from this example, the Brownian Motion model predicts a certain pattern of relative similarity – how similar are close relatives compared to the similarities among distant relatives. It is not a statement about the absolute rate of evolution, and how much the actual trait values have diverged over time. You can rescale the trait axes in the figure below so each clade spans much less or more much actual trait space, and than doesn't change the fit of the data to Brownian Motion. This is because of the unbounded assumption of the model – evolving more quickly does not result in more convergence and loss of phylogenetic signal – and in this regard continuous and discrete characters are very different.

As we discussed earlier, the variance of the Brownian Motion process is a direct measure of evolutionary rate – how quickly are trait values diverging from each other over time. Ackerly (2009) proposed using this rate parameter with standardized measures of trait variation – using log-transformed data – and branch lengths scaled in millions of years as a measure of the rate of character evolution based on comparative data (without fossils).



There are many other possible models that one can simulate and use to fit empirical trait distributions for continuous characters, which have received more or less attention in the literature.

Mode	Model	Verbal explanation
Brownian	$x_{t+1} = x_t + N(0,s)$	random walk
Brownian + trend	$x_{t+1} = x_t + N(t,s), t \neq 0$	random walk with a trend
Bounded Brownian	$x_{t+1} = x_t + N(0,s)$ if x_{t+1} < minX or > maxX, resample	random wak with lower and/or upper limits to trait values
Proportional	$x_{t+1} = x_t * LN(\log mean = 0, \log sd = s)$	multiplicative random walk (equivalent to a random walk on log of trait)
ACDC	$x_{t+1} = x_t + N(0,s)$ $s = s_0 \gamma^{-t}$ $\gamma < 1$, accelerating $\gamma > 1$, decelerating	random walk with an accelerating or decelerating rate parameter
OU-1 (Ornstein-Uhlenbeck)	$x_{t+1} = \mu + \theta(x_t - \mu) + N(0,s)$ μ = trait optimum $\theta = 1$: brownian $\theta = 0$: complete stabilizing selection	stabilizing selection; random walk with a 'pull' towards the optimum for the clade; theta is strength of the pull
OU-2+ (Ornstein-Uhlenbeck)	$x_{t+1} = \mu + \theta(x_t - \mu) + N(0,s)$ μ_i = trait optimum on branch i $\theta = 1$: brownian $\theta = 0$: complete stabilizing selection	stabilizing selection; random walk with a 'pull' towards different optima on different branches, to test a priori hypotheses of adaptive evolution
Speciational	Brownian, with change only at time of speciation	random walk, but only one change per speciation event (i.e. branch length = 1 between each speciation event)
Punctuational	Brownian, with change only at time of speciation and only in one daughter lineage	like speciational, but change only occurs in one daughter lineage and the other exhibits stasis

There are many evolutionists who believe that stabilizing selection is the overarching cause of evolutionary stasis and the maintenance of similarity among close relatives. The OU model has been termed a stabilizing selection model, because trait values are ‘pulled’ back towards an optimum value. Paradoxically, an OU process with a single optimum generates traits with K<1, so stabilizing selection reduces phylogenetic signal, which is not intuitive at first! Estes and Arnold (2007, American Naturalist) offer an important discussion of stabilizing selection and apply it to a large data set compiled by Gingerich (1983, Science) on rates of morphological evolution (mostly from fossils, not comparative data).

OU models can be used in a very different way when we have an a priori hypothesis that evolution has favored different trait values in different clades, or on different branches of a tree. These are two-rate (or more) OU models that can use a nested-likelihood ratio approach to test

whether a model with two (or more) different optima are significantly better than one optimum. This is a powerful alternative to independent contrasts, because it doesn't focus on divergence between sister taxa; rather, the focus is on the degree of similarity among taxa hypothesized to be experiencing similar adaptive regimes. We will look at the first example (Hansen 1997).

Literature:

Abouheif, E. 1999. A method for testing the assumption of phylogenetic independence in comparative data. Evol. Ecol. Res. **1**:895-909.

Ackerly, D. D. 1999. Phylogeny and the comparative method in plant functional ecology. Pages 391-413 *in* M. Press, J. Scholes, and M. G. Barker, editors. Physiological plant ecology. Blackwell Scientific, Oxford, United Kingdom.

Ackerly, D. D., and M. J. Donoghue. 1995. Phylogeny and ecology reconsidered. J. Ecol. **83**:730-732.

Ackerly, D. D., and M. J. Donoghue. 1998. Leaf size, sapling allometry, and Corner's rules: a phylogenetic study of correlated evolution in maples (Acer). Amer. Nat. **152**:767-791.

Ackerly, D. D. 2009. Conservatism and diversification of plant functional traits: Evolutionary rates versus phylogenetic signal. Proceedings of the National Academy of Sciences USA 106:19699-19706. doi: 10.1073/pnas.0901635106.

Ackerly, D.D. 2009. Phylogenetics and comparative methods. Pages 117-125 in S. Levin, ed. Guide to Ecology. Princeton Univ. Press.

Ackerly, D.D. 2009. Phylogenetic methods in ecology. Encyclopedia of Life Sciences (ELS), John Wiley & Sons, Chichester, http:www.els.net. [DOI: 10.1002/9780470015902.a0021223]

Archie, J. 1989. Homoplasy excess ratios: New indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. SystZool. **38**:253-269.

Archie, J. 1996. Measures of homoplasy. Pages 153-188 *in* M. J. Sanderson and L. Hufford, editors. Homoplasy: the recurrence of similarity in evolution. Academic Press, San Diego.

Blomberg, S. P., and T. Garland, Jr. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. J. Evol. Biol. 15:899-910.

Blomberg, S. P., T. Garland, Jr, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution **57**:717-745.

Burt, B. B. 2001. Evolutionary stasis, constraint and other terminology describing evolutionary patterns. Biol. J. Linn. Soc. **72**:509-517.

Coddington, J. 1994. The roles of homology and convergence in studies of adaptation. Pages 53-78 *in* P. Eggleton and R. Vane-Wright, editors. Phylogenetics and ecology. Academic Press, London.

de Queiroz, K. 1996. Including the characters of interest during tree reconstruction and the problems of circularity and bias in studies of character evolution. AmerNat. **148**:700-708.

Desdevises, Y., P. Legendre, L. Azouzi, and S. Morand. 2003. Quantifying phylogenetically structured environmental variation. Evolution **57**:2647-2652.

Diaz-Uriarte, R. & Garland, T. (1996) Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. *Systematic Biology,* **45,** 27-47.

Diniz-Filho, J. A. F., C. E. Ramos de Sant'Ana, and L. M. Bini. 1998. An eigenvector method for estimating phylogenetic inertia. Evolution **52**:1247-1262.

Estes S and Arnold SJ. 2007. Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence at all timescales. Amer. Nat. 169: 227-244.

Farris, J. 1989. The retention index and the rescaled consistency index. Cladistics **5**:417-419.

Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of the evidence. Amer. Nat. **160**:712-726.

Garland, T., Jr., P. H. Harvey, and A. R. Ives. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst. Biol. **41**:18-32.

Gittleman, J., and H.-K. Luh. 1992. On comparing comparative methods. Ann. Rev. Ecol. gittSyst.

23:383-404.
Gittleman, J., and H.-K. Luh. 1994. Phylogeny, evolutionary models and comparative methods: a simulation study. Pages 103-122 *in* P. Eggleton and R. Vane-Wright, editors. Phylogenetics and ecology. Academic Press, London.
Gittleman, J., and M. Kot. 1990. Adaptation: statistics and a null model for estimating phylogenetic effects. Syst. Zool. **39**:227-241.
Givnish, T. 1997. Adaptive radiation and molecular systematics: issues and approaches. Pages 1-54 *in* T. J. Givnish and K. J. Sytsma, editors. Molecular evolution and adaptive radiation. Cambridge UnivPress, New York.
Hansen, T. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341-1351.
Hansen, T. F., and S. H. Orzack. 2005. Assessing current adaptation and phylogenetic inertia as explanations of trait evolution: The need for controlled comparisons. Evolution 59:2063-2072.
Harmon, L. J., J. A. Schulte II, A. Larson, and J. B. Losos. 2003. Tempo and mode of evolutionary radiation in Iguanian lizards. Science 301:961-964.
Harvey, P. H., and T. H. Clutton-Brock. 1985. Life history variation in primates. Evolution **39**:559-581.
Harvey, P., A. Read, and S. Nee. 1995. Further remarks on the role of phylogeny in comparative ecology. JEcol. **83**:733-734.
Harvey, P., A. Read, and S. Nee. 1995. Why ecologists need to be phylogenetically challenged. JEcol. **83**:000-000.
Kochmer, J., and S. Handel. 1986. Constraints and competition in the evolution of flowering phenology. EcolMonogr. **56**:303-325.
Leroi, A. M., M. R. Rose, and G. V. Lauder. 1994. What does the comparative method reveal about adaptation? Amer. Nat. **143**:381-402.
Lord, J., M. Westoby, and M. Leishman. 1995. Seed size and phylogeny in six temperate floras: constraints, niche conservatism, and adaptation. Amer. Nat. **146**:349-364.
Martins, E. 1996. Phylogenies, spatial autoregression, and the comparative method: a computer simulation test. Evolution **50**:1750-1765.
McKitrick, M. C. 1993. Phylogenetic constraint in evolutionary theory: Has it any explanatory power? Ann. Rev. Ecol. Syst. **24**:307-330.
Morales, E. 2000. Estimating phylogenetic inertia in Tithonia (Asteraceae): A comparative approach. Evolution **54**:475-484.
Orzack, S. H., and E. Sober. 2001. Adaptation, phylogenetic inertia, and the method of controlled comparisons. Pages 45-63 *in* S. H. Orzack and E. Sober, editors. Adaptationism and optimality. Cambridge University Press, Cambridge.
Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. Science **285**:1265-1267.
Rees, M. 1995. EC-PC comparative analyses. JEcol. **83**:891-892.
Revell, L. J., L. J. Harmon, and D. C. Collar. 2008. Phylogenetic signal, evolutionary process, and rate. Systematic Biology 57:591-601.
Revell, L. J., L. J. Harmon, R. B. Langerhans, and J. J. Kolbe. 2007. A phylogenetic approach to determining the importance of constraint on phenotypic evolution in the neotropical lizard Anolis cristatellus. EVOLUTIONARY ECOLOGY RESEARCH 9:261-282.
Revell, L. J., and L. J. Harmon. 2008. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. EVOLUTIONARY ECOLOGY RESEARCH 10:311-331.
Ridley, M. 1992. Darwin sound on comparative method. TREE **7**:37.
Sanderson, M. 1991. In search of homoplastic tendencies: Statistical inference of topological patterns in homoplasy. Evolution **45**:351-358.
Starck, J. 1998. Non-independence of data in biological comparisons a critical appraisal of current concepts assumptions and solutions. Theory Biosci. **117**:109-138.
Wanntorp, H.-E. 1983. Historical constraints in adaptation theory: traits and non-traits. Oikos **41**:157-160.

Westoby, M., M. Leishman, and J. Lord. 1995. Further remarks on phylogenetic correction. JEcol. **83**:727-729.
Westoby, M., M. Leishman, and J. Lord. 1995. Issues of interpretation after relating comparative datasets to phylogeny. J. Ecol. **83**:892-893.
Westoby, M., M. Leishman, and J. Lord. 1995. On misinterpreting the "phylogenetic correction". JEcol. **83**:531-534.

Appendix. More on independent contrasts and phylogenetic signal

Be sure to review lecture notes from last time on independent contrasts - portions are cut and paste below, but not all

• Tests of adaptive hypotheses are often conducted using comparative, interspecific data, e.g. do larger animals have larger home range? Do larger trees have larger seeds? Is small leaf size associated with long leaf life span? In some cases these hypotheses are explicitly evolutionary, which is to say the underlying hypothesis addresses patterns of trait change arising from natural selection: as body size increases (evolutionarily) does range size also increase. In this case, one can restate the hypothesis as a test of correlations in evolutionary changes (if we knew the history of trait change).

In other cases, the hypothesis is more ecological and functional: among present day species, is seed dispersal greater for taller trees? In this case, it is less obvious that an evolutionary or phylogenetic framework is needed. However, even here if these traits have evolved through descent with modification, such that close relatives do exhibit some degree of trait similarity, then tests using species data may encounter statistical issues arising from the non-independence of species values. In either case, comparative methods, especially independent contrasts, are an important tool to incorporate phylogenetic knowledge and use it to improve hypothesis tests, from both historical and statistical perspectives.

One of the basic principles of standard statistical tests is that the residuals in the data, after fitting to a statistical model, are independently and identically distributed (IID). The relatedness among species creates problems for the independence part, and the differences in branch lengths connecting the species affect the identically distributed part.

The degree of non-independence in species values is a function of how much shared history species exhibit, in the context of the clade being studied. Shared history is simply measured by the ratio of shared branch lengths to the total length from root to tips. This can be visualized on the tree and represented in matrices of phylogenetic distances (D, distance down to MRCA and back up to another species) and phylogenetic covariance (shared history = C = 1 - D/max(D)).

• Conventional statistical hypothesis testing requires that the residuals of the observations, after fitting a given statistical model, are *identically and independently distributed (IID)*

• The phylogenetic covariance resulting from shared history clearly violates the independent part.

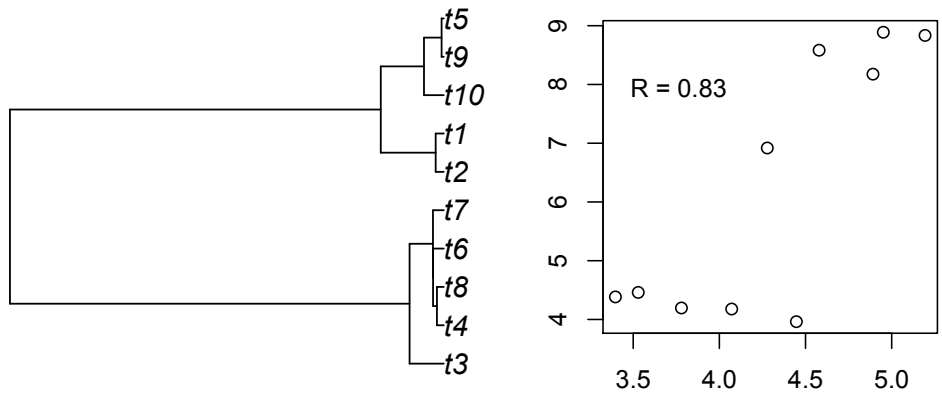
• So, if you calculate a correlation of (X,Y) where X and Y are trait values for a set of species, the shared history means that there is an elevated chance (of some magnitude) that related

traits evolved
independently

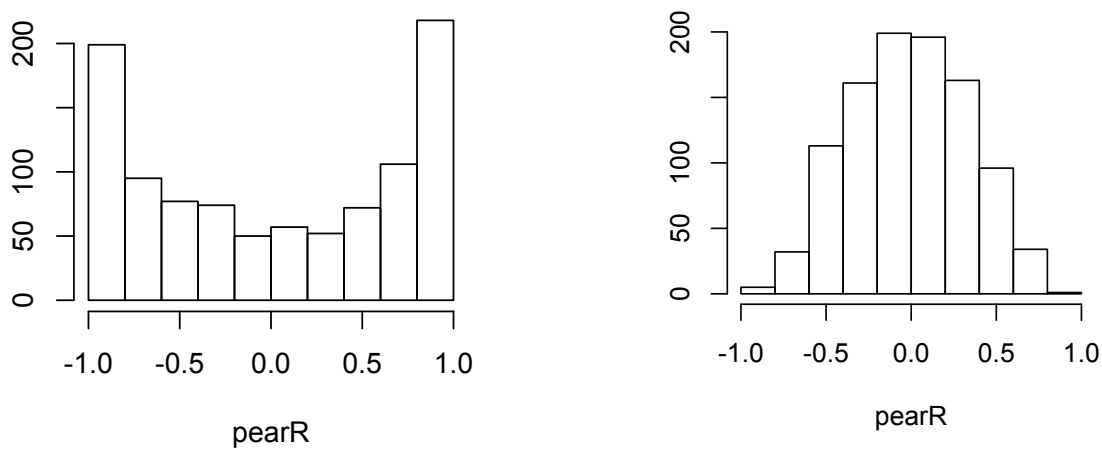
species will share similar trait values *for both traits*, even if they evolved independently on the tree, with no underlying correlation. Since they share trait values for both traits, we may incorrectly reject the null hypothesis leading to elevated type I error.

• Example:
1) Simulate the independent evolution of 2 traits on this tree:

2) Calculate the pearson correlation coefficient between the two traits. Here is one example:

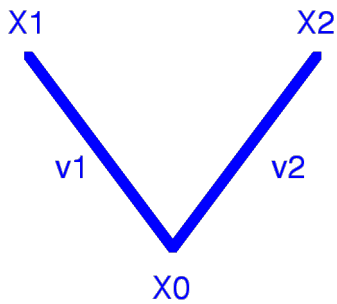


3) Repeat 1000 times and look at the distribution of correlation coefficients. In contrast, the expected distribution, for random data sets with N = 10 is shown on the right. For N = 10, the critical significance value at p <= 0.05 is 0.63; the type I error rate is a whopping 60%!



• Independent contrasts to the rescue!

The independent contrasts method is derived from the Brownian Motion model. Let's start with a single divergence:



Assume a trait of interest starts with value X_0 at the ancestor. After speciation, the trait evolves independently along each branch, with a Brownian Motion parameter s^2 (the variance of expected character change per unit branch length). If the length of the subtending branches are v_1 and v_2, then the changes that occur on each branch will be drawn from normal distributions with mean = 0 and variance = $s^2 \bullet v_1$ and $s^2 \bullet v_2$, respectively. The resulting trait values will be:

$$X1 = X0 + N\{0, s^2 \bullet v1\}$$
$$X2 = X0 + N\{0, s^2 \bullet v2\}$$

The shared X0 value in the distribution for both traits reflects the covariance issue between trait values for related species.

So, if we could somehow address the *changes* that occur along each branch, rather than the observed trait values of the species themselves, then we would have independent estimates of evolutionary change, and we could ask if there are correlated changes between two traits. In other words, we would like to know:

$$X1 - X0$$
$$X2 - X0$$

But, we don't know X0. And if we use only comparative data, we can only estimate it from the values of X1 and X2, so any estimate we got would not be independent of those data. Also, a tree with N species has 2N-2 internal branches. So it seems there's something wrong with increasing the number of data values than we start with.

Felsenstein's solution is to calculate the difference between X1 and X2, reflecting the cumulative divergence in trait values from their common ancestor, without necessarily knowing how much of the change occurred along each branch. So we want to know the expected distribution of the difference or contrast in the resulting trait values. We can take advantage of the fact that the variance of the sum *or* difference of two random variates is the sum of the variances. Therefore,

$$C_U = (X1 - X2) = N\{0, s^2 (v1+v2)\}$$

In other words, the mean value of the contrast will be zero, on average. And the variance around this mean will be the rate parameter s^2, times the sum of the branch lengths connecting the two taxa.

This value is termed the unstandardized contrast, because it is not standardized for the branch lengths. If two taxa are connected by long branches, we expect larger contrasts, and if they share

a more recent common ancestor, we expect smaller contrasts. Statistical theory is developed on the assumption that the variates are independent, and are drawn from normal distributions with equal variance or standard deviation. To achieve this, we now standardize the contrast, by dividing by the standard deviation of the expected distribution (since the rate parameter beta is constant over the tree, and only relative values of the contrasts are important, it is usually dropped at this step and the contrast is just standardized by the summed branch lengths):

$$C_S = \frac{C_U}{\sqrt{v_1 + v_2}}$$

In theory, if the trait has evolved following Brownian Motion, the values of C_S now represent a set of independent, normal variates drawn from distributions of equal variance, and these meet the assumptions for use in standard statistical tests such as correlations, regressions, and analysis of variance.

The calculation of contrasts, as illustrated above, require trait values at each pair of nodes sharing a common ancestor. Starting at the tips, these are the species values provided by the data set. Moving down the tree, however, we need to calculate a set of estimated internal values to obtain the appropriate contrasts. These are based on the local (top-down) maximum likelihood algorithm—essentially a weighted average of the nodes above, using the inverse of the branch lengths as weighting factors:

$$X_0 = \frac{\left(1/v_1\right)X_1 + \left(1/v_2\right)X_2}{1/v_1 + 1/v_2}$$

As a final step, the branch length below node 0 is lengthened, reflecting the increased uncertainty associated with deeper nodes (since a longer branch results in higher variance):

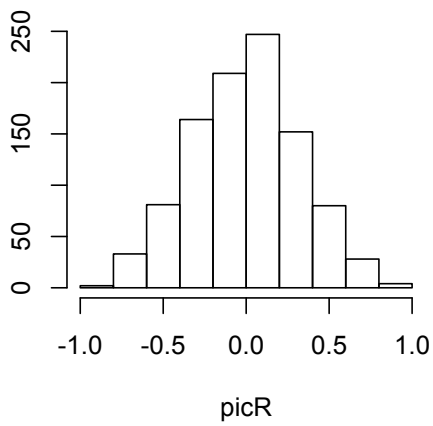
$$v'_0 = v_0 + v_1 v_2 (v_1 + v_2)$$

With the new internal node values calculated, the calculation of contrasts proceeds down the tree. These internal values are essentially the local (not global) maximum likelihood ancestral states based only on information from descendents, not the totality of information obtained when considering both descendents and other taxa in the tree, as in the full ancestral state algorithms we discussed previously.

II.C. Correlations of independent contrasts: Before we conduct a statistical test with independent contrasts, note that they have one unusual property. Each contrast is based on subtraction of one value from another. Clearly, the direction of subtraction is arbitrary, as long as it is kept the same for each trait in the study. As a result, each contrast has a mirror image of the opposite sign. So clearly the average value of all contrasts must be zero, since each one could be flipped around. As a result, all correlations and regression analyses must be calculated through the origin, and anovas would have to be conducted without a grand mean term. The formula for the correlation coefficient through the origin is a bit different than the familiar one in a stats textbook:

$$r_{xy} = \frac{\sum C_x C_y}{\left\{\sum C_x^{\,2} \sum C_y^{\,2}\right\}^{1/2}}$$

where C_x and C_y are the standardized contrasts for traits x and y. Because we are not estimating the mean, the degrees of freedom for independent contrasts is N_c - 1, where N_c is the number of contrasts. Remember that $N_c = N_t - 1$, where N_t is the number of taxa, assuming it's a fully bifurcating tree. See Garland et al. 1992 for additional discussion.



Returning to our example above with the 10 taxon tree, here is the distribution of independent contrasts under the null hypothesis. Type I error for 1000 reps is 0.051% - perfect!

II. General linear models

There is an alternative approach to this problem, which we will not address in detail, but it is important for you to be aware of it. Conventional linear models that we use in statistics are a special case of a more general class of general linear models, or generalized least squares. In these general models, one can relax the assumption that residuals are identically and independently distributed, and instead include in the model a correlation matrix that specifies the expected pattern of non-independence among residuals. The model is then 'corrected' for this expected non-independence, and provides appropriate results for hypothesis tests. For the phylogenetic case, the phylogenetic covariance matrix (p. 2 above) can be used for the error covariance structure. This then allows one to use the full array of linear models, including anova, regression, ancova, etc., opening up possibilities beyond what can be done with independent contrasts. For the case of two continuous variables, it can be shown (though not by me!) that the results converge and the independent contrasts and generalized least squares approaches are the same.

In *ape*, the gls approach looks like this, where *x* and *y* are the dependent and independent variables, contained in data.frame *dat*:

```
m1 <- gls(y~x,data=dat, correlation=corBrownian(1,phy))
```

For the null model on the ten taxon tree in the example above, the type I error rate was exactly the same as the independent contrasts.