

Phylogenetic Pseudoreplication Scenarios

Excerpt from:

Uyeda, J. C., Zenil-Ferguson, R. & Pennell, M. W. (2018) Rethinking phylogenetic comparative methods. *Systematic Biology* **67**, 1091–1109.

Every so often, evolution comes up with something totally new and unexpected, a so-crazy-it-just-might work set of adaptations that is the stuff of nature documentaries. Many biologists likely have a favorite example of a lineage that has evolved something spectacular such as devilishly horned lizards that squirt blood from their eye sockets or marine sloths that grazed ancient seabeds.

As macroevolutionary researchers, it is hard to know what to do with these types of events (Vermeij 2006). Their singular and unreplicated nature seems incompatible with models that we typically use to describe change over time, such as Brownian motion (BM; Felsenstein 1973) or the Mk model (Pagel 1994; Lewis 2001). Such models presume continuity, whereas one-off events, such as the evolution of novel nutritive function in exocrine glands leading to mammalian milk, have no clear precedent in history. The evolution of such traits may set in motion a cascade of changes across an organism, such that descendant lineages may look very different in many ways from their more distant relatives. Or alternatively, a suite of traits may just happen to change at the same time. In either case, it is these sorts of idiosyncratic and unreplicated events that we often think of when we think of the need to consider phylogeny in analyses of comparative data. And this is not an abstract concern; a wide breadth of macroevolutionary data suggest that abrupt shifts and discontinuities have been a major feature of life on Earth (Uyeda et al. 2011, 2017; Landis and Schraiber 2017; Jablonski 2017). But as recent controversies in phylogenetic comparative biology have highlighted, our current methods (reviewed in O’Meara 2012; Pennell and Harmon 2013; Garamszegi 2014) are not designed to deal with such dynamics.

For example, Maddison and FitzJohn (2015) recently demonstrated that common statistical tests (e.g., Maddison 1990; Pagel 1994) for the evolutionary correlation of discrete characters are prone to reporting a significant association even when the pattern is driven by a single (or, very few) independent transition(s) from one character state to another. Maddison and FitzJohn (2015) referred to such scenarios as cases of “phylogenetic pseudoreplication” (see also Read and Nee 1995; Nee et al. 1996).

We will argue that this unresolved challenge permeates not just tests for discrete character correlations, but nearly every method of finding associations in comparative methods (Fig. 1). For example, Rabosky and Goldberg (2015) show that applying trait-dependent diversification models (e.g., BiSSE, Maddison et al. 2007) to real-world phylogenies, which are usually not shaped like trees resulting from simulations of a birth–death stochastic process (Mooers and Heard 1997), often leads to support for trait-dependent diversification models regardless of whether traits are actually affecting speciation and extinction. [...]

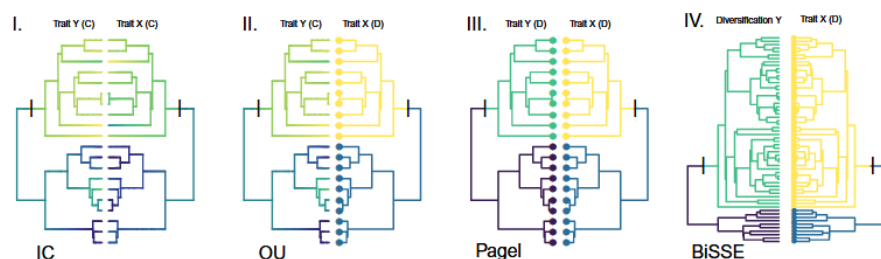


FIGURE 1. Singular, unreplicated events (vertical dashes) can generate apparently significant associations across several types of comparative analyses. Case Studies I–III are indicated in panels I–III, and though we do not consider diversification models such as BiSSE in our examples, they are similarly affected (panel IV). In each case, we map (in some cases, arbitrarily) the dependent variable (Y) on the phylogeny on the left and the predictor trait on the same phylogeny to the right (X), and indicate whether the trait is a continuous trait (C), a discrete trait (D) or a diversification rate. Colors on the branches indicate the state of the character on the phylogeny—either continuous trait value, discrete character state, or diversification rate regime. Panels I and III correspond to variations of “Felsenstein’s worst-case scenario” and “Darwin’s scenario,” respectively. We also suggest a common method used to analyze such associations: IC = Independent Contrasts (Felsenstein 1985); OU = Ornstein–Uhlenbeck models (Butler and King 2004); Pagel = Pagel’s correlation test (Pagel 1994).

More than anything else, it was the famous series of figures depicting the “worst-case scenario” (Figs. 5, 6, and 7 in the original; our Fig. 2) from Felsenstein’s iconic 1985 article “Phylogenies and the comparative method” that awakened biologists to the need for tree-thinking and started a revolution in modern comparative biology. The idea is simple: as a result of shared ancestry, measurements taken on one species will not be independent from those collected on another and especially so, if the two species are closely related. This nonindependence can create apparent correlations between traits that, are in truth, evolving independently. To illustrate the effect of nonindependence of characters, Felsenstein generated a scenario in which two clades are separated by long branches (our Fig. 2). He then evolved traits according to a BM process along the phylogeny; he recovered a significant regression slope using Ordinary Least Squares (OLS) despite there being no evolutionary covariance between the traits. [...]

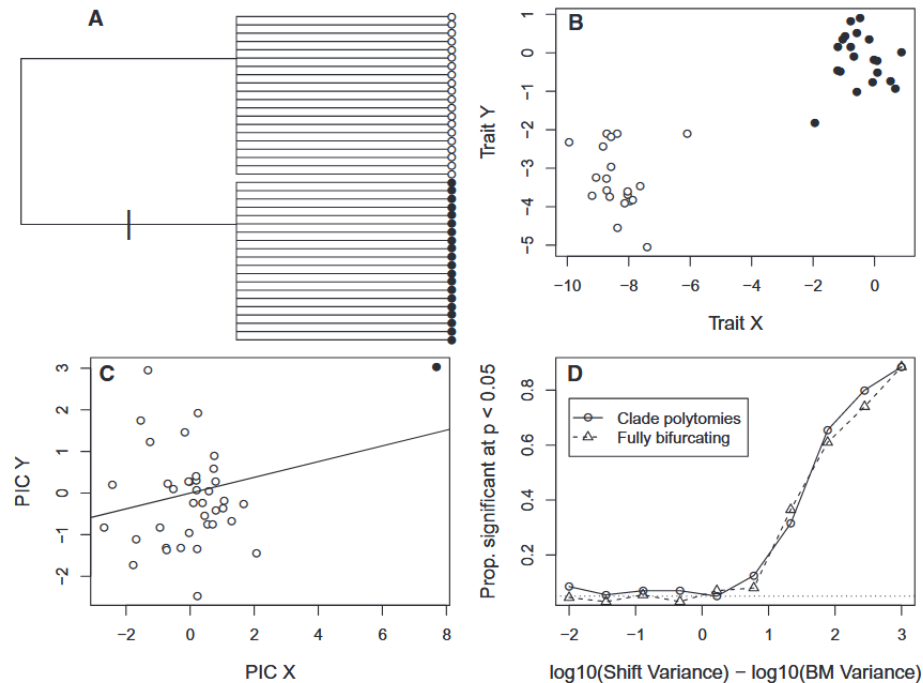


FIGURE 2. Felsenstein’s worst-case scenario (Felsenstein 1985) illustrates a problem quite like that identified by Maddison and FitzJohn. Here we modify Felsenstein’s original generating process from simple BM, to A) BM with a single burst occurring on the stem branch of one of the two clades (indicated by vertical dash). B) The distribution of trait values produces a figure very similar to Felsenstein’s original scenario, but results in C) a single contrast (black) that is not well-described by the estimated BM process, and thereby generates a significant regression of PIC Y and PIC X (dotted line) despite both X and Y in the shift and BM distributions being uncorrelated. D) As the ratio of the shift variance to the BM variance increases, the proportion of contrast regressions that return a significant result increases dramatically (each point represents 200 simulations for a fixed phylogeny, with both the BM process and the random draw from the shift distribution being uncorrelated with equal variance for both traits). While IC corrects for singular events consistent with BM, it does not correct for the more general phenomenon of dramatic singular events driving significant results in comparative analyses. Note that the nonindependence of species is not the issue.

As we mention above, Maddison and FitzJohn (2015) recently demonstrated that commonly used methods return significant correlations all the time—and in scenarios that seem to defy our statistical intuition. For example, Pagel’s (1994) correlation test would find the phylogenetic codistribution of milk production and middle ear bones highly statistically significant even though they both are a defining characteristic of mammals, an inference so obviously dubious that even Darwin (1872) warned against it. This seems to be a clear case of phylogenetic pseudoreplication (Read and Nee 1995; Maddison and FitzJohn 2015). More broadly, Maddison and FitzJohn describe the goal of correlation tests as finding the “weak” conclusion that “the two variables of interest appear to be part of the same adaptive/functional network, causally linked either directly, or indirectly through other variables” (p. 128). They assert that with our current approaches, we cannot even clear this (arguably low) bar. [...]

Maddison and FitzJohn highlight two hypothetical situations, that they refer to as “Darwin’s scenario” and an “unreplicated burst.” They argue that these scenarios provide little evidence for an adaptive/functional relationship between two traits because the patterns of codistribution only reflect singular evolutionary events

(Fig. 1). In Darwin's scenario, two traits are coextensive on the phylogeny, meaning that in every lineage where one trait is in the derived character state, the other trait is as well. As an example, consider the aforementioned phylogenetic distribution of middle ear bones and milk production in animals; all mammals (and only mammals) have middle ear bones and produce milk. These traits (depending on how they are defined) have only appeared once on the tree of life and both occurred on the same branch (the stem branch of mammals). The unreplicated burst scenario is identical to Darwin's scenario except that rather than a single transition occurring in both traits, there is a single transition in the state of one trait (e.g., the gain of middle ear bones) and a sudden shift in the transition rates in another trait (e.g., the rates by which external testes are gained and lost across mammals). Note that these scenarios do not differ qualitatively from Felsenstein's worst-case scenario [...] (Fig. 1). In all three scenarios, something novel and interesting happened on a single branch and the distribution of traits at the tips of the phylogeny reflects this.