LETTER

Morphogenera, monophyly, and macroevolution

Jablonski and Finarelli (1) suggest that morphogenera, even when they are nonmonophyletic, serve as good representatives for large-scale evolutionary studies. We feel there are two issues that warrant further discussion. First, the test used to evaluate the effect of using nonmonophyletic groups for macroevolutionary studies was not conservative and thus does not provide strong evidence about the impact of nonmonophyly on evolutionary studies. Their test examined whether the median trait value for the species in a nonmonophyletic genus correlated with the median trait value for those same species plus the additional species needed to make the set monophyletic. Correlation of the median value of a set of measurements with the median value of a superset of those measurements is to be expected (as is briefly mentioned in the methods section of ref. 1). To demonstrate this, we correlated the median values of *ln* body mass from random sets of 3, 7, and 15 mammal species and 0–30 additional species for 45 simulated genera, using data from ref. 2 (Fig. 1*A*). Even when the median of 3 species is correlated with those 3 plus 15 additional species, the correlation is significant. Moreover, slightly better than random assignment of species to genera increases the expected correlation. To show this, we compare the results of figure 3*A* from ref. 1 with simulations where 45 genera of 3 species each were created from species chosen randomly from (*i*) the same family (Fig. 1*B*) and (*ii*) the same order (Fig. 1*C*) (data from refs. 2 and 3). Correlations between morphogenus median and the smallest clade containing the morphogenus species were significant ($P < 0.0001$). Thus, the observed correlations presented in ref. 1 are not surprising.

Second, ref. 1 also appears agnostic as to whether the morphogenera should be used together with a phylogeny when

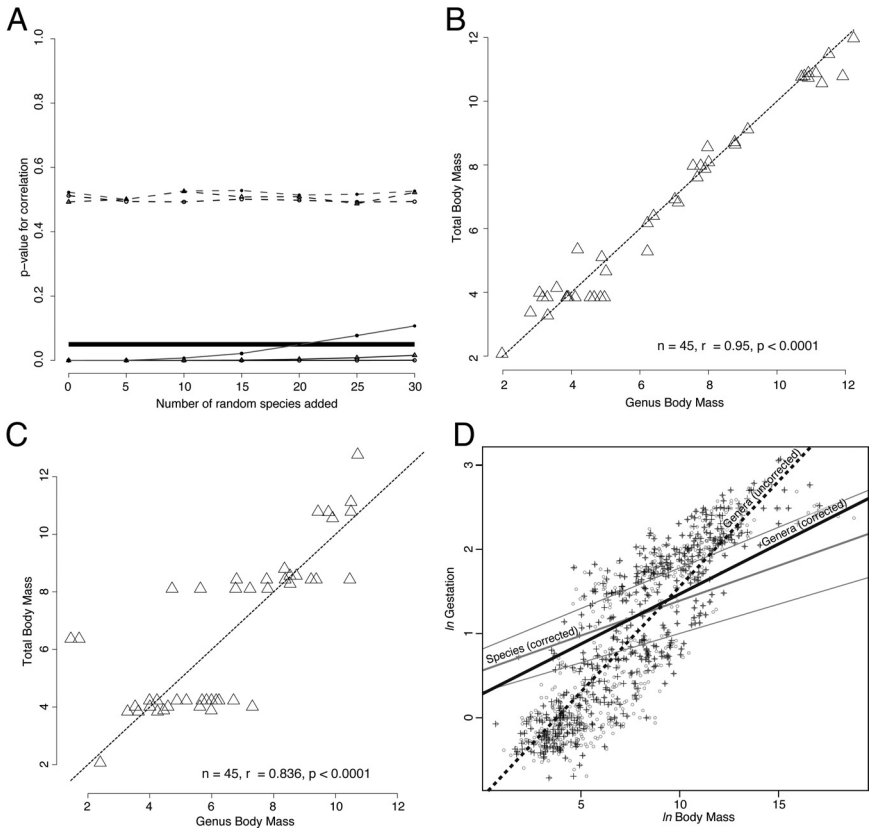

Fig. 1. Correlation analyses of morphogenera and comparative analysis of body mass and gestation. (*A*) Significance (Spearman correlation, as in ref. 1) of correlation of means of *ln* body mass of a set of *X* species with a set of $X + Y$ species (solid lines), where *X* is 3 (closed circles), 7 (triangles), or 15 (open circles) and *Y* is 0, 5, 10, 15, 20, 25, 30, where the larger set includes all the species in the smaller set. Dotted lines correspond to correlating a set *X* with a set $Z + Y$ species, where the size of *Z* and *X* are the same, but the species differ. The dark solid line represents $P = 0.05$. For each point, 2,000 replicates, each with 45 randomly generated sets, were used. The masses came from ref. 2, pruned down to the set of species also in the tree of ref. 3. (*B* and *C*) Plots of *ln* body mass 45 simulated morphogenera, as in figure 3A of ref. 1, where genera consist of 3 species chosen at random from within mammal families (*B*) or orders (*C*) (species chosen without replacement within genera but with replacement between genera, and genera simulated until 45 nonmonophyletic ones were recovered). The *x* axis shows the median value for the morphogenus, the *y* axis shows the median for the smallest clade containing all of the species (based on the tree of ref. 3). No distinction between polyphyletic or paraphyletic genera was attempted (open triangle used for both). Otherwise, plot and analysis as in figure 3A of ref. 1. (*D*) Generalized least-squares analyses of *ln* gestation period against *ln* body mass for all mammal species and genera. The thick black line shows the fit of these variables for all genera. The medium grey line shows the fit for all mammal species, and the light grey lines show the 95% confidence interval. The thick dashed line shows the fit for the genera without correcting for the phylogeny. Grey dots represent data using species-level values and black crosses represent generic-level data.

conducting large-scale evolutionary analyses or whether the morphogenera would themselves be independent data points (as in some of the papers it cites). As discussed in ref. 4, accounting for the bias of shared ancestry in analyses of biological data is essential. We illustrate the importance of using phylogenies by examining the correlation of body size and gestation in mammals (data from refs. 2 and 3). First, we conducted a generalized least-squares analysis (taking into account phylogeny) of gestation on body size by using species-level data [slope, 0.083 (0.07–0.096, 95% confidence interval); intercept, 0.561 (0.306–0.815, 95% confidence interval); Fig. 1*D*]. Then, we examined the same relationship by using the phylogenetically corrected generic-level data (slope, 0.119; intercept, 0.282) and phylogenetically uncorrected generic level data (slope, 0.244; intercept, −0.851). Both the phylogenetically corrected and uncorrected generic-level analyses result in intercepts and slopes outside the 95% confidence interval of the species level dataset. This demonstrates that generic-level data may mislead large-scale evolutionary studies, and so using morphogenera rather than species should be done with caution. More importantly, the worse results achieved by the generic-level analysis when not accounting for phylogeny (a slope 3 times greater than the slope estimated from all the data) suggest that, even if morphogenera are used, correction based on phylogeny must be considered.

Stephen A. Smith and Brian C. O'Meara[1]
National Evolutionary Synthesis Center, 2024 West Main Street A200, Durham, NC 27705

1. Jablonski D, Finarelli JA (2009) Congruence of morphologically defined genera with molecular phylogenies. *Proc Natl Acad Sci USA* 106:10057–10060.
2. Ernest SKM (2003) Life history characteristics of placental nonvolant mammals. *Ecology* 84:3402.
3. Bininda-Emonds, et al. (2007) The delayed rise of present-day mammals. *Nature* 446:507–512.
4. Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1–15.

Author contributions: S.A.S. and B.C.O. designed research, performed research, contributed new reagents/analytic tools, and wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: sasmith@nescent.org.