How should functional relationships be evaluated usingphylogenetic comparative methods? A case study using metabolicrate and body temperature

This repository contains code and data to reproduce all analyses in our note response letter "How should functional relationships be evaluated usingphylogenetic comparative methods? A case study using metabolicrate and body temperature". As background, our case study re-examines the results of the article entitled: [The decoupled nature of basal metabolic rate and body temperature in endotherm evolution].(https://www.nature.com/articles/s41586-019-1476-9).

Supplementary methods

I. Phylogenetic analyses

The main file implementing the data analysis is in this section is R/phylogenetic_analyses.Rmd. We examined simple models of trait evolution for three traits from the Bird and Mammal datasets used by ALEA to determine whether the variable-rates regression model (VRRM), which is a variant of Brownian Motion, is suitable for these data. We examined 3 focal traits, Tb, Ta and mass-specific BMR. The former two are analyzed in degrees Celsius, but the latter is log transformed and taken as the residuals of the phylogenetic regression of lnBMR ~ lnMass, or in the case of mammals, lnBMR ~ lnMass + lnMass2, as previous work has shown curvature in this scaling relationship (Kolokotrones et al. 2010, Uyeda et al. 2017).

I.A Phylogenetic signal

For each of the 3 traits, we then fit a single optimum Ornstein-Uhlenbeck model to measure phylogenetic signal. We take the phylogenetic half-life, (Ln(2)/alpha) value where alpha is the "rate of adaptation" parameter of the OU model. If alpha = 0 as in BM, then the the half-life is infinite. If alpha is very large, then the half-life will approach 0. Practically, the half-life can be interpreted relative to tree height as the total amount of time it takes for phylogenetic signal to be eroded by the OU process, which as a mean-reverting process, will erase phylogenetic covariance over time. Specifically, half-life is intpreted as the amount of time it takes for a lineage to be expected to get halfway to the optimum. If this is very rapid, then two closely related species will not appear any more similar than any other pair of species. Whereas if it is very slow, then it is likely that these two species will share considerable similarity relative to the clade as a whole. We interpret half-life values from 0-0.25 units of tree height as very weak phylogenetic signal and/or strong evidence of constraint. Values from 0.25-0.75 we intrepret as moderate levels of phylogenetic signal. And values > 0.75 begin to approach BM-like evolution. We also compare the fit of the OU model, fit with the R package `phylolm` (using the `OUfixedRoot` option) to a Brownian Motion model using AIC.

In all cases we find strong evidence in favor of an OU model, with at least 46 AIC units separating the BM and OU models. Furthermore, all half-life values are <0.26, with all bird values actually appearing at their lower bounds of 0.01. This indicates that BM models are inappropriate to capture the process of evolution in these groups. In general, birds had much less variation and much less signal than equivalent traits in mammals, which have more variation across tree.

I.B Phylogenetic regressions showing "state-coupling"

We verified using phylogenetic regressions that Tb and lnBMR are positive correlated in both birds and mammals using models that include lnMass, and in the case of mammals, lnMass2. We found the best model of the set BM, `lambda`, `OUfixedRoot` and `OUrandomRoot` and fit these regressions in phylolm. In mammals, we find very strong correlations between Tb and lnBMR and significant negative correlations of Ta and Tb. In birds, the former is marginally significant and also positive, and the latter is non-significant. We note that very little variation exists for Avian Tb and that proper analysis of these correlations would likely require careful accounting of measurement error.

I.C Correlations of rates as evidence of "rate-coupling"

We then turn our attention to evolutionary rates. Rather than using the VRRM of ALEA, we take the absolute value of the independent contrasts at each node as an estimate of evolutionary rate, and log transform these for all 3 traits. We then test for a phylogenetic association between them. In general, we find that these relationships are weak, if present at all. However, depending on how the data are treated (e.g. how much of an offset is used for the log transformation of rates when the rate is estimated to be 0 change between species) we can obtain a p-value ranging from <0.10 to <0.05 in mammals. We take this as marginal evidence of a weak association of evolutionary rates between Tb and lnBMR.

I.D Pathwise-rates and branch lengths

Because we find strong support for an OU model, it is unsurprising that the node-height test of all traits supports a strongly positive slope between the recency of a node, and it's absolute contrast value (i.e. its rate). Under BM we expect this slope to be 0. However, because of constraints, older nodes and longer branches will be estimated to have lower rates. We note that OU models for ultrametric trees are mathematically equivalent to "Accelerating change" Brownian Motion models, in which the rate of evolution increases toward the present, for the same reasons. Thus, a VRRM model will assign higher rates to short terminal branches, simply as a mechanism for explaining what is likely OU-like evolution. If short terminal branches are more common in some state than others (either by trait-dependent diversification or because of bias in taxonomic sampling) then rates will be estimated to be faster for taxa with those states. We think this is likely for Ta, where we know that faster turnover rates in temperate zones results in shorter terminal branches than the "museum-like" tropics. Thus, even under a constant OU-process, we would estimate faster rates for low Ta. Turning to the numerator of rate, asymmetries in constraints also can result in faster rates at lower temperature, as we examine in more detail in our Figure 2. We emphasize that the accelerating change model is extremely unlikely for these traits, and that OU models and stabilizing selection is much more strongly supported as a mechanism for trait change from multiple lines of evidence.

II.A Causal models and "state-" vs. "rate-coupling"

To generate Figure 1 in our manuscript, we use simulations of causal models that we think are plausible hypotheses for interpreting the outcomes depicted in ALEA's Figure 1. The scripts for recreating these figures are in R/Figure1.Rmd. We use the full mammal phylogeny and simulate 3 scenarios. The First corresponds to a multivariate BM or PGLS model, which are known to be equivalent in estimation, but distinct in causal structure. We simulate no rate variation, so that all branch scalars are 1. However, we simulate a correlation of 0.8 and rates of 1 for each trait, and obtain a significant state-coupling. Next, we simulate uncorrelated states by changing the correlation in the evolutionary rate matrix to 0. However, we introduce correlations in rate by drawing shared branch scalars for both traits from a gamma distribution with shape parameter 1.1. This results in collinear evolutionary rates for both traits, but no correlation in states. We consider this type of "rate-coupling" evidence that a 3rd factor, Z, affects both traits, but believe it is not any sort of evidence of state-coupling, or functional relationship, between the two traits. Finally, we examine a third scenario. Here we first simulate Tb with branch scalars from a gamma distribution with parameters (10, 10) and a Brownian Rate of 1, and Z with independent branch scalars from the same distribution, but with a Brownian Rate of 20. lnBMR is generated by adding Z and Tb. Thus, Z has substantially greater variation in Tb and therefore a stronger effect on lnBMR. We view this as very likely. Tb is a trait known to be under stasis in endotherms and that is relatively non-responsive to environmental change and changing biology, likely due to the multivariate effects it has on enzymatic performance and physiology across an organism. On the other hand, traits such as body size and other residual effects on Metabolic Rate are exposed to strong varying selection forces in different clades. Thus, we expect that even if Tb is a direct cause of Metabolic Rate, rate correlations will be hard to detect. We reanalyze the empirical datasets using the full phylogenetic regression models described above and add them to our figure in Row 4 (see below, II.B).

II.B Empirical rate correlations

To validate that our approach for evaluating evidence or signal for rate correlations can detect shared rate heterogeneity to compare to the Variable-Rates Regression Model, we performed a correlation test on the log

absolute contrasts for lnBMR and Tb. Code and methods for performing this are presented in the following document. In brief, we simulated 50 replicates for 11 values of correlation (over a sequence of values from 0 to 1 by 0.1 units) between rate scalars, for a total of 550 simulations. We used the R package bayou to simulate shift locations with from a uniform distribution from 0 to the number of branches in the Fritz et al. mammal phylogeny. We then drew log rate scalars from a multivariate normal distribution with mean = (1.5, 1.5) and variances = (1.5, 1.5). These rate scalars were then exponentiated and mapped onto the shift locations to generate a phylogeny upon which independent Brownian Motion processes were used to simulate X and Y. The contrasts of these traits were then calculated on the original phylogeny, and the Pearson-correlation coefqficient was calculated. Because the branch scalars are reweighted number of branches that inherit their values, the exact correlation of the simulated values will not match the value used to simulate from the multivariate normal distribution. We therefore reestimate this value from the true rate scalars. We find that rate correlations of the magnitude we observed for mammals correspond to weak or moderate levels of evolutionary correlation (see below).
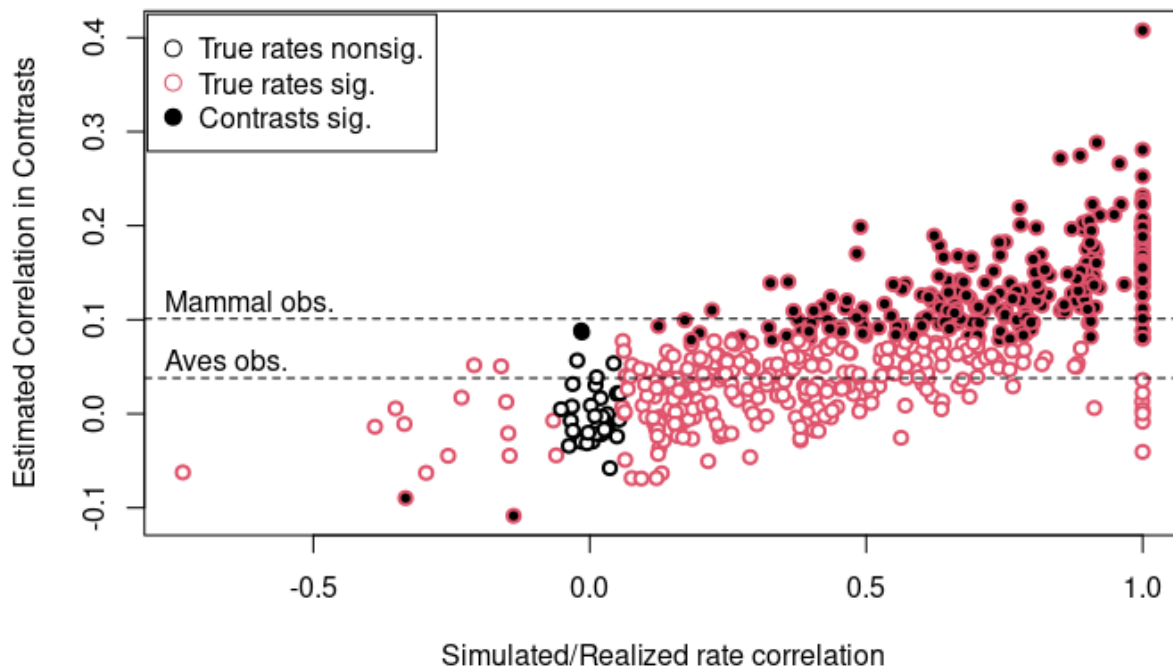
Supplementary Figure 1. Simulated values of contrast correlations from phylogenetically-distributed rate scalars demonstrates reasonable power to detect significant relationships between rates directly from contrasts in the presence of rate variation. Open circles indicate that the contrast Pearson-correlation coefficient was non-significant, whereas filled black circles indicate it is significantly correlated in X and Y. Red outlines indicate that the true correlation coefficient among branch rates are significant, whereas black outlines indicate even with perfect knowledge of the branch rates, the obesrved correlation was non-significant. Observed values from the empirical data are shown, while note that the data are simulated to match the Mammalian, not bird data.

III. Pathwise-rates are not proxies for ancestral state estimation for constrained traits

We address the conclusions drawn from ALEA Figure 4 through simulation. ALEA conclude from this analysis that body temperature and environmental temperature have decreased over the course of mammal and bird evolution due to the observation that the fastest pathwise rates are estimated from the lowest values of these traits. To demonstrate how biases in trait evolution or diversificaiton can generate such patterns, we simulate using the R package BBMV on a constrained macroevolutionary landscape for 10 simulations and

reanalyze with `bayestraits` and a customized implementation of the VRRM in the R package `bayou` for comparison. We initialize BBMV simulations with a starting point of 25C and set the macroevolutionary landscape peak at 38C. Thus, lineages evolve on average from 25C to 38C. We use a Sharpe-Schoolfield thermal performance curve to generate the macroevolutionary landscape under the assumption that the overall macroevolutionary landscape will reflect features and the general shape of the within-species thermal performance curves (but broader). We generate datasets where all values are >25C in all descendent species. In this BBMV simulation, bounds are set but only as a convenience for analysis and simulation. If lineages hit the bounds, we discard those runs. We parameterize the Sharpe-Schoolfield equation so that the resulting species roughly align with the empirical distribution of mammalian body temperatures, ranging from 29-42C. The salient feature of the Sharpe-Schoolfield equation is that the upper bounds of body temperature have a much steeper and harsh boundary than the lower limits, reflecting that high temperatures impose a much stronger constraint on enzymatic performance than low temperatures. Thus, more change accumulates at low temperatures, resulting in faster rates; despite an overall increasing trend. We verify this in our analyses by showing we can recover similar slopes as in the empirical analyses in our Figure 2.

We provide custom scripts to implement a variant of the VRRM in the R package `bayou`. This was done simply so that we could verify what was occurring in the analysis, and because initial testing with `bayestraits` was hindered by the requirement of using a server to process our batch results, which resulted in some difficulty. However, we were able to obtain comparable analyses between the two approaches that broadly aligned, with pathwise rates being strongly correlated (0.93) in the empirical analyses. We set a uniform prior on the number of shifts ranging from 0 to the number of branches on the tree. We set a half-Cauchy prior on the evolutionary rate parameter with a scale of 0.1 and a Normal prior on the root of 35C with a standard deviation of 10C. Thus, we allow for inference of both increasing and decreasing trends over time. We use Reversible-Jump MCMC estimated for 1 to 2 million generations and discard the first 30% of samples as burnin. We then take the median branch scalars and standardize them as in `bayestraits` to compare pathwise rates and trait values.