



# Causes and Consequences of Apparent Timescaling Across All Estimated Evolutionary Rates

Luke J Harmon, Matthew W Pennell, L. Francisco Henao-Diaz, Jonathan Rolland, Breanna Sipley, Josef C Uyeda

## ► To cite this version:

Luke J Harmon, Matthew W Pennell, L. Francisco Henao-Diaz, Jonathan Rolland, Breanna Sipley, et al.. Causes and Consequences of Apparent Timescaling Across All Estimated Evolutionary Rates. Annual Review of Ecology, Evolution, and Systematics, 2021, 52 (1), pp.587-609. 10.1146/annurev-eolsys-011921-023644 . hal-03822949

**HAL Id: hal-03822949**

**<https://hal.science/hal-03822949>**

Submitted on 20 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The causes and consequences of the ubiquitous time-scaling of all evolutionary rates

Luke J. Harmon<sup>1</sup>, Matthew W. Pennell<sup>2</sup>, L. Francisco Henao-Diaz<sup>3</sup>, Jonathan Rolland<sup>4</sup>, Breanna Siple<sup>5</sup>, and Josef C. Uyeda<sup>6</sup>

<sup>1</sup>Institute for Bioinformatics and Computational Biology and Department of Biological Sciences, University of Idaho, Moscow, ID, USA; email: [lukeh@uidaho.edu](mailto:lukeh@uidaho.edu); ORCID: <https://orcid.org/0000-0002-4985-5750>

<sup>2</sup>Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada; email: [pennell@zoology.ubc.ca](mailto:pennell@zoology.ubc.ca)

<sup>3</sup>Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada; email: [fhenao@zoology.ubc.ca](mailto:fhenao@zoology.ubc.ca); ORCID: <https://orcid.org/0000-0002-1975-4908>

<sup>4</sup>CNRS, UMR5174, Laboratoire Evolution et Diversité Biologique, Université Toulouse 3 Paul Sabatier, Bâtiment 4R1, 118 Route de Narbonne, 31062 Toulouse, France; email: [jonathan.rolland@yahoo.fr](mailto:jonathan.rolland@yahoo.fr); ORCID: <https://orcid.org/0000-0002-7141-0870>

<sup>5</sup>Program for Bioinformatics and Computational Biology, University of Idaho, Moscow, ID, USA; email: [breanna.siple@gmail.com](mailto:breanna.siple@gmail.com); ORCID: <https://orcid.org/0000-0002-2475-5630>

<sup>6</sup>Department of Biological Sciences, Virginia Tech University, Blacksburg, VA, US; email: [juyeda@vt.edu](mailto:juyeda@vt.edu)

## Abstract

Evolutionary rates play a central role in connecting micro- and macroevolution. All evolutionary rate estimates, including rates of molecular evolution, trait evolution, and lineage diversification, share a similar scaling pattern with time: the highest rates are those measured over the shortest time interval. This creates a disconnect between micro- and macroevolution, although the pattern is the opposite that some might expect: patterns of change over short time scales predict that evolution has tremendous potential to create variation, and that potential is barely tapped by macroevolution. In this review, we discuss this shared scaling pattern across evolutionary rates. We break down possible explanations for scaling into two categories, estimation error and model misspecification, and discuss how both apply to each type of rate. We discuss the consequences of this ubiquitous pattern, which can lead to unexpected results when comparing rates over different time scales. Finally, we discuss a few possibilities for a shared underlying explanation across the three types of rates.

# Keywords

Speciation, extinction, genetics, phenotypes, microevolution, macroevolution

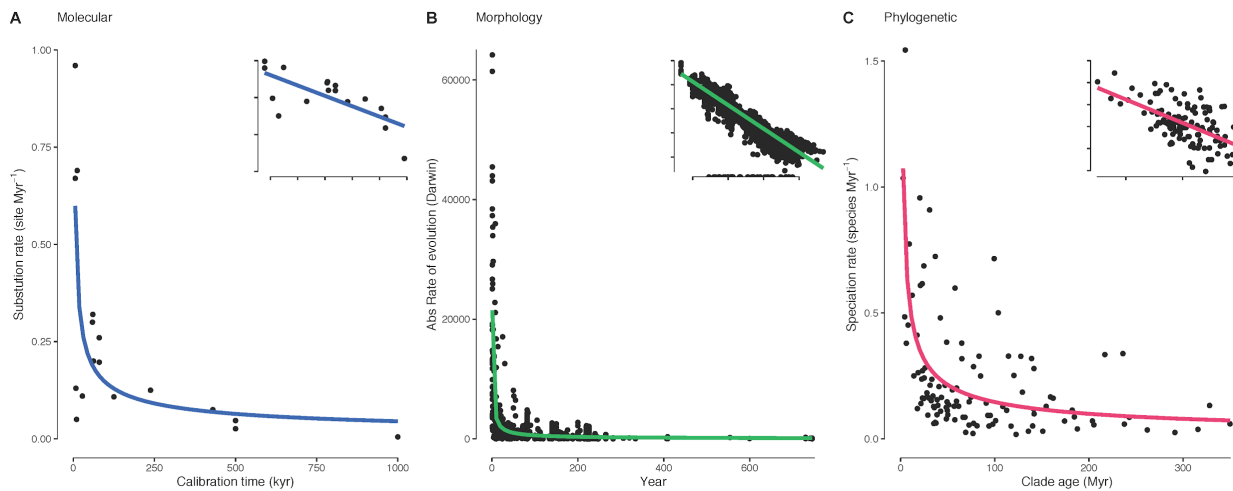
## Introduction

Evolutionary rates play a central role in the quest to bridge micro- and macroevolution (Gingerich 1983, 2009; Hansen and Martins 1996; Rolland et al. 2018). The central question is whether rates of microevolutionary changes that we measure over a few generations are sufficient to explain the diversity of life that has accumulated on the planet over the past 4 billion years (Kinnison and Hendry 2001). Or, in slightly different terms, can we explain all of the diversity of life via models of population and quantitative genetics, extrapolated along the branches of the tree of life (Charlesworth et al. 1982; Estes and Arnold 2007)? The opposing point of view is that we need special macroevolutionary phenomena, such as species selection or times of dramatically accelerated diversification, to explain broad-scale phenomena (Gould and Eldredge 1972; Erwin 2000; Hendry and Kinnison 2001; Reznick and Ricklefs 2009; Rolland et al. 2018). Measuring and comparing rates of evolution across time scales could provide some resolution for this debate.

Empirical measurements of evolutionary rates have revealed a remarkable consistent pattern of scaling across micro- and macro-evolutionary time. Rate estimates depend predictably on the time scale over which they are measured, with the highest rates measured over the shortest time interval (Ho et al. 2011; Uyeda et al. 2011; Francisco Henao Diaz et al. 2019). Across rates of molecular, trait, and lineage diversification, the fastest rates tend to be measured over short time intervals, while rates measured over longer time intervals tend to be slower (Figure 1). As a consequence, patterns of change over short time scales predict that evolution has tremendous potential to create variation, and that potentially is barely tapped by the present-day diversity of life (Gould 1985). In other words, the short-term potential of evolution, in terms of rates of change of DNA sequences, traits, and the accumulation of species, is much

greater than what is manifest over long time scales. The causes of this ubiquitous scaling pattern remain mysterious, and pose a challenge to the unification of evolutionary studies.

There is one obvious candidate for an explanation of this scaling pattern: perhaps evolution is constantly accelerating. In a world where evolution is speeding up through time, we are making our measurements when all evolutionary rates are at their long-term maximum. There are a number of unique aspects of the world today, from human impacts to maximal levels of biodiversity, and these could provide potential explanations for long-term patterns of increasing rates. Indeed, some evolutionary dynamics seem to be



changing in the Anthropocene (Otto 2018). However, we are skeptical of this as an explanation for rate scaling. One primary counterargument is that the patterns of rate scaling are not limited to extant taxa, and also appear in the fossil record (Gingerich 1983; Uyeda et al. 2011; Francisco Henao Diaz et al. 2019). Instead, we view this pattern as an artifact caused by our failure to accurately measure and model the world. Most of our paper, then, is an analysis of this artifact - although we do return to this acceleration idea later.

In this review, we promote a view of evolutionary rates as model parameters whose interpretation depends on the accuracy and adequacy of the model used to estimate rates (Hunt 2012). To that end, rate scaling can come from two main sources: estimation error, involving statistical error in the data or the way the data is fit to a model; and model misspecification, where the model does not capture the main behavior of the evolutionary process. Both of these phenomena can lead to rate scaling patterns, and authors sometimes argue past one another due to a lack of understanding of these potential sources of scaling patterns. We advocate, then, a direct attempt to consider estimation error and model misspecification when considering rates (see Box 1).

Our goal is to provide a unified statistical and biological perspective on the scaling of evolutionary rates of molecular evolution, trait evolution, and diversification. We first review how each type of rate is typically measured, then review the literature on rate scaling for that type of rate. We discuss common explanations for scaling patterns, separating them into explanations based on estimation error and model

misspecification. We end the review with two sections. In the first, we discuss the practical consequences of rate-scaling patterns. Scaling patterns have strong impacts on testing hypotheses about evolutionary rates, especially when comparing different time periods. It is common for researchers to compare speciation rates, for example, across lineages of different ages, and all such comparisons are compromised by ubiquitous rate scaling patterns. Finally, we attempt to characterize and identify the potential causes of rate scaling. While a definitive mechanism is still lacking, we highlight a range of explanations that are likely candidates, and discuss how they can be evaluated. We hope that this review will help clarify work comparing rates across time scales, suggest ways to improve models used to estimate rates, and, eventually, help researchers identify and describe the biological processes that affect long-term evolutionary patterns.

## Box 1: Statistical issues and misconceptions associated with scaling of rates

Analyses of rate scaling with time intervals is subject to a particular type of estimation error, variously referred to as “spurious self-correlation” (Kenney 1982; Prairie and Bird 1989; Jackson and Somers 1991) or as a “mathematical coupling of data” (Moreno et al. 1986). The issue arises when one variable is plotted or regressed against another, but both share a common mathematical term (Long 1980). In this example, an evolutionary rate (amount of change / time) is plotted against its own denominator (time) (Sheets and Mitchell 2001)). The shared component of these coupled variables can then lead to peculiar scaling relationships (Long 1980).

One might be tempted to simply avoid coupled variables. However, in many cases, correlations between ratios and their own denominators are of interest, even in spite of their mathematical coupling (Long 1980). The clearest cases are when the ratio itself has particular meaning, as here when comparing evolutionary rates (Prairie and Bird 1989). However, we cannot ignore the presence of shared terms in our regression and scaling plots (Long 1980).

One can find four important take-home points in the extensive literature on this topic. First, we should expect a relationship between rates (change/time;  $Y/X$ ) and time ( $X$ ) —a scaling rule—whenever the original variables (time and amount of change) show no relationship (Kenney 1982). Such a finding redirects the question to “why is there no correlation between  $X$  and  $Y$ ?” In the case of evolutionary rates, we can evaluate this possibility by analysing the amount of evolutionary change (or species diversity accumulated) versus time. We find that in some cases—most commonly in the case of lineage diversification—there is little or no relationship between the original variables, so that the scaling rule emerges as a consequence of their shared term (Sheets and Mitchell 2001; Wiens 2011). This lack of a relationship between the original variables, though, does not provide a mechanistic answer, or even tell us whether the result is due to estimation error or model misspecification. Additionally, sometimes there

is a relationship between X and Y - and, even in that case, peculiar scaling patterns can still follow due to coupling.

Second, measurement and sampling error, particular types of estimation error, can have a nonrandom effect on comparisons of rates across time, resulting in spurious relationships under some circumstances (Long 1980). Focusing on the case where we relate  $Y/X$  to X, when measurement and/or sampling error is particularly high in X compared to Y, then negative scaling relationships can emerge due to error alone. Therefore, it is particularly important to explicitly account for the measurement of both variables when considering scaling relationships. In the specific case of rates versus time, we should expect problems to emerge when the error in estimating time is greater than the error in estimating the amount of change that has taken place (Long 1980). We will see echoes of this explanation in all three sections and conclude that this factor plays a prominent role in explaining the ubiquity of scaling across rates.

Third, the fundamental question of any particular scaling pattern is why does the relationship take the particular form that it does? That is, why does the scaling plot show the particular shape and slope (Estes and Arnold 2007; Gingerich 2019)? We believe that this answer is best addressed via assessment of model adequacy (Brown and Thomson 2018). For evolutionary rates, a wide range of evolutionary models predict change to accumulate with time. The lack of relationship between change and time discussed in point 1, above, deviates dramatically from many simple evolutionary models, whether they be based on constant-rate birth-death (diversification), time-homogeneous stochastic substitution models (genetics), or Brownian motion (traits). We can explore this deviation by considering the rate scaling properties of macroevolutionary models as a consequence of model misspecification, and proceed either analytically or via simulation (Pennell et al. 2015). For example, due to model misspecification, negative relationships may occur that are stronger than expected given the statistical properties of ratios even in cases where X and Y are uncorrelated. In those cases, one can still detect deviations from a model—model inadequacy—by comparing the observed slope of the scaling with that expected under the null model (Francisco Henao Diaz et al. 2019). One must then search for a different model that is adequate to explain the observed pattern. This type of approach also allows one to separate scaling due to estimation error—that is, error due to the act of measuring nature to estimate evolutionary rates—from model misspecification. This distinction reflects the extremely common desire in the literature to uncover the role of biological processes (as opposed to measurements and statistical analyses) in driving observed patterns.

<< END BOX 1 >>

## Estimating rates of evolution

Estimating an evolutionary rate at first seems straightforward. After all, isn't a rate simply the amount of change divided by the time over which that change occurred? In this simplistic view, rates are simply summary statistics of the net change divided by time. However, while the measurements themselves are descriptors of net rate over a certain timescale, there is no particular reason to think that rates will be directly comparable over different time scales (Gingerich 2019). Indeed, it is important to remember that

a particular pattern of rate scaling may be explained by many different models and dynamics of change, both biological and otherwise. Thus, inference of process from scaling patterns of net change over time must be approached with caution, with careful consideration given to alternative sources of rate-scaling.

We draw a distinction between these summary statistics of net rate and parameter estimates of rates obtained by fitting process-based models (Hunt 2012). Process-based models generally assume constant (or at least locally constant), instantaneous rates of change that are thus independent of time scale—so long as the model assumed is adequate and uniquely identifiable. We discuss this in more detail in the sections below.

## Estimating rates of molecular evolution

Following previous literature (Ho et al. 2005), when we discuss rates of molecular evolution, we are focused on the nucleotide substitution rate, defined as the rate at which new mutations become fixed in populations (Holmes et al. 2016). Rates of molecular evolution could be measured simply by comparing two sequences, counting the number of differences between them, and dividing by time. This would underestimate rates, especially between older sequences, by undercounting when multiple substitutions have occurred at a single site (Yang 2006). Even in the early days of modeling sequence evolution, this was recognized as a potential problem (Jukes et al. 1969). Since then, rates of molecular evolution have generally been based on fitting continuous-time Markov models (Gillespie 1986).

Continuous-time Markov models work by considering transitions among possible character states at each locus independently. For DNA, these states are represented by A, C, T, and G. Transitions between those states are described by a rate matrix (usually denoted  $Q$ ) that contains, as off-diagonal elements, the rates of transition from each state to the three others.

$$Q = \begin{bmatrix} -q_A & q_{AG} & q_{AC} & q_{AT} \\ q_{GA} & -q_G & q_{GC} & q_{GT} \\ q_{CA} & q_{CG} & -q_C & q_{CT} \\ q_{TA} & q_{TG} & q_{TC} & -q_T \end{bmatrix}$$

By convention, the diagonal elements are set so that the rows sum to zero. That is, for example  $q_A = q_{AG} + q_{AC} + q_{AT}$  (Yang 2006).

In the molecular evolution literature, various restricted forms of this  $Q$ -matrix go by different names (Yang 2006). For example, if all of the rates are set equal to one another, the model is called the Jukes-Cantor model (Jukes et al. 1969); if each pair of states has a unique rate, but the forward and reverse rates are equal (e.g. the rate of  $A \rightarrow T$  is equal to the rate of  $T \rightarrow A$ ), then this is called the General-Time Reversible (GTR) model (Yang 2006). Modifications also allow rates to vary across sites or across branches in the tree.

To obtain an overall rate, one estimates the best multiplier,  $r$ , such that the tree branch length times  $r$ , applied to a normalized Q matrix best predicts patterns of sequence divergence. If time is measured in absolute units, say millions of years, then,  $r$  gives average substitution rates as the expected number of substitutions per site per million years. Using this protocol, rates can be calculated for any specified model of molecular evolution (Yang 2006). Note that the procedure here also works for any set of discrete characters evolving on a phylogeny under a Markov process, including phenotypic traits (Lewis 2001). However, rate estimates for morphology are more often measured from continuously-valued data, which we address in the next section.

## Estimating rates of trait evolution

In paleontology and evolutionary biology, it is common to study the rates of evolution of continuously-valued traits (Hunt 2007; Gingerich 2009). There are two main ways to estimate evolutionary rates: either directly by using one of two expressions that measure change in trait values over time, or by estimating a model parameter that represents the instantaneous rate of change (Harmon 2014).

There are two “classic” units for measuring evolutionary rates, darwins and haldanes (Haldane 1949, Gingerich 1993). These can be viewed as either summary statistics of net rate over a time interval, or an estimate of instantaneous evolutionary rate under a model of strict linear directional change. Darwins express change in proportional change per million years, while haldanes express change in units of within-population phenotypic standard deviation per generation—the scale which is more relevant when considering genetic constraints on evolutionary change (Harmon 2014). These two metrics have been widely used to quantify rates across a range of scales, from single generations to millions of years over phylogenies (Hendry and Kinnison 1999). Importantly, however, both darwins and haldanes can only be interpreted as an instantaneous rate under the highly restrictive scenario of a linear trend in trait change over time. Under many other models, both of these rate metrics are expected to scale with the time interval over which they are measured.

In most modern studies rates of evolution are calculated as the parameters of a model of evolution. The most common of these models is Brownian motion, a model where trait change follows a random walk through time (Felsenstein 1973). Although sometimes equated with genetic drift, Brownian motion can result from a range of evolutionary scenarios including randomly varying selection, drift, and other factors (Hansen and Martins 1996). For example, if populations are under strong selection following a peak on the adaptive landscape, and that peak is changing via Brownian motion, then we expect traits to also evolve following that same Brownian motion model (Hansen and Martins 1996). Brownian motion can be described with two parameters: the starting value, often denoted  $z_0$ , and the “rate” parameter,  $\sigma^2$ . The latter parameter,  $\sigma^2$ , predicts the differences that one will find among pairs of species after a certain amount of time has elapsed, and it is this parameter that is often described as the rate of evolution. It is worth noting, though, that the Brownian motion rate is not strictly comparable to haldanes and darwins. For example, since both darwins and haldanes are only constant under a linear trend, when these metrics are calculated from data evolved under a Brownian motion model they both will show a negative



dependence on the time interval over which they are measured (Gingerich 1983; Hendry and Kinnison 1999). Likewise, if species change following a constant linear trend, rates of evolution estimated assuming a Brownian motion model will have rate-scaling properties (Gingerich 2019).

## Estimating rates of diversification

The simplest intuitive way to measure rates of speciation and extinction through time is to count the number of such events over some interval and divide by the length of time. Although such an approach does have some precedence in paleontology (Foote 2003), simply counting events ignores the fact that the number of speciation and extinction events in an interval almost certainly depends on the number of lineages alive in that interval (Raup 1985). Thus, most modern estimates of speciation and extinction rates are measured on a per-lineage basis and based on birth-death models of diversification (Raup 1985; Foote 2003).

Under a birth-death model, each lineage has an origination rate  $\lambda$  and extinction rate  $\mu$  (and, thus, a collection of  $n$  such lineages has overall rates of  $n\lambda$  and  $n\mu$  (Kendall 1948). Speciation and extinction events are assumed to occur randomly in time (i.e., are the result of a Poisson process) such that the waiting times between events are exponentially distributed. If  $\lambda$  and  $\mu$  are not equal, then birth-death models predict that species numbers will change exponentially through time. For this reason, some paleontological estimates are based on dividing the log-transformed number of events by the time interval, which then estimates  $\lambda$  and  $\mu$ . Related metrics for phylogenetic trees involve a method of moments estimator (Magallon and Sanderson 2001) that accounts for the age and diversity of a clade. This estimator, in its simplest form for a pure-birth model with stem age  $t$  and extant diversity  $n$ , is  $\lambda = \log(n)/t$ . Modifications for including extinction, correcting for sampling, and accounting for crown rather than stem age are also possible (Magallon and Sanderson 2001). The similar form of these equations from paleontology and phylogenetics emphasizes the shared assumption of a birth-death model.

In both paleobiology and phylogenetic biology, most current diversification rate estimation takes the form of estimating the parameters of such birth-death models using maximum likelihood or Bayesian inference (Nee 2006; Harmon 2019). This approach has the main advantage of using all of the information about births and deaths stored in the phylogenetic tree. Likelihood-based approaches for birth-death models have an additional complexity that the proper equation—and, thus, one's result—depends precisely on what one conditions: the age of the crown (the most recent common ancestor of extant lineages), the age of the stem (the most recent common ancestor of all lineages, extant and extinct, in a clade), and the survival (or not) of the stem or crown lineage(s) to the present day (Stadler 2013; MacPherson et al. 2020). This conditioning is essential because the process of extinction and the lineages that we sample for phylogenetic reconstruction are intimately linked (Höhna et al. 2011).

Many variations of birth-death models have been developed including those that allow variation in rates through time (Rabosky and Lovette 2008; Morlon et al. 2011; Etienne et al. 2012; Rabosky 2014; Magee et al. 2020), across clades (Alfaro et al. 2009; Rabosky 2014; Höhna et al. 2019), or allowing rates to depend on organismal traits (Maddison et al. 2007; FitzJohn 2010) and/or the environment (Goldberg et

al. 2011; Cantalapiedra et al. 2014; Rolland et al. 2014; Condamine et al. 2019). Each of these process-based models include rate parameters that can potentially be estimated from the branching structure of phylogenetic trees.

## Time-dependence of all evolutionary rates

All three rates of evolution, molecular, trait, and diversification, have been found to scale with time, with the most rapid rates observed over the shortest time periods (Figure 1). This pattern has generated discussion in the literature in all three cases, with controversy over both the existence of and explanation for rate scaling patterns. Despite the striking similarity of all three relationships, we are unaware of any paper giving a collective view of all three scaling patterns, so we present that here.

### Time-dependence for rates of molecular evolution

Rates of molecular evolution have been observed to scale with time. The earliest studies comparing divergence and time suggested complex patterns (reviewed in Gingerich 1986). The first suggestion of a consistent overall trend towards negative rate scaling (that we are aware of) traces to Wayne et al. (1991), who measured faster rates of molecular evolution when comparing mammal lineages with younger divergence times. Later, similar patterns were found in other organisms (e.g. Garcia-Moreno 2004 in birds; (Albà and Castresana 2005) in mammals, see also (Elhaik et al. 2005)). The systematic pattern of molecular rates scaling with time was described by Ho et al. (Ho et al. 2005, 2011). These papers show that estimates of rates are accelerated for comparisons of very recent divergence. In contrast to the other rate-scaling patterns discussed below, rates of molecular evolution are thought to follow a “J-shaped curve,” flattening out after approximately two million years rather than continuing to decline over the longest time scales (Penny 2005) .

This main result of rate scaling pointed out by Ho et al. (2005) has attracted controversy, with one review calling it a “tempest in a teapot” (Bandelt 2008). Much of this argument has centered around whether the scaling patterns reflect statistics or interesting biological properties of the system. For example, Emerson (Emerson 2007) argues that errors in estimating sequence divergence, biased sampling, and errors in calibration together explain the apparent scaling. In other words, Emerson attributes the entirety of the scaling shown by Ho et al. (Ho et al. 2005) to various types of estimation error, rendering the pattern ultimately uninteresting in terms of learning about biology. A follow-up from Ho et al. (Ho et al. 2011) used substantially better data and improved analyses, again arguing in favor of rate scaling as evidence for interesting and general biological processes acting over long time scales; in other words, as a consequence of model misspecification. This was in turn disputed by Emerson and Hickerson (Emerson and Hickerson 2015), again arguing for the primacy of estimation error. It is worth emphasizing, though, that all rebuttals do not deny the empirical pattern; instead, much of this discussion relates to whether this pattern is an “artifact” (i.e. driven by estimation error) or reflects some non-modeled biological process of interest (i.e. driven by model misspecification). The bulk of the discussion of the latter centers

around the distinction between mutation rates and long-term substitution rates, and, in turn, relates to purifying selection and the time needed to rid populations of segregating neutral and deleterious mutations (Penny 2005). This debate, to our knowledge, remains unresolved.

## Time-dependence for rates of trait evolution

In a classic study on the rate-scaling of trait evolution in macroevolution, Gingerich (Gingerich 1983) collated a large dataset on evolutionary rates spanning a wide range of time intervals from the fossil record to contemporary rates of evolution in the lab. Plotting all of those rate estimates together revealed a scaling pattern that spanned seven orders of magnitude of time. Gingerich showed that evolutionary rates correlated strongly with the time interval over which they were measured, with the highest rates estimated over the shortest intervals; later analyses have confirmed these results (Uyeda et al. 2011).

Since this key paper, there have been several reviews and compilations of rates of evolution published in the literature. There is a large amount of overlap in the data used for these reviews, as one original data compiled by Gingerich (Gingerich 1983) has been reused and added to over time. In general, follow-up studies confirm the basic pattern but differ in interpretation. For example, Kinnison and Hendry (Kinnison and Hendry 2001) combined the Gingerich data with information about evolutionary rates gathered from more recent literature. They confirmed the negative scaling pattern and used a suite of randomization tests to suggest that the pattern was stronger than expected due to estimation error. Likewise, Uyeda et al. (Uyeda et al. 2011) compiled data on evolutionary rates both from the Gingerich data set and phylogenetic data in the tree of life. This paper showed clear patterns of rate scaling. Uyeda et al. (Uyeda et al. 2011) emphasize an explanation based on model misspecification, demonstrating that trait divergence among species shows fairly consistent differences, regardless of time since divergence, for time scales less than a million years, only accumulating major differences beyond that.

## Time-dependence of diversification rates

Finally, diversification rates scale with the time interval over which they are measured, in a strikingly similar way to rates of molecular and trait evolution. This pattern has been described in the literature several times (Magallon and Sanderson 2001; Ricklefs 2006; McPeck and Brown 2007; Linder 2008; Scholl and Wiens 2016). Two recent papers focused particular attention on the scaling pattern, but came to opposing conclusions about the cause. First, Marin and Hedges (Marin and Hedges 2018) identified an apparent acceleration in rates of diversification in young clades across a diversity of groups. They attributed this pattern to estimation error, and in particular the undersampling of genomes and consequent underestimation of the ages of young nodes in the tree of life. Second, Henao Diaz et al. (Francisco Henao Diaz et al. 2019) showed a similar pattern of apparent rapid diversification among young clades. This paper also showed that the pattern extended to the fossil record, with rates of formation and extinction of genera inversely related to the time interval over which they were measured (Foote 1994, 2005). This inclusion of fossils precludes any common explanation based on estimation error associated with inferring phylogenetic trees from gene sequence data.

## Rate scaling as a consequence of estimation error

We consider estimation error to be statistical errors that arise during analysis even when the assumed biological process adequately describes nature. We focus in particular on three sources of estimation error: biased sampling, measurement/sampling error, and parameter non-identifiability. With biased sampling, the entities that are compared to calculate rates are not a random subset of the total population. Measurement and sampling error, on the other hand, can result from several types of error, such as finite sample size or instrumental error, while measuring a desired quantity from observed data, whether species traits, species richness, or the timing of evolutionary events. Finally, parameter non-identifiability occurs when a rate parameter is misestimated because it is not uniquely identified or only weakly estimable for a given dataset and model. All three of these phenomena can affect rate estimates, regardless of the type.

DNA sequences can be reconstructed with error due to artifacts from amplification and/or sequencing. Since sequencing errors should be random, they might not show any relationship to time since divergence, in turn leading to negative scaling of rates with time (Ho et al. 2011); see Box 1). Such arguments are particularly compelling for ancient DNA, which might be especially susceptible to errors in sequence reconstruction (Ho et al. 2011). Likewise, divergence times are also estimated with error (Arbogast et al. 2003). As argued by Weir and Schluter (Weir and Schluter 2008), this error will tend to lead to scaling artifacts when errors in estimating divergence times (X in Box 1) are greater than errors in estimating sequence divergence (Y in Box 1).

Both measurement error and biased sampling can lead to errors in estimating rates of trait evolution that can again, in some cases, result in scaling patterns with time. Ignoring measurement error will, on average, inflate differences among species, thereby elevating evolutionary rates. This effect will be most pronounced over the shortest time intervals, resulting in highly inflated rates when age differences are small (Silvestro et al. 2015; Cooper et al. 2016). Thus, failing to account for measurement error can lead to rate-scaling patterns. This argument applies equally well to time-series and phylogenetic analyses, which frequently consider species means without considering measurement error (Silvestro et al. 2015; Kostikova et al. 2016). Measurement error can also affect our calculation of elapsed time and the timing of branching events in the phylogenetic tree. Similar to trait differences, ages are estimated with error, and that error is sometimes not accounted for in calculation of rates of trait evolution. Unlike the case of measurement error for species' traits, random error in branch lengths should not lead to negative scaling of rates through time. Rate scaling could result if ages were systematically biased (underestimated for younger comparisons, or overestimated for older ones, or some combination of the two). Rate scaling can also arise when there is error in both the numerator and the denominator of a rate, but the error in the denominator is greater than that in the numerator (Long 1980). This seems highly likely for trait evolution, and almost certainly makes some contribution to the observed scaling patterns.

Biased sampling can also have an impact on scaling for rates of trait evolution. We can see one example in Gould (Gould 1984), who called the pattern of trait rates scaling with time a “psychological and

mathematical artifact”. Gould’s main argument was that paleontologists were biased in what they measured when calculating rates. In the fossil record, two lineages might only be connected in a time-series dataset if they are within some threshold of similarity. Any lineages that are “too different,” then, are not used to calculate evolutionary rates. This could lead to rate scaling through systematically underestimated rates over long time scales, as old, highly divergent lineages are not likely to be included in rate calculations. Any pattern of biased sampling that affects the expected distribution of species’ differences could lead to scaling, whether rates are calculated from phylogenetic trees or fossil data.

Finally, time-scaling in diversification rates may be due to estimation error, either of the rates themselves or of the divergence times/clade ages. As mentioned in the trait evolution section above, one possible explanation for this scaling relationship between rates and time is estimation error—either of the rates themselves or of the divergence times/clade ages. For instance, it is well recognized that even with a lot of sequence data, there is still considerable uncertainty in divergence time estimates, particularly deep in the tree (Graur and Martin 2004; Revell et al. 2005). Henao Diaz et al. (Francisco Henao Diaz et al. 2019) show in their supplementary materials that such errors could lead to scaling patterns, echoing an early point by Ricklefs (Ricklefs 2006). A similar argument applies to misestimation of time intervals in the fossil record. However, two considerations suggest that misestimated branch lengths are not the full explanation for the observed patterns. First, the pattern remains despite substantial improvements in divergence time estimation, even as the size of analyzed clades has greatly increased (Francisco Henao Diaz et al. 2019). Second, simulations from Henao Diaz et al. (Francisco Henao Diaz et al. 2019) suggest that the scaling patterns due to estimation error of branch lengths are not sufficient to explain the negative slope as observed. Fortunately, it appears that estimated rates of diversification are generally quite robust to misspecification of the clock model used to date nodes (Wertheim and Sanderson 2010, Sarver et al. 2019).

Diversification rate scaling can also be influenced by biased sampling. For phylogenetic trees, one must decide on the clades for analysis, breaking up a single tree of life into monophyletic clades. We do not yet know how this selection of clades affects diversification rates, although the scaling pattern seems to persist through alternative selection schemes in Henao Diaz et al. (2019). One clear effect is due to the exclusion of small clades: trees are unlikely to be constructed or considered for particularly species poor (or monotypic) clades. If these depauperate clades are not included in analyses, we should expect a biased view of rates across time. Other effects on scaling are possible based on how clades are defined by crown or stem age (Stadler et al. 2014). These issues with sampling are clearly present with paleontological data (Hopkins et al. 2018). One expects these biases to lead to scaling to the extent that it breaks the association between the length of the time interval and the number of observed diversification events.

The most promising path forward is to improve our ability to both estimate measurement error and account for it in our modeling approaches (Houle et al. 2011). Significant progress has been made in nearly every modeling type to account for measurement error. For example, several methods for estimating evolutionary rates from Brownian motion models use external estimates of observational measurement error to provide unbiased estimates of rate (Ives et al. 2007; Felsenstein 2008); other approaches exist to account for sequencing error in rates of molecular evolution (Johnson and Slatkin 2008; Ho et al. 2015)

and unsampled lineages in rates of diversification (Höhna 2014). While the statistical procedure can be straightforward, it is worth noting that determining what values to use as measurement error can be confused when “biological error” and statistical sampling properties become conflated (Hansen and Bartoszek 2012). This issue will be explored more fully in the model misspecification sections below. Sampling bias can be more challenging, but some approaches exist that account for sampling when fitting evolutionary models (FitzJohn 2010). While it is still often omitted from comparative analyses, we believe there is growing and widespread recognition such biases and errors must always be included in a comparative analysis (Silvestro et al. 2015), and, in general, we think researchers are largely aware and cautious of this sort of model inadequacy.

Even when data are measured without error and models are adequate, there are limits to parameter estimation. For a simple example, consider a set of molecular sequences for which every site is completely saturated so as to remove all phylogenetic signal. This set of sequences, regardless of branch length or time interval, is consistent with any sufficiently high substitution rate (Mossel and Vigoda 2005). The resulting flat likelihood surface will estimate rate parameters subject to the vagaries of a particular optimization algorithm and largely invariant to time, resulting in slower rates estimated over longer time intervals. Extending this problem to model selection, we would be remiss to not mention the dark corridors of uncertainty in model-fitting stemming from model non-identifiability. Non-identifiable models are defined as two or more models that produce observationally equivalent distributions of data, and result in rate parameters that cannot be uniquely identified. For example, Louca and Pennell (Louca and Pennell 2020) have recently challenged the entire enterprise of estimating rates of diversification from molecular phylogenies. They prove that for the case of any homogeneous birth-death process, in which the process is the same for all lineages at any given time point but where the process can change through time, neither  $\lambda$  nor  $\mu$  nor any combination thereof (e.g., net diversification rate  $r = \lambda - \mu$ ) can be uniquely defined or estimated (Louca and Pennell 2020). This is due to the inherent non-identifiability of birth-death processes: for any time variable process, where  $\lambda$  and  $\mu$  vary through time according to some mathematical functions, there are an infinite number of alternative functions (called a congruence class) that generate an identical expected distribution of tree sizes and branching patterns which cannot be distinguished using likelihood or any other method using phylogenetic trees (Louca and Pennell 2020). This result suggests that interpreting model fit should be done with caution. In particular, current methods select, somewhat arbitrarily, particular parameter values from the congruence class. This selection alone could lead to scaling patterns. Similar congruence classes are known to exist for continuous trait models on ultrametric phylogenies (e.g. Brownian Motion vs. Brownian Motion with a trend (Felsenstein 1988); Ornstein-Uhlenbeck vs. accelerating change models, etc (Hansen and Martins 1996; Slater et al. 2012; Uyeda et al. 2015)). It is likely that similar congruence classes exist across all macroevolutionary models, not only diversification models. These problems blend into our next section, in which we discuss the consequences of such model misspecification.

# Rate scaling as a consequence of model misspecification

There are countless ways that our macroevolutionary models might be incorrect, and many of these could lead to scaling patterns. Some of these are well-summarized in the previous literature on this topic (Penny 2005; Ho et al. 2011; Marin and Hedges 2018). In this review, rather than comprehensively cover all the ways that our simple models may be wrong, we will focus on four possible categories of model misspecification. We direct our focus on shared explanations that may explain all three rate-scaling patterns in a way that is consistent with contemporary evolutionary theory.

## Evolution is accelerating towards the present day

Perhaps evolution is speeding up, with all rates peaking at the present day. Such a pattern could explain why everything that we observe over short time scales from our present-day point of view appears rapid, but including longer time scales stretching back into the distant past results in slower rates. This explanation is potentially viable for datasets that rely on the present day for sampling, such as analyses based on phylogenetic trees.

In our view, this explanation fails on two counts. First, there is no expectation that evolution should speed up uniformly through time from evolutionary theory. Although some authors have suggested that diversity begets diversity (Emerson and Kolm 2005), and perhaps might anticipate that evolution is faster in places where diversity is higher, one would still not expect an increasing pattern in all places over all time. Furthermore, this explanation does not account for the presence of similar scaling patterns for trait evolution and diversification using purely paleontological data (Uyeda et al. 2011; Francisco Henao Diaz et al. 2019).

## Evolutionary entities are highly heterogeneous

We know that evolutionary lineages are highly heterogeneous in a number of ways. All of the processes that we are concerned with here vary tremendously across clades, with some lineages evolving rapidly and others more slowly. This heterogeneity is why, in molecular evolution, the relaxed clock is often used, modeling rate variation across branches in the tree of life; why the shape of that tree is unbalanced, reflecting variation in speciation and extinction rates across clades (Mooers and Heard 1997; Ng and Smith 2014); and why the tempo of trait evolution is highly variable among even closely related clades.

It might not be immediately obvious that rate heterogeneity will lead to time-scaling for the resulting rate estimates. Indeed, if we are able to accurately account for the heterogeneity, then rate estimates should vary, but independently of time interval. However, in practice, heterogeneity is a type of model misspecification where we fit a sort of “average” rate across heterogeneity across clades (Yang 2006). In such a case, older clades will lead to rate estimates that adequately sample across the heterogeneity, leading to rates that match the long-term average; but, as clades get younger, we are more and more likely to sample a part of the tree of life with unusually high rates, unpolluted by the tempering influence

of slow lineages. There will be, by chance, slow young clades as well - but the unusually fast young clades will dominate the pattern, leading to the classic negative relationship between rates and time.

## Evolutionary events are clustered in time

All of the models we use to estimate evolutionary rates assume that events occur independently in time. For example, for discrete event models, like sequence evolution or birth-death models, the waiting time to the next event is independent of whether or not an event has just occurred (Yang 2006; Harmon 2019). Likewise, Brownian motion models assume that variance accumulates through time independently of what changes have recently occurred in the population (Felsenstein 1973). Although these assumptions allow the use of Markov models and ease mathematical analysis, they may not be realistic or accurate for modeling evolution.

We suggest that macroevolution may commonly violate this independent events assumption, and that in particular macroevolutionary events might be clustered in time, occurring in a “pulsed” pattern (Foote 2005; Landis and Schraiber 2017). For example, species might experience stasis over relatively long time intervals, and then periods of rapid change, followed by further stasis (Estes and Arnold 2007). Likewise, speciation and extinction events might be clustered in intervals, perhaps associated with rapid environmental change or major alterations in species interactions (Foote 2005). Clustered events lead to rate scaling because some short time intervals will occur during these periods of concentrated change, leading to elevated rates (Sadler 1981). Consistent with this explanation, rate scaling plots often show tremendous scatter over short time intervals, suggesting that evolution is often, but not always, rapid over short time scales (Figure 1). Finally, different patterns of clustering through time could lead to different scaling patterns, and potentially could be estimated from the shape of the rate scaling pattern (Foote 2005). Nonhomogeneous event models (Hawkes 2018; Goldberg and Foo 2020) might be particularly useful in investigating the clustering of evolutionary events. Indeed, models of pulsed evolution show a better fit to some empirical datasets for both trait evolution (Landis and Schraiber 2017) and speciation and extinction rates (Foote 2005), and the success of relaxed molecular clock models (Drummond et al. 2006) suggests pulsed rates of molecular evolution.

## Evolution is constrained

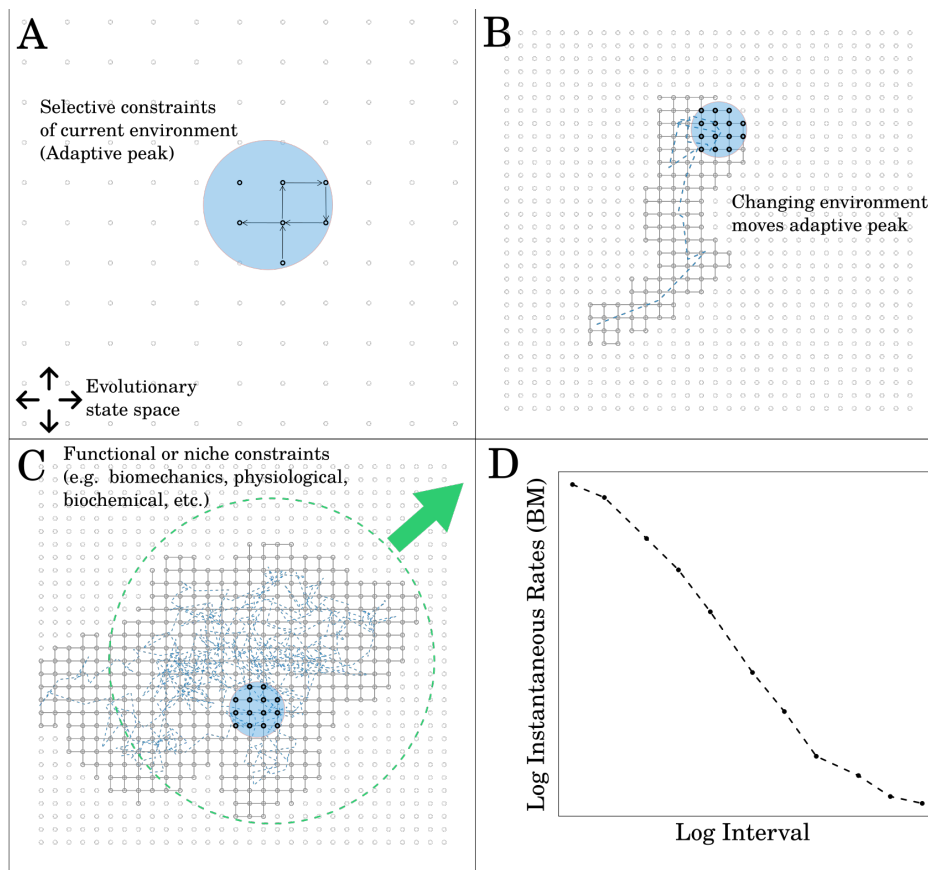
Constraints on evolutionary change set bounds that limit the long-term effects of rapid evolution over short time scales (Schwenk 1994). This logic can be applied to all three rates. For molecular evolution, sequences can evolve rapidly over short time scales by accumulating neutral or nearly-neutral substitutions (Ohta 1992). However, there are a limited number of such mutations in the vicinity of the current sequence, resulting in rapid saturation of sequence-space due to the constraints of natural selection (DeSalle et al. 1987). Some of the most common genes used to estimate evolutionary rates are key metabolic components of cells, critical for function and likely under long-term constraints for functionality. Given this constrained state space, rate-scaling will emerge for rate parameters if the constraints are not adequately modeled, or as a consequence of our inability to estimate rate once complete saturation occurs (Elhaik et al. 2005).



The argument for rate scaling driven by constraint is perhaps most developed for trait evolution. Here, a constrained model of evolution, called Ornstein-Uhlenbeck model, is commonly used for analysis (Hansen 1997). This model is often described as Brownian motion, but with a “central tendency” that tends to make evolution go towards some optimal value (Butler and King 2004). OU models, in general, have four parameters: a rate parameter ( $\sigma^2$ ), the constraint parameter ( $\alpha$ ), an optimal value ( $\Theta$ ), and a starting value  $z_0$ . OU models are commonly selected as the best fitting for comparative data compared to various alternatives (Harmon et al. 2010; Pennell et al. 2015; Cooper et al. 2016). If we analyze data generated under a constrained OU model by assuming unconstrained Brownian motion, we will observe rapid change over short time scales, approaching the OU sigma-squared parameter (Uyeda et al. 2015). We will also see slower and slower apparent rates over longer times as constraints begin to dampen variance among species compared to the Brownian expectation (Estes and Arnold 2007; Uyeda et al. 2011).

Finally, constrained models are also commonly invoked for diversification. In this case, many authors have suggested that speciation and extinction are diversity dependent, with net diversification rates slowing as species richness nears some carrying capacity (Harmon et al. 2003; McPeck 2008; Etienne et al. 2012; Harmon and Harrison 2015; Rabosky and Hurlbert 2015; Condamine et al. 2019). This pattern is fundamentally different than the other two constraints in that such a process would leave a clear mark in phylogenetic data, with branching events concentrated at the base of the tree of life. However, it remains unclear whether such a pattern is common or rare in comparative data (Pannetier et al. 2020). One can construct scenarios where a complex pattern of diversity-dependence can lead to rate scaling. For example, imagine that all of life was generated under a multi-regime diversity-dependent model, where lineages were subject to carrying capacities based on their close relatives such that diversification rates slow as species accumulate. Imagine also that lineages sometimes escaped those constraints and entered into new regimes with independent carrying capacities (e.g. new adaptive zones). In that case, you would observe high rates of diversification only when considering clades that have just moved into new regimes—and these clades would be disproportionately young (Schluter 2000). Thus, we would see the expected negative scaling of rates with clade age. In any case, there is some debate and uncertainty in the literature about the amount of support for diversity-dependence of speciation and extinction.

In all cases, failure to account for, or accurately estimate, the constraints on a process can result in rate scaling for model parameters (Figure 2). Models that account for the long-term dynamics of these constraints might lead the way to deeper understanding of the state space of evolutionary change, and how the evolutionary process leads to convergence and saturation (Figure 2). This evolutionary state space may itself change over time (e.g. Butler and King 2004, Beaulieu et al. 2012). Such models may also behave as the “clustered” models described above, with rapid evolution concentrated near peak shift events; our potential explanations for rate-scaling need not be mutually exclusive.



## Evolution is hierarchical

All three models follow macroevolutionary tradition in taking the species as the unit, and modeling changes that accrue among species over deep time (Stanley 1979). This ignores the hierarchical nature of evolution—and, in particular, can conflate change happening among populations within species with change occurring among species (Jablonski 2000, 2007; Eldredge et al. 2005). In each case of evolutionary rates, there is good reason to suspect that this failure of models to “cross the species boundary” is a problem for estimating evolutionary rates. In all three cases, there are also models and methods that can potentially directly address this limitation..

For molecular evolution, the core idea that we should expect substitutions to accumulate linearly with respect to time stems from the neutral theory of molecular evolution (Li 1997). However, this theory applies to the substitution rate and assumes that all measured differences are substitutions. If, instead, some of the differences we observe represent variation within species—that is, alleles that have not yet reached fixation—then we have misspecified our model (Penny 2005). Our model does not account for the dynamics among individuals within species and populations, and only considers fixed differences between species. For recent divergence, differences that we see reflect both the substitution rate and the mutation rate, while over long time scales, most differences are substitutions. This leads directly to a predicted pattern of rates of molecular evolution scaling negatively with the time interval over which they are measured (Ho et al. 2011).

Under selection, scenarios are more complex. First, selection by itself does not necessarily lead to rate scaling; in fact, under nearly-neutral models of molecular evolution we can still expect substitutions to accumulate linearly through time (Ohta 1992). However, issues emerge again when we consider processes that occur among individuals within species. As most mutations are deleterious, selection filters ephemeral polymorphism generated at present. This exacerbates the pattern described above for neutral models. In simple terms, comparisons over short time scales include actual substitutions, neutral segregating variants, and deleterious mutations that have not been purged from the population. Over long time scales, differences eventually reflect mostly substitutions, both neutral and non-neutral (Penny 2005). Again, failure to account for the presence of both deleterious and neutral segregating variants within species populations—which can be considered as a model misspecification—will lead to a pattern of negative rate-scaling. This leads to problems analogous to measurement error but cannot be corrected simply with estimates of instrument repeatability or sampling variance, rather it is “biological error” that must be included into any adequate process model of evolution (Felsenstein 1985; Hansen and Bartoszek 2012).

For trait evolution, a similar type of hierarchical explanation lies in Futuyma’s ephemeral divergence model (Futuyma 1987, 2010). Under this model, most evolutionary trait divergence takes place in small populations, for example, at the periphery of a species’ range where selective optima may be different. This divergence is mostly ephemeral, as the ultimate fate of many such populations is extinction or re-assimilation (Futuyma 1987; Eldredge et al. 2005). According to Futuyma (2010), then, novel traits that diverge persist only rarely, when these otherwise ephemeral changes are locked in by speciation and the evolution of reproductive isolation, or spread among the whole species. Under the ephemeral divergence model, we expect to see rapid, but ephemeral, trait evolution over short time scales, and seemingly much slower rates over long time scales (Eldredge et al. 2005).

A similar explanation for diversification comes in the form of “ephemeral” or “incipient” speciation models (Rosindell et al. 2010; Etienne and Rosindell 2012; Rosenblum et al. 2012). The idea here is that the initiation of speciation, where populations begin to become independent lineages, occurs at a high rate, but that these ephemeral forms only rarely persist as full species over longer time scales (Rosenblum et al. 2012; Dynesius and Jansson 2014). Such a process will result in higher rates of speciation initiation over short time scales and slower rates over longer time scales (Rabosky 2016; Harvey et al. 2019).

All three of the above hierarchical explanations predict that the rate-scaling pattern will be non-linear, with an inflection arising from the transition from within-species to between-species changes. Within species, population-level comparisons include within-species constraints, along with standing or ephemeral variation, while between-species comparisons consider only such changes that persist and become fixed at the species-level). This is a prediction that can be directly tested. Indeed, most empirical patterns of rate scaling in evolution appear to have the predicted non-linear shape with an inflection between one and ten million years ago (Ho et al. 2011; Uyeda et al. 2011). Whether or not rate scaling disappears asymptotically at this “micro to macro” inflection point, or continues with a different scaling factor is variable among the three rates examined and the model under which they are estimated.

Ultimately, however, the hierarchical explanation relies on saturation of variation and constraint at successive levels of hierarchy causing the data to lose information about time—either because change has not yet accumulated or because the state space has become saturated (Figure 2). Notably, this is also a definition of “phylogenetic signal” that generalizes across the biological hierarchy (Hansen and Orzack 2005; Jablonski 2007). Since successive levels of biological organization, from species to genera above, may be subject to their own constraints that may depend on factors that themselves change at rates ranging from generations to eons, patterns of rate scaling may be expected to similarly occur at successive levels of the biological hierarchy (Figure 2). Adequately describing such dynamics will require models that likewise can accommodate the hierarchical nature of evolution (Reitan et al. 2012; Kostikova et al. 2016; Reitan and Liow 2019).

## Consequences of rate scaling

It remains unclear what the relative importance of each of these explanations for scaling of evolutionary rates among our three identified scaling patterns. However, and regardless of the cause, the prevalence of rate scaling means that this pattern will have effects on many of our analyses, and should be considered when interpreting evolutionary rates.

The most obvious effect that pervasive rate-scaling patterns will have is that they will tend to cause us to confuse the young with the fast. For example, when comparing evolutionary rates—whether speciation, extinction, trait evolution, or molecular evolution—across time periods of different length, we will tend to estimate the highest rates for the shortest interval. This can be expressed in different ways. For example, rates of trait evolution will tend to seem fastest in younger clades; speciation and extinction rates will tend to be estimated as higher in lineages from younger habitats; and rates of molecular evolution will seem faster when comparing shallow versus deeply divergent pairs of taxa. This is not to say that the younger or shorter comparisons will always be uniformly faster; in fact, in all known scaling relationships there is enough scatter around the regression line that we should expect variation. Still, the scaling pattern should be taken into consideration when comparisons are made across different time scales.

A more subtle problem that also stems immediately from rate scaling is that parameters that all scale with the same covariate will tend to correlate with one another. In this context, then, we should expect to see correlations among evolutionary rates, since all are related to time. Indeed, extensive papers have documented putative relationships between all three possible pairs of rates: molecular evolution and speciation rates (Pagel et al. 2006; Lanfear et al. 2010); trait evolution and speciation rates (Rabosky and Adams 2012; Rabosky et al. 2012, 2013; Cooney and Thomas 2021); and molecular evolution and trait evolution (Omeland 1997; Berv and Field 2018). While we are not suggesting that any of these studies are incorrect, *per se*, we argue that future papers in this area should strive to account for the potential effects of time scaling in such analyses.

# Conclusion

There is no lack of explanations for rate scaling, which range from purely statistical to biologically meaningful. We hope we have provided a clear guide as to the ubiquity of these patterns, how such patterns can emerge from statistical *and* biological sources, and how they will affect our analyses. While it seems likely that many of the mechanisms we describe play a major role in rate scaling, there remains considerable mystery in the relative importance of each as ultimate explanations for these patterns. While estimation error certainly explains some of these patterns, substantial evidence suggests that at least a portion of the apparent pattern likely results from the inadequacy of our models in representing evolutionary processes. Ultimately, we think the ubiquitous nature of rate scaling in evolution reveals much about our ignorance of how microevolutionary change accumulates from micro- to macroevolutionary scales, or at least how to adequately model it (Houle et al. 2017). Furthermore, we lack a mechanistic understanding of how to interpret apparent constraints on macroevolution (Boucher et al. 2017, Houle et al. 2017). Indeed, much of our discussion and the relevant literature on rates are intimately tied to Gould and Eldredge's (Gould and Eldredge 1977) motivation in proclaiming that "stasis is data", and the enduring "paradox of stasis" (Hansen and Houle 2004). While spirited debate has ensued, and continues to this day, as to the causes of this pattern, it is unquestionable that this reframing revealed the great biases that evolutionary biologists often have in the questions they ask and the methods that they use. Indeed, many of the great mysteries in evolution are not why change occurs, but why it fails to occur. Traditionally, evolutionary biologists have framed the Earth's incredible biodiversity as an outstanding mystery to be solved. Perhaps, though, we should be surprised not by how much—but rather how little—biodiversity there is, considering the tremendous potential for evolution to generate new genes, traits, and species, even over short time scales. It is by understanding the underlying mechanisms of these limits and how they evolve over scales of time and space that hold the most promise for adequately describing evolutionary processes with meaningful parameters that characterize the tempo and mode of evolutionary change.

# Acknowledgments

M.W.P. was funded by a NSERC Discovery Grant. J.R. received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie (Grant No. 785910). This work also benefited from Investissement d'Avenir grants managed by the Agence Nationale de la Recherche (CEBA: ANR-10-LABX-25-01;TULIP: ANR-10-LABX-0041). B.S. was funded by NSF GRFP 1832399. J.C.U. was funded by NSF DEB-1942717. Some ideas in this manuscript were discussed in a working group funded by University of British Columbia (UBC) through the GCRC program.

## References

- Albà, M. M., and J. Castresana. 2005. Inverse relationship between evolutionary rate and age of mammalian genes. *Molecular biology and evolution* 22:598–606.
- Alfaro, M. E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, G. Carnevale, et al. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences* 106:13410–13414.
- Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2003. Estimating Divergence Times from Molecular Data on Phylogenetic and Population Genetic Timescales.
- Bandelt, H.-J. 2008. Clock debate: when times are a-changin': Time dependency of molecular rate estimates: tempest in a teacup. *Heredity*.
- Berv, J. S., and D. J. Field. 2018. Genomic Signature of an Avian Lilliput Effect across the K-Pg Extinction. *Systematic biology* 67:1–13.
- Brown, J. M., and R. C. Thomson. 2018. Evaluating Model Performance in Evolutionary Biology. *Annual review of ecology, evolution, and systematics*.
- Butler, M. A., and A. A. King. 2004. Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution. *The American naturalist* 164:683–695.
- Cantalapiedra, J. L., R. G. Fitzjohn, T. S. Kuhn, M. H. Fernández, D. DeMiguel, B. Azanza, J. Morales, et al. 2014. Dietary innovations spurred the diversification of ruminants during the Caenozoic. *Proceedings. Biological sciences / The Royal Society* 281:20132746.
- Charlesworth, B., R. Lande, and M. Slatkin. 1982. A NEO-DARWINIAN COMMENTARY ON MACROEVOLUTION. *Evolution; international journal of organic evolution* 36:474–498.
- Condamine, F. L., J. Rolland, and H. Morlon. 2019. Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecology letters* 22:1900–1912.
- Cooney, C. R., and G. H. Thomas. 2021. Heterogeneous relationships between rates of speciation and body size evolution across vertebrate clades. *Nature ecology & evolution* 5:101–110.
- Cooper, N., G. H. Thomas, C. Venditti, A. Meade, and R. P. Freckleton. 2016. A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. *Biological journal of the Linnean Society. Linnean Society of London* 118:64–77.
- DeSalle, R., T. Freedman, E. M. Prager, and A. C. Wilson. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *Journal of molecular evolution* 26:157–164.
- Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS biology* 4:e88.
- Dynesius, M., and R. Jansson. 2014. Persistence of within-species lineages: a neglected control of speciation rates: PERSPECTIVE. *Evolution; international journal of organic evolution* 68:923–934.
- Eldredge, N., J. N. Thompson, P. M. Brakefield, S. Gavrilets, D. Jablonski, J. B. C. Jackson, R. E. Lenski, et al. 2005. The dynamics of evolutionary stasis. *Paleobiology* 31:133–145.
- Elhaik, E., N. Sabath, and D. Graur. 2005. The “Inverse Relationship Between Evolutionary Rate and Age of Mammalian Genes” Is an Artifact of Increased Genetic Distance with Rate of Evolution and Time of Divergence. *Molecular biology and evolution* 23:1–3.
- Emerson, B. C. 2007. Alarm bells for the molecular clock? No support for Ho et al.’s model of time-dependent molecular rate estimates. *Systematic biology* 56:337–345.
- Emerson, B. C., and M. J. Hickerson. 2015. Lack of support for the time-dependent molecular evolution hypothesis. *Molecular ecology* 24:702–709.
- Emerson, B. C., and N. Kolm. 2005. Species diversity can drive speciation. *Nature* 434:1015–1017.
- Erwin, D. H. 2000. Macroevolution is more than repeated rounds of microevolution. *Evolution &*

- development 2:78–84.
- Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. *The American naturalist* 169:227–244.
- Etienne, R. S., B. Haegeman, T. Stadler, T. Aze, P. N. Pearson, A. Purvis, and A. B. Phillimore. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings. Biological sciences / The Royal Society* 279:1300–1309.
- Etienne, R. S., and J. Rosindell. 2012. Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Systematic biology* 61:204–213.
- Felsenstein, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. *American journal of human genetics* 25:471–492.
- Felsenstein, J. 1985. Phylogenies and the Comparative Method. *The American naturalist* 125:1–15.
- Felsenstein, J. 1988. Phylogenies and Quantitative Characters. *Annual review of ecology and systematics* 19:445–471.
- Felsenstein, J. 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. *The American naturalist* 171:713–725.
- FitzJohn, R. G. 2010. Quantitative traits and diversification. *Systematic biology* 59:619–633.
- Foote, M. 1994. Temporal variation in extinction risk and temporal scaling of extinction metrics. *Paleobiology* 20:424–444.
- Foote, M. 2003. Origination and extinction through the Phanerozoic: a new approach. *The Journal of geology*.
- . 2005. Pulsed origination and extinction in the marine realm. *Paleobiology*.
- Francisco Henao Diaz, L., L. J. Harmon, M. T. C. Sugawara, E. T. Miller, and M. W. Pennell. 2019. Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences of the United States of America* 116:7403–7408.
- Futuyma, D. J. 1987. On the role of species in anagenesis. *The American naturalist* 130:465–473.
- . 2010. Evolutionary constraint and ecological consequences. *Evolution; international journal of organic evolution* 64:1865–1884.
- Gillespie, J. H. 1986. Rates of molecular evolution. *Annual review of ecology and systematics* 17:637–665.
- Gingerich, P. D. 1983. Rates of evolution: effects of time and temporal scaling. *Science* 222:159–161.
- Gingerich, P. D. 2009. Rates of Evolution. *Annual review of ecology, evolution, and systematics* 40:657–675.
- . 2019. *Rates of evolution: A quantitative synthesis*. Cambridge University Press.
- Goldberg, E. E., and J. Foo. 2020. Memory in Trait Macroevolution. *The American naturalist* 195:300–314.
- Goldberg, E. E., L. T. Lancaster, and R. H. Ree. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic biology* 60:451–465.
- Gould, N. E.-S. J., and N. Eldredge. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pages 82–115 in T. J. M. Schopf, ed. *Models in Paleobiology*. Johns Hopkins University Press Baltimore.
- Gould, S. J. 1984. Smooth curve of evolutionary rate: a psychological and mathematical artifact. *Science* 226:994–996.
- Gould, S. J. 1985. The paradox of the first tier: an agenda for paleobiology. *Paleobiology* 11:2–12.
- Gould, S. J., and N. Eldredge. 1977. Punctuated Equilibria: The Tempo and Mode of Evolution Reconsidered. *Paleobiology* 3:115–151.
- Graur, D., and W. Martin. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in genetics: TIG* 20:80–86.
- Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution*;

- international journal of organic evolution 51:1341–1351.
- Hansen, T. F., and K. Bartoszek. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. *Systematic biology* 61:413–425.
- Hansen, T. F., and D. Houle. 2004. Evolvability, stabilizing selection, and the problem of stasis. Phenotypic integration: Studying the ecology and evolution of complex phenotypes 130–150.
- Hansen, T. F., and E. P. Martins. 1996. Translating Between Microevolutionary Process and Macroevolutionary Patterns: The Correlation Structure of Interspecific Data. *Evolution; international journal of organic evolution* 50:1404–1417.
- 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; international journal of organic evolution* 59:2063–2072.
- Harmon, L. 2019. Phylogenetic comparative methods: learning from trees. *Ecoevorxiv*.
- Harmon, L. J. 2014. Macroevolutionary Rates. *The Princeton Guide to Evolution*. Princeton University Press.
- Harmon, L. J., and S. Harrison. 2015. Species Diversity Is Dynamic and Unbounded at Local and Continental Scales. *The American naturalist* 185:584–593.
- Harmon, L. J., J. B. Losos, T. Jonathan Davies, R. G. Gillespie, J. L. Gittleman, W. Bryan Jennings, K. H. Kozak, et al. 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution; international journal of organic evolution* 64:2385–2396.
- Harmon, L. J., J. A. Schulte, A. Larson, and J. B. Losos. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. *Science* 301:961–964.
- Harvey, M. G., S. Singhal, and D. L. Rabosky. 2019. Beyond Reproductive Isolation: Demographic Controls on the Speciation Process. *Annual review of ecology, evolution, and systematics*.
- Hawkes, A. G. 2018. Hawkes processes and their applications to finance: a review. *Quantitative Finance* 18:193–198.
- Hendry, A. P., and M. T. Kinnison. 1999. PERSPECTIVE: THE PACE OF MODERN LIFE: MEASURING RATES OF CONTEMPORARY MICROEVOLUTION. *Evolution; international journal of organic evolution* 53:1637–1653.
- Hendry, A. P., and M. T. Kinnison. 2001. An introduction to microevolution: rate, pattern, process. *Genetica* 112-113:1–8.
- Höhna, S. 2014. Likelihood inference of non-constant diversification rates with incomplete taxon sampling. *PloS one* 9:e84184.
- Höhna, S., W. A. Freyman, Z. Nolen, and J. P. Huelsenbeck. 2019. A Bayesian approach for estimating branch-specific speciation and extinction rates. *bioRxiv*.
- Höhna, S., T. Stadler, F. Ronquist, and T. Britton. 2011. Inferring speciation and extinction rates under different sampling schemes. *Molecular biology and evolution* 28:2577–2589.
- Holmes, E. C., G. Dudas, A. Rambaut, and K. G. Andersen. 2016. The evolution of Ebola virus: Insights from the 2013-2016 epidemic. *Nature* 538:193–200.
- Hopkins, M. J., D. W. Bapst, C. Simpson, and R. C. M. Warnock. 2018. The inseparability of sampling and time and its influence on attempts to unify the molecular and fossil recordsHopkins, Bapst, Simpson, and Warnocksampling and time in fossil and molecular records. *Paleobiology* 44:561–574.
- Ho, S. Y. W., S. Duchêne, M. Molak, and B. Shapiro. 2015. Time-dependent estimates of molecular evolutionary rates: evidence and causes. *Molecular ecology*.
- Ho, S. Y. W., R. Lanfear, L. Bromham, M. J. Phillips, J. Soubrier, A. G. Rodrigo, and A. Cooper. 2011. Time-dependent rates of molecular evolution. *Molecular ecology* 20:3087–3101.
- Ho, S. Y. W., M. J. Phillips, A. Cooper, and A. J. Drummond. 2005. Time dependency of molecular rate



- estimates and systematic overestimation of recent divergence times. *Molecular biology and evolution* 22:1561–1568.
- Houle, D., C. Pélabon, G. P. Wagner, and T. F. Hansen. 2011. Measurement and meaning in biology. *The Quarterly review of biology* 86:3–34.
- Hunt, G. 2007. The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proceedings of the National Academy of Sciences of the United States of America* 104:18404–18408.
- . 2012. Measuring rates of phenotypic evolution and the inseparability of tempo and mode. *Paleobiology*.
- Ives, A. R., P. E. Midford, and T. Garland Jr. 2007. Within-species variation and measurement error in phylogenetic comparative methods. *Systematic biology* 56:252–270.
- Jablonski, D. 2000. Micro- and Macroevolution: Scale and Hierarchy in Evolutionary Biology and Paleobiology. *Paleobiology* 26:15–52.
- . 2007. SCALE AND HIERARCHY IN MACROEVOLUTION. *Palaeontology* 50:87–109.
- Jackson, D. A., and K. M. Somers. 1991. The spectre of “spurious” correlations. *Oecologia* 86:147–151.
- Johnson, P. L. F., and M. Slatkin. 2008. Accounting for bias from sequencing error in population genetic estimates. *Molecular biology and evolution* 25:199–206.
- Jukes, T. H., C. R. Cantor, and Others. 1969. Evolution of protein molecules. *Mammalian protein metabolism* 3:21–132.
- Kendall, M. G. 1948. Rank correlation methods.
- Kenney, B. C. 1982. Beware of spurious self-correlations! *Water Resources Research*.
- Kinnison, M. T., and A. P. Hendry. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* 112-113:145–164.
- Kostikova, A., D. Silvestro, P. B. Pearman, and N. Salamin. 2016. Bridging Inter- and Intraspecific Trait Evolution with a Hierarchical Bayesian Approach. *Systematic biology* 65:417–431.
- Landis, M. J., and J. G. Schraiber. 2017. Pulsed evolution shaped modern vertebrate body sizes. *Proceedings of the National Academy of Sciences of the United States of America* 114:13224–13229.
- Lanfear, R., S. Y. W. Ho, D. Love, and L. Bromham. 2010. Mutation rate is linked to diversification in birds. *Proceedings of the National Academy of Sciences of the United States of America* 107:20423–20428.
- Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic biology* 50:913–925.
- Linder, H. P. 2008. Plant species radiations: where, when, why? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 363:3097–3105.
- Li, W. 1997. *Molecular evolution*. Sinauer associates incorporated.
- Long, S. B. 1980. The Continuing Debate over the Use of Ratio Variables: Facts and Fiction. *Sociological methodology* 11:37–67.
- Louca, S., and M. W. Pennell. 2020. Extant timetrees are consistent with a myriad of diversification histories. *Nature* 580:502–505.
- MacPherson, A., S. Louca, A. McLaughlin, J. B. Joy, and M. W. Pennell. 2020. A General Birth-Death-Sampling Model for Epidemiology and Macroevolution. Cold Spring Harbor Laboratory.
- Maddison, W. P., P. E. Midford, S. P. Otto, and T. Oakley. 2007. Estimating a Binary Character’s Effect on Speciation and Extinction. *Systematic biology* 56:701–710.
- Magallon, S., and M. J. Sanderson. 2001. Absolute Diversification Rates in Angiosperm Clades. *Evolution; international journal of organic evolution* 55:1762–1780.
- Magee, A. F., S. Höhna, T. I. Vasylyeva, A. D. Leaché, and V. N. Minin. 2020. Locally adaptive Bayesian birth-death model successfully detects slow and rapid rate shifts. *PLoS computational biology*

- 16:e1007999.
- Marin, J., and S. B. Hedges. 2018. Undersampling genomes has biased time and rate estimates throughout the tree of life. *Molecular biology and evolution*.
- McPeck, M. A. 2008. The ecological dynamics of clade diversification and community assembly. *The American naturalist* 172:E270–84.
- McPeck, M. A., and J. M. Brown. 2007. Clade age and not diversification rate explains species richness among animal taxa. *The American naturalist* 169:E97–106.
- Mooers, A. O., and S. B. Heard. 1997. Inferring Evolutionary Process from Phylogenetic Tree Shape. *The Quarterly review of biology* 72:31–54.
- Moreno, L. F., H. H. Stratton, J. C. Newell, and P. J. Feustel. 1986. Mathematical coupling of data: correction of a common error for linear calculations. *Journal of applied physiology* 60:335–343.
- Morlon, H., T. L. Parsons, and J. B. Plotkin. 2011. Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences of the United States of America* 108:16327–16332.
- Mossel, E., and E. Vigoda. 2005. Phylogenetic MCMC algorithms are misleading on mixtures of trees. *Science* 309:2207–2209.
- Nee, S. 2006. Birth-death Models in Macroevolution. *Annual review of ecology, evolution, and systematics* 37:1–17.
- Ng, J., and S. D. Smith. 2014. How traits shape trees: new approaches for detecting character state-dependent lineage diversification. *Journal of evolutionary biology* 27:2035–2045.
- Ohta, T. 1992. The nearly neutral theory of molecular evolution. *Annual review of ecology and systematics* 23:263–286.
- Omland, K. E. 1997. CORRELATED RATES OF MOLECULAR AND MORPHOLOGICAL EVOLUTION. *Evolution; international journal of organic evolution* 51:1381–1393.
- Otto, S. P. 2018. Adaptation, speciation and extinction in the Anthropocene. *Proceedings. Biological sciences / The Royal Society* 285.
- Pagel, M., C. Venditti, and A. Meade. 2006. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* 314:119–121.
- Pannetier, T., C. Martinez, L. Bunnefeld, and R. S. Etienne. 2020. Branching patterns in phylogenies cannot distinguish diversity-dependent diversification from time-dependent diversification. *Evolution; international journal of organic evolution*.
- Pennell, M. W., R. G. FitzJohn, W. K. Cornwell, and L. J. Harmon. 2015. Model Adequacy and the Macroevolution of Angiosperm Functional Traits. *The American naturalist* 186:E33–E50.
- Penny, D. 2005. Relativity for molecular clocks. *Nature*.
- Prairie, Y. T., and D. F. Bird. 1989. Some misconceptions about the spurious correlation problem in the ecological literature. *Oecologia* 81:285–288.
- Rabosky, D. L. 2014. Automatic Detection of Key Innovations, Rate Shifts, and diversity-dependence on Phylogenetic Trees. *PloS one* 9:e89543.
- . 2016. Reproductive isolation and the causes of speciation rate variation in nature. *Biological journal of the Linnean Society. Linnean Society of London* 118:13–25.
- Rabosky, D. L., and D. C. Adams. 2012. Rates of morphological evolution are correlated with species richness in salamanders. *Evolution; international journal of organic evolution* 66:1807–1818.
- Rabosky, D. L., and A. H. Hurlbert. 2015. Species richness at continental scales is dominated by ecological limits. *The American naturalist* 185:572–583.
- Rabosky, D. L., and I. J. Lovette. 2008. Density-dependent diversification in North American wood warblers. *Proceedings. Biological sciences / The Royal Society* 275:2363–2371.
- Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation.

- Nature communications 4:1958.
- Rabosky, D. L., G. J. Slater, and M. E. Alfaro. 2012. Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS biology* 10:e1001381.
- Raup, D. M. 1985. Mathematical Models of Cladogenesis. *Paleobiology* 11:42–52.
- Reitan, T., and L. H. Liow. 2019. layeranalyzer: Inferring correlative and causal connections from time series data in r. *Methods in ecology and evolution / British Ecological Society* 10:2183–2188.
- Reitan, T., T. Schweder, and J. Henderiks. 2012. Phenotypic evolution studied by layered stochastic differential equations. *The annals of applied statistics* 6:1531–1551.
- Revell, L. J., L. J. Harmon, and R. E. Glor. 2005. Under-parameterized model of sequence evolution leads to bias in the estimation of diversification rates from molecular phylogenies. *Systematic biology* 54:973–983.
- Reznick, D. N., and R. E. Ricklefs. 2009. Darwin’s bridge between microevolution and macroevolution. *Nature* 457:837–842.
- Ricklefs, R. E. 2006. Global variation in the diversification rate of passerine birds. *Ecology* 87:2468–2478.
- Rolland, J., F. L. Condamine, F. Jiguet, and H. Morlon. 2014. Faster speciation and reduced extinction in the tropics contribute to the Mammalian latitudinal diversity gradient. *PLoS biology* 12:e1001775.
- Rolland, J., D. Silvestro, G. Litsios, L. Faye, and N. Salamin. 2018. Clownfishes evolution below and above the species level. *Proceedings. Biological sciences / The Royal Society* 285.
- Rosenblum, E. B., B. A. J. Sarver, J. W. Brown, S. Des Roches, K. M. Hardwick, T. D. Hether, J. M. Eastman, et al. 2012. Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary biology* 39:255–261.
- Rosindell, J., S. J. Cornell, S. P. Hubbell, and R. S. Etienne. 2010. Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology letters* 13:716–727.
- Sadler, P. M. 1981. Sediment Accumulation Rates and the Completeness of Stratigraphic Sections. *The Journal of geology* 89:569–584.
- Schluter, D. 2000. *The ecology of adaptive radiation*. Oxford University Press, Oxford.
- Scholl, J. P., and J. J. Wiens. 2016. Diversification rates and species richness across the Tree of Life. *Proceedings. Biological sciences / The Royal Society* 283.
- Schwenk, K. 1994. A utilitarian approach to evolutionary constraint. *Zoology* 98:251–262.
- Sheets, H. D., and C. E. Mitchell. 2001. Uncorrelated Change Produces the Apparent Dependence of Evolutionary Rate on Interval. *Paleobiology* 27:429–445.
- Silvestro, D., A. Kostikova, G. Litsios, P. B. Pearman, and N. Salamin. 2015. Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution*.
- Slater, G. J., L. J. Harmon, and M. E. Alfaro. 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution; international journal of organic evolution* 66:3931–3944.
- Stadler, T. 2013. How can we improve accuracy of macroevolutionary rate estimates? *Systematic biology* 62:321–329.
- Stanley, S. M. 1979. *Macroevolution, pattern and process*.
- Uyeda, J. C., D. S. Caetano, and M. W. Pennell. 2015. Comparative Analysis of Principal Components Can be Misleading. *Systematic biology* 64:677–689.
- Uyeda, J. C., T. F. Hansen, S. J. Arnold, and J. Pienaar. 2011. The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences of the United States of America* 108:15908–15913.
- Weir, J. T., and D. Schluter. 2008. Calibrating the avian molecular clock. *Molecular ecology* 17:2321–2328.
- Wiens, J. J. 2011. The Causes Of Species Richness Patterns Across Space, Time, And Clades And The Role Of “Ecological Limits.” *The Quarterly review of biology* 86:75–96.
- Yang, Z. 2006. *Computational Molecular Evolution*. Oxford University Press.

## Figure legend:

**Figure 1.** Rate scaling patterns showing rate (x-axis) versus time interval (y-axis); insets are on a log-log scale. A) Rates of molecular evolution versus time interval; estimated mean mitochondrial substitution rates compared with their calibration intervals or oldest sampled tip; rates were calculated on contemporaneous and ancient sequences using Bayesian or coalescent methods under the best substitution model; data compiled from Ho et al. (2007). B) Rates of trait evolution versus time interval; absolute morphometric divergence rate in darwins compared with its measurement time interval estimated for 30 orders among tetrapods, arthropods, molluscs, radiolarians and foraminifera (Uyeda et al. 2011). C) Rates of speciation versus time interval; estimated speciation rates from 104 time-calibrated molecular phylogenies compared with clade's crown age (Hénao Díaz et al. 2019); diversification rates were calculated using BAMM (Rabosky 2014) on representatives of seed plants, ferns, vertebrates and invertebrates.

**Figure 2.** Hierarchical constraints on divergence at multiple scales. A) The state space for evolution is illustrated as a grid, with points representing either phenotypic states or sequences. Neighboring points are 1 evolutionary step away from each other. The red circle represents the set of phenotypic or sequence states that are adapted to a particular environment. Evolution occurs among these allowable states but are constrained by stabilizing selection. Evolution among these states becomes saturated, resulting in a loss of identifiability of evolutionary rate and rate scaling. B) Over longer timescales, the set of evolutionary states that is adaptive may change in response to changing abiotic or biotic conditions. This could represent, for example, change in the adaptive peak. C) These adaptive peaks themselves may encounter further constraints (green circle), for example, representing the limits of an ecological niche, an adaptive zone, or the biomechanical/physiological/biochemical limits for a particular organismal function. If the function or niche of an organism also changes, then the green circle may also change on even longer timescales (Green arrow), where they may eventually encounter still other constraints. These "Russian nested dolls" of selective constraints that change on ever increasing macroevolutionary timescales can ensure that evolutionary rates measured over shorter timescales are always faster than those measured over longer timescales, and that model-based estimates of evolutionary rates that fail to account for these hierarchical constraints will be incapable of accounting for saturation at all levels simultaneously. D) Instantaneous rates of the process in (D) estimated using Brownian motion with increasingly large intervals results in ever decreasing rate estimates.