Phylograms better fit neutrally evolving traits than chronograms

Thibault Latrille[1], **Théo Gaboriau**[1], **Nicolas Salamin**[1]
[1]Department of Computational Biology, University of Lausanne, Quartier Sorge, 1015 Lausanne, Switzerland
thibault.latrille@ens-lyon.org

March 27, 2026

Abstract

At the species level, the evolution of traits is driven by a combination of selective and neutral forces. To disentangle these processes, different scenarios of evolution are modelled and compared. For quantitative traits under selection, the species is usually considered to track a trait optimum, and such an optimum can change along the branches of the species tree. On the other hand, neutral evolution is modelled with a trait changing randomly around the ancestral trait value along the different branches of the species tree. Regardless of the intricacy of modelling trait changes, the species tree is assumed to be known and obtained independently of the trait data analysed. Branch lengths are usually assumed to be in units proportional to time, and the tree is represented by a chronogram. The rationale is that time correlates with trait changes because of its direct connection with the number of generations that have occurred. However, since the generation time of species can also vary along the phylogenetic tree, we argue that their use introduces biases. In contrast, if the phylogenetic tree is represented by a phylogram with branch lengths in units of sequence divergence, this will account for the effect of changing generation time. In this study, we show using simulations that, for a trait evolving neutrally, the fit of a random evolution model has more support on a phylogram than on a chronogram. However, comparing models and testing different scenarios of selection using a phylogram leads to incorrect predictions. Given these results, we argue that we should use phylograms instead of chronograms when claiming that a trait is evolving under drift. Nevertheless, we support the generally accepted use of chronograms to model selection acting on a quantitative trait.

***K*eywords** Comparative studies · Morphological evolution · Genetic drift · Macroevolution · Chronogram · Phylogram

Teaser text

Many species, are characterised by trait differences between each other. For example, the brain size of different primate and human lineages varies. The question we ask is: Does the time that passed explain the differences in their traits? Or are genetic differences between species better at explaining the differences? The standard in evolutionary biology is to use time to explain differences in traits. However, if the trait is not under natural selection and instead is changing due to the accumulation of neutral substitutions (because of genetic drift), the genetic differences between species should better explain the differences in traits than the time since divergence. As a result, we argue that using both genetic differences and species divergence time simultaneously provides a better understanding of how traits evolve across species.

Introduction

By measuring changes in a phenotypic trait along a lineage, it is possible to infer the type of selection acting on it or whether this change is only due to random genetic drift. At a larger scale, from the observed variations of traits across species, regimes of evolution are typically assessed using phylogenetic comparative methods, where traits are modelled as evolving along the branches of the species tree (Felsenstein, 1985, 1988; Harmon, 2018). For example, to model neutral evolution, the mean trait value is said to follow a Brownian motion (BM), branching and evolving independently after each speciation event (Felsenstein, 1985; Hansen and Martins, 1996; Lynch and Hill, 1986). In other words, for each branch of the tree, the value at the descendant node is normally distributed around the ancestral value, with a variance proportional to the branch length. In this framework, reconstructing trait variation along the whole phylogeny as a BM process can thus constitute a null model of neutral trait evolution, although as we discuss later, BM can also approximate non-neutral evolutionary processes.

Alternatively to a simple BM, adding a trend parameter in the BM is interpreted as a signature of directional selection at the phylogenetic scale, reflecting a consistent bias in trait evolution across lineages (Silvestro et al., 2019). However, such trends can be difficult to distinguish from BM in analyses using only extant taxa without fossil calibration (Felsenstein, 1988). More complex models to detect selection have been proposed, notably the Ornstein-Uhlenbeck (OU) processes, where trait variation is constrained around an optimum value, which is often interpreted as a signature of stabilizing selection (Beaulieu et al., 2012; Butler and King, 2004; Hansen, 1997). Methodologically, this alternative model of evolution raises issues since an OU process might be statistically preferred over BM due to sampling artifacts (Cooper et al., 2016; Price et al., 2022; Silvestro et al., 2015). Adding biological realism, the OU process can be relaxed to allow for multiple optima along the phylogenetic tree, which is interpreted as a few abrupt changes in the environment along the phylogeny (Grabowski et al., 2023; Ingram and Mahler, 2013; Khabbazian et al., 2016; Mitov et al., 2020; Uyeda and Harmon, 2014). Finally, the optimum can also be allowed to change continuously along the tree, which results in the trait itself reflecting the movement of the optimum along the lineages due to the constant adaptive evolution of the trait toward the optimum (Hansen, 2024; Hansen and Martins, 1996). One special case of a continuously changing optimum is again the BM, where changes in the optimum are a consequence of the randomly changing environment (Hansen et al., 2008), named fluctuating selection (Holstad et al., 2024) or also diversifying selection due to the diversity of generated phenotypes at the clade level (Latrille et al., 2024). This modelling raises an issue in disentangling selection from neutral evolution: the fact that BM is the best-supported model in a set does not necessarily imply that a trait is evolving neutrally. In other words, Brownian evolution can also result from non-neutral evolutionary processes, such as stabilizing selection on an optimum that moves as a BM in time, or such as fluctuating natural

selection. More generally, due to the central limit theorem, the additive effects of many different neutral and non-neutral random factors might result in evolution that appears Brownian.

To disentangle neutral evolution from different regimes of selection, another approach is to contrast the observed rate of evolution to the neutral expectation (Lande, 1980). The neutral expectation can be obtained if the underlying genetic architecture of the trait is known and the trait is encoded by many loci of additive effects (Barton et al., 2017; Sella and Barton, 2019). Additionally, we can also derive the expected changes in mean trait value along a lineage given the trait variance at the population scale (Felsenstein, 1988; Lynch, 1990; Turelli, 1984, 1988). When generalized to many species at the phylogenetic scale, contrasting between and within-species variations allows us to disentangle neutral evolution from selection tracking a moving optimum. Specifically, Latrille et al. (2024) showed that the ratio of between-species to within-species trait variation should scale linearly with nucleotide divergence for a neutral trait, providing a neutrality test that requires both types of variation data. However, such within-species variation data are not always available, motivating the current study, which focuses solely on between-species variation and tree branch length units. More generally, comparing the rate of evolution at different timescales allows inferring the regime of evolution (Hansen, 2024; Holstad et al., 2024). However, trait variation at different timescales is not always available, and the genetic architecture of the trait is often unknown. Instead, to disentangle neutral evolution from selection tracking a moving optimum in the context of phylogenetic comparative methods, we hereby focus on the underlying tree and the unit of its branch lengths (Fig. 1A).

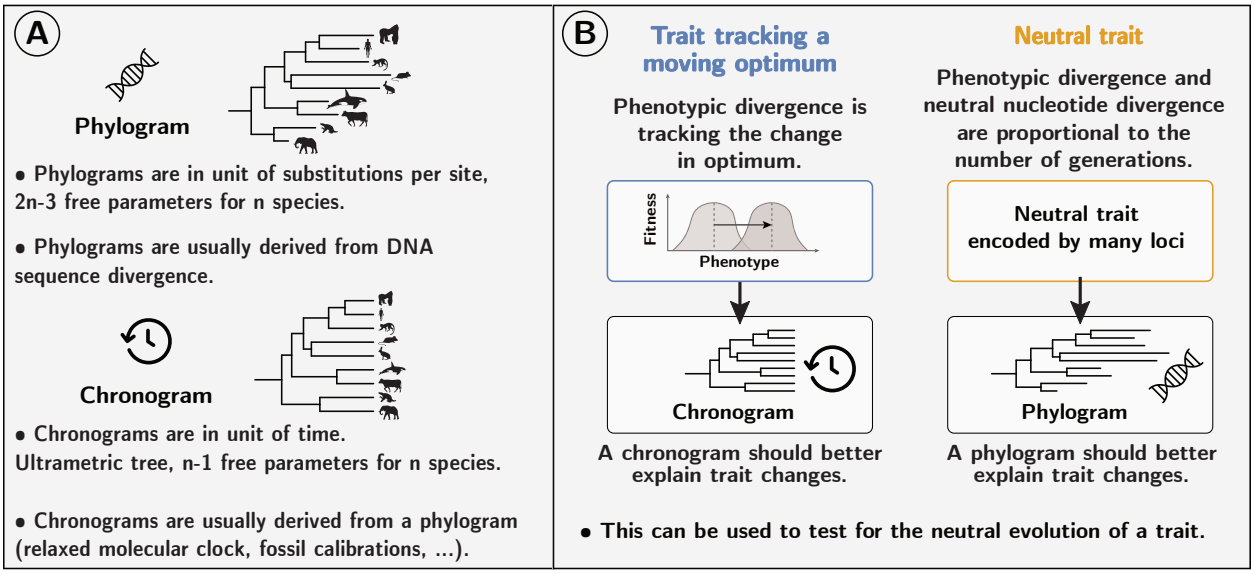

Figure 1: Panel A: Across many species, the evolution of continuous traits can be modelled as a stochastic process evolving along the branches of a phylogenetic tree. The branches of such a phylogenetic tree can be measured in either time (chronogram) or number of substitutions (phylogram). Panel B: Theoretically, changes of a trait under a moving optimum should be better predicted by a chronogram than by a phylogram, since the optimum is expected to change with time rather than with the number of generations. Conversely, changes of a neutral trait should be better predicted by a phylogram than by a chronogram, since neutral trait changes are proportional to the number of generations (and thus nucleotide divergence), not time. Comparing the fit of a BM process on both trees can help disentangle neutral evolution from selection tracking a moving optimum. Alt text: Schematic representation of the rationale to use a phylogram or a chronogram to model trait evolution.

Typically, the tree is assumed to be known and obtained independently, and the *de facto* standard in phylogenetic comparative methods is to use a chronogram, where the branch lengths are proportional to time (Felsenstein, 1985; Harmon, 2018). Theoretically, for a neutrally evolving trait, trait changes depend directly on the number of generations (Hansen and Martins, 1996), which is proportional to time only if the average time between two consecutive generations, called generation time, is constant and equal for all species. Since generation time depends on the time for an individual to reach sexual maturity, it can vary considerably between species, even within a single clade. For example, in mammals, generation time ranges from less than a year in mice to over 20 years in elephants (De Magalhães and Costa, 2009), representing variation that could substantially impact inferences of neutral trait evolution. As a result, variations of generation time between species suggest that modelling neutral evolution as BM on a chronogram induces biases (Litsios and Salamin, 2012).

Alternatively, on a phylogram, branch lengths represent another quantity than time, which can be used as the backbone to model trait evolution and can potentially absorb changes in generation time. For example, it is known that using phylograms in units of morphological distance can be used to improve ancestral trait reconstruction for a discrete character (Wilson et al., 2022). From a genomic perspective, phylograms can also be in units of nucleotide divergence, that is, depicting the number of nucleotide substitutions occurring along a branch. Because the number of neutral nucleotide substitutions accumulating along a branch is proportional to the number of generations (Kimura, 1968, 1983), nucleotide divergence would in this case absorb the effect of changing generation time (i.e. longer branches for lineages with shorter generation time). Such a phylogram, if obtained from neutrally evolving genomic loci would be more appropriate to model a trait evolving neutrally (Latrille et al., 2024). More precisely, under the assumption of a constant genetic architecture, the covariance in mean trait value between a pair of species increases linearly with the number of shared generations, and thus with shared nucleotide divergence for neutral sequences (see Supplementary Material section S1 for the theoretical derivation). Additionally, the rationale to use nucleotide divergence to absorb changes in generation time is also valid for changes in mutation rate (per loci and per generation) under the assumption of a constant genetic architecture of the trait and that changes in mutation rate impact the whole genome (Latrille et al., 2024). Empirically, such changes in mutation rate along a phylogenetic tree are also observed, although to a lesser extent than changes in generation time (Bergeron et al., 2023).

Under the alternative scenario involving selection, shorter generation time and higher mutation rate can also allow for the species to track the changing optimum faster and not lag behind (Hansen et al., 2008; Lande, 1976). However, the timescale for this adaptive lag (typically hundreds to thousands of generations) is negligible compared to the timescale on which the optimum changes at the level of clades (millions of years) in a first approximation (Hansen, 2024). While this lag timescale is difficult to measure empirically, theoretical considerations suggest it should be much shorter than the macroevolutionary timescale. Thus, at the phylogenetic scale, since shifts in the optimum fitness peak are extrinsic to the species and are dependent on the environment which varies with time, using a chronogram would be more accurate to model mean trait changes. As a result, for a trait under selection, mean trait changes should be better explained by the time rather than the number of substitutions that occurred along a branch (Fig. 1B).

In this study, we test this hypothesis, namely whether using a phylogram (in units of nucleotide divergence for neutral loci) instead of a chronogram (in units proportional to time) can allow to assess the evolutionary regime of a trait, by evaluating the fit of BM on both trees. More precisely, we seek to test whether the changes of a neutral trait are better predicted by a phylogram. Conversely, we also seek to test whether changes of a trait under a moving optimum are better predicted by a chronogram. Using genomic information, we seek

to provide additional approaches to model trait evolution, particularly to disentangle neutral evolution from selection tracking a moving optimum.

Methods

To assess whether a phylogram is better suited to model neutral evolution, we first performed simulations of a trait evolving along a phylogenetic tree under different regimes of selection (neutral, moving optimum, multiple optima). Second, we fitted a single-rate Brownian motion (BM) to the simulated data, and we compared the fit of the BM on a phylogram versus a chronogram. Third, based on the single-rate BM, we derived a model to estimate the support of a phylogram over a chronogram that we apply to the simulated dataset. Fourth, we fitted several models of trait evolution to the simulated datasets: a multi-rate BM, an Ornstein-Uhlenbeck (OU) process, and a relaxed OU with multiple optima, tested on a phylogram versus a chronogram. Finally, we gathered an empirical dataset of body and brain masses from mammals, including both a chronogram and a phylogram (in units of nucleotide divergence from neutral loci), on which we tested the support of a phylogram for the single-rate BM. To assess which tree type (phylogram versus chronogram) better supports neutral evolution, we evaluate both the statistical fit of the BM and the accuracy of ancestral trait reconstruction under each tree.

Simulations along a phylogenetic tree

We performed simulations under different selective regimes (neutral, moving optimum, multiple optima). Simulations were individual-based and followed a Wright-Fisher model with mutation, selection and drift for a diploid population including speciation along a predefined ultrametric phylogenetic tree. We used the same simulation framework as in Latrille et al. (2024), with parameters detailed in the Supplementary Material (section S2). The parameters of simulations were chosen to mimic an empirical dataset of mammals. In summary, the trait was encoded by L independent loci (Fig. S1), with each locus contributing additively, and mutations were drawn from a Poisson distribution at each generation. Parents were selected for reproduction according to their phenotypic value, with a probability proportional to their fitness. Flattening the fitness landscape resulted in neutral evolution, which meant that each individual had the same probability of being sampled at each generation regardless of its trait value. Alternatively, for a trait under selection around a moving optimum, we modelled stabilizing selection acting on the trait, with the optimum value changing as a geometric Brownian motion (BM) along the phylogenetic tree (Hansen, 1997; Hansen and Martins, 1996). Finally, for a trait under selection around multiple optima (Uyeda and Harmon, 2014), we draw a Bernoulli variable to test whether there is a switch of optimum along a branch; if a switch occurs, the sliding of the optimum is drawn from a reflected exponential distribution (symmetric positive and negative values).

At each internal node of the tree (i.e. speciation event), the population is split into two daughter populations running independently on each of the two branches, and the process is repeated until the tips of the tree are reached. As a control, we performed simulations with constant mutation rates, generation times and effective population sizes (N_e). Alternatively, we performed simulations with fluctuating mutation rates, generation times and N_e, where we used BM to model the long-term changes along the phylogenetic tree, and we overlaid short-term changes in N_e (see Supplementary Material section S2).

The phylogram is obtained from the same simulation setting on a set of 30,000 independent neutral loci, and the nucleotide divergence is computed as the number of substitutions along each branch. The chronogram is the generating ultrametric time tree used in the simulation. For the empirical datasets, the chronogram is

derived from the phylogram by fitting a relaxed molecular clock model (correlated rate model with penalized likelihood, default value) to the phylogram with the *R* package *ape* (Paradis et al., 2004).

Brownian motion (BM)

On the simulated dataset, we fitted a single-rate BM using either a phylogram or a chronogram as the underlying tree. We used *RevBayes* (Höhna et al., 2016) to fit a Brownian motion (BM) on continuous traits evolving along the branches of a phylogenetic tree (Felsenstein, 1985, 1988). The data consist of the mean trait value for each extant species, and the tree topology is fixed. Along each branch, the value of the trait is drawn from a normal distribution with mean equal to the parent node value and variance equal to the branch length (Felsenstein, 1985). Formally, the BM process is described by the equation:

$$\mathrm{d}P_t = \sigma \cdot \mathrm{d}W_t, \tag{1}$$

where P_t is the trait value at time t, σ is the rate of the BM, and $\mathrm{d}W_t$ is a Wiener process (i.e. a standard Brownian motion). We used a log-uniform prior for σ. Using Markov chain Monte Carlo (MCMC), we reconstructed the ancestral trait value at each node of the tree as the posterior mean estimate (burn-in of 1000 gen., running of 10,000 gen., 2 chains), and we compared the accuracy of the reconstruction using either a phylogram or a chronogram as the underlying tree. Both trees are scaled such that the sum of all the branch lengths is equal to one in each case. Importantly, the data and the tree topology are the same in both analyses; only the branch lengths are different. In this simulation setting, for traits simulated under neutral evolution, we expect more accurate trait reconstruction on a phylogram, and conversely, for traits under selection (moving optimum), more accurate trait reconstruction on a chronogram.

BM with a switch

To test the support of a phylogram over a chronogram, we implemented a model based on a single-rate BM. The model was implemented in *RevBayes* and contains a switch variable, denoted as π that allows the model to switch the branch lengths in units of time to units of substitutions (Fig. 2A). Formally, π is a Bernoulli random variable ($\pi \in \{0, 1\}$) with a prior probability of 0.5. If π is 0, the branch lengths are those of the chronogram, and if π is 1, the branch lengths are those of the phylogram. The tree topology is fixed, meaning both trees have the same branching structure, but because branch lengths are different and not necessarily on the same scale, both trees are re-scaled by dividing branch lengths by the total tree length. As in the previous section, BM is fitted to the data by modelling trait changes as normal distributions (*dnNormal* in *RevBayes*) running along the branches of the tree, with variance proportional to the branch length and the squared value of the Brownian rate parameter (σ, log-uniform prior). Alternatively for large trees or for faster computation, the likelihood can also be estimated by the REML method (*dnPhyloBrownianREML* in *RevBayes*). Altogether, the input data is the mean trait value for extant species and both the chronogram and phylogram with the same tree topology. The posterior mean of π (burn-in of 1,000 gen., running of 10,000 gen., 2 chains) indicates support for the phylogram; we consider the phylogram to be favoured when $\pi > 0.5$. We expect $\pi = 1$ for a trait under neutral evolution, and $\pi = 0$ for a trait under selection (moving optimum).

From a maximum likelihood perspective, the support of a phylogram over a chronogram can also be assessed by model selection, by comparing the maximum likelihood of the data given each tree, and computing the Akaike Information Criterion (AIC) weights of each model (see Supplementary Material section S5).

Alternative models of evolution

In addition to a single-rate BM, we also fitted BM with multiple rate parameters (Eastman et al., 2011), using either a phylogram or a chronogram as the underlying tree to the simulated dataset. The likelihood of the data is estimated by the REML method (*dnPhyloBrownianREML* in *RevBayes*), and the number of rate shifts is estimated by reversible-jump MCMC (*dnReversibleJumpMixture* in *RevBayes*). For each branch, we draw a rate-multiplier either equal to 1 (no rate shift), or drawn from a log-normal distribution with a median of 1, and a standard deviation such that rate shifts range over about one order of magnitude. We set our prior to the expected number of rate shifts of 1 across the whole tree. The number of rate shifts as well as the variance of rate parameters are estimated as the posterior mean (burn-in of 1,000 gen., running of 50,000 gen., 2 chains).

Also, we compared the fit of an Ornstein-Uhlenbeck (OU) process (eq. 2) to BM (Butler and King, 2004; Hansen, 1997), using either a phylogram or a chronogram as the underlying tree. The OU process is described by the equation:

$$\mathrm{d}P_t = -\alpha\,(P_t - \theta)\,\mathrm{d}t + \sigma \cdot \mathrm{d}W_t, \tag{2}$$

where the parameter α in eq. 2 is the strength of the pull towards the optimum, θ is the optimum value and the other parameters are defined as in eq 1. The likelihood of the data is estimated by the REML method (*dnPhyloOrnsteinUhlenbeckREML* in *RevBayes*), and because the OU process has more parameters than the BM, we used a reversible-jump MCMC switch between the two models (0 for BM and 1 for OU, *dnReversibleJumpMixture* in *RevBayes*). The prior for θ is uniform between the minimum and maximum trait value for extant species. The prior for α is a reversible-jump mixture distribution: a value of 0 or drawn from an exponential distribution with mean equal to half the root age divided by $\ln(2)$, meaning that we expect a phylogenetic half-life of half the crown age. Thus, when $\alpha = 0$ the OU process (eq. 2) is equivalent to BM (eq. 1). The half-life, $t_{1/2}$, is defined as $\ln(2)/\alpha$ is also estimated as the posterior mean estimate of finite values (burn-in of 1,000 gen., running of 5,000 gen., 2 chains), representing the time needed for the expected trait value to move half-way toward the optimum (Grabowski et al., 2023; Hansen, 1997),

Finally, we fitted a relaxed OU with multiple optima (Uyeda and Harmon, 2014), using either a phylogram or a chronogram as the underlying tree. The likelihood of the data given the tree is estimated by the REML method (*dnPhyloOrnsteinUhlenbeckREML* in *RevBayes*), and the number of optimums is also estimated by the reversible-jump MCMC method (*dnReversibleJumpMixture* in *RevBayes*) as the posterior mean estimate (burn-in of 1,000 gen., running of 5,000 gen., 2 chains). For each branch, we draw a shift in optimum either equal to 0 (no shift), or drawn from a uniform distribution centred on 0 with breadth spanning the whole range of observed trait value. We set our prior to the expected number of optimum shifts of 1 across the whole tree.

Empirical dataset

We analysed a dataset of body and brain masses from mammals. The log-transformed values of body and brain masses were taken from Tsuboi et al. (2018). We removed individuals not marked as adults and split the data into males and females due to sexual dimorphism in body and brain masses. We also extracted brain and body masses from the *COMBINE* dataset (Soria et al., 2021). The mammalian genomic data are gathered from the Zoonomia project (Genereux et al., 2020). More specifically, the phylogram in units of nucleotide divergence is estimated on a set of neutral markers in Foley et al. (2023). The chronogram is derived from the phylogram by fitting a relaxed molecular clock model (correlated rate model with penalized likelihood, default value) to the phylogram with the *R* package *ape* (Paradis et al., 2004).

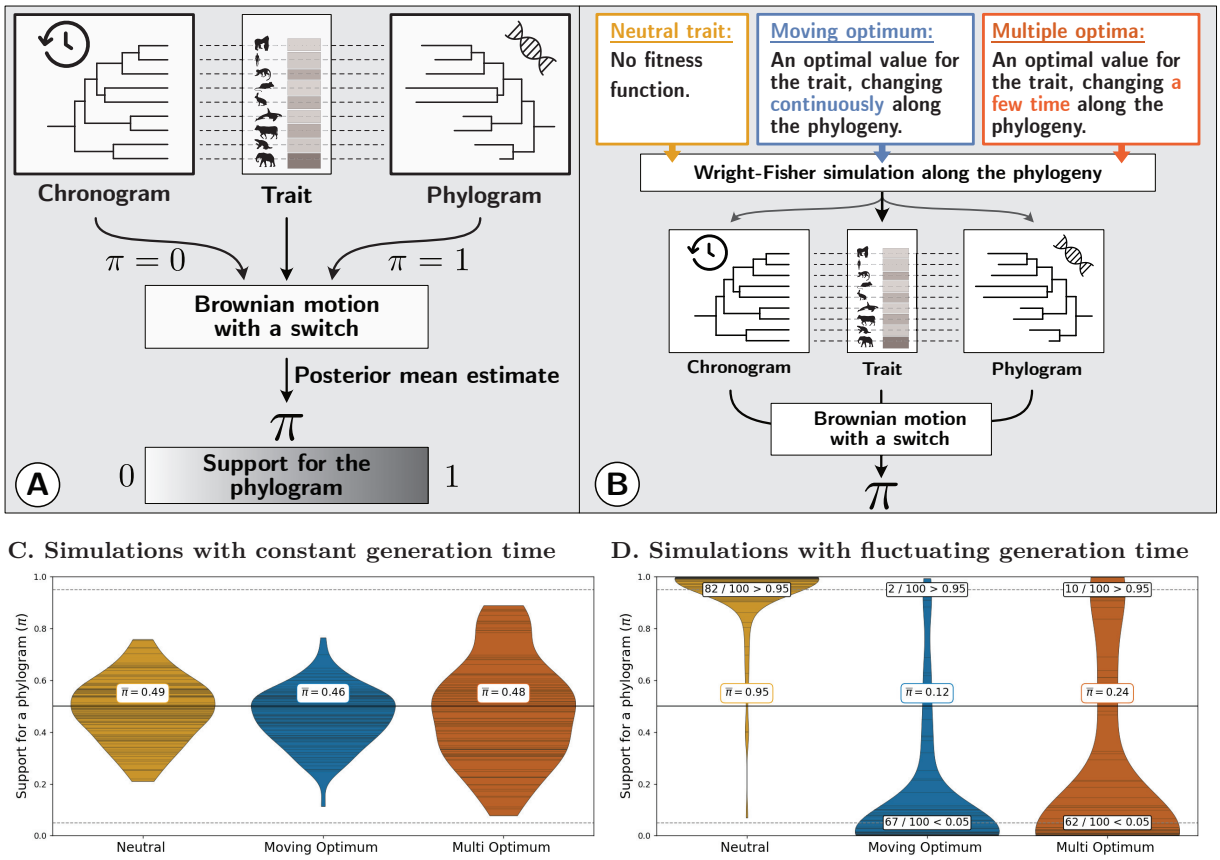

Figure 2: **Testing the support for a phylogram over a chronogram** Panel A: Brownian motion (BM) is used to model trait changes along a phylogeny, given a mean trait value for extant species. The model contains a switch variable, denoted π, that switches the branch lengths in units of time ($\pi = 0$, chronogram) to units of nucleotide divergence for neutral sites ($\pi = 1$, phylogram). The posterior mean estimate of π indicates whether BM supports the phylogram better than the chronogram ($0 \leq \pi \leq 1$). Panel B: Wright-Fisher simulations of trait evolution along a phylogeny. At each node of the tree a speciation event creates two descendant species evolving independently. Neutral evolution is modelled as a constant fitness regardless of the phenotype of individuals (yellow). Selection is modelled by stabilizing selection around an optimum value, which is changing along the phylogenetic tree: as a moving optimum (blue) or as multiple optima (red). The output of the simulation is: 1) the mean trait value for each extant species, 2) the phylogram as obtained from independent neutral markers and 3) the chronogram derived from the phylogram assuming a relaxed molecular clock. The posterior mean estimate of π is inferred from the output of the simulation. Panels C & D: Violin plot of the posterior mean of π across 100 replicates for the different simulated regimes of evolution with mean π values at the center ($\overline{\pi}$). Horizontal lines inside the violins are the estimated π of each replicate simulation. Results of simulations with constant generation time (Panel C) and with fluctuating generation time, mutation rate and effective population size (Panel D). Alt text: Results of simulations testing the support for a phylogram over a chronogram under different regimes of evolution.

Results

Support for a phylogram over a chronogram

We tested the hypothesis that a phylogram is better suited to model neutral evolution than a chronogram. We simulated traits evolving on a phylogenetic tree under different regimes of selection: 1) neutrally evolving, 2) under a moving optimum, and 3) under multiple optima (see Methods). We first assessed the accuracy of ancestral trait reconstruction on a phylogram versus a chronogram, under the assumption that changes follow BM. For a neutral trait, ancestral trait reconstruction was more accurate on a phylogram (Fig. S2A, $r^2 = 0.95$) than on a chronogram (Fig. S2B, $r^2 = 0.83$; Fisher's z-test, $p_{\text{value}} = 9.0 \times 10^{-11}$). In contrast, for a trait evolving under a moving optimum, ancestral trait reconstruction was less accurate on a phylogram (Fig. S2C, $r^2 = 0.79$) than on a chronogram (Fig. S2D, $r^2 = 0.96$; Fisher's z-test, $p_{\text{value}} = 0$).

We then tested for the BM support of phylogram over a chronogram, and we developed a statistic denoted as π (Fig. 2A), which ranges from 0 (full support for the chronogram) to 1 (full support for the phylogram). For simulations with a constant generation time, the support is similar for both a phylogram and a chronogram regardless of the regime of evolution (Fig. 2B-C), which is expected since the branch lengths are equivalent in both trees. Next, we simulated changing generation time, mutation rate (per generation) and effective population size (N_{e}) along the phylogenetic tree, mimicking a mammalian range of changes. In this case, for a simulated neutral trait, the BM was better fitted on a phylogram than on a chronogram, with an average support for phylogram of $\overline{\pi} = 0.95$ across 100 replicate simulations, and 82 of the 100 replicates above the 0.95 threshold (Fig. 2D, yellow). Conversely, for a simulated trait under selection, the fit of BM on a chronogram was better than on a phylogram with an average support for phylogram of $\overline{\pi} = 0.12$ for multiple optima (67 out of 100 replicates below the 0.05 threshold, Fig. 2D, blue) and $\overline{\pi} = 0.24$ for a moving optimum (62 of the 100 below the 0.05 threshold, Fig. 2D, red). Finally, from a statistical perspective, the equivalent of π using the maximum likelihood approach (AIC weights, see Supplementary Material section S5) is highly correlated with our Bayesian estimate ($r^2 = 0.99$, Fig. S3).

On the empirical mammalian dataset from Tsuboi et al. (2018), for body mass the support for the phylogram is $\pi = 0.0$, when sex is not taken into account. When splitting the dataset into males and females, the support for the phylogram is $\pi_{♀} = 0.85$ and $\pi_{♂} = 0.70$. For brain mass, the support for the phylogram is $\pi = 0.0035$ on the mixed dataset, with $\pi_{♀} = 0.56$ and $\pi_{♂} = 0.51$. On the *COMBINE* dataset, which does not distinguish between males and females, the support for the phylogram is $\pi = 0.0$ for body mass and $\pi = 0.0015$ for brain mass (Table S1). We also computed π using an approximate likelihood computation (REML), a method that allow for faster computation and scale better for larger trees, and found similar results (Table S1).

Fitting alternative models of evolution

Besides testing the fit of BM with a single-rate, we also fitted a multi-rate BM (see Methods) to simulated datasets with changing generation time, mutation rate and effective population size (N_{e}). When fitting a multi-rate BM, the estimated variance of rate parameters (v) was lower on a phylogram than on a chronogram for a trait evolving neutrally (Fig. 3A, yellow violins, Wilcoxon paired test with $p_{\text{value}} = 4.5 \times 10^{-13}$), and the number of rate shifts (n) was reflecting the prior on the phylogram while not on the chronogram, showing a bias (Fig. 3B, yellow violins). Conversely, for a trait under a moving optimum, v was higher on a phylogram than on a chronogram (Fig. 3A, blue violins, Wilcoxon paired test with $p_{\text{value}} = 1.3 \times 10^{-14}$), and n was reflecting the prior on the chronogram while not on the phylogram, showing a bias (Fig. 3B, blue violins).

Altogether, fitting BM on a phylogram is more accurate for a trait evolving neutrally, but results in less accurate estimates for a trait under selection.

Moreover, we tested alternatives to the BM with OU models: with a single optimum and with multiple optima (see Methods). Following recommendations by Grabowski et al. (2023), we report the phylogenetic half-life ($t_{1/2} = \ln(2)/\alpha$), which represents the time for a trait to evolve halfway toward a new optimum. The higher the half-life, or in other words, the smaller the α, the more the OU process becomes effectively indistinguishable from BM. For a trait evolving neutrally, the estimated half-life was longer on a phylogram than on a chronogram (Fig. 3C, yellow violins, Wilcoxon paired test with $p_{\text{value}} = 4.9 \times 10^{-5}$), confirming BM-like dynamics on phylograms as expected for neutral evolution. Conversely, for a trait under a moving optimum, the half-life was shorter on a chronogram than on a phylogram (Fig. 3C, blue violins, Wilcoxon paired test with $p_{\text{value}} = 1.5 \times 10^{-12}$), indicating faster adaptation on the appropriate tree. When fitting a multi-OU process on a trait evolving neutrally, the number of optimum shifts (m) was reflecting the prior on the phylogram, but not on the chronogram (Fig. 3D, yellow violins). Conversely, for a trait under a moving optimum, (m) was reflecting the prior on the chronogram but not on the phylogram (Fig. 3D, blue violins).

Discussion

In phylogenetic comparative methods, the evolution of continuous traits is typically modelled as a stochastic process evolving along the branches of a phylogenetic tree. The null model of neutral trait evolution, genetic drift, is thought of as a Brownian motion (BM) running on the tree (Felsenstein, 1988; Lynch and Hill, 1986), and deviations from this model are interpreted as selection acting on the trait (Butler and King, 2004). As a result, a trait on which the BM is favoured over alternative models are sometimes interpreted as a trait evolving neutrally (Catalán et al., 2019; Khaitovich et al., 2006). However, the BM is not only the null model of neutral evolution, but can also model selection, with a trait tracking a moving optimum, where the optimum itself is following BM (Hansen and Martins, 1996; Latrille et al., 2024). More fundamentally, the central limit theorem implies that trait evolution can appear Brownian whenever the trait is influenced by a large number of independent random factors acting additively (Felsenstein, 1985). These factors may include fluctuating genetic architecture, mutational pressures, changing population size, randomly changing environment, genetic drift, and natural selection. Some of these influences may vary in a manner more directly connected to generation time, while others may track real time, making it difficult to predict a priori which tree type should provide the better fit. Therefore, a good fit of BM alone cannot distinguish between neutral evolution and stabilizing selection on an optimum that moves as a BM, and the comparison between phylogram and chronogram provides a crucial additional diagnostic.

Distinguishing these scenarios is not trivial, but we here showed that the underlying backbone tree to model trait evolution can help disentangle neutral evolution from selection tracking a moving optimum. The standard in phylogenetic comparative methods is a chronogram, where the branch lengths are measured in time (Felsenstein, 1985; Harmon, 2018), while phylograms, where the branch lengths are measured in number of substitutions instead, have been generally overlooked. For a simulated trait under selection we assessed that the BM has indeed a better fit on a chronogram than on a phylogram. However, we argue that to model a trait evolving neutrally, the backbone tree should not be a chronogram but instead a phylogram. Phylograms better represent the number of generations that occurred, and in turn should explain best the differences in traits between species (Latrille et al., 2024) In practice, we show that for simulations of a trait evolving neutrally, the fit of BM supports a phylogram rather than a chronogram. Additionally, for a neutral trait, both ancestral state reconstruction and estimation for the number of rate changes for a multi-rate BM

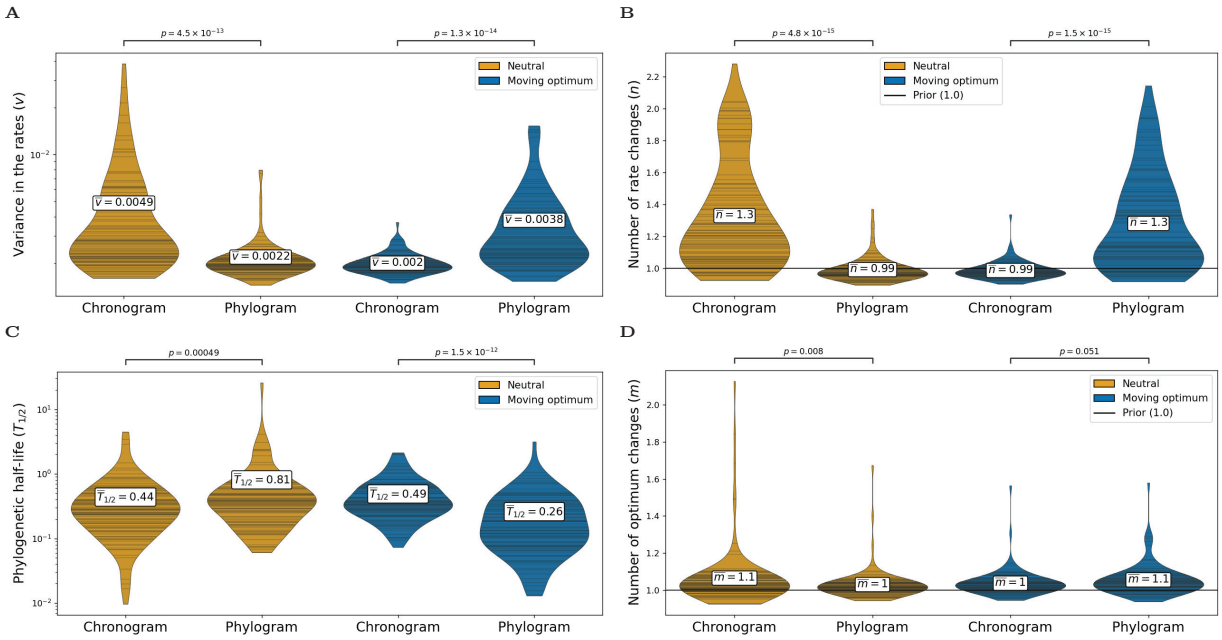

Figure 3: **Mis-classification of trait evolution on a phylogram.** Violin plot of posterior parameter estimates for different models of trait evolution and different simulated regimes of selection. Horizontal lines inside the violin show the results of each replicate simulation. Simulations of 100 replicates per regime: trait evolving under a neutral regime (yellow) and under a moving optimum (blue). Wilcoxon rank-test are performed between the paired estimates on the phylogram and the chronogram. Panel A & B: Fit of a relaxed Brownian motion (BM) with multiple rate parameters, using either a phylogram or a chronogram. Posterior estimate for the variance in rate parameters (Panel A) and number of rate changes (Panel B). Panel C: Estimated phylogenetic half-life ($t_{1/2} = \ln(2)/\alpha$) from an Ornstein-Uhlenbeck (OU) process, using either a phylogram or a chronogram. The half-life represents the time for a trait to evolve halfway toward the optimum; longer half-life indicates BM-like dynamics. Panel D: Fit of a relaxed OU with multiple optima, using either a phylogram or a chronogram. Posterior estimate for the number of optimum changes. Alt text: Violin plot of posterior parameter estimates for different models of trait evolution and different simulated regimes of selection.

was more accurate on a phylogram. Altogether, the combined use of a phylogram and a chronogram can help support the hypothesis of neutral evolution, or to rule it out.

We applied this method to different empirical datasets of body and brain masses in mammals, and we showed that the chronogram was favoured over the phylogram, ruling out the neutral model of evolution for these traits (Latrille et al., 2024). The mammalian dataset is particularly relevant, since changes in generation time are expected to occur along the phylogenetic tree, from 6 months for mice to 50 years for bowhead whales (De Magalhães and Costa, 2009; Nielsen et al., 2016). When analysing each sex separately, the dataset with annotated sex available is smaller (column Taxa in Table S1). Only 46 species have annotated sexes, compared to a range of 125–217 species when not specifying sex. On this dataset of 46 species, neither the phylogram nor the chronogram is favoured, with undecided support ($0.497 \leq \pi \leq 0.848$). This undecided result could reflect insufficient statistical power due to the small number of species, but other explanations are also possible: the true evolutionary process may not conform well to BM on either timescale, or it may

involve a combination of selective and neutral processes that does not clearly favour one tree type over the other.

More generally, our method is also relevant when changes in mutation rate (per generation) and effective population size (N_e) are expected. First, normalizing by nucleotide divergence also accounts theoretically for variations in effective population size (N_e) since the substitution rate of neutral mutations is equal to the mutation rate, such that N_e has no effect on the rate of neutral evolution (Kimura, 1968; Ohta, 1972). Moreover, if population structure were to impact the probability of fixation for neutral mutations (which would also impact a neutral trait), the phylogram would automatically absorb these changes. This argument was partially tested in our simulation by modelling changes in both N_e and generation time (although not modelling population structure *per se*), and the phylogram was favoured over the chronogram for a trait evolving neutrally. Second, under certain conditions, the use of phylograms can also absorb changes in mutation rate (per loci per generation), which in mammals can vary up to 10 fold (Bergeron et al., 2023). As in the case of our simulations, we assumed that the genetic architecture of the trait is constant and the mutation rate is homogeneous across the genome. However, if the mutation rate would only change for a subset of the genome that is encoding the trait, the phylogram would not absorb these changes in mutation rate. Finally, because chronograms are usually derived from phylograms by calibrating the tree with molecular clocks, using phylograms directly has the additional advantage of removing model assumptions required to calibrate node ages (Litsios and Salamin, 2012).

These examples show that while phylograms are useful to model neutral evolution, caution is still warranted. In our simulations, we generally found that using a phylogram tends to favour the OU process over BM. This suggests that trees in units of nucleotide divergence may not always provide a reliable framework for modelling trait evolution, especially when selection is involved. We recommend continuing to use chronograms for modelling selection acting on a trait, as they generally provide more accurate and reliable results. Moreover, for a trait under selection for which the pace of optimum changes is also tracking the changes in generation time, the support for a phylogram over a chronogram would be misleading (Litsios and Salamin, 2012). Therefore, we argue that while phylograms can offer valuable insights (Wilson et al., 2022), they should be used with caution and in conjunction with chronograms to provide a more comprehensive understanding of trait evolution. Contrarily to our cautious note, Wilson et al. (2022) suggested that phylograms are more robust for studying discrete character evolution than chronogram. This apparent contradiction is due to the different units of measurement used in the studies, as they measure branch lengths in units of morphological distance rather than substitutions per site (Wilson et al., 2022). We thus argue that the use of a phylogram, measured in units of substitutions, could also be useful for studying the evolution of discrete characters, but the same caution should be applied.

Caution is indeed warranted, as there are several limitations to our approach. First, we simulated only a subset of different regimes of selection. Notably, we did not simulate a static OU process (with a constant optimum), which would presumably favour an OU model; such simulations are outside the scope of this study since our focus is on distinguishing neutral evolution from selection tracking a moving optimum. As emphasized by Grabowski et al. (2023), the main value of OU models lies in multi-optimum variants that can detect shifts between selective regimes. But more complex scenarios could be included, for example with primary and secondary moving optima (Hansen, 2024).

Additionally, the simulated genetic architecture of a single trait is assumed to be constant across the phylogeny. A more realistic genetic architecture should include pleiotropy, epistasis and linkage disequilibrium, while here we only considered non-linked loci contributing additively to a single trait. Second, while the tested alternative models of evolution (multi-rates BM, OU, multi-optimum OU) are standard, they do

not encompass the full breadth of available methods to model trait evolution (Höhna et al., 2016; Pennell et al., 2014). Finally, the mammalian dataset (brain and body) may not be representative of all traits or species and even though our analysis showcases the utility of phylograms, more empirical studies are needed to replicate the findings.

Gene expression levels (both mRNA and protein levels) represent one class of traits where distinguishing neutral evolution from different regimes of selection is particularly important. While the polygenic nature of expression traits can vary, testing whether their evolution is better explained by phylograms or chronograms can help determine whether changes are driven by accumulation of neutral substitutions or by selection responding to environmental changes. Indeed, the regime of selection acting on gene expression level is the focus of intense debate (Bertram et al., 2023; Dimayacyac et al., 2023; Price et al., 2022; Signor and Nuzhdin, 2018). Typically, datasets on which BM is favoured over other models is interpreted as mRNA expression level evolving neutrally (Catalán et al., 2019; Dimayacyac et al., 2023; Khaitovich et al., 2006). Phenotypic distance is also plotted as a function of time while claiming that a trait is evolving under drift (Jiang et al., 2023). Instead, we argue that phenotypic distance should be expressed in units of substitutions instead of time, using the square root of branch lengths. More generally, from an empirical perspective, the use of nucleotide changes at neutral loci as a normalizing factor brings new perspectives into trait evolution (Latrille et al., 2024).

Altogether, our study supports the use of a chronogram when testing for selection acting on a trait. We propose the phylogram-chronogram comparison as a new tool in the comparative methods toolbox, particularly useful when assessing whether a trait is evolving neutrally or under selection tracking a moving optimum. In such cases, comparing the support for a phylogram versus a chronogram provides an additional diagnostic to assess the regime of evolution acting on a trait. We provide a Bayesian model implemented in *RevBayes* to perform this comparison, which estimates whether BM supports a phylogram or a chronogram ($0 \leq \pi \leq 1$), given the data at the tip of the tree and both trees with the same topology. For empiricists more familiar with maximum likelihood approaches, we also provide an implementation that yields comparable results through AIC weights (see supplementary material section S5). Only if the phylogram has more support than the chronogram (π close to 1) should one consider the neutral model of evolution as a plausible explanation for the evolution of the trait, but claims of neutral evolution should still be made cautiously.

Supplementary material

Supplementary material is available online at *Evolution Letters*.

Data and code availability

The materials that support the findings of this study are openly available in GitHub at github.com/ThibaultLatrille/ChronoPhylogram. Snakemake pipeline, simulator, analysis scripts and *RevBayes* scripts and documentation are available in the repository to replicate the study.

Author contributions

Original idea: T.L.; Model conception: T.L., T.G. and N.S.; Code: T.L.; Data analyses: T.L. and T.G.; Interpretation: T.L., T.G. and N.S.; First draft: T.L.; Editing and revisions: T.L., T.G. and N.S. Project management and funding: N.S.

Funding

This work was funded by Faculté de Biologie et de Médecine, Université de Lausanne (https://www.unil.ch; to TL, TG and NS) and Swiss National Science Fund (https://www.snf.ch; grant 315230-219757 to TL and NS). The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

We gratefully acknowledge the help of Lucy M. Fitzgerald, Diego A. Hartasánchez, Anna Marcionetti and Julien Joseph for their advice and reviews concerning this manuscript. We also thank an anonymous reviewer of our previous study (neutrality index for a trait) for commenting on the use of phylogram to test for neutral evolution, which sparked the idea of this study. Finally, we thank the editor and two anonymous reviewers for their constructive comments and suggestions.

References

Barton, N. H., Etheridge, A. M., and Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, 118:50–73.

Beaulieu, J. M., Jhwueng, D.-C., Boettiger, C., and O'Meara, B. C. (2012). Modeling stabilizing selection: Expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution*, 66(8):2369–2383.

Bergeron, L. A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M. F., Quintard, B., Hoffman, J. I., Li, Z., St. Leger, J., Shao, C., Stiller, J., Gilbert, M. T. P., Schierup, M. H., and Zhang, G. (2023). Evolution of the germline mutation rate across vertebrates. *Nature*, pages 1–7.

Bertram, J., Fulton, B., Tourigny, J. P., Peña-Garcia, Y., Moyle, L. C., and Hahn, M. W. (2023). CAGEE: Computational Analysis of Gene Expression Evolution. *Molecular Biology and Evolution*, 40(5):msad106.

Butler, M. A. and King, A. A. (2004). Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *The American naturalist*, 164(6):683–695.

Catalán, A., Briscoe, A. D., and Höhna, S. (2019). Drift and directional selection are the evolutionary forces driving gene expression divergence in eye and brain tissue of Heliconius butterflies. *Genetics*, 213(2):581–594.

Cooper, N., Thomas, G. H., Venditti, C., Meade, A., and Freckleton, R. P. (2016). A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. *Biological Journal of the Linnean Society*, 118(1):64–77.

De Magalhães, J. P. and Costa, J. (2009). A database of vertebrate longevity records and their relation to other life-history traits. *Journal of Evolutionary Biology*, 22(8):1770–1774.

Dimayacyac, J. R., Wu, S., Jiang, D., and Pennell, M. (2023). Evaluating the Performance of Widely Used Phylogenetic Models for Gene Expression Evolution. *Genome Biology and Evolution*, page evad211.

Eastman, J. M., Alfaro, M. E., Joyce, P., Hipp, A. L., and Harmon, L. J. (2011). A Novel Comparative Method for Identifying Shifts in the Rate of Character Evolution on Trees. *Evolution*, 65(12):3578–3589.

Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*, 125(1):1–15.

Felsenstein, J. (1988). Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, 19(1):445–471.

Foley, N. M., Mason, V. C., Harris, A. J., Bredemeyer, K. R., Damas, J., Lewin, H. A., Eizirik, E., Gatesy, J., Karlsson, E. K., Lindblad-Toh, K., Zoonomia Consortium, Springer, M. S., and Murphy, W. J. (2023). A genomic timescale for placental mammal evolution. *Science*, 380(6643):eabl8189.

Genereux, D. P., Serres, A., Armstrong, J., Johnson, J., Marinescu, V. D., Murén, E., Juan, D., Bejerano, G., Casewell, N. R., Chemnick, L. G., Damas, J., Di Palma, F., Diekhans, M., Fiddes, I. T., Garber, M., Gladyshev, V. N., Goodman, L., Haerty, W., Houck, M. L., Hubley, R., Kivioja, T., Koepfli, K.-P., Kuderna, L. F. K., Lander, E. S., Meadows, J. R. S., Murphy, W. J., Nash, W., Noh, H. J., Nweeia, M., Pfenning, A. R., Pollard, K. S., Ray, D. A., Shapiro, B., Smit, A. F. A., Springer, M. S., Steiner, C. C., Swofford, R., Taipale, J., Teeling, E. C., Turner-Maier, J., Alfoldi, J., Birren, B., Ryder, O. A., Lewin, H. A., Paten, B., Marques-Bonet, T., Lindblad-Toh, K., Karlsson, E. K., and Zoonomia Consortium (2020). A comparative genomics multitool for scientific discovery and conservation. *Nature*, 587(7833):240–245.

Grabowski, M., Pienaar, J., Voje, K. L., Andersson, S., Fuentes-González, J., Kopperud, B. T., Moen, D. S., Tsuboi, M., Uyeda, J., and Hansen, T. F. (2023). A Cautionary Note on "A Cautionary Note on the Use of Ornstein Uhlenbeck Models in Macroevolutionary Studies". *Systematic Biology*, 72(4):955–963.

Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, 51(5):1341–1351.

Hansen, T. F. (2024). Three modes of evolution?Remarks on rates of evolution and time scaling. *Journal of Evolutionary Biology*, page voae071.

Hansen, T. F. and Martins, E. P. (1996). Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. *Evolution*, 50(4):1404–1417.

Hansen, T. F., Pienaar, J., and Orzack, S. H. (2008). A comparative method for studying adaptation to a randomly evolving environment. *Evolution*, 62(8):1965–1977.

Harmon, L. (2018). Phylogenetic comparative methods: Learning from trees.

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., Huelsenbeck, J. P., and Ronquist, F. (2016). RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Systematic Biology*, 65(4):726–736.

Holstad, A., Voje, K. L., Opedal, Ø. H., Bolstad, G. H., Bourg, S., Hansen, T. F., and Pélabon, C. (2024). Evolvability predicts macroevolution under fluctuating selection. *Science*, 384(6696):688–693.

Ingram, T. and Mahler, D. (2013). SURFACE: Detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods in Ecology and Evolution*, 4(5):416–425.

Jiang, D., Cope, A. L., Zhang, J., and Pennell, M. (2023). On the Decoupling of Evolutionary Changes in mRNA and Protein Levels. *Molecular Biology and Evolution*, 40(8):msad169.

Khabbazian, M., Kriebel, R., Rohe, K., and Ané, C. (2016). Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. *Methods in Ecology and Evolution*, 7(7):811–824.

Khaitovich, P., Enard, W., Lachmann, M., and Pääbo, S. (2006). Evolution of primate gene expression. *Nature Reviews Genetics*, 7(9):693–702.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30(2):314–334.

Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics*, 94(1):203–215.

Latrille, T., Bastian, M., Gaboriau, T., and Salamin, N. (2024). Detecting diversifying selection for a trait from within and between-species genotypes and phenotypes. *Journal of Evolutionary Biology*, 37(12):1538–1550.

Litsios, G. and Salamin, N. (2012). Effects of Phylogenetic Signal on Ancestral State Reconstruction. *Systematic Biology*, 61(3):533–538.

Lynch, M. (1990). The Rate of Morphological Evolution in Mammals from the Standpoint of the Neutral Expectation. *The American Naturalist*, 136(6):727–741.

Lynch, M. and Hill, W. G. (1986). Phenotypic evolution by neutral mutation. *Evolution*, 40(5):915–935.

Mitov, V., Bartoszek, K., Asimomitis, G., and Stadler, T. (2020). Fast likelihood calculation for multivariate Gaussian phylogenetic models with shifts. *Theoretical Population Biology*, 131:66–78.

Nielsen, J., Hedeholm, R. B., Heinemeier, J., Bushnell, P. G., Christiansen, J. S., Olsen, J., Ramsey, C. B., Brill, R. W., Simon, M., Steffensen, K. F., and Steffensen, J. F. (2016). Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (Somniosus microcephalus). *Science*, 353(6300):702–704.

Ohta, T. (1972). Population size and rate of evolution. *Journal of Molecular Evolution*, 1(4):305–314.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20(2):289–290.

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., Alfaro, M. E., and Harmon, L. J. (2014). Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30(15):2216–2218.

Price, P. D., Palmer Droguett, D. H., Taylor, J. A., Kim, D. W., Place, E. S., Rogers, T. F., Mank, J. E., Cooney, C. R., and Wright, A. E. (2022). Detecting signatures of selection on gene expression. *Nature Ecology & Evolution*, pages 1–11.

Sella, G. and Barton, N. H. (2019). Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annual Review of Genomics and Human Genetics*, 20(1):461–493.

Signor, S. A. and Nuzhdin, S. V. (2018). The Evolution of Gene Expression in cis and trans. *Trends in Genetics*, 34(7):532–544.

Silvestro, D., Kostikova, A., Litsios, G., Pearman, P. B., and Salamin, N. (2015). Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution*, 6(3):340–346.

Silvestro, D., Tejedor, M. F., Serrano-Serrano, M. L., Loiseau, O., Rossier, V., Rolland, J., Zizka, A., Höhna, S., Antonelli, A., and Salamin, N. (2019). Early Arrival and Climatically-Linked Geographic Expansion of New World Monkeys from Tiny African Ancestors. *Systematic Biology*, 68(1):78–92.

Soria, C. D., Pacifici, M., Di Marco, M., Stephen, S. M., and Rondinini, C. (2021). COMBINE: A coalesced mammal database of intrinsic and extrinsic traits. *Ecology*, 102(6):e03344.

Tsuboi, M., van der Bijl, W., Kopperud, B. T., Erritzøe, J., Voje, K. L., Kotrschal, A., Yopak, K. E., Collin, S. P., Iwaniuk, A. N., and Kolm, N. (2018). Breakdown of brain–body allometry and the encephalization of birds and mammals. *Nature Ecology & Evolution*, 2(9):1492–1500.

Turelli, M. (1984). Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology*, 25(2):138–193.

Turelli, M. (1988). Phenotypic Evolution, Constant Covariances, and the Maintenance of Additive Variance. *Evolution*, 42(6):1342–1347.

Uyeda, J. C. and Harmon, L. J. (2014). A Novel Bayesian Method for Inferring and Interpreting the Dynamics of Adaptive Landscapes from Phylogenetic Comparative Data. *Systematic Biology*, 63(6):902–918.

Wilson, J. D., Mongiardino Koch, N., and Ramírez, M. J. (2022). Chronogram or phylogram for ancestral state estimation? Model-fit statistics indicate the branch lengths underlying a binary character's evolution. *Methods in Ecology and Evolution*, 13(8):1679–1689.