Testing eco-evolutionary predictions using fossil data: Phyletic evolution following ecological opportunity

Kjetil Lysne Voje\textsuperscript{1,2}

\textsuperscript{1}Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway
\textsuperscript{2}E-mail: k.l.voje@ibv.uio.no

Received March 1, 2019
Accepted July 15, 2019

Fossil sequences provide observations of phenotypes within a lineage over time and represent essential data for increasing our understanding of phyletic evolution beyond microevolutionary timescales. I investigate if fossil time series of the diatom \textit{Stephanodiscus niagarae/yellowstonensis} follow evolutionary dynamics compatible with hypotheses for how the adaptive landscape changes when a population enters a new environment. The lineage—which has a remarkably detailed stratigraphic record—invaded Yellowstone Lake immediately after recession of ice from the basin 14,000 years ago. Several phyletic models portraying different types of evolutionary dynamics—both compatible and not compatible with changes in the adaptive landscape following ecological opportunity—were fitted to the fossil times-series of \textit{S. niagarae/yellowstonensis}. Different models best describe the three analyzed traits. Two of the models (a new model of decelerated evolution and an Ornstein–Uhlenbeck model) capture trait dynamics compatible with an event of ecological opportunity, whereas the third model (random walk) does not. Entering a new environment may accordingly affect trait dynamics for thousands of years, but the effects can vary across phenotypes. However, tests of model adequacy reveal shortcomings in all three models explaining the trait dynamics, suggesting model development is needed to more fully understand the phyletic evolution in \textit{S. niagarae/yellowstonensis}.

**KEY WORDS:** Adaptive landscape, decelerated evolution, morphology, time series, OU model.

The adaptive landscape is an important concept that aids our thinking on how evolution across a handful of generations can be connected to evolutionary patterns observed on macroevolutionary timescales (Arnold et al. 2001; Hansen 2012; Arnold 2014). The number of models portraying dynamics of the adaptive landscape on million-year timescales using phylogenies has steadily grown in the last decade (e.g., Harmon et al. 2010; Bartoszek et al. 2012; Beaulieu et al. 2012; Ingram and Mahler 2013; Pennell et al. 2014; Uyeda and Harmon 2014; Khabbazian et al. 2016; Caetano and Harms 2017; Landis and Schraiber 2017; Tarasov 2018). Still, only a handful of models allow for the investigation of how changes in the adaptive landscape affect single lineages (Hunt 2006; Hunt et al. 2008; Reitan et al. 2012; Hunt et al. 2015). Phylogenetic comparative methods and analyses of fossil time series have different strengths in developing our understanding of evolutionary trait dynamics beyond microevolutionary timescales. Phylogenetic comparative methods are powerful tools for investigating the causes and consequences of phenotypic evolution at the level of whole clades (O’Meara 2012; Pennell and Harmon 2013; Garamszegi 2014), but trait dynamics along branches of a phylogenetic tree are rarely observed directly as models are commonly parameterized based on trait means that are (usually) available only at the tips of a phylogeny. Fossil time series provide on the other hand more direct observations of phenotypes over time in a lineage/branch, but are less useful for clade-wide questions. There is often also a difference in time scale. Fossil sequences are most often covering time intervals that ranges from thousands (e.g., Hunt et al. 2008; Van Boxelier and Hunt 2013; Firmat et al. 2014; Chiba 2015; Spanbauer et al. 2018) up to a few million years (e.g., Clyde and Gingerich 1994; Pearson and Ezard 2014), whereas phylogenetic comparative analyses can cover time intervals of several tens and even hundreds of millions of years (e.g., Uyeda
Analyses of fossil time series may therefore provide complementary insight into how microevolution links to macroevolution through detailing how changes in the adaptive landscape shape trait dynamics within lineages across time intervals unattainable by most studies of contemporary populations and comparative studies of species.

Interactions between organisms and their environment shape the adaptive landscape of populations (Fear and Price 1998; Schluter 2000; Svensson and Calsbeek 2012; Voje et al. 2015). This tangled nature of ecological and evolutionary processes has long been recognized (Darwin 1859; Fisher 1930), one example being “ecological opportunity,” a term defined in different ways, but usually referring to a scenario where a lineage is suddenly presented to niche space with a wide range of underexploited resources (Schluter 2000; Rundell and Price 2009; Losos and Mahler 2010; Yoder et al. 2010; Wellborn and Langerhans 2015; Stroud and Losos 2016). Ecological opportunity can arise when a lineage (1) enters a new environment (e.g., geographic colonization of islands, lakes etc.), (2) evolves a key innovation, (3) loses important predators or competitors in its present environment, or (4) gets access to new resources in their current environment (Losos and Mahler 2010; Yoder et al. 2010; Stroud and Losos 2016), scenarios expected to lead to substantial evolution in lineages due to changes in selection regimes (ecological release sensu Yoder et al. 2010).

Ecological opportunity is often linked to evolutionary consequences on the level of clades. Adaptive radiations are a prime example: Ecological opportunity has been invoked as an explanation for how higher taxonomic clades come about through invasion of distinct adaptive zones made available by evolution of key adaptations (Simpson 1944, 1953; Van Valen 1971). Adaptive radiations at lower taxonomic levels (species and genera) have also been explained as a consequence of ecological opportunities when lineages are offered “a wealth of evolutionarily accessible resources little used by competing taxa” (Schluter 2000, p. 69), resulting in a rapid diversification of lineages to different peaks in the adaptive landscape (Benkman 2003; Mahler et al. 2013; Martin and Wainwright 2013). Adaptive radiations predict a decaying rate of phenotypic diversification within clades over time as a consequence of ecological niches filling up with lineages (Blomberg et al. 2003; Harmon et al. 2003, 2010) and the presence of such a signal has been investigated by testing for an “early burst” of morphological evolution within clades (e.g., Cooper and Purvis 2010; Harmon et al. 2010; Slater and Pennell 2013; Pennell et al. 2015; Désamoré et al. 2018). Contemporary evolution has also been documented in many natural populations colonizing new environments (e.g., Berry 1964; Johnston and Selander 1964; Reznick and Ghalambor 2001; Sætre et al. 2017), but the extent single lineages experiencing ecological opportunity follow predicted evolutionary dynamics on time intervals beyond what is possible to observe across microevolutionary timescales has received little attention.

Conceptualizing how ecological opportunity may cause the adaptive landscape to change can aid in depicting scenarios and develop models for testing hypotheses. The appearance of previously unavailable resources can, for example, move or produce new peaks in the adaptive landscape, potentially displacing a population from a peak in the adaptive landscape (i.e., the population becomes maladapted). The consequence of such an adaptive mismatch between important ecological traits and the ecological conditions of the new environment is therefore immediate and strong directional selection on key ecological traits toward a new optimum (Hunt et al. 2008; Lahti et al. 2009). A lineage experiencing this scenario is expected to climb the nearest adaptive peak and stay there due to stabilizing selection, an evolutionary trajectory captured by an Ornstein–Uhlenbeck (OU) model (Hansen 1997; Hunt et al. 2008). A related, but different, scenario that might follow an event of ecological opportunity is mainly relaxations of stabilizing selection on ecologically important traits, causing an initial flattening of the adaptive landscape that allows a wider range of permissible phenotypes (Roughgarden 1972; Lahti et al. 2009). A flattening of the adaptive landscape may cause population variation to increase, making it easier to explore a broader part of the adaptive landscape (Wellborn and Langerhans 2015). A lineage experiencing ecological opportunity may therefore show an initial higher rate of evolution when a wider range of permissible phenotypes are explored before the lineages settles in the new environment and evolution slows down due to the build-up of new stabilizing selection pressures (i.e., the adaptive landscape gets more “hilly” over time). A model of phyletic evolution capturing such trait dynamics is currently lacking.

Analyses of series of ancestor-descendant populations (morphological time series) have contributed to our understanding of evolutionary trait dynamics across various time scales (e.g., Hunt 2007; Hopkins and Lidgard 2012; Voje 2016; Brombacher et al. 2017; Sætre et al. 2017; Voje et al. 2018), but have rarely been used to test predictions from theory on ecological opportunity (but see Hunt et al. 2008). There are good reasons for this. Testing predictions from theory on ecological opportunity demand fossil sequences with a fine-scaled time-resolution spanning a relative short time interval, properties that are relatively rare for fossil data. Additional evidence, like information on when a lineage entered a new habitat, the effects of migration and gene flow from neighboring populations, etc. (Polly 2018), is also needed to convincingly argue that a fossil sequence represents a lineage that has experienced ecological opportunity—evidence that is often lacking for fossil data. Models portraying different types of trait dynamics predicted in a lineage experiencing ecological opportunity have also been lacking. This study aims to meet these challenges.
I analyze trait dynamics in three fossil time series of the diatom *Stephanodiscus niagarae* yellowstonensis, a lineage that invaded Yellowstone Lake soon or immediately after the recession of ice from the basin at the end of the last ice age 14,000 years ago (Theriot 1992; Theriot et al. 2006). The postglacial Yellowstone Lake represented a depauperate community when it was invaded by *S. niagarae* (Theriot et al. 2006), which likely created an ecological opportunity for the lineage due to niche availability that allowed the population to persist in the new environment. The fossil sequence of *S. niagarae* yellowstonensis has a remarkably detailed stratigraphic record (Theriot et al. 2006), allowing a detailed investigation of whether evolution happened according to changes in the adaptive landscape compatible with ecological opportunity. I fit a set of candidate models that portray different types of trait dynamics, both compatible and less compatible with trait dynamics that may follow an event of ecological opportunity. One of the models is a new model of phyletic evolution where the net rate of morphological change slows exponentially through time as evolution proceeds, a model compatible with an initial flattening of the adaptive landscape before the build-up of a more “hilly” landscape with time.

**Materials and Methods**

**DATA ON PHYLETIC EVOLUTION**

The fossil sequence of *S. niagarae* yellowstonensis was published in Theriot et al. (2006) and a detailed description of the methodology to obtain the data can be found there. Here follows a brief summary of the nature of the fossil data.

The fossil sequence of *S. niagarae* yellowstonensis was collected from an 8.55-m long core from the central basin of Yellowstone Lake. A total number of 63 sediment samples were taken from the core, most of them at 16 cm intervals, with additional samples studied at 4–8 cm intervals in parts of the core where the lineage showed the largest evolutionary change. For each sampled level, 50 specimens of *S. niagarae* yellowstonensis were measured, but a lower number were obtained for 14 out of the 63 samples (median number of specimens in the samples not reaching 50 specimens: 25, range 8–42). Three traits were measured on all specimens sampled from each interval: valve diameter (µm), the number of costae per valve, and the number of spines per valve. All three traits are considered to be of ecological importance in diatoms. The size of diatoms strongly influences their growth rate and ability to harvest light and nutrients, and size evolution in diatoms has been linked to environmental forcing due to temperature, nutrient availability, physical mixing, and predation/grazing (Sörhannus et al. 1991; Finkel et al. 2005; Litchman et al. 2009). Spines are common in planktonic diatoms and have been suggested to enhance capture of rare nutrient by increasing rotation and tumbling of the cells by microscale turbulence in the water, and to control sinking and floating in the water column to ensure enough light for photosynthesis (see Van Den Hoek et al. 1995, p. 135, for a discussion of different hypotheses). The structure of the frustule (the hard and silica-rich cell wall) is hypothesized to provide mechanical protection against predators (e.g., Smetacek 2001) and costae (ribs) have been shown to absorb external pressures in such a way that cells avoid being deformed when being predated (Hamm et al. 2003). Evolution of the three traits was analyzed on an additive scale, that is, the traits were not log-transformed prior to analysis. This allows (and restricts) interpreting trait dynamics and rates of evolution using the original scale (micrometer for valve diameter and numbers of spines and costae), which makes it possible to compare the results of the present study to the evolutionary dynamics of these traits as they were reported and interpreted in the original study by Theriot et al. (2006). How these traits change on a proportional scale can also be of interest, however. For sake of completeness, I therefore repeated all analyses on log-transformed traits and report the results in the Supporting Information Material.

The chronology of the Yellowstone Lake sediments were constructed based on age–depth relationships derived from radiocarbon dating of terrestrial macrofossils and ages of two volcanic ashes (Theriot et al. 2006). The time interval covered by the fossil data starts at 13,728 years before present and the lineage was on average sampled every 217 years up until present. *Stephanodiscus niagarae*, the ancestor of *S. yellowstonensis*, invaded Yellowstone Lake right after recession of ice from the basin 14,000 years ago and developed anagenetically into *S. yellowstonensis* during the time interval between 12.4 Ka (kilo-annum) to about 10.2 Ka (Theriot 1992; Theriot et al. 2006). The invasion of Lake Yellowstone likely placed *S. niagarae* in a new environment with low species richness but sufficient resources to support additional species. The fact that *S. yellowstonensis* is endemic to Lake Yellowstone (Theriot 1992) suggests gene flow from other populations is unlikely to have affected the observed evolutionary trait trajectories.

**MODELS OF PHYLETIC EVOLUTION**

I compare the fits of five different models of phyletic evolution to the sequence of ancestor-descendant populations of the *S. niagarae* yellowstonensis lineage. The models stasis, random walk, directional trend, and OU have already been developed for phyletic time series and are implemented in the R package paleoTS (Hunt 2006; Hunt et al. 2008, 2015), whereas a model of decelerated evolution is presented for the first time in this contribution. All five models assume that the population means in the fossil sequence have a joint distribution that is multivariate normal with an expected mean vector and covariance matrix that are functions of the parameters of each model, the time
intervals separating the fossil samples in the sequence, and the sampling variances of the trait means calculated for each fossil sample.

The tree models stasis, random walk, and directional change portray trait evolution not readily predicted for a lineage experiencing ecological opportunity. The models random walk and directional change (Hunt 2006) are described by parameters defining the ancestral trait value at the start of the sequence and the variance ($\sigma^2_{\text{step}}$) and mean ($\mu_{\text{step}}$) of the normal distribution (i.e., $N(\mu_{\text{step}}, \sigma^2_{\text{step}})$) from which evolutionary steps are drawn at each evolutionary transition. The mean of the normal distribution is defined as zero for the random walk so all transitions are assumed to belong to a normal distribution $N(0, \sigma^2_{\text{step}})$. This means the evolutionary dynamics is entirely determined by the variance of the normal distribution, a parameter commonly referred to as the step variance in studies of fossil sequences (Bookstein 1987; Hunt 2006, 2012), which is also an estimate of the rate of evolution in this model (Hunt 2012). The random walk model can be interpreted as a “drift” model, either where the population is assumed to track an optimum that is moving in a random fashion on the adaptive landscape through time (Estes and Arnold 2007; Hunt 2007) or as a description of the effects of genetic drift on the phenotype (Lande 1976). The directional trend model differs from the random walk model in that the mean of the normal distribution is different from zero, that is, transitions come from the distribution $N(\mu_{\text{step}}, \sigma^2_{\text{step}})$ and ($\mu_{\text{step}} \neq 0$). The mean of the distribution reflects the direction of evolution in phenotype space over time, while the estimated variance represents the fluctuations around the trend. The rate of evolution according to the directional trend model is therefore not a single parameter, as the trait dynamics depends on the mean and the variance of the estimated normal distribution (Hunt 2012). The directional trend model can be interpreted as describing trait dynamics where there is on average directional selection in one direction in phenotype space over time, causing a long-term trend in the data (Hunt 2006). The stasis model describes trait fluctuations with a variance ($\omega$) around an optimal/fixed phenotype ($\theta$), basically portraying trait variation as a white noise process, with uncorrelated normally distributed trait values around a fixed mean (Hunt 2006). The stasis model has been interpreted as a case of stabilizing selection around a constant optimum (Hunt 2006; Voje et al. 2018). Note, however, that the stasis model makes assumptions about the trait values coming from a normal distribution ($N(\theta, \omega)$), whereas the first two models are about the trait transitions. Since time does not affect the expected trait divergence in the stasis model, it is not possible to estimate a rate of evolution describing changes between discrete time units (Hunt 2012; Voje et al. 2018).

Ecological opportunity may cause an adaptive mismatch between traits in the population relative to the new environment. One type of evolutionary dynamics predicted from ecological opportunity is therefore directional selection toward a new optimum followed by stabilizing selection after the new optimum is approached. Such a scenario is portrayed by the OU model where a population ascends a peak in the adaptive landscape and shows fluctuating evolution around the peak after the optimum has been reached (Lande 1976; Martins and Hansen 1996; Hansen 1997; Butler and King 2004; Hansen et al. 2008; Hunt et al. 2008). The model assumes the ancestral population is displaced from the optimum (following Lande (1976), see Uyeda et al. (2011) for a relaxation of this assumption) and the strength of selection toward the optimum is increasing with the distance between the optimum and the population mean. The OU model fit is determined by four parameters: the ancestral trait mean ($z_0$), the optimal phenotype ($\theta$) defining the peak in the adaptive landscape the trait evolves toward, the strength of directional selection toward the optimal phenotype ($\alpha$), and the variance of the normal distribution ($\sigma^2_{\text{step}}$), which represents trait fluctuations around the trend toward the optimum. The rate of evolution according to an OU model is time variant as it depends on how displaced the population is from the optimal state ($\theta$) in combination with the strength of the pull toward the optimum ($\alpha$) and the parameter summarizing the trait variance around the optimum ($\sigma^2_{\text{step}}$) (Hunt 2012). The expected time it takes the population to evolve halfway to the new optimum in a OU process is given by $\ln(2)/\alpha$ (Hansen 1997).

THE DECELERATED-EVOLUTION MODEL

A second evolutionary outcome compatible with ecological opportunity is an initial flattening of the adaptive landscape. This may lead to an initially high capacity for evolution as a large range of phenotypes are permissible, followed by a decrease in the rate of change when the lineage is settling in the new niche. Here, I develop a decelerated model of phyletic evolution, which I model as a random walk where the step variance (the variance of the normal distribution from which evolutionary steps are drawn) is reduced exponentially through time. The decelerated-evolution model is closely related to the early burst model developed for phylogenetic comparative data (Blomberg et al. 2003; Harmon et al. 2010), but their interpretations differ. The decelerated-evolution model describes trait dynamics that is compatible with a single lineage entering a new environment (see Discussion for alternative interpretations of this model), whereas the early burst model portrays a slowdown in diversifying morphological evolution at the clade level when ecological opportunities are exhausted. As for the random walk model, the expected evolutionary divergence between ancestor and descendant populations is always zero in the model of decelerated evolution, meaning the expected trait value for each population in the fossil sequence ($z_i$) is identical to the ancestral trait mean ($z_0$):

$$E[z_i] = z_0.$$
The variance around this expectation depends on two parameters: the initial value for the step distribution ($\sigma_{\text{step.in}}^2$) and a parameter describing the exponential decay in the $\sigma_{\text{step.in}}^2$ parameter through time ($r$):

$$\text{Var}[z_i] = \sigma_{\text{step.in}}^2 e^{rt_i} + \epsilon_i,$$

where $t_i$ is the time interval from the ancestral population mean (the start of the fossil sequence, which has a time of 0) to the $i$th population mean. Estimation (sampling) error of the population means contribute to the expected variance between two population means. This estimation error ($\epsilon_i$) is equal to the population sample variance divided by the number of measured specimens for that sample and is added to the diagonal of the variance–covariance matrix. The rate of evolution will decrease as a function of time according to the model and the rate of evolution (the width of the normal distribution from which evolutionary steps are drawn) has a half-life given by $-\ln(2)/r$, given that $r$ is constrained to be negative. The covariance among population means is given by

$$\text{Cov}[z_i, z_j] = \sigma_{\text{step.in}}^2 e^{r t_{\text{min}}},$$

where $t_{\text{min}}$ is the time interval from the ancestral population to the oldest of the two populations $z_i$ and $z_j$. Given the assumption of multivariate normality of population means in the sequence of ancestor descendant populations, their expected distribution is given by their first, second, and mixed moments (covariance).

Optimization and fitting of the decelerated model of evolution is implemented in R and the code has been written so that this model is compatible with how models are optimized and compared using the paleoTS package (where the models stasis, random walk, directional change, and OU are implemented). In short, the paleoTS package uses the built-in optimization routines for estimating maximum likelihood parameter estimates. The default hill-climbing optimization technique used in paleoTS is a quasi-Newton method that constrains the optimization of certain parameters. For example, the optimization of the $\sigma_{\text{step}}^2$ (and the $\sigma_{\text{step.in}}^2$) parameter is constrained to be zero or larger for all models, whereas the $r$ parameter (unique to the decelerated model of phylogetic evolution) is constrained to be negative (to force a reduction in the rate of evolution with time).

I ran simulations to investigate to what extent Akaike information criterion (AICc) is able to distinguishing among the five candidate models (stasis, directional change, random walk, OU, and decelerated evolution) when the decelerated-evolution model is generating the trait dynamics (procedure described in the Supporting Information Material). The support for the decelerated model is steadily increasing with sequence length. The model is well supported (>90%) for time series with a sequence length larger than 30 and is always supported for time series with a sequence length >60 (Fig. S1). I also investigated to what extent the estimation procedure was able to correctly estimate the $r$ parameter (procedure described in the Supporting Information). I find that the average and median parameter estimates of $r$ tend to be somewhat overestimated for shorter sequence lengths, but are approaching the true value for increasing sequence lengths (Fig. S2).

**MODEL ADEQUACY**

AICc, likelihood ratio tests, etc., lack the ability to reject all candidate models even when all of them represent poor descriptions of the observed trait dynamics. Tests of model adequacy may therefore help us avoid making meaningless interpretations of model parameters that do not describe the data well (Pennell et al. 2015; Voje et al. 2018). Tests of model adequacy for the phyletic models stasis, random walk, and directional trend have been developed and implemented in the R package adePEM (Voje 2018). Here, I describe three statistical tests that enable assessment of the extent to which the OU model and the decelerated-evolution model represent adequate descriptions of phylogetic trait dynamics. The procedure for assessing model adequacy is to evaluate how likely it is that a particular model X with parameters Y can produce trait dynamics similar to what is observed in a time series Z. This is assessed by a parametric bootstrap approach: A model is judged as an adequate statistical representation of the trait dynamics in a particular dataset if the results of statistical tests on the observed data are similar to the same test statistics calculated on a large number of simulated time series generated using the investigated model. The confidence in a model as an accurate description of the trait dynamics is increased if the model is able to reproduce properties of the trait dynamics observed in the real data.

Two of the adequacy tests for the decelerated-evolution model and the OU model are similar to the adequacy tests already developed for the models random walk, stasis and trend (Voje 2018; Voje et al. 2018). These two tests investigate whether the data contain the expected levels of autocorrelation and the number of runs (one run is a sequence of consecutive numbers with the same sign on the residuals from the expected trait value) assuming either the OU model or the decelerated-evolution model produced the data. The third test investigates if the distance traveled in morphospace is higher in the beginning of the time series compared to later in the time series, an expectation shared by the OU and the decelerated-evolution models (but not expected in the models stasis, random walk, and directional trend, see Fig. S3).

The Supporting Information Material contains a more detailed description of the three adequacy tests and their rationale. The Supporting Information Material also contains a description of a simulation study conducted to assess if the type I error rates of the adequacy test for the decelerated-evolution and OU models are of an expected magnitude. Figure S4 shows that the three test statistics work as intended, with the runs test potentially having a slightly larger type I error than 0.05.
A detailed description of the parametric bootstrap procedure to assess model adequacy of phyletic models can be found in Voje (2018), but the following example illustrates how the approach works: We want to assess if the decelerated-evolution model represents an adequate statistical description of a fossil time series by investigating if the level of autocorrelation in the data is within the expected range. We first estimate the parameters in the decelerated-evolution model using maximum likelihood by fitting it to the time series. Second, we calculate the autocorrelation on the detrended time series (see Supporting Information Material for information on how and why the data are detrended). We then proceed by simulating a large number of time series using the estimated model parameters from the true data and calculate the test statistic (level of autocorrelation) on all the detrended simulated datasets. Finally, we investigate if the autocorrelation we calculated on the observed fossil time series fall within the distribution of autocorrelations calculated on the simulated data. The model passes the adequacy test if the test statistic (autocorrelation) calculated on the real data falls inside 95% of the calculated test statistic on the simulated time series. This criterion is also applied on the runs test, while the rate of evolution test is considered passed if 80% of the simulated time series have a positive sum of residuals (see the Supporting Information Material for detail on this point).

The adequacy tests for the OU and the decelerated-evolution models have been implemented in the r package adePEM, available on GitHub (github.com/kivoje/adePEM).

ANALYSES OF DATA

I fitted the models stasis, random walk, directional trend, and OU to each of the three fossil time series using the paleoTS package (version 05-1), and fitted the decelerated evolution model using my own R code, specifying joint parameterization for all models. To account for variation in model complexity in the different candidate models, I estimated relative model fit using the small sample-corrected version of the AICc (Akaike 1974; Anderson et al. 2000). I also calculated Akaike weights, which is standardized AICc scores such that scores across all candidate models sum to unity. AICc weights correspond to the relative likelihood of each of the candidate models. Time from first to last population in the time series was scaled to unit length before fitting the models to ease parameter estimation and interpretation of model parameters. To investigate to what extent temporal trait changes in the three traits are correlated, I calculated Spearman Rho of first differences between all pairwise trait combinations.

The extent to which the models showing the best relative fit according to their AICc scores also showed an absolute fit to the data was assessed by running adequacy test implemented in the adePEM package version 1.1. The number of simulated datasets in the bootstrap routine was set to 1000.

R code and data to reproduce all analyses have been uploaded to Dryad (https://doi.org/10.5061/dryad.80725jn/1).

Results

Three different models show the best relative fit to each fossil time series (Fig. 1, Table 1), despite the fact that trait evolution is correlated in these traits (Spearman Rho of first differences between costae and diameter = 0.89; costae and spines = 0.73; diameter and spines = 0.76). The decelerated-evolution model has a much better AICc score compared to the candidate models in describing the trait dynamics in the number of spines per valve. The random walk model and the OU model have the best relative fit to the evolution of costa and valve diameter, respectively, but other models have similar (yet worse) degree of fit describing the dynamics in these two traits.

Parameter estimates for the three models with the best support are shown in Table 2. The rate of evolution in the decelerated-evolution model describing changes in the number of spines is halved about 1700 years after the S. niagarae yellowstonensis lineage entered Lake Yellowstone (Fig. 1). The expected amount of time it takes the population to evolve halfway to the new optimum in a OU process is given by ln(2)/alpha (Hansen 1997), which is 1740 years for the valve diameter. All model parameters need to be interpreted with care, however, as the adequacy tests indicate that key assumptions of the best models according to AICc are violated (Fig. 2, Table 3). For example, the number of runs is too large in the time series on spines compared to the number of runs predicted by the decelerated model. Likewise, the number of runs in the valve diameter is deviating from the number predicted by the OU model (the adequacy test is passed, but only barely), whereas the random walk model fails the autocorrelation test for the data on costae. Although all fossil time series pass most adequacy tests, these results indicate that none of the models showing best relative fit can fully account for the observed trait dynamics in an absolute sense. I therefore also investigated the models showing second best relative fit to the trait dynamics in the number of costae and valve diameter using adequacy tests, because these second-best models only had a slightly worse fit compared to the best model based on AICc. The random walk model did not pass the autocorrelation test for the valve diameter, whereas the directional trend model passed all tree adequacy tests for the evolution of the number of costae. The parameter estimates of the trend model N(tstep = 14, tstep = 2433) show that the step size is about 175 times larger than the mean of the normal distribution. This means the trait dynamics is almost exclusively driven by the variance around the trend (Hunt 2012), which explains why a more trend-like behavior is not clear from eyeballing the data (see Fig. 2).
Figure 1. Plots of population mean trait values over time (in years). Year zero marks the time when the \textit{S. niagarae/yellowstonensis} lineage invaded Lake Yellowstone. Vertical error bars through the population means represent one standard error. The box plots show the AICc weights for each model (see Table 1 for more details on the relative fit of each model). The broken vertical line in the upper plot denote the time where the rate of evolution has been reduced by 50% compared to the initial rate at the start of the fossil sequence according to the decelerated model of evolution. The expected evolutionary trajectory for the valve diameter based on the OU model is also presented. Table 2 contains parameter estimates for the three models showing the best relative fit to each of the fossil time series.

**Discussion**

Evolutionary consequences of ecological opportunity have commonly been investigated on the level of clades, often in connection to adaptive radiations (Simpson 1944, 1953; Schluter 2000; Runnell and Price 2009; Losos and Mahler 2010; Yoder et al. 2010; Wagner et al. 2012; Stroud and Losos 2016; Slater and Friscia 2019). However, ecological opportunity is also predicted to have consequences for the evolutionary dynamics of single lineages (Roughgarden 1972; Hunt et al. 2008; Lahti et al. 2009) and adaptations have indeed been documented in contemporary populations entering a new environment (reviewed in Reznick and Ghalambor 2001). Testing evolutionary dynamics predicted from ecological opportunity in single lineages demands an excellent stratigraphic record with large sample sizes of individuals from each population (fossil sample). Both these demands are met for the fossil sequence of \textit{S. niagarae/yellowstonensis}, a lineage that invaded a depauperate community in the postglacial Yellowstone Lake soon or immediately after the ice cap covering the lake retreated about 14,000 ago (Theriot et al. 2006). Here, I tested whether phenotypic evolution in the \textit{S. niagarae/yellowstonensis} lineage, covering a time interval rarely accessible for detailed
Table 1. Relative model fit of the five candidate models.

| Model                  | LogL    | K | AICc    | Akaike weight |
|------------------------|---------|---|---------|---------------|
| **Number of spines per valve** |         |   |         |               |
| Directional trend      | −119.37 | 3 | 245.15  | 0.000         |
| Random walk            | −119.54 | 2 | 243.28  | 0.001         |
| Stasis                 | −166.59 | 2 | 337.38  | 0.000         |
| OU                     | −117.86 | 4 | 244.42  | 0.001         |
| **Decelerated evolution** | −111.66 | 3 | 229.72  | 0.998         |
| **Number of costae per valve** |         |   |         |               |
| Directional trend      | −211.76 | 3 | 429.92  | 0.221         |
| Random walk            | −211.81 | 2 | 427.81  | 0.636         |
| Stasis                 | −245.45 | 2 | 495.10  | 0.000         |
| OU                     | −211.75 | 4 | 432.20  | 0.071         |
| **Valve diameter**     |         |   |         |               |
| Directional trend      | −168.25 | 3 | 342.91  | 0.086         |
| Random walk            | −168.31 | 2 | 340.82  | 0.245         |
| Stasis                 | −206.83 | 2 | 417.86  | 0.000         |
| OU                     | −165.20 | 4 | 339.09  | 0.579         |
| Decelerated evolution  | −168.20 | 3 | 342.81  | 0.090         |

The best models are given in bold. \(\text{logL} = \text{log-likelihood}; K = \text{number of model parameters}; \text{AICc} = \text{small sample-corrected Akaike information criterion}\).

Table 2. Maximum-likelihood parameter estimates for the models showing best relative fit to the three fossil time series.

| Model   | Trait               | \(z_0\)         | \(\theta\) | \(\alpha\) | \(\sigma^2_{\text{step/init}}\) | \(r\)         |
|---------|---------------------|-----------------|------------|------------|-------------------------------|-------------|
| Decel   | No. of spines per valve | 18.28 (0.79)    | 145.56 (0.00)  | −5.63 (1.42) |
| RW      | No. of costae per valve | 55.14 (3.34)    | 2380.96 (0.00) |                  |
| OU      | Valve diameter (\(\mu\) m) | 40.69 (1.77)    | 49.17 (2.22)  | 691.60 (182.81) | 12.03 (5.46) |

The parameters are the ancestral trait value at the start of the fossil sequence \(z_0\), the position of the adaptive peak in phenotype space in the OU model \(\theta\), the step variance \(\sigma^2\), the strength of the pull toward the optimum in the OU model \(\alpha\), the (initial) value for the step distribution \(\sigma^2_{\text{step/init}}\), and the parameter describing the exponential decay in the rate of evolution through time \(r\) in the decelerated-evolution model. Numbers in parentheses are approximate standard errors based on inverse of the negative Hessian matrix. *, the approximated standard error is < 0.0001 and should not be trusted.

analyses of phyletic changes, is compatible with trait dynamics predicted from ecological opportunity.

Different phyletic models are favored describing the evolutionary dynamics in three traits in the fossil sequence of the diatom lineage \(S.\ niagareayellowstonensis\). The two models describing changes in valve diameter and the number of spines are compatible with predicted changes in the adaptive landscape following an event of ecological opportunity. The decelerated-evolution model showed the best relative fit to the evolution of spines, a model compatible with a scenario in which the lineage experienced relaxation of selection on ecologically important traits due to an initial flattening of the adaptive landscape following ecological opportunity (Roughgarden 1972; Lahti et al. 2009). This flattening of the adaptive landscape will reduce the strength of directional selection, but may lead to an initially high capacity for evolution as a large range of phenotypes are permissible, before the lineage is settling in its new environment (and stabilizing selection is building up as a consequence of a more “hilly” adaptive landscape). The OU model best described the evolution of valve diameter and this model is consistent with a scenario that should be common during an event of ecological opportunity: The new environment displaced the population from its previous position on a peak to a “hillside” on the adaptive landscape, causing the trait to evolve toward its new and closest optimum. Factors that were different in Lake Yellowstone compared to the ancestral environment of \(S.\ niagarae\) and that caused the evolutionary changes in the lineage are difficult to know with certainty. Theriot and colleagues (2006) are cautious in pointing to specific selective factors driving the morphological changes in the \(S.\ niagareayellowstonensis\) lineage, but note that the assemblages of diatom species in the lake, sediment lithology, biogenic silica concentrations, and regional vegetation underwent substantial changes during the first
Figure 2. Histograms showing the distribution of test statistics calculated on the simulated time series. Red dashed vertical lines indicate the values of the test statistics from the observed fossil time series. Bar plots A–C (first row) are the results of the adequacy tests of the decelerated-evolution model in explaining the trait dynamics of the number of spines. In the case of the runs test, the test statistic calculated on the observed time series (red dashed vertical lines) is in the 2.5% tail of the distribution of the simulated test statistic, which means this test of model adequacy failed. The model passed the two other adequacy tests. The bar plots D–F (second row) are the results of the adequacy tests of the random walk model in explaining the evolution of the number of costae. This model passed all adequacy tests except the autocorrelation test. Bar plots G–I (bottom row) show the results of the adequacy tests of the OU model in explaining the trait dynamics of the valve diameter. The OU model passes all tests, but is very close to fail the runs test. See Table 3 for more details on each adequacy test. Note that the adequacy test for initial rapid change (test applies to the OU model and the decelerated-evolution model) is passed if 80% of the simulated time series have a positive sum of residuals (see Supporting Information Material for more information on the three adequacy tests).
Table 3. Results of adequacy tests for each model with the best relative model fit to each of the three fossil time series.

| Trait                        | Model | Autocorrelation | Runs | Fixed variance | Initial rapid change |
|------------------------------|-------|-----------------|------|----------------|----------------------|
| No. of spines per valve      | Decel | 0.110           | 0.008| NA             | 1.000                |
| No. of costae per valve      | RW    | 0.014           | 0.140| 0.052          | NA                   |
| Valve diameter               | OU    | 0.318           | 0.030| NA             | 0.836                |

The P-values indicate the fraction of simulated test statistics that is larger (or smaller) than the calculated test statistic on the observed data, divided by 0.5. A value of 1 means 50% of the test statistics on the simulated data is larger and smaller than the calculated statistic on the observed data. A P-value larger or smaller than 0.975 and 0.025 indicates a failed adequacy test (bold), except for the “initial rapid change” adequacy test where a P-value smaller than 0.800 indicates a failed test.

Decel, decelerated-evolution model; RW, random walk; OU, Ornstein–Uhlenbeck model.

thousand years after the retreat of the ice cap covering the lake (Theriot et al. 2006). Concentrations of diatoms are low at the basal sediments of the investigated core from Yellowstone Lake, but the absolute abundance of diatoms increase in the lake around 12,000 years before present (see Fig. 9 in Theriot et al. 2006). Fluctuations in the abundance of three investigated diatom lineages (S. minutulus, S. niagarae/yellowstonensis, and Aulacoseria subarctica) are large between 12,000 years and 8000 years, but stabilize after that, with S. niagarae/yellowstonensis being most abundant. Theriot et al. (2006) mention that this increased abundance of S. niagarae/yellowstonensis coincides with increased nitrogen limitation during the early Holocene, and point out that S. yellowstonensis is a better competitor for nitrogen than S. niagarae. Presumably, some of the changes in the environment might have constituted both relaxed and new selection pressures on the S. niagarae/yellowstonensis lineage, implicating a role of ecological release (sensu Yoder et al. 2010) on the observed trait dynamics. Ecological opportunity may therefore have important effects on the shape of the adaptive landscape of a population for thousands of years. A similar conclusion was argued for in a study by Hunt et al. (2008), where the authors showed that the reduction in armor traits in evolutionary time series of a stickleback lineage best fitted an OU model. The absence (or presumed low density) of predators in the environment was suggested as a likely explanation of the evolutionary changes in the traits.

The third trait analyzed in the S. niagarae/yellowstonensis lineage is best described by the random walk model. This is a model commonly found to fit fossil time series (Hunt 2007; Voje 2018) and describes trait dynamics not directly predicted from an event of ecological opportunity. Different traits in the same lineage can evolve according to different modes of evolution that have been thoroughly documented (Hopkins and Lidgard 2012). The different evolutionary dynamics among traits observed in the S. niagarae/yellowstonensis lineage therefore suggest that ecological opportunity have distinctive effects on different part of the phenotype, that is, ecological opportunity may shape the adaptive landscape differently for different traits. Note, however, that changes in all the tree traits are rather strongly correlated, which implies rather similar evolutionary dynamics across the three traits. This suggests that small differences in trait dynamics can cause a different model to fit best according to AICc.

The nested nature of some of the models can partly explain why this happen, for example, both the OU and the decelerated model become a random walk when the alpha and the \( r \) parameter are zero, respectively. The level of correlated changes in the three investigated traits in the S. niagarae/yellowstonensis lineage may also suggest that natural selection did not act directly on all three traits individually and that parts of the trait evolution may be due to indirect selection. It is also possible that the reduced rate of evolution with time of the number of spines according to the decelerated-evolution model is partly due to analyzing the data on an absolute scale as the analysis of the data on a proportional scale is found to fit an OU-model (see Supporting Information).

Models may have different interpretations based on the context they are applied. For example, the OU model is often interpreted differently depending on the time scale of the data analyzed (see, e.g., Hansen et al. 2008). Likewise, a random walk is commonly used to model effects of genetic drift on the phenotype (Lande 1976), but has also been interpreted as random changes in a peak on the adaptive landscape on macroevolutionary timescales (Estes and Arnold 2007; Hunt 2007). It seems unlikely that an event of ecological opportunity should have detectable effects on the evolution of a lineage across millions of years, which suggests the interpretation of the decelerated model depends on the time interval covered by the data and the ecological and environmental context of the lineage. An evolutionary time series that fits the decelerated model well and covers a time interval of millions of years may still make sense, but would most likely demand a different interpretation than dynamics predicted from ecological release. For example, it can be argued that trait dynamics is expected to slow down over time if biotic and abiotic parts of
the environment become more stable (vary less). It should be noted that Theriot and colleagues (2006) point out that the period of rapid morphological change in *S. niagarae/yellowstonensis* coincides with rapid environmental change in Yellowstone Lake.

The models fitted to the *S. niagarae/yellowstonensis* data are relatively simple, which limits the detail in the interpretation of the inferred trait dynamics. Also, tests of model adequacy suggest that the observed trait dynamics in the *S. niagarae/yellowstonensis* lineage are not perfectly captured by any of the models with best AICc score. That the relatively simple OU model and the even simpler decelerated-evolution and random walk models fail to fully capture trait dynamics in a very well-sampled fossil sequence like the *S. niagarae/yellowstonensis* lineage is not surprising, but the way these models fail some of the adequacy tests might indicate which model developments it might be worthwhile exploring. The failure of the OU model to pass the runs test—and the observation that a random walk showed the best relative fit to one of the traits—might suggest that a static optimum is a too simple assumption in explaining these data. One possible development of the OU model could therefore be to assume a moving optimal state instead of assuming it fixed. Hunt et al. (2008) suggested to model the optimum in the OU model as a random walk and a similar model has been implemented in a phylogenetic comparative method (Hansen et al. 2008). Also, given that we find evidence for a decay in the rate of evolution with time for two of the traits, it might also be interesting to explore an OU model where the random walk part of either the optimum or the trait decrease as a function of time. However, even though the fitted models failed some of the adequacy tests, the relative fit of the candidate models still represents positive evidence for trait dynamics compatible with a lineage experiencing ecological opportunity. The fossil record represents an important source of data to better understand detailed phenotypic evolution beyond microevolutionary timescales.

**AUTHOR CONTRIBUTIONS**

Kjetil Lysne was involved in all aspects of analysis and writing.

**ACKNOWLEDGMENTS**

I thank Edward C. Theriot and the others who generated the fossil time series data and Gene Hunt, Melanie J. Hopkins, and Scott Lidgard for making the data easily accessible via their publications. Suggestions and input from Jostein Starrfelt, Jo Skeie Hermansen, and two anonymous reviewers improved the manuscript. The work was supported by grant number 249961 from the Norwegian Research Council.

**DATA ARCHIVING**

Data and r code to reproduce all analyses have been submitted to Dryad, https://doi.org/10.5061/dryad.80725jn/1.

**LITERATURE CITED**

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19:716–723.

Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems, prevalence, and an alternative. J. Wildl. Manag. 64:912–923.

Arnold, S. J. 2014. Phyletochronology: the ongoing synthesis. Am. Nat. 183:729–746.

Arnold, S. J., M. E. Pfenninger, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112–113:9–32.

Bartoszek, K., J. Pienaar, P. Mostad, S. Andersson, and T. F. Hansen. 2012. A phylogenetic comparative method for studying multivariate adaptation. J. Theor. Biol. 314:204–215.

Beaulieu, J. M., D. C. Jhwueng, C. Boettiger, and B. C. O’Meara. 2012. Modelling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. Evolution 66:2369–2383.

Benkman, C. W. 2003. Divergent selection drives the adaptive radiation of crossbills. Evolution 57:1176–1181.

Berry, R. J. 1964. The evolution of an island population of the house mouse. Evolution 18:468–483.

Blomberg, S. P., T. Garland Jr, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745.

Bookstein, F. L. 1987. Random walk and the existence of evolutionary rates. Paleobiology 13:446–464.

Brombacher, A., P. A. Wilson, I. Bailey, and T. H. G. Ezard. 2017. The breakdown of static and evolutionary allometries during climatic upheaval. Am. Nat. 190:350–362.

Butler, M. A., and A. A. King. 2004. Phylogenetic comparative analysis: a modelling approach for adaptive evolution. Am. Nat. 164:683–695.

Caetano, D. S., and L. J. Harmo. 2017. ratematrix: an R package for studying evolutionary integration among several traits on phylogenetic trees. Methods Ecol. Evol. 8:1920–1927.

Chiba, S. 2015. A 40,000-year record of discontinuous evolution of island snails. Paleobiology 22:177–188.

Clyde, W. C., and P. D. Gingerich. 1994. Rates of evolution in the dentition of early Eocene Cantius: comparison of size and shape. Paleobiology 20:506–522.

Cooney, C. R., J. A. Bright, E. J. R. Capp, A. M. Chira, E. C. Hughes, C. J. A. Moody, L. O. Nouri, Z. K. Varley, and G. H. Thomas. 2017. Mega-evolutionary dynamics of the adaptive radiation of birds. Nature 542:344–347.

Cooper, N., and Purvis, A. 2010. Body size evolution in mammals: complexity in tempo and mode. Am. Nat. 175:727–738.

Darwin, C. R. 1859. On the origin of species. John Murray, London.

Désamoré, A., Laenen, B., Miller, K. B., and Bergsten, J. 2018. Early burst in body size evolution is uncoupled from species diversification in diving beetles (Dytiscidae). Mol. Ecol. 27:979–993.

Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. Am. Nat. 169:227–244.

Fear, K. K., and T. Price. 1998. The adaptive surface in ecology. Oikos 82:440–448.

Finkel, Z. V., M. E. Katz, J. D. Wright, O. M. E. Schofield, and P. G. Falkowski. 2005. Climatically driven macroevolutionary patterns in the size of marine diatoms over the Cenozoic. Proc. Natl. Acad. Sci. USA 102:8927–8932.

Firmat, C., I. Lozano-Fernández, J. Agustí, G. H. Bolstad, G. Cuenca-Bescós, T. F. Hansen, and C. Pélabon. 2014. Walk the line: 600000 years of
