Inferring the effect of competition on trait evolution

Liang Xu[1], Sander van Doorn[1], Hanno Hildenbrandt[1]& Rampal S. Etienne[1]

[1]Groningen Institute for Evolutionary Life Sciences, University of Groningen, PO Box 11103, Groningen 9700 CC, The Netherlands

Correspondence: liang.xu@rug.nl

Abstract

Models of trait evolution usually assume that abiotic factors pull species toward an optimal trait value, whereas competitive interactions drive the trait values apart. However, these models do not consider population dynamics and dynamics of the trait variance and they oversimplify competition. Here we develop a coherent trait evolution model, with abundance-dependent competitive interactions, against a macroevolutionary background encoded in a phylogenetic tree. We use Approximate Bayesian Computation to fit the model to baleen whale body sizes and compare it to a model without population dynamics and a model where competition depends on the total metabolic rate of the competitors. All models suggest that baleen whales undergo weak environmental attraction and strong competition. However, they differ in their predictions of the abundance distribution. Data on abundance distriubutions therefore allow us to distinguish the models from one another, and infer the nature of competitive interactions.

INTRODUCTION

While it is generally acknowledged that "nothing in evolution or ecology makes sense except in the light of the other." (Pelletier et al. 2009), how exactly this mutual interaction can be understood is an area of active research (Schoener 2011). We are starting to understand how ecology, i.e. species interactions and population dynamics, depends on evolutionary history (Schoener 2011) and how evolutionary history, such as morphological trait evolution and phylogenetic information, explains ecological processes and patterns (Ives and Godfray 2006; Rezende et al. 2007; Rafferty and Ives 2013; Pigot and Etienne 2015; Clarke et al. 2017). However, the relative roles of biotic and abiotic factors in driving evolution remain elusive. This is in part because establishing a more explicit and mechanistic connection between the processes of trait evolution and community dynamics remains a difficult task (Narwani et al. 2015). More specifically, we do not fully understand how population dynamics influence trait evolution and vice versa. This is particularly important if evolution is fast (Reznick and Ghalambor 2001; Reznick et al. 2004; Hairston et al. 2005).

Species tend to evolve towards a phenotype that best exploits the most abundant resource (Darwin 1859), but also tend to diverge, in order to avoid competition with other species feeding on the same resources. The balance between these opposing tendencies is governed by two factors, niche width and population size, which ultimately determine how many species coexist in a given environment and how they partition the available ecological niches among them. Larger populations exert a larger impact on other species competing with them in the same part of niche space. Conversely, how much competition is experienced by a species ultimately determines its population size and may lead to a shift in niche space.

Classic methods attempting to infer the relative roles of the forces shaping trait evolution from trait distribution data use models that describe trait dynamics with Brownian motion or Ornstein–Uhlenbeck processes (Raup et al. 1973; Garland et al. 2000), with the latter

describing evolution towards an optimum. These models assume independent evolution for each species and thus do not account for species interactions (Pennell and Harmon 2013). Adaptive dynamics models do explicitly take species interactions into account but are not commonly used for inference from data. Recently, several phylogenetic comparative tools have been developed to study how the abiotic environment imposing a trait optimum and competition among species jointly drive trait dynamics. Specifically, Nuismer and Harmon (Nuismer and Harmon 2015) derived an analytical likelihood formula based on Lande's work (Lande 1976) to investigate how phylogenetic relationships influence trait evolution and rates of interaction among species. This method was extended to incorporate biogeographical factors (Drury et al. 2016), and generalized to other ecological interactions (Manceau et al. 2017). However, none of these studies took into account population dynamics thereby ignoring eco-evolutionary feedbacks of population size on trait evolution. Moreover, to obtain a tractable expression for the likelihood, they relied on a linear approximation of the fitness gradient function that has the following consequence: the more different species are, the more they repel each other. This is unrealistic because we expect that for species diverging more than their niche width, the rate of trait divergence will decrease.

Here, we develop a model that explicitly accounts for population dynamics by making fitness, the per capita growth rate, explicitly dependent on population dynamics. We explicitly allow for an effect of population size on trait evolution, and refer to this model as the abundance-weighted competition (AWC) model. Furthermore, our model assumes an intermediate trait distance (related to the competition intensity) where repulsion is most intense, which reflects a nonlinear relationship between competition and trait similarity. Trait evolution also depends on the variance in traits, and hence we also model the dynamics of intraspecies trait variance. We assume that the traits in our model do not drive the evolution of reproductive isolation (speciation), but we let species evolve along a given phylogenetic tree. The model is not amenable to analytical treatment, so we used Approximate Bayesian Computation to investigate parameter inference. We find that our method is capable of recovering the

generating parameters from simulated data sets under certain conditions; if these conditions are not satisfied, very different parameter sets can yield very similar patterns, so we expect that no method will be able to infer the parameters under these conditions. We illustrate our approach using data on the evolution of total body length in baleen whales (Mysticeti) and compare our model with two variants: one that ignores population dynamics, i.e., a model with unweighted competition (the UWC model) and one that assumes that competition is weighted by total metabolic rate (abundance multiplied by per capita metabolic rate) rather than only abundance, i.e., a model with metabolic-rate-weighted competition (the MWC model). We find that competition has played an important role in shaping trait distribution patterns in baleen whales. We discuss how our model sets the stage for a new generation of models that allow determining the relative roles of abiotic and biotic drivers of trait evolution.

MODEL DESCRIPTION AND ANALYSIS

Trait evolution and population dynamics on trees

We derive a model for the population size, mean and variance of the trait of each species in Box 1. We consider trait and population dynamics along given phylogenetic trees. The phylogenetic trees can be either reconstructed trees, containing only extant species, or full trees, containing extinct species as well. We assume that data may be available on the final trait distribution and species abundances at the tips of the tree. We used simulations to explore these distributions. We initialized the simulations with two ancestral species of identical trait means equal to 0 and variances equal to 0.5. The initial population sizes were drawn from a normal distribution with a mean of 500 individuals and a variance of 100.

Without loss of generality, we assumed zero as the trait optimum set by the environment. To be in line with the previous models of phenotype evolution along a given phylogeny, we assumed that at the branching points of the tree the two daughter species inherit the trait value from their parent. The abundance of the mother lineage is divided over the two daughter species following a binomial distribution that is truncated at the bottom and the top, so that the daughter species always have positive abundance. When a species goes extinct according to the phylogeny, we just remove the species from the simulation but keep its trait means and variances and abundance stored for the time they were extant.

Parameter inference using ABC-SMC

The complexity of our model precludes analytical approaches to fit the model to data. Hence, we developed an inference framework using Approximate Bayesian Computation based on Sequential Monte Carlo (ABC-SMC). In ABC-SMC, first introduced by Toni et al. 2009, one starts with a large number of parameter sets sampled from the prior (these are called particles in the terminology of the field), which are then used to simulate many data sets. We then evaluate the similarity of the simulated data to the empirical data (measured by one or more summary statistics). This similarity is the goodness-of-fit (GOF). The GOFs for all these data sets are used as weights to sample parameter sets in the next generation, with some random noise added to it. After a few iterations (generations in the terminology of the field), the parameter sets will form the posterior distribution. The details of the algorithm, including the computation of the GOF, can be found in the supplementary material.

The choice of an efficient summary statistic is crucial to evaluate the similarity between simulation and data. In the simulation study we use the Euclidean distance between simulated and observed trait values. Because of the stochasticity of the trait inheritance after speciation, traits of a focal species can differ substantially across replicate simulations. Nevertheless,

the gap between adjacent traits regardless of species identities reflects the true strength of environmental stabilizing selection and competition. Therefore, we do not label the species in our simulation and sort both the empirical traits and the simulated traits in an increasing order before computing the Euclidean distance of these two vectors. We also compute the Euclidean distance of the variance vectors corresponding to the reordered trait means.

Choosing the Euclidean distance of sorted traits may not be the best way to fit our model to empirical data, because we ignore information on the empirical order of trait values across the phylogenetic tree. Therefore, in the empirical study, we considered phylogenetic independent contrasts (PICs) (Felsenstein 1985) as an alternative set of summary statistics. The PICs are designed to transform the original n traits of species to $n-1$ independently and identically distributed contrasts between pairs of related species or estimated ancestral nodes (Garland 2005). Because the PICs have one dimension less than the trait data, we combined the PICs with the unsorted mean trait distance as a third set of summary statistics. We compared results between the three sets of summary statistics.

Simulation setup

To assess the behavior of our model, we first simulated data for known parameter sets and explored whether the parameter values can be correctly inferred. We considered six different values $(0, 0.001, 0.01, 0.1, 0.5, 1)$ for both the stabilizing selection coefficient γ and the competition coefficient α leading to a total of 36 parameter combinations for a given phylogenetic tree. We set $R_0 = e$ (i.e. the mathematical constant 2.7183), $\beta_0 = 10^9$ and the mutation rate $\nu = 10^{-11}$ for all simulations. We focused on the inference of three parameters, namely γ, α, ν.

To study how the phylogenetic information influences the evolution process, we generated

several phylogenetic trees, including extinct branches, under the diversity-dependent diversification model (Etienne et al. 2011) for various parameter settings of this macroevolutionary model (see Table. S1). In addition, to mimic the fact that in practice complete trees with extinct species are often not available, we reconstructed phylogenies of only extant species by pruning the extinct species. This means that we generated the trait data under the full tree, but estimated the parameters of our trait evolution model using only the reconstructed tree. Comparing this inference to inference using the complete tree informs us to what extent the loss of information of extinct species affects parameter estimation.

The ratio of the time scale of trait evolution and population dynamics to the time scale set by the phylogeny (i.e. the number of time-steps of trait and population size dynamics in each unit of time of the phylogeny) is a crucial factor. We denote this ratio in our model by the scaling parameter s. For instance, given a phylogenetic tree with a crown age of 15 million years, trait and population dynamics involves $15 \times s$ time steps. The value of s may influence our parameter estimates. So to assess how not knowing the true number of time steps (i.e. the time scale of trait evolution to the time scale of speciation) affects parameter inference, we generated data under $s = 10000$ and then ran our inference algorithm under $s = 10000$ and $s = 20000$ and compared their performance in parameter estimation (see Table. S1).

In summary, we generated a total of 14 phylogenetic trees and pruned these trees when extinction rates were nonzero, resulting in 22 trees in total (see Table. S1). We designed 30 scenarios to investigate the influence of tree size, speciation rate, extinction rate, removal of extinct species and number of time steps. We simulated our model for 36 parameter combinations for each scenario. We applied our inference algorithm on the simulated data and examined if the generating parameters could be recovered correctly. In the inference process, we set 30 iterations and 20000 particles for each iteration. For the analysis ofa single scenario, we exploited a cluster of 36 high performance computers with 32 threads running on each computer. Each parameter combination for each scenario analysis took between 2 and 80 hours, depending on

the number of evolutionary events and tree size of the specific scenario. All the code is available on Github (https://github.com/xl0418/The_trait_population_coevolution_model_code).

Applying the model to baleen whale body size evolution

Baleen whales represent the largest extant animal species and are distributed globally. They are filter-feeders on small fish and crustaceans. Body mass is an ideal trait that responds both to the abiotic factors (Smith et al. 2010) and biotic competitors but measurements of body mass are rarely available. Data on total length are available, however, and because it has been shown that whale total length scales with body mass$^{1/3}$ (Lockyer 1976), we used the total length as a proxy for body mass (Slater et al. 2017). We log-transformed (base 10) the body length, because the log scale is a more natural scale on which evolution takes place (Gingerich 2019) (Results based on the untransformed total length can be found in the supplementary material, see Fig. S56-S58 and Fig. S62-S64). We fitted our model to mean trait data given a reconstructed phylogeny with 15 extant species (Slater et al. 2017). We did not use abundance or trait variance data, as these were not available.

We designed 8 scenarios to fully assess the effects of environmental stabilizing selection and competition: four values of the time scaling parameter s (20000, 40000, 60000 and 80000) corresponding to four reasonable generation times (50, 25, 16.7, 12.5 years/generation, respectively) and two heritability values ($h^2 = 0.5, 1$). In contrast to the simulation study, we also estimated the variance due to mutation and segregation, V_m, and the trait optimum, θ. The remaining parameter settings were identical to the simulation study. In the ABC-SMC algorithm, we set 40000 particles for each iteration and in total 40 iterations for each scenario.

We further developed two additional models for comparison. The first model (the UWC model) assumes that competition does not depend on species abundance and hence does not need population dynamics. The UWC model is similar to Drury *et al.*'s nonlinear extension

(Drury et al. 2017) of Nuismer & Harmon's model (Nuismer and Harmon 2015). However, it differs in the competition kernel, i.e. it models pairwise competition as $(\mu_i - \mu_j) \cdot e^{-\alpha(\mu_i - \mu_j)^2}$ which is a natural consequence of our trait evolution model, instead of as Drury *et al*'s *ad hoc* choice of $\text{sign}(\mu_i - \mu_j) \cdot e^{-\alpha(\mu_i - \mu_j)^2}$ (Eq. 1 in Drury et al. 2017). The second additional model (the MWC model) assumes that competition depends on total metabolic rate, in which abundance is multiplied by the per capita metabolic rate, which depends on body length (see Eq. S32-S33 & S35-S37 in the supplementary material). That is, the pairwise competition is $e^{-\alpha(\mu_{i,t} - \mu_{j,t})^2} B_{j,t}$ instead of $e^{-\alpha(\mu_{i,t} - \mu_{j,t})^2} N_{j,t}$, where $B_{j,t}$ is the total metabolic rate of species j at the $t-$th generation. Because the logarithms of body length and body mass of whales are strongly correlated with a slope of $\frac{1}{3}$ (Lockyer 1976) and per capita metabolic rate has a scaling with body mass of $\frac{3}{4}$ (Brody and Procter 1932; Brody 1945; Kleiber 1947; Etienne et al. 2006), the total metabolic rate depends on body size and abundance as follows: $B_i = N_i \cdot B_0 \cdot \mu_i^{9/4}$. Here B_0 is a basal metabolic rate per kg (BMR/kg) that is assumed to be constant across whale species, and therefore drops out of our equations because only the relative metabolic rate matters. Thus, large-bodied species have more competitive power than small-bodied species. This asymmetry makes the trait distribution at present highly dependent on the optimum trait. We therefore also estimated the optimum trait θ. Because more free parameters require more simulations in each iteration in the algorithm to obtain sufficient sampling, we here used 40000 particles per iteration instead of 20000 in the simulation study. For these two additional models we again estimated five parameters $(\gamma, \alpha, \nu, V_m, \theta)$. but we considered only one scenario of heritability and time scaling ($h^2 = 1; s = 20000$), because the analyses are computationally extremely demanding, and because we found that the scenarios were similarly supported for the AWC model (see Results).

We used three alternative sets of summary statistics, i.e. the sorted mean trait distance (SMTD), the phylogenetic independent contrasts only (PICs) and the unsorted mean trait distance with the phylogenetic independent contrasts (UMTD+PICs). To compare the goodness-of-fit of the eight scenarios (for the AWC model) among each other, we took the

simulations with the 5% highest GOF-values across all scenarios and computed the percentage of simulations represented by each scenario in these 5% best fitting simulations as a measure of the support of that scenario. We did this for each of the three sets of summary statistics.For comparing the three models we used the exact same procedure; support of a model is thus measured by its representation among the 5% best GOF-values across all three models. Lastly, for each model we used the parameter estimates to generate 1000 data sets to compare the predicted phylogenetic independent contrasts with the empirical observations.

RESULTS

Inference in the simulation study

We find the most accurate inference when $\sqrt{\frac{\alpha}{\gamma}}$, is roughly equal to the number of tips in the phylogeny. When $\alpha > \gamma$, the traits are informative for parameter inference because the traits are sufficiently separated. Thus, the parameter space where $\alpha > \gamma$ will be called the informative parameter domain and otherwise the uninformative parameter domain. Tree size has a large impact on the estimation of α and mutation rate ν, but not on the prediction of γ (Fig. 1 and Figs. S2-S4). A high speciation rate tends to bias the estimation of α and ν and increases the variance of all the estimates for the scenarios with nonzero extinction (Figs. S7-S9) while for the pure birth scenarios it improves the estimation (Figs. S4-S6 for the scenarios with extinction). The median parameter inference remains good even when extinction rate increases but the variance increases (Fig. S10-S15).

There is no substantial difference in parameter estimates when using the full tree or the pruned tree for most of the parameter combinations in the informative parameter domain (Figs. S16-S23). In addition, choosing a larger value of the scaling parameter ($s = 20000$)

in estimation than the one used to generate the data ($s = 10000$) generally does not lead to significant differences (Fig. S24-S27), nor does a larger value of the scaling parameter in the simulation and in inference improve the performance of parameter inference (Figs. S28-S31). This indicates that the scaling parameter $s = 10000$ is sufficiently large in the study so that the trait evolution is close to equilibrium. . We refer to the supplementary material for further details of the simulation study.

Environmental stabilizing selection and competition in baleen whales

To examine the phylogenetic signal in the whale body length data, i.e. to study to what extent the correlation in traits reflects their shared evolution history, we computed Blomberg's K (Blomberg et al. 2003) and Pagel's λ (Pagel 1999). These two statistics represent two different quantitative measures of the phylogenetic signal, measuring the correlations between species relative to what is expected under Brownian motion and a scaled ratio of the variance among species over the contrasts variance, respectively. Blomberg's K for the baleen whales is 0.751 with a p-value of 0.002, implying a moderate degree of phylogenetic signal close to an O-U process with a weak adaptation effect (Blomberg et al. 2003; Diniz-Filho et al. 2012). Pagel's λ ($= 0.99$) suggests resemblance among close relative species (close to Brownian motion process which has $\lambda = 1$) although not significant ($p-$value $= 0.02$). These findings suggests that there is some phylogenetic signal in the body length data, making it reasonable to apply our model.

In all scenarios, the estimates of γ are much smaller than the estimates of α, implying weak environmental stabilizing selection and fairly strong competition in baleen whale body size evolution. Heritability tends to influence on the inference of environmental stabilizing selection coefficient γ, competition coefficient α but not so much on the mutation rate ν, the maximum variance by mutation V_m and the optimum trait θ (see Fig. 3). With a larger

heritability ($h^2 = 1$), the estimates of γ and α are smaller and less variable than for $h^2 = 0.5$ across all three summary statistics. Alternative sets of summary statistics lead to similar results in the estimation of the parameters ν, V_m and θ, but for γ and α, the inferences using PICs and UMTD+PICs differ from those using SMTD particularly when $h^2 = 0.5$ (see Fig. 3 for SMTD, Fig. S54 for PICs and Fig. S55 for UMTD+PICs). The estimated γ is smallest, around 6, when using PICs. The value increases when the summary statistics include absolute trait information (SMTD and UMTD+PICs), reaching 9 for the algorithm using SMTD only. In contrast, the estimations for α using SMTD are much smaller than using the other two summary statistics. The mutation rate ν is consistently inferred to be 0.001 except that when using PICs increasing s to 80000 increases the estimate to 0.002. We did not find substantial differences in the estimates under different time scaling parameters, except that with only PICs as the set of summary statistics, the predicted α decreases when increasing the time scaling parameter and/or heritability. The estimation of θ is estimated to be around 3.05 across all three summary statistics, although a larger the time scaling parameter leads to a larger variance when using only PICs.

Comparing the supports of the scenarios across h^2 and s (computed from the GOF-values) tells us that when using SMTD as summary statistic $h^2 = 0.5$ better fits the data than $h^2 = 1$ while all values of s are similarly supported. Using PICs and UMTD+PICs as summary statistics leads to equally good fits for different h^2 and s (see Fig. 5).

We note that the quantitative values of the estimates for stabilizing selection and competition coefficients are not comparable among the three models of trait evolution (AWC, UWC and MWC) because these models assume different factors affecting competition. Nevertheless, the estimations all suggest a weak environmental stabilizing selection coefficient and a strong competition coefficient (see Fig. S59-S61 for the three sets of summary statistics). When using the summary statistics SMTD, the best fitted value of the optimum trait θ is close to the mean of the extant species traits (3.087) for the AWC model and the UWC model,

and around 2.9 for the MWC model. A similar pattern is also found in the estimation using UMTD+PICs but a large variance emerges for the MWC model. Using PICs (but not SMTD and UMTD+PICs) leads to large variance to estimations of θ in all three models. The abundance distribution substantially differs for the three models (Fig. S67) regardless of the choice of summary statistics. The AWC model generates a symmetric unimodal abundance distribution around the optimum trait. For the MWC model, the abundance distribution shows a decrease of abundance with increasing body length. The prediction of the intraspecific variance also differs substantially among the three models (see Fig. S68). The AWC model produces higher intraspecific variance than the UWC and MWC models. Comparing the supports of the three models we find that the UWC model is favored when using SMTD as summary statistic while all three models show a similar fit when using the other two summary statistics (5).

Using the estimated parameters, we generated 1000 replicate trait data sets for the three models and compared the predicted phylogenetic independent contrasts to the empirical data (see Fig. 4 for the results using UMTD+PICs and Fig. S65-S66 for the results using PICs and SMTD). We observe a good fit for most of the contrasts but substantial differences for some pairs of species and clades. For example, the contrasts between the pairs of *B.musculus* and its daughter clade, *M.novaeangliae* and *B.physalus* are underestimated.

Discussion

We have developed a novel model of trait evolution and population dynamics to explore how environmental stabilizing selection and competition shape trait patterns. We employed Approximate Bayesian Computation with a Sequential Monte Carlo algorithm to infer the

parameters of interest and measure the performance of our method for simulation data. Our analysis reveals that the trait data at the present is generally informative for parameter inference, particularly when the number of species allowed by the strength of stabilizing selection and competition is similar to the number of species present in the phylogeny, which implies that the competition coefficient needs to be larger than the stabilizing selection coefficient. In contrast, when environmental stabilizing selection is greater than competition, our inference approach is limited due to an uninformative trait pattern that is highly compressed to the optimum trait. Our empirical study shows a weak environmental stabilizing selection and a relatively strong competition, suggesting a rapid repulsion among whale species.

Our model advances previous studies (Nuismer and Harmon 2015; Drury et al. 2016; Manceau et al. 2017; Drury et al. 2017) in three ways. First, our model is based on an explicit definition of fitness based on population dynamics, instead of chosen *ad hoc*. This property makes our model more coherent. Second, we have relaxed the assumption of a linear competition kernel, and thereby captures a realistic mechanism in species interactions: when two populations have very similar trait values, the substantial overlap in their intraspecific trait distributions neutralizes directional selection although competition is very intense there. Consequently, the rate of trait evolution to evolve apart is also low. However, once stochastic mutation produces imbalance, competition leads to character displacement. Thus, the force of directional selection dramatically increases at the onset of a division in traits. When species interactions decrease due to increasing dissimilarity, the force of directional selection drops. This process lasts until the two populations evolve apart in traits to the extent that competition is balanced by stabilizing attraction. The third advantage of our model is that we have made population size a factor in the force of competition (which seems more realistic), and modeled the population dynamics.

Our simulation study reveals that phylogenetic information, i.e. tree size, branching times

and extinct branches, is informative for the inference of stabilizing selection but carries little information on species interactions in phylogenies with nonzero extinction rate (Nuismer and Harmon 2015). Large trees show more variance in estimations. On the one hand, this is because large trees have more speciation events that may result in species branching before they can evolve to equilibrium. On the other hand, the potential mismatch between the species carrying capacity K used to generate phylogenetic trees and the number of coexisting species that is allowed by the strength of environmental stabilizing selection and competition could play a role in the performance of the inference method. Large environmental stabilizing selection results in a narrow width of natural resources, which may sustain only a few species. However, our simulation conditioning on a prescribed individual carrying capacity and a given phylogenetic tree forces more species to survive than stabilizing selection would naturally allow. This conflict may produce more variance in parameter estimation. High speciation rate generally improves parameter inference in phylogenies resulting from pure birth process, because the community can reach the clade-level carrying capacity faster, allowing more time to form a specific pattern in traits. Conversely, extinction prevents the community from reaching the carrying capacity and generates more macroevolutionary events, allowing less time for trait evolution. Hence, bias in parameter inference is expected for a tree with high extinction, even when speciation rate is high.

Reconstructing trees by pruning extinct species has little influence on parameter inference if environmental stabilizing selection is present and the number of time steps is large. One might expect that extinction leaves gaps in the trait distribution suggesting a wider effective competition zone corresponding to a smaller competition coefficient α, which is observed in our results. However, any possible bias that extinction may cause will disappear if given enough time steps because extant species eventually take over empty niches left by species that went extinct. Thus, if the rate of evolution is fast enough, the trait pattern of extant species only can still inform about species interactions. That is to say, under these conditions phylogenetic information is not necessarily needed for the inference of species interactions.

This conclusion is consistent with Nuismer & Harmon's finding for the phenotype differences model (Nuismer and Harmon 2015). This result increases the robustness of our inference approach when applied to empirical data, where we often have only molecular phylogenies and morphological trait data for extant species.

In our example of baleen whales we find evidence for strong competition. However, we also find some differences between observed and predicted contrasts (regardless of the set of summary statistics used in ABC), showing a limitation in the ability of our model to predict body sizes. That is, some close relatives show larger contrasts than predicted by our model. This suggests that (1) other traits than body size also determine the strength of competitive interactions, and/or (2) some of the baleen whales explore different niche dimensions (e.g. they utilize different types of resources), i.e. they evolve to adapt to new niches to avoid competition. For example, the largest underestimation of the contrasts is found in the pair of *B.musculus* and its daughter clade and the pair of *B.physalus* and its daughter species which seems to imply that the two largest species find niches that favor gigantism (Slater et al. 2017). This additional mechanism in trait evolution also reduces the phylogenetic signal, as suggested by Blomberg's K and Pagel's λ test. This does not mean that competition has not played a role in baleen whale trait evolution: competition may have led to diversification in different niche dimensions. Both explanations involve multiple traits, and hence extending our model to incorporate evolution of multiple traits presents an interesting avenue for further research.

With the three models we studied in this paper we have introduced additional mechanisms in the competition kernel. The formulation of the pairwise competition can be written in the general form $e^{-\alpha(\mu_{i,t}-\mu_{j,t})^2}F(N_{j,t},\mu_{j,t})$, where $F(N_{j,t},\mu_{j,t})$ describes the mechanism of interest. For example, the UWC model assumes equal competitive power ($F(N_{j,t},\mu_{j,t})=1$) to all species while the AWC model assumes an abundance related competition ($F(N_{j,t},\mu_{j,t})=N_{j,t}$) and the MWC model sets a metabolism related competition ($F(N_{j,t},\mu_{j,t})=B_{j,t}$). The AWC

model agrees with the intuition that the species that can utilize the most available natural resource is the most abundant even if competition is most intense there. The MWC model captures the fact that relationship between body size and abundance is generally negative (Damuth 1981; Peters and Wassenberg 1983; Peters and Raelson 1984; Damuth 1987). Our general formulation allows for other ways of weighing the effect of competitive interactions on trait evolution.

Our model can be extended in various directions. For example, our model currently only incorporates demographic stochasticity; the noise term decreases rapidly with population size. One can easily add environmental noise, which is independent of population size. Other possible extensions would be to replace the single optimum of a single trait by multiple optima, and to relax the assumption that all species interact with each other to only local interactions in a spatially explicit model (see e.g. Manceau et al. 2017; Drury et al. 2017). One of the most challenging issues in spatial modeling is how to keep track of widespread species that can utilize a wealth type of resources and reside in multiple regions (Xu and Etienne 2018). Another extension of our model would be to model the effect of trait evolution on diversification, along the lines of Aristide & Morlon (Aristide and Morlon 2019) who integrated trait evolution in a macroevolutionary model to study the effect of competition on biodiversity and phenotypic diversity (but without population dynamics). Finally, our model and the ABC-approach also allow for joint inference of macroevolutionary (speciation and extinction) dynamics and microevolutionary (trait) dynamics. We hope that the assumption of nonlinear interspecific interactions and association with population dynamics will spur progress of model-based approaches to questions involving trait evolution.

References

Abrams, P. a. (2001). A world without competition. *Nature*.

Aristide, L. and H. Morlon (2019). Understanding the effect of competition during evolutionary radiations : an integrated model of phenotypic and species diversification. *Ecology Letters*.

Barnett, S. A. and G. G. Simpson (1955). *The Major Features of Evolution*, Volume 46. New York: Columbia University Press.

Blomberg, S. P., T. G. Jr, A. R. Ives, T. Garland, and A. R. Ives (2003). Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits are More Labile. *Evolution 57*(4), 717–745.

Brody, S. (1945). *Bioenergetics and growth, with special reference to the efficiency complex in domestic animals*. Reinhold, New York.

Brody, S. and R. Procter (1932). *Relation between basal metabolism and mature body weight in different species of mamals*. Univ.Mo.Agr.Exp.Sta.Res.Bull.

Clarke, M., G. H. Thomas, and R. P. Freckleton (2017). Trait Evolution in Adaptive Radiations: Modeling and Measuring Interspecific Competition on Phylogenies. *The American Naturalist 189*(2), 121–137.

Damuth, J. (1981). Body size in mammals. *Nature 290*(April), 699.

Damuth, J. (1987). Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy-use. *Biological Journal of the Linnean Society 31*, 193–246.

Darwin, C. R. (1859). *On the Origin of Species by means of Natural Selection; or the Preservation of Favoured Races in the Struggle for Life*, Volume 5. John Murray, London, UK.

Diniz-Filho, J. A. F., T. Santos, T. F. Rangel, and L. M. Bini (2012). A comparison of metrics for estimating phylogenetic signal under alternative evolutionary models. *Genetics and Molecular Biology 35*(3), 673–679.

Drury, J., J. Clavel, M. Manceau, and H. Morlon (2016). Estimating the effect of competition on trait evolution using maximum likelihood inference. *Systematic Biology 65*(4), 700–710.

Drury, J. P., G. F. Grether, T. Garland, H. Morlon, T. Garland, Jr., H. Morlon, T. Garland, H. Morlon, T. Garland, Jr., H. Morlon, T. Garland, and H. Morlon (2017). An Assessment of Phylogenetic Tools for Analyzing the Interplay Between Interspecific Interactions and Phenotypic Evolution. *Systematic Biology 67*(May), 413–427.

Etienne, R. S., M. E. F. Apol, and H. Olff (2006). Demystifying the West, Brown & Enquist model of the allometry of metabolism. *Functional Ecology 20*(2), 394–399.

Etienne, R. S., B. Haegeman, T. Stadler, T. Aze, P. N. Pearson, A. Purvis, and A. B. Phillimore (2011). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proceedings of The Royal Society B: Biological Sciences 279*(1732), 1300–1309.

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

Garland, Ives, T. Garland, Jr.,, A. R. Ives, T. Garland, A. R. Ives, T. Garland, Jr.,, A. R. Ives, T. Garland, and A. R. Ives (2000). Using the Past to Predict the Present: Confidence Intervals for Regression Equations in Phylogenetic Comparative Methods. *The American Naturalist 155*(3), 346–364.

Garland, T. (2005). Phylogenetic approaches in comparative physiology. *Journal of Experimental Biology 208*(16), 3015–3035.

Gingerich, P. D. (2019). *Rates of Evolution: A Quantitative Synthesis*. Cambridge University Press.

Hairston, N. G., S. P. Ellner, M. A. Geber, T. Yoshida, and J. A. Fox (2005). Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters 8*(10), 1114–1127.

Harmon, L. J., C. S. Andreazzi, F. Débarre, J. Drury, E. E. Goldberg, A. B. Martins, C. J. Melián, A. Narwani, S. L. Nuismer, M. W. Pennell, S. M. Rudman, O. Seehausen, D. Silvestro, M. Weber, and B. Matthews (2019, apr). Detecting the Macroevolutionary Signal of Species Interactions. *Journal of Evolutionary Biology*, jeb.13477.

Ives, A. R. and H. C. J. Godfray (2006). Phylogenetic Analysis of Trophic Associations. *168*(1).

Kimura, M. and J. F. Crow (1964). The number of alleles that can be maintained in a finite population. *Genetics 49*(4), 725–738.

Kleiber, M. (1947). Body size and metabolic rate. *Physiological Reviews 27*(4).

Lande, R. (1976). Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution 30*(2), 314–334.

Lenormand, T., D. Roze, and F. Rousset (2009). Stochasticity in evolution. *Trends in Ecology and Evolution 24*(3), 157–165.

Lockyer, C. (1976). Body weights of some species of large whales. *ICES Journal of Marine Science 36*(3), 259–273.

Mahler, D. L., T. Ingram, L. J. Revell, and J. B. Losos (2013). Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science 341*(6143), 292–295.

Manceau, M., A. Lambert, and H. Morlon (2017). A Unifying Comparative Phylogenetic Framework Including Traits Coevolving Across Interacting Lineages. *Systematic Biology 66*(4), 551–568.

Narwani, A., B. Matthews, J. Fox, and P. Venail (2015). Using phylogenetics in community assembly and ecosystem functioning research. *Functional Ecology 29*(5), 589–591.

Nuismer, S. L. and L. J. Harmon (2015). Predicting rates of interspecific interaction from phylogenetic trees. *Ecology Letters 18*(1), 17–27.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature 401*(6756), 877–884.

Pelletier, F., D. Garant, and A. P. Hendry (2009). Eco-evolutionary dynamics. pp. 1483–1489.

Pennell, M. W. and L. J. Harmon (2013). An integrative view of phylogenetic comparative methods: Connections to population genetics, community ecology, and paleobiology. *Annals of the New York Academy of Sciences 1289*(1), 90–105.

Peters, R. H. and J. V. Raelson (1984). Relations between individual size and mammalian population density. *124*(4), 498–517.

Peters, R. H. and K. Wassenberg (1983). The effect of body size on animal abundance. *Oecologia 60*(1), 89–96.

Pigot, A. and R. S. Etienne (2015). A new dynamic null model for phylogenetic community structure. *Ecology Letters 18*(2), 153–163.

Rafferty, N. E. and A. R. Ives (2013). Phylogenetic trait-based analyses of ecological networks. *Ecology 94*(10), 2321–2333.

Raup, D. M., S. J. Gould, T. J. M. Schopf, and D. Simberloff (1973). Stochastic models of phylogeny and the evolution of diversity. *J. Geol. 81*, 525–542.

Rezende, E. L., P. Jordano, and J. Bascompte (2007). Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. (June), 1919–1929.

Reznick, D., H. Rodd, and L. Nunney (2004). *Evolutionary Conservation Biology*. Cambridge Univ. Press.

Reznick, D. N. and C. K. Ghalambor (2001). The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution. *Genetica 112-113*(1956), 183–198.

Schoener, T. W. (2011). The newest synthesis: Understanding the interplay of evolutionary and ecological dynamics. *Science 331*(6016), 426–429.

Slater, G. J., J. A. Goldbogen, and N. D. Pyenson (2017). Independent evolution of baleen whale gigantism linked to Plio-Pleistocene ocean dynamics. *Proceedings of the Royal Society B: Biological Sciences 284*(1855).

Smith, F. A., A. G. Boyer, J. H. Brown, D. P. Costa, T. Dayan, M. Ernest, A. R. Evans, M. Fortelius, J. L. Gittleman, J. Marcus, L. E. Harding, K. Lintulaakso, S. K. Lyons, C. Mccain, J. G. Okie, J. J. Saarinen, R. M. Sibly, P. R. Stephens, J. Theodor, and M. D. Uhen (2010). Supporting Online Material for The Evolution of Maximum Body Size of Terrestrial Mammals. *Science 1216*(November), 1216–1220.

Toni, T., D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface 6*(31), 187–202.

Van Valen, L. (1969). Variation Genetics of Extinct Animals. *The American Naturalist 103*(931), 193–224.

Xu, L. and R. S. Etienne (2018). Detecting local diversity-dependence in diversification. *Evolution 72*(6), 1–12.

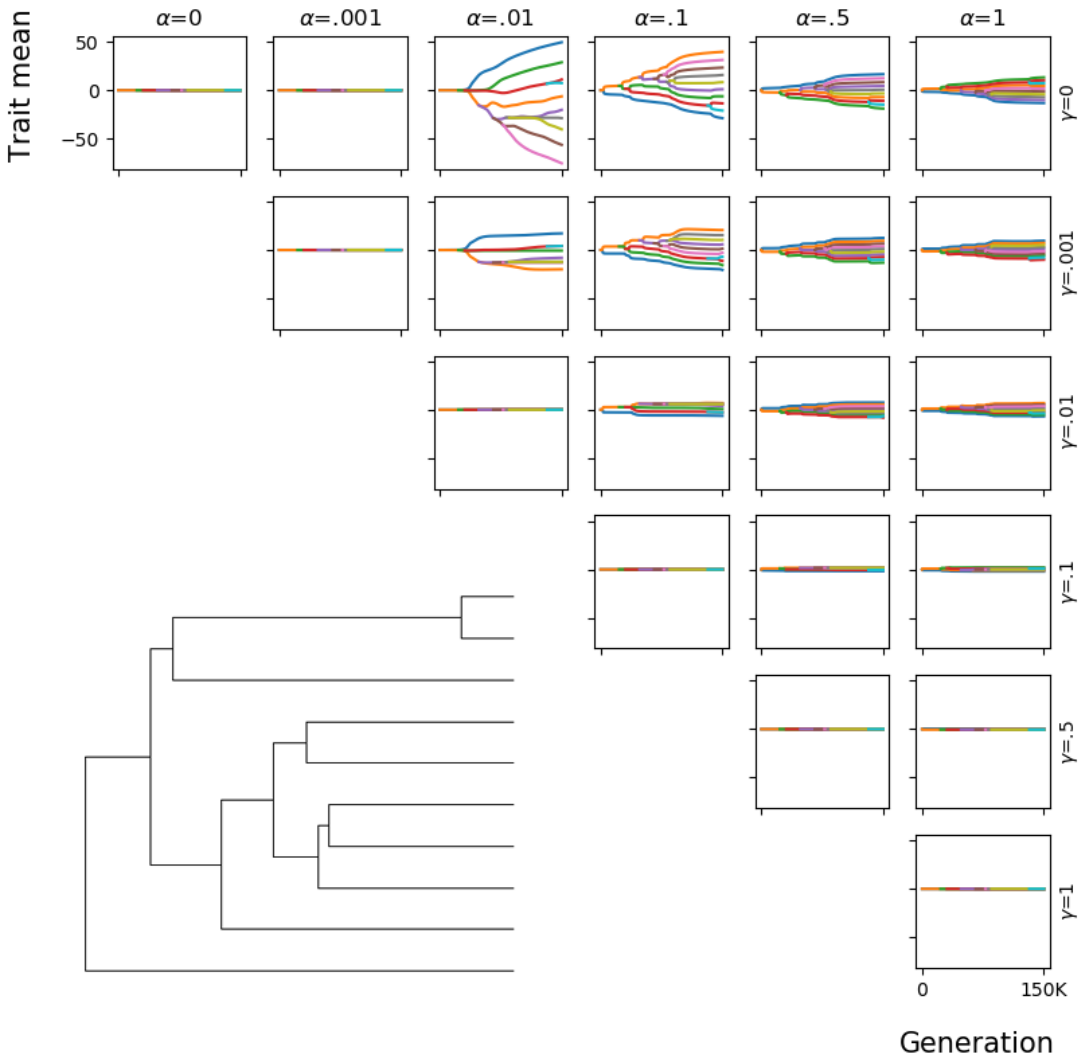

Figure 1 – The phylogenetic tree and the corresponding trait trees along Tree 1 across all parameter combinations. Different colors denote the traits of different species. The simulation is initialized with two ancestral species with identical trait mean 0 and variance 1/2 and the population size is randomly chosen from a normal distribution centering 500 with variance of 100. The parameters used to generate the phylogenetic tree are $\lambda = 0.4, \mu = 0, K = 10$ with a crown age of 15 Myr. The time scaling parameter is $s = 10000$ indicating a total of 150,000 time steps in the trait simulation. Trees are shown only for the informative parameters domain ($\gamma < \alpha$); when $\gamma > \alpha$, the trait trees are all almost identical: a highly compressed line such as the pattern on the diagonal.

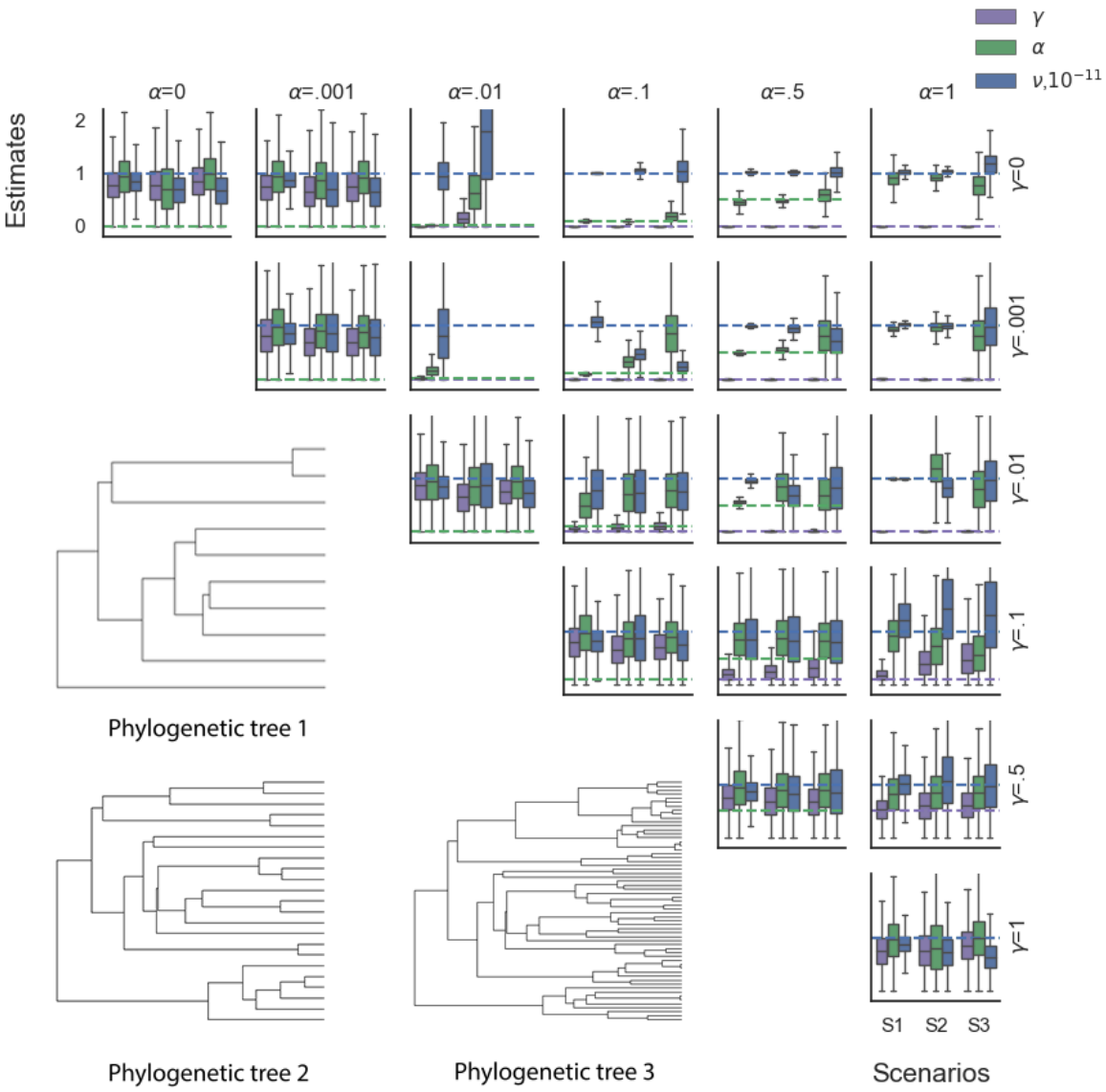

Figure 2 – Parameter inference for the 36 parameter combinations under Scenarios 1, 2 and 3 (the corresponding phylogenetic trees are shown at the bottom left). The dashed lines in three colors denote three parameter values used to generate the data. Note that the scale of mutation rate is 10^{-11}. Box plots indicate the distribution of inferred parameter values, where the whiskers extend from the minimum to the maximum value. Some plots show no estimates for part of the parameter combinations, because there is no complete simulation under these parameter combinations after 10,000 attempts. The shared parameters used to generate the phylogenetic trees are $\lambda = 0.4, \mu = 0$ with a crown age of 15 Myr. The clade-level carrying capacity is $K = 10, 30, 100$ for trees 1, 2, 3, respectively. Data are shown only for the informative parameter domain ($\gamma \leq \alpha$); when $\gamma > \alpha$, the parameter estimation show major bias and substantial variance.

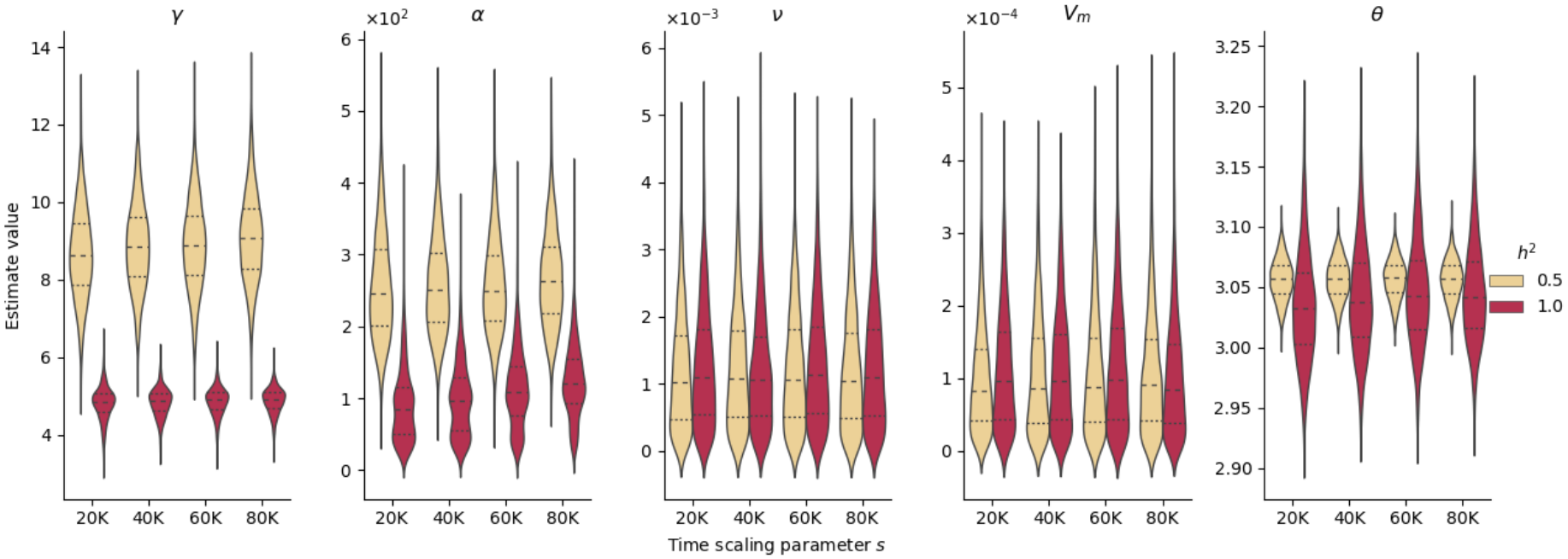

Figure 3 – Parameter inference under 8 scenarios to test the influence of the number of time steps and heritability on parameter estimation using SMTD as the summary statistic (for the results using PICs and UMTD+PICs, see Fig. S54-S55). The three dashed lines in the violin plot are 25th percentile, median and 75th percentile quantile of the samples in the last iteration of the ABC algorithm that produce the 5% best fits to the baleen whale data.

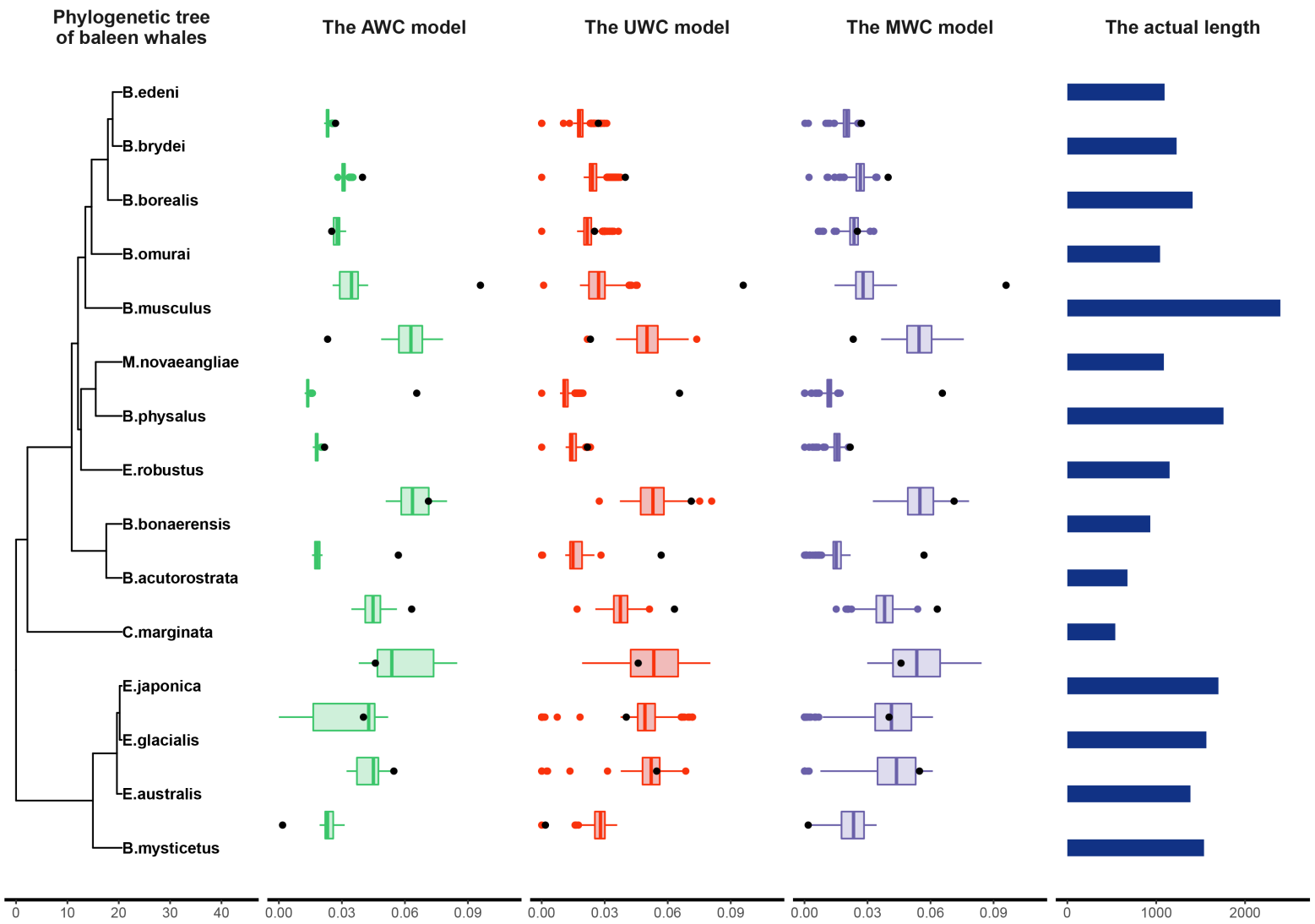

Figure 4 – Prediction of the phylogenetic independent contrasts (PICs) of baleen whale log-transformed body length, simulated using the parameters estimated using UMTD+PICs . The time scaling parameter s is $20,000$, corresponding to 50 years per generation and heritability $h^2 = 1$. The phylogeny is the reconstructed tree of the Mysticeti with the x-axis in units of million years. The box plots show the distributions of PICs across 1000 simulations under the three models against the true data (the black dots) with the x-axis denoting the absolute phylogenetic independence contrasts. The bars on the right show the untransformed body length of the species.

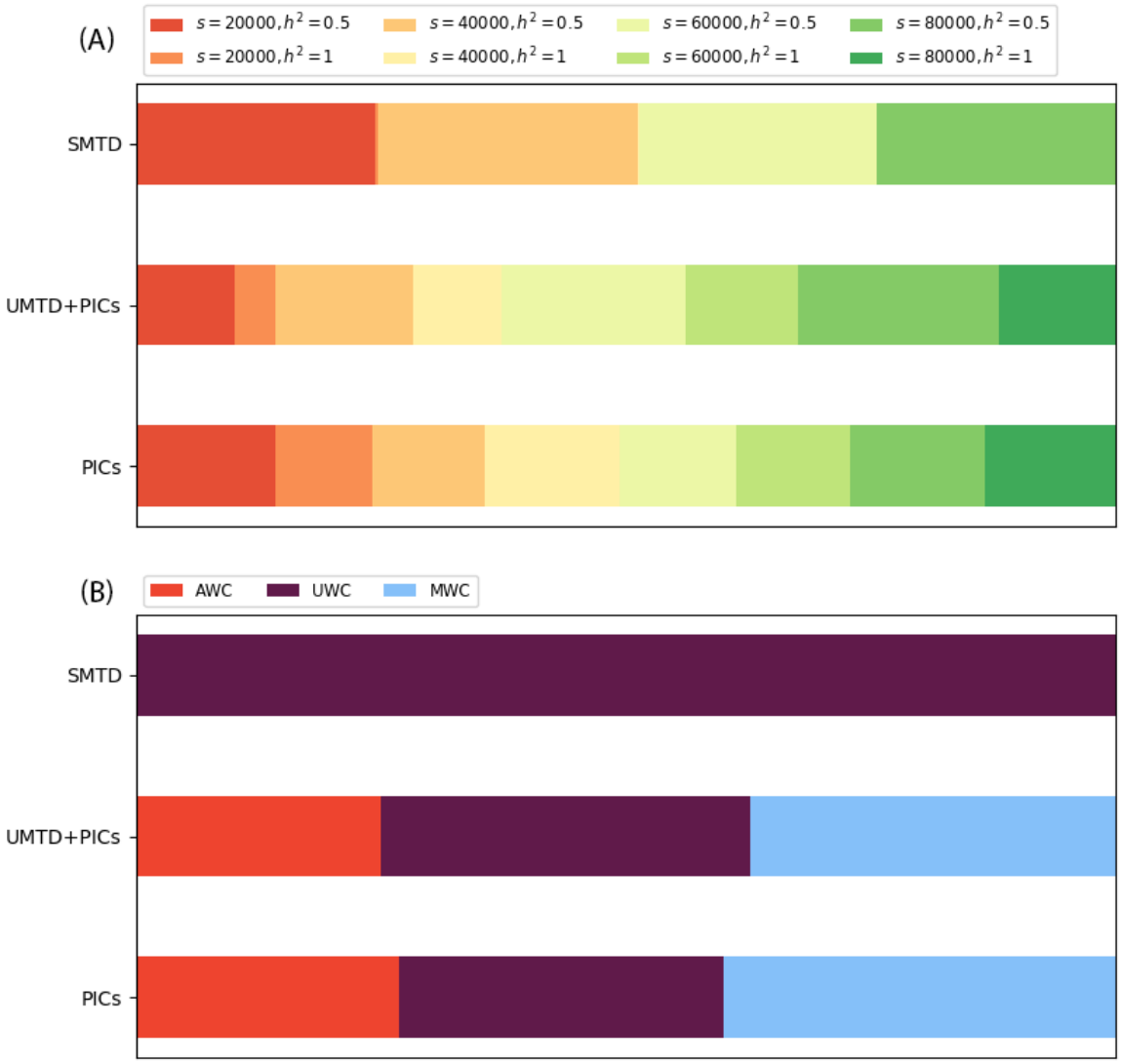

Figure 5 – Model comparison. Panel A shows support of various h^2 and s for the three summary statistics under the AWC model.The width of the bar denotes the support of the corresponding parameter combination. Panel B shows support across the three candidate models (AWC, UWC and MWC) for the three summary statistics. The height of the bar denotes the supports of the corresponding model.

Box 1. The derivation of the trait evolution model with population dynamics.

We consider n species competing for the same spectrum of resources with a fixed and unimodal distribution (Mahler et al. 2013).
We define fitness for an individual of species i with trait z_i through the per capita growth function

$$\omega(z_i) = \frac{N_{i,t+1}}{N_{i,t}} \frac{f_{i,t+1}(z_i)}{f_{i,t}(z_i)}. \tag{1}$$

where $N_{i,t}$ denotes the population size at the t-th generation and $f_{i,t}$ and $f_{i,t+1}$ are the densities of traits of the two generations respectively. After applying simplifying assumptions such as weak selection, we can derive a model on the species level that describes population dynamics, and dynamics of the mean and intraspecific variance of each species' trait in terms of the fitness function (see the supplementary material for the full derivation) (Harmon et al. 2019):

$$N_{i,t+1} = N_{i,t} \cdot \omega(\mu_{i,t}) \tag{2}$$

$$\mu_{i,t+1} = \mu_{i,t} + h^2 \cdot V_{i,t} \cdot \frac{\omega'(\mu_{i,t})}{\omega(\mu_{i,t})} \tag{3}$$

$$V_{i,t+1} = (1 - \frac{1}{2}h^2)V_{i,t} + \frac{1}{2}h^2 \cdot V_{i,t}^2 \cdot \frac{\omega''(\mu_{i,t})}{\omega(\mu_{i,t})} + \frac{2N_{i,t}\nu V_m}{1 + 4N_{i,t}\nu} \cdot h^2, \tag{4}$$

where $V_{i,t}$ denotes the variance of the trait of species i at the t-th generation. In the model, $\omega'(\mu_{i,t})$ and $\omega''(\mu_{i,t})$ are the first and second derivatives of the fitness function with respect to the trait value, evaluated at the species mean of the trait $\mu_{i,t}$. The first derivative of the fitness function is proportional to the rate at which the mean trait in the population evolves uphill on the fitness landscape. The second derivative of the fitness function determines whether selection is stabilizing or disruptive, which influences the variance of the trait and, thereby, the rate of evolution. . Eq. 2, which is is essentially Lande's formula (Eq. 7 in Lande 1976), links the population size to a user-specified fitness function. In the literature, the trait variance is usually assumed to be constant (Lande 1976; Nuismer and Harmon 2015; Drury et al. 2016, 2017), but this may be an oversimplification (Barnett and Simpson 1955; Van Valen 1969). Therefore, we allow variance dynamics (Eq. 4) in trait evolution which consists of three components that capturea reduction of variance due to sexual reproduction ($(1 - \frac{1}{2}h^2)V_{i,t}$), the effect of selection ($\frac{1}{2}h^2 \cdot V_{i,t}^2 \cdot \frac{\omega''(\mu_{i,t})}{\omega(\mu_{i,t})}$), and an inflow of new variance (Kimura and Crow 1964) due to segregation and mutation ($\frac{2N_{i,t}\nu V_m}{1+4N_{i,t}\nu} \cdot h^2$) . Here, ν is the average rate of mutation of the alleles existing in a diploid population and V_m represents the maximum variance of the trait that can be supported by mutation in an infinitely large population, in the absence of selection on the trait. In the simulation study we set V_m equal to 1and the environmental contribution to phenotypic variance to 0, such that trait heritability h^2 is equal to 1. In the empirical application we

estimate V_m as a free parameter, and we consider two distinct values of heritability h^2 . We allow for stochastic evolutionary trait change due to drift by adding a noise term $\eta_{i,t}$ to Eq. 3 (Lenormand et al. 2009; Nuismer and Harmon 2015) that follows a normal distribution with mean 0 and a variance that is inversely proportional to the effective population size (which we set equal to the actual population size $N_{i,t}$), i.e. $\eta_{i,t} \sim N(0, \frac{1}{2}\frac{V_{i,t}}{N_{i,t}})$.We incorporate demographic stochasticity by drawing species abundances from a zero-truncated Poisson distribution with a mean determined by Eq. 2. Hence, we do not allow for extinction due to demographic stochasticity, but species can become extinct if the phylogeny tells us so (see below).
. Here, we define the fitness of the mean phenotype $\mu_{i,t}$ at the t-th generation using a Ricker–type form of discrete-time population dynamics (see the supplementary material):

$$\begin{aligned}\omega(\mu_{i,t}) &= \frac{N_{i,t+1}}{N_{i,t}} \\ &:= R(\mu_{i,t})e^{-\beta(\overrightarrow{\mu},\overrightarrow{N})/\beta_0}. \end{aligned} \tag{5}$$

Here, $R(\mu_{i,t})$ is the growth factor that depends on the trait value and the parameter θ that represents the optimum trait favored by abiotic stabilizing selection as follows:

$$R(\mu_{i,t}) = R_0 e^{-\gamma(\theta-\mu_{i,t})^2}. \tag{6}$$

where R_0 is the optimal growth factor and γ determines the strength of stabilizing selection towards the optimum. Furthermore, the function β in Eq. 5 quantifies the intensity of competition. Assuming a Gaussian competition kernel, we define β as

$$\beta(\overrightarrow{\mu}, \overrightarrow{N}) = \sum_{j=1}^{n}(e^{-\alpha(\mu_{i,t}-\mu_{j,t})^2}N_{j,t}). \tag{7}$$

Eq. 7 states that the strength of competition between two species with trait means $\mu_{i,t}$ and $\mu_{j,t}$ increases with similarity in these trait means, and that the effect of competition on species i increases with population size of species j. The competition coefficient α scales the strength of the interaction and determines the effective interaction length. Finally, the parameter β_0 in Eq. 5 has a similar interpretation as an individual-scale carrying capacity of each species (Abrams 2001), as it sets the scale at which competitive interactions start to strongly impact the growth of the population. Because Eq. 5 is an increasing function of β_0, the ecological equilibrium $\omega(\mu_{i,t}) = 1$ is reached at a carrying capacity set by the equilibrium condition $\beta = \beta_0 \cdot \ln R$ where environmental stabilizing selection and competition balance each other.
Inserting our fitness function Eq. 5 in Eq. 2, 3, 4 and adding stochasticity leads to

$$N_{i,t+1} \sim \mathbf{Pois}\left(N_{i,t}R_0 e^{-\gamma(\theta-\mu_{i,t})^2} \cdot e^{-\sum_{j=1}^{n}(e^{-\alpha(\mu_{i,t}-\mu_{j,t})^2}N_{j,t})/\beta_0}\right) \tag{8}$$

$$\mu_{i,t+1} = \mu_{i,t} + h^2 \cdot V_{i,t}\left(2\gamma(\theta-\mu_{i,t}) + \Omega\right) + \eta_{i,t} \tag{9}$$

$$V_{i,t+1} = (1-\frac{1}{2}h^2)V_{i,t} + \frac{2N_{i,t}\nu V_m}{1+4N_{i,t}\nu} \cdot h^2 + \frac{1}{2}h^2 \cdot V_{i,t}^2\Big[2\gamma(-1+2\gamma\theta^2) + \frac{\partial\Omega}{\partial\mu_{i,t}} - (2\gamma(2\theta-\mu_{i,t})+\Omega)\cdot(2\gamma\mu_{i,t}-\Omega)\Big] \tag{10}$$

where $\Omega = \frac{2\alpha}{\beta_0}\sum_j(\mu_{i,t}-\mu_{j,t})e^{-\alpha(\mu_{i,t}-\mu_{j,t})^2}N_{j,t}$ and $\mathbf{Pois}(\cdot)$ denotes the zero-truncated Poisson distribution. Eq. 2-, 4 advances Nuismer and Harmon 2015's model by relaxing the simplification of a linear species interaction, as in Drury et al. 2017's model. Moreover, in contrast to Drury's models (Drury et al. 2016, 2017) it takes into account the abundance of competitors in the community as a weight in the competition kernel, and includes the dynamics of trait variance. Furthermore, the full derivation of our model is based on a coherent definition of fitness.