Evolutionary models with ecological interactions

By:

Magnus Clarke

A thesis submitted in partial fulfilment of the requirements for the degree
of
Doctor of Philosophy

The University of Sheffield

Faculty of Science

Department of Animal and Plant Sciences

September 2017

Abstract

Models for the evolution of species' traits and speciation rates usually ignore ecology: most comparative analyses of evolution are ecologically neutral and ignore ecological mechanisms such as competition and limiting similarity. However, such processes can impact profoundly on the distribution of traits across a group of species if they are ecologically similar (e.g. as in adaptive radiations). Here, two new models are introduced to explicitly include the effects of lineage-lineage interactions, one generating trait evolution and the other addressing lineage diversification. The new model for trait evolution is fitted to a wide range of existing animal datasets, using a simulation approach. Evidence was found of clade-wide character displacement patterns in some adaptive radiations, including Darwin's finches; however, these patterns are not prevalent across animal clades as a whole. Three types of diversification model are also compared here, including the novel interaction-based diversification model. This new model links trait space densities to diversification rates, supressing diversification among closely packed species. Although these models have a similar conceptual basis, in terms of the accumulation of filled niches, they produce quite different evolutionary histories. The implications of ecological interactions between species are discussed, both for data interpretation and for future modelling approaches.

List of contents

Abstract 2

List of contents 3

List of figures 7

List of tables 8

Chapter 1: Introduction and overview of the literature 9

Background 9

Project overview 9

Models of continuous evolution 10

Data *10*

Evolution on branches *11*

Evolution on phylogenies *13*

Stretching branches *14*

Multiple traits *14*

Models of speciation-based evolution 15

Evolution on phylogenies *15*

Speciation and phylogeny shape *16*

Diversification rate models 17

Comparing models 19

Relating trait models to data *19*

Model likelihoods *21*

Empirical tests *23*

Software 24

Research questions 25

Trait divergence under competition *26*

Character displacement across animal clades *26*

    Speciation and extinction ... *26*
  Thesis structure ... 27
Chapter 2: Trait evolution in adaptive radiations: modelling and measuring interspecific competition on phylogenies ... 30
  Abstract ... 30
  Introduction ... 31
  Methods ... 34
    The model ... *34*
    Simulation framework ... *36*
    Model comparisons and likelihoods ... *38*
    Case study: Darwin's finches ... *41*
  Results ... 41
    Example of clade evolution under the competition model ... *41*
    Trait distributions across tree tips ... *43*
    Effects of tree size ... *44*
    Effects of competition on correlated traits and phylogenetic signal ... *44*
    Comparisons with alternative models ... *47*
    Power ... *48*
    Case study: Darwin's finches ... *48*
  Discussion ... 51
  Acknowledgements ... 54
Chapter 3: Testing for clade-wide patterns of character displacement in animal clades .. 55
  Abstract ... 55
  Introduction ... 55
  Methods ... 57
    Modelling interspecific competition ... *57*
    Quantifying range overlaps ... *59*
    Data source and preparation ... *59*
  Results ... 60

Mammals ... *60*

Lizards ... *64*

Cichlids ... *65*

Discussion ... 66

Chapter 4: Using phylogenetic comparative models to infer differences in trait evolution between hummingbird clades ... 69

Abstract ... 69

Introduction ... 69

Phylogenetic trait models ... *69*

Hummingbirds ... *70*

Applying phylogenetic trait models to hummingbird clades ... *71*

Methods ... 72

Results ... 75

Trait distributions ... *75*

Trait evolution modelling ... *76*

Character displacement ... *77*

Discussion ... 78

Brilliants and coquettes ... *80*

Bees and mountain gems ... *80*

Hermits and emeralds ... *81*

Character displacement ... *82*

Conclusion ... *83*

Chapter 5: Diversification rates and lineage trait densities ... 84

Abstract ... 84

Introduction ... 84

Diversity dependence ... *85*

Trait evolution and diversification ... *86*

A new model: trait density dependent diversification ... *87*

Goals ... *87*

Methods 88

Model for trait density dependent diversification *88*

Simulations *90*

Summary statistics *91*

Results 92

Tree imbalance *93*

Diversification rate changes *94*

Phylogenetic signal *97*

Effects of tree size *97*

Discussion 98

Theory and assumptions for the three diversification models *98*

Comparing predictions of the three models: *99*

Inferring ecological processes *100*

Conclusion *101*

Chapter 6: General conclusion 102

Summaries of project work 102

Findings 103

Limitations and future directions 104

References 106

Appendices 122

Appendix 2A: Estimating competition strength and effect width simultaneously 122

Appendix 2B: Ornstein-Uhlenbeck model with competition 123

Appendix 2C: Comparing models 123

Appendix 2D: Allopatry in Darwin's finches 125

Appendix 3A: results for competition model + limits 127

Appendix 4A 129

Appendix 5A 131

List of figures

Figure 1.1: An example of evolution under the BM and OU models. 12

Figure 1.2: Accumulation of difference between two lineages 13

Figure 1.3: An illustration of change-scaled trees 15

Figure 2.1: Examples of a single trait evolving under the competition model 42

Figure 2.2: Effects of interspecific competition on trait value distributions 45

Figure 2.3: Traits and contrasts for two traits 47

Figure 2.4: Power to detect competition versus Brownian Motion 50

Figure 3.1: Competition strength, likelihoods and average clade sympatry for mammals. 61

Figure 3.2: Competition strength, likelihoods and average clade sympatry for lizards. 64

Figure 4.1: Hummingbird phylogeny 73

Figure 4.2: Body size and culmen length distributions for 7 hummingbird clades 75

Figure 5.1: Trait-diversity dependence (TDD) model illustration 89

Figure 5.2: Tree imbalance 95

Figure 5.3: Diversification speedup / slowdown 96

List of tables

Table 1.1: Models and parameters for evolution on a known phylogeny 18
Table 2.1: Traits and likelihood ratio test statistics for Galapagos finches. 51
Table 3.1: Competition model estimates and mean sympatry for mammal genera 61
Table 3.2: Competition model estimates and mean sympatry for lizard genera 65
Table 3.3: Results for Neotropical cichlids 66
Table 4.1: evolutionary models fitted to hummingbird body size 76
Table 4.2: evolutionary models fitted to hummingbird beak (culmen) length 77
Table 4.3: character displacement models fitted to hummingbird body size 77
Table 4.4: character displacement models fitted to hummingbird beak length 77
Table 5.1: Model simulations and average summary statistics 92
Table 5.2: Summary statistics for model simulations using different tree sizes. 97
Table 5.3: Matrix of the qualitative predictions of the three diversification models. 99
Table 2C.1: Parameter values and summary statistics 124
Table 2D.1: Sympatry matrix for Darwin's finches 126
Table 2D.2: Results for Darwin's finches with partial allopatry 127
Table 3A.1: Competition model with limits: mammal genera. 127
Table 3A.2: Competition model with limits: lizard genera 129
Table 3A.3: Competition model with limits: neotropical cichlids 129
Table 4A.1: character displacement models fitted to hummingbirds 130
Table 5A.1: Summary statistics for model simulations using different tree sizes 131

Chapter 1: Introduction and overview of the literature

Background

Evolution describes the change in animal species over time. Ecology describes the relationships between these species, and the effects they have on one another. Evolution and ecology are known to affect one another, and some progress has been made in understanding the roles of ecological processes in shaping trait evolution (Johnson and Stinchcombe 2007, Cavender-Bares et al. 2009, Schoener 2011, Pennell and Harmon 2013, Hadfield et al. 2014, Price et al. 2014, Pigot and Etienne 2015). The exponential increase in molecular sequencing data in recent years means that more detailed pictures of clades' evolutionary histories are available, and this helps to test more sophisticated models of evolution and ecology (e.g. Rabosky and Lovette 2008, Harmon et al. 2010). These models link plausible evolutionary processes to their results in patterns of diversification or trait values.

Competition between species is the root of some key ideas at the interface between evolution and ecology: character displacement (Grant 1972, Schluter 2000) and limiting similarity (MacArthur and Levins 1963). Character displacement is the evolution of species traits to become more different, reducing competition for any particular niche. It is most commonly detected by observing a given pair of species to be more different in regions where they coexist than elsewhere (Pfennig and Pfennig 2010). Similarly, limiting similarity, the idea that species cannot coexist if their niches are too similar (Abrams 1973), is often thought to slow the rate of accumulation of new species, since empty niches may be needed for new species to fill (e.g. Phillimore and Price 2008).

Project overview

There are many statistical models of trait evolution and species diversification (reviewed below; see also O'Meara 2012 for a technical overview), but these models do not typically include ecological mechanisms such as competition and limiting similarity. At the same time, ecological effects are often expected across groups of species when they are ecologically similar, for example ecological opportunity across adaptive radiations (Yoder et al. 2010). The overall purpose of this project was to develop evolutionary models with explicit assumptions about the underlying ecology, and to develop tests that allow us to distinguish between these models using phylogenies and trait data.

The ecological focus for the project was on competition between species. Character displacement is a familiar and well tested idea, but there are comparatively few studies looking for signs of it on a clade-wide scale (but for an example see Davies et al. 2012). Also, although the idea is familiar, it has not traditionally been included in models of trait evolution on phylogenies. Instead, trait evolution is assumed to be independent for each lineage. There is a good reason for making this assumption: it makes it straightforward to generate expressions for model likelihoods. New likelihood methods are just beginning to appear for comparative phylogenetic models with lineage-lineage interactions (Drury et al. 2016, Manceau et al. 2017, Bartoszek et al. 2017). The approach I took to this problem was to create software for simulating data under a novel, more complex model that explicitly includes interactions, and then estimating the likelihood by comparing these simulations to observed datasets.

I also looked at the effects of competition on diversification. Models exist to make diversification a function of the total number of species in a group, so that diversification rates can slow down as the species count increases and niches are filled up (Walker and Valentine 1984, Etienne et al. 2012). I compared the predictions of different models of this type, and introduced a new, related diversification model. By contrasting these similar models, I show how the details of their assumptions manifest in different predictions about phylogeny shape and trait distributions.

Models of continuous evolution

Data

Models of continuous trait change on a phylogeny can be used to generate simulated datasets, or can be fitted to observed datasets. Fitting these models requires three types of data:

- Tree topology;
- Branch lengths;
- Trait values of tree tips.

The models give instantaneous rates of evolution as a function of position on the phylogeny branch. Then, the difference between two species is predicted from the length of branches separating them. The tendency for phylogenetically close species to be phenotypically similar is referred to as ‘phylogenetic signal’ (Blomberg et al. 2003). Phylogenetic niche conservatism (PNC) is a related concept, with no universal definition (Cooper et al. 2010). Loosely, it is phylogenetic signal where the trait is the species’ niche. Phylogenetic signal is often assumed to be particularly strong for such traits, but this assumption may be frequently unsatisfied (Losos 2008).

The best data for comparative analyses consists of many closely related pairs of species that differ in trait value, with substantial time between speciation events for each lineage (Garland et al. 2005). Ancestral trait values can be very useful in fixing non-tip phylogeny nodes, but are rare and generally not incorporated into phylogenetic comparative analyses (Harmon et al. 2010a).

Evolution on branches

The need to account for phylogeny in comparative analyses was first made clear by Felsenstein (1985). He used a model where the change in trait value in a short period of time is drawn from a normal distribution: this is Brownian Motion (BM). Other models of trait evolution along a branch are derived from the BM model by adding parameters. Here, the various models are listed in time-differential form in table 1.1.

The BM model has a trait value X evolving at random, at a rate σ:

$$dX(t) = \sigma dW(t), \qquad \text{Equation 1.1}$$

where $W(t)$ is the integral of the continuous white noise function, such that $\Delta W_t \sim N(0, \Delta t)$. If used to predict trait values at a time T, the BM model has two free parameters: the evolutionary rate σ and the root trait value $X(0)$.

The parameter λ (Pagel 1997, 1999) measures the goodness of fit of trait data to those expected under a given BM model. A value of 1 indicates a consistent fit to the BM model; $\lambda < 1$ indicates less phylogenetic signal and $\lambda > 1$ indicates more signal than expected. λ can be represented as a transformation of the tree into one with internal branches rescaled relative to tip branches.

We know that rates of evolution vary. How do we add plausible patterns of rate-change to the BM model? Looking at equation 1.1, we seem to have two options: we can elevate the factor σ to the status of a function, or we can add a separate, deterministic term to the equation. Time-dependency of σ could be linear ($\sigma \rightarrow (t_{max} - t)\sigma$), polynomial ($\sigma \rightarrow \kappa t^{\kappa-1}\sigma$), or exponential ($\sigma \rightarrow g^{-t}\sigma$).

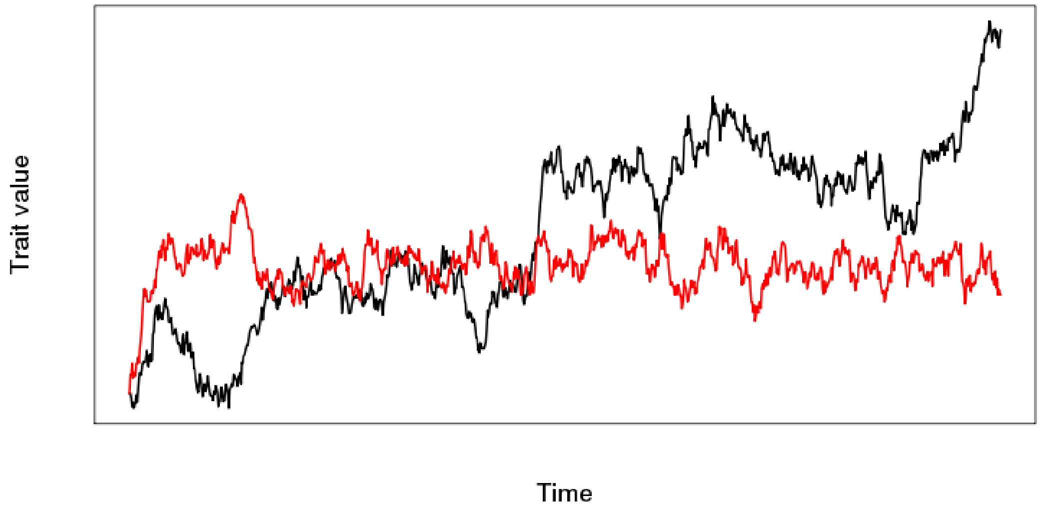

Figure 1.1: An example of evolution under the BM (black) and OU (red) models.

Polynomial time dependence κ is first seen in Pagel (1997). Exponential models, meanwhile, are termed Accelerating/Decelerating (ACDC) models (Blomberg et al. 2003). Each of these models adds one free parameter to the BM model. One more alternative exists and has been implemented: the evolutionary rate σ may be a step function of time; these steps (i.e. discrete rate changes) may represent sudden environmental changes, or transitions into or out of ecological niches (Thomas et al. 2006, O'Meara et al. 2006). If some extrinsic event is known about, then the positions of the corresponding steps can be built into the model before the model is used. Alternatively, maximum likelihood (ML) positions and sizes of discrete rate changes can be estimated from the data (Eastman et al. 2011, Thomas and Freckleton 2012, Revell et al. 2012, Rabosky et al. 2013).

The second way to modify the BM model of equation 1.1 is to add an additional term. In the OU model (Hansen 1997) the trait X is drawn towards a central value ('primary optimum') ψ with a strength proportional to its distance from the optimum:

$$dX(t) = -\alpha(X(t) - \psi)dt + \sigma dW(t) \qquad \text{Equation 1.2}$$

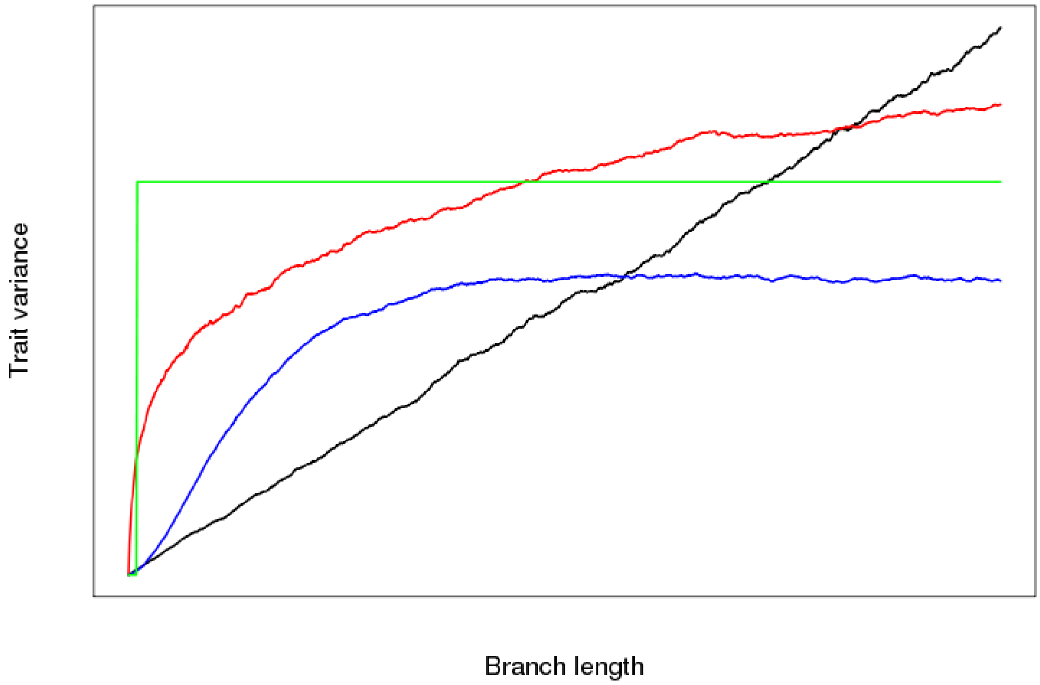

Figure 1.2: Accumulation of difference between two lineages under different models of evolution. Branch lengths in units of time. The black line is the BM model; the red line is κ<1; the blue line the OU model, with a different optimum for each branch; the green line is a NF model.

For this reason, the simplest implementation of the OU model is sometimes called the 'single stationary peak' (SSP) model. It is perhaps easier to imagine this as an ecological process of avoiding extreme trait values: as a species gets further from the clade mean, it becomes more likely to evolve towards the clade mean. The OU model can be extended by adding discrete optima shifts based on prior knowledge of extrinsic events, or with methods to estimate ML positions of optima shifts (Hansen 1997, Hansen and Orzack 2005). Additionally, the optima may themselves evolve according to a BM or OU model (Hansen and Houle 2008, Hansen and Bartoszek 2012). An example of the accumulation of difference between two lineages is shown for the BM, OU, κ and NF models in figure 1.2.

Evolution on phylogenies

Most of the BM-like models of evolution can be applied separately to each branch, but we can also make evolutionary rates dependent on position in the whole phylogeny. Firstly, we might test for a phylogeny-wide change in the rate of evolution, analogous to the within-branch parameter κ (Pagel 1997). This parameter takes the same form as κ, and is referred to as δ

(Pagel 1997, Thomas and Freckleton 2012) (See table 1.1). We can also permit rate parameters to be branch-specific. Having an independent σ for each branch is possible (Mooers et al. 1999), but will tend to lead to too many free parameters (Thomas and Freckleton 2012). ML estimation of a limited number of discrete rate changes, however, can be applied clade-wise as well as time-wise, so that a change appears on just one branch in a phylogeny, but is inherited by subsequent 'offspring'.

Stretching branches

The evolutionary models in the previous section allow us to calculate the amount of change along each tree branch, assuming that the length of the branch corresponds to amount of time passed. However, an alternative perspective is to apply only the BM model to the tree, but implement alternative models by first transforming the tree's branch lengths. The evolutionary models effectively become maps from time-scaled trees into change-scaled trees. Some models were originally presented as such a transformation. An example tree transformed under each of several trait models is shown in figure 1.3.

Multiple traits

The BM and BM-like models generalise immediately to multiple traits: the trait value X becomes a vector of trait values $\boldsymbol{X}$, and the rate parameter σ becomes a covariance matrix Σ. Nonzero nondiagonal elements of Σ indicate correlation between traits. This is important, since selection acts on many traits simultaneously and a particular selective force may act on some combination of measured traits.

One of the main uses for phylogenetic methods is to measure correlations between traits while controlling for the phylogeny. OU model variants have been developed to account for coadaptation between traits, with trait optima either fixed or evolving by the BM model. With multiple traits, α becomes a matrix, and off-diagonal elements can permit one trait's optimum to change according to another trait's value, even if the second trait is simply evolving according to BM (Bartoszek et al. 2012).

It may be useful to note a potential source of confusion: many methods seek to separate phylogenetic and ecological effects. However, ecology may be heritable. If two species are labile but have phylogenetic signal due to shared inherited ecology, then they are independent data points with respect to the details of that ecology, but non-independent with respect to general ecological principles.

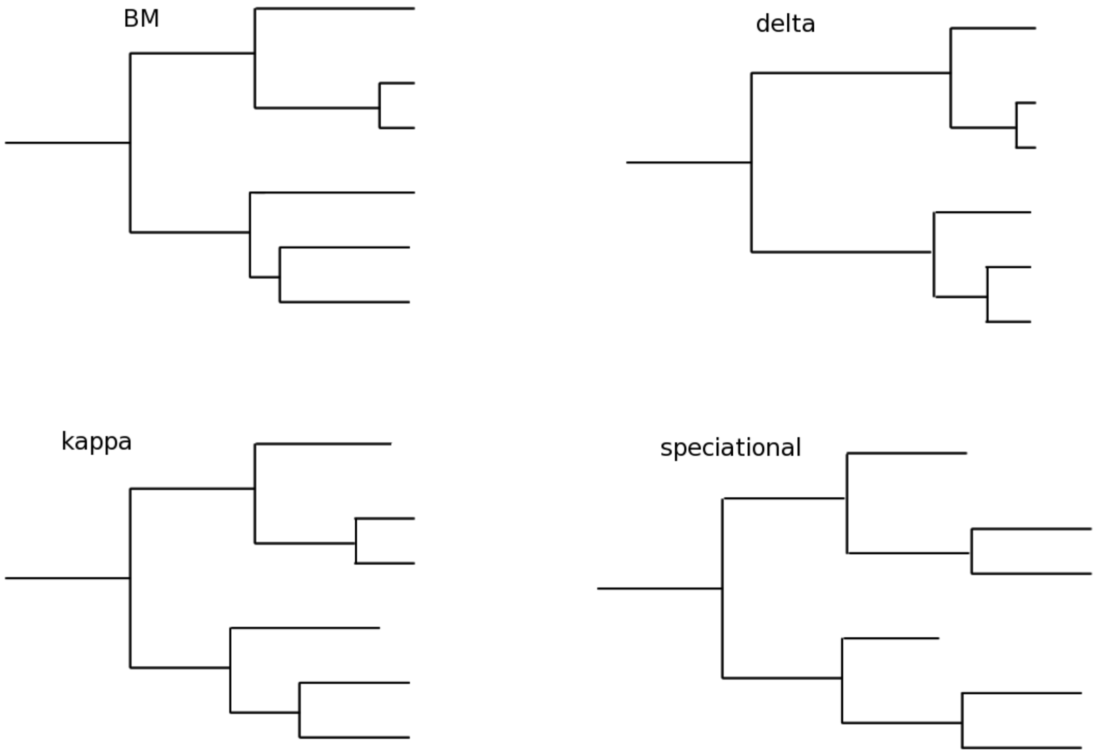

Figure 1.3 : An illustration of change-scaled trees. The BM tree is identical to a time-scaled tree, and different evolutionary models are different transformations into other trees. The values of kappa (κ) and delta (δ) are between 0 and 1.

Models of speciation-based evolution

Evolution on phylogenies

Branch-evolution parameters such as κ, discussed above, represent one type of test for speciational evolution. An alternative is simply to postulate a 'lump' of trait change at each speciation event, with the lump drawn from $N(0, \sigma_c)$ (Bokma et al. 2008, Ingram 2010). If the speciation rate is ξ, then we can measure the degree to which rapid evolution is linked to speciation events with the value of

$$\frac{\xi\sigma_c^2}{\sigma_{total}^2} \qquad \text{Equation 1.3}$$

This model requires extinction rate estimates, which can be estimated from the phylogeny itself (Rabosky 2009, Pybus and Harvey 2000). If the amount of Brownian Motion is set to zero, then the Ingram (2010) model simply reduces to random amounts of trait change at speciation events and none in between speciation events. This corresponds to a pure ‘punctuated equilibrium’ model of evolution.

Niche-filling (NF) models (Price 1997, Harvey et al. 2000, Freckleton and Harvey 2006) also have purely speciational modes of evolution, but have a few differences from the Ingram model. In niche-filling models, species’ trait values are constant in time, and the position of each niche in niche-space is constant in time. New niches are filled by new species, branching from whichever species is closest in niche-space. Freckleton and Harvey (2006) distinguish between NF models where every species is equally likely to speciate and form a new and filled niche, and NF models where niches appear at a constant rate, randomly positioned in niche-space. The two are equivalent when all the species have evenly distributed trait values, but have different predictions otherwise; for example, a species particularly distant from all others in trait-space will be more likely to speciate under the randomised NF model.

In the Ingram model speciation is a splitting process, where each of the two species shifts to a new trait value. However, in NF models speciation is a branching process, where only the ‘new’ species moves to a new trait value. This leads to a difficulty for NF models, however: if a lineage splits and the ’original’ lineage subsequently goes extinct, then a phylogeny built from extant species will have unexplained mid-branch evolutionary change.

In NF models, species respond very quickly to environmental changes (via speciation), but phylogenetic signal exists and persists, because new niches are filled from nearby niches. Hence, sibling species resemble one another. The BM model achieved this signal by the contrary postulate that evolution is purely random and unconstrained.

A further different model combines phylogenetic signal from a BM process with a postulate of trait variance increasing linearly with spatial separation between the species (Freckleton and Jetz 2009). The relative importance of the two terms of the expression for the variance then measures the importance of physical locale in the evolution of the trait.

Speciation and phylogeny shape

BM-like models require a phylogeny as input data, and provide no insights on the question of when and why speciation should occur. If $b(t)$ is the birth rate and $d(t)$ is the death rate, then a ‘Yule process’ has $d = 0$ and b constant so that the number of species increases linearly with time. The γ statistic tests for deviations from the Yule model; it is a measure of acceleration in the rate of speciation, with standard normal distribution under the Yule model (Pybus and Harvey 2000).

A reduction in speciation rate through time is consistent with NF models that accommodate a limited number of niches. However, this pattern can be seen even in models which do not have such ecological effects; therefore, it is important to compare model predictions and not just test a null model of constant speciation rate (Rabosky 2009). One solution is to look for co-occurring slowdowns in speciation rate and trait evolution rate (Harmon et al. 2010a).

Diversification rate models

Molecular phylogenies can also be used to estimate rates of diversification, that is, the rate of change of total number of species. Although a molecular phylogeny will usually include only extant species, it can in principle be used to infer both speciation and extinction rates (Nee et al. 1994). The overall diversification rate, equal to the speciation rate b minus the extinction rate d, is given by the total accumulation of species and the lifetime of the clade. The extinction rate can then be computed from the diversification rate b-d and the 'relative extinction rate' d/b. The relative extinction rate is essentially the apparent change in diversification rate through time. When the extinction rate is small, the diversification rate will approach the speciation rate b. But when the relative extinction rate is high, the diversification rate will appear to increase through time, meaning that extant species are generally younger. The rate of diversification therefore appears to accelerate with time (Rabosky 2006). From this framework, likelihoods for speciation and extinction rates can be calculated for phylogenetic data. The reliability of extinction rate estimates, however, relies on assumptions, such as equality of diversification rates across taxa, which are typically broken in real data (Rabosky 2010).

There are many models of variation in diversification rates. The effect of binary traits on speciation and extinction rates can be estimated using BiSSE (Maddison et al. 2007, Magnuson-Ford and Otto 2012), a likelihood model. BiSSE has been extended to multiple discrete variables in the MuSSE model (Fitzjohn 2012). QuaSSE (Fitzjohn 2010) is a similar approach, estimating the effect of continuous traits on rates and generating phylogeny likelihoods. QuaSSE assumes no interactions between species and no diversification rate diversity dependence. There are methods for detecting discrete rate changes whose time is not known a priori (Rabosky 2014).

Table 1.1: Models and parameters for evolution on a known phylogeny. Multi-trait parameter counts are upper limits assuming free covariance between all traits.

Model	Equation	parameters	parameter count
BM	$dX(t) = \sigma dW(t)$	root value, $\sigma = rate$	2
	$var(X) = T\sigma^2$	T is time from common ancestor	
κ	$dX(t) = \kappa\sigma t^{\kappa-1} dW(t)$	$\kappa > 0$, t from branch start,	3
	$var(X) = \Sigma_i T_i \sigma^2$	i labels each branch, and T_i is its length	
δ	$dX(t) = \delta\sigma t^{\delta-1} dW(t)$	$\delta > 0$, t from tree root	3
	$var(X) = \sigma^2\left(T^\delta - T_0^\delta\right)$	T is time over tree, T_0 is time of MRCA	
Step	$dX(t) = \sigma(t, clade) dW(t)$	δ step function with m steps	$1 + 2m$
ACDC	$dX(t) = \sigma g^{-t} dW(t)$	$g > 0$	3
OU	$dX(t) = -\alpha(X(t) - \psi)dt + \sigma dW(t)$	ψ is optimum, strength α	4
	$var(X) = \frac{\delta^2}{2\alpha}(1 - e^{-2\alpha T})$		
stepOU	$dX(t) = -\alpha(X(t) - \psi_i)dt + \sigma dW(t)$	i is number of fitted optima	4
Ingram2010	BM + $N(0, \sigma_c^2)$ per speciation		3

A related source of slowdowns is diversity dependence (DD). This is where the diversification rate for each lineage depends on the total number of lineages in the clade. A fixed number of niches are considered to be available to the whole clade (Walker and Valentine 1984, Etienne et al. 2012), and as they get filled up, the speciation rate declines. Diversity dependence has been modelled using this number as a hidden Markov variable, with extinction and speciation generating continual species turnover (Etienne et al. 2012). The equation for this process is

$$\lambda_n = \max\left(0, \lambda_0\left(1 - \frac{n}{K}\right)\right), \qquad \text{Equation 1.4}$$

where n is the number of species, λ_0 is the initial speciation rate, and K is the maximum number of available niches.

NF models (Price 1997, Harvey and Rambaut 2000, Freckleton et al. 2000, Price et al. 2014) were discussed earlier in the context of trait evolution. However, these models integrate speciation and trait evolution, so that they produce both the traits and the corresponding phylogeny. NF models are conceptually related to density-dependent diversification, since it is based on the gradual filling up of available niches; however, the models operate differently, and the NF model predicts—by definition—the creation of new species and at regular time intervals, and therefore a slowdown in per-lineage diversification rates. Observed slowdowns are sometimes, therefore, taken as evidence of niche filling (e.g. Phillimore and Price 2008).

Comparing models

Relating trait models to data

Given a tree and an evolutionary model, we want to know the probability of obtaining the observed phenotypic tip data. One popular method is independent contrasts, developed by Felsenstein (1973, 1985). The difference in trait value between sibling species depends only on the branch length (in units of expected evolutionary change) separating them, and is therefore independent of shared evolutionary history. Since trees bifurcate, each node has exactly one sibling, so there are n-1 independent contrasts in a tree with n tips. The uncertainty in non-tip node trait values is accommodated by lengthening their branch by an amount $v_i v_j/(v_i + v_j)$, where v_i, v_j are the branch lengths of the species descended from that node. Methods exist to use trees with unresolved nodes (Pagel 1992).

The rate σ^2 of Brownian evolution can be estimated with a generalised least squares (GLS) method (Pagel 1997). This method, equivalent to the independent contrasts method, uses regression such that each tip has a trait value

$$X_i = \alpha + \sigma^2 \Sigma v_j + e_i, \qquad \text{Equation 1.5}$$

where v_j is the total branch length separating tip i from tip j, and e_i is a residual. The evolutionary rate can then be estimated as

$$\sigma^2 = \frac{1}{n-1} X^T V^{-1} X, \qquad \text{Equation 1.6}$$

where X is a vector of the tip trait values, and V is the covariance matrix, i.e. a matrix of the branch lengths (or for non-BM models, the transformed branch lengths) shared by each pair of tips. Non-phylogenetic methods are the subset of GLS methods with V diagonal.

The autocorrelation method (Cheverud et al. 1985), like GLS, uses regression, partitioning the between-species trait variance into heritable (phylogenetic) and specific components. Then, covariance between traits of specific, but not heritable, components is evidence for coadapted traits. A covariance matrix is used, but not derived. The autocorrelation method generally performs less well than the independent contrasts method (Diaz-Uriarte 1996).

The animal model, or mixed model, used in quantitative genetics, can be adapted for comparative phylogenetic analysis (Lynch et al. 1991). A species' trait value is multiply regressed on phylogenetic effect and the values of other traits, with additional residual terms. Efficient calculational methods exist, and within-species variation is readily included (Hadfield and Nakagawa 2010). The latter point is important: two identical populations imperfectly sampled will look different, causing independent contrasts to be overestimated. Restricted ML techniques now exist to correct for this within the independent contrasts method (Ives et al. 2007, Felsenstein 2008). This will be particularly important when comparing the BM and OU models, since the variation around a 'primary optimum' present in the OU model will resemble this bias (Cooper et al. 2016).

The independent contrasts method removes phylogenetic effects without estimating them. This makes it computationally faster, and generally better performing when assumptions are broken. Methods which use the covariance of each tip with every other tip apply the evolutionary model to every point in the tree, not just the section containing the taxa under comparison. However, if the phylogenetic component of a species' trait is what we want to know, then regression methods are more appropriate. As evolutionary models become more complicated, and rates of evolution are modelled as functions of trait values, it seems likely that computer simulation will become preferred to the above methods. The main limitation of simulation is simply computational time (Garland 1999).

When studying correlations between traits, it is not obvious that phylogenetic methods are superior to simple analyses of raw data. This is because phylogenetic methods, including independent contrasts, make assumptions about the evolutionary process. If these assumptions are false, there are conditions under which analyses of raw data can be more accurate than those with erroneous phylogenetic corrections (Price 1997, Harvey and Rambaut 2000). This is one reason for using tests for phylogenetic signal (Pagel 1997, Bjorklund 1997). However, phylogenetic methods in general, and independent contrasts in particular, are usually well supported and robust to perturbation away from their assumptions (Harvey and Rambaut 1998, Martins and Hansen 1997, Diaz-Uriarte 1996).

One difficulty with most approaches is that they assume perfect knowledge of the phylogeny. Methods exist to account for unresolved nodes (Pagel 1992, Slater 2012). However, this does not make use of the likelihood data which is usually generated by sequence-data phylogeny building. Blackburn et al. (2013) demonstrated the possibility of using a posterior distribution of phylogenies to account for phylogenetic uncertainty.

Model likelihoods

The likelihood of hypothesis H given data D is $L(H|D) = \frac{P(D|H)P(H)}{P(D)}$. A ratio of hypothesis likelihoods is then $\frac{L(H_1|D)}{L(H_0|D)} = \frac{P(D|H_1)P(H_1)}{P(D|H_0)P(H_0)}$. With no prior expectations of model likelihoods, the ratio becomes $\frac{P(D|H_1)}{P(D|H_0)}$. Models with more free (fitted) parameters should fit the data better. To avoid overparameterisation, we therefore have to require a ’significant’ improvement in fit from the more complex model than the simpler model. There are various approaches to determining this significance, including likelihood ratio tests (LRT), the Akaiki information criterion (AIC) and Bayesian methods.

Using LRTs to compare two models results in a test statistic which is the log of the ratio of their likelihoods. When H_0 is a special case of H_1, so that the models are ’nested’, then the test statistic forms (0.5 times) a χ^2 distribution. This distribution, however, also assumes large samples, that one of the models is true, and that parameters are normally distributed; these assumptions may sometimes be significantly violated by phylogenetic methods (Freckleton 2009). To avoid these difficulties, we can take the parameter maximum likelihood estimates (MLEs) from the null model, and simulate new datasets from those parameters. For each dataset, new MLEs are generated, and the log-likelihood ratio calculated. The distribution of LRTs then allows us empirically to map LRT values to P-values.

The AIC is a number assigned to each model: the difference between the maximised log-likelihood and the number of free parameters K:

$$AIC = -2l + 2K. \qquad \text{Equation 1.7}$$

'Akaiki weights' then represent relative likelihoods of models:

$$w_i = \frac{e^{-\Delta_i}}{\sum_{r=1}^{R} e^{-\Delta_r}} \; \Delta_i = AIC_i - \min(AIC) \qquad \text{Equation 1.8}$$

Like the LRT, the AIC assumes a large sample size with parameters that are multivariate normal (Posada and Buckley 2004). However, the AIC has a key advantage over LRTs in that non-nested models can be tested without the need for simulation and bootstrapping.

When likelihoods themselves are difficult to calculate, we can estimate them by using the model to generate new simulated datasets, and comparing this distribution of datasets with the observed data. We can then choose to use the likelihood for the best-fit model parameters, as in the LRT, or to integrate over all model parameters according to a prior distribution of parameter values, chosen before fitting the model. One implementation of the latter approach is approximate Bayesian computation (ABC). In ABC, a set of parameter values is sampled from the prior distribution, and some data $\widehat{D}$ is simulated. For observed data D and tolerance ε, we accept $\widehat{D}$ if

$$\rho\left(\widehat{D}, D\right) < \varepsilon, \qquad \text{Equation 1.9}$$

where ρ is the discrepancy, or distance in solution space, between $\widehat{D}$ and D. The set of parameter values which produce accepted instances of $\widehat{D}$ are then taken to be a sample from the posterior distribution of parameter values. To compare models, each model's likelihood is taken to be proportional to the fraction of simulations accepted. Then we can use the Bayes factor:

$$K = \frac{P(D|M_1)}{P(D|M_2)} = \frac{\int P(\theta_1|M_1)P(D|\theta_1,M_1)d\theta_1}{\int P(\theta_2|M_2)P(D|\theta_2,M_2)d\theta_2}, \qquad \text{Equation 1.10}$$

where θ is the set of parameter values.

Fitted-model likelihood ratios and integrated-model likelihood ratios represent two different measures of model usefulness, and it is probably advisable to calculate both and compare in order to learn more about the truth. By generating data under MLE model parameters, we can also visualise the distribution of modelled data alongside the observed data, to gain an idea of the model's adequacy in describing real data.

Empirical tests

Diffusion models of trait evolution have been applied to observational data covering a wide range of animal clades. A review of comparative studies found that λ is typically high, consistent with strong phylogenetic signal (Freckleton et al. 2002). This suggests limited applicability of the OU model, which predicts decay of signal. The presence of phylogenetic signal suggests that a BM or NF model will typically be best, but does not automatically distinguish between them (Cooper et al. 2010). This question depends on evolutionary gradualism; for example, some studies have found that two-thirds of variation in body mass is speciational (Mattila and Bokma 2008, Bokma 2008). However, it is important to make phylogenetic signal estimates jointly with ecological models, not prior to fitting the ecological models (Hansen and Orzack 2005). An additional source of phylogenetic signal can be spatial effects: if closely related species also tend to be geographically closer then they may share adaptations to that local environment (Garland et al. 2005).

BM models have successfully been rejected in favour of NF models for data on warbler birds using two tests (Freckleton and Harvey 2006). Firstly, tests for correlation between independent contrasts (Felsenstein 1985) and phylogenetic positions of the contrasted species reveal links between divergence rates and phenotype (i.e. position in niche-space). Secondly, testing for an overall slowdown of evolution across the phylogeny can reveal constraints arising from the available niches getting 'filled up'. Other studies also find speciation rate slowing with time, consistent with a limited number of niches being filled, but find that phenotypic evolution does not share this slowdown (Burbrink et al. 2012). This pattern might be consistent with a NF model with niche positions evolving randomly in trait space. Alternatively, it could be that the niche is defined by a complex combination of traits, such that that combination is conserved while individual trait values are not (Crisp and Cook 2012). NF models can also be tested from a diversification perspective, as a cause of slowdowns in diversification rate (Price et al. 2014, Pigot et al. 2016).

Ecological release (the removal of selective constraints) is sometimes linked to adaptive radiations, but frequently is not (Yoder et al. 2010). The chances of adaptive radiations may depend on fluctuations in population size and density, in turn dependent on fluctuations in selection strength and direction (Siepielski et al. 2009, Futuyma 2010). Since long-term stasis can arise from short-term fluctuations, evolution can be the result not of environmental change but of environmental stability (Futuyma 2010). Yoder et al. (2010) conclude that many factors affect the link between ecological release and speciation, and that further study should assess the commonness of these factors, and their ability to reinforce or cancel each other. They recommend that population-genetic parameters such as effective population size and trait variance be included in models of long-term evolution.

Any overall change in diversification rate with time can be quantified by the statistic gamma, where $\gamma = 0$ under a Yule prediction, and negative γ is indicative of a slowdown. The median γ across 45 bird clades was found to be around -1.2, and 15 of the 45 clades had significantly negative gamma ($\gamma < -1.645$) (Phillimore and Price 2008). Diversification rates have also been shown to vary by key morphological innovations (Dumont et al. 2012), or more generally by species' trait values (Paradis 2005, Maddison et al. 2007, Freckleton et al. 2008, Fitzjohn 2010) or rates of trait evolution (Adams et al. 2009, Rabosky et al. 2013). For example, bird speciation rates vary with ecological factors such as annual dispersal and feeding generalisation (Owens et al. 1999, Phillimore et al. 2006). These rates are thought to be diversity dependent, i.e. dependent on the total number of species (Phillimore and Price 2008, Rabosky and Lovette 2008), resulting from competition for ecological resources (Price et al. 2014).

Software

Software exists to visualise, simulate and fit evolutionary models to phylogenies and tip data. Most of this software exists as packages for the R platform (R Team 2015). Trees can be stored as Newick files, with all tips extant and nested branch lengths. Nexus files can contain various data including Newick trees. The *ape* format (Paradis et al. 2004) is an alternative which also allows nodes to be labelled, and arbitrary tip dates to be set, by labelling each branch by its parent and offspring species and specifying branch lengths explicitly.

GEIGER (Harmon et al. 2008) is used to generate simulated phylogenies and trait data. It can randomly 'prune' clades to mimic incomplete sampling. The trees are created with birth-death models, and the tip data are generated under the BM model, with discrete and continuous traits. Multiple continuous traits can be simulated given a covariance matrix. *GEIGER* can perform AIC tests for significantly nonzero rate-change parameters.

The *caper* package (Orme 2013) allows model fitting to trees and tip data using the independent contrasts and the GLS methods. It uses *ape* data, and requires the *ape* package. The transformation parameters $\delta,\ \kappa,\ \lambda$ can be estimated and tested with ANOVA or AIC. Trees can be simulated with birth-death models, and tip data can be simulated with a BM model.

The *ouch* package (King et al. 2012) fits an OU model to tree and tip data, with α and σ as free parameters. Clades can be chosen to have independent estimations. Multivariate estimates can be made with a symmetric α-matrix and a lower-diagonal σ-matrix. Phylogenetic signal can be measured using the function *gls* in the R package *ape*, for computing λ, and the function *Kcalc* in the R package *picante* (Kemble et al. 2010) for computing K (Blomberg et al. 2003).

MOTMOT (Thomas and Freckleton, 2012) is a package used to fit a specified number of discrete rate changes in evolutionary rate, and to compare likelihood with the pure BM model. Both the size and position of the changes are estimated. This contrasts with the *ouch* package, where clades must be chosen for independent parameter fitting beforehand. *MOTMOT* can also be used for AIC tests of nonzero continuous rate change parameters, and for Ingram's (2010) ψ parameter. *MOTMOT* calculates the likelihood in closed-form solution for a rate-change at each tree node. Another package, phytools (Revell 2012, Revell et al. 2012) fits discrete rate changes with an MCMC approach, permitting mid-branch rate changes.

Similarly, software packages exist for the purpose of measuring phylogeny shape, and fitting diversification models. The shape statistic γ can be computed using the R package *ape* (Paradis et al. 2004). Discrete diversification rate shifts can be estimated with software such as BAMM (Rabosky 2014) and the R package *TreePar* (Stadler 2011). BAMM, for example, is used to link morphological innovations to diversification rate changes in scincids in Rabosky et al. (2014). Gradual, rather than discrete, changes in rate can be fitted using the R package *LASER* (Rabosky and Lovette 2008).

Explicit dependence of diversification rates on traits can be modelled for binary traits with the BISSE model (Maddison et al. 2007, Fitzjohn et al. 2009), and for quantitative traits with QUASSE (Fitzjohn 2010). Both these models can be fitted using the R package *Diversitree* (Fitzjohn 2012). Traits dependence models are reviewed in Ng and Smith (2014).

Research questions

The overall research goal for the project was to investigate how competition between species for niches affects the evolution of those species. This relates to both their trait evolution and the rates of speciation and extinction. The goal breaks down into a few broad research questions, which are discussed in further detail below. Firstly, I ask what patterns can be expected in phylogenies and trait data if competition drives trait divergence. Second, I investigate the prevalence of character displacement in animals, and its correlation with observed sympatry. Finally, I consider the effect that niche competition has on speciation and extinction.

Trait divergence under competition

In Chapter 2, I introduce a novel phylogenetic comparative model. The main goals here include:

1. Develop a new phylogenetic comparative model that explicitly includes character displacement interactions.
2. Characterise this model in comparison with other models (such as Brownian and Ornstein-Uhlenbeck trait diffusion models), using simulations.
3. Develop the means for fitting this model to datasets.
4. Determine the statistical power for detecting character displacement patterns (compared with a Brownian null model).

This new model is also applied to an example dataset: Darwin's finches. This is an adaptive radiation whose evolution has been well characterised in many previous studies, so this study can be placed in the context of past results. The aim was to demonstrate the use of the new model with a small, well studied adaptive radiation.

Character displacement across animal clades

Chapters 3 and 4 apply the model from chapter 3 to a range of animal datasets. First, I perform a broad survey of datasets available from the literature and fit the model to each. Then in chapter 4, this and other models are used with American hummingbirds: a geographically widely distributed group with extensive range overlap among subsets of species. I also fit a range of other phylogenetic comparative models to these hummingbird clades, and assess variation in evolutionary mode within the hummingbirds. The key research questions for these studies are:

1. How prevalent is clade-wide character displacement among animal clades?
2. What relationship is there between sympatry levels across a clade, and evidence of character displacement?
3. How much variation is there in the mode of trait evolution among hummingbird clades? (The hummingbirds are often studied as one whole.)
4. Are the evolutionary processes that have governed each hummingbird clade correlated with those clades' main environment types?

Speciation and extinction

In Chapter 5, I introduce a new model of diversification. This model has a similar conceptual background to other 'diversity dependence' models, but with a different implementation. I explore how small differences in approach between these models can produce very different results. Tree shape and trait distributions are often used to infer past processes such as diversity dependence, so it is important to clarify exactly which assumptions generate what patterns in these datasets. The goals for the diversification study include:

1. Develop a new diversification model, where lineage diversification rates are coupled to trait values, and diminish as traits become more tightly packed.
2. Compare the predictions of this new model with those of other, conceptually related models.
3. Link ecological processes to the particular assumptions of these models, and show how they can best be inferred from phylogenies and trait datasets.

Thesis structure

There are four main pieces of work in the thesis, contained in the four main chapters (2, 3, 4 and 5). These are preceded by this introduction and literature review. The literature review is included for completeness, but each main chapter is self-contained and has an introduction reviewing the appropriate literature. Likewise, each chapter has its own discussion; however, a brief conclusion to the whole thesis is given after the main chapters.

Chapter 1: Introduction and literature overview

This chapter sets forth the research questions for the project, and describes the structure of the thesis. It also reviews the existing literature review for models of trait evolution and diversification using phylogenies. These models are contrasted with one another, and some examples are given of their applications. Various software packages are available for fitting these models and simulating datasets using them; some of these programs and libraries are also discussed, with focus on the tools that are used in later chapters.

Chapter 2: Trait evolution in adaptive radiations: modelling and measuring interspecific competition on phylogenies

Chapter 2 presents a new model of trait evolution, with an explicitly parameterised character displacement effect. This new model is fitted to a small example dataset: the Darwin’s finches. This chapter was published in American Naturalist as:

- Clarke, Magnus, Gavin H. Thomas, Robert P. Freckleton. 2017. Trait Evolution in Adaptive Radiations: Modeling and Measuring Interspecific Competition on Phylogenies. The American Naturalist 189:121-137.

The methods for fitting the new model are found to have good power to detect clade-wide character displacement patterns, and evidence of such patterns is found in the beak size of Darwin's finches.

Chapter 3: Testing for clade-wide patterns of character displacement in animals

The trait evolution model that was developed in chapter 2 is applied to a large range of datasets sourced from the literature, including various mammal and lizard genera, as well as cichlid fish. The goal is to use the new model to characterise the prevalence of clade-wide character displacement patterns. Overall, this model is not well supported across animal clades, although there is some evidence of competition among cichlids, widely regarded as being composed of adaptive radiations.

Chapter 4: Using phylogenetic comparative models to infer differences in trait evolution between hummingbird clades

This chapter is a case study of 279 species of American hummingbirds. Their trait distributions are compared between several major subclades, and a range of trait evolution models are fitted, including the new model introduced in Chapter 2. Different parts of the hummingbird phylogeny turn out to have evolved quite differently, in terms of rates, interactions, and phylogenetic signal. Environment is shown to affect evolutionary mode; this is consistent with previous community based studies. These results are used to make inferences about the different processes that shaped the hummingbirds' evolution, and to demonstrate the importance of phylogenetic scale when studying their traits.

Chapter 5: Diversification rates and lineage trait densities

This chapter is an examination of diversification models, focusing on diversity dependence models that are premised on the idea of a finite set of available niches being filled up over time. A novel 'trait-density dependence' model is introduced, where diversification rates are coupled to trait evolution, so that lineages with more unusual traits have greater opportunity to diversify. Three models, including this new one, are compared. Each model makes subtly different assumptions about how niches are distributed and how they fill up, and I examine the differences between the phylogenies and trait distributions that these models generate. Although conceptually similar, the three models result in quite different observations.

Conclusion

The results from the four main chapters are brought together, and general conclusions are drawn. Limitations of this work are discussed, and directions for future research are suggested.

References

This is the bibliography for the whole thesis.

Appendices

All appendices are provided at the end of the thesis, numbered such that, for example, appendix 3B is the second appendix relating to chapter 3 of the thesis. Tables and figures are then numbered within each appendix as e.g. 3B.1.

Chapter 2: Trait evolution in adaptive radiations: modelling and measuring interspecific competition on phylogenies

This chapter was published in American Naturalist as:

> Clarke, Magnus, Gavin H. Thomas, Robert P. Freckleton. 2017. Trait Evolution in Adaptive Radiations: Modeling and Measuring Interspecific Competition on Phylogenies. The American Naturalist 189:121-137.

Abstract

The incorporation of ecological processes into models of trait evolution is important for understanding past drivers of evolutionary change. Species interactions have long been thought to be key drivers of trait evolution. However, models for comparative data that account for interactions between species are lacking. One of the challenges is that such models are intractable and difficult to express analytically. Here we present phylogenetic models of trait evolution that include interspecific competition amongst chosen species. Competition is modelled as a tendency of sympatric species to evolve towards difference from one another, producing trait overdispersion and high phylogenetic signal. The model predicts elevated trait variance across species and a slowdown in evolutionary rate both across the clade and within each branch. The model also predicts a reduction in correlation between otherwise correlated traits. We use an Approximate Bayesian Computation (ABC) approach to estimate model parameters. We find reasonable power to detect competition in sufficiently large (20+ species) trees, compared with Brownian trait evolution and with OU and Early-Burst models. We apply the model to examine the evolution of bill morphology of Darwin's finches, and find evidence that competition affects the evolution of bill length.

Introduction

There is an increasing drive to combine evolutionary and ecological perspectives in order to fully capture the long-term dynamics of ecological communities (Johnson and Stinchcombe 2007, Cavender-Bares et al. 2009, Schoener 2011, Pennell and Harmon 2013, Hadfield et al. 2014, Price et al. 2014, Pigot and Etienne 2015). This has led to insights into the roles of ecological processes such as competitive exclusion and character displacement in shaping trait evolution and today's distributions of traits (Webb et al. 2002, Kraft et al. 2007, Emerson and Gillespie 2008, Vamosi et al. 2009). However, linking such patterns in data to underlying processes is difficult, since any given pattern could be the outcome of several processes (Dayan and Simberloff 2005, Mayfield and Levine 2010).

Evidence that competition has shaped trait evolution has been generated using two main approaches. The first is the observation of character displacement, i.e. a tendency for species with overlapping ranges to exhibit increased phenotypic differences where they coexist (Schluter and McPhail 1992, Dayan and Simberloff 2005, Pfennig and Pfennig 2010, Stuart and Losos 2013). The second line of evidence for competitive effects makes use of a phylogeny to measure the distribution of species trait values relative to a null model (Webb et al. 2002, Freckleton and Harvey 2006, Vamosi et al. 2009). This is especially useful for adaptive radiations, where typically several similar species are confined to the same geographical area. Distributions that are more even than expected by chance (Webb et al. 2002, Dayan and Simberloff 2005, Davies et al. 2012) are taken as evidence that past competition caused species to seek unique ecological niches.

Convergent evolution of sets of species in separate clades has also been observed and interpreted as evidence of interspecific competition (Moen and Wiens 2009). With close niche packing interspecific competition can reduce evolutionary rates, even with a changing environment (De Mazencourt et al. 2008). Phylogenetic comparative models of adaptive radiations have slowing rates of phenotypic evolution, implicitly assuming that competition for finite niche space is an underpinning mechanism (e.g. the 'early burst' model, Harmon 2010a). Despite much study, however, the importance of competition remains uncertain (Gillespie et al. 2001, Cavender-Bares et al. 2009) and, importantly, direct tests for evidence of past competition in phylogenetic data are lacking.

One approach could be to explicitly model the evolution of traits in systems of species in which competition is occurring. In general, evolutionary models use some combination of continuous random change through time (Felsenstein 1973), possibly with changes of rate (Garland et al. 1992, Pagel 1997, Freckleton et al. 2002, Blomberg et al. 2003, Eastman et al. 2011, Revell et al.

2012, Thomas and Freckleton 2012), discrete random changes at speciation events (Ingram 2010), or shifts in shared adaptive optima (Uyeda and Harmon 2014). However phylogenetic models of trait evolution are ecologically neutral, since they are stochastic models that depend on the independent evolution of each species to be statistically well behaved (Pennell and Harmon 2013). Processes such as competition between species are typically not accounted for. In previous models species interactions have been assumed to generate phenomenological outcomes. For example models may assume rate slowdowns associated with competition among lineages either implicitly by modelling through time (Harmon et al. 2010a) or explicitly (Mahler et al. 2010). Several models include clade-wide non-random effects (Hansen 1997, Price 1997, Harvey and Rambaut 2000, Freckleton and Harvey 2006, Bartoszek et al. 2012), reflecting the interaction of species with their environments, but none of these models permits trait values to be influenced by interspecific interactions.

Phylogenetic datasets have been simulated with competitive interactions (Freckleton et al. 2003, Nuismer and Harmon 2015). However, direct parameterisation with data is difficult because of the complexity of accounting for interspecific interactions. Niche-filling models of trait evolution on trees (Price 1997, Harvey and Rambaut 2000, Freckleton and Harvey 2006) are models of adaptive radiations where new species move discretely to the nearest of a random set of points (niches) in trait-space. Simulations under these models show that such ecological processes affect inferences drawn from comparative analyses. The most important conclusion from the analysis of such models is that methods based on Brownian motion are inappropriate or even misleading, when applied to traits evolving in such systems. However, the problem of modelling such data has never been satisfactorily resolved (Freckleton and Harvey 2006), largely because of the complexity of statistically describing the traits of a set of interacting species.

In terms of fitting complex models to data one potential approach is Approximate Bayesian Computation (ABC; see Beaumont 2010). This provides a simple method for generating posterior probabilities of models, provided we can simulate them. It is therefore well suited to fitting complex models, where it is not possible to compute a likelihood function. In this way species interactions could be incorporated into evolutionary models, thus permitting better inference of the ecology underlying trait evolution. Processes such as character displacement and mutualism affect trait values, and ABC is a means of comparing models that explicitly include these processes. ABC has been explored for simple phylogenetic trait evolutionary models (Kutsukake and Innan 2013) including birth-death models (Slater et al. 2012), but its flexibility has not previously been used for including complex effects like interspecific interactions.

In this paper we introduce a new model for the evolution of interacting species within phylogenetic data. The objective is to create a model that includes character-displacement interactions and makes realistic predictions, but that also may be fitted to real data. We do not assume that all species are interacting with each other, but instead allow competitive interactions to be turned on and off at different times and for different pairs of species. We use this flexibility in two ways. Firstly, interactions can be turned on for sympatric species and turned off for allopatric species, so that the model can encompass large groups of species with variously overlapping or non-overlapping ranges. Secondly, we can add a delay, after each new species arises, before it begins interacting with the other species. This corresponds to a scenario where speciation occurs in allopatry and a subsequent range expansion brings the new species back into contact after a significant time interval. These two uses for 'optional sympatry' can of course be combined.

There are two main diffusion models of trait evolution: Brownian motion (BM, Felsenstein 1973, 1985) and the Ornstein-Uhlenbeck (OU) model (Hansen 1997). The OU model is based on BM, with the addition of an overall 'optimum' trait value to which all lineages are attracted. The strength of attraction adds a further parameter. However, we chose to build our model on the BM model. The BM model is a very simple, neutral model, while the OU model produces, in some respects, results that are the opposite of those produced by our competition model. The ability to reject BM in favour of either OU or competition may therefore be a useful aid in interpreting data.

The model predictions are compared with those of the Brownian motion and rate-change models for sympatric clades. We then outline how ABC methods may be used to detect competition effects and we show that the model is readily fitted to data. Finally, we apply these methods to a simple case study, the adaptive radiation of Galapagos finches.

Methods

The model

Under the Brownian motion (BM) model of trait evolution (Felsenstein 1973), for each species i, a trait value x_i evolves according to the differential equation,

$$dx_i = \sigma dW_i(t) \qquad \text{Equation 2.1}$$

where $W(t)$ is the integral of the continuous white noise function, such that over a finite time it has a normal probability distribution: $W(t) \sim N(0, \Delta t)$. The BM model has two free parameters, the evolutionary rate σ and the root trait value $x_{root}(0)$. The expected variance between tips is proportional to the branch length separating them.

Many models for comparative data are based on modifying this model by adding additional parameters (Pagel 1997, Blomberg et al. 2003, Eastman et al. 2011, Revell et al. 2012, Thomas and Freckleton 2012, Boucher and Démery 2016). For example, exponential rate change models replace σ with $g^{-t}\sigma$ (Blomberg et al. 2003), so that $g < 1$ corresponds to an increasing rate of trait evolution and $g > 1$ means a decreasing rate: an 'evolutionary slowdown'. Notably these models all assume that the evolutionary trajectories of species traits are independent and assume that there are no interactions between different species.

Our competition model is based on the BM model, with a term added to account for interspecific interactions. Competition is modelled such that species with similar trait values tend to evolve away from each other, while species with dissimilar trait values have little influence on each other. To achieve this we assume a flat fitness surface for trait values in the absence of other species. In effect we assume that if the trait in question has a one-to-one correspondence with some resource, e.g. body size and prey size, then the distribution of resources is flat. We assume that a species with a given trait value has a corresponding 'ideal' resource but also uses up other resources such that the distribution of resource types used is normal and centred on the 'ideal' resource type. Therefore a Gaussian curve is associated with each species along a single trait axis representing this resource use and consequently its amount of influence on other species as a function of the difference in trait value between the two species (Doebeli and Dieckmann 2003, Pigolotti et al. 2010, Liemar et al. 2013, Liemar et al. 2008).

The repulsion between two species in trait space is assumed to be proportional to the overlap of each of their associated curves. For the evolution of a single trait *x* in a species i, we get a deterministic term, scaled by a parameter a:

$$dx_i = a \sum_j S_{ij}(t) overlap\left(x_i, x_j\right) dt + \sigma \, dW_i(t) \qquad \text{Equation 2.2}$$

The area of overlap of two normal curves is related to the cumulative normal function Φ of minus the distance (in standard deviations) between them, such that the overlap is equal to $2\,\Phi(-distance\,/\,2)$ (Inman and Bradley 1989). The overlap of two curves very far from each other is $2\Phi(-\infty) = 0$, whereas the overlap of two curves with the same centre is $2\Phi(0) = 1$. **S** is a 'sympatry matrix', with elements S_{ij} each equal to either 0 or 1. If species i and j inhabit the same geographical area and have the opportunity to interact then we can set $S_{ij} = 1$; otherwise, $S_{ij} = 0$. The matrix **S** can be a function of time, permitting any given pair of species to spend time effectively in allopatry and in sympatry.

The relative intensity of competition is measured by the value of the competition parameter a relative to the Brownian rate parameter σ. Ideally we would have chosen to make the kernel width an additional parameter of the model. However in practical terms it would not have been possible to distinguish this effect from that of the competition parameter a. Appendix A shows that to a linear approximation the effects of the two are the same, and so they are likely to be statistically indistinguishable.

The instantaneous change of the trait value x_i of species i is given by Equation 2.3:

$$dx_i \;=\; a\sum_j S_{ij}(t)e_{ij}\;\; 2\Phi(-|x_i - x_j|)\,dt \;+\; \sigma\,dW_i(t) \qquad \text{Equation 2.3}$$

Each x_j is a vector in trait space; the index j denotes species. The right-hand-side has two terms: the first is a deterministic 'competition' term, which pushes apart species that are nearby in trait-space. e_{ij} is the unit vector pointing from species j to species i in trait-space. Thus, $a\,e_{ij}\,\Phi(-|x_i - x_j|)$ is a vector in trait-space pointing from species i to species j, proportional to the model parameter a and depending on the closeness in trait-space of species i and j. What distinguishes this model from previous ones is that in the competition term of the equation all traits are linked: the evolution of two species away from each other in trait space depends on the Euclidian distance between them, as well as their distances to all other species. We largely concentrate here on single resources and traits. However, more generally a multivariate normal curve in trait-space may be associated with each species in order to model interactions along several resource axes.

In both the BM and competition models, trait variance increases without bound as time progresses. In reality there are limits that will be driven by ecology or by developmental and physiological constraints. We therefore adapted the model by imposing hard limits on trait-space, such that species can evolve up to a chosen extreme value but no further. This model was simulated alongside the limitless model and hence we obtain a new model with constrained trait/niche space. We assume that the limits are symmetric about the root trait value and equal to the most extreme value L.

Simulation framework

In diffusion models such as BM or OU, trait evolution may be modelled readily and quickly because species are assumed to be independent. However our model requires that we simulate evolution over interacting branches, which makes it far more computationally demanding. The approach we used was to take a discrete approximation to the continuous differential process in Equation 2.3. A large number of time steps were used, and trait values computed sequentially for each step using the discrete approximation to Equation 2.3,

$$\Delta x_i \ = \ a \sum_j S_{ij}(t) e_{ij} \ \ 2\Phi(-|x_i - x_j|)\, \Delta t \ + \ \sigma\, W(\Delta t), \qquad \text{Equation 2.4}$$

where $W(\Delta t) \sim N(0, \Delta t)$, i.e. a normal distribution with mean 0 and variance Δt. For every time step, pairwise trait value differences must be computed between all species, giving $|x_i - x_j|$ for all species i and j. Then the trait values are all updated according to Equation 2.4. For a tree with 20 tips, divided into 1000 time steps, a typical modern consumer processor takes about 0.008 seconds to complete a simulation. The simulation of traits on a single clade is single threaded.

Simulations were performed on random ultrametric trees generated under a Yule process (TESS, Hoehna 2013) with between 2 and 100 tips. These tree sizes are large enough to demonstrate the effects of competition, and provide a range over which our power to detect competition varies from low to high. Because the competition model is designed for sympatric, interacting sets of species having undergone adaptive radiation, it is unlikely that numbers of species will be very large. For example, the average clade size in Harmon et al.'s (2010a, 2010b) survey of adaptive radiations was 30 tips. We first estimated some summary statistics for a single 10-tip tree: the fitted Brownian rate σ, phylogenetic signal K (Blomberg et al. 2003) and the rate change parameters κ and δ (Pagel 1997). We did this for a range of competition strengths from $a = 0$ to $a = 5$, recording trait values through time to generate figure 2.1. We also repeated this with competition and limits, and with the two main subclades set to be allopatric (non interacting) with one another.

To compare tip trait value distributions, we generated a single 40-tip random tree, and collected trait data from many simulations under each of three evolutionary regimes: BM, competition and competition with limits. The distributions for each regime included all the corresponding datasets. To assess the accumulation of trait variance, we did the same again, but with two different values of the competition parameter a.

Tree size has an effect on the variance of tip trait values in our competition model, unlike BM. To assess this we generated trees that had numbers of tips from 5 to 100. For each possible number of tips within this interval, we generated 50 random trees. For each tree we then simulated a trait dataset under BM, competition and competition with limits. The mean tip trait value variance was plotted as a function of the number of tips.

Competition also affects correlations between pairs of co-evolving traits. We simulated traits whose Brownian evolution was strongly correlated (80% weighting for a single Brownian process and a 20% weighting for additional independent Brownian processes; under pure BM this results in a correlation of tip trait values of approximately $r^2 = 0.8$) under a wide range of competition strengths ($a = 0$ to 5). We did this for a single small tree (10 tips) and a single larger tree (40 tips). We also repeated these steps for four fixed values of a (0, 1, 4, 8) while varying the BM dependence, i.e. the amount of shared change between traits, and observed the resultant correlation between tip trait values.

The sympatry matrix $S(t)$ in Equations 2.2 and 2.3 allows us to control which lineages interact and when. This means we can set up simulations where lineages start to have competitive interactions a fixed time after they start. This corresponds to the scenario of speciation in allopatry and subsequent range expansion so that all lineages eventually come back into contact. We simulated data with a range of delay periods: 0.1, 0.2, 0.5 and 1.0 times the mean time between speciation events, and generated power estimates for detecting competition in these cases. We also generated an example data set with a delay period of 0.5 times the mean time between speciation events, on a single 16 tip tree for comparison with other models (figure 2.1E). It is important to note that we do not fit our sympatry matrix to data. The sympatry or allopatry of any given set of species is assumed to be known from independent data on species' ranges. A delay between speciation and competitive interactions can also be added, but again this is pre-set by the user, not fitted.

The reasons for basing the competition model on BM rather than the OU model are detailed in the introduction. However, for the sake of comparison we did run some simulations of an 'OU + competition' model. The model is detailed in Appendix B, and an example simulation is shown in figure 2.1D.

Finally, we assessed the effect of competition on phylogenetic signal. A single 100-tip tree was generated, and trait values simulated for a range of competition parameter strengths ($a = 0$ to 5). Blomberg's K (Blomberg et al. 2003) was then computed using the Kcalc function in the R package picante (Kemble et al. 2010).

Model comparisons and likelihoods

We fitted the model to data using Approximate Bayesian Computation (ABC) (reviewed in Beaumont 2010, Csilléry et al. 2010, Hartig et al. 2011). ABC can be used for comparing the probabilities of datasets under different models when these probabilities are difficult to compute directly. This is because the only requirement to perform ABC is that we can simulate new datasets using the model. The dataset probabilities are approximated by simulating a large number of datasets, and 'accepting' only those simulations that are very similar to the observed dataset. This similarity can be judged either from the data values themselves, or using summary statistics. The proportion of simulations that are accepted is then assumed to be proportional to the dataset probability. When the model contains continuous parameters, we sample across these parameters and obtain an approximate probability density for the observed data under any point in a range of parameter values. This can be used to estimate the likelihood curves of fitted models. ABC relies on the likelihood being a fairly smooth function of the model parameters (Hartig et al. 2011). Since our model progressively changes the trait distribution and phylogenetic signal smoothly (with respect to the competition strength a), we expect the likelihood to be a smooth function.

To apply ABC to phylogenies (e.g. Beaumont 2010) we sample the parameters of the evolutionary model randomly many times from a prior parameter distribution. Here we choose the prior distribution to be uniform, with the model necessitating a hard limit at zero for both the Brownian rate and the competition strength. For each set of parameters, trait datasets are then simulated for the known phylogeny. Summary statistics are generated for the simulated data, and only those simulations for which the summary statistics are within a small value ε of the observed data's summary statistics are accepted. Thus, for observed data D and tolerance ε, we accept some simulated data D' if

$$\rho(D', D) \ < \ \varepsilon, \qquad \text{Equation 2.5}$$

where ρ is the discrepancy, or distance in summary statistic space, between D' and D. In practice we chose the tolerance ε based on the size of posterior sample that we wanted to obtain, so we might simulate a million datasets and choose ε such that we accept the best 500 simulations. By plotting acceptance rate against parameter values, we get an estimated likelihood surface.

To compare simulated and observed datasets, it is necessary to compare summary statistics. We chose to use three summary statistics: the mean and the variance of the differences between each species and its closest neighbour in trait space, and the overall phylogenetic signal as measured by Blomberg's K (Blomberg et al. 2003). The rationale for using these three statistics was to capture the overall amount of evolution, the overdispersion of trait values, and the phylogenetic structuring of the trait values. There is no well-established procedure for choosing summary statistics for ABC. High sufficiency is needed to compare models, but the ABC method quickly loses accuracy and stability with large numbers of summary statistics (Csilléry et al. 2010). Our summary statistics were chosen on a pragmatic basis, since they capture the important aspects of the model's behaviour, namely increased divergence between sibling species, and an even overall distribution of traits across the phylogeny.

We chose to compare the competition model with the BM model using maximum approximated likelihood, because the BM model is embedded in the competition model. The null and alternative ABC acceptance rates A give an estimate of the likelihood L(H|D) of the observed dataset under the various model parameters. This assumes a smooth probability distribution with an overall peak and a separate peak on the $a = 0$ (Brownian motion) plane. Since small changes to the model parameters should create small changes in average trait distributions, the assumption of smoothness should be met, provided many simulations are used. The two peaks define the null and alternative model parameters. The log-likelihood ratio statistic for the comparison of two models H_0 and H_1, when there is no prior difference in model likelihood expectation, is then given by:

$$-2\,log\frac{L(H_1|D)}{L(H_0|D)} = -2\,log\frac{A_1}{A_0}. \qquad \text{Equation 2.6}$$

When the models are nested, this test statistic approximates a χ-squared distribution given certain assumptions: large samples and normally distributed parameters. However, these assumptions may be significantly violated by phylogenetic methods (Freckleton 2009). For instance, in the BM model σ is bounded at zero, and in our competition model a is also bounded at zero. To correctly interpret the test statistic, therefore, the null distribution of the log-likelihood ratio test statistic was assessed with a parametric bootstrap.

The bootstrap was undertaken by performing the model comparison analysis on datasets generated under BM, to create a null distribution of likelihood ratios. Then if, for example, we want to know the likelihood ratio corresponding to a p-value of 0.05, we simply look at the 95th percentile of the null distribution. The resulting Type I error rate is therefore chosen by design: if a significant likelihood ratio is one that corresponds to a p-value of 5%, then the Type I error rate is 5%. To estimate typical significance thresholds, we performed this procedure for random trees using 1000 random datasets.

The power to reject BM in favour of the competition model was assessed by using random ultrametric Yule trees (20, 40, 60 and 80 tips). The bootstrap process was performed to determine the significance threshold for that tree. Then, for a given value of the competition parameter a, we simulated a large number of datasets and determined the likelihood ratio (between the BM and competition models) for each one. The proportion of these datasets which showed significant support for competition effects defines the power of the model for that value of a. We repeated this process for a range of competition strengths from $a = 0$ to $a = 5$. This range covers evolution from a Brownian process with no interspecific interaction ($a = 0$) to a largely deterministic regime with high phylogenetic structuring of trait values ($a = 5$). We also assessed power as a function of the delay period before new species start interacting.

To evaluate the simulated data produced by the competition model, other comparative models were fitted to the data: the Brownian model itself, the δ-model, which measures temporal acceleration and deceleration in rate, the κ-model, which measures the degree to which evolution is speciational rather than gradual (Pagel 1997), and K, a measure of phylogenetic signal (Blomberg et al. 2003). Parameter estimates were generated using the R packages geiger (Harmon et al. 2008) and picante (Kembel et al. 2010).

We compared two further common models to the competition model. Trait distributions and phylogenetic signal were computed for a single 25-tip tree, using the competition model as well as the OU model (Hansen 1997) and the Early-Burst (EB) model (Blomberg et al. 2003, Harmon et al. 2010a). Traits were simulated for the OU and EB models using the function rTraitCont in the R package ape (Paradis et al. 2004). For the EB simulations, however, the tree is first time-transformed according to

$$t \rightarrow \frac{e^{gt}-1}{g} \qquad \text{Equation 2.7}$$

where t is time, and g parameterises the model. When g is negative, this represents a slowdown over time. As g approaches 0, the transformation approaches the identity, and we recover the Brownian model. We applied this transformation using the function transfBranchLengths in the R package phylolm (Ho and Ané 2014).

The simulations were written in C++. Scripts for using these datasets for likelihood estimation were written in R (R Core Team, 2015), using ape (Paradis et al. 2004) and TESS (Hoehna 2013) for tree generation (code available at https://github.com/mcshef/treecomp/).

Case study: Darwin's finches

The phylogeny of Galapagos finches was taken from Lamichhaney et al. (2015). We used a dataset from Harmon (2010b), using individual data originally from Lack (1947), with species values for five traits: wing length, tarsus length, bill length (culmen), bill depth and bill width (gonys). We computed likelihood ratios for each trait individually, as well as for combined pairs of beak traits. The final datasets used are deposited in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.3sk15.

After simulating data on the phylogeny to determine likelihood cutoffs for rejecting BM, a likelihood comparison between the competition model and the nested BM model was run for each of the five traits separately. We performed the tests twice, once including and once excluding the phylogenetic summary statistic K, to judge the importance of signal in favouring the competition model.

Results

Example of clade evolution under the competition model

Illustrative examples of evolution under the competition model are shown in figure 2.1. Estimates of phylogenetic signal K and phenotypic rate change transformation parameters δ and κ (Pagel 1997) are also shown for the simulated data. The evolution of each species is tracked through time from left to right. It can be seen that competition increases phylogenetic signal while giving the appearance of an overall slowdown in rates of trait evolution.

Competition tends to increase the overall variance in traits amongst the species in a phylogeny, as is clear from the increase in range and variation of traits moving from figure 2.1A to figure 2.1C. This is because species experiencing competition from other species are more likely to evolve extreme trait values to become more different and escape competition.

As the strength of competition is increased, the differences between species become more clearly defined, with them occupying distinct positions in niche space. There are fewer intersections of traits' evolutionary paths over time between species, and the phylogenetic signal K exceeds the neutral BM prediction of $K = 1$. Competition thus increases phylogenetic signal above that expected under the BM model, while presenting the appearance of a considerable tree-wide evolutionary slowdown. This slowdown can be seen in figure 2.1 where the rate change model δ is fitted to datasets simulated under competitive effects. This means that a species' trait values map more directly onto its position in the tree. For sympatric clades, there is thus a prediction of traits being more phylogenetically conserved than under BM.

Figure 2.1. Examples of a single trait evolving under the competition model with different strengths of the competition effect parameter a. A single random 16-tip tree is used throughout. The BM parameter $\sigma = 1$ is used for all the simulations. The parameter values listed by each plot are the estimates obtained using the data shown; these are subject to substantial stochastic variation between simulations. A, When $a = 0$ we recover the Brownian Motion model. B, Simulated using $a = 1$. C, Simulated using $a = 3$. D, Competition plus OU model, with the OU parameter α set to 5. Strong competition and OU effects cancel each other out in some respects. E, Delayed competition, where each new lineage starts to experience competitive interactions a fixed time after it appears. This corresponds to a scenario of allopatric speciation and subsequent range expansion. F, Here the two main subclades of the tree do not interact with each other, though there is strong competition within the subclades. This corresponds to a mixture of sympatric and allopatric species, e.g. a radiation on two different islands.

Estimates for commonly used branch transformation parameter κ from these datasets are also shown in figure 2.1. κ measures the rate change along branches, and overall measures the degree to which change is speciational (Pagel 1997). A transformation parameter δ models the overall changes in evolutionary rate across the tree, with lower values corresponding to evolutionary slowdowns (Pagel 1997). We find that the δ parameter diminishes very rapidly as competition is increased. It should be noted here that the δ parameter is biased such that the expectation values for BM are greater than 1 (Freckleton et al. 2002). Nevertheless, this reflects an apparent slowdown of evolutionary rate, which becomes more pronounced as the value of a increases. Species competing for unoccupied niche-space thus evolve more rapidly early on in their development, when they are more similar to one another and the effects of competition are stronger, as one would expect in an adaptive radiation (Yoder et al. 2010).

Figure 2.1D shows a simulation using a combined competition and OU model (see Appendix 2B for details). It can be seen that the repulsive effect of competition and the central attraction of the OU process have somewhat opposite effects, with competition tending to increase phylogenetic signal and rates of trait evolution, while the OU process reduces signal and the overall rate of trait change. This is part of the reason why we chose to base the competition model on the BM rather than the OU process. Even in real cases where both processes are present, the dominant process can likely be judged by the relative support for the competition and OU models considered separately. It seems unlikely that both components could be accurately fitted simultaneously, since their effects are so confounded.

Finally, we considered two cases of partial sympatry. In the first (figure 2.1E), new lineages come into interaction with pre-existing lineages only after a fixed delay period. This reduces phylogenetic signal relative to complete sympatry. The distribution of tip trait values is less affected, unless a substantial number of lineages remain in allopatry at the tree tips, i.e. at the present time. The second case is analogous to two parallel adaptive radiations on isolated islands: all the descendants of one ancestral species interact with each other, but not with the descendants of the second ancestral species. This removes the raised phylogenetic signal effect and most of the change in trait distribution. However, we can still detect competition provided we have a priori knowledge of the sympatry or allopatry of each species pair.

Trait distributions across tree tips

The distribution of trait values of the phylogeny tips is flattened in the competition model compared with BM models, which predict normal distributions for large trees (figure 2.2A). This outcome is expected when competition shapes trait values (Davies et al. 2012). The impact of competition on trait distributions is even more pronounced where hard limits are placed on the available range of trait values (also figure 2.2A).

In addition to creating a more even trait distribution, competition increases the overall amount of trait divergence, given equal BM rates (figure 2.2B). This is consistent with the expectation that equivalent species sets should be more diverged in sympatry than in allopatry when there is competition (Schluter 2000). From a biological perspective there is thus a prediction that competition leads to a wider range of morphological variation in a clade, reflecting the increased tendency towards extreme traits when there is lots of competition.

Effects of tree size

We used trees normalised to the same total length, regardless of the number of tips. Given this normalisation, under BM and rate-change models, the variance of tip trait values shows no change with increasing the number of tips (in agreement with Ricklefs 2004). In the competition model larger trees have greater variance, since a greater number of species are 'pushing' each other away; this is shown in figure 2.2C. This relationship seems to be approximately linear for the unbounded competition model. When hard limits are imposed, the variance reaches a maximum corresponding to the positions of the extremes.

Effects of competition on correlated traits and phylogenetic signal

For pairs of traits, in which the evolutionary changes in trait values are correlated, the correlation between the traits decays rapidly with increasing competition strength. This is even more pronounced when there are limits on extreme trait values. Figure 2.2D and figure 2.2E show how the correlation decays. By de-correlating traits, competition forces the trait space to be occupied more evenly.

Phylogenetic signal is increased by competition because species tend to remain adjacent in trait space to their close relatives (figure 2.2F), and so their trait values are unlikely to 'cross over' with time. Plots of traits through time therefore become more defined and tree-like. This can be seen for example in the sample simulations of figure 2.1. Correlation between traits has little effect on the phylogenetic signal exhibited by the individual traits under either the BM model or the competition model. Limits reduce the phylogenetic signal, since there is less trait-space for distantly related species to diverge. Indeed, without competition driving the signal up (i.e. when $a = 0$), the model with limits predicts reduced signal compared with the BM model, with K estimated to be less than 1.

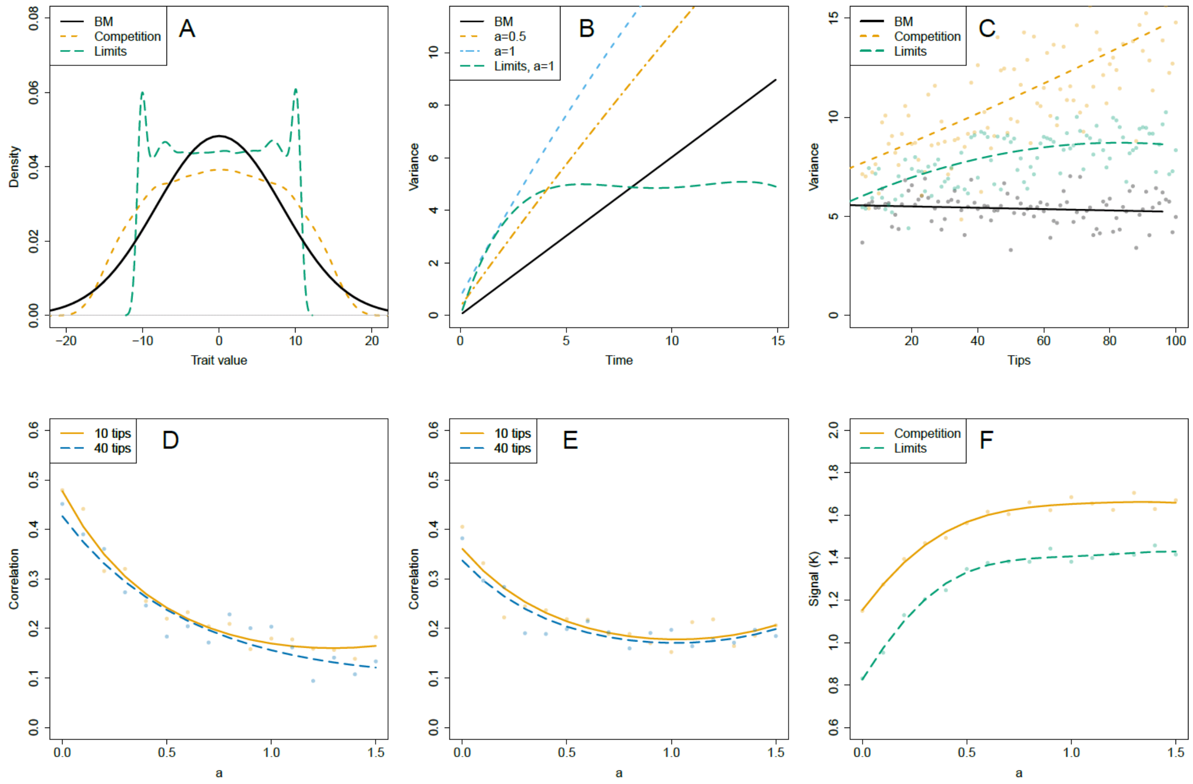

Figure 2.2. A and B show effects of interspecific competition on trait value distributions across a clade. The BM model results in a normal distribution, whose variance increases linearly with time. The competition model predicts a flattened distribution, whose variance increases with time, initially faster than for BM, but slowly slowing to the BM rate. When hard limits are placed on trait space, the competitive distribution is further flattened, with probability peaks at the limits, since competition will tend to push species to those limits. C, Effect of tree size on variance of trait values of an evolving clade. Under BM the tree size has no effect, but with competition across the clade, a more numerous clade results in a greater amount of trait variation in that clade. D, The correlation between two traits across the species in a clade, as a function of the strength of interspecific competition parameter. The BM rate parameter is set to 1 throughout. E, Correlation for the model with limits. F, The signal (Blomberg's *K*) as a function of competition strength. The traits are evolving with their random (BM) evolution strongly correlated; the pressure from competition to be dissimilar acts against this natural correlation.

Price’s (1997) model of adaptive radiations makes the unique prediction that when two traits have correlated evolution, the correlation between phylogenetically independent contrasts is potentially different in form than that between the traits themselves (Price 1997, Rambaut and Harvey 2000, Freckleton and Harvey 2006). In contrast, a Brownian model predicts equal correlation for both traits and contrasts.

We compared trait and contrast correlations under the competition model presented here. Competition tends to reduce correlation between traits, as discussed above, but we set the Brownian evolution of the traits to have very high correlation (figure 2.3). We found that contrasts had higher correlations than traits. This probably reflects the fact that competition tends to have a greater effect earlier in the evolutionary history of any particular species. However, the general principle seems to hold, namely that when species interact the correlations between traits and trait changes (i.e. contrasts) are not expected to be equal.

Comparisons with alternative models

We generated trait distributions and computed phylogenetic signal for a single 25-tip tree, using the competition model as well as the OU model (Hansen 1997) and the Early-Burst (EB) model (Blomberg et al. 2003, Harmon et al. 2010a). (We in fact simulated on a range of trees, with sizes up to 100 tips, but report on a single tree here for the sake of comparison. Tree size makes no qualitative difference to the observed patterns.) The expected distributions for OU and EB trait values remain normal, although the OU model reduces the variance relative to BM. The actual trait distribution for a typically sized tree, however, was often multimodal for the EB model. In contrast, the competition model predicts a flattened, regular distribution of trait values. In this respect the competition and EB predictions differ. Phylogenetic signal, on the other hand, is raised by both the competition model and the EB model, although more strongly by competition. The OU model reduces phylogenetic signal to lower values than expected under BM, and to very low values when strong. These results are tabulated in Appendix 2C.

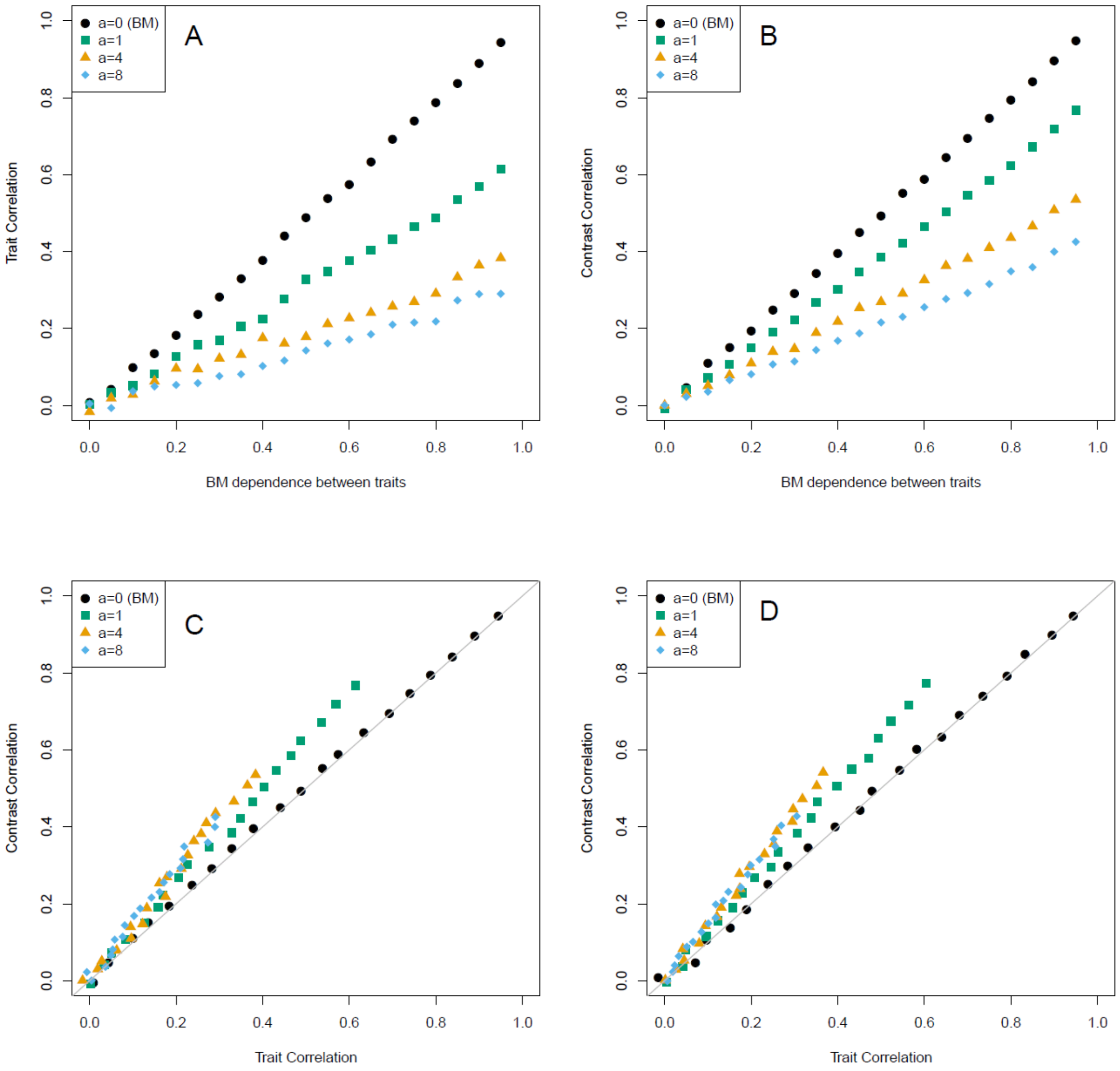

Figure 2.3. Traits and contrasts for two traits, where one trait has a dependence on the other. For each step in time, the dependent trait has an evolutionary change that depends on the change to the other trait. If the dependence is 1, then these changes are equal; if the dependence is 0.5, then the change in the dependent trait is 0.5 of the change in the other trait, and the remaining change is random. A, Trait correlation as a function of the intrinsic trait dependence. B, Contrast correlation as a function of the intrinsic trait dependence. C, Correlation between contrasts is slightly greater than correlation between traits for competitive evolution. D, Contrast and trait correlations for a model of competition with traitspace limits.

Power

The power to detect competition effects against a background of BM evolution is shown in figure 2.4A-D for trees of various sizes. We define the power as the frequency with which simulated datasets show significant support for competition effects as opposed to the (nested) BM null model. Power is greatest for large trees with high competition strength relative to BM rate. This can be interpreted as the relative contribution to overall evolutionary change of competitive effects versus other, effectively random, effects.

We also computed the power to distinguish data generated under a competition model from OU and EB models. This is shown in figure 2.4F. There is excellent power to detect even weak competition compared with these other models. This is not surprising, since a large amount of variation in trait distributions and phylogenetic signal can be generated under BM, and both the EB and OU models represent deviations from BM that differ from the deviations caused by competition. Competition raises signal and reduces the variance in differences between neighbouring tip trait values, whereas EB increases tip value difference variance, and OU reduces signal.

Figure 2.4E shows power as a function of delay. 'Delay' here means the time from a lineage originating to the time at which it starts interacting with all the pre-existing lineages, measured as a proportion of the mean time between speciation events. While this delay reduces power, substantial power to detect strong competition remains even for delay periods of similar length to the time between speciation events.

In this context a significant dataset is one for which the Type I error rate is estimated to be ≤ 0.05. This is the frequency with which data from null model simulations display model likelihood ratios that equal or exceed the ratio for the 'observed' dataset. This is determined via a parametric bootstrap.

Case study: Darwin's finches

The simulations described above demonstrate two things: first that the model we describe successfully captures behaviour that we would expect to be observed in systems of interacting species. And second, that it may be applied to data and used to infer the presence of competitive interactions. In order to use the model in a real-world example we applied the competition model to an example dataset, using trait measurements collected in Harmon et al. (2010a; originally Grant and Grant 2002, Lack 1947; repository in Harmon 2010b http://dx.doi.org/10.5061/dryad.f660p), and a recent molecular phylogeny (Lamichhaney et al. 2015). We used the Galapagos finches (*Geospiza spp.*), because they are a well-studied adaptive radiation, and ecological effects were anticipated to be of importance. The effect of character displacement on intraspecific variation among these finches is well documented (e.g. Grant and Grant 2006). Here we are looking to see whether evidence for this mechanism can be detected in the overall distribution of traits across the clade.

The parameter estimates and model likelihood ratios are shown in table 2.1. The beak traits showed greater support for competition compared with Brownian evolution than the other traits. This appears to point to an ecological effect: the competition model implies a tendency towards well-differentiated niches that don't cross, and the beak shape is an 'ecological' trait, in the sense that it corresponds strongly to feeding habits (Grant and Grant 2011). Multi-trait analyses are also shown in table 2.1, for pairs of beak traits. Competition tends to reduce correlation between traits in our model, however (see figure 2.1). Since beak measurements are likely to be correlated (due to overall size), the model may have a poor fit without adding correlated Brownian evolution as a further fitted parameter, which we have not done. The model as it stands may be better suited to Principle Component data, where correlations have already been accounted for and removed before the analysis. Figure 2.4 shows illustrative plots of simulated trait evolution using the model parameters that were estimated for the culmen length. Compared with BM, shown in figure 2.4A, the tree becomes very well defined, with strong phylogenetic signal.

One of the beak traits, culmen length, favoured the competition model when signal was not used, but less so when it was included. Brownian rate parameter estimates under the model with limits are higher than those for the non-limited competition model. This higher rate does not result in a greater total amount of evolutionary change, due to the hard limits that are reached either way. This result does, however, suggest that the niche landscape may be the limiting factor in the finches' evolution: the finch traits are capable of evolving rapidly, but their values are constrained by the combination of interactions between species and environmental limits on niche space.

It is worth noting that none of these results takes into account intraspecific variation or uncertainty in the finch phylogeny. Species interactions will be somewhat independent on different islands, as illustrated by the character displacement seen in intraspecific trait variation (Grant and Grant 2006).

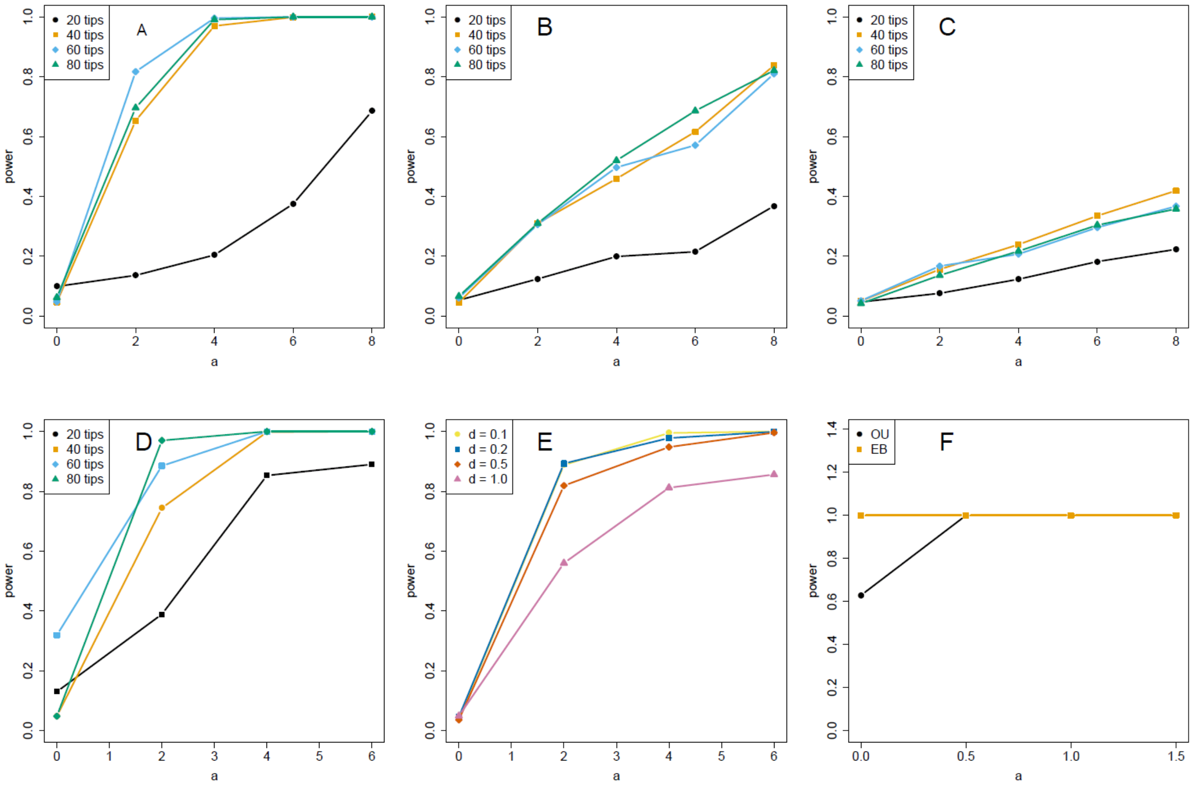

Figure 2.4. Power to detect competition versus Brownian Motion as tree size, rate parameter (σ) and intensity of competition (a) are varied. A, σ = 1. B, σ = 1.5. C, σ = 2. D, power as a function of competition strength a for data simulated under a model with imposed limits, effectively restricting trait values to half the range they would otherwise have. E, the effect of delayed interactions on power; d is the time delay before a new lineage starts interacting with pre-existing lineages, measured as a proportion of the mean time between speciation events; tree size is fixed as 60 tips. F, Power to distinguish our model from the OU and EB models for a fixed tree size of 60 tips.

Table 2.1: Traits and likelihood ratio test statistics for model comparisons for Galapagos finches.

Trait	**σ**	**a**	**LRTS**
Wing length	1.48	0.28	1.01
Tarsus length	1.64	0.48	1.04
Culmen length	1.24	0.96	4.34*
Beak depth	1.40	1.92	2.73
Gonys width	1.48	1.32	1.87
Culmen + beak depth	1.12	3.76	3.11
Culmen + gonys width	1.12	4.20	3.24

Note: the finch trait dataset is that given in Harmon et al. (2010b). The competition model is compared with the nested BM model. The competition model has one extra parameter compared with the BM model. * indicates significance based on simulated parameter distributions.

Discussion

There have been several recent approaches to integrating ecological processes into phylogenetic models of evolution (Cavender-Bares et al. 2009, Hadfield et al. 2014, Pennell and Harmon 2013). As a step in this direction, we have created a model of interspecific competition on phylogenies of coexisting species. The model provides a process-based picture of competitive evolution, linking statistical patterns directly to the underlying ecology. It generates the patterns we expect to see in situations where interspecific competition is important.

Competition and niche overlap have a complex relationship. According to one scenario, niches form a continuum, with similar fitnesses along a niche axis. Species compete for empty regions of the niche axis and evolve minimally-overlapping, evenly-spaced trait values, consistent with the ecological idea of character displacement (Grant 1972, Strong Jr et al. 1979, Dayan and Simberloff 2005). Our model accommodates this process in a phylogenetic context. However, according to a second scenario, there is an optimum position on a niche axis, and species will compete with each other to occupy it. Consequently they will evolve similar trait values tailored to that optimum (Colwell and Futuyma 1971), a form of convergent evolution. Traits that have evolved like this are likely to be poorly represented by our competition model, and we would be unlikely to find evidence supporting the model. Instead, such a process would be better modelled by a process such as an OU model (Hansen 1997, Butler and King 2004). The model we have introduced here corresponds better to a process of occupancy of distinct niches distributed across trait-space.

One prediction of the competition model is a flattened distribution of trait values among contemporary species of a single sympatric clade. Indeed, competition is often inferred from such evenness (Dayan and Simberloff 2005, Davies et al. 2012). The same pattern can however be caused by competition at the community assembly level rather than in situ trait evolution (Cavender-Bares et al. 2009, Stuart and Losos 2013), or by geographical structure in speciation and extinction (Pigot and Etienne 2015). For this reason the model presented here is best applied to complete clades, rather than local paraphyletic communities. Since complete clades are rarely entirely sympatric, we have included the ability to turn competitive interactions on and off on a pairwise basis depending on sympatry. The model can similarly include a delay between speciation events and the start of competitive interactions involving the new species, corresponding to allopatric speciation and subsequent range expansion. This is important since there is a wealth of variation between clades in species' geographic structure and opportunity to interact (Fitzpatrick et al. 2008). Strong niche-conservatism is, nevertheless, predicted for largely sympatric clades. Developing summary statistics for testing our model on data with substantial allopatry – e.g. multiple island radiations – will be an important future development.

The pattern of non-Brownian trait distributions and high phylogenetic signal is also generated by an alternative, but related, mechanism, where instead of there being a continuum of possible niches, instead they are discrete and new species arise by jumping to a nearby niche (Price 1997, Harvey and Rambaut 2000, Freckleton and Harvey 2006). In these models, niches appear at random, and existing species that are nearby in niche-space can speciate to occupy a new niche. Determining a method to distinguish this type of model from diffusion models will be a useful future development. Specifically, this approach differs from the other models discussed here in that the tree topology is not fixed but interacts with species' trait values as the clade evolves.

There are numerous speciation/extinction models for phylogenies (Nee et al. 1994, Pybus and Harvey 2000, Rabosky 2006, Freckleton et al. 2008, FitzJohn 2010), including some that are expected to correspond to clades with interspecific competition (Harmon et al 2010a, Etienne et al. 2012). Our model is concerned only with trait evolution. Trait evolution and diversification rates may be coupled in nature, however, and both may vary with factors such as interspecific competition. Building models of adaptive radiations that simultaneously predict trait evolution and diversification will be key in the future.

Most phylogenetic models of trait evolution are modifications of the random BM model. As noted above, adaptive radiations are generally consistent with a tree-wide gradual slowdown in rates of phenotypic evolution (delta-model: Pagel 1997; ACDC model: Blomberg et al. 2003).

Speciational evolution can be modelled as a gradual branch-wise slowdown (κ model, Pagel 1997), or by partitioning evolution into gradual and speciational parts (Bokma 2008, Ingram 2011). Discrete shifts in evolutionary rate can be modelled to detect, for example, adaptive radiations embedded in a larger tree (O'Meara et al. 2006, Thomas et al. 2006). Slowdowns in evolutionary rate have also been observed as a function not of time but directly of a clade's size (Mahler et al. 2010). The results for our competition model suggest that it reproduces the appearance of a strong tree-wide slowdown. During a radiation, though, competition is predicted to cause overall trait variance to increase much more rapidly. Our results for the competition model also demonstrate raised phylogenetic signal when most species are sympatric and have opportunity to interact, in agreement with similar results in Nuismer and Harmon (2015).

In all analyses we used a fixed competition kernel width. The fact that this width is not distinguishable from the competition strength itself suggests that the amount of variation possible within a single niche is not readily ascertained from a phylogeny and trait data. Measurements of intraspecific variation will be more suited to this question. In fact, the competition kernel widths could be set empirically before analysis, if data on intraspecific variation were available.

Our results for the Galapagos finches support the well-known presence of character displacement in that clade (Grant and Grant 2006), and further suggest that interspecific competition is a significant force comparable with other, effectively random, sources of evolutionary change for the Galapagos finches. For some beak measurements, the Galapagos finches exhibit the elevated phylogenetic signal predicted by the competition model, and for beak length we find strong support for the model.

As phylogenetic methods continue to be used to infer evolutionary processes, it will be important to include specific ecological mechanisms (Vamosi et al. 2009). Competition for ecologically distinct roles is often implicitly or explicitly assumed in adaptive radiations, but its prevalence and importance remain uncertain (Schluter 2000, Stuart and Losos 2013). We have developed an explicit model of competition on phylogenies, to detect competitive effects in sympatric adaptive radiations and to enable measurement of competition strength. The predictions of this model may help in understanding the roles ecological processes play in shaping trait evolution.

Acknowledgements

Our thanks to Natalie Cooper and two anonymous referees for extensive useful feedback on the manuscript. GHT was supported by the Natural Environment Research Council (grant number NE/G012938/1).

Chapter 3: Testing for clade-wide patterns of character displacement in animal clades

Abstract

Character displacement is an increase in the difference in trait value between two species in order to reduce competition for resources. Evidence for character displacement is usually found in sympatric sibling species pairs; however, there is no universal agreement on its prevalence. Here, I use a previously introduced model of trait evolution to search for character displacement patterns across entire clades. A range of animal phylogenies and trait datasets are sourced from the literature, and several evolutionary models, including our character displacement model, are fitted to them. Overall, I find that clade-wide character displacement patterns are not common. I also used range data to compute a measure of overall sympatry for each clade, and found that this was uncorrelated with evidence for character displacement. The importance of the geographic and phylogenetic scale of analysis is discussed in this context.

Introduction

Stochastic models of trait evolution on a phylogeny were originally proposed with the aim of reconstructing the tree from trait values (Felsenstein 1973). Today, the same methods are typically used as a way of correcting for phylogenetic correlations between species when looking for correlations between traits (Felsenstein 1985, Freckleton et al. 2002). However, these models have become useful for inferring the rate and mode of evolution through the history of a clade. Phylogenetic methods are often used to test hypotheses such as discrete shifts in rates of trait evolution (O'Meara 2006, Thomas and Freckleton 2012, Revell et al. 2012) or effects of character states on patterns of speciation and extinction (Maddison et al. 2007, Freckleton et al. 2008, Fitzjohn et al. 2010).

Some of these hypotheses correspond to mechanisms studied in ecology, for example evolutionary slowdowns as clades become more diverse (Rabosky and Lovette 2008). However, other mechanisms have not been widely incorporated into stochastic models of trait evolution, despite being important topics in ecology. One example is character displacement. It has been inferred from phylogenetic studies of community structure (Webb et al. 2002, Kraft et al. 2007), but not widely included as a mechanism in models of trait evolution (but see Davies et al. 2012 for a test for it).

Character displacement can be defined as the shift in character state, and therefore niche, of two sympatric populations to minimise interpopulational competition for resources (Pfennig and Pfennig 2010). There is no consensus on the prevalence of character displacement patterns arising across clades due to interspecific competition. Competition is expected to generate divergence between sibling species (Cavender-Bares et al. 2009, Yoder et al. 2010), and this is most commonly observed as character displacement: elevated differences between species where they coexist (Schluter and McPhail 1992). Surveys of character displacement disagree about the evidence for its importance and frequency, concluding either that mechanisms causing character displacement are widespread and important (Schluter 2000, Dayan & Simberloff 2005), or that we yet have little evidence for them (Stuart and Losos 2013, Tobias et al. 2014).

Evidence for character displacement is traditionally sought in sibling species pairs, by comparing coexisting pairs and geographically separated pairs (Schluter and McPhail 1992, Dayan and Simberloff 2005, Pfennig and Pfennig 2010, Stuart & Losos 2013). Tobias et al. (2014) found no evidence of raised divergence rates due to sympatry in a study on ovenbirds. They instead found that apparent differences were due to the tendency for coexisting lineages to be older than noncoexisting lineages. Alternative methods detect character displacement across whole clades, using a phylogeny to measure the distribution of species trait values relative to a null model (Webb et al. 2002, Freckleton and Harvey 2006, Vamosi et al. 2009). For example, Davies et al. (2012) looked at trait distribution evenness across a phylogeny's tips and compared it to a null expectation under Brownian motion, inferring character displacement in a small mammal community. This approach makes greater use of the available trait data, by examining the allocation of evolutionary change across the tree, as well as overall rates.

Here I use a method for looking at character displacement patterns in adaptive radiation data, using a phylogeny and tip trait data (described in detail in Chapter 2, also Clarke et al. 2017). According to this method, this data is used to fit a model of trait evolution where the usual assumption of independent evolution in each branch of the phylogeny is relaxed. Through the history of a clade, the competition model permits species to interact and repel one another in trait-space. In this way, past competition and character displacement may be inferred from a tree and contemporary trait data. Approximate Bayesian Computation (ABC) is a simulation-based method that can be used to generate likelihood estimates when an analytic expression for the model likelihood is difficult to obtain. I use ABC to compare the predictions of Brownian evolution and evolution with character displacement. I look for effects on phylogenetic signal and tip trait value distribution shape, and hence obtain an estimate of the strength of the competition effect and a likelihood ratio for the two models.

The competition model incorporates interspecific ecological interactions into a stochastic evolutionary model, and in doing so makes some strong assumptions. In particular, it assumes that the clade has evolved with opportunity for ecological interactions between all the species that constitute it. The proxy I use for this is sympatry: the co-occurrence of two species in the same geographical area. I compute a measure of mean range overlap for each of the clades under study here, for use as an approximation of the degree to which the species exist in sympatry. Of the clades under study, some clades fulfil the assumption of sympatry well, while others are widely geographically distributed, occupying different environments and having many species that have no opportunities to interact.

Here, I use the methods from Chapters 1 and 2 to search for evidence of character displacement across a wide range of animal clades, mostly divided at the genus level. This includes 75 mammal genera and 17 lizard genera. I also analyse a cichlid clade, both as a whole and by looking at its three main subclades. The trait data and phylogenies for these clades were sourced from the literature. There were two broad objectives: (1) to assess the overall prevalence of clade-wide character displacement patterns in animals; (2) to determine whether coexistence makes character displacement more likely among groups of related species.

Methods

Modelling interspecific competition

Completely random, gradual evolution is modelled with Brownian motion (BM, Felsenstein 1985, Freckleton et al. 2003). The differential change in trait x for a given species is given in terms of white noise dW by equation 3.1 below. The expectation value of the change in a finite time t is therefore 0, with variance $\sigma^2 t$.

$$dx = \sigma dW. \qquad \text{Equation 3.1}$$

The competition model is a modification of the BM model, with an additional term representing interspecific interactions. Character displacement patterns are generated by the tendency of species with similar trait values to evolve away from each other in trait-space. We effectively assume that a species with a particular trait value has a corresponding normal distribution of resource types, and competition occurs when these distributions overlap. Therefore a Gaussian curve is associated with each species. Along a single trait axis, this normal distribution curve represents the species' resource use and consequently the amount of influence on other species with trait values that are similar (Doebeli and Dieckmann 2003, Pigolotti et al. 2010, Leimar et al. 2013, Leimar et al. 2008).

For the competition model we therefore add a deterministic term to the BM stochastic differential equation such that each species i in existence at a given moment in time interacts with all the other species present, as per Equation 3.2. This is described in more detail in Clarke et al. (2017) (See also chapter 2).

$$dx_i \ = \ a\sigma \sum_j e_{ij} \ 2\Phi(-|x_i - x_j|) \, dt \ + \ \sigma \, dW. \qquad \text{Equation 3.2}$$

Here Φ is the cumulative normal function, e_{ij} is a unit vector in trait-space pointing from species j to species i, and a is a parameter corresponding to the strength of interspecific competition. In the limit as $a \rightarrow 0$, we recover BM.

I also assess a variant of this model, where hard limits are imposed on the available trait values. The evolution of each tree branch follows equation 3.2, but if a trait value reaches a limit at any time, then it is prevented from further evolution in that direction, i.e. towards more extreme trait values. These limits are not estimated by the model, but are set, fixed assumptions.

To fit the model to datasets, Approximate Bayesian Computation (ABC, reviewed in Beaumont 2010, Csilléry et al. 2010, Hartig et al. 2011) is used, a method for using simulations to generate likelihood estimates when an analytic expression for the model likelihood is difficult to obtain. Model parameters are sampled from a prior distribution, and many simulations performed. Summary statistics are computed for each simulation, and compared with those for the observed data. The model parameters of the simulations whose summary statistics are within a small distance of those of the observed data then form the posterior parameter distribution. I then obtain a likelihood ratio estimate for the BM and competition models. For this approach I use three summary statistics: the mean and the variance of the differences between each species and its closest neighbour in trait space, and the overall phylogenetic signal as measured by Blomberg's K (Blomberg et al. 2003). These summary statistics contain information on the overall amount of evolution, the overdispersion of trait values, and the phylogenetic covariance structure of the trait values. For each trait in each clade, one million simulations were performed to generate the likelihood estimates.

Quantifying range overlaps

To assess the effect of sympatry on the results, I used a measure of overall clade sympatry. This was computed as the mean of the range overlaps of each pair of extant species in the clade. The measure of range overlap used was the area of overlap divided by the area of the smaller of the two ranges (Barraclough & Vogler 2000). Therefore, a sympatry value of 1 indicates that all the species have the same range, while a value of 0 indicates no geographical overlap between any ranges. To assess the relationship between present day sympatry and evidence of past competition, I computed the correlation of sympatry level and the competition strength parameter *a*.

Mammal and lizard ranges were obtained as polygons from the IUCN Red List of Threatened Species online database (IUCN 2014), and analysed using the R packages *sp* (Pebesma 2015), *rgdal (*Bivand et al. 2014), and *maptools* (Bivand and Lewin-Koh 2013).

Data source and preparation

A wide range of animal clades are considered here, each representing a single genus expanded to be monophyletic. Among these are 77 mammal genera and 10 lizard genera, using data from Jones et al. (2009), Fritz et al. (2009) and Bergmann and Irschick (2012). For the mammal dataset, I also sampled random clades for similar analyses, allowing them to be performed on larger clades and on clades with a wider range of mean sympatry than the genera.

Mammal bodysize data was extracted from the PanTHERIA database (Jones et al. 2009). Data was available for 5417 species, covering most known mammal species. The phylogeny is from Fritz et al. (2009), a modified version of the tree in Bininda-Edmonds et al. (2007). I subdivided the tree by genus, in each case taking the most recent common ancestor of that genus and expanding the analysis to encompass all of its descendants. I then discarded clades with <15 species, since I expected to have little power to detect evidence of competition for very small clades. This left 77 clades, with a total of 2164 species. The average size of these clades was 28.5 species; the largest had 73 species. A further test performed using the mammal data was to divide the complete clade into random subclades of between 50 and 250 tips, and compute the mean sympatry and estimate model parameters for each subclade. These trees were larger than typical genera, giving more power to compare models and a greater range of available timescales to study.

A database of squamate morphometrics and corresponding phylogeny has been collected in Bergmann and Irschick (2012). I used this tree and extracted lizard bodysize data from this database. Again, I subdivided the tree and data by genera with >15 species, to leave 10 clades suitable for analysis. Of the 1375 squamate species in the full tree, this procedure left 455 for analysis.

These analyses are also performed on a set of Neotropical cichlids and 3 major clades therein, following Lopez-Fernandez et al. (2013), who regard the entire clade as an ancient adaptive radiation. A morphometric dataset for Neotropical cichlids is available in Lopez-Fernandez et al. (2013), with a corresponding phylogeny in Lopez-Fernandez et al. (2010). I used these resources to run the analysis on the complete clade, as well as on three major subclades: Cichlasomatini, Geophagini and Heroini, estimated to have around 70, 250 and 150 species respectively (Lopez-Fernandez et al. 2010). The phylogeny has 160 tips, thus representing around one third of the extant species. I chose not to use African cichlids from e.g. Lake Victoria, since they may have considerable geneflow and uncertainty in the molecular phylogenies (Wagner et al. 2013), and the character displacement model assumes an accurate phylogeny and rapid speciation.

Results

Mammals

The complete results for mammal genera are shown in table 3.1. I studied 77 mammal genera. Of these, 3 had significantly better support ($P < 0.05$) for competitive effects than for Brownian motion; these three genera were *Pseudantechinus*, *Artibeus* and *Psudomys*: marsupials, bats and rodents respectively. However, significant results for 3 clades is no more than might be expected by chance. The distribution of likelihood ratio statistics for mammal genera is shown in figure 3.1A, along with a similar distribution for data simulated under a Brownian motion model on the same trees. Both distributions are roughly normal, the mammal data in fact showing a slightly lower average likelihood ratio statistic for the presence of character displacement effects than expected under a random BM model. Results for the competition model with trait-space limits are given in Appendix A; they are similar to the results without limits, and there are no clades for which limits are supported over competition with no limits.

The distribution of genus-wide sympatry across mammal genera is approximately normal, with few clades having mean sympatry greater than 0.8. There was no discernible relationship between overall clade sympatry and competition strength, as seen in figure 3.1B (estimated effect -1.15, P=0.45). The three clades with significant support for competition effects had sympatry 0.28, 0.62 and 0.52 respectively.

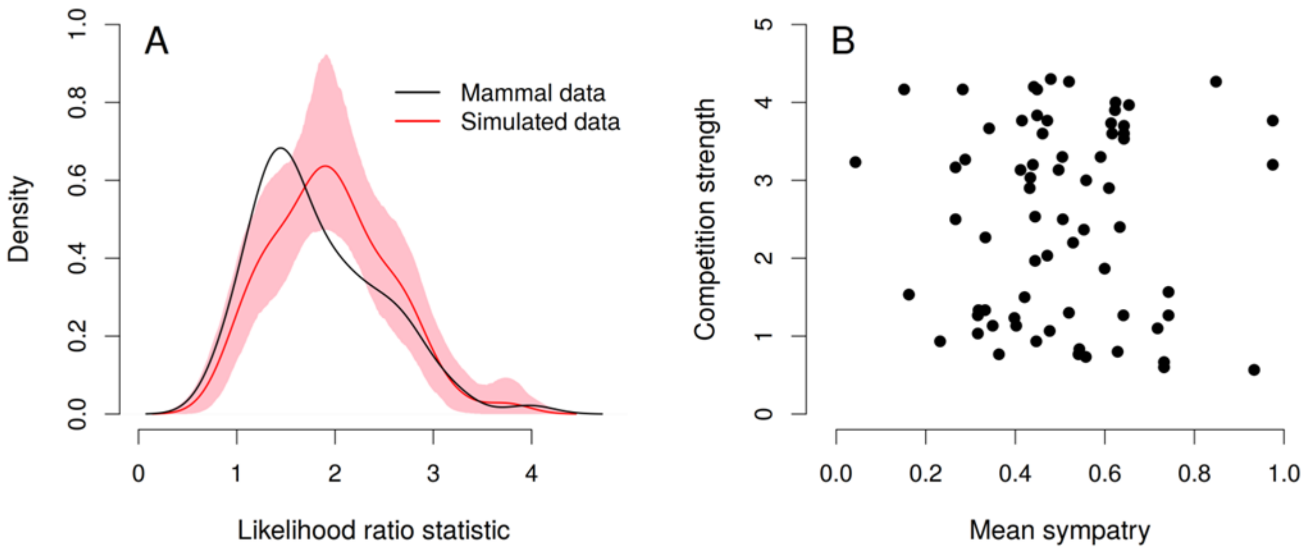

Figure 3.1: A, Distribution of likelihood ratio statistics from each mammal genus. The red line shows the corresponding results for BM datasets simulated on the same trees. A 95% confidence interval for the simulated results is shaded, based on resampling of the density distribution. B, Estimated competition strength and average clade sympatry for mammals.

Table 3.1: Competition model estimates and mean sympatry for mammal genera. Sigma and a are the parameters of the model, corresponding to the rate of trait value diffusion and the strength of lineage-lineage interactions respectively. The likelihood ratio is for this model as compared to a null Brownian motion model. Sympatry is computed as the average amount of pairwise range overlap across all the species pairs in each clade.

Clade	Sigma	a	Likelihood ratio	Sympatry
Stenella	2.266667	3.833333	2.607327	NA
Trinomys	0.133333	0.666667	1.153897	0.732139
Trachypithecus	1.533333	2.5	2.648321	0.266445
Talpa	2.266667	1.266667	1.444433	0.3163
Thomasomys	4.6	3.7	1.914608	0.642574
Sundasciurus	1.866667	3.766667	1.928703	0.974973
Tupaia	3.233333	1.266667	1.472369	0.742125
Abrothrix	2.233333	3.6	1.990246	0.460886

Acomys	2.033333	2.266667	1.693553	0.333138
Aotus	2.366667	1.533333	1.421866	0.162371
Sylvilagus	1.866667	4.166667	2.098439	0.448654
Akodon	3.466667	1.133333	1.119375	0.401949
Dipodillus	2.166667	0.766667	1.170109	NA
Cratogeomys	2.4	4.166667	1.508262	0.151676
Echimys	2.566667	2.533333	1.425115	NA
Cercopithecus	1.333333	2.9	1.572199	0.609196
Chaetodipus	2.466667	1.333333	1.257967	0.317848
Galago	2.1	3.766667	1.907503	0.414853
Dipodomys	0.7	4.2	2.848566	0.441388
Cephalophus	0.6	2.4	2.309238	0.633412
Genetta	2.933333	3.3	1.59248	0.590515
Ctenomys	2.533333	0.766667	1.032995	0.363437
Gerbillus	1.233333	3	1.999667	0.558156
Cryptotis	1.9	3.133333	1.66393	0.496652
Chaerephon	2.266667	1.333333	0.930371	0.332483
Artibeus*	0.4	4	3.150412	0.623748
Lagenorhynchus	0.4	2	1.342643	NA
Makalata	2.966667	3.3	1.636448	0.505029
Lophuromys	2.1	1.1	1.364932	0.717734
Hylopetes	1.633333	0.8	1.336264	0.628476
Mogera	1.633333	3.233333	2.310501	0.042962
Mesoplodon	2.266667	4.1	2.745201	NA
Marmosa	2.1	0.833333	0.983286	0.54299
Monodelphis	3.1	3.133333	1.598352	0.411785
Meriones	1.233333	1.3	1.274966	0.519683
Mops	2.366667	0.933333	1.698779	0.446762
Melomys	1.7	0.733333	1.334248	0.557428
Maxomys	1.766667	0.566667	0.88797	0.933469
Tamias	2.566667	2.533333	1.438865	0.444142
Kerivoula	1.233333	1.266667	2.082395	0.641405
Macaca	2.133333	1.5	1.668266	0.420748
Lasiurus	0.833333	3.533333	2.11846	0.642597

Otomys	2.133333	3.666667	1.537459	0.341533
Paramelomys	3.066667	3.2	1.448469	0.439175
Phyllomys	1	3.9	2.666692	0.622628
Pseudantechinus*	0.933333	4.166667	3.259516	0.282593
Oecomys	0.8	3.966667	2.984403	0.653803
Neotoma	2.8	1.233333	1.311317	0.397958
Ochotona	2.633333	2.033333	1.289976	0.471339
Oryzomys	4.666667	1.133333	2.290858	0.349467
Proechimys	4.266667	4.3	1.815649	0.479153
Phalanger	1.733333	3.733333	2.331557	0.613949
Murina	1.266667	3.766667	2.321878	0.471471
Pseudomys*	0.866667	4.266667	3.983908	0.519907
Niviventer	2.7	2.2	1.474011	0.528945
Nycteris	1.533333	3.6	1.098765	0.616444
Reithrodontomys	1.9	2.366667	1.260778	0.553157
Rhipidomys	3	0.766667	0.907891	0.541145
Saguinus	3.5	0.933333	1.591801	0.231991
Sminthopsis	2.566667	1.066667	1.405091	0.476755
Stenella	1.4	4.033333	2.076977	NA
Oligoryzomys	2.8	2.9	1.379409	0.432121
Trinomys	0.133333	0.6	0.811055	0.732139
Trachypithecus	1.533333	3.166667	1.754339	0.266445
Sundasciurus	1.5	3.2	1.937148	0.974973
Thomasomys	4.533333	3.6	2.839939	0.642574
Talpa	2.233333	1.033333	1.233688	0.3163
Tupaia	2.966667	1.566667	1.375111	0.742125
Sylvilagus	1.533333	3.833333	1.747396	0.448654
Spermophilus	0.9	3.266667	2.471116	0.288026
Myodes	0.533333	1.866667	2.40561	0.599524
Rousettus	2.8	4.266667	2.586106	0.848278
Tamias	2.666667	1.966667	2.0984	0.444142
Peromyscus	1.133333	2.5	2.554674	0.506081
Sciurus	1.5	3.033333	2.723497	0.433537

Note: Star denotes significant support for competition versus Brownian motion.

Lizards

I also considered 10 lizard genera. Of these, only 1 had evidence of competitive effects: *Egernia* skinks. Data was available for 10 species in this clade. However, the appearance of considerable character displacement-like patterns in *Egernia*, which does not consist of species all sympatric with one another, may rather be the result of many one-off speciations into new environments (Heatwole and Taylor 1987). The distribution of likelihood ratio test statistics for the competition model had a slightly lower mean than expected under Brownian motion (see figure 3.2A). *Egernia* was a noticeable outlier in this distribution. The average sympatry across the 10 clades was 0.65. As was the case for mammals, I didn't find any relationship between competition strength a and sympatry, as seen below in figure 3.2B (effect estimate = -1.01, $P = 0.45$). The complete results for lizards are tabulated in table 3.2, with additional results for the competition+limits model given in Appendix 3A.

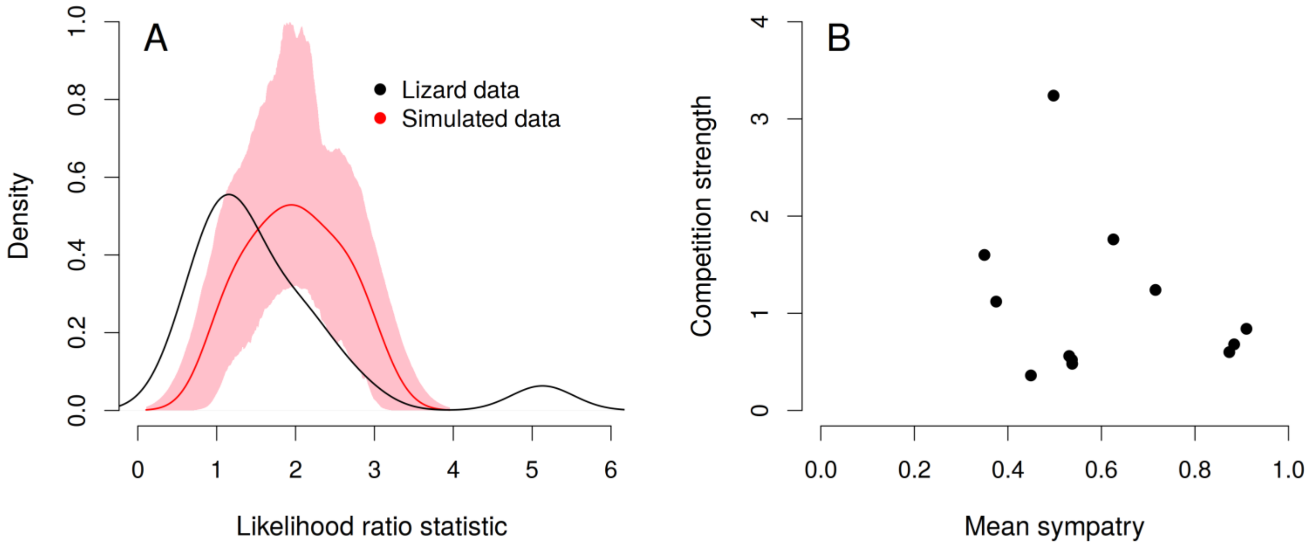

Figure 3.2: A, Distribution of likelihood ratio statistics from each lizard genus. The red line shows the corresponding results for BM datasets simulated on the same trees. A 95% confidence interval for the simulated results is shaded, based on resampling of the density distribution. B, Estimated competition strength and average clade sympatry for lizards.

Table 3.2: Competition model estimates and mean sympatry for lizard genera. Sigma and a are the parameters of the model, corresponding to the rate of trait value diffusion and the strength of lineage-lineage interactions respectively. The likelihood ratio is for this model as compared to a null Brownian motion model. Sympatry is computed as the average amount of pairwise range overlap across all the species pairs in each clade.

Clade	Sigma	a	Likelihood ratio	Sympatry
Amphiglossus	1.72	3.12	1.528665	0.672849
Acanthodactylus	0.76	0.52	0.713051	0.536915
Ameiva	1.44	1.24	1.139811	0.715558
Ctenophorus	0.8	0.68	1.129544	0.883716
Chalcides	3.28	1.6	1.468827	0.349796
Egernia*	0.32	5.44	5.13022	NA
Mabuya	1.24	1.76	1.605079	NA
Trachylepis	0.24	4.4	2.028256	1.144031
Lerista	2.48	1.12	1.244732	0.374923
Liolaemus	0.84	0.56	1.023522	0.531074
Cnemidophorus	0.12	1.4	2.744123	NA
Draco	0.28	1.76	2.186701	0.625564
Varanus	0.72	3.24	2.121207	0.497624
Sphenomorphus	1.12	0.84	1.46932	0.910035
Emoia	1.24	0.6	1.029718	0.873417
Sceloporus	0.72	0.36	0.634488	0.449219
Anolis	1.36	0.48	0.908225	0.537528

Note: Star denotes significant support for competition versus Brownian motion.

Cichlids

Finally, I fitted the competition model to the complete Neotropical cichlid dataset and three major clades therein, following the subsets used in Lopez-Fernandez et al. (2013). The results are shown in table 3.3. I found significant support for character displacement effects in bodysize across the Neotropical cichlids as a whole, and stronger support particularly in the Cichlasomatini subclade. In contrast, however, I found no such evidence for character displacement in jaw morphology. Results for the competition+limits model are given in Appendix 3A.

Table 3.3: Results for Neotropical cichlids. Sigma and a are the parameters of the model, corresponding to the rate of trait value diffusion and the strength of lineage-lineage interactions respectively. The likelihood ratio is for this model as compared to a null Brownian motion model.

Clade	Trait	Sigma	a	likelihood ratio
Cichlid*	bodysize	1.07	4.85	3.17
	gape	3.17	0.70	0.93
Cichlasomatini*	bodysize	0.79	5.55	4.20
	gape	1.31	3.31	2.75
Geophagini	bodysize	1.68	0.84	1.20
	gape	1.77	1.07	1.57
Heroini	bodysize	2.71	1.26	1.58
	gape	1.17	4.06	1.98

Notes: Star denotes significant support for competition versus Brownian motion (not corrected for multiple tests).

Discussion

I fitted a model of clade-wide character displacement to a range of mammal and lizard clades, as well as a group of Neotropical cichlids. The results for mammals and lizards showed little sign of clade-wide character displacement, and simulations showed that the distribution of competition model likelihoods was similar to that expected under a Brownian model. Indeed, the support for character displacement was overall slightly lower than expected under a Brownian model, perhaps reflecting convergent evolution among species in larger clades, where the same niche is occupied by related but distinct species in different environments.

The results for cichlid fishes showed much greater signs of competition and character displacement effects. As well as significant results for the Cichlasomatini subclade, I found evidence of interspecific competition in the cichlid clade as a whole, suggesting that they may form a single adaptive radiation, and that their trait values may be strongly influenced by historical resource competition between species and adaptations to minimise this competition. These results were found only for bodysize, however jaw morphology showed no sign of character displacement effects for the whole cichlid clade. This is consistent with the findings of Lopez-Fernandez et al. (2013), who inferred convergent evolution of head shape in the three subclades, with opposite effect to competition as I have modelled it. I did not find similar support for the competition model with limits on trait space, suggesting the notion that the adaptive radiation has not exhausted the range of viable trait values.

Broad searches for character displacement in the literature have focused on individual clades where character displacement was claimed. Tests for character displacement therefore tend to be performed where it is expected to be found. The evidence is usually judged according to a set of 6 criteria first compiled by Schluter and McPhail (1992). These criteria are: (1) differences in phenotype are due to genetics rather than, for instance, phenotypic plasticity; (2) apparent patterns of character displacement are not due to chance; (3) the species have evolved in situ, so that the apparent patterns of character displacement are not due to species sorting; (4) the trait under study is linked to resource use; (5) where species in sympatry and allopatry are compared, the different environments are similar or can be controlled for; (6) independent evidence exists that individuals with similar phenotypes compete for resources. Despite these agreed criteria for evidence, different conclusions have been drawn about the amount of evidence we have for widespread character displacement, ranging from a lot (Schluter 2000, Dayan & Simberloff 2005) to not very much (Stuart & Losos 2013, Tobias et al., 2014). These differences rest largely on the interpretation of cases that satisfy most but not all of the six criteria, a category that includes most of the individual studies in the literature. In either case, however, the evidence under review consists mostly of studies performed on clades that are promising a priori candidates for character displacement patterns.

Individual cases with strong support for character displacement (satisfying all six criteria) include, for example, Darwin's finches (Grant and Grant 2006) and 3-spine sticklebacks (Schluter 2000). These two examples are quite different: Darwin's finches form a clade of around 15 species adapted to different niches on a small set of islands, while the sticklebacks have diverged into two ecologically distinct forms repeatedly and separately in several different lakes. However, in both cases the species undergoing character displacement are known to be largely sympatric and to have evolved *in situ* with plenty of opportunity for interspecific interactions. In this way they meet assumptions of models of evolution with clade-wide character displacement in a way that larger-scale datasets cannot. Indeed, I have previously found significant support for interspecific competition effects on clade-wide models of beak size evolution in Darwin's finches (Chapter 2).

I also compared model fits with estimates of average levels of range overlap across each clade, using range overlap as a proxy for sympatry. There was no evidence of a relationship between character displacement and sympatry using this method. One possible explanation for this result is that the simple method of using large-scale range overlaps to judge sympatry may not capture the actual opportunities species have had for interacting over evolutionary timescales (Connell 1980). This is especially likely since species ranges are known to change over time (Sexton et al. 2009). A second plausible explanation is that character displacement is simply not common enough in the datasets to establish any relationships with other variables; this would be consistent with other studies such as Tobias et al. (2014), who found that sympatric species pairs diverge at the same rate as allopatric pairs, for traits such as bird beak shape, across many species. Similarly, Monroe (2012) used range overlap as a proxy for the opportunity for competition, as I have done here, and found that divergence rates among mammal species were unaffected by range overlaps.

From the lack of evidence for competition effects in mammal and lizard clades, I can infer that character displacement is not one of the main mechanisms shaping patterns of character states across large clades. Of course, these large clades are not confined to single geographical areas, but overall many of them still have extensive sympatry. However, the model of competitive evolution has high power to detect competition when it is the predominant force shaping species' relative trait values, so if competitive exclusion were widespread among mammals and lizards we would expect to detect signs of it in the clades with greater overall sympatry. Further, we would expect to see some relationship between sympatry levels and character displacement patterns. Mechanisms other than character displacement are therefore likely the main driving forces behind evolution in most clades. Competitive exclusion may still affect small subclades of a few species, as can be seen in *Anolis* lizards (Losos et al. 1998, Mahler et al. 2010) where a few particular niche forms developed several times over in separate locations; across the whole *Anolis* clade, there is little evidence of character exclusion, and overall sympatry is low. This also does not mean that character displacement is not a primary factor in specific cases such as adaptive radiations of the Darwin's finches type.

It will be interesting in further work to contrast the results given here for large clades with studies focussed on smaller clades that are promising candidates for character displacement. An example, following the Darwin's finches example, would be further study of island adaptive radiations with data for all extant species in the clade using the same model I used here. The model of clade-wide character displacement would give useful insights into the evolutionary mechanisms that shaped such clades, and provide a potential contrast to the types of data analysed here.

Future models will likely benefit from addressing the interactions between geographical dynamics (e.g. Losos and Glor 2003, Pearman et al. 2008, Freckleton and Jetz 2009), speciation and extinction rates (modelling links to trait values in e.g. Paradis 2005, Freckleton et al. 2008, Fitzjohn 2010, Magnuson-Ford and Otto 2012), as well as intra- and inter-specific evolutionary interactions like those investigated here. Prior knowledge of the adaptive landscape can also be included in future models (Arnold et al. 2001, Aguilée et al. 2013).

Chapter 4: Using phylogenetic comparative models to infer differences in trait evolution between hummingbird clades

Abstract

Hummingbirds provide a useful application for evolutionary models, as they have strong phylogenetic structure, often with large numbers of species coexisting in sympatry. Here, data on 279 species of American hummingbirds is analysed. Their trait distributions are compared between seven major subclades, and a range of trait evolution models are fitted, including the new model introduced in Chapter 2. I use these results to make inferences about the different processes that shaped the hummingbirds' evolution. Overall, there is strong phylogenetic signal across the tree for both body size and beak length, but little evidence of character displacement. There is evidence of considerable variation in evolutionary history for different traits and clades, and at different geographical and phylogenetic scales.

Introduction

Phylogenetic trait models

Phylogenetic comparative models can be used to make inferences about the evolutionary history of a group of species (Felsenstein 1985, Freckleton et al. 2002, Pagel and Harvey 2002). Natural selection has a directional or stabilising effect on trait values, and in turn has many possible sources, many of them stemming from interactions between the animal and its environment. By looking at the distribution of trait across today's species, inferences can be made about these evolutionary pressures and processes (Webb et al. 2002, Kraft et al. 2007, Emerson and Gillespie 2008, Vamosi et al. 2009).

Adaptive radiations are consistent with a tree-wide gradual slowdown in rates of phenotypic evolution (Pagel 1997, Blomberg et al. 2003). Character displacement, on the other hand, is a tendency for species with overlapping ranges to exhibit increased phenotypic differences where they coexist, due to competition for niches (Schluter and McPhail 1992, Dayan and Simberloff 2005, Pfennig and Pfennig 2010, Stuart and Losos 2013). For whole clades, this outcome can also be inferred from the distribution of species trait values: an observed distribution of traits can be compared to that of a null model given a phylogeny for the clade (Webb et al. 2002, Freckleton and Harvey 2006, Vamosi et al. 2009). This is especially useful for adaptive radiations, where often a number of similar species are confined to the same geographical area. Distributions that are more even than expected by chance (Webb et al. 2002, Dayan and Simberloff 2005, Davies et al. 2012) are taken as evidence that past competition caused species to seek unique ecological niches.

This chapter uses the trait evolution model developed in chapter 2. This model explicitly includes these competitive effects, so that the strength of competitive effects can be quantified for a given clade (see chapter 2, also Clarke et al. 2017). To fit this competition model to datasets, Approximate Bayesian Computation (ABC, reviewed in Beaumont 2010, Csilléry et al. 2010, Hartig et al. 2011) is used, a method for using simulations to generate likelihood estimates when an analytic expression for the model likelihood is difficult to obtain. Model parameters are sampled from a prior distribution, and many simulations performed. Summary statistics are computed for each simulation, and compared with those for the observed data.

Hummingbirds

Hummingbirds provide a useful application for evolutionary models, as they have significant phylogenetic structure, often with large numbers of species coexisting in sympatry (Graham et al. 2009). Hummingbirds are a clade of approximately 338 species, composing the family Trochilidae. These are distributed across many communities in the Americas, with the greatest numbers around the Andes in central and South America. Hummingbirds provide a useful application for evolutionary models, as they have significant phylogenetic community structure, often with large numbers of species coexisting in sympatry (Graham et al. 2009). Recent molecular phylogenies put the clade age at around 22 million years (McGuire et al. 2014). Their morphology is thought to be closely shaped by available niches in any given community (Feinsinger and Colwell 1978). There are nine main hummingbird clades: topazes, patagona, hermits, mangoes, brilliants, coquettes, gems, bees and emeralds. The brilliants and coquettes originated in the area where the Andes now stand. The hermits are thought to have originated in the lowlands to the east, and most of the other clades in the lowlands to the west. The bees and mountain gems are the only clades thought to have radiated in North America.

The different hummingbird clades are thought to have been formed by a combination of radiation and colonisation, with emeralds and hermits radiating in the Amazonian lowlands and later colonising the Andes, and mangoes, coquettes and brilliants doing the reverse. These histories were inferred in Parra et al. (2011), by looking at how under- or over-represented each of these clades is, in many communities in and around the Andes. Parra et al. (2011)

hypothesise that species from over-represented clades are the products of in situ radiations, while species from under-represented clades have colonised from other geographical areas. Bees are thought to have radiated in Northern dry lowlands.

The highest diversification rates are seen in the bees. The clade consisting of bees and mountain gems is thought to be unique among the hummingbirds in having radiated in North America; some emeralds have since colonised North America, but the common ancestor of the bees and gems was the first invader, and most of the extant species remain there (McGuire et al. 2014). The bees could have acquired their increase in diversification rate from a recent, rapid radiation. However, a similar pattern could have arisen from a continuous rapid turnover of species throughout the history of the clade, i.e. high and approximately constant speciation and extinction rates. Analysis of the phylogenetic structure of the bees' trait values may provide some insight into this question.

The mechanisms shaping hummingbird evolution have previously been studied from a community composition perspective (McGuire et al. 2007, Graham et al. 2009, 2012, Parra et al. 2011, Lessard et al. 2016). Warm, wet lowland communities in Ecuador were found to be phylogenetically overdispersed, i.e. composed of more distantly related species than expected by chance, whereas the more "challenging" environments of the cooler highlands have communities that are phylogenetically clustered (Graham et al. 2012). This suggests that environmental filtering is a key mechanism shaping community composition, with competitive exclusion playing a greater role in milder environments.

Applying phylogenetic trait models to hummingbird clades

Here, I take a clade-based rather than community-based approach. The competition model described in chapter 2 is used to assess how much evidence there is of character displacement among hummingbirds as a whole. I also fit a range of other phylogenetic comparative models, and investigate how much variation exists in evolutionary mode among the different hummingbird clades. I consider whether evolutionary history differs consistently between highland and lowland groups of species.

The analysis is divided between the 7 main hummingbird clades, so that conclusions can be drawn about the differences between their evolutionary histories. The idea behind this is to distinguish different evolutionary process acting on traits in different clades; since they diversify according to different processes, and frequently occupy largely different environments, it makes sense to fit trait evolution models to individual clades rather than the whole hummingbird tree. There is an evolutionary mechanism that is related to, but distinct from, competitive exclusion: species may evolve to become more distinct via character displacement after entering the community. This alternative is not explicitly available to community building models, but can be included in phylogenetic trait evolution models.

I fit hummingbird data to a range of stochastic evolutionary models, including random diffusion, rate change models, and the character displacement model. Comparing the results of these 7 main clades with results for the whole of the hummingbirds may also help clarify the levels at which processes (such as evolutionary constraints and character displacement) act. Traits such as mass, wingspan, beak length have been found to have high phylogenetic signal (Graham et al. 2012) for the whole hummingbird tree; I investigate how they vary at lower taxonomic levels. I compare these results with previous community studies of hummingbirds, and draw conclusions about their evolutionary history.

Methods

The phylogeny and trait data used here were collated by Lessard et al. (2015; http://dx.doi.org/10.5061/dryad; phylogeny from McGuire et al. 2014). The phylogeny is shown in figure 4.1. This dataset includes 279 extant species, with body mass and culmen length values for each. These measurements were averages across all individual measurements available in the literature. Both males and females are included in the data. There are 338 known species of hummingbird, so this dataset is around 80% complete. I fitted evolutionary models to each of the 7 main hummingbird clades: hermits, mangoes, brilliants, coquettes, gems, bees and emeralds. Five species are omitted from these analyses, as they are members of other clades that are too small to be useful with these methods (these species are *Topaza pyra*, *Topaza pella*, *Florisuga fusca*, *Florisuga mellivora*, and *Patagona gigas*).

To fit Brownian Motion (BM) models and transformations thereof, I used the R packages ape (Paradis 2005), picante (Kemble et al. 2010) and geiger (Harmon et al. 2008). In the BM model, the expectation of variance in trait values between tree branches accumulates linearly with time at a rate σ^2. This is also the rate at which any given lineage moves in a random direction in trait-space. This rate can be made a function of time by 'tree transformation' parameters such as κ and δ (Pagel 1997) and the early burst (EB) model (Blomberg et al. 2003, Harmon et al. 2010a). The kappa model describes gradual changes in rate of an individual branch as it ages, whereas δ models gradual clade-wide changes as the whole clade ages. An alternative, slightly more complex model than BM is the Ornstein-Uhlenbeck (OU) model. In the OU model, species undergo random BM evolution but are also drawn towards a central 'optimum' value, with a strength proportional to the distance from this value (Hansen 1997). In this way, the OU model is analogous to the concept of stabilising selection. It is worth noting that maximum likelihood fits of the OU model are known to be biased (Cooper et al. 2016).

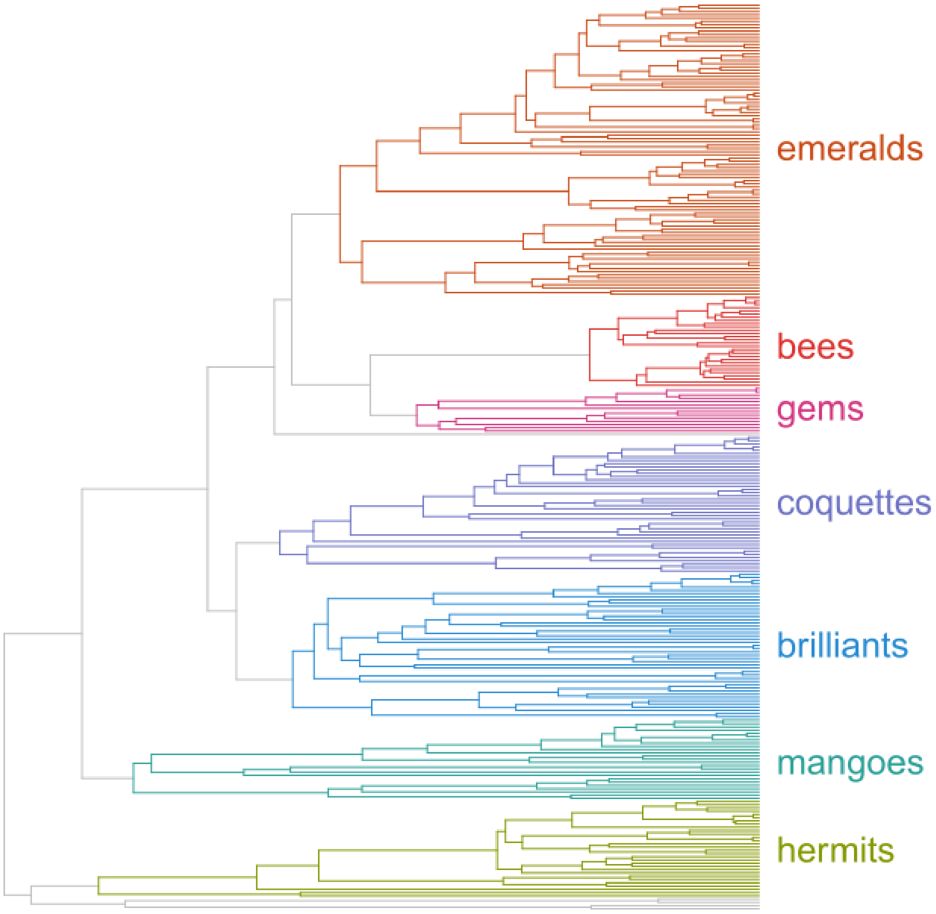

Figure 4.1: hummingbird phylogeny, with seven major clades marked.

The parameters K (Blomberg 2003) and λ (Pagel, 1997, 1999) measure the phylogenetic structure of trait data compared with that expected under a given BM model. A value of $K = 1$ indicates a perfect fit to the BM model; $K < 1$ indicates less phylogenetic signal and $K > 1$ indicates more signal than expected. In the EB model, the rate of evolution across the whole phylogeny undergoes an exponential slowdown. The differential equation for the evolution of a branch is

$$dX(t) = g^{-t} dW(t). \qquad \text{Equation 4.1}$$

This model is fitted to each hummingbird clade and trait, and P-values obtained for the null model where g is equal to 1. A value of g significantly greater than 1 implies an early burst of trait evolution, and describes an exponential slowdown over time.

I also applied a previously developed model and simulation method for detecting character-displacement patterns in phylogenetic data. This model is described by equation 4.2:

$$dX_i(t) = \sigma dW(t) + a \, \Sigma_j \, e_{ij} \, \Phi\left(-\left(\mathrm{X_i} - \mathrm{X_j}\right)\right). \qquad \text{Equation 4.2}$$

This adds a further term to equation 4.1, representing the effect of species with similar trait values on one another. Φ is the cumulative normal function, e_{ij} is a unit vector in trait-space pointing from species$j$ to species i, and a is a parameter corresponding to the strength of interspecific competition. In the limit as $a \to 0$, we recover BM. Species that are nearby in trait-space “push” each other apart, tending to make their trait values more distinct. Therefore, this process predicts an overdispersion of trait values across the clade, and high phylogenetic signal. Simulated data sets are generated for this model using a discrete approximation to equation 4.2. For each time step, random Brownian motion is assigned to each phylogeny branch, and the effect of branches on each other are computed. We are left with a set of trait values for the phylogeny tips, i.e. the extant species. The competition interaction between any given pair of species can be turned on and off at predetermined times.

Because this character displacement model has many interacting parts, it is not obvious how to calculate a likelihood function for it. An Approximate Bayesian computation (ABC) method is therefore used to generate an approximate likelihood distribution for model parameters (rate of random change and strength of interactions) via simulation. This statistical approach is reviewed in an ecological context in Beaumont (2010) and Csilléry (2010). For a given phylogeny, many data sets are simulated, and summary statistics generated for each one. Then, for a given trait data set, we accept simulations with summary statistics that are closest to those of the true data set. The distribution of summary statistics for this ‘nearby’ set of simulations is then considered an approximation to the likelihood distribution. This process can be repeated for different models, or the likelihood can be maximised only along given planes in parameter space, to fit nested models. Likelihood ratio tests can then be used in the usual way to generate a P-value for generating the observed summary statistics under a null model. Here, that null model is typically Brownian motion. I chose three summary statistics: the phylogenetic signal K, and the mean and standard deviation of the distances in trait space between neighbouring species. These summary statistics reflect the intended features of the model. For each hummingbird clade, fitted model parameters and P-values were obtained. The P-value here is the likelihood of obtaining the observed summary statistics from data simulated under a BM model.

Blomberg's *K* (Blomberg et al. 2003) is a measure of phylogenetic signal. It measures whether more closely related species have more similar trait values, relative to the predictions of the BM model. It is expected to be 1 under BM, and less than 1 for OU processes. Values of *K* greater than 1 imply that trait values are even more phylogenetically structured than in BM. Such an excess of signal is predicted by the character displacement model. *K* was computed using the function Kcalc in the R package picante (Kemble et al. 2010), for each hummingbird clade and trait.

Results

Trait distributions

The body size distributions and beak length distributions for each of the 7 main hummingbird clades are shown in figure 4.2. The hermits have the widest range of values for both traits. Both the bees and mountain gems have relatively narrow size distributions, but for beak length the gems' distribution is very wide, while the bees remain tightly packed. The overall hummingbird trait distributions are approximately normal, but the distributions of individual clades often appear multimodal.

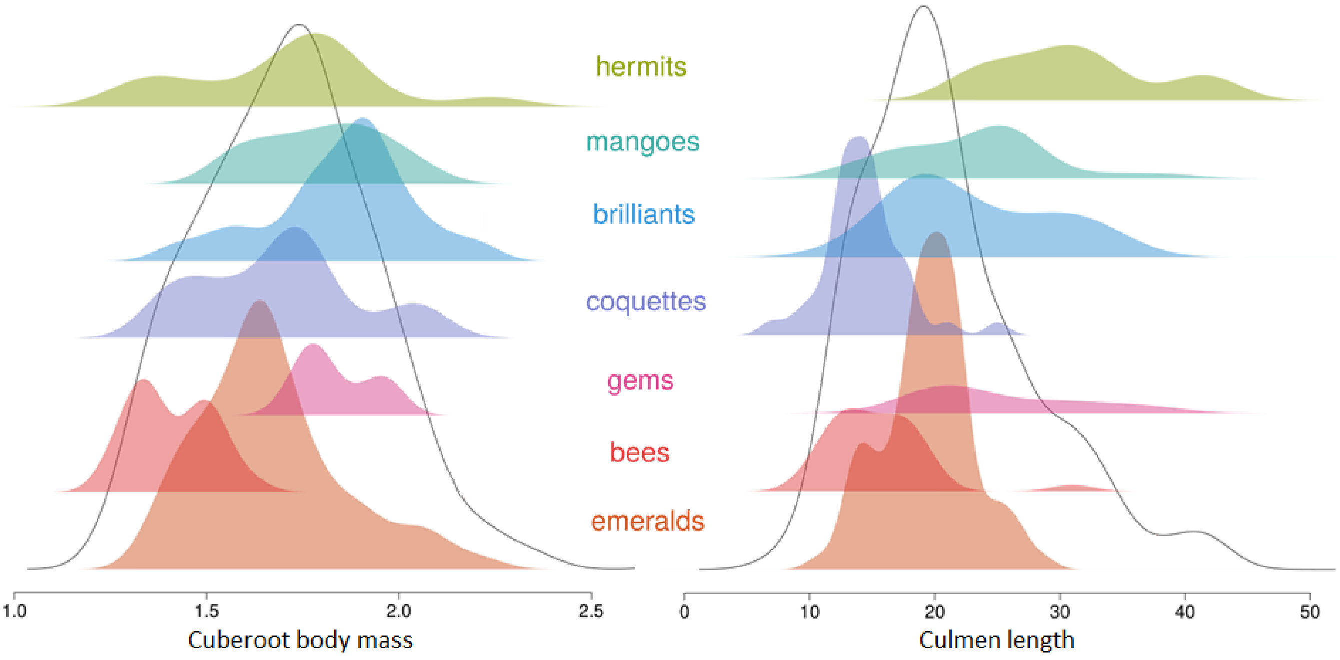

Figure 4.2: body size and culmen length distributions for 7 hummingbird clades. The area under each curve is proportional to the number of species in that clade.

Trait evolution modelling

Brownian, Ornstein-Uhlenbeck (OU) and early burst (EB) models were fitted to the seven main clades, as well as to the whole hummingbird tree. Maximum likelihood parameter estimates and null P-values are shown in table 4.1 for body size and in table 4.2 for beak length. Also shown are estimates of the phylogenetic signal K.

For body mass, the OU model is significantly preferred to BM for bees and mangoes. These are also the clades with particularly low K. There is no significant evidence for the EB model for any clade, although the whole hummingbird tree shows weak signs of a slowdown in trait evolution ($P = 0.24$). The whole tree also has high signal, with $K = 1.22$, whereas we would have $K = 1$ for Brownian evolution.

The results for beak (culmen) length are quite different. Here, the early burst model is strongly supported for brilliants and for the whole hummingbird tree. This is despite a weaker signal of $K = 0.84$ for the whole tree. The only clade with particularly high signal was the gems, with $K = 1.56$. The maximum likelihood estimate of slowdown for the gems was also high ($g = -0.44$), but not significant ($P = 0.132$).

Table 4.1: evolutionary models fitted to hummingbird body size. σ is the Brownian motion rate. K is Blomberg's K, measuring phylogenetic signal. α is the OU model parameter representing the strength of attraction towards a central 'optimum'. The early burst model parameter is also given, with negative values corresponding to slowdowns in trait change. P-values are shown for the OU model and EB model fits, as compared to the null BM model. Significant results are highlighted with a star.

Clade	BM: σ^2	Signal K	OU: α	OU: P-value	EB	EB: P-value
all	0.383	1.216	0.000	1.000	-0.044	0.238
emeralds	0.331	0.929	0.000	1.000	-0.046	0.629
bees	0.404	0.225	2.718*	0.001*	0.000	1.000
gems	0.232	0.446	0.985	0.067	0.000	1.000
coquettes	0.305	0.896	0.024	0.743	0.000	1.000
brilliants	0.289	0.909	0.000	1.000	-0.009	0.933
mangoes	0.403	0.394	0.238*	0.023*	0.000	1.000
hermits	0.261	3.433	0.000	1.000	-0.019	0.818

Table 4.2: evolutionary models fitted to hummingbird beak (culmen) length. σ is the Brownian motion rate. K is Blomberg's K, measuring phylogenetic signal. α is the OU model parameter representing the strength of attraction towards a central 'optimum'. The early burst model parameter is also given, with negative values corresponding to slowdowns in trait change. P-values are shown for the OU model and EB model fits, as compared to the null BM model. Significant results are highlighted with a star.

Clade	BM: σ^2	Signal K	OU: α	OU: P-value	EB	EB: P-value
all	0.383	0.842	0.000	1.000	-0.100*	0.016*
emeralds	0.331	0.530	0.072	0.215	0.000	1.000
bees	0.404	0.476	0.808	0.110	0.000	1.000
gems	0.232	1.559	0.000	1.000	-0.443	0.132
coquettes	0.305	0.380	0.285*	0.027*	0.000	1.000
brilliants	0.289	1.080	0.000	1.000	-0.353*	0.001*
mangoes	0.403	0.681	0.000	1.000	-0.031	0.847
hermits	0.261	0.695	0.067	0.314	0.000	1.000

Character displacement

The character displacement model adds a parameter to BM, as per equation 4.2. This additional parameter, a, is the strength of character displacement interactions. I fitted this model to each of the 7 main clades, as well as to the whole hummingbird tree. Body size in hermits had significant evidence of character displacement patterns. Beak length in gems showed some weaker support for competition. Overall, competition strength parameter values were slightly higher for body size than for beak length.

Table 4.3: character displacement models fitted to hummingbird body size.

Clade	σ^2	a	LRTS
all	2.60	0.76	1.35
bees	2.52	1.92	1.34
brilliants	1.76	0.40	1.27
coquettes	1.60	1.00	2.53
emeralds	2.24	1.20	1.51
gems	1.36	0.28	1.16
hermits*	1.52	4.04	4.06
mangoes	1.28	0.24	0.96

Notes: Star denotes significant support for competition versus Brownian motion (not corrected for multiple tests).

Table 4.4: character displacement models fitted to hummingbird beak length.

Clade	σ^2	a	LRTS
all	3.63	3.97	1.53
bees	2.20	0.37	0.98
brilliants	3.70	3.17	1.45
coquettes	1.67	0.17	0.70
emeralds	1.60	0.10	0.67
gems	1.68	0.53	3.04
hermits	0.97	0.07	0.91
mangoes	1.43	0.13	0.95

Discussion

I divided the hummingbird tree into several main clades. Overall, I found little evidence of character displacement patterns, with most clades favouring either a BM or OU model. The results indicate that the evolutionary processes shaping any given clade cannot easily be inferred from either phylogenetic closeness or historical environment. For example, the coquettes and brilliants are sibling clades, both radiating in the Andes, and, while both have similar results for body size, they show completely different patterns for beak length. Inference of these processes for whole groups of species is likely to be subject to a lot of individual variation and exceptions. This is known to be true for geographic scale (González-Caro et al. 2012), and I found a similar conclusion for phylogenetic scale. This inconsistency of phylogenetic signal and evolutionary mode across the whole hummingbird clade makes it important that studies of hummingbirds are cautious in their overall assumptions.

The assemblage composition of South American hummingbirds has previously been studied using differences in traits and relatedness between neighbouring communities (β diversity; Anderson et al. 2011, Weinstein et al. 2014). Similar methods have linked composition to aspects of environment; for example, there is greater phylogenetic clustering at higher altitudes in the Andes, suggestive of greater environmental filtering in harsher conditions (Graham et al. 2009). These studies use the scheme from Webb et al. (2002), where phylogenetic overdispersion implies competition between close relatives (limiting similarity), and phylogenetic clustering implies environmental filtering, where close relatives have some key trait required for a given environment. Graham et al. (2012) found that the phylogenetically clustered communities (the ones living in harsher environments) have the more even trait distributions. This may mean that while harsh environments filter for certain key traits, causing phylogenetic clustering, while competitive exclusion operates on the other, measured traits leaving them overdispersed (i.e. flatter distribution than a normal distribution) in the community. This is supported by the fact that Graham et al. (2012) also found body mass to have a reduced variance in harsh environments, but at the same time a more even distribution than expected under random null models. Phylogenetic methods have also been used to show that mechanisms like limiting similarity and environmental filtering can affect species in the same community differently, such that the forces acting on a species differ by its environment and its ancestry simultaneously (Parra et al. 2010, Lessard et al. 2016).

The phylogeny and trait data I used (Lessard et al. 2015) includes 279 extant species, with body mass and culmen length values for each of them. I found high phylogenetic signal in both traits for the whole tree. Signal was slightly higher for body size than for beak length. I also found highly varying signal between clades. Overall, the signal for the whole hummingbird clade is, particularly for body size, greater than the median of the signal for the individual clades. This suggests that there is more phylogenetic structure at large scales—e.g. between genera—than at the finer scale of individual species. This could be due to clades living in different places and having different mean body sizes according to their general environment, while beak length is more affected by local competition. Then, similar beak sizes would be found in distant locations, but geographically neighbouring species will compete for different niches and have more phylogenetically structured beak lengths.

Our results for phylogenetic signal are broadly consistent with the results of Graham et al. (2009, 2012). However, while I found higher signal for body size than beak length, Graham et al. (2012) found the opposite to be true for a more localised collection of 126 hummingbird species from Ecuador and Columbia. The contrast between these results supports the idea that body size may be more phylogenetically structured at larger scales. It is worth noting that Graham et al. (2012) go on to do analyses which assume that high phylogenetic signal is present across the hummingbird tree; these analyses may be less valid for clades where signal is much lower, such as the bees and mangoes.

A mild slowdown was significantly supported ($P = 0.016$) for the whole hummingbird clade. Since diversification rate has also slowed across the hummingbirds (McGuire et al. 2014), this result is consistent with a gradual filling up of available niches or new environments. The early burst model was not significantly supported (relative to BM) for body size in any of the clades. Overall, the early burst model is generally better supported for beak length, and the OU model for body size. Since body size is associated with warmth retention, and beak length is associated with diet and feeding habits, this overall result is consistent with ecological expectations.

Brilliants and coquettes

Brilliants and coquettes together form the ‘Andes clade’. They radiated where the Andes now stand, and most of their extant species live there now. Hummingbird body size is positively correlated with altitude overall, and the brilliants have the highest mean body size of all the main clades (Stiles 2008). They might therefore be expected to have a smaller range of body sizes, or to have their sizes constrained to be close to a high optimum for the harsher conditions at higher altitudes. If this were so, then the body size of this clade would likely fit an OU model best. I did not find that the OU model was a better fit than Brownian motion, however. In fact, figure 4.2 shows that these clades don’t have a particularly narraw distribution of size values, although the brilliants have the highest mean size of the major clades. Indeed, the coquette distribution appears to be multimodal. This may be due to species that have colonised areas outside the Andes and then evolved varying sizes. Or, it may be that body size is unconstrained by the harsh environment. Subsetting the clade by area or by altitude might help clarify this point in future analyses.

Both clades had high phylogenetic signal ($K = 0.90$ and 0.91) for body size, but for beak length the brilliants had far more phylogenetic structure than the coquettes ($K = 1.1$ and 0.4). The early burst model was not significantly supported (relative to BM) for body size in any of the hummingbird clades. For beak length, however, the EB model was strongly supported in the brilliants clade, with a large estimated slowdown parameter, $g = -0.353$. This suggests a rapid divergence of beak lengths early in the brilliants’ evolutionary history. The brilliants also have the highest mean beak length.

Though they are adapted to similar environments, brilliants and coquettes are very divergent (Stiles 2008). Both have large wings relative to their mass—an adaptation to a thinner atmosphere—but the brilliants are larger than the coquettes, and have longer bills. Coquettes have much bigger feet than any clade, suggesting that they do less hovering and more perching compared with the brilliants. The differences between these two clades seem to be reflected in distinct evolutionary histories, especially for beak length.

Bees and mountain gems

The bees and mountain gems are the hummingbird clades thought to have radiated in North America. I found that both have low signal for the body size trait. For beak length, the bees also have low signal ($K = 0.48$), but signal for the mountain gems is very high ($K = 1.56$). The mountain gems are almost entirely larger in size than all of the bees. This could reflect a greater generalism, in terms of feeding, among the bees, which are much smaller than the gems. It would be interesting to compare this result with a detailed study of the feeding behaviour of the two clades.

The maximum likelihood slowdown estimate (EB model) for the mountain gems is high ($g = -0.443$), suggesting an early burst of trait evolution consistent with an adaptive radiation. Despite having a stronger estimated slowdown than the brilliants, the mountain gems result is not significant; this may reflect the fact that the gems are simply a smaller clade. For the bees, however, the OU model is significantly preferred for body size. The bees can be seen in figure 4.2 to have a much smaller range of body sizes than most clades. Along with the support for the OU model, this suggests that the bees' sizes may be subject to relatively tight evolutionary constraints. If so, this shows that such constraints do not necessarily coincide with harsh environments: although some bees live in mountainous areas, the clade is not composed exclusively of high-altitude species.

The bees have the highest diversification rate of all the main clades, and, along with the rest of the hummingbirds, a significant recent slowdown (McGuire et al. 2014). This suggests the possibility that the bees may be a recent radiation. However, their trait evolution turns out to be OU-like, more so than any other clade. This implies that the bees may not be a recent radiation, but rather an older clade with rapid, continual species turnover.

Hermits and emeralds

The hermits are thought to have been the first hummingbirds to have occupied the lowlands east of the Andes, and are considered to have radiated there (Parra et al. 2010). This clade had very high signal ($K = 3.43$) for body size, but only moderate signal ($K = 0.70$) for culmen length. The hermits and the emeralds have the two highest measures of signal for body size; both are considered lowland species, so this is consistent with increasing evolutionary constraints on body size at higher altitudes. The hermits also had significant support for the character displacement model. This model predicts a rapid expansion into trait-space, so that species' trait values are well differentiated. Evolutionary constraints, on the other hand, are likely to erode phylogenetic signal and produce an OU-like pattern over enough time.

Character displacement

I found little support overall for character displacement patterns in hummingbird clades. This does not exclude the possibility that character displacement and niche-packing mechanisms have an important role in hummingbird evolution; however, it suggests that they don't operate at the phylogenetic and geographic level of the analysis. It would be interesting to compare these results to an application of the model to smaller clades of species that either coexist or share geographical borders.

Although the competition model has character displacement interactions, it does not necessarily assume sympatric speciation. Similar patterns can arise from adaptation to alternative niches in neighbouring environments as well as in the same environment. For example, in North-American wood warblers, local coexistence has been found to increase with the time since the last common ancestor of the species (Lovette and Hochachka 2006). This suggests that competitive exclusion acts on new sibling species, but that after time passes, those species diverge phenotypically, and become different enough to come back into coexistence. This sort of process produces the same patterns we expect to see from sympatric speciation and character displacement.

The exception to the results for this model is the body size of hermit species. For this clade, the competition model was strongly supported, with a large effect size ($a = 4.04$). This clade was a lowland radiation, initially isolated to the east of the Andes. This isolation from the other hummingbirds may go some way towards explaining these results; however, it is then not clear why beak length has far less phylogenetic structure. A more detailed investigation of this clade, combining phylogenetic and geographic data with evolutionary modelling, may help clarify this point. With coexistence data, for example, a coexistence matrix could be added to the competition model. This could establish whether the phylogenetic structure in hermit body size is due to character displacement in sympatry, or, for instance, sequential adaptation to environments at distinct altitudes.

Stochastic models of trait evolution have been fitted to the whole hummingbird phylogeny: an OU model was preferred for mean temperature, mean precipitation and range size, while a BM model was preferred for body mass and beak length (Lessard et al. 2016). These models were not fitted to any subdivisions of the hummingbird phylogeny. Diversification models have been found to vary greatly among hummingbird clades (Jetz et al. 2012, McGuire et al. 2014), with for example high rates in bees and some of the emeralds, and low rates in the mangoes. Since diversification rates vary so much between clades, these different clades may be expected to have undergone different regimes of trait evolution also.

Conclusion

For the analyses, I divided the hummingbird tree into several main clades. The results indicate that the evolutionary processes shaping any given clade cannot easily be inferred from either phylogenetic closeness or historical environment. For example, the coquettes and brilliants are sibling clades, both radiating in the Andes, and, while both have similar results for body size, I find that they show completely different patterns for beak length. Inference of these processes for whole groups of species is likely to be subject to a lot of individual variation and exceptions. This is known to be true for geographic scale (González-Caro et al. 2012), and I find a similar conclusion for phylogenetic scale.

This high level of variance between subclades means that care must be taken in choosing the assumptions that are made in further analysis. For example, a common assumption for the whole hummingbird tree is phylogenetic signal consistent with Brownian evolution (e.g. in Graham et al. 2012). But signal varies a lot between clades and between scales, becoming much more varied at finer scales. Hence, analyses making this assumption for the whole hummingbird tree may be valid only for certain portions of the tree.

I found little evidence of character displacement patterns in hummingbirds. However, this study was restricted to body size and beak length. A better place to look for evidence of competition may be in beak shape. The character displacement model could be applied, for example, to beak length, width and curvature simultaneously. This would provide a characterisation of the beak that would be better related to its ecological niche. Using this kind of data would therefore be a good next step.

Chapter 5: Diversification rates and lineage trait densities

Abstract

Diversification based on the ‘filling up’ of niches can be modelled in several different ways. I introduce a new model that couples diversification rates to trait packing densities, and compare the predictions of this model to those of two other, related models. With similar conceptual starting points, these three models of diversification leads to very different results in terms of tree shapes and trait distributions. I characterise these differences using simulated data from each model, and discuss the implications for diversification modelling and the inference of ecological effects from phylogenies.

Introduction

Diversification rates

A fundamental goal of evolutionary study is to understand when and why species multiply. Diversification rates vary between clades (Jetz et al. 2012), between different sizes of clade (Phillimore and Price 2008, Etienne et al. 2012), between climates (Kozak and Wiens 2010), across geographic domains (Moen and Morlon 2014), through time (Rabosky 2006) and between ecological groups (Martin and Wainwright 2011). Rates can also be affected by key morphological innovations (Dumont et al. 2012), or more generally by species’ trait values (Paradis 2005, Maddison et al. 2007, Freckleton et al. 2008, Fitzjohn 2010) or rates of trait evolution (Adams et al. 2009, Rabosky et al. 2013).

A molecular phylogeny containing only extant species and their history can be used to infer speciation and extinction rates (Nee et al. 1994, Paradis 1997). The overall diversification rate, equal to the speciation rate b minus the extinction rate d is given by the total accumulation of species and the lifetime of the clade. The extinction rate, meanwhile, is computed from the diversification rate $b-d$ and the ‘relative extinction rate’ d/b (Nee et al. 1994). The latter ratio is obtained via the apparent change in diversification rate through time. When the relative extinction rate is low, the diversification rate will simply approximate the speciation rate b. But when the relative extinction rate is high, the diversification rate will appear to increase through time, meaning that extant species are generally younger. The rate of diversification therefore appears to accelerate with time (Rabosky 2006). From this framework, likelihoods for speciation and extinction rates are computed as functions of phylogenetic data.

In these simplest cases, rates of extinction and diversification are assumed to be independent of ecological or biological factors. However, this ignores a suite of possible processes that shape trait evolution and consequently models have been developed for diversification rate change as a function of binary traits (Maddison et al. 2007, Magnuson-Ford and Otto 2012, Rabosky and Goldberg 2017), quantitative traits (Fitzjohn 2010), or discrete rate changes at unknown times (Rabosky 2014).

Diversity dependence

Another potential source of diversification rate changes is diversity dependence, where the diversification rate for each lineage depends on the total number of lineages in the clade. This is the focus here. In diversity dependent models, the diversification rate for each lineage depends on the total number of lineages in the clade. A fixed number of niches are assumed to be available to the whole clade (Walker and Valentine 1984, Etienne et al. 2012), and as they become filled, the speciation rate declines. This approach is rooted in the idea of diversification as a function of ecological opportunity (Schluter 2000), and is consistent with slowdowns in diversification rate through time. Slowdowns are common in studied clades (Phillimore and Price 2008), and are often interpreted as evidence of adaptive radiation (e.g. Price 2008).

Diversity dependence has been modelled using the number of niches as a hidden Markov variable, with extinction and speciation generating continual species turnover (Etienne et al. 2012). The equation for this process is

$$\lambda_n = \max\left(0, \lambda_0\left(1-\frac{n}{K}\right)\right), \qquad \text{Equation 5.1}$$

where n is the number of species, λ_0 is the initial speciation rate, and K is the maximum number of available niches.

Under a pure birth Yule model, the expected diversification rate of each lineage is the same, and the total number of lineages grows exponentially with time. Any overall change in diversification rate with time can be quantified by the statistic γ, which has a null expectation value 0 for a pure-birth model. Negative values of γ are indicative of a diversification slowdown (Pybus and Harvey 2000). The median value of γacross 45 bird clades was found to be around -1.2, and 15 of the 45 clades had significantly negative gamma ($\gamma < -1.645$) (Phillimore and Price 2008), suggesting that diversification slowdowns are fairly common among bird clades. Another aspect of phylogeny shape is imbalance: the degree to which historical sibling species have differing numbers of descendants. Imbalance is important because it indicates that different lineages have diversified at different rates, and observed trees tend to have substantial imbalance (e.g. Purvis et al. 2011). The processes responsible for generating diversification rates are therefore likely to leave their mark in the resulting imbalance of the tree. Models that generate different diversification rates between lineages may produce more realistically imbalanced trees than models where all lineages have an equal rate, even if that rate changes through time. However, this will depend on the details of the models.

Diversity dependent models of rate variation are based on the idea that the number of ecological niches is finite and that these niches fill up as a clade evolves and diversifies (Walker and Valentine 1984, Etienne et al. 2012). If traits are directly linked to niches, this means the models are implicitly connected to the species' trait values, implying a diversification rate that slows as trait-space becomes more densely occupied. This effect has been demonstrated experimentally with bacteria in a lab experiment (Bailey et al. 2013). However, the diversity dependence model does not include any explicit mode of trait evolution.

Trait evolution and diversification

Methods have been developed to model traits under a variety of different processes. Traits are commonly assumed to follow a diffusion process such as Brownian motion (BM, Felsenstein 1985, Freckleton et al. 2002) or an Ornstein-Uhlenbeck process (OU, Hansen 1997). A BM model predicts that traits follow random trajectories, becoming on average ever more different from each other. Given enough time, traits of a given set of species become less densely occupied. However, the density of traits may increase or decrease depending on the relative rates of diversification and trait evolution. The OU model is similar to the BM model, but adds an attraction to a central 'optimum' value; this means traits are likely to stay more tightly packed over time, and increasing numbers of species should increase trait densities. A BM process preserves phylogenetic signal, so that species are more similar to their close relatives, whereas OU processes tend to erase this signal over time.

Models also exist for the simultaneous, linked effects of trait (or niche) evolution and speciation events, where all the evolution occurs at the speciation event and a taxon is more likely to speciate if it is close to empty niches (Price 1997, Harvey and Rambaut 2000, Freckleton et al. 2000, Price et al. 2014). In these 'niche filling' (NF) models, species' trait values are constant in time, and the position of each niche in niche-space is constant in time. The available niches are randomly positioned in niche-space, typically following a multivariate normal distribution. They are then gradually filled by new species, branching from whichever species is closest. The NF model is conceptually related to diversity dependent diversification, since it is based on the gradual filling up of available niches; however, the models operate differently, and the NF model predicts—by definition—the creation of new species at regular time intervals, and therefore a slowdown in per-lineage diversification rates. Like BM evolution, the NF model predicts strong phylogenetic signal.

A new model: trait density dependent diversification

I introduce a new model, where the diversification rate changes with the trait packing density of the species. This approach is consistent with the ecological idea of competitive exclusion, usually invoked for community assembly studies (e.g. Webb et al. 2002). Because trait evolution and trait packing are tracked explicitly, this model generates differences in diversification rate between lineages. This is in contrast to the diversity dependence approach, where the diversification rate for the whole clade is slowed by the accumulation of new species anywhere within the clade. The detailed equations and simulation process for the trait density dependence model are explained in the methods section below.

Goals

In this chapter, I examine the predictions of diversity dependent diversification models, and consider their compatibility with common trait evolution models. I simulate data under diversity dependence models, as well as the NF model (Price 1997). I also demonstrate a new model: density-dependent diversification based on trait-packing, rather than the total number of species in the clade. I show how this new model works, how it can be simulated, and characterise the datasets it generates.

These three models—diversity dependence, trait density dependence, and niche-filling—all have closely related assumptions, and are plausible under similar conditions: adaptive radiations where the total number of species is increasing towards an ecologically constrained maximum. I compare the predictions of these three models, and discuss their implications for the inference of ecological processes from diversification patterns.

Methods

Model for trait density dependent diversification

The new model treats diversification rate as a linear function of distances between neighbouring traits. For each new species, the Euclidian distance to its nearest neighbour in trait-space is calculated, and used as a parameter in determining the time until the next speciation event on that lineage. If the trait value for species *i* is denoted q_i, and the nearest species in trait-space is species j, then the speciation rate of lineage *i* is

$$\lambda_i = \max\left(0,\ \lambda_0 + \frac{b}{|q_i - q_j|}\right), \qquad \text{Equation 5.2}$$

where b is a parameter representing the strength of the effect of trait values of speciation rate. This equation is plotted in figure 5.1, showing the effect of b on the diversification rate. Negative values of b cause reduced diversification rates. For trait distances smaller than $-b$, no speciation can occur. For distances larger than $-b$, the rate of diversification gradually increases up to a limit of λ_0 when the distance is very large relative to $-b$. In this sense, b is a 'characteristic trait difference' of the model: the point at which speciation becomes impossible, and which suppresses speciation when approached. Positive values of b indicate that common trait values promote speciation. If there were also a linear dependence on the trait value q_i itself, then we could add a parameter a for the effect of trait value, and write

$$\lambda_i = \max\left(0,\ \lambda_0 + b\frac{1}{|q_i - q_j|}\right) + aq_i. \qquad \text{Equation 5.3}$$

To simulate this model, I constructed trees and trait datasets branch by branch, using functions in the R package *ape* (Paradis et al. 2004). The speciation times were generated from the rates described above, by sampling randomly from an exponential distribution with rate λ, using the R function *rexp*. The trait values on each branch were evolved between speciation events according to one of the common diffusion models: Brownian Motion (BM; Felsenstein 1985) or the Ornstein-Uhlenbeck (OU) process (Martins 1994, Hansen 1997, Butler and King 2004; but see Cooper et al. 2016). These are stochastic functions, where the branches are all independent of one another. Therefore, although the diversification rate of a lineage is affected by other lineages, its trait evolution is not.

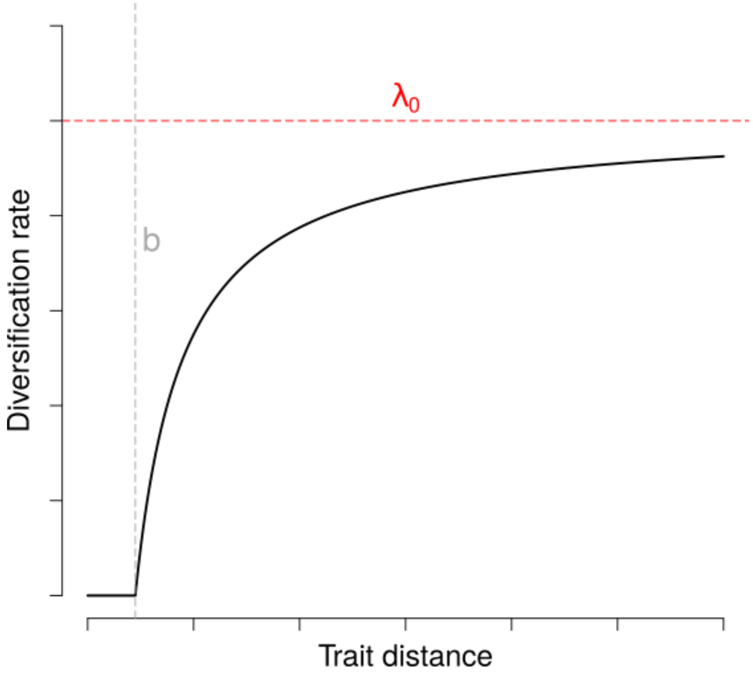

Figure 5.1: Trait-diversity dependence (TDD) model illustration. λ_0 is the diversification rate when it is unaffected by trait packing. b is the strength of the trait packing effect, as per Equation 5.2.

The procedure for generating simulations starts with a single root branch and a single root trait value. The following steps are then followed:

1. Take the youngest tip and split it so that at its position we now have two tips with zero branch length and equal trait value.
2. Compute the speciation rate of the splitting tip according to Equation 5.2 or Equation 5.3.
3. Set the two new branch lengths l_i equal to each of two random draws from an exponential distribution with rate λ. The probability density for any given branch length (time to next speciation event) t is

$$p(t) = \lambda e^{-\lambda t} \qquad \text{Equation 5.4}$$

4. Compute the trait values of the two new tips. If the model of trait evolution is BM with rate parameter σ, then the new trait values will be draws from a normal distribution with mean equal to the trait value of the start node and variance $l_i\sigma^2$. If the trait model is OU, then in addition to random Brownian diffusion, the lineage is attracted to a 'central optimum' with a strength parameterised by α.
5. Repeat steps 1 to 4 until the tree has at least the desired number of tips.

6. Trim the tree by imposing a time cutoff such that the tree has the desired number of tips and is ultrametric.

This implementation was done entirely in R, using the packages *ape* and *mvtnorm*. For OU trait evolution, I used the function *rTraitCont* in the R package *ape* (Paradis et al. 2004). This was applied separately for each edge of the phylogeny—that is, for each lineage section, as divided by speciation events.

Simulations

I generated simulated trees and trait datasets using each of the models: diversity dependent diversification (DD, Etienne 2012), niche filling (NF, Price 1997), and the new model presented here, trait-density dependent diversification (TDD). Using a wide range of model parameters, with and without extinction, allows us to compare their behaviour. The trees were largely 100 tips in size (though I also assessed the effect of tree size; see below), since this is similar to that of larger adaptive radiation datasets (e.g. Harmon et al. 2010a). Also, the relative behaviour of the models was largely unaffected by total tree size: the effect of larger trees, in the case of the diversity dependence model, for example, was equivalent to reducing the total number of available niches in the model; both γ and I_c tend to be inflated as trees get larger (see Appendix 5A). These simulations were done for both a speciation-only scenario and a speciation with extinction scenario, with a constant rate of extinction set to 0.3.

For the standard diversity dependence model (DD), as described in Rabosky and Lovette (2008) and Etienne et al. (2012), I obtained 100-tip trees from scenarios where the total number of available niches, k, was either 120 or 240. In each case, I used extintion (μ) values of 0, 0.1 and 0.4. For each set of model parameters, 100 simulations were made.

With the trait-density dependence model (TDD), I generated simulations using density effect strengths (b) of 0.0, -0.1, -0.2 and -0.5. I used both BM trait evolution and OU evolution, with $b = -0.1$ and $\alpha = 0.1$ in the latter case; such small values of the OU and TDD parameters were necessary because, when used in combination, they suppress speciation very strongly, and larger parameter values result in trees that never reach 100 lineages. Further simulations used higher effect strengths ($b = -0.6, -0.8, -0.9$) with both BM trait evolution. Most of these simulations had no extinction, but I also ran simulations with nonzero extinction rates ($\mu = 0.1, 0.4$) for $b = -0.5$. For each set of model parameters, 100 simulations were made.

The niche-filling model of Price (1997) (see also Harvey and Rambaut 2000, Freckleton et al. 2000) has no parameters to be varied, except for the multivariate normal distribution from which niches are drawn. The correlation structure of different traits was not the focus here, so I set all the niche-filling simulations to use a 2-dimensional normal niche distribution, with unit variance and no covariance.

Finally, I compared results for different tree sizes, to check whether the qualitative outcomes were affected by clade size. I simulated trees with 20, 50 and 100 tips. For the diversity dependence model, I used a maximum niche count k that was 20% larger than the number of tips: 24, 60 and 120 niches respectively. For the trait density depencence model, I used $b = -0.5$.

Summary statistics

I computed a range of descriptive statistics for simulated trees. First, tree imbalance was measured using Colless' index I_c (Mooers and Heard 1997), obtained using the function *colless* in the R package *apTreeshape* (Bortolussi et al. 2006). This statistic is the sum over nodes of the differences in the number of tips between the two clades that are descendent from each node. I used a Yule-normalisation (Bortolussi et al. 2006) so that the expectation value under a Yule process is 0, and positive values indicate greater imbalance than expected. The Colless index I_c is computationally simple and has well established statistical properties (Mooers and Heard 1997), making it well suited to a simulation study like the one presented here.

Bias of internal node positions towards the root or the tips can be measured by the γ-statistic (Pybus and Harvey 2000), and I obtained it using the function *gammaStat* in the R package *ape* (Paradis et al. 2004). Extinction can negatively bias the γ statistic, i.e. create the appearance of diversification slowdowns where they don't truly exist (Phillimore and Price 2008, Pennell et al. 2012). However, it is the most commonly used slowdown statistic, so I have chosen to categorise the models here using it. Also, I explicitly consider extinction rates in the simulations. It is important to remember that significantly negative values of γ must be interpreted with caution, however.

I also looked at the distribution of traits across the tips. Under a diffusion model, tip trait values are expected to be (multivariate) normally distributed. However, the trait-density dependent model has speciation rates that vary with trait values, and this may mean that tip traits are no longer normally distributed.

Phylogenetic signal can be measured using Blomberg's K (Blomberg et al. 2003). Although λ is perhaps more frequently used for quantifying signal, K was preferable here because it can measure both supressed and elevated signal compared with BM. I computed K for each set of simulated data and phylogeny using the function *Kcalc* in the R package *picante* (Kembel et al. 2010).

Results

The model and model parameters for each simulation, along with the mean and standard deviation of the various summary statistics, are shown in table 5.1.

Tree imbalance

Tree imbalance, as measured by the Colless index I_c (Mooers and Heard 1997), is shown in figure 5.1, for each of the three models, as a function of the strength of ecological effects. For the DD model, this 'strength' is how close the number of tree tips is to the total number of available niches. In the NF model, the situation is similar, but the number of niches is shown not explicitly but as a proportion of the number of tree tips. In the TDD model, the effect strength is an explicit model parameter, b.

I found that trees simulated using DD were slightly more balanced than random trees: they had $I_c < 0$, corresponding to sibling subclades with more similar numbers of tips than expected by random chance. The TDD and NF models had strongly positive I_c, meaning that they produced much more imbalanced trees. The strongest imbalance ($I_c = 12.41$) was produced by the NF model, while the TDD trees were most imbalanced when they had high extinction rates and OU trait evolution.

Table 5.1: Model simulations and average summary statistics. PB the ‘null’ pure-birth diversification model, with Brownian trait evolution. b is the strength of the effect of trait packing on diversification rates in the trait-density dependence model. k is the total number of niches available in the DD model. μ is extinction rate. α is the strength of attraction to the global ‘optimum’ trait value in the OU trait evolution process; $\alpha = 0$ corresponds to BM trait evolution. γ is a measure of diversification rate change through time. I_c is the Colless index of tree imbalance. K is a measure of phylogenetic signal in trait values. The stated values of these summary statistics are the means across 100 simulations; the standard deviation is also shown.

b	k	μ	α	γ	$sd(\gamma)$	I_c	$sd(I_c)$	K	$sd(K)$
Pure birth diversification model with Brownian trait evolution:									
0	NA	0	0	0.08	1.06	-0.03	0.77	1.06	0.62
Diversity dependence:									
NA	120	0	NA	-2.30	0.92	-0.18	0.79	NA	NA
NA	120	0.1	NA	-2.32	1.03	-0.33	0.66	NA	NA
NA	120	0.4	NA	-2.34	1.02	-0.26	0.77	NA	NA
NA	240	0	NA	-1.01	0.86	-0.10	0.69	NA	NA
NA	240	0.1	NA	-0.93	1.00	-0.04	0.72	NA	NA
NA	240	0.4	NA	-0.84	1.02	0.02	0.76	NA	NA
Niche filling:									
NA	NA	NA	NA	-5.60	0.00	12.41	3.99	1.50	0.53
Trait density dependence:									
-0.1	NA	0	0	-2.14	1.05	0.06	0.79	0.16	0.09
-0.2	NA	0	0	-2.92	1.22	0.40	0.88	0.21	0.11
-0.5	NA	0	0	-3.48	1.29	1.16	1.24	0.24	0.11
-0.5	NA	0.1	0	-2.75	1.42	2.23	1.53	0.22	0.11
-0.2	NA	0.2	0	-2.12	1.54	1.30	1.14	0.21	0.13
-0.1	NA	0	0.1	-3.77	1.40	0.12	0.81	0.28	0.17
-0.1	NA	0	0.1	-4.05	1.31	0.12	0.90	0.29	0.15
-0.1	NA	0	0.2	-5.23	2.04	0.43	1.01	0.34	0.17
-0.1	NA	0	0.3	-5.05	2.00	0.39	1.01	0.30	0.15
-0.1	NA	0	0.4	-6.92	2.29	0.44	0.96	0.37	0.17
-0.6	NA	0	0	-3.73	1.36	1.47	1.22	0.25	0.12
-0.7	NA	0	0	-3.34	1.37	1.55	1.32	0.23	0.11
-0.8	NA	0	0	-3.41	1.28	1.83	1.28	0.23	0.11
-0.9	NA	0	0	-3.43	1.34	2.04	1.46	0.21	0.11

It can be seen that the NF model creates highly imbalanced trees regardless of the number of available niches, and that this number has little effect on I_c ($R^2 = 0.01,\ P = 0.27$ for linear correlation). The number of available niches in the DD model also has little or no effect on I_c, the tree imbalance being close to the null (random pure birth) expectation of $I_c = 0$ in all cases ($R^2 = 0.01,\ P = 0.25$ for linear correlation). In the TDD model, however, the strength of the coupling between trait packing and diversification has a large effect on tree imbalance ($R^2 = 0.58,\ P < 10^{-15}$). $b = 0$ corresponds to a pure-birth model, and the tree imbalance corresponds to the null expectation $I_c = 0$. As b becomes more negative, however, close trait packing suppresses diversification, and trees become progressively more imbalanced. When $b = -2$, the trees are about as imbalanced as they are under the NF model. When an OU trait evolution process is used in the TDD model rather than a BM process, the relationship between trait-diversification coupling and tree imbalance is no longer present ($R^2 = 0.00,\ P = 0.50$).

Diversification rate changes

The overall speedup or slowdown of diversification rate in a phylogeny is measured by γ. I find γ to be negative for all three models, with moderate effects in DD models and the strongest effects in the NF model and TDD models with OU trait evolution (i.e. $\alpha > 0$). Extinction rates do not appear to affect the γ-statistic in the DD and TDD models.

For the NF model, γ is a constant, because new species are created at regular intervals, unaffected by any other parameters of the model. This can be seen in figure 5.2, where the NF slowdown is unaffected by the ratio of available niches to number of tree tips. When the DD model is used, trees show a significant slowdown in diversification rate when the number of tree tips is close to the total number of available niches, k ($R^2 = 0.35,\ P = 7 \times 10^{-11}$). As k becomes large relative to the tree size, γ approaches 0. The TDD model produces trees with strong slowdowns when b is strongly negative ($R^2 = 0.04,\ P = 0.04$ for BM process), especially when the trait evolution follows an OU process ($R^2 = 0.64,\ P < 10^{-15}$). As b approaches 0, so too does γ.

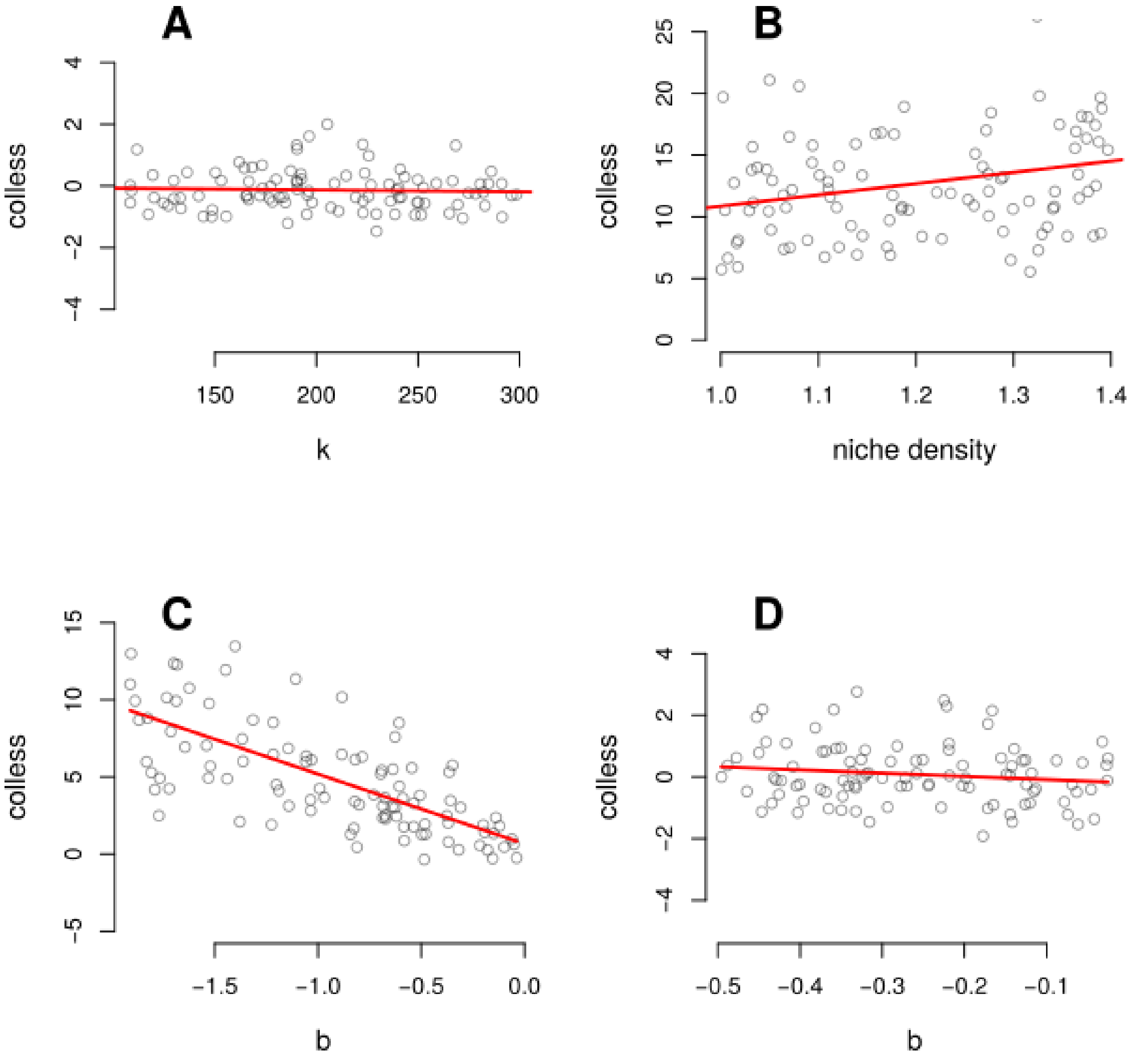

Figure 5.2: Tree imbalance, as measured by the Colless statistic I_c, as a function of niche packing measures for various models. Each point is an average over 100 simulations. A, diversity dependent diversification (Etienne 2012); k measures the total number of niches. B, niche-filling model (Price 1997); niche density corresponds to the number of niches available relative to the number of tips on the phylogeny. C, trait-density dependent diversification with BM trait evolution; b is the trait effect strength parameter. D, trait-density dependent diversification with OU trait evolution.

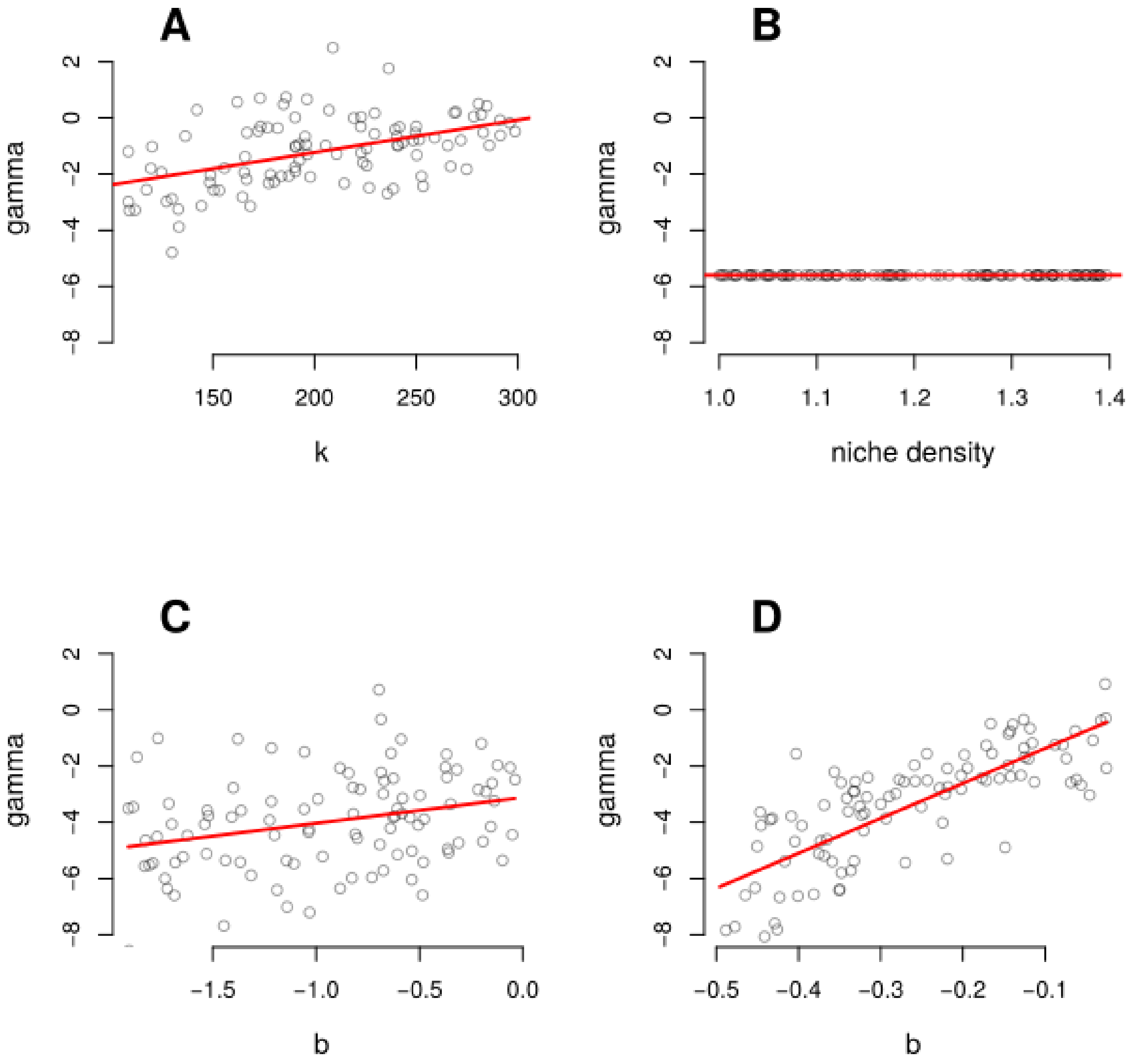

Figure 5.3: Diversification speedup / slowdown. The gamma statistic γ measures the rate of change of diversification rate with time. Negative γ corresponds to a diversification slowdown. Each point is an average over 100 simulations. A, diversity dependent diversification (Etienne 2012); *k* measures the total number of niches. B, niche-filling model (Price 1997); niche density corresponds to the number of niches available relative to the number of tips on the phylogeny. C, trait-density dependent diversification with BM trait evolution; b is the trait effect strength parameter. D, trait-density dependent diversification with OU trait evolution

Phylogenetic signal

Phylogenetic signal in trait values was measured by K. DD is purely a diversification model, so it does not generate any trait values, only trees. If trait evolution were simulated on these trees using BM, then the expected signal would be $K = 1$; if an OU process were used instead, then K would be lower. The other two models produce trees and trait values together, since they couple trait evolution to the diversification process. In the NF model, signal is increased (to $K = 1.42$) relative to the null (BM) expectation of $K = 1$. In the TDD model, however, signal is substantially reduced, down to $K = 0.22$ for Brownian trait evolution and no extinction.

Effects of tree size

I found that tree size did not affect the relative result from the three tested models. Overall, however, larger trees resulted in greater values of γ and I_c – that is, greater apparent slowdowns and greater imbalance. Representative results for different tree sizes are shown in table 5A.1, in appendix 5A.

Table 5.2: Summary statistics for model simulations using different tree sizes.

Tips	b	k	μ	α	γ	$sd(\gamma)$	I_c	$sd(I_c)$	K	$sd(K)$
Diversity dependence:										
100	NA	120	0	NA	-2.34764	1.19332	-0.26084	0.722059	NA	NA
50	NA	60	0	NA	-1.66691	1.081896	-0.12049	0.719663	NA	NA
20	NA	24	0	NA	-1.1135	1.148504	-0.1448	0.55432	NA	NA
Niche filling:										
100	NA	NA	NA	NA	-5.60341	0	11.51796	3.836061	1.459877	0.39711
50	NA	NA	NA	NA	-3.84615	0	5.869109	2.711725	1.360428	0.407361
20	NA	NA	NA	NA	-2.22681	0	2.713199	1.657943	1.254571	0.418132
Trait density dependence:										
100	-0.5	NA	0	0	-3.19561	1.332053	1.084761	1.186597	0.230155	0.093587
50	-0.5	NA	0	0	-2.16324	1.224179	0.422709	0.995987	0.281092	0.173794
20	-0.5	NA	0	0	-1.01211	1.135204	-0.0168	0.674144	0.468933	0.353014

Discussion

I examined three diversification models that all have a conceptual basis in species ‘filling up’ the available niches. I show how similar conceptual ideas can lead to very different results, depending on the implementation.

Theory and assumptions for the three diversification models

Diversification rates are known to vary greatly in nature (Phillimore and Price 2008, Kozak and Wiens 2010, Jetz et al. 2012, Moen and Morlon 2014). Apparent slowdowns in diversification rates are often attributed to diversity dependent diversification, where speciation rates vary with the number of extant species present in the clade. The basis for this is the idea that only so many niches are available to the clade, and, as they become filled up, lineages become less likely to split (Rabosky 2006, Etienne et al. 2012).

The diversity dependence model (DD; Phillimore and Price 2008, Etienne 2012) and the new ‘trait density dependence’ (TDD) model effectively differ in the way they distribute niches in niche-space. The DD model assumes that every new species reduces the ecological opportunities available to other species, regardless of which particular niche it occupies. So, for example, if the smallest species in a clade diversifies into two species, then the probability of a new species splitting from any other lineage in the clade – even the largest – is reduced. Niches are thus assumed to all interact equally.

The TDD model, on the other hand, has niches evolving in accordance with a diffusion model, as is commonly assumed for traits (e.g. Felsenstein 1985, Pagel et al. 1997). Interactions between species occur only when their niches are nearby in niche-space. This approach is perhaps more conceptually intuitive, in that new lineages suppress each other’s diversification to a greater extent when they are more similar. However, it also implies that niches can be measured, and that they evolve in a manner similar to traits. This may be justified by the evidence that morphological traits often correspond closely to ecological niches (e.g. Pigot et al 2013). Since the TDD model has a diversity dependence that affects different lineages differently, I can expect it to generate greater imbalance in diversity between different parts of a clade.

The niche-filling model (NF; Price 1997, Harvey and Rambaut 2000, Freckleton et al. 2000) is included in this study because it is conceptually similar to the TDD model, although it does not fit a diversification rate to datasets. Rather, the NF model simply assumes that each new lineage arises after a fixed time interval. Contrary to the DD model, the probability of a new species arising in a particular niche is not affected by how filled up the surrounding niches are. However, the probability of a particular lineage splitting is affected by the surrounding niches, since, as they get filled up, the number of available nearby niches diminishes. Therefore, the NF model may potentially generate different patterns of both trait distributions across tree tips, and tree imbalance (the relative diversity of different parts of the clade).

Comparing predictions of the three models:

The average predictions from simulations of each model are summarised in table 5.3.

Table 5.3: Matrix of the qualitative predictions of the three diversification models.

Model	Tree imbalance	Apparent diversification slowdown	Phylogenetic signal
DD	none	slight	NA
NF	strong	strong	high
TDD (BM)	moderate	moderate	low
TDD (OU)	none	strong	low

The summary statistics observed from simulations of the three models—DD, NF and TDD—are quite different. Firstly, the DD model produces trees that are slightly more balanced than pure-birth trees, whereas the TDD and NF models produce significantly imbalanced trees. The latter correspond better to observed patterns: real trees tend to be imbalanced (Blum et al. 2006). Of course, there are many possible sources of tree imbalance: any source of differences in diversification rates is also a source of tree imbalance (Mooers and Heard 1997). A tree may also appear to be imbalanced due to nonrandom omission of species, or because of inconsistent approaches to defining species. While trees are imbalanced for many different reasons, a statistical model should generate realistic imbalance if its fit to diversification rates is in any way meaningful.

All three models produced trees with slowdowns in diversification rate through time (negative γ). Extinction did not affect the slowdown in the DD model, but greater extinction rates caused less pronounced slowdowns in the TDD model. The greatest slowdowns (most negative values of γ) were obtained for the NF model. However, this result cannot be considered as an ecological effect, since the model simply produces new species at regular intervals. The NF model cannot, therefore, be fitted to an observed rate of change in diversification rate. That said, this model could readily be extended by allowing the speciation times of different niches to depend on their position in the niche distribution and on the occupancy rates of nearby niches.

The NF and TDD models generate trait data along with each phylogeny. The DD model does not generate trait data, since trait values and diversification are completely decoupled; therefore, the expected phylogenetic signal for the DD model is simply $K = 1$ under a BM trait evolution process, and $K < 1$ under an OU process. Phylogenetic signal is raised in the NF model, to be significantly stronger than Brownian evolution. Under the TDD model, on the other hand, signal is significantly reduced, with values as low as $K = 0.22$ when simulated with Brownian trait evolution and strong ($b = -0.5$) trait packing effects.

Inferring ecological processes

Since these three models are all based in the idea of diversity limited by the available niches, they all predict diversification slowdowns. Such slowdowns are expected to occur when the clade is an adaptive radiation (Harmon et al. 2010a), and this pattern has been found in many bird clades that resemble adaptive radiations (Phillimore and Price 2008). However, a study of a complete tree of birds found accelerating diversification (Jetz et al. 2012). Care should be taken, however, in interpreting significantly negative values of γ as a sign of adaptive radiation, since the γ statistic is sensitive to recent rate changes, and not good at detecting early bursts of evolution (Fordyce 2010).

If trait values correspond to niches, then these results suggest that DD models are not consistent with Brownian trait evolution, nor with any continuous trait evolution process that preserves phylogenetic signal. If we want to generate high phylogenetic signal, as is typically observed in adaptive radiations, then we need a different form of trait evolution, such as the non-continuous mode found in the NF model. Either way, the DD model does not generate realistic levels of tree imbalance, whereas NF and to some extent TDD models do. This suggests that DD models will need changes if they are to generate realistically shaped phylogenies. For example, if diversification rates are allowed to vary with both time and trait value, then realistic tree shapes can be obtained (Purvis et al. 2011). These models could also be extended, for example, to make different numbers of niches available to different subclades. Alternatively, NF models might be used; however, these models cannot be fitted to observations of changing diversification rates, because they simply create new species at regular intervals. Useful extensions to the NF model may include diversification rates that change through time, or that vary with the distance in trait-space that a new species must breach. We know what the phylogeny and trait distribution of an adaptive radiation look like, but density dependence alone is not sufficient to generate similar data.

Conclusion

The three diversification models studied here all have a similar theoretical justification in niche filling, but they produce quite different results. This is on top of the wide variance in tree shape for even a single given model, which makes interpretation already difficult (Gascuel et al. 2015).

The only case in which different models produce similarly shaped trees is in the similarity between strong diversity dependence and milder trait-density dependence with an OU process. This is noteworthy, because OU trait evolution suppresses phylogenetic signal. If DD models are justified by the filling up of available niches and, therefore, of trait-space, then they seem to be inconsistent with the commonly assumed Brownian mode of trait evolution. Further, if DD emerges from the packing of continuously evolving traits, then it requires those traits to evolve under a constrained OU-like process. In an OU model, traits are evolving towards a common optimum, causing signal to be erased over time (Hansen 1997, Cooper et al. 2016).

Observed phylogenetic signal tends to be quite high, though on average less than the Brownian expectation, for traits like body size (Freckleton et al. 2002, Blomberg et al. 2013). If diversity dependence models are to represent adaptive radiations, then we must also assume a mode of trait evolution that is quite different from the usual continuous Brownian and OU models.

I have shown how the notion of diversification as a ‘filling up’ of the available niches can be implemented in a number of different ways. By making subtly different assumptions from the same conceptual starting point, three different models of diversification lead to very different results in terms of tree shape and trait distributions. The differences between these models suggest that observations of diversity dependence in real trees should be interpreted with care.

Chapter 6: General conclusion

This project was concerned with the integration of ecological processes into models of trait evolution. Such models are often used to help understand a clade's evolutionary history using a molecular phylogeny. I developed new models to include the effects of interspecific interactions, and tested these models using datasets from a range of animal clades. Both the effects on trait values and on diversification rates were considered, and I studied and modelled each of these effects. Detailed discussions of the results of each chapter are given in each chapter's discussion section, but a brief overview is also provided here.

Summaries of project work

For investigating trait values, I developed a model of interspecific competition driving trait differences in an evolving phylogeny. This is conceptually based on character displacement, but I looked at the whole clade, not just species pairs. This is a complex and nonlinear system, making the model difficult to fit using traditional likelihood methods. I therefore used a Monte Carlo approach, simulating large numbers of datasets and fitting the model via Approximate Bayesian Computation (ABC). Running millions of non-vectorisable simulations required an efficient implementation. The simulations were therefore written in C++ and compiled to a shared object library so that the simulations could then be run from R. The model fitting was then done using custom scripts in R. ABC is reviewed in Beaumont (2010) and Csilléry et al. (2010). It is gaining popularity in ecological research (e.g. Hartig 2011), but has not previously been applied to this problem.

Chapters 3 and 4 cover applications of this model. First, I fitted it to a wide range of freely available animal datasets and phylogenies. This way, I could assess the overall prevalence of clade-wide character displacement patterns, both in clades regarded as adaptive radiations and not. Then, in chapter 4, I use the model, along with other trait evolution models, to investigate a particular group in detail: American hummingbirds. I compared the results for 7 major hummingbird subclades, and interpreted the results in the context of the existing literature around the hummingbirds.

To study the effect of competition on diversification, I contrasted different approaches to the idea of density dependence. This is where diversification in a clade slows as its species become more numerous, since this is conceptually linked to the filling up of available niches. I also created a new density dependence model, this time coupling the differences between species' traits to their speciation rates rather than their trait evolution. I considered the conceptual relations between these different models, and compared the differences in the shapes of phylogenetic trees that they produce, along with the corresponding trait distributions. These analyses were done in R, using existing phylogenetics packages such as *ape* (Paradis et al. 2004).

Findings

Predicted phylogenies and trait data where competition drives trait divergence

Our model of competition-driven trait evolution predicts more even trait distributions across the clade than would arise at random, along with high phylogenetic signal. I found that the trait evolution approach had good power to detect competition-driven trait divergence across a phylogeny, provided the phylogeny had around 12 tips or more. This doesn't assume universal sympatry, but it does assume that the species are actively competing for niches which may include a geographic axis. I found that in adaptive radiations where competition is expected to be strong—e.g. for Darwin's finches, where limiting similarity has long been observed between species pairs—there is evidence of competition-driven trait divergence.

Prevalence of clade-wide character displacement patterns across animal clades

Evidence of clade-wide character displacement in body size does not seem to be prevalent across animal genera. In chapter 3, I considered the clades of 75 mammal genera, 17 lizard genera and 4 cichlid genera. For mammals and lizards, the competition model was preferred to Brownian evolution no more often than would be expected by chance. For cichlids, there was greater evidence of competition, and also the model was significantly preferred for the cichlid clade as a whole. Cichlids are considered an adaptive radiation (Kocher 2004, Takahashi and Koblmüller 2011), so this supports the idea that character displacement plays an important role in adaptive radiation but is not prevalent across other clades. Overall, I found no connection between competition and sympatry; however, this may be due to the large-scale range-based method I used to estimate sympatry.

I used the previously developed competition model, along with other phylogenetic comparative methods, to investigate one clade in greater detail: the American hummingbirds. The results were quite different for body size and beak length, with significant support for character displacement in body size for hermits, but in beak length for gems. The overall hummingbird clade did not have strong evidence of competition, but it did significantly support an early burst trait evolution model, where trait evolution slows down through time (Harmon et al. 2010a). I found that there was lots of variance in these results between hummingbird sub-clades; looking at trait evolution on a variety of phylogenetic scales should be important in future work (e.g. Graham et al. 2012).

Diversification rates and competition for niches

I also investigated the effects of niche-packing density dependent diversification on phylogeny shape and trait distributions. This involved quantifying phylogeny shape and trait distributions for existing models and for a new ‘trait-density dependence’ model. I found that the effects of density dependence on phylogeny shape depend very much on the details of the niche-packing process. This work shows that care must be taken in interpreting signs of density dependence in real phylogenies. I also found that, if traits map directly on to niches, then commonly assumed trait evolution models like Brownian Motion (BM) and the Ornstein-Uhlenbeck process (OU) are typically inconsistent with density dependence diversification models.

Limitations and future directions

The implications and limitations of the modelling approaches I used are considered in detail in the discussion sections of chapters 2, 3, 4 and 5. The main points, however, are briefly considered here. For example, all the approaches assume that the species in a clade have opportunities to interact. They may not be sympatric, but they are in some way competing, even if just for geographic space. This means they are not totally isolated. However, the means of judging which species do and do not have the opportunity to interact is not obvious. This is even more difficult for ancestral species, whose range is unlikely to be known.

A simple means of avoiding this strong assumption is achieved by using a ‘sympatry matrix’, described in Chapter 2 (page 35), where species range data is used to turn interactions on and off. However, this method is coarse-grained and necessarily makes assumptions about ancestral lineages’ interactions when the likelihood is that these will depend strongly on unknown historical details. Also, this approach treats interactions as being either present or absent, whereas in fact the strength of interactions is likely to vary greatly between species pairs and even through time. The approaches in this thesis can only show whether interactions have had a significant overall effect on a clade. Future models will likely benefit from addressing the interactions between geographical dynamics (e.g. Losos and Glor 2003, Pearman et al. 2008, Freckleton and Jetz 2009) and speciation and extinction rates (modelling links to trait values in e.g. Paradis 2005, Freckleton et al. 2008, Fitzjohn 2010, Magnuson-Ford and Otto 2012), as well the competitive interactions have studied here.

Diversification models face different issues. Essentially these models are about categorising differences in diversification rate: differences between lineages result in tree imbalance, while differences through time affect rate-change statistics such as γ (Pybus and Harvey 2000). But in real datasets, lineages are expected to diversify at different rates for many different reasons, often depending on the taxon and trait in question (Chira and Thomas 2016). Therefore, fitting diversification models may be uninformative in terms of detailed ecological processes, or at least difficult to interpret. Also, these analyses are again sensitive to scale (e.g. González-Caro et al. 2012), and this should perhaps be incorporated into studies on diversification mechanisms.

An overall difficulty for any study of competition and character displacement is the mapping of niches onto traits or 'morphospace'. Approaches such as the 'niche filling' models (Price 1997) and the trait competition model presented in chapter 2 assume that there is a close relationship between the two, but actually mapping them onto each other is an important area that has received relatively little focus (e.g. Miles and Ricklefs 1984, Pigot et al. 2016). A further unknown when fitting these models is the adaptive landscape itself: how the available niches are positioned in morphospace. The trait competition model effectively assumes an even distribution of niches, while Price's (1997) niche-filling (NF) model assumes a multivariate normal distribution of niches. It might be possible to include prior knowledge of the true adaptive landscape for a particular clade in future models (Arnold et al. 2001, Aguilée et al. 2013).

Inferring past ecology from present day molecular and trait data is a complex task, with many potentially compounded processes. Understanding the details is an ongoing challenge, integrating a variety of phylogenetic, geographical and community approaches. Here, new tools were developed and novel analyses performed, contributing to the development of this knowledge.

References

Abrams, P. 1983. The theory of limiting similarity. Annual Review of Ecology and Systematics 14:359–376.

Adams, D. C., C. M. Berns, K. H. Kozak, and J. J. Wiens. 2009. Are rates of species diversification correlated with rates of morphological evolution? Proceedings of the Royal Society of London B: Biological Sciences 276:2729–2738.

Aguilée, R., D. Claessen, and A. Lambert. 2013. Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. Evolution 67:1291–1306.

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

Bailey, S. F., J. R. Dettman, P. B. Rainey, and R. Kassen. 2013. Competition both drives and impedes diversification in a model adaptive radiation. Proc. R. Soc. B 280:20131253.

Barraclough, T. G., and A. P. Vogler. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. The American Naturalist 155:419–434.

Bartoszek, K., S. Glémin, I. Kaj, and M. Lascoux. 2017. Using the Ornstein–Uhlenbeck process to model the evolution of interacting populations. Journal of Theoretical Biology 429:35-45.

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

Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. Annual Review of Ecology, Evolution, and Systematics 41:379–406.

Bergmann, P. J., and D. J. Irschick. 2012. Vertebral evolution and the diversification of squamate reptiles. Evolution 66:1044–1058.

Bininda-Emonds, O. R. P., M. Cardillo, K. E. Jones, R. D. E. MacPhee, R. M. D. Beck, R. Grenyer, S. A. Price, et al. 2007. The delayed rise of present-day mammals. Nature 446:507–512.

Bivand, R., and N. Lewin-Koh. 2013. maptools: Tools for reading and handling spatial objects. R package version 0.8 27. Available at: http:// CRAN.R-project.org/package=maptools. Last accessed 1 March 2017

Bivand, R., T. Keitt, and B. Rowlingson. 2014. rgdal: Bindings for the geospatial data abstraction library. R package version 0.8-16. Available at: http:// CRAN.R-project.org/package=rgdal. Last accessed 1 March 2017

Björklund, M. 1997. Are “comparative methods” always necessary? Oikos 80:607–612.

Blackburn, D. C., C. D. Siler, A. C. Diesmos, J. A. McGuire, D. C. Cannatella, and R. M. Brown. 2013. An adaptive radiation of frogs in a southeast Asian island archipelago. Evolution 67:2631–2646.

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

Blum, M. G., O. François, and S. Janson. 2006. The mean, variance and limiting distribution of two statistics sensitive to phylogenetic tree balance. The Annals of Applied Probability 2195–2214.

Bokma, F. 2008. Detection of “punctuated equilibrium” by Bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. Evolution 62:2718–2726.

Bortolussi, N., E. Durand, M. Blum, and O. François. 2006. apTreeshape: statistical analysis of phylogenetic tree shape. Bioinformatics 22:363–364.

Boucher, F. C., and V. Démery. 2016. Inferring bounded evolution in phenotypic characters from phylogenetic comparative data. Systematic Biology 65:651–661.

Burbrink, F. T., X. Chen, E. A. Myers, M. C. Brandley, and R. A. Pyron. 2012. Evidence for determinism in species diversification and contingency in phenotypic evolution during adaptive radiation. Proceedings of the Royal Society B: Biological Sciences 279:4817–4826.

Butler, M. A., and A. A. King. 2004. Phylogenetic comparative analysis: A modeling approach for adaptive evolution. The American Naturalist 164:683–95.

Cavender-Bares, J., K. H. Kozak, P. V. A. Fine, and S. W. Kembel. 2009. The merging of community ecology and phylogenetic biology. Ecology Letters 12:693–715.

Charmantier, A., R. H. McCleery, L. R. Cole, C. Perrins, L. E. B. Kruuk, and B. C. Sheldon. 2008. Adaptive phenotypic plasticity in response to climate change in a wild bird population. Science 320:800–803.

Cheverud, J. M., M. M. Dow, and W. Leutenegger. 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight among primates. Evolution 39:1335–1351.

Chira, A. M., and G. H. Thomas. 2016. The impact of rate heterogeneity on inference of phylogenetic models of trait evolution. Journal of Evolutionary Biology 29:2502-2518.

Clarke, Magnus, Gavin H. Thomas, Robert P. Freckleton. 2017. Trait evolution in adaptive radiations: modeling and measuring interspecific competition on phylogenies. The American Naturalist 189:121-137.

Colwell, R. K., and D. J. Futuyma. 1971. On the measurement of niche breadth and overlap. Ecology 52:567–576.

Connell, J. H. 1980. Diversity and the coevolution of competitors, or the ghost of competition past. Oikos 131–138.

Cooper, N., W. Jetz, and R. P. Freckleton. 2010. Phylogenetic comparative approaches for studying niche conservatism. Journal of Evolutionary Biology 23:2529–2539.

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

Crisp, M. D., and L. G. Cook. 2012. Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? New Phytologist 196:681–694.

Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010 Approximate Bayesian computation (ABC) in practice. Trends in Ecology and Evolution 7:410–18

Davies, T. J., N. Cooper, J. A. F. Diniz-Filho, G. H. Thomas, and S. Meiri. 2012. Using phylogenetic trees to test for character displacement: a model and an example from a desert mammal community. Ecology 93:S44–S51.

Dayan, T., and D. Simberloff. 2005. Ecological and community-wide character displacement: the next generation. Ecology Letters 8:875–894.

De Mazancourt, C., E. Johnson, and T. G. Barraclough. 2008. Biodiversity inhibits species' evolutionary responses to changing environments. Ecology Letters 11:380–388.

Díaz-Uriarte, R., and T. Garland. 1996. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from brownian motion.

Systematic Biology 45:27–47.

Doebeli, M., and U. Dieckmann. 2003. Speciation along environmental gradients. Nature 421:259–264.

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

Dumont, E. R., L. M. Dávalos, A. Goldberg, S. E. Santana, K. Rex, and C. C. Voigt. 2012. Morphological innovation, diversification and invasion of a new adaptive zone. Proceedings of the Royal Society of London B: Biological Sciences 279:1797–1805.

Eastman, J. M., M. E. Alfaro, P. Joyce, A. L. Hipp, and L. J. Harmon. 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. Evolution 65:3578–3589.

Emerson, B. C., and R. G. Gillespie. 2008. Phylogenetic analysis of community assembly and structure over space and time. Trends in Ecology and Evolution 23:619–630.

Etienne, R. S., B. Haegeman, T. Stadler, T. Aze, P. N. Pearson, A. Purvis, and A. B. Phillimore. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. Proceedings of the Royal Society B: Biological Sciences 279:1300–1309.

Feinsinger, P., and R. K. Colwell. 1978. Community organization among neotropical nectar-feeding birds. American Zoologist 18:779–795.

Felsenstein, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. American Journal of Human Genetics 25:471–492.

Felsenstein, J. 1985. Phylogenies and the comparative method. The American Naturalist 125:1–15.

Felsenstein, J. 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. The American Naturalist 171:713–725.

FitzJohn, R. G. 2010. Quantitative traits and diversification. Systematic Biology 59:619–633.

FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods in Ecology and Evolution 3:1084–1092.

FitzJohn, R. G., W. P. Maddison, and S. P. Otto. 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Systematic Biology 58:595–611.

Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrilets. 2008. What, if anything, is sympatric speciation? Journal of Evolutionary Biology 21:1452–1459.

Fordyce, J. A. 2010. Interpreting the γ statistic in phylogenetic diversification rate studies: a rate decrease does not necessarily indicate an early burst. PLoS One 5:e11781.

Freckleton, R. P. 2009. The seven deadly sins of comparative analysis. Journal of Evolutionary Biology 22:1367–1375.

Freckleton, R. P., A. B. Phillimore, and M. Pagel. 2008. Relating traits to diversification: a simple test. The American Naturalist 172:102–115.

Freckleton, R. P., and P. H. Harvey. 2006. Detecting non-Brownian trait evolution in adaptive radiations. PLoS Biology 4:e373.

Freckleton, R. P., and W. Jetz. 2009. Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. Proceedings of the Royal Society B: Biological Sciences 276:21–30.

Freckleton, R. P., M. Pagel, and P.H. Harvey. 2003. Comparative methods for adaptive radiations. Pages 391–407 in T. M. Blackburn and K. J. Gaston, eds. Macroecology: concepts and consequences. Blackwell Scientific, Oxford.

Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. The American Naturalist 160:712–726.

Fritz, S. A., O. R. P. Bininda-Emonds, and A. Purvis. 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. Ecology Letters 12:538–549.

Futuyma, D. J. 2010. Evolutionary Constraint and Ecological Consequences. Evolution 64:1865–1884.

Garland, T., A. F. Bennett, and E. L. Rezende. 2005. Phylogenetic approaches in comparative physiology. Journal of Experimental Biology 208:3015–3035.

Garland, T., P. E. Midford, and A. R. Ives. 1999. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. American Zoologist 39:374–388.

Garland, T., P. H. Harvey, and A. R. Ives. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Systematic Biology 41:18–32.

Gascuel, F., R. Ferrière, R. Aguilée, and A. Lambert. 2015. How ecology and landscape dynamics shape phylogenetic trees. Systematic Biology 64:590-607.

Gillespie, R. G., F. G. Howarth, and G. K. Roderick. 2001. Adaptive radiation. Encyclopedia of Biodiversity 1:25–44.

Gomulkiewicz, R., and D. Houle. 2009. Demographic and genetic constraints on evolution. The American Naturalist 174:218–229.

Graham, C. H., J. L. Parra, B. A. Tinoco, F. G. Stiles, and J. A. McGuire. 2012. Untangling the influence of ecological and evolutionary factors on trait variation across hummingbird assemblages. Ecology 93:S99–S111.

Graham, C. H., J. L. Parra, C. Rahbek, and J. A. McGuire. 2009. Phylogenetic structure in tropical hummingbird communities. Proceedings of the National Academy of Sciences 106:19673–19678.

Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science 296:707–711.

Grant, P. R., and B. R. Grant. 2006. Evolution of character displacement in Darwin's finches. Science 313:224–226.

Grant, P. R., and B. R. Grant. 2011. How and why species multiply: the radiation of Darwin's finches. Princeton University Press, Princeton.

Grant, Peter R. 1972. Convergent and divergent character displacement. Biological Journal of the Linnean Society 4:39–68.

Hadfield, J. D., and S. Nakagawa. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. Journal of Evolutionary Biology 23:494–508.

Hadfield, J. D., B. R. Krasnov, R. Poulin, and S. Nakagawa. 2014. A tale of two phylogenies: comparative analyses of ecological interactions. The American Naturalist 183:174–187.

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

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology 21:1201–1219.

Hansen, T. F., and K. Bartoszek. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. Systematic Biology 61:413–425.

Hansen, T. F., and S. H. Orzack. 2005. Assessing Current Adaptation and Phylogenetic Inertia as Explanations of Trait Evolution:the Need for Controlled Comparisons. Evolution 59:2063–2072.

Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. B. Jennings, K. H. Kozaket al. 2010a. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64:2385–2396.

Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. B. Jennings, K. H. Kozak, et al. 2010b. Data from: Early bursts of body size and shape evolution are rare in comparative data. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.f660p

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hartig, F., J. M. Calabrese, B. Reineking, T. Wiegand, and A. Huth. 2011. Statistical inference for stochastic simulation models – theory and application. Ecology Letters 14:816–827.

Harvey, P. H., and A. Rambaut. 2000. Comparative analyses for adaptive radiations. Philosophical Transactions of the Royal Society B: Biological Sciences 355:1599–1605.

Heatwole, H., and J. A. Taylor. 1987. Ecology of reptiles, 2nd ed. Surrey Beatty & Sons, Chipping Norton, NSW. ISBN: 0949324140

Ho, L. S. T., and C. Ané. 2014. A Linear-Time Algorithm for Gaussian and Non-Gaussian Trait Evolution Models. Systematic Biology 63:397–408.

Hoehna, S. 2013. Fast simulation of reconstructed phylogenies under global, time-dependent birth-death processes. Bioinformatics 29:1367-1374

Ingram, T. 2010. Speciation along a depth gradient in a marine adaptive radiation. Proceedings of the Royal Society B: Biological Sciences 278:613-618

Ingram, T., L. J. Harmon, and J. B. Shurin. 2009. Niche Evolution, Trophic Structure, and Species Turnover in Model Food Webs. The American Naturalist 174:56–67.

Inman, H. F., and E. L. Bradley. 1989. The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. Communications in Statistics - Theory and Methods 18:3851–3874.

IUCN. 2014. IUCN Red List of Threatened Species. Version 2014.2. <www.iucnredlist.org>. Downloaded on 6 November 2014.

Ives, A. R., P. E. Midford, and T. Garland. 2007. Within-species variation and measurement error in phylogenetic comparative methods. Systematic Biology 56:252–270.

Jetz, W., G. H. Thomas, J. B. Joy, K. Hartmann, and A. O. Mooers. 2012. The global diversity of birds in space and time. Nature 491:444–448.

Johnson, M. T. J., and J. R. Stinchcombe. 2007. An emerging synthesis between community ecology and evolutionary biology. Trends in Ecology and Evolution 22:250–257.

Jones, K. E., J. Bielby, M. Cardillo, S. A. Fritz, J. O'Dell, C. D. L. Orme, K. Safi, et al. 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 90:2648–2648.

Jungck, J. R. 2001. BioQUEST Library. Vol. 6. Academic Press.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and Campbell O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, et al. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Butler, M. A., and A. A. King. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. The American Naturalist 164:683-695.

Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nature Reviews Genetics 5:288–298.

Kozak, K. H., and J. J. Wiens. 2010. Accelerated rates of climatic-niche evolution underlie rapid species diversification. Ecology Letters 13:1378–1389.

Kraft, N. J. B., W. K. Cornwell, C. O. Webb, and D. D. Ackerly. 2007. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. The American Naturalist 170:271–283.

Kutsukake, N., and H. Innan. 2013. Simulation-based likelihood approach for evolutionary models of phenotypic traits on phylogeny. Evolution 67:355–367.

Lack, D. 1947. The significance of clutch size. Ibis 89.2:302-352.

Lamichhaney, S., J. Berglund, M.S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, et al. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518: 371-375.

Leimar, O., A. Sasaki, M. Doebeli, and U. Dieckmann. 2013. Limiting similarity, species packing, and the shape of competition kernels. Journal of Theoretical Biology 339:3–13.

Leimar, O., M. Doebeli, and U. Dieckmann. 2008. Evolution of phenotypic clusters through competition and local adaptation along an environmental gradient. Evolution 62:807–822.

Lessard, J.-P., B. G. Weinstein, M. K. Borregaard, K. A. Marske, D. R. Martin, J. A. McGuire, J. L. Parra, et al. 2015. Process-based species pools reveal the hidden signature of biotic interactions amid the influence of temperature filtering. The American Naturalist 187:75–88.

López-Fernández, H., J. H. Arbour, K. O. Winemiller, and R. L. Honeycutt. 2013. Testing for Ancient Adaptive Radiations in Neotropical Cichlid Fishes. Evolution 67:1321–1337.

López-Fernández, H., K. O. Winemiller, and R. L. Honeycutt. 2010. Multilocus phylogeny and rapid radiations in Neotropical cichlid fishes (Perciformes: Cichlidae: Cichlinae). Molecular Phylogenetics and Evolution 55:1070–1086.

Losos, J. B. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecology Letters 11:995–1003.

Losos, J. B., and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. Trends in Ecology & Evolution 18:220–227.

Losos, J. B., T. R. Jackman, A. Larson, K. de Queiroz, and L. Rodríguez-Schettino. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science 279:2115–2118.

Lovette, I. J., and W. M. Hochachka. 2006. Simultaneous effects of phylogenetic niche conservatism and competition on avian community structure. Ecology 87.

Lynch, M. 1991. Methods for the analysis of comparative data in evolutionary biology. Evolution

45:1065–1080.

MacArthur, R., and R. Levins. 1967. The limiting similarity, convergence, and divergence of coexisting species. American naturalist 377–385.

Maddison, W. P., P. E. Midford, and S. P. Otto. 2007. Estimating a Binary Character's Effect on Speciation and Extinction. Systematic Biology 56:701–710.

Magnuson-Ford, K., and S. P. Otto. 2012. Linking the Investigations of Character Evolution and Species Diversification. The American Naturalist 180:225–245.

Mahler, D. L., L. J. Revell, R. E. Glor, and J. B. Losos. 2010. Ecological opportunity and the rate of morphological evolution in the diversification of greater antillean anoles. Evolution 64:2731–2745.

Manceau, M., A. Lambert, and H. Morlon. 2017. A unifying comparative phylogenetic framework including traits coevolving across interacting lineages. Systematic Biology 66:551–568.

Martin, C. H., and P. C. Wainwright. 2011. Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of Cyprinodon pupfish. Evolution 65:2197–2212.

Martins, E. P. 1994. Estimating the rate of phenotypic evolution from comparative data. The American Naturalist 144:193–209.

Martins, E. P., and T. F. Hansen. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. The American Naturalist 149:646–667.

Mattila, T. M., and F. Bokma. 2008. Extant mammal body masses suggest punctuated equilibrium. Proceedings of the Royal Society of London B: Biological Sciences 275:2195–2199.

Mayfield, M. M., and J. M. Levine. 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology Letters 13:1085–1093.

McGuire, J. A., C. C. Witt, D. L. Altshuler, and J. V. Remsen. 2007. Phylogenetic systematics and biogeography of hummingbirds: Bayesian and maximum likelihood analyses of partitioned data and selection of an appropriate partitioning strategy. Systematic biology 56:837–856.

McGuire, J. A., C. C. Witt, J. V. Remsen, A. Corl, D. L. Rabosky, D. L. Altshuler, and R. Dudley. 2014.

Molecular phylogenetics and the diversification of hummingbirds. Current Biology 24:910–916.

Miles, D. B., and R. E. Ricklefs. 1984. The correlation between ecology and morphology in deciduous forest passerine birds. Ecology 65:1629–1640.

Moen, D. S., and J. J. Wiens. 2009. Phylogenetic evidence for competitively driven divergence: body-size evolution in caribbean treefrogs (hylidae: osteopilus). Evolution 63:195–214.

Moen, D., and H. Morlon. 2014. Why does diversification slow down? Trends in Ecology & Evolution 29:190–197.

Monroe, M. J. 2012. Does competition drive character differences between species on a macroevolutionary scale? Journal of Evolutionary Biology 25:2341–2347.

Mooers, A. Ø., and S. B. Heard. 1997. Inferring evolutionary process from phylogenetic tree shape. The Quarterly Review of Biology 72:31–54.

Mooers, A. Ø., Vamosi Steven M., and D. Schluter. 1999. Using phylogenies to test macroevolutionary hypotheses of trait evolution in Cranes (Gruinae). The American Naturalist 154:249–259.

Nee, S., E. C. Holmes, R. M. May, and P. H. Harvey. 1994. Extinction rates can be estimated from molecular phylogenies. Philosophical Transactions of the Royal Society B: Biological Sciences 344:77–82.

Ng, J., and S. D. Smith. 2014. How traits shape trees: new approaches for detecting character state-dependent lineage diversification. Journal of evolutionary biology 27:2035–2045.

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

O'Meara, B. C., C. Ané, M. J. Sanderson, and P. C. Wainwright. 2006. Testing for different rates of continuous trait evolution using likelihood. International Journal of Organic Evolution 60:922–933.

Orme, D., R. Freckleton, G. Thomas, T. Petzoldt, S. Fritz, N. Isaac, and W. Pearse. 2012. Caper: Comparative analyses of phylogenetics and evolution in R. R package version 0.5.

Owens, I. P. F., P. M. Bennett, and P. H. Harvey. 1999. Species richness among birds: body size, life history, sexual selection or ecology? Proceedings of the Royal Society of London B: Biological Sciences 266:933–939.

Pagel, M. 1992. A method for the analysis of comparative data. Journal of Theoretical Biology 156:431–442.

Pagel, M. 1997. Inferring evolutionary processes from phylogenies. Zoologica Scripta 26:331–348.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.

Paradis, E. 1997. Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. Proceedings of the Royal Society of London B: Biological Sciences 264:1141–1147.

Paradis, E. 2005. Statistical analysis of diversification with species traits. Evolution 59:1–12.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Parra, J. L., C. Rahbek, J. A. McGuire, and C. H. Graham. 2011. Contrasting patterns of phylogenetic assemblage structure along the elevational gradient for major hummingbird clades. Journal of Biogeography 38:2350–2361.

Pearman, P. B., A. Guisan, O. Broennimann, and C. F. Randin. 2008. Niche dynamics in space and time. Trends in Ecology & Evolution 23:149 – 158.

Pebesma, E. 2014. Classes and methods for spatio-temporal data in R. Journal of Statistical Software.

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 1289:90–105.

Pennell, M. W., B. A. J. Sarver, and L. J. Harmon. 2012. Trees of unusual size: biased inference of early bursts from large molecular phylogenies. PLOS ONE 7:e43348.

Pfennig, D. W., and K. S. Pfennig. 2010. Character displacement and the origins of diversity. The American Naturalist 176:S26–S44.

Phillimore, A. B., and T. D. Price. 2008. Density-dependent cladogenesis in birds. PLOS Biol 6:e71.

Phillimore, A. B., R. P. Freckleton, C. D. L. Orme, I. P. F. Owens, A. E. J. J. Wiens, and E. J. B. Losos. 2006. Ecology predicts large-scale patterns of phylogenetic diversification in birds. The

American Naturalist 168:220–229.

Pigolotti, S., C. López, E. Hernández-García, and K. H. Andersen. 2010. How Gaussian competition leads to lumpy or uniform species distributions. Theoretical Ecology 3:89–96.

Pigot, A. L., and R. S. Etienne. 2015. A new dynamic null model for phylogenetic community structure. Ecology Letters 18:153–163.

Pigot, A. L., C. H. Trisos, and J. A. Tobias. 2016. Functional traits reveal the expansion and packing of ecological niche space underlying an elevational diversity gradient in passerine birds. Proc. R. Soc. B 283:20152013.

Posada, D., and T. R. Buckley. 2004. Model Selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Systematic Biology 53:793–808.

Price, T. D. 1997. Correlated evolution and independent contrasts. Philosophical Transactions of the Royal Society B: Biological Sciences 352:519–529.

Price, T. D., D. M. Hooper, C. D. Buchanan, U. S. Johansson, D. T. Tietze, P. Alström, U. Olsson, et al. 2014. Niche filling slows the diversification of Himalayan songbirds. Nature 509:222–25.

Purvis, A., S. A. Fritz, J. Rodríguez, P. H. Harvey, and R. Grenyer. 2011. The shape of mammalian phylogeny: patterns, processes and scales. Philosophical Transactions of the Royal Society of London B: Biological Sciences 366:2462–2477.

Pybus, O. G., and P. H. Harvey. 2000. Testing macro–evolutionary models using incomplete molecular phylogenies. Proceedings of the Royal Society of London. Series B: Biological Sciences 267:2267–72.

R Core Team. 2015. R: A language and environment for statistical computing. Vienna, Austria. https://www.R-project.org/

Rabosky, D. L. 2006. Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60:1152–1164.

Rabosky, D. L. 2009. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. Ecology letters 12:735–743.

Rabosky, D. L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution 64:1816–1824.

Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS ONE 9:e89543.

Rabosky, D. L., and E. E. Goldberg. 2017. FiSSE: A simple nonparametric test for the effects of a binary character on lineage diversification rates. Evolution 71:1432-1442.

Rabosky, D. L., and I. J. Lovette. 2008. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution 62:1866–1875.

Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nature communications 4:1958. doi: 10.1038/ncomms2958.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.

Revell, L. J., D. L. Mahler, P. R. Peres-Neto, and B. D. Redelings. 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. Evolution 66:135–146.

Ricklefs, R. E. 2004. Evolution: how do characters evolve? (reply). Nature 432. Schluter, D. 2000. Ecological character displacement in adaptive radiation. The American Naturalist 156:S4–S16.

Schluter, D., and J. D. McPhail. 1992. Ecological character displacement and speciation in sticklebacks. The American Naturalist 140:85–108.

Schoener, T. W. 2011. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. Science 331:426–429.

Sexton, J. P., P. J. McIntyre, A. L. Angert, and K. J. Rice. 2009. Evolution and ecology of species range limits. Annual Review of Ecology, Evolution, and Systematics 40.

Siepielski, A. M., J. D. DiBattista, and S. M. Carlson. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. Ecology Letters 12:1261–1276.

Slater, G. J., L. J. Harmon, D. Wegmann, P. Joyce, L. J. Revell, and M. E. Alfaro. 2012. Fitting models of continuous trait evolution to incompletely sampled comparative data using approximate Bayesian computation. Evolution 66:752–762.

Snodgrass, R. E., and E. H. 1904. Papers from the Hopkins-Stanford Galápagos Expedition, 1898-1899. In Proceedings of the Washington Academy of Sciences, 5:231–372.

Stadler, T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. Proceedings of the National Academy of Sciences 108:6187–6192.

Stiles, F. G. 2008. Ecomorphology and phylogeny of hummingbirds: divergence and convergence in adaptations to high elevations. Ornitologia Neotropical 19:511–519.

Strong Jr, D. R., L. A. Szyska, and D. S. Simberloff. 1979. Test of community-wide character displacement against null hypotheses. Evolution, 33:897–913.

Stuart, Y. E., and J. B. Losos. 2013. Ecological character displacement: glass half full or half empty? Trends in Ecology and Evolution 28:402–408.

Swarth, H. S. 1931. The avifauna of the Galapagos islands. California Academy of Sciences, 18.

Takahashi, T., and S. Koblmüller. 2011. The adaptive radiation of cichlid fish in lake tanganyika: a morphological perspective. International Journal of Evolutionary Biology 2011: 620754.

Thomas, G. H., and R. P. Freckleton. 2012. MOTMOT: models of trait macroevolution on trees. Methods in Ecology and Evolution 3:145–151.

Thomas, G. H., R. P. Freckleton, and T. Székely. 2006. Comparative analyses of the influence of developmental mode on phenotypic diversification rates in shorebirds. Proceedings of the Royal Society B: Biological Sciences 273:1619–1624.

Tobias, J. A., C. K. Cornwallis, E. P. Derryberry, S. Claramunt, R. T. Brumfield, and N. Seddon. 2014. Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. Nature 506:359–363.

Uyeda, J. C., and L. J. Harmon. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Systematic Biology, July, 63:902-918.

Uyeda, J. C., T. F. Hansen, S. J. Arnold, and J. Pienaar. 2011. The million-year wait for macroevolutionary bursts. Proceedings of the National Academy of Sciences 108:15908–15913.

Vamosi, S. M., S. B. Heard, J. C. Vamosi, and C. O. Webb. 2009. Emerging patterns in the comparative analysis of phylogenetic community structure. Molecular Ecology 18:572–592.

Wagner, C. E., I. Keller, S. Wittwer, O. M. Selz, S. Mwaiko, L. Greuter, A. Sivasundar, et al. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular ecology 22:787–798.

Walker, T. D., and J. W. Valentine. 1984. Equilibrium models of evolutionary species diversity and the number of empty niches. The American Naturalist 124:887–899.

Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33:475–505.

Yoder, J. B., E. Clancey, S. Des Roches, J. M. Eastman, L. Gentry, W. Godsoe, T. J. Hagey, et al. 2010. Ecological opportunity and the origin of adaptive radiations. Journal of Evolutionary Biology 23:1581–1596.

Appendices

Appendix 2A: Estimating competition strength and effect width simultaneously

The overlap between species *i* and *j* is proportional to $\Phi\left(-\Delta X_{ij}/\sigma\right)$ where Φ is the cumulative normal distribution. Integrating by parts yields the following approximate function for the overlap between two species:

$$\Phi\left(\frac{-|\Delta X_{ij}|}{\sigma}\right) = 0.5 - \frac{1}{\sqrt{2\pi}} e^{\frac{-|\Delta X_{ij}|^2}{2\sigma^2}} \left[\frac{|\Delta X_{ij}|}{\sigma} + \frac{|\Delta X_{ij}|^3}{3\sigma^3} + \cdots\right] \qquad \text{Equation 2A1}$$

To get the evolutionary rate I multiply this by a, giving:

$$\mathrm{a}\Phi\left(\frac{-|\Delta X_{ij}|}{\sigma}\right) = 0.5a - \frac{a}{\sqrt{2\pi}} e^{\frac{-|\Delta X_{ij}|^2}{2\sigma^2}} \left[\frac{|\Delta X_{ij}|}{\sigma} + \frac{|\Delta X_{ij}|^3}{3\sigma^3} + \cdots\right] \qquad \text{Equation 2A2}$$

At first glance it might appear that changing a and σ would have different effects because the former changes evolutionary rates in a linear manner, whilst the effect of the latter is non-linear. However, if there are a large number of species within a limited niche space, then distances between species will be low, i.e. $\left|\Delta X_{ij}\right|$ is small. Consequently, we can use the following approximation by the Maclaurin series expansion of e^x:

$$e^{\frac{-|\Delta X_{ij}|^2}{2\sigma^2}} \approx 1 - \frac{1}{2\sigma^2}|\Delta X_{ij}|^2$$

Substituting into equation 2A2 and ignoring higher than squared terms we get:

$$\mathrm{a}\Phi\left(\frac{-|\Delta X_{ij}|}{\sigma}\right) \approx 0.5a - \frac{a}{\sqrt{2\pi}}\left(1 - \frac{1}{2\sigma^2}|\Delta X_{ij}|^2\right)\left[\frac{|\Delta X_{ij}|}{\sigma}\right] \approx 0.5a\left(1 - \frac{|\Delta X_{ij}|}{\sigma^2\sqrt{2\pi}}\right) \qquad \text{Equation 2A3}$$

Overall, the rate of evolution is given by the overlap, $2\Phi\left(-\Delta X_{ij}/\sigma\right)$ multiplied by a, yielding:

$$2\mathrm{a}\Phi\left(\frac{-|\Delta X_{ij}|}{\sigma}\right) \approx a\left(1 - \frac{|\Delta X_{ij}|}{\sigma^2\sqrt{2\pi}}\right)$$

This equation is linear in both a and σ^{-2}. Thus from a statistical perspective a and σ will be non-identifiable if the species are interacting strongly. If species are not interacting strongly, i.e. ΔX_{ij} is large, then the data will contain no information on interactions between species and hence it will not be possible to fit the model and we cannot estimate either a or σ.

Appendix 2B: Ornstein-Uhlenbeck model with competition

We based the competition model on BM for reasons detailed in the introduction. However, for the sake of comparison we also simulated under a model that included both OU and character displacement processes. The result is random diffusion with attraction of all lineages to a central ‘optimum’ but repulsion between lineages.

The instantaneous change in the OU model of the trait value x for any given lineage is given in differential form by

$$dx(t) \ = \ -\alpha(x(t) - \psi) \ + \ \sigma \, dW_i(t) \qquad \text{Equation 2B1}$$

where ψ represents the optimum trait value to which lineages are attracted. The ‘OU + competition’ model combines Equation 3 and Equation 2B1, giving

$$dx_i \ = \ -\alpha(x(t) - \psi) + \ a \sum_j S_{ij}(t) e_{ij} \ \ 2\Phi(-|x_i - x_j|) \, dt \ + \ \sigma \, dW_i(t) \qquad \text{Equation 2B2}$$

Appendix 2C: Comparing models

We simulated datasets for a range of random trees, for Brownian motion (BM), Ornstein-Uhlenbeck (OU), Early-burst (EB) and the competition model presented here. Two summary statistics are shown in table 2C.1: a measure, K, of phylogenetic signal, and the standard deviation of differences between ordered tip trait values. These standard deviations are scaled by the mean difference between ordered tip trait values, so that they reflect the shape of the resultant trait distribution rather than its width, which depends primarily on the overall evolutionary rate. Similar results can be obtained for a range of tree sizes; the results given below are the average summary statistics obtained for 100 different simulations, each on a different 100-tip tree.

It can be seen in table 2C.1 that the OU parameter α reduces phylogenetic signal without significantly affecting the variance of trait values, while EB increases signal and increases the trait value variance, and the competition model increases signal and reduces trait value variance.

Table 2C.1: Parameter values and summary statistics for BM, OU, EB and competition models

	σ	a	α	g	K	sd(trait difference)
BM	1	0	0	0	1.025988	1.088189
	2	0	0	0	1.19753	1.155985
	3	0	0	0	0.945132	1.13056
	4	0	0	0	0.989281	1.043405
	1	0	0	0	1.01629	1.144142
OU	1	0	0.5	0	0.625919	1.104392
	1	0	1	0	0.446008	1.14508
	1	0	1.5	0	0.374059	1.150754
	1	0	2	0	0.323968	1.132398
	1	0	2.5	0	0.310235	1.178088
	1	0	3	0	0.299937	1.072554
EB	1	0	0	-0.5	1.452419	1.190265
	1	0	0	-1	2.171459	1.363512
	1	0	0	-1.5	2.761718	1.569579
	1	0	0	-2	4.091478	1.760675
	1	0	0	-2.5	4.143498	2.023122
	1	0	0	-3	4.570352	2.098729
Comp.	1	0	0	0	1.024409	1.223923
	1	0.5	0	0	2.051598	0.801998
	1	1	0	0	2.255849	0.664316
	1	1.5	0	0	2.526787	0.602482
	1	2	0	0	2.597334	0.529352
	1	2.5	0	0	2.699814	0.493765
	1	3	0	0	2.525436	0.476312

Note: σ is the BM rate parameter; a is the strength of the character displacement process in the competition model; α is the strength of attraction to the OU model's 'optimum'; g is the rate change parameter of the EB model. The two summaries of simulated data are the phylogenetic signal K, and the standard deviation of the differences between ordered tip trait values, scaled by the mean difference between these ordered values.

Appendix 2D: Allopatry in Darwin's finches

In the analysis of Darwin's finches presented in the main text, the competition model used the assumption that all species had the opportunity to interact competitively. However, if some species pairs are only found on different islands from one another, then they are allopatric and may have no opportunities for interaction. We therefore repeated the analysis accounting for species pairwise sympatry/allopatry. This information is contained in the 'sympatry matrix' S in the model equation:

$$dx_i \;=\; a \sum_j S_{ij}(t) e_{ij}\;\; 2\Phi(-|x_i - x_j|)\, dt \;+\; \sigma\, dW_i(t) \qquad \text{Equation 2.3}$$

To construct this matrix, we used a dataset that includes morphological, species and island data for 2826 individuals. This dataset was from the BioQuest library (Jungck 2001), and includes data from Snodgrass and Heller (1904), Swarth (1931) and Lack (1947). Each species pair was marked as sympatric if individuals of both species were recorded at any one island, and allopatric otherwise. The resultant matrix S is given in table 2D.1 below. The final datasets used for this analysis, including sympatry matrix, are deposited in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.3sk15 (for paper Clarke et al. 2017).

The analysis results are given in table 2D.2. They are broadly similar to the results in table 2.1, with a slightly greater difference between a significant culmen length and non-significant results for the other traits.

Table 2D.1: Sympatry matrix for Darwin's finches

	fusca	Platyspiza	Pinaroloxias	difficilis	pauper	psittacula	parvulus	pallida	fuliginosa	fortis	scandens	magnirostris	conirostris
fusca	1	1	0	0	0	1	0	0	1	1	1	1	0
Platyspiza	1	1	0	1	1	1	1	1	1	1	1	1	0
Pinaroloxias	0	0	1	0	0	0	0	0	0	0	0	0	0
difficilis	0	1	0	1	0	1	1	1	1	1	1	1	1
pauper	0	1	0	0	1	1	1	1	1	1	1	1	0
psittacula	1	1	0	1	1	1	1	1	1	1	1	1	0
parvulus	0	1	0	1	1	1	1	1	1	1	1	1	0
pallida	0	1	0	1	1	1	1	1	1	1	1	1	0
fuliginosa	1	1	0	1	1	1	1	1	1	1	1	1	1
fortis	1	1	0	1	1	1	1	1	1	1	1	1	1
scandens	1	1	0	1	1	1	1	1	1	1	1	1	0
magnirostris	1	1	0	1	1	1	1	1	1	1	1	1	1
conirostris	0	0	0	1	0	0	0	0	1	1	0	1	1

Note: 1 denotes sympatry, and 0 denotes allopatry.

Table 2D.2: Results for Darwin's finches with partial allopatry

Trait	σ	a	LRTS
Wing length	1.72	0.64	0.94
Tarsus length	1.88	0.68	1.05
Culmen length	1.16	5.12	4.62*
Beak depth	2.08	3.64	1.46
Gonys width	1.84	4.96	1.63

Note: the finch trait dataset is that given in Harmon et al. (2010b). The competition model is compared with the nested BM model. The competition model has one extra parameter compared with the BM model.

Appendix 3A: results for competition model + limits

Table 3A.1: Competition model with limits: estimates and mean sympatry for mammal genera.

Clade	Sigma	a	Likelihood ratio	Sympatry
Stenella	3.6	3.233333	1.766651	NA
Trinomys	0.1	0.833333	1.271742	0.732139
Trachypithecus	3.8	4.333333	1.871537	0.266445
Talpa	3.9	2.866667	1.616459	0.3163
Thomasomys	2.966667	1.066667	1.454605	0.642574
Sundasciurus	1.7	4.066667	1.919317	0.974973
Tupaia	2.366667	4.2	1.357854	0.742125
Abrothrix	3.266667	4.4	1.550806	0.460886
Aotus	2.966667	0.966667	1.465372	0.162371
Sylvilagus	3.766667	1.133333	2.319126	0.448654
Akodon	3.666667	0.7	0.928456	0.401949
Dipodillus	4.4	4	1.646361	NA
Cratogeomys	4.266667	2.2	1.670301	0.151676
Echimys	3.8	2.866667	1.838684	NA
Cercopithecus	2.466667	1.066667	1.939631	0.609196
Chaetodipus	2.266667	1.833333	1.145847	0.317848
Galago	2.533333	2.866667	1.632804	0.414853
Dipodomys	2.933333	4.133333	2.894492	0.441388
Cephalophus	0.666667	2.633333	1.70207	0.633412
Genetta	3.8	4.066667	2.492198	0.590515
Ctenomys	3.566667	3.666667	2.179411	0.363437
Gerbillus	1.5	1.433333	1.27206	0.558156
Cryptotis	2.433333	1.233333	1.629153	0.496652
Chaerephon	2.066667	2.333333	1.188823	0.332483
Artibeus	0.566667	3.733333	2.426456	0.623748
Lagenorhynchus	0.633333	0.666667	1.056957	NA
Makalata	2.666667	3.733333	2.08894	0.505029

Hylopetes	1.7	4	1.550101	0.628476
Mogera	0.5	1.7	1.599588	0.042962
Mesoplodon	4.1	2.033333	2.099837	NA
Marmosa	3.733333	0.666667	1.106698	0.54299
Monodelphis	3.933333	3.566667	2.119344	0.411785
Meriones	4.333333	1.466667	1.415611	0.519683
Mops	3.733333	1.6	1.570416	0.446762
Melomys	3.533333	3.366667	1.55891	0.557428
Maxomys	1.766667	1	1.49568	0.933469
Tamias	2.866667	2.8	1.686982	0.444142
Kerivoula	3.2	1.633333	1.316345	0.641405
Macaca	3.333333	3.1	1.265699	0.420748
Lasiurus	1.1	3.4	1.361457	0.642597
Otomys	3.8	4.4	1.618978	0.341533
Paramelomys	4.333333	1.4	1.46282	0.439175
Phyllomys	1.5	4.433333	2.309888	0.622628
Pseudantechinus	0.9	4.166667	2.015085	0.282593
Oecomys	0.766667	3.366667	2.206531	0.653803
Neotoma	2.633333	3.6	1.972166	0.397958
Ochotona	3.4	2.366667	1.570928	0.471339
Oryzomys	3.466667	0.8	1.091168	0.349467
Proechimys	2.866667	0.9	1.355575	0.479153
Phalanger	4.366667	3.766667	2.277605	0.613949
Murina	4.4	3.033333	1.9557	0.471471
Pseudomys	1.066667	4	2.096022	0.519907
Niviventer	4.333333	3.233333	2.158819	0.528945
Nycteris	4.366667	1.3	1.353137	0.616444
Reithrodontomys	2.8	1.233333	2.132297	0.553157
Rhipidomys	2.4	3.433333	1.135112	0.541145
Saguinus	4.233333	3.633333	0.965606	0.231991
Sminthopsis	2.266667	2.3	1.665739	0.476755
Stenella	3.6	3.233333	1.766651	NA
Oligoryzomys	3.433333	2.833333	2.198648	0.432121
Trinomys	0.1	0.833333	1.271742	0.732139
Trachypithecus	3.8	4.333333	1.871537	0.266445
Sundasciurus	1.7	4.066667	1.919317	0.974973
Thomasomys	2.966667	1.066667	1.454605	0.642574
Talpa	3.9	2.866667	1.616459	0.3163
Tupaia	2.366667	4.2	1.357854	0.742125
Sylvilagus	3.766667	1.133333	2.319126	0.448654
Spermophilus	1.333333	1.833333	1.16088	0.288026
Myodes	0.6	1.5	2.437767	0.599524
Rousettus	3.5	3.6	1.564568	0.848278
Tamias	2.866667	2.8	1.686982	0.444142
Peromyscus	1	3.266667	1.46114	0.506081
Sciurus	1.6	4.133333	1.694363	0.433537

Table 3A.2: Competition model (with limits) estimates and mean sympatry for lizard genera.

Clade	Sigma	a	Likelihood ratio	Sympatry
Ctenophorus	0.633333	2.733333	1.070521	0.883716
Ameiva	3.433333	0.666667	1.068647	0.715558
Egernia	0.666667	4.266667	2.965008	NA
Cnemidophorus	0.133333	1.466667	1.840725	NA
Lerista	2.933333	1.566667	1.528421	0.374923
Liolaemus	2.666667	0.733333	1.346785	0.531074
Mabuya	2.133333	2.5	1.747425	NA
Acanthodactylus	0.833333	0.866667	1.233354	0.536915
Amphiglossus	1.133333	1.9	1.809286	0.672849
Trachylepis	1	2.4	1.919612	1.144031
Chalcides	3.3	0.966667	1.384771	0.349796
Sphenomorphus	1.966667	0.9	1.473225	0.910035
Anolis	1.333333	0.533333	1.256849	0.537528
Emoia	1.4	0.6	0.908681	0.873417
Sceloporus	0.766667	0.5	1.010852	0.449219
Draco	0.333333	2.2	2.178584	0.625564
Varanus	0.933333	2.566667	1.835185	0.497624

Table 3A.3: Results for neotropical cichlids; model with limits.

Clade	Trait	Sigma	a	likelihood ratio
cichlid	bodysize	1.49	3.45	2.02
	gape	4.01	2.19	0.83
cichlasomatini	bodysize	5.60	4.15	1.87
	gape	1.12	4.99	1.98
geophagini	bodysize	1.87	0.65	0.96
	gape	5.69	1.45	0.90
heroini	bodysize	2.85	1.40	1.20
	gape	2.05	1.77	1.17

Appendix 4A

I also performed analyses using the character displacment trait evolution model on both traits simultaneously. The results are shown in table 4A.1 below. Competition is not supported in any of the clades; this is probably due to beak length and body size being correlated, while the model suppresses correlation between traits. A more complete dataset of beak shape, using (uncorrelated) principle components would be an interesting future analysis.

Table 4A.1: character displacement models fitted to hummingbird body size and beak length.

Clade	σ^2	a	LRTS
all	3.68	3.00	1.20
bees	2.72	1.52	1.49
brilliants	4.60	1.28	1.57
coquettes	1.68	0.44	0.87
emeralds	1.28	0.36	1.24
gems	1.64	0.68	0.96
hermits	1.20	0.68	1.68
mangoes	1.68	1.4	1.12

Appendix 5A

Tree size did not affect the relative result from the three tested models. Overall, however, larger trees resulted in greater values of γ and I_c – that is, greater apparent slowdowns and greater imbalance. Representative results for different tree sized are shown in table 5A.1.

Table 5A.1: Summary statistics for model simulations using different tree sizes.

Tips	b	k	μ	α	γ	$sd(\gamma)$	I_c	$sd(I_c)$	K	$sd(K)$
Density dependence:										
100	NA	120	0	NA	-2.34764	1.19332	-0.26084	0.722059	NA	NA
50	NA	60	0	NA	-1.66691	1.081896	-0.12049	0.719663	NA	NA
20	NA	24	0	NA	-1.1135	1.148504	-0.1448	0.55432	NA	NA
Niche filling:										
100	NA	NA	NA	NA	-5.60341	0	11.51796	3.836061	1.459877	0.39711
50	NA	NA	NA	NA	-3.84615	0	5.869109	2.711725	1.360428	0.407361
20	NA	NA	NA	NA	-2.22681	0	2.713199	1.657943	1.254571	0.418132
Trait density dependence:										
100	-0.5	NA	0	0	-3.19561	1.332053	1.084761	1.186597	0.230155	0.093587
50	-0.5	NA	0	0	-2.16324	1.224179	0.422709	0.995987	0.281092	0.173794
20	-0.5	NA	0	0	-1.01211	1.135204	-0.0168	0.674144	0.468933	0.353014