Evolving entities: towards a unified framework for understanding diversity at the species and higher levels

Timothy G. Barraclough*

*Division of Biology and NERC Centre for Population Biology, Imperial College London,
Silwood Park Campus, Ascot, Berkshire SL5 7PY, UK*

Current approaches to studying the evolution of biodiversity differ in their treatment of species and higher level diversity patterns. Species are regarded as the fundamental evolutionarily significant units of biodiversity, both in theory and in practice, and extensive theory explains how they originate and evolve. However, most species are still delimited using qualitative methods that only relate indirectly to the underlying theory. In contrast, higher level patterns of diversity have been subjected to rigorous quantitative study (using phylogenetics), but theory that adequately explains the observed patterns has been lacking. Most evolutionary analyses of higher level diversity patterns have considered non-equilibrium explanations based on rates of diversification (i.e. exponentially growing clades), rather than equilibrium explanations normally used at the species level and below (i.e. constant population sizes). This paper argues that species level and higher level patterns of diversity can be considered within a common framework, based on equilibrium explanations. It shows how forces normally considered in the context of speciation, namely divergent selection and geographical isolation, can generate evolutionarily significant units of diversity above the level of reproductively isolated species. Prospects for the framework to answer some unresolved questions about higher level diversity patterns are discussed.

Keywords: species; speciation; higher taxa; diversity; adaptation

1. INTRODUCTION

Understanding how diversity evolves is a central goal for evolutionary biology, and one that has gathered pace recently thanks to methodological advances. However, current approaches are inconsistent with respect to the treatment of species versus higher level diversity patterns. Most studies on the evolution of diversity rely on the concept of species—discrete genetic and phenotypic units, evolving independently from one another—and the theory of how and why species originate is well known (Coyne & Orr 2004; Gavrilets 2004). However, in most groups, species are still delimited by qualitative methods familiar to Linnaeus (Sites & Marshall 2003; Rieseberg *et al.* 2006; Wiens 2007). Quantitative methods for delimiting species from genetic and other data, based on evolutionary theory, have been developed recently, but most still assume that a simple entity called species exists and focus on a particular signature to delimit those entities (Pritchard *et al.* 2000; Hey *et al.* 2003; Sites & Marshall 2003; Knowles & Carstens 2007). While this may hold true in many cases, it makes an assumption about the true pattern of diversity, namely that it falls into simple, discrete units, that

has rarely been tested (Hutchinson 1968; Coyne & Orr 2004; Rieseberg *et al.* 2006).

Contrast this with the treatment of higher levels of diversity. Prior to the 1960s, the focus was on higher taxa, which are units judged by taxonomists to be distinct based on morphology. Species had special status as arenas of recombination, selection and drift, but higher taxa were used as equally valid units of diversity (Simpson 1953; Stanley 1979; Lee 2003). While still used for convenience, higher taxa are now largely dismissed as arbitrary units of diversity, lacking any special evolutionary significance (Coyne & Orr 2004). Taxa of the same rank may differ substantially in their age and hence cannot be compared as equivalent units (Barraclough *et al.* 1998a). The growth of phylogenetic approaches has led to an alternative paradigm, in which the only evolutionarily significant property of higher groupings is whether they comprise a monophyletic clade or not. Clades are evolutionarily significant in the sense of comprising all living descendants of a common ancestor but no level in the hierarchy is more biologically significant than any other. If higher taxa such as genera appear distinct from one another, it is just because of chance gaps in the branching of the tree of life. A wealth of studies has explored patterns and processes of the origin of species richness (Nee *et al.* 1994; Pybus & Harvey 2000; Barraclough & Nee 2001; Nee 2001; Ree 2005; Rabosky 2006; Ricklefs 2007a), and of diversification in phenotypic traits (Pagel *et al.* 2004; Adamowicz & Purvis 2006; Freckleton & Harvey

*t.barraclough@imperial.ac.uk

2006; Thomas *et al.* 2006; Maddison *et al.* 2007). These efforts have been very successful in generating null models for diversity patterns, but often ineffective at explaining real diversity patterns (Davies *et al.* 2004*a*; Freckleton & Harvey 2006; Ricklefs 2007*a*,*b*).

The different treatment of species and higher taxa reveals a weakness at the heart of current approaches to the evolution of biodiversity. Species are held to be evolutionarily significant units, yet they are mostly delimited by subjective rather than quantitative methods. Only recently have quantitative methods come under the spotlight (Sites & Marshall 2003; Wiens 2007). In contrast, higher levels of diversity are subjected to rigorous quantitative study, but theory that adequately explains the observed patterns has been lacking. At a general level, patterns of diversity in terms of species numbers, the pattern of relatedness among extant species and their trait similarities must depend on the combined processes of natural selection and geographical isolation, mediated by the effects of species interactions or limits on space or resource (Williams 1992; Rosenzweig 1996). These effects must also be mediated via their influence on speciation and extinction. Yet, a general dynamical framework is lacking. Moreover, higher level studies often focus on 'neutral' aspects of diversity, such as identifying isolated populations and rates of cladogenesis, rather than on the role of natural selection in shaping diversity patterns.

I will first review the causes of diversification and the methods for delimiting evolutionarily significant units of diversity from broad systematic samples. I will then argue that some of the processes causing distinctness at the level of species can have similar effects at higher levels, potentially generating distinct evolutionarily significant units at levels above reproductively isolated species. This points towards a framework for evolutionary diversity studies, already advocated by some authors (Rosenzweig 1996; Ricklefs 2007*a*,*b*), based on limits, turnover and adaptation to the environment, rather than on rates. The framework may offer one explanation for the widely observed imbalance in diversity among groups.

2. CAUSES OF DIVERSIFICATION

Consider a set of individuals descended from a common ancestor. After collecting a large sample of individuals from the clade, sequencing a series of genes and measuring a number of morphological characters, distinct clusters are detected in genotype and phenotype space (figure 1). 'Distinct clusters' are taken to imply greater genotypic or phenotypic clustering than expected under a null model of birth, death and mutation within a single species. In reality, the nature of sampling will also affect patterns of variation (discussed further below), but for now assume that the sample includes several individuals from each of any distinct clusters found in the clade, and that individuals are sampled at random from within each cluster. The question then is what processes cause diversification into these distinct forms.

(a) *Reproductive isolation*
In a sexual clade, distinct clusters are expected to evolve if there are sets of individuals that interbreed within sets

but not between sets. This occurs because mutations arising in different reproductively isolated sets cannot be recombined together. Such isolation can be caused by geographical separation or, in sympatry, by isolating barriers such as different mate preferences or post-zygotic isolation. Without reproductive isolation, balanced polymorphism at a single locus or linked loci (Veliz *et al.* 2006) or adaptive divergence between the sexes (Butler *et al.* 2007) could arise, but not distinct forms differing in quantitative traits determined by several unlinked loci. Note that, for clusters to emerge, gene flow need not be zero, but low enough not to counteract other diversifying forces such as divergent selection (Butlin *et al.* 2008; Mallett 2010).

(b) *Independent limitation or demographic non-exchangeability*
This has been identified as a key requirement for strongly distinct clusters to arise in asexual clades, but has often been neglected as a separate issue in sexual clades (Templeton 1989; Barraclough *et al.* 2003). Independent limitation means that, for two or more sets of individuals, the chance of an individual contributing genes to subsequent generations depends only on the contribution of individuals within its own set, not on the contribution of individuals in the other set. It can be caused by geographical isolation or by ecological differences in sympatry. The outcome is that drift and natural selection operate separately in each population (Fisher 1930; Cohan 2001), leading to coalescence within populations and the appearance over time of distinct genetic and phenotypic clusters (Barraclough *et al.* 2003).

Although normally only considered in the context of asexual organisms, independent limitation also acts in sexuals as a distinct process from reproductive isolation (Barraclough *et al.* 2008), as demonstrated by the following thought experiment. Imagine two reproductively isolated sets of individuals, A and B, but in which survival and recruitment occurs irrespective of whether individuals belong to A or B. This situation is equivalent to the reproductively isolated but ecologically neutral species proposed by Hubbell (2001), who also proposed a mechanism for coexistence of such species. Drift, and natural selection affecting neutral loci or those involved in survival and recruitment rather than sexual traits, occurs within an arena defined by the entire clade, not separately within A and B (Etienne & Olff 2004; Etienne *et al.* 2007). Recombination only occurs within A or within B, but gene combinations in A can out-compete gene combinations in B and vice versa. Under this scenario, no distinct genetic or phenotypic clusters are expected to evolve for neutral markers or non-reproductive traits. Only traits or genes directly associated with reproductive characteristics are expected to diverge into distinct clusters.

Individual-based simulation models of three different scenarios confirm these predictions. In the first, all individuals within the clade belong to a single population and variation at a neutral gene reflects the balance between mutation and drift. The expected variation within the population is twice the product
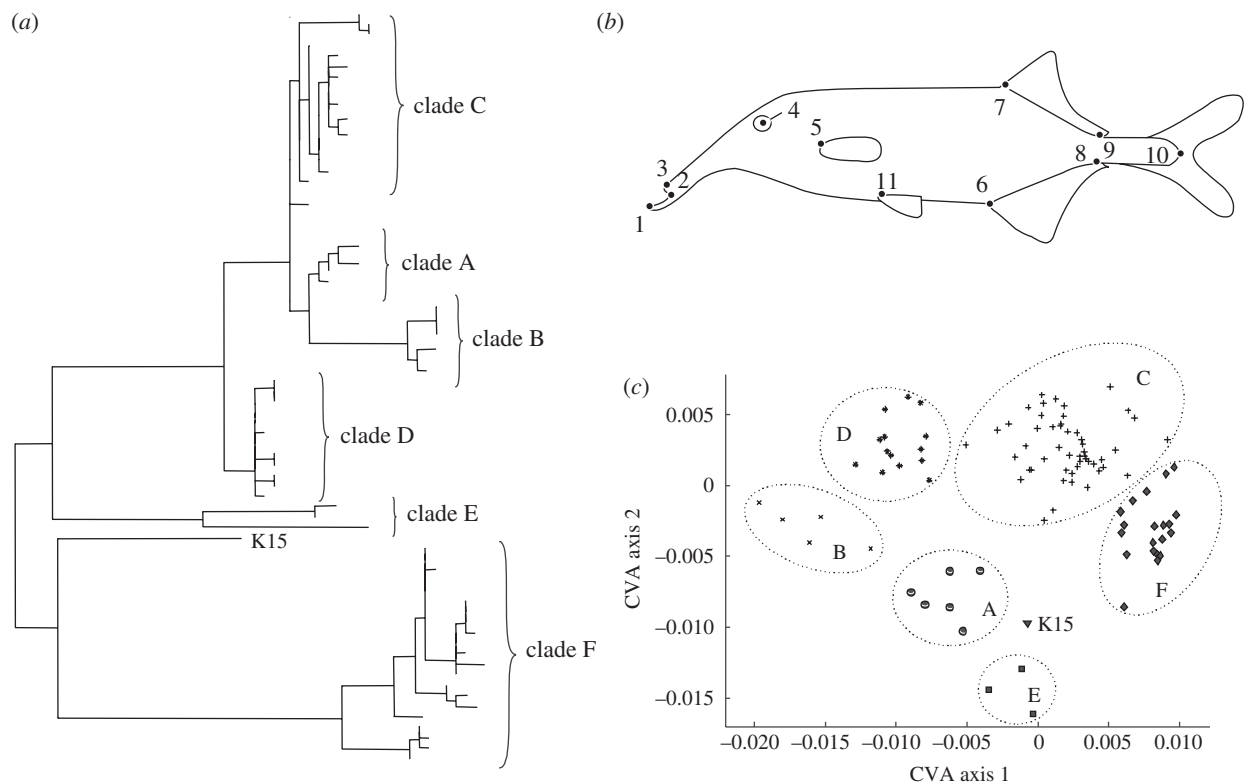
Figure 1. An example of a discrete pattern of genetic and morphological variation taken from a study of African weakly electric fish in the genus *Campylomorphus* (Feulner *et al.* 2007). (*a*) Consensus Bayesian tree based on the mitochondrial cytochrome *b* and nuclear S7 genes, showing the presence of six distinct genetic clusters (A–F) plus a singleton (called K15). Individuals from clade C have been pruned for present illustrative purposes. (*b*) Landmarks used for canonical variates analysis of shape variation. (*c*) A plot of individuals' locations along the first two axes from the canonical variates analysis. Distinct morphological groups are present that correspond to the genetic clusters. Re-drawn with permission.

of the effective population size and the mutation rate (figure 2, black lines). In the second scenario, the clade is split into five reproductively isolated species, each of which is limited independently with a separate effective population size. Species diverge over time (figure 2, dashed lines), leading to the appearance of discrete genetic clusters (figure 3b). In the final scenario, the clade is split into five reproductively isolated species, such that individuals can only mate with other individuals from the same species, but population size is limited at the level of the entire clade. To prevent species richness from declining over time because species are lost by drift, a constant speciation rate per generation is introduced at a level chosen to maintain an average of five species across the course of the simulations. As predicted, distinct genetic clusters do not evolve (figure 3c): the level of genetic variation within the clade is constrained by the same process of turnover as in the first scenario (figure 2, grey lines), albeit acting across a metacommunity of reproductively isolated species.

Although the biological reality of neutral models of metacommunities, and their ability to explain real patterns, is questionable (Ricklefs 2006), these simulations demonstrate that independent limitation, not reproductive isolation, is the cause of genotypic clusters in neutral markers. For example, the observation that single locus sequence variation from broad scale surveys, such as DNA barcoding initiatives, tends to fall into well-defined clusters (Hebert *et al.* 2003, 2004; Acinas *et al.* 2004; Meyer & Paulay 2005;

Pons *et al.* 2006) is not a straightforward consequence of reproductive isolation. Instead, it indicates that independently limited units of diversity are present, either caused by geographical isolation or by ecological distinctiveness in sympatry. Under the alternative scenario that reproductively isolated species are present, but all individuals are ecologically equivalent and interact within the same geographical region (as in Hubbell's models), no such distinctness for non-reproductive genes is expected. For example, new beneficial mutations in a marker might arise and spread, but in a neutral metacommunity *sensu* Hubbell, they would spread or dwindle across the metacommunity as a whole' the reproductively isolated species within which they arose wiping out descendents of all the others if the mutation spreads to fixation. In a scenario in which numbers of individuals are limited jointly across species, clustering would only be expected for genes involved in reproductive differences among the species. Therefore, the only way that the clustering of mitochondrial DNA variation could be a direct consequence of reproductive isolation is if mitochondrial genes themselves played a functional role in reproductive isolation, for example, in post-zygotic isolation (Bolnick *et al.* 2008).

(c) *Divergent selection*
Divergent selection plays an important role promoting the origin of both reproductive isolation and independent limitation (Schluter 2001), but has an additional
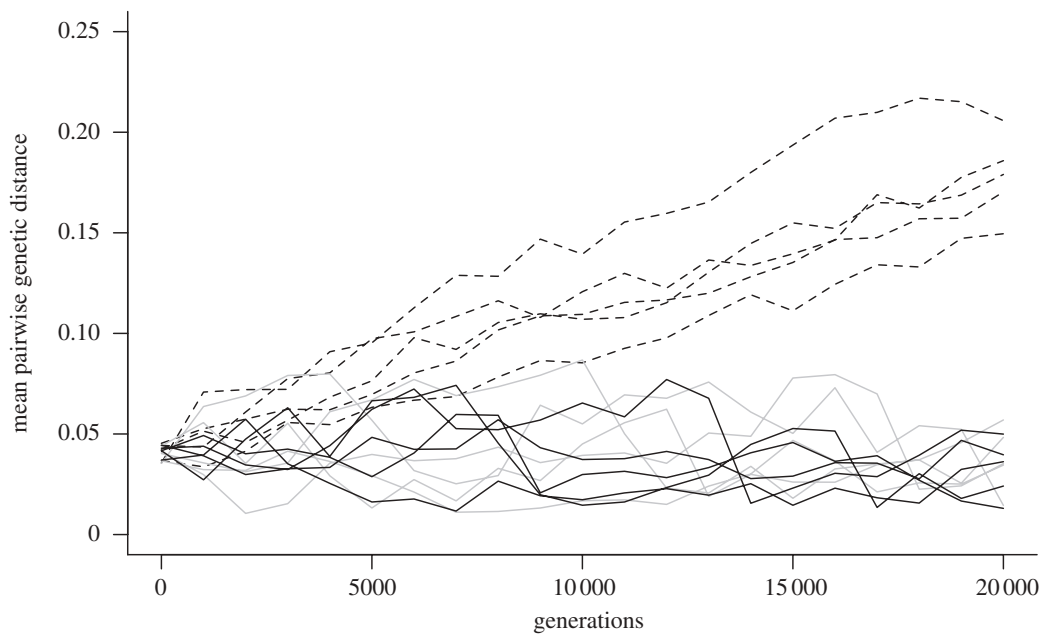
Figure 2. Mean pairwise genetic distance for a neutral haploid DNA sequence marker evolved under three scenarios. Null model (solid lines) = five replicates of a model in which all individuals belong to a single species with constant population size of 5000 individuals. Independent limitation (dashed lines) = five replicates of a model in which an ancestral population is split into five reproductively isolated and independently limited species at time zero, each of constant population size 1000. The mothers of offspring for the next generation were chosen at random with replacement from the 1000 individuals of each species in turn. Joint limitation (grey lines) = five replicates of a model of a metacommunity of 5000 individuals, which is jointly limited but split into five reproductively isolated species at time zero. In this case, the 5000 offspring contributing to the next generation were chosen at random with replacement from mothers irrespective of species membership. The joint limitation model has a constant per generation speciation rate = 0.00178, chosen to maintain an average of five species over time, equivalent to the independent limitation scenario. At each speciation event, a random individual was chosen and its species split into two daughter species (the fraction of individuals assigned to each daughter being chosen from a binomial distribution with probability 0.5). Thus, more abundant species are more likely to speciate (Etienne *et al.* 2007). The simulations ran for 20 000 generations with a mutation rate of 5×10^{-6}. The code is available in the electronic supplementary material.

effect in driving genetic and phenotypic divergence. For example, a clade with geographically isolated sets of individuals has automatic reproductive isolation and independent limitation, yet units of diversity will be more distinct if divergent selection acts on the genes or traits being measured rather than if divergence occurs solely through drift or uniform selection pressures (Fontaneto *et al.* 2007). Once again, divergent selection need not act at the same level as reproductive isolation or independent limitation: selection on a given trait or set of traits could be uniform across a clade of several reproductively isolated species, being divergent only between higher clades (Barraclough *et al.* 2008).

Often, all three conditions will coincide, for example, species isolated in different areas are reproductively isolated and independently limited, but also likely to experience divergent selection. Similarly, ecological divergence entails both divergent selection and independent limitation. The primacy of reproductive isolation in sexual organisms is also clear, as independent limitation and divergent selection cannot cause divergence unless reproductively isolated entities are present or emerge as a consequence of divergent selection (Coyne & Orr 2004; Mallet 2010). However, both independent limitation and divergent selection can operate on broader taxonomic units subsuming several reproductively isolated species. In this case, patterns of

diversity could be more complex than expected under the model that simple species units exist as the basal unit for all diversifying forces. Indeed, the arguments made here could apply to lower levels than species as well: populations might display clustering in some traits. Ultimately, for studies into the fundamental nature of diversification, a framework in which terms such as population, species and higher levels are used to denote units exhibiting idealized evolutionary properties could be replaced with one in which the strength of each diversifying force is quantified across all sampled individuals within a clade, allowing more complex patterns of diversity to be specified directly.

3. SIGNATURES OF DIVERSIFICATION
Various methods have been developed to test for the effects of the processes outlined above, and these could be used to test whether the different processes act on the same units of diversity or at different levels.

(a) *Reproductive isolation*
Crossing experiments or observations of reproductive interactions can be used to survey reproductive isolation, but these are impractical across a large clade and have been limited to problem species or particular groups, such as fungi (Dettman *et al.* 2003). An alternative approach is to infer limits of recombination
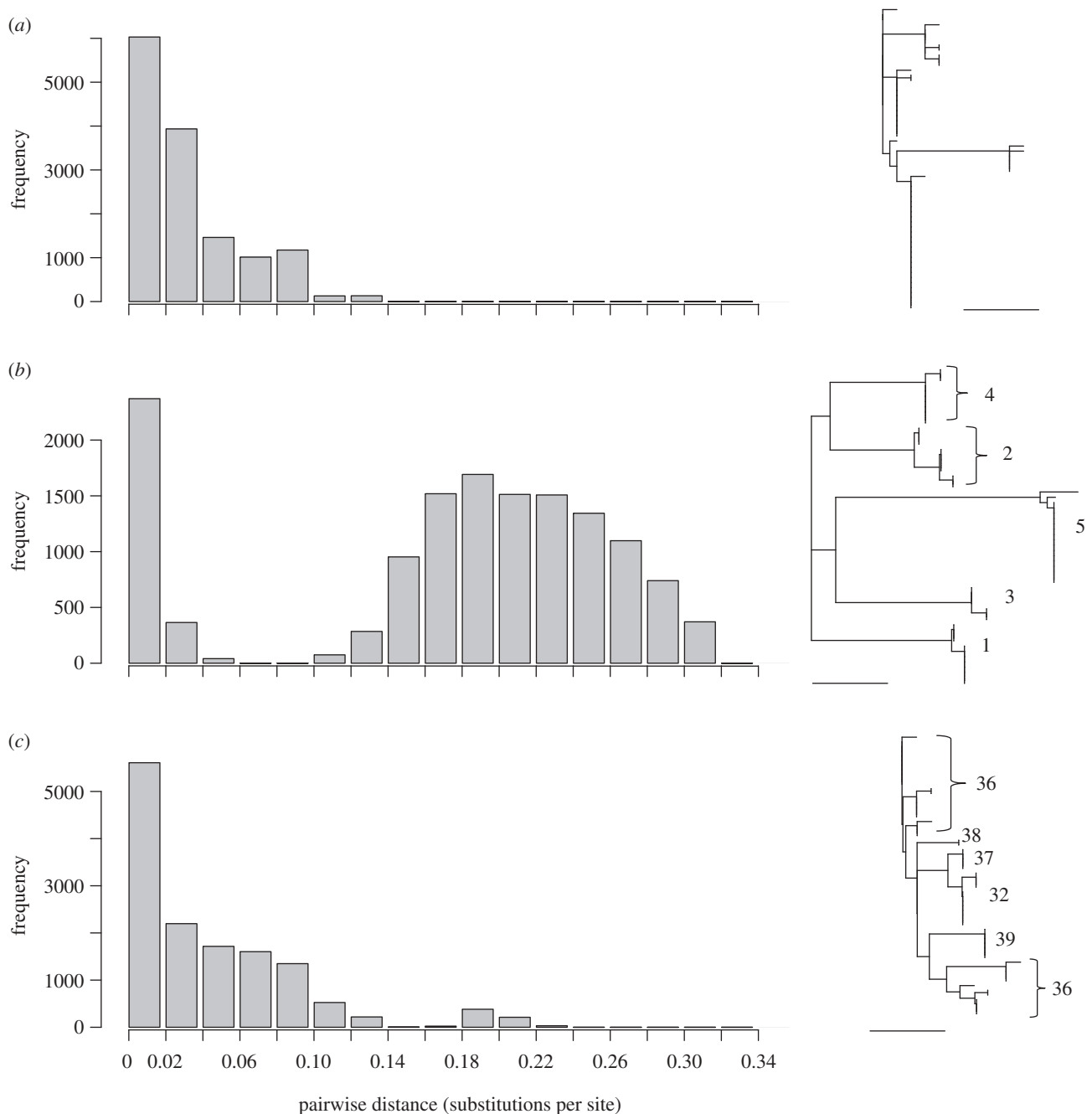
Figure 3. Histograms of pairwise genetic distances pooled from generation 20 000 across five replicates for each of the three scenarios shown in figure 2. The final sampled gene tree from one replicate of each scenario is shown, with species labelled in scenarios (*b*) and (*c*). Scale bars show a genetic distance of 0.05 substitutions per site. Only the model with independent limitation shows evidence for divergence into distinct genetic clusters, apparent as a bimodal plot reflecting intraspecific and interspecific divergences. Species still tend to be monophyletic in the jointly limited model, because of chance extinction of multiple lineages over time, but there is a higher incidence of paraphyly and species are not separated by long internal branches as in the independently limited model. (*a*) Null model, (*b*) independent limitation and (*c*) joint limitation.

using multi-locus DNA data (Koufopanou *et al*. 1997; Hanage *et al*. 2006). A simple approach has been to compare genealogies among different loci and infer-ring the number of recombining populations based on congruence. Within a sexual clade, genealogies among unlinked loci are expected to be discordant within an interbreeding population, but concordant between isolated populations. Hence, a consensus tree can be used to count the number of isolated popu-lations. A major limitation of this approach is the time taken for reciprocal monophyly to evolve: for a set of five nuclear loci, it takes 11.8 times the effective

population size generations before there is a 95 per cent chance of all loci displaying reciprocal monophyly (Hudson & Coyne 2002; Rosenberg 2003), hence many recently diverged populations could go unde-tected. Recently, it has been shown that recently diverged populations can be detected by considering the joint likelihood of alternative species histories given the set of gene trees across loci (Edwards *et al*. 2007; Knowles & Carstens 2007). The set of gene trees can support isolation even if none of them has evolved reciprocal monophyly. At present, this approach has been applied to relatively small datasets

with prior hypotheses for species limits based on geographically or morphologically defined populations. Ideally, it could be extended to cases in which no prior information on limits of reproductive isolation is available, but to optimize the boundaries defining reproductively isolated populations based just on the genetic data themselves (Pritchard *et al.* 2000; O'Meara 2010). Additionally, such methods can be extended to more complex cases, for example, in the presence of rare gene exchange among otherwise non-recombining bacterial groups (Ochman *et al.* 2005; Didelot & Falush 2007).

(b) *Independent limitation*
The signature of independent limitation is that coalescence occurs separately in different populations, leading over time to the appearance of discrete genetic clusters, separated from each other by longer internal branches (Barraclough *et al.* 2003; Acinas *et al.* 2004; Birky *et al.* 2005). Branching within clusters reflects population genetic processes, whereas branching between clusters reflects the timing of events that caused population divergence (Barraclough *et al.* 2003). Various authors have recognized this signature, for example, through identifying gaps in plots of the distribution of pairwise distances from a sample (Higgs & Derrida 1992; Meyer & Paulay 2005) or through applying different thresholds to classify operational taxonomic units (Blaxter *et al.* 2005; Hong *et al.* 2006). Recently, Pons *et al.* (2006) developed a more direct approach that uses likelihood models of branching patterns to compare a null model of no diversification (i.e. a single population) with one of diversification into independently limited populations. The approach allows optimization of the location and number of genetic clusters if the null model is rejected. It can also identify the set of alternative solutions within 95% confidence intervals, i.e. defining uncertainty in the optimal solution. The approach has been used to delimit units of diversity from single locus data of Australian tiger beetles (Pons *et al.* 2006), bdelloid rotifers (Fontaneto *et al.* 2007), South African plants and Costa Rican orchids (Lahaye *et al.* 2008), and bacteria (Barraclough *et al.* 2009).

(c) *Divergent selection*
The detection of divergent selection has received considerable attention in the population genetics literature, but has been neglected as a source of information for delimiting of units of diversity. Robust methods exist for identifying divergent selection acting on protein-coding DNA sequences (Nielsen 2005). For example, the McDonald & Kreitman (1991) test compares the ratio of polymorphism (i.e. within populations) and divergence (i.e. between populations) of coding changes with the same ratio for silent changes (assumed to be neutral). An excess of amino acid divergence relative to polymorphism is indicative of divergent selection. Uniform negative selection against coding variation across the entire sample is indicated by a low amino acid divergence relative to polymorphism.

In principle, this approach could be extended to delimit units of diversity experiencing divergent selection on sampled loci, for example, by fitting models allowing the ratio of coding to silent changes (dn/ds ratio) to vary within and between different units and selecting the optimal allocation yielding the highest likelihood score. However, although many widely used DNA markers are protein coding, they are typically genes chosen to be variable markers, rather than because of any particular interest for their role in diversification. For example, cytochrome oxidase I plays a role in metabolism that could, in principle, be a target for divergent selection, either as a target for thermal adaptation or perhaps in epistasis with nuclear genes (Bolnick *et al.* 2008). However, overall, mitochondrial DNA appears to be under uniform, purifying selection in most clades (Barraclough *et al.* 2007; Meiklejohn *et al.* 2007). Although it is now possible to uncover candidate genes behind key traits, such as those behind resource use or mate choice, these remain hard if not impossible to sample at a broad systematic scale.

An alternative is to use morphology or other aspects of phenotype to survey traits with presumed roles in resource use or mate choice. In the past, morphology has been used for species delimitation in a descriptive way, rather than to test evolutionary hypotheses. Also, where morphology and genetic data have been sampled together, they are typically used for cross-corroboration, rather than integrated analyses. However, the same rationale as the McDonald–Kreitman test has been used to test for selection on quantitative traits among wild populations (Leinonen *et al.* 2008), by comparing Qst (the proportion of variation in a quantitative trait occurring among populations) and Fst (the proportion of variation in neutral genetic markers occurring among populations). A similar approach could be applied at taxonomic scales, either testing for low rates of phenotypic change within pre-defined units versus between them, using genetic markers to provide a baseline, or delimiting units of adaptive diversity by identifying units experiencing uniform selection within them but divergent selection between them. Fontaneto *et al.* (2007) used this approach to show that bdelloid rotifers, a renowned group of asexual animals, had diversified through divergent selection on their feeding morphology. In this case, comparison of wild and laboratory-reared animals confirmed strong heritability of the study trait. Direct evidence of heritability is unlikely to be available for many systematic studies, an important limitation for this approach. Nonetheless, finding uniform phenotype within a set of individuals but divergence between different sets, with reference to a baseline provided by genetic data, would be an important step to uncovering the pattern of adaptive diversity within clades.

For all of these methods, sampling is an important issue. In principle, incomplete sampling of partially isolated populations, for example, from extreme ends of a species range, might yield the appearance of distinct clusters that would break down if intermediate populations had been sampled. In some applications, such as delimiting species from a single environmental

sample such as a sample of sea water (Venter *et al.* 2004), this problem does not apply. In other cases, apart from ensuring a broad geographical sample from across putative species ranges, conclusions need to be conditioned on the nature of the sample. Because biological variation, and our ability to sample it, is not distributed randomly, it is hard to derive model-based approaches to deal with sampling issues. One approach might be to develop simulation models of alternative possible biases and to explore their effect on inferred patterns, similarly to approximate Bayesian computation (ABC) methods developed for population genetics (Beaumont *et al.* 2002; see also Rabosky & Lovette 2008).

4. HIGHER LEVEL DIVERSITY PATTERNS

Methods for inferring the patterns and processes of evolution above the species level have focused on statistical models describing stochastic processes, such as the branching of a phylogenetic tree or neutral trait evolution, in terms of the underlying rates of those processes (Nee *et al.* 1994; Pagel 1999, 1997). The benefits of this approach include the wealth of statistical theory available for such processes and the ease of generating simple null models against which to compare real patterns of diversity. However, at the core of these models is the assumption that present day diversity patterns, in terms of numbers of species and their characteristics, are determined by non-equilibrium processes: diversity patterns depend on *rates* of cladogenesis and trait divergence (Ricklefs 2007*a*). While this does not prohibit the use of such methods as statistical tools, research has often focused on non-equilibrium explanations as well. For example, the observation that phylogenetic trees tend to be imbalanced, such that sister clades often differ greatly in their diversity and more so than expected under simple birth–death models (Blum & Francois 2006), led to the search for traits or key innovations that cause faster diversification rates in some lineages than in others. While some traits have been found that do appear to speed up diversification rates (Barraclough *et al.* 1998*b*), on the whole the success of non-equilibrium models in explaining diversity patterns is extremely limited. Models that include ecological characteristics have proved much more successful (Davies *et al.* 2004*b*, 2005; Phillimore *et al.* 2006) and analyses of branching rate often support a density-dependent mode of diversification rather than exponential growth (Harmon *et al.* 2003; Weir 2006; Phillimore & Price 2008; Rabosky 2009). Together with studies of ecological trait variation in adaptive radiations—clades that have diversified rapidly in terms of both species numbers and ecological niches occupied (Losos *et al.* 1998; Pinto *et al.* 2008; Thomas *et al.* 2009)—these results argue for a conceptual framework based on ecological limits and equilibria, rather than one focused on rates of cladogenesis (Ricklefs 2007*b*).

A simple change is to consider diversification above the species level in terms of independent limitation and divergent selection as outlined above. First, assume that the number of species is limited, rather than growing infinitely (Ricklefs 2007*a*,*b*). This seems a reasonable assumption for a finite world and could arise in two ways. First, clades occupy finite geographical regions, mostly of a continental or lower scale. We know from species-area curves that geographical area places a limit on species numbers (Rosenzweig 1996). Second, although ecological characteristics can vary among closely related species, clades often occupy a recognizable broadly defined niche. For example, in mammals, carnivores display varying degrees of carnivory with just one herbivorous species, the panda, whereas most artiodactyls are herbivores. Therefore, within a given geographical region, different clades may face limits on the availability of resources within their broadly defined niches. Of course, resource use does not always match so neatly with clade membership, but, where it does, there is the potential for those limits to generate higher level units of diversity (explained further below).

A second feature of this modified framework is turnover. While it is possible that clades could grow to fill their available space and then stagnate with no further speciation or extinction, more realistic is that they would persist in a dynamic equilibrium with continued turnover in species (Rosenzweig 1996; Ricklefs 2007*a*,*b*). Every species has a finite chance of extinction and, when extinction does occur, it leaves space to be filled by new species formed by speciation.

The combination of limitation and turnover has interesting consequences for the pattern of diversity, equivalent to processes outlined for diversification at the species level. For example, imagine a clade occupying a given continental region, A (figure 4). At some point, a member of the clade colonizes a new region, B. Assuming that sufficient resources are available, the clade diversifies to occupy available space in the new region. Initially, the clade in region B will be nested within the clade in region A (figure 4*b*). However, after a sufficient period of time (determined by the rate of species turnover and the equilibrium number of species), assuming no subsequent dispersal, all of the species in region A will be descended by chance from a common ancestor more recent than their ancestor with those from region B (figure 4*c*), and vice versa, in a manner equivalent to coalescence at the individual level between isolated populations. This occurs because the limit on the number of species operates separately in the two regions. In addition, the date of the node separating the two clades will tend to pre-date the actual dispersal event, because of standing diversity of species in region A at the time of dispersal, in the same way that dates of speciation events from gene trees are overestimates unless population variation is taken into account (Edwards & Beerli 2000).

Individual-based simulation models confirm the prediction that independent limitation, caused by geographical isolation, causes evolutionary independence and clustering above the level of reproductively isolated species. The third scenario simulated previously, of species turnover within a metacommunity, was repeated but this time defining two independently limited metacommunities, representing separate geographical regions. In this scenario, genetic variation within each region is again constrained by the turnover
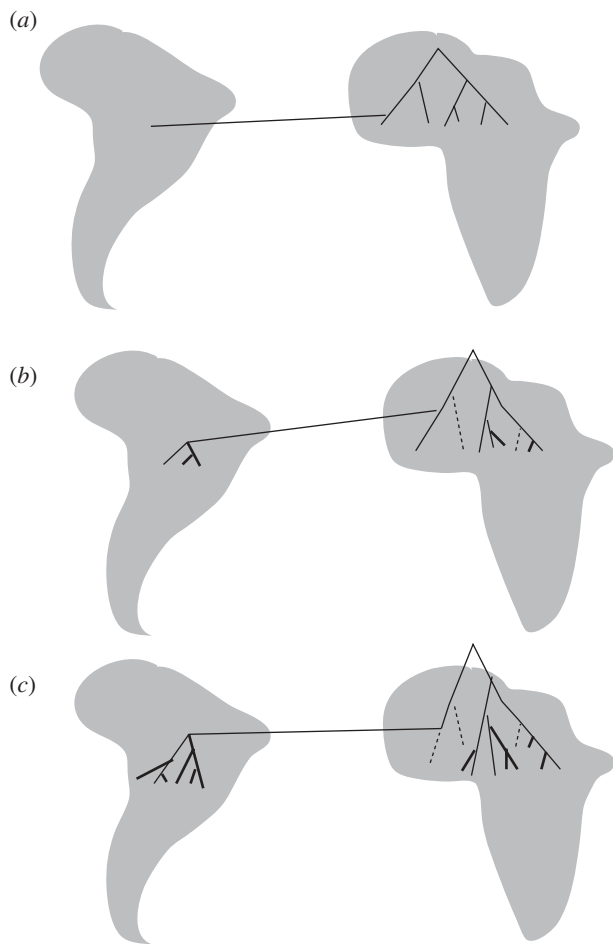
Figure 4. Hypothetical example showing the consequences of species turnover within isolated geographical regions. (*a*) Individuals from Africa colonize South America ($t = 0$). (*b*) Speciation occurs in South America to fill available space (thick lines); speciation (thick lines) and extinction (dotted lines) occurs in Africa, i.e. species turnover ($t = 10$). (*c*) After a sufficient period of time, turnover leads to reciprocal monophyly of species between the two regions. Note the divergence time between these two clades at the end depends on the time elapsed since the dispersal event *plus* the divergence at the time of the dispersal event between the colonizer of South America and the ancestors of those lineages surviving to the present in Africa ($t = 20$). In this case, by chance the final date for the divergence time is more than twice the actual time since the dispersal event.

process within each region, but genetic divergence occurs between the regions (figure 5*a*), producing a signature of distinct genetic clusters above the level of species (figure 5*b*). Note that non-neutral models would be expected to generate similar patterns, as long as the conditions of limits on species numbers and turnover within regions were met.

A possible example of this scenario is the case of sister clades in South America and Africa. Such cases were often interpreted as evidence of vicariance, but many have been re-evaluated in the light of evidence that nodes are too recent to match the break-up of the two continents (Lavin *et al.* 2004). Instead, sister clades may reflect a process of very rare dispersal followed by coalescence within each region owing to species turnover. A similar scenario could occur with the invasion of a new, broadly different

niche by one species, leading in time to two distinct clades in sympatry if species numbers within each niche are limited independently (Lavin *et al.* 2004; but see Alroy 1996). In effect, these are evolutionarily significant higher taxa: their members share not only evolutionary history but interact in a way that maintains cohesion of relatedness and phenotypic traits.

One process that could enhance the coalescence at the level of clades is turnover driven by environmental change (Jablonski & Sepkoski 1996; Barnosky *et al.* 2003; Mayr 2005; Jaramillo *et al.* 2006). Consider a regional assemblage of species adapted to particular conditions, say mesic conditions. Upon a general drying of the climate, there will be selection to adapt to the new conditions, for example, an increase in sclerophylly or succulence within the flora. This might promote evolutionary responses in parallel among all of the species present (Pulido & Berthold 2004). However, alternatively, some species might be pre-adapted to the new conditions (either generalist species or those living in habitats most similar to the new environment) and those species might thrive and diversify at the expense of others. If so, this could enhance species turnover, causing all species in the region at a later time to be descended from just one or a few species present prior to environmental change (Tilman & Lehman 2001). This process is equivalent to a selective sweep but through replacement of genotypes rather than just the spread of a single gene region within a recombining population.

While it might seem far-fetched to extend population genetics to the scale of entire faunas and floras, there is good reason to expect that species assemblages should respond to change in this way. Much of the genetic and phenotypic variation within an assemblage is distributed among species, rather than within them; therefore ecological changes in abundance and distribution should occur more readily than evolutionary changes within species (Ackerly 2003; de Mazancourt *et al.* 2008; Johansson 2008). For example, Crisp *et al.* (2004) showed that many taxa in Australia comprise relatively recent radiations descended from long stem branches. They argued that this reflects the signature of species turnover associated with the shift from an aseasonal wet biome to an arid biome and the reversal of fortunes between clades adapted to mesic conditions and those suited to arid environments.

To conclude, the consequences of the modified framework for patterns of diversity, which assumes that numbers of species are limited and that turnover occurs within limited groups of species, are twofold. First, as long as rates of dispersal between geographical regions or shifts between broad niches are sufficiently low, then limitation of species numbers and turnover within sets of species sharing the same limit can create distinct units of diversity at levels above the species. This will be apparent as distinct clades separated by relatively long branches in the phylogeny (Lavin *et al.* 2004). We also expect traits reflecting use of the broad niche to be conserved within the distinct clades but strongly divergent between them (cf. Ackerly *et al.* 2006; Silvertown *et al.* 2006; Ackerly & Cornwell 2007).
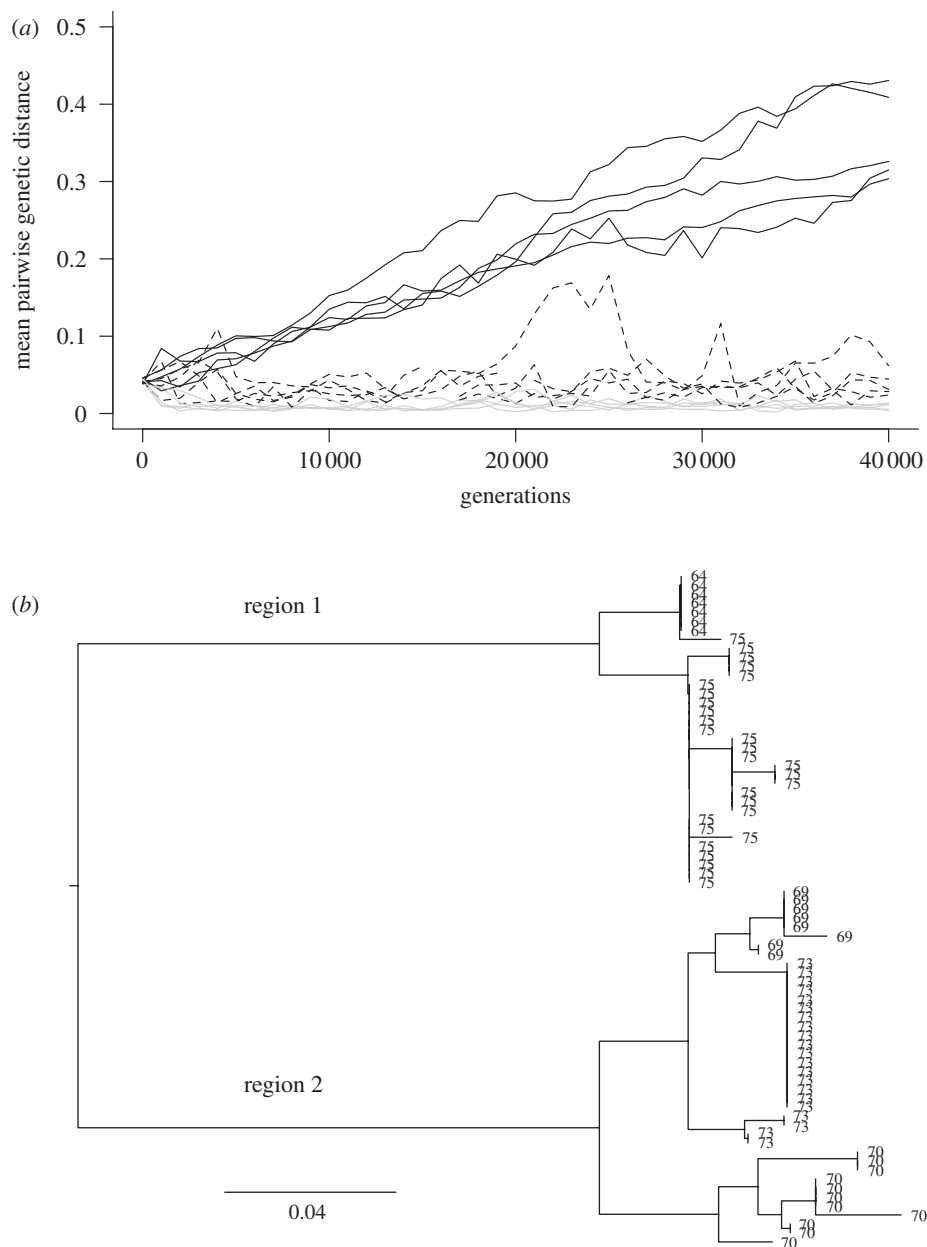
Figure 5. (a) Pairwise genetic distances between two independently limited regions (solid black lines), between jointly limited species within regions (dashed lines) and within species (grey lines) evolved across five independent replicates. Both regions had constant population size of 2500 and the mothers of offspring for the next generation were chosen at random with replacement from individuals in the two regions in turn, but irrespective of species membership. (b) Sampled gene tree at the end of the simulation from one replicate, with species labelled by numbers. Clades corresponding to individuals from the two regions are indicated. Simulations ran for 40 000 generations, starting with four species present, and a speciation rate of 0.00178.

Second, a process of turnover driven by ongoing environmental change, in which a limited set of species survive and contribute to radiations in subsequent new environmental conditions, will create a pattern of imbalance among clades at any snapshot in time. The diverse clades are those thriving in current (or recently past) conditions, but only a few species in those clades will thrive in future environments and foster new radiations. In other words, imbalance is a signature of the waxing of groups that are either pre-adapted to new conditions or lucky to be in the right place, and the waning of groups remnant from previous times but poorly adapted to current conditions or trapped in the wrong place, e.g. a small geographical region (Davies *et al.* 2004a). Possible examples

include the pattern of diversity among recent radiations into arid environments and 'relict' mesic taxa in Australia (Crisp *et al.* 2004; Rabosky *et al.* 2007) and the imbalance among flowering plant families, with recent radiations such as grasses being linked to global cooling and drying (Davies *et al.* 2004a).

The pattern of diversity will depend on the number of dispersal events between different geographical areas and on the frequency of shifts between different broad niches. Similarly to the effects of migration in population genetics, there might be a threshold of dispersal below which higher level clustering emerges. For example, if evolutionary shifts between mesic and xeric plants were frequent, then the entire flora would be limited as a single unit and no higher level

patterns of diversity would emerge, beyond those expected through turnover within a single interacting species assemblage. However, if switches are rare, and there are separate limits on the number of species in the two habitats, then we would expect distinct clades to emerge. In addition, a hierarchical scale of limitation is likely: species may share a broad niche yet have diversified to occupy distinct finer niches or space within that. Thus, significant discreteness may be found at higher levels for some niche traits but not others. Therefore, characterizing the pattern of diversity shifts towards looking at the 'shape' of diversity patterns in terms of phylogenetic relatedness and phenotypic similarity: do higher level units of diversity exist and, if so, what can we infer regarding the dynamics of the processes leading to their formation?

5. TESTING FOR HIGHER LEVEL UNITS OF DIVERSITY

Methods already developed for studying diversity patterns can be readily applied in this new framework. For example, to test for distinct taxa indicative of independent coalescence occurring at higher levels, we can retain the null model of random birth and death (Nee *et al.* 1994), but test against an alternative model of significant clusters, equivalent to the one used for diversification at species level (Pons *et al.* 2006). If ecological and geographical limits act independently of clade membership, or if no such limits are in effect, then no such clustering is expected beyond levels under the null model. If there are hierarchical levels of evolutionary independence, more complex models could be devised; for example, discriminating between branching processes within populations, between populations within species and between species within evolutionary guilds.

Another possible approach is to test for evolutionary guilds of species sharing ecological characteristics and traits. This equates to identifying units within which there is uniform selection on traits but between which there is divergent selection leading to different phenotypic optima. Intuitively, we know this, in terms of phylogenetically conserved versus variable traits (Owens & Bennett 1995), but now we can test patterns of trait diversity against suitable null models of constant rates: are conserved traits just evolving slowly across the entire tree or are they constrained within evolutionary guilds narrower than the entire sample? Gaining answers will not be trivial—as with all observational data, the same pattern often has several alternative and equally plausible explanations (Barraclough & Nee 2001). However, different broad scenarios do generate predictions that are distinguishable given sufficient data.

At present, large datasets that are complete at the level of species are rare. There is still no species-level tree for all birds and the complete mammalian supertree lacks empirical estimates of dates for 37 per cent of nodes (Bininda-Emonds *et al.* 2007), not a problem for most uses but limiting the application of some tests outlined above. In addition, ideally analyses should go beyond assumptions of traditional taxonomic species and incorporate individual sampling allowing objective

tests of the pattern of diversification. DNA barcoding initiatives offer the best prospects for achieving this goal, with comprehensive sampling of individuals across broad samples of species from higher taxa (Hebert *et al.* 2003, 2004; Meyer & Paulay 2005). While some authors have criticized the reliance on a single locus (Will & Rubinoff 2004), single-locus data can be used to address many questions concerning the shape of diversification patterns. Although preferable where feasible, multi-locus data are currently not attainable on the same taxonomic scale as single locus data in most taxa. One strength of barcoding initiatives is the level of curation of samples: additional loci or morphological measurements could be readily assembled in future for a chosen sample of individuals curated within their data system (Ratnasingham & Hebert 2007). A combination of a robust deeper level phylogenetic framework, with comprehensive samples at finer levels extending below the traditional species level, offers the potential for the kinds of tests outlined here to be applied at unprecedented scales.

6. CONCLUSIONS

Higher level diversity can be considered in the same framework as species divergence. In particular, shifting the emphasis from 'non-equilibria' and 'rates' to 'equilibria' and 'limits' leads to a radical change in the predictions concerning diversity patterns. The same processes, such as geographical isolation and divergent selection, can exert their effects at multiple scales. In principle, the net effect could be to generate complex diversity patterns in which hierarchical levels of evolutionary independence and clustering are present, preventing simple assignment of populations, species and higher levels. Alternatively, the different processes might tend to coincide and generate simple patterns consistent with the traditional view of simple species units (Rieseberg *et al.* 2006). Standard opinion suggests that the former scenario might apply in many prokaryote groups, whereas the latter will apply in many eukaryote groups, especially well-studied vertebrate, plant and invertebrate taxa. However, data that comprehensively distinguish these scenarios are currently lacking. The approaches outlined here offer the potential for deeper insights into the causes of diversification.

REFERENCES

Acinas, S. G., Klepac-Ceraj, V., Hunt, D. E., Pharino, C., Ceraj, I., Distel, D. L. & Polz, M. F. 2004 Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**, 551–554. (doi:10.1038/nature02649)

Ackerly, D. D. 2003 Community assembly, niche conservatism, and adaptive evolution in changing environments. *Int. J. Plant Sci.* **164**, S165–S184. (doi:10.1086/368401)

Ackerly, D. D. & Cornwell, W. K. 2007 A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecol. Lett.* **10**, 135–145. (doi:10.1111/j.1461-0248.2006.01006.x)

Ackerly, D. D., Schwilk, D. W. & Webb, C. O. 2006 Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology* **87**, S50–S61. (doi:10.1890/0012-9658(2006)87[50:NEAART]2.0.CO;2)

Adamowicz, S. J. & Purvis, A. 2006 From more to fewer? Testing an allegedly pervasive trend in the evolution of morphological structure. *Evolution* **60**, 1402–1416.

Alroy, J. 1996 Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 285–311. (doi:10.1016/S0031-0182(96)00100-9)

Barnosky, A. D., Hadly, E. A. & Bell, C. J. 2003 Mammalian response to global warming on varied temporal scales. *J. Mammal.* **84**, 354–368. (doi:10.1644/1545-1542(2003)084<0354:MRTGWO>2.0.CO;2)

Barraclough, T. G. & Nee, S. 2001 Phylogenetics and speciation. *Trends Ecol. Evol.* **16**, 391–399. (doi:10.1016/S0169-5347(01)02161-9)

Barraclough, T. G., Nee, S. & Harvey, P. H. 1998a Sister-group analysis in identifying correlates of diversification—comment. *Evol. Ecol.* **12**, 751–754. (doi:10.1023/A:1017125317840)

Barraclough, T. G., Vogler, A. P. & Harvey, P. H. 1998b Revealing the factors that promote speciation. *Phil. Trans. R. Soc. Lond. B* **353**, 241–249. (doi:10.1098/rstb.1998.0206)

Barraclough, T. G., Birky, C. W. & Burt, A. 2003 Diversification in sexual and asexual organisms. *Evolution* **57**, 2166–2172.

Barraclough, T. G., Fontaneto, D., Ricci, C. & Herniou, E. A. 2007 Evidence for inefficient selection against deleterious mutations in cytochrome oxidase I of asexual bdelloid rotifers. *Mol. Biol. Evol.* **24**, 1952–1962. (doi:10.1093/molbev/msm123)

Barraclough, T. G., Fontaneto, D., Herniou, E. A. & Ricci, C. 2008 The evolutionary nature of diversification in sexuals and asexuals. In *Speciation and ecology* (eds R. Butlin, J. Bridle & D. Schluter). Cambridge, MA: Cambridge University Press.

Barraclough, T. G., Hughes, M., Ashford-Hodges, N. & Fujisawa, T. 2009 Inferring evolutionarily significant units of bacterial diversity from broad environmental surveys of single-locus data. *Biol. Lett.* **5**, 425–428. (doi:10.1098/rsbl.2009.0091)

Beaumont, M. A., Zhang, W. & Balding, D. J. 2002 Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025–2035.

Bininda-Emonds, O. R. P. *et al.* 2007 The delayed rise of present-day mammals. *Nature* **446**, 507–512. (doi:10.1038/nature05634)

Birky, C. W., Wolf, C., Maughan, H., Herbertson, L. & Henry, E. 2005 Speciation and selection without sex. *Hydrobiologia* **546**, 29–45. (doi:10.1007/s10750-005-4097-2)

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R. & Abebe, E. 2005 Defining operational taxonomic units using DNA barcode data. *Phil. Trans. R. Soc. B* **360**, 1935–1943. (doi:10.1098/rstb.2005.1725)

Blum, M. G. B. & Francois, O. 2006 Which random processes describe the tree of life? A large-scale study of phylogenetic tree imbalance. *Syst. Biol.* **55**, 685–691. (doi:10.1080/10635150600889625)

Bolnick, D. I., Turelli, M., López-Fernández, H., Wainwright, P. C. & Near, T. J. 2008 Accelerated mitochondrial evolution and 'Darwin's corollary': asymmetric viability of reciprocal F_1 hybrids in centrarchid fishes. *Genetics* **178**, 1037–1048. (doi:10.1534/genetics.107.081364)

Butler, M. A., Sawyer, S. A. & Losos, J. B. 2007 Sexual dimorphism and adaptive radiation in Anolis lizards. *Nature* **447**, 202–205. (doi:10.1038/nature05774)

Butlin, R. K., Galindo, J. & Grahame, J. W. 2008 Sympatric, parapatric or allopatric: the most important way to classify speciation? *Phil. Trans. R. Soc. B* **363**, 2997–3007. (doi:10.1098/rstb.2008.0076)

Cohan, F. M. 2001 Bacterial species and speciation. *Syst. Biol.* **50**, 513–524. (doi:10.1080/10635150118398)

Coyne, J. A. & Orr, H. A. 2004 *Speciation*. Sunderland, Matts: Sinauer Associates.

Crisp, M., Cook, L. & Steane, D. 2004 Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Phil. Trans. R. Soc. Lond. B* **359**, 1551–1571. (doi:10.1098/rstb.2004.1528)

Davies, T. J., Barraclough, T. G., Chase, M. W., Soltis, P. S., Soltis, D. E. & Savolainen, V. 2004a Darwin's abominable mystery: insights from a supertree of the angiosperms. *Proc. Natl Acad. Sci. USA* **101**, 1904–1909. (doi:10.1073/pnas.0308127100)

Davies, T. J., Savolainen, V., Chase, M. W., Moat, J. & Barraclough, T. G. 2004b Environmental energy and evolutionary rates in flowering plants. *Proc. R. Soc. Lond. B* **271**, 2195–2200. (doi:10.1098/rspb.2004.2849)

Davies, T. J., Savolainen, V., Chase, M. W., Goldblatt, P. & Barraclough, T. G. 2005 Environment, area, and diversification in the species-rich flowering plant family Iridaceae. *Am. Nat.* **166**, 418–425. (doi:10.1086/432022)

de Mazancourt, C., Johnson, E. & Barraclough, T. G. 2008 Biodiversity inhibits species' evolutionary responses to changing environments. *Ecol. Lett.* **11**, 380–388. (doi:10.1111/j.1461-0248.2008.01152.x)

Dettman, J. R., Jacobson, D. J., Turner, E., Pringle, A. & Taylor, J. W. 2003 Reproductive isolation and phylogenetic divergence in Neurospora: comparing methods of species recognition in a model eukaryote. *Evolution* **57**, 2721–2741.

Didelot, X. & Falush, D. 2007 Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**, 1251–1266. (doi:10.1534/genetics.106.063305)

Edwards, S. V. & Beerli, P. 2000 Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54**, 1839–1854.

Edwards, S. V., Liu, L. & Pearl, D. K. 2007 High-resolution species trees without concatenation. *Proc. Natl Acad. Sci. USA* **104**, 5936–5941. (doi:10.1073/pnas.0607004104)

Etienne, R. S. & Olff, H. 2004 A novel genealogical approach to neutral biodiversity theory. *Ecol. Lett.* **7**, 170–175. (doi:10.1111/j.1461-0248.2004.00572.x)

Etienne, R. S., Apol, M. E. F., Olff, H. & Weissing, F. J. 2007 Modes of speciation and the neutral theory of biodiversity. *Oikos* **116**, 241–258. (doi:10.1111/j.0030-1299.2007.15438.x)

Feulner, P. G. D., Kirschbaum, F., Mamonekene, V., Ketmaier, V. & Tiedemann, R. 2007 Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae: Campylomormyrus): a combined molecular and morphological approach. *J. Evol. Biol.* **20**, 403–414. (doi:10.1111/j.1420-9101.2006.01181.x)

Fisher, R. A. 1930 *The genetical theory of natural selection*. Oxford, UK: Oxford University Press.

Fontaneto, D., Herniou, E. A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C. & Barraclough, T. G. 2007

Independently evolving species in asexual bdelloid rotifers. *PloS Biol.* **5**, 914–921. (doi:10.1371/journal.pbio.0050087)

Freckleton, R. P. & Harvey, P. H. 2006 Detecting non-Brownian trait evolution in adaptive radiations. *PloS Biol.* **4**, 2104–2111. (doi:10.1371/journal.pbio.0040373)

Gavrilets, S. 2004 *Fitness landscapes and the origin of species.* Princeton, NJ: Princeton University Press.

Hanage, W. P., Fraser, C. & Spratt, B. G. 2006 Sequences, sequence clusters and bacterial species. *Phil. Trans. R. Soc. B* **361**, 1917–1927. (doi:10.1098/rstb.2006.1917)

Harmon, L. J., Schulte, J. A., Larson, A. & Losos, J. B. 2003 Tempo and mode of evolutionary radiation in iguanian lizards. *Science* **301**, 961–964. (doi:10.1126/science.1084786)

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. 2003 Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B* **270**, S96–S99. (doi:10.1098/rsbl.2003.0025)

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. 2004 Identification of birds through DNA barcodes. *PloS Biol.* **2**, 1657–1663. (doi:10.1371/journal.pbio.0020312)

Hey, J., Waples, R. S., Arnold, M. L., Butlin, R. K. & Harrison, R. G. 2003 Understanding and confronting species uncertainty in biology and conservation. *Trends Ecol. Evol.* **18**, 597–603. (doi:10.1016/j.tree.2003.08.014)

Higgs, P. G. & Derrida, B. 1992 Genetic distance and species formation in evolving populations. *J. Mol. Evol.* **35**, 454–465. (doi:10.1007/BF00171824)

Hong, S. H., Bunge, J., Jeon, S. O. & Epstein, S. S. 2006 Predicting microbial species richness. *Proc. Natl Acad. Sci. USA* **103**, 117–122. (doi:10.1073/pnas.0507245102)

Hubbell, S. P. 2001 *The unified neutral theory of biodiversity and biogeography.* Princeton, NJ: Princeton University Press.

Hudson, R. R. & Coyne, J. A. 2002 Mathematical consequences of the genealogical species concept. *Evolution* **56**, 1557–1565.

Hutchinson, G. E. 1968 When are species necessary? In *Population biology and evolution* (ed. R. C. Lewontin). Syracuse, NY: Syracuse University Press.

Jablonski, D. & Sepkoski, J. J. 1996 Paleobiology, community ecology, and scales of ecological pattern. *Ecology* **77**, 1367–1378. (doi:10.2307/2265534)

Jaramillo, C., Rueda, M. J. & Mora, G. 2006 Cenozoic plant diversity in the Neotropics. *Science* **311**, 1893–1896. (doi:10.1126/science.1121380)

Johansson, J. 2008 Evolutionary responses to climate changes: how does competition affect adaptation? *Evolution* **62**, 421–435. (doi:10.1111/j.1558-5646.2007.00301.x)

Knowles, L. L. & Carstens, B. C. 2007 Delimiting species without monophyletic gene trees. *Syst. Biol.* **56**, 887–895. (doi:10.1080/10635150701701091)

Koufopanou, V., Burt, A. & Taylor, J. W. 1997 Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. *Proc. Natl Acad. Sci. USA* **94**, 5478–5482. (doi:10.1073/pnas.94.10.5478)

Lahaye, R. *et al.* 2008 DNA barcoding the floras of biodiversity hotspots. *Proc. Natl Acad. Sci. USA* **105**, 2923–2928. (doi:10.1073/pnas.0709936105)

Lavin, M., Schrire, B. P., Lewis, G., Pennington, R. T., Delgado-Salinas, A., Thulin, M., Hughes, C. E., Matos, A. B. & Wojciechowski, M. F. 2004 Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Phil. Trans. R. Soc. Lond. B* **359**, 1509–1522. (doi:10.1098/rstb.2004.1536)

Lee, M. S. Y. 2003 Species concepts and species reality: salvaging a Linnaean rank. *J. Evol. Biol.* **16**, 179–188. (doi:10.1046/j.1420-9101.2003.00520.x)

Leinonen, T., O'Hara, R. B., Cano, J. M. & Merila, J. 2008 Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J. Evol. Biol.* **21**, 1–17.

Losos, J. B., Jackman, T. R., Larson, A., deQueiroz, K. & Rodriguez-Schettino, L. 1998 Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**, 2115–2118. (doi:10.1126/science.279.5359.2115)

Maddison, W. P., Midford, P. E. & Otto, S. P. 2007 Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710. (doi:10.1080/10635150701607033)

Mallet, J. 2010 Group selection and the development of the biological species concept. *Phil. Trans. R. Soc. B* **365**, 1853–1863. (doi:10.1098/rstb.2010.0040)

Mayr, G. 2005 The Paleogene fossil record of birds in Europe. *Biol. Rev.* **80**, 515–542. (doi:10.1017/S1464793105006779)

McDonald, J. H. & Kreitman, M. 1991 Adaptive protein evolution at the adh locus in *Drosophila*. *Nature* **351**, 652–654. (doi:10.1038/351652a0)

Meiklejohn, C. D., Montooth, K. L. & Rand, D. M. 2007 Positive and negative selection on the mitochondrial genome. *Trends Genet.* **23**, 259–263. (doi:10.1016/j.tig.2007.03.008)

Meyer, C. P. & Paulay, G. 2005 DNA barcoding: error rates based on comprehensive sampling. *PloS Biol.* **3**, 2229–2238. (doi:10.1371/journal.pbio.0030422)

Nee, S. 2001 Inferring speciation rates from phylogenies. *Evolution* **55**, 661–668. (doi:10.1554/0014-3820(2001)055[0661:ISRFP]2.0.CO;2)

Nee, S., May, R. M. & Harvey, P. H. 1994 The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* **344**, 305–311. (doi:10.1098/rstb.1994.0068)

Nielsen, R. 2005 Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**, 197–218. (doi:10.1146/annurev.genet.39.073003.112420)

Ochman, H., Lerat, E. & Daubin, V. 2005 Examining bacterial species under the specter of gene transfer and exchange. *Proc. Natl Acad. Sci. USA* **102**, 6595–6599. (doi:10.1073/pnas.0502035102)

O'Meara, B. C. 2010 New heuristic methods for species delimitation and species tree inference. *Syst. Biol.* **59**, 59–73. (doi:10.1093/sysbio/syp077)

Owens, I. P. F. & Bennett, P. M. 1995 Ancient ecological diversification explains life-history variation among living birds. *Proc. R. Soc. Lond. B* **261**, 227–232. (doi:10.1098/rspb.1995.0141)

Pagel, M. 1997 Inferring evolutionary processes from phylogenies. *Zool. Scr.* **26**, 331–348. (doi:10.1111/j.1463-6409.1997.tb00423.x)

Pagel, M. 1999 The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* **48**, 612–622. (doi:10.1080/106351599260184)

Pagel, M., Meade, A. & Barker, D. 2004 Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* **53**, 673–684. (doi:10.1080/10635150490522232)

Phillimore, A. B. & Price, T. D. 2008 Density-dependent cladogenesis in birds. *PloS Biol.* **6**, 438–439. (doi:10.1371/journal.pbio.0060071)

Phillimore, A. B., Freckleton, R. P., Orme, C. D. L. & Owens, I. P. F. 2006 Ecology predicts large-scale patterns of phylogenetic diversification in birds. *Am. Nat.* **168**, 220–229. (doi:10.1086/505763)

Pinto, G., Mahler, D. L., Harmon, L. J. & Losos, J. B. 2008 Testing the island effect in adaptive radiation: rates and patterns of morphological diversification in Caribbean and mainland Anolis lizards. *Proc. R. Soc. B* **275**, 2749–2757. (doi:10.1098/rspb.2008.0686)

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D. & Vogler, A. P. 2006 Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* **55**, 595–609. (doi:10.1080/10635150600852011)

Pritchard, J. K., Stephens, M. & Donnelly, P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Pulido, F. & Berthold, P. 2004 Microevolutionary response to climatic change. *Birds Clim. Change* **35**, 151–183. (doi:10.1016/S0065-2504(04)35008-7)

Pybus, O. G. & Harvey, P. H. 2000 Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. Lond. B* **267**, 2267–2272. (doi:10.1098/rspb.2000.1278)

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

Rabosky, D. L. 2009 Ecological limits on clade diversification in higher taxa. *Am. Nat.* **173**, 662–674. (doi:10.1086/597378)

Rabosky, D. L. & Lovette, I. J. 2008 Density-dependent diversification in North American wood warblers. *Proc. R. Soc. B* **275**, 2363–2371. (doi:10.1098/rspb.2008.0630)

Rabosky, D. L., Donnellan, S. C., Talaba, A. L. & Lovette, I. J. 2007 Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proc. R. Soc. B* **274**, 2915–2923. (doi:10.1098/rspb.2007.0924)

Ratnasingham, S. & Hebert, P. D. N. 2007 BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Mol. Ecol. Notes* **7**, 355–364. (doi:10.1111/j.1471-8286.2007.01678.x)

Ree, R. H. 2005 Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution* **59**, 257–265.

Ricklefs, R. E. 2006 The unified neutral theory of biodiversity: do the numbers add up? *Ecology* **87**, 1424–1431. (doi:10.1890/0012-9658(2006)87[1424:TUNTOB]2.0.CO;2)

Ricklefs, R. E. 2007a Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* **22**, 601–610. (doi:10.1016/j.tree.2007.06.013)

Ricklefs, R. E. 2007b History and diversity: explorations at the intersection of ecology and evolution. *Am. Nat.* **170**, S56–S70. (doi:10.1086/519402)

Rieseberg, L. H., Wood, T. E. & Baack, E. J. 2006 The nature of plant species. *Nature* **440**, 524–527. (doi:10.1038/nature04402)

Rosenberg, N. A. 2003 The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution* **57**, 1465–1477.

Rosenzweig, M. L. 1996 *Species diversity in space and time.* Cambridge, MA: Cambridge University Press.

Schluter, D. 2001 Ecology and the origin of species. *Trends Ecol. Evol.* **16**, 372–380. (doi:10.1016/S0169-5347(01)02198-X)

Silvertown, J., Dodd, M., Gowing, D., Lawson, C. & McConway, K. 2006 Phylogeny and the hierarchical organization of plant diversity. *Ecology* **87**, S39–S49. (doi:10.1890/0012-9658(2006)87[39:PATHOO]2.0.CO;2)

Simpson, G. G. 1953 *The major features of evolution.* New York, NY: Columbia University Press.

Sites, J. W. & Marshall, J. C. 2003 Delimiting species: a renaissance issue in systematic biology. *Trends Ecol. Evol.* **18**, 462–470. (doi:10.1016/S0169-5347(03)00184-8)

Stanley, S. M. 1979 *Macroevolution: pattern and process.* San Francisco, CA: W.H. Freeman.

Templeton, A. 1989 The meaning of species and speciation: a population genetics approach. In *Speciation and its consequences* (eds D. Otte & J. Endler), Sunderland, MA: Sinauer Associates.

Thomas, G. H., Freckleton, R. P. & Szekely, T. 2006 Comparative analyses of the influence of developmental mode on phenotypic diversification rates in shorebirds. *Proc. R. Soc. B* **273**, 1619–1624. (doi:10.1098/rspb.2006.3488)

Thomas, G. H., Meiri, S. & Phillimore, A. B. 2009 Body size diversification in *Anolis*: novel environment and island effects. *Evolution* **63**, 2017–2030. (doi:10.1111/j.1558-5646.2009.00694.x)

Tilman, D. & Lehman, C. 2001 Human-caused environmental change: impacts on plant diversity and evolution. *Proc. Natl Acad. Sci. USA* **98**, 5433–5440. (doi:10.1073/pnas.091093198)

Veliz, D., Duchesne, P., Bourget, E. & Bernatchez, L. 2006 Stable genetic polymorphism in heterogeneous environments: balance between asymmetrical dispersal and selection in the acorn barnacle. *J. Evol. Biol.* **19**, 589–599. (doi:10.1111/j.1420-9101.2005.01000.x)

Venter, J. C. *et al.* 2004 Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74. (doi:10.1126/science.1093857)

Weir, J. T. 2006 Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. *Evolution* **60**, 842–855.

Wiens, J. J. 2007 Species delimitation: new approaches for discovering diversity. *Syst. Biol.* **56**, 875–878. (doi:10.1080/10635150701748506)

Will, K. W. & Rubinoff, D. 2004 Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* **20**, 47–55. (doi:10.1111/j.1096-0031.2003.00008.x)

Williams, G. C. 1992 *Natural selection: domains, levels and challenges.* Oxford, UK: Oxford University Press.