

The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten

James Rosindell^{1,2}, Stephen P. Hubbell^{3,4} and Rampal S. Etienne^{5,6}

¹Institute of Integrative and Comparative Biology, University of Leeds, Leeds, UK, LS2 9JT

²Institute of Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA

³Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

⁴Center for Tropical Forest Science, Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002-0948, Republic of Panama

⁵Community and Conservation Ecology Group, Centre for Ecological and Evolutionary Studies, University of Groningen, Box 11103, 9700 CC Groningen, The Netherlands

⁶Department of Soil and Physical Sciences, Faculty of Agriculture and Life Sciences, Lincoln University, Box 84, Lincoln 7647, Christchurch, New Zealand

A decade has now passed since Hubbell published *The Unified Neutral Theory of Biodiversity and Biogeography*. Neutral theory highlights the importance of dispersal limitation, speciation and ecological drift in the natural world and provides quantitative null models for assessing the role of adaptation and natural selection. Significant advances have been made in providing methods for understanding neutral predictions and comparing them with empirical data. In this review, we describe the current state-of-the-art techniques and ideas in neutral theory and how these are of relevance to ecology. The future of neutral theory is promising, but its concepts must be applied more broadly beyond the current focus on species–abundance distributions.

The mystery of biodiversity

Imagine yourself deep in a tropical rain forest or floating over a coral reef surrounded by thousands of interacting species. How can so many species coexist? The traditional answer is that all species differ in important ways [1], so that each species is limited by a unique set of factors. Adaptive trade-offs prevent the evolution of super-species that are better at doing everything. Classic ecological niche theory captures the unique roles of species and, consequently, is often complex and parameter rich, but is all this detail really necessary? Niche assembly is supported by abundant case studies, but how universal is it? These are the kinds of question that motivated Hubbell to write *The Unified Neutral Theory of Biodiversity and Biogeography* [2]. Here, we discuss the past, present and future of neutral theory, with a particular focus on recent advances, directions for future research and the utility of the theory in a general ecological context.

The history of neutral theory

Neutral theory makes a controversial ‘neutrality assumption’: all individuals within a particular trophic level have the same chances of reproduction and death regardless of

Glossary

Beta-diversity: also known as distance-decay; the probability as a function of distance between two individuals that they will be of the same species.

Coalescence: a technique developed in population genetics for simulating or analytically solving properties of a sample of individuals by tracing their ancestry back to their common ancestors.

Dispersal kernel: a statistical distribution describing stochastic dispersal events by giving the probability of dispersal as a function of distance.

Dispersal limitation: a process that causes the location of an individual to be restricted in some sense by the location of its parent. In the common application of neutral models to tropical forest trees, this could be regarded as ‘seed limitation’. It has been further noted that ‘recruitment limitation’ might be a better term because only individuals of adult reproductive age are usually modelled. In the context of the spatially implicit neutral model, dispersal limitation is given by parameter m , which measures the relative importance of regional dispersal processes from the metacommunity compared with local processes of birth and death in the local community [17].

Fat-tailed distribution: a statistical distribution favouring rare, yet erratic draws very far from the mean. Mathematically, the tails of a ‘fat-tailed’ dispersal kernel have power-law decay (rather than exponential decay) and can have an infinite variance.

Fundamental biodiversity number: a parameter $\theta = J_M \frac{v}{1-v}$ describing the behaviour of the metacommunity in a neutral model. Sometimes approximated to $\theta \approx J_M v$ when v (the per capita speciation rate) is small. Where generations are not overlapping, so that all individuals in the model die and are replaced together, $\theta \approx 2J_M v$.

Fundamental immigration number: captures dispersal limitation in a more logical way than with the parameter m . The fundamental immigration number I is given by $I = \frac{m(J-1)}{1-m}$.

Individual-level neutral model: all individuals are equivalent and distinguished only by their species labels. Species can be different insofar as they might have different abundances.

Metacommunity: in the classic neutral model, the metacommunity is a well-mixed source pool of individual organisms, each with potential to give rise to offspring in the local community patch being studied. The parameter J_M is typically used to describe the number of individuals in the metacommunity.

Non-spatial: a model is non-spatial if the spatial structure of the system is not considered; in the context of neutral theory, this means that all organisms are well mixed in space and their positions are unknown and/or unimportant.

Protracted speciation: a mode of speciation where speciation is a gradual process rather than an instantaneous event.

Sampling theory: a theory that takes into account the effect of sampling a small proportion of the much larger natural system.

Spatially explicit: a model is spatially explicit if spatial structure is fully accounted for in the model. In the context of neutral theory, the position of each individual in space would be known explicitly and the dynamics of the model would reflect this spatial structure.

Spatially implicit: a spatially implicit model is in between non-spatial and spatially explicit: there is some limited spatial structure. For example, the classic neutral model is spatially implicit and does not take into account the exact spatial location of each individual, but it does classify individuals as

belonging to one of two distinct spatial scales: the local community or the larger metacommunity.

Species–area relationships (SARs): the relationship between area and the species richness observed in that area.

Species-level neutral model: all species are equivalent; they have the same chance of immigration, extinction and speciation.

Symmetric model: take individual organisms labelled as belonging to species X or species Y. In a symmetric model, all Xs and Ys can be switched without having any effect; in such models, species have no unique properties. Symmetric models could be individual or species based, but in the former case the species labels would have to be swapped for all individuals of the given species.

Tri-phasic SAR: a form of SAR with three distinct phases (in log-log space) a curve down at small scales, a straight line at intermediate scales and a curve up at large scales to a limiting gradient of unity.

their species identity [2]. Kimura pioneered the idea in a different context with his neutral theory of molecular evolution [3]. His intent was to model changes in allele frequency at a locus where all the alleles were selectively ‘neutral’, so that substituting one allele for another did not affect the fitness of an individual. The ecological counterpart [4–6] assumed neutrality of individuals within communities rather than of alleles within populations, but the parallel with neutral models of alleles later proved useful for developing methods for neutral theory in ecology. From here on, ‘neutral theory’ refers to that of ecology.

The next significant advancement came 20 years after publication of the early papers [4–6] on neutral theory and began to consider spatial structure [7–9], which led ultimately to the publication of neutral theory in its classic form in 2001 [2]. The classic neutral theory has a meta-

community (well-mixed source pool of potential immigrants) and a local community. Within the local community, local extinction and immigration from the metacommunity are in equilibrium (Box 1). This spatially implicit (see Glossary) structure was inspired by MacArthur and Wilson’s theory of island biogeography [10], which proposed that species richness on an island was determined by a dynamic equilibrium: a balance between local extinction and immigration.

Interpreting fits to empirical data

Most empirical evaluations of neutral models have focused on fitting the classic neutral model to species abundance distributions [2] (Box 1). Substantial improvements in the fitting methods have been developed since the early work (Box 2); in fact, one great benefit to the classic neutral model is that its simplicity allows a likelihood-based framework for fitting data [11–14]. Here, we do not attempt to enumerate the many fits that have been carried out [15], but instead focus on what they mean in a general ecological context.

Strikingly similar species abundance distributions and species–area relationships (SARs) appear across very different communities. The processes responsible for these patterns are probably those that are common to a wide variety of different systems; those driven by species-specific qualities do not seem probable candidates because they would change in different communities. Neutral theory is thus a good starting point for an explanation of the

Box 1. The classic neutral model

The classic neutral theory models a local community containing J individuals. One of these individuals, chosen at random, dies and is replaced at every time step. The replacement is an offspring of another randomly chosen individual from the local community, occurring with probability $(1 - m)$, or offspring from a randomly chosen individual from an outside pool of individuals known as the metacommunity with probability m . The parameter m is thus a measure of dispersal limitation (Figure 1). One problem with m is that it does not translate into dispersal limitation in the most logical way: a small local community also involves a much smaller flux of immigrants for the same value of m . An alternative parameter $l = \frac{m(J-1)}{1-m}$ accounts for this and is therefore often used instead of m .

The metacommunity itself contains J_M individuals and the theory assumes that its species abundance distribution does not change at the local community timescales. A process similar to that in local

communities determines the distribution of species abundance in the metacommunity, but replacement individuals are offspring from a randomly chosen individual in the metacommunity, except in the case of a speciation event. Speciation occurs with probability ν ; a so-called ‘point mutation’ speciation event, such that the replacement individual founds a new species. Neutral theory characterises the sampling distribution of species abundances in a metacommunity of size J_M and speciation rate ν entirely by a single composite parameter $\theta = J_M \frac{\nu}{1-\nu}$, known in neutral theory as the fundamental biodiversity number (Box 2). This is often written simply as $\theta = J_M \nu$, which is an approximation for small speciation rates. When the entire community rather than a single individual dies and is replaced (non-overlapping generations), $\theta \approx 2J_M \nu$ [93]. Several variations of the model exist that generate mathematically exact predictions for macroecological patterns, such as species abundance distributions and species accumulation curves (Box 2).

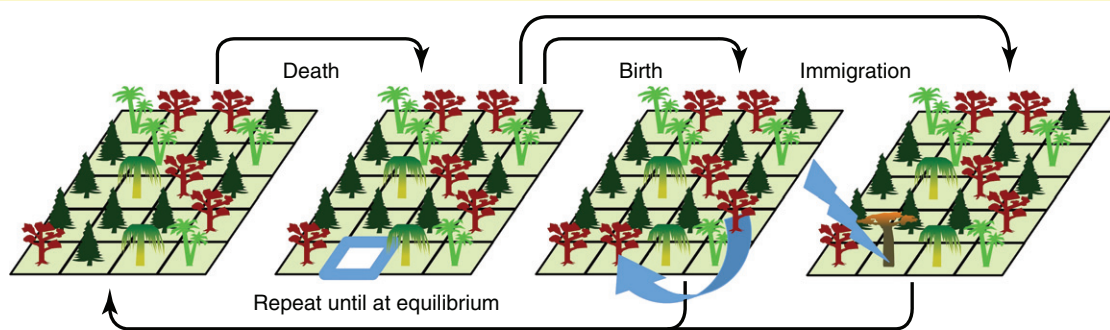


Figure 1. The mechanics of neutral theory in the local community. In this example, $J = 20$ so the probability of the indicated individual dying is $\frac{1}{20}$, the probability of birth happening within the local community, as indicated in the third panel, is $(1 - m) \frac{1}{20}$ for each living individual and the probability of immigration is m . Exactly the same dynamics occurs in the metacommunity, except that the process of speciation replaces immigration and, thus, ν is used instead of m .

Box 2. Analysis and fitting of the classic neutral model

Because neutral theory is a sampling theory, it makes predictions about species abundances in samples of communities. This is often more useful than making predictions for entire communities, which are rarely fully sampled in reality. Much of the earlier mathematical work calculated and then fitted the expected values predicted by the neutral model. This was done by summarising the classic neutral model framework mathematically in a master equation [94–96]. The master equation describes the dynamics of abundance of an arbitrary species in the community as a function of the transition probabilities between different abundances of the species. Because all individuals are demographically identical, these dynamics also characterise the distribution of abundances of all species in the community. The most rigorous comparisons with abundance data require more involved multivariate likelihood methods, which did not appear until later [13]. These methods, based on coalescence, gave the likelihood of the model producing exactly the empirically observed species abundance distribution and represent the most accurate method to fit the classic neutral model with empirical abundance data.

In the special case of no dispersal limitation, the likelihood equations reduce to Ewens' sampling formula [97], which was developed in population genetics and pre-dates ecological neutral models. In this case, species richness S is a sufficient statistic because the fundamental biodiversity number θ can be found from Equation (1):

$$S = \sum_{k=1}^J \frac{\theta}{\theta + k - 1} \quad (1)$$

The abundance distribution has no influence and thus provides a parameter-free test because species abundances conditioned on sample size and richness must follow a prescribed distribution that one can compare against real data. An equivalent test in the dispersal limited case ($m < 1$) has recently been perfected [53]. One must first find optimal values of the parameters that maximise the likelihood of reproducing the data, then generate simulated data sets using these same parameters (Box 4). The likelihood of each generated data set is then compared with the likelihood of the real data; if the likelihood of this empirical data is around the median of the artificially generated

likelihoods, the fit is good. Similar procedures have since been developed for more complex versions of the model, such as when multiple local communities are linked to the same metacommunity [53,55] or when other speciation mechanisms are considered [34,36,37,98].

The statistical properties of the likelihood formula for the classic neutral model indicate there are often two local maxima in parameter space and a ridge corresponding to an isocline of species richness in (θ, m) -space (where m is dispersal limitation) along which the likelihood is relatively constant (Figure 1). This illustrates limitations to the inferences that one can make from species abundance distributions alone.

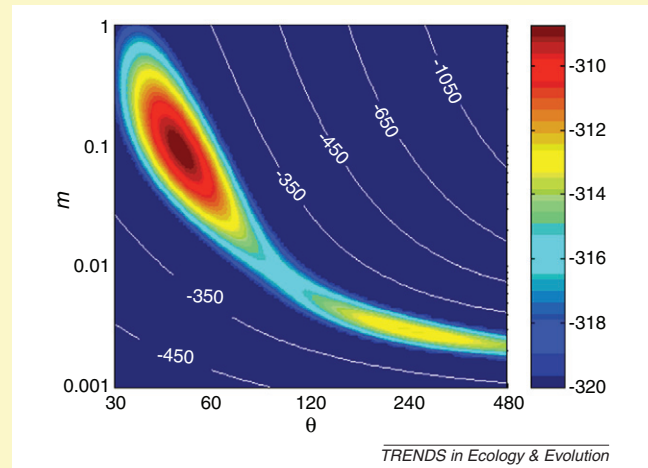


Figure 1. The likelihood surface shown in colours as a function of dispersal limitation m and the fundamental biodiversity number θ . Red and yellow colours represent parameter combinations that are most likely to produce the observed data. The uncertainty in these parameters, especially along the isocline of constant species richness, is clear. Reproduced from figure 3 in [93] using data for tropical forest trees from Barro Colorado Island. With kind permission from Springer Science and Business Media.

universal patterns in species abundance distributions and it has indeed been successful at this [2]. It does not follow that species-specific qualities are absent in the real world; instead, it suggests that their effects do not penetrate the foggy lens of the summary statistics being studied. This might explain how a neutral model can effectively fit data from a non-neutral world.

Many different models can explain the same sampled species abundance data [16–18]. Even models based on neutral theory, but introducing a particular niche structure, give mathematically identical results to the classic neutral model at large spatial scales [19,20]. On the one hand, this implies that neutral models are robust to the introduction of certain niche structures, but on the other hand, it means that species abundance data will not always be able to distinguish various models.

Neutral theory has an advantage over the many alternative models of species abundances: it can also predict other ecological patterns. To make more powerful inferences from neutral theory, an important step will be to exploit this ability and combine fitting of species abundances with temporal dynamics of abundance [21,22], SARs [23,24], or phylogenetic data [25]. It is sometimes found that the classic neutral model fails to match patterns; for example, spatially explicit patterns. However, in the next section, we show that this can often

be solved by relaxing the other assumptions of the model while retaining its neutrality. If a new factor introduced to the model produces a dramatic improvement in its performance, then it is likely that the new factor is important. For example, fits of species–area curves by neutral models were not realistic [26,27] until long-distance dispersal was incorporated [28] (Box 3). This supports the general concept that long-distance dispersal is important [29].

The other assumptions of the model

Classic neutral theory includes several ‘auxiliary assumptions’, unrelated to neutrality [30], which can cause the model to fail for reasons other than its neutrality and draw attention away from the key issues. It is therefore desirable to remove these auxiliary assumptions.

Zero-sum assumption

The classic neutral model assumes that replacement of dead individuals occurs immediately: the zero-sum assumption [2]. However, when species instead have independent dynamics, the same species abundance distribution and the same relationship between sample size and species richness still arise [17,31,32]. Species independence even causes the same relationships when a specific form of niche structure is introduced [20].

Box 3. Neutral model predictions

Neutral models make many rich predictions, here we discuss the best known of these. Although neutral models can make extremely varied predictions, some patterns are outside their scope. For example, distributions of the traits and interactions of species between trophic levels require non-neutral models because they fundamentally involve species differences. If species differences are not themselves the object being studied, then the first attempt at a simple model need not include species differences; similarly, if commonness and rarity of species is not the subject of a study, then the first attempt at a simple model need not be individual based (e.g. MacArthur and Wilson's model of island biogeography [10]).

Relative species abundances

Relative species abundances are the best known predictions of neutral theory. Most empirical species abundance distributions are either log-series like or log-normal like. In the spatially implicit version of neutral theory, the metacommunity follows a log-series distribution and the local community follows a zero-sum multinomial distribution, (or dispersal-limited multinomial distribution). This distribution entails fewer rare species than does the log-series and it looks like a skewed log-normal distribution (Figure 1). It is this flexibility of the classic neutral model to reconcile the log-series and log-normal distributions that allows it to fit empirical data well and arguably with greater flexibility than some spatially explicit versions of the model [52].

SARs

The relationship between area A and species richness S (SAR) typically exhibits a power-law relationship $S \propto A^\eta$ at intermediate spatial scales, forming a straight line in log-log space. The power law is not expected at the smallest scales, where the SAR curves down, or at the largest scales, where the SAR curves up, so that a 'tri-phasic' S-shape is formed overall [2]. The voter model reproduces the larger scale

behaviour of empirical SARs [63,64]. Voter models cannot, however, capture the behaviour at scales similar to the dispersal distance because the dispersal distance is just one grid space, which is also the smallest length explicitly modelled: the model has no resolution finer than the dispersal distance. The primary reason for a decelerating curve at the smallest scales in nature is the differential sampling rate of common and rare species, which can be seen even in a spatially implicit neutral model [2]. To explain the complete tri-phasic SAR requires a spatially explicit neutral model with dispersal beyond the nearest neighbours. This was first shown for intermediate dispersal distances in relatively small communities [99], then more generally for all but the smallest dispersal distances in an infinite community [26]. This model is qualitatively correct, but the fit to empirical data is relatively poor and cannot be resolved even with small speciation rates [27]. A good fit to data (including an appreciable region of the SAR that follows the power law) requires not only smaller speciation rates, but also 'fat-tailed' dispersal kernels [28] that exhibit a greater probability of dispersal over long distances and are generally considered more realistic than the alternatives [29] (Figure 1).

Beta-diversity

Analytical results for spatially explicit neutral theory were derived for beta-diversity, the probability that two individuals will be conspecific as a function of their separation [65]. These analytical results use an infinite community size and predict a steady decay with distance. Observations, however, show a bi-phasic decay curve [100], with fast decay of similarity on short distances (<0.1 km) and slow decay on large distances (to >1000 km). Zillio *et al.* [64] showed that these bi-phasic curves can easily be explained by a symmetric model that includes local conspecific density dependence that weakens to neutrality beyond the local area.

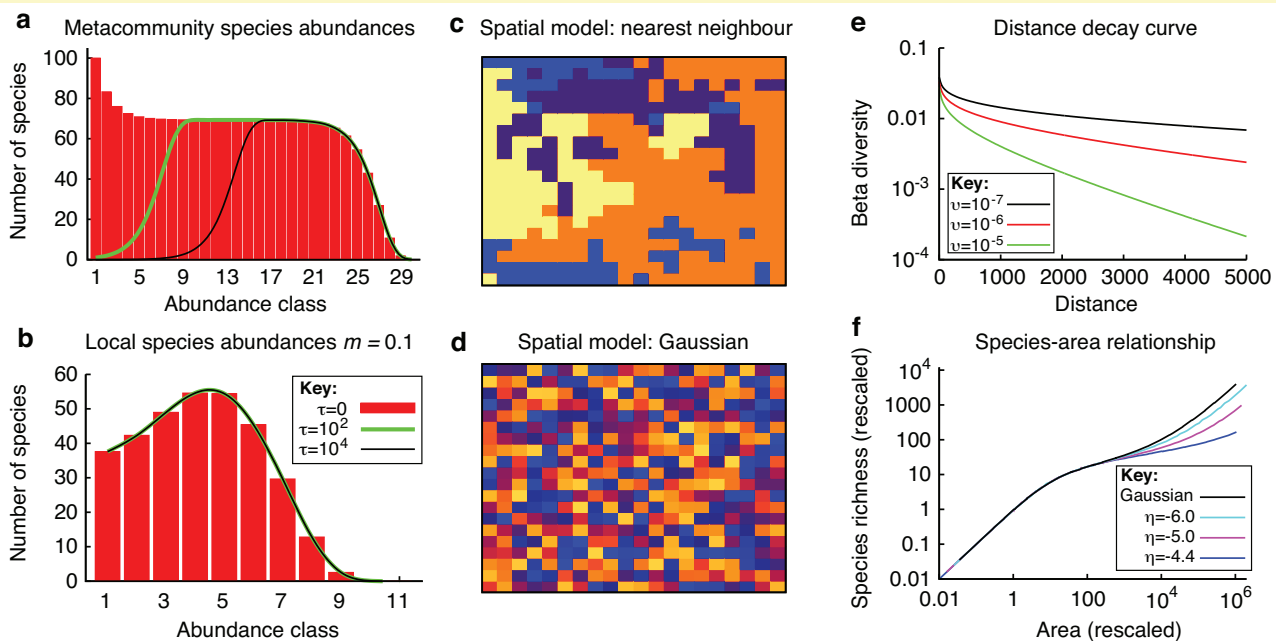


Figure 1. Some of the varied predictions from neutral models: (a,b) show species abundance distributions from the spatially implicit neutral model local community and metacommunity respectively, under both point mutation speciation ($\tau = 0$) and protracted speciation (with transition times τ given by $\tau = 10^2, 10^4$). (c,d) show example maps of species for the spatially explicit neutral model with nearest neighbour dispersal and Gaussian dispersal, respectively. (e) shows distance decay (beta-diversity) for a spatially explicit model. (f) shows SARs for the spatially explicit neutral model with Gaussian dispersal kernel and 'fat-tailed' dispersal kernels. Here, η represents how fat tailed the dispersal kernel is with $\eta = -4.4$ representing the greatest proportion of long-distance dispersal events. (a,b) was produced using formulas from [18], (c,d) by simulation (Box 4) inspired by [99], (e) using the formulas from [65] and (f) replotted with permission from [28]

Point mutation speciation assumption

The way in which speciation is modelled in neutral theory has received much attention [33]. The original neutral model assumes a 'point mutation' mode of speciation,

where, with each birth in the metacommunity, there is a small probability ν that the newborn founds a new species [2]. Without speciation to replenish species, gradual extinctions owing to drift guarantee that only a single species

remains; the point mutation model is the simplest way to avoid this. An apparently more realistic alternative is random fission speciation, in which populations split into two daughter species at random [2]. Etienne and Haegeman [34] recently noted that this produces identical results to MacArthur's broken stick model [35], which was motivated by random apportioning of limiting resources to niches that would then dictate relative species abundances. Surprisingly, point mutation speciation fits empirical data much better than do random fission and a variety of other alternatives [34,36,37]. This suggests that actual species may arise as fairly small populations [38].

Point mutation speciation produces many species whose lifetimes are far too short, whereas random fission speciation produces mean species lifespans that are too long [39]. These problems were initially linked to the neutrality assumption [39]. Hubbell and Lake resolved this by bridging the gap between point mutation and random fission speciation with 'peripheral isolate' speciation, which allowed new species to arise with founding abundances given by a normal distribution [40,41]. Although peripheral isolate speciation has never been fitted to data, a special case introduced new species with a fixed founding population size greater than one and was successfully fitted to Foraminifera metacommunity data [42].

Rosindell *et al.* recently proposed a model of 'protracted speciation' [18], in which speciation is not instantaneous, but a gradual process beginning with the creation of an 'incipient species' or 'variant'. Variants are considered conspecific with their parent species until a transition time has passed, after which they become good species. Protracted speciation provides fits to local species abundance distributions at least as good as those of point mutation speciation, while allowing for realistic mean species lifetimes by effectively increasing the founding abundances of new species.

The fact that the most common species would be too old in a neutral model [6,43–45] is attributable primarily to the slow pace of random drift, rather than to the mode of speciation and is very different from issues regarding mean species lifetimes. Environmental stochasticity, which causes the total community size to fluctuate, increases the turnover of common species, and suggests a promising solution to the problem [42]. It has been noted that selection is required for speciation to occur at the realistic rates in neutral models (Desjardins-Proulx, P. and Gravel, D., unpublished data); species-level neutral models of speciation might therefore be regarded as phenomenological, or even as implicitly including some effects of selection at the within-species level. Neutral theory can be used as a foundation for more advanced speciation models that include explicit genetic details [46,47]. This work forms a useful link between community ecology and genetics, although complexity of the more detailed models is presently too great to allow simulations of sufficient size for accurate quantitative comparisons with data [46].

Spatially implicit spatial structure

Dispersal limitation is central to neutral theory and important for determining species abundance distributions even in niche-based models [48]. Although there are species

differences in dispersal ability [49], a neutral model can be used conveniently to study the effect of species being dispersal limited where all species have the same dispersal abilities. The further effects of species having differing (rather than identical) dispersal abilities can only really be quantified in comparison with the neutral case.

When trying to explain non-spatial patterns, such as species abundance distributions, it is natural to start with spatially simple models. The spatially implicit neutral theory captures dispersal limitation in a mathematically tractable way, but only approximates the spatially explicit reality of dispersal in actual communities. In the classic neutral model, the local community is dispersal limited from a panmictic metacommunity, but the metacommunity is, in reality, just a collection of local communities, so it seems inconsistent to assume a homogeneously mixed metacommunity [30]. The parameters of the classic neutral model have biological meaning, but performing a fit of the model to data remains the only way to measure the parameters easily. To convert between the dispersal limitation parameter m and the dispersal kernel, one must draw repeatedly from the dispersal kernel, counting the number of immigrant individuals that enter the local community from outside [13,50]. The fundamental biodiversity number $\theta \approx J_M \nu$ depends on the speciation rate ν and the metacommunity size J_M . The metacommunity size, in particular, cannot easily be measured (but see [51]) and might not really be indicative of the total number of individuals that could disperse into, and survive within, the local community. Spatially explicit models are a significant improvement as they have speciation rate ν , spatial structure and dispersal probabilities (most often just a dispersal kernel) as the only parameters and, thus, direct biological interpretations come more naturally, although ν will be difficult to quantify. Fitting the spatially implicit neutral model to data produced by a spatially explicit neutral model provides a way to put the spatially implicit parameters θ and m in context [52]. Given that the data being fitted were artificially created, the parameters for speciation rate ν and dispersal distance D are known and can be compared with the spatially implicit model parameters that produced the best fit. Surprisingly, both the fundamental biodiversity number θ and dispersal limitation parameter m depend strongly on dispersal distance D and only weakly on speciation rate ν .

Spatially explicit models

There is a trade-off between the generality of the predictions of a model and the inclusion of extra details, which requires careful consideration, particularly when developing spatially explicit neutral models. One simple and generic extension to the classic neutral model was the joint consideration of more than one local community connected to a single metacommunity [53–55]. Volkov *et al.* [56] similarly considered several local communities, but they combined them to provide a more realistic description of the metacommunity. A different approach to the artificial distinction between the metacommunity and local community scales was to introduce a continuum model [16]. Further developments in the form of networks of patches are more realistic extensions to these ideas [57–59]. This

Box 4. Coalescence simulations

To simulate neutral models traditionally, one starts with an initial condition and then applies the rules of the model methodically, taking the system forward until one reaches an equilibrium. This is extremely slow and the total size of the simulation is severely limited by computer memory. One can study neutral models more efficiently by using simulation methods from coalescence theory [67] borrowed from population genetics (Box 2). Although coalescence has long been known to be useful for simulating neutral models [63,65,99], it is only in recent years that its full power as a simulation method in the neutral theory of ecology has been realised [68]. Coalescence begins with a present-day sample from the community and works backwards in time, applying the rules of the model in reverse. This efficiently uncovers the ancestry of contemporary individuals in the sample and can usually only be applied to neutral models. Crucially, coalescence simulations are now limited not by the size of the landscape being simulated, but by the size of the sample being taken from it. This is extremely valuable because empirical data often sample a tiny proportion of an extremely large system. Coalescence thus provides a tractable way to simulate individual-based theory for very large and complex spatial structures; no known alternative model would be tractable in these situations. The algorithm proceeds as follows:

- (i) For each individual in the desired sample, define one lineage and store it in memory with the spatial coordinates of the individual, which are usually restricted to a grid or patch structure.

- (ii) Identify the individual that was most recently born. If one works back in time, one identifies this individual as the one that will disappear (be un-born) first. Choose the individual according to a (discrete) uniform distribution from all the lineages being traced (initially the sample size, but later less than this). Deaths of other lineages are of no relevance. This is equivalent to the forwards-in-time step in which one chooses an individual at random to die and be replaced.
- (iii) One is now interested in the location of the maternal parent of the chosen focal individual. This is chosen according to the dispersal rules of the model, which usually means drawing from the dispersal kernel, or choosing a patch based on a network topology.
- (iv) If two lineages share the same position in space, then they have coalesced; a common ancestor has been identified between two groups of individuals in the ancestry of the sample and there is now one fewer lineage to trace.
- (v) With each birth on a lineage, there is a probability ν of speciation. If speciation happens, tracing the lineage ceases as it is now known that all of its decedents in the sample are conspecific with each other and heterospecific with all others.

The algorithm terminates once no further lineage is left to trace; at this point, the exact species identity of each individual in the sample should be known. If the algorithm does not terminate, steps (ii) through (v) are repeated and the results then reassessed.

type of model has been applied to riverine fish communities, where it produced excellent fits to species richness data [60]. A disadvantage of the network approach is the difficulty in defining the network itself, but in the case of a river and its tributaries, a clear topology is available.

The first fully spatially explicit neutral models where the unique position of each individual in space was explicitly modelled, were ‘voter models’. These were based on a discrete grid of individuals that could disperse only to neighbouring sites [61], and have yielded some analytical solutions [62]. Whereas voter models could predict patterns such as SARs [63,64], their predictions could not capture the behaviour of the model at scales smaller than the dispersal distance. More recent extensions of the concept use dispersal kernels that allow dispersal over greater distances; these can explain spatially explicit patterns such as beta-diversity [65] and SARs [26,28] (Box 3). Although most applications of spatially explicit theory are to predict patterns that cannot be approached with spatially implicit theory, there is still a need to study spatially explicit predictions of species abundance distributions.

Spatially explicit models present a greater challenge for analysis than do spatially implicit models. A promising technique based on quantum field theory was recently introduced [66]. This approach provides an analytical solution to the SARs of a fully spatially explicit neutral model. Unfortunately, it cannot yet deal with either species abundance distributions or fat-tailed dispersal kernels. If no analytical solutions are available, the only alternative is simulation, which is best done by coalescence (Box 4) based on a population genetics framework [67,68]. Coalescence can comfortably simulate large (even infinite) and extremely complex spatially explicit and temporally changing environments under neutral theory (Box 4) and can output species richness, abundance and many other patterns (Box 3). To our knowledge, no other individual-based model

remains tractable for such large numbers of individuals and high spatial complexity.

The future of neutral theory

The field has progressed considerably during the past 10 years, but what developments do the next 10 years hold in store? In this section, we discuss the possibilities.

Island biogeography

MacArthur and Wilson’s theory of island biogeography introduced the revolutionary idea that community species composition is not static, but in continual dynamic turnover. The neutral theory of biodiversity can replicate these results, but in addition incorporate species abundances [22]. Recent work used an advanced form of protracted speciation, where duration of speciation was linked to dispersal, to model the distributions and abundances of endemic and non-endemic species on islands [69]. The new mechanistic model supports existing work based on verbal models and correctly predicts the richness of endemic and immigrant avifauna species on islands. Just as ecological neutral theory has encouraged an artificially high interest in the species abundance distribution, MacArthur and Wilson’s island biogeography theory has focused too much attention on species richness. New models are thus needed that look beyond species richness on islands and neutral models have the potential to fulfil this need [69].

Dynamics

Neutral theory also makes explicit predictions about the time evolution of communities, but again little work has been done in this area despite much potential ([70–72]). Keil *et al.* studied population dynamics in a patch-based neutral model [72] and showed that neutral theory can reproduce Taylor’s power law relating the mean and variance in population size for a given species. The rate at which species richness accumulates as the period of observation increases

produce species–time relationships that have also been predicted by neutral theory [73]. Vanpeteghem *et al.* [59,71] found an analytical solution to the time evolution of Simpson's diversity index under neutral theory. However, more work is possible, especially studying the effects of neutrality in dynamically changing environments.

Palaeobiology

The temporally explicit nature of neutral theory might make it useful in palaeontological applications [74]. Neutral theory could help to address the complex taxonomic and other sampling issues that palaeontologists routinely face. For example, Tomasovych and Kidwell recently used the neutral theory framework, calibrated with data from living taxa, to investigate the effects of time averaging on fossil data [75]. Alternative modes of speciation can be investigated where speciation rates depend on species abundance rather than on species richness [36]. Furthermore, spatially explicit coalescence methods have the potential to investigate the effects of variation in sample size over time and space at large scales, including the possibility of accounting for preservation bias. An important question is the extent to which biological changes result from geographical interferences; neutral theory, with the use of coalescence, has the unique ability to incorporate geographical events at large scales in the absence of selection, a valuable tool that is not used perhaps only because it is not widely known (Box 4).

Phylogenies

The birth–death model of diversification [76] and variations on it have dominated phylogenetic research for nearly 40 years. Surprisingly little has been done with neutral theory to model phylogenies since its original publication [2]. Neutral theory is able to make predictions about phylogenetic patterns and the abundances of the species involved. Methods are available for comparing these with empirical data [25] based on Approximate Bayesian Computation; these show that incorporating the balance of the phylogenetic tree (a summary statistic encompassing the evenness of splitting) into the fitting process is sufficient to change substantially the estimated parameter values of best fit. Other research shows that the phylogenetic trees predicted by neutral theory are, similar to real data, more unbalanced than that predicted by the standard species-level birth–death model [77,78]. However, neutral models seem unable to capture the rate of diversification and its variation with time in a reconstructed phylogeny [78]. This problem therefore requires further attention.

Conservation

Relatively few publications on neutral theory appear in conservation journals, suggesting that neutral theory has little relevance to useful conservation applications [49]. We suggest that the publication trend instead indicates that the value of neutral theory in conservation has gone unrecognised. Indeed, two papers appeared very recently each showing different uses of neutral theory in conservation modelling [79,80]. Neutral models provide a baseline expectation for important questions, such as the effect on biodiversity of habitat fragmentation [70,79,81]. Halley

and Iwasa [80] recently tested neutral expectations of avifaunal extinctions following habitat loss and found that the neutral model performs extremely well. Approximate predictions of species loss under different scenarios of habitat destruction can also be derived from the theory [82], although we note that very different results might have been obtained had protracted speciation instead of point mutation speciation been used.

Neutral theory fosters stronger science by challenging existing paradigms and their assumptions. For example, it recently showed that seemingly identical communities might not be redundant in the way that was previously thought: their removal leads to a gradual cascade of extinctions owing to dispersal limitation and ecological drift [79]. Although important conservation decisions should not be based only on neutral models, neutral predictions could be considered in addition to existing alternatives as they provide a different perspective and highlight the influence of other processes that are easily neglected.

Discussion

History has shown, not only in physics, but also in ecology, that one can make considerable progress with simple models. The ideal gas law, for example, is a good approximation, but there are no 'ideal gases' in reality. The success of the birth–death model of diversification [76], MacArthur and Wilson's theory of island biogeography [10] and Levins' metapopulation model [83] are all testimony to the utility of simple models in ecology and evolution. Neutral theory follows this tradition.

Researchers in this field have been far too focused on snapshots of species abundance distributions, and the heated debate about the validity of the neutrality assumption. This vigorous debate has now spanned a decade (see [14,15,30,49,84–89] for reviews and opinions on both sides). We suggest that the time has now come to put both the theory and its criticisms into perspective and move forward. To use the words of the statistician George Box 'All models are wrong, but some are useful'. No ecologist believes the world is really neutral. Proponents of neutral theory simply assert that neutral models are of value as another addition to the ecologist's toolbox.

A measure of the usefulness of a theory is its ability to fail in informative ways. Neutral theory provides a valuable null model [2,30], a baseline to compare with empirical data and other models. If a neutral model is inconsistent with data, one can learn something about what missing biological factors are needed to improve it. It is harder to draw such conclusions from complex models where, if a rigorous comparison with data is even possible, failure could equally result from inclusion of incorrect details. If neutral theory does in fact fit observations, this does not imply that the world is neutral, but it does call into question both more complex explanations for the data and the power of the data itself. For example, neutral theory has fitted many empirical species abundance distributions and helped to highlight the limitations of species abundance data.

Neutral theory can also be regarded as an approximation of biodiversity patterns and it remains tractable even when faced with extremely large numbers of individual

organisms and high levels of spatial complexity [68]. It can thus be used as a first tractable attempt at modelling an otherwise impenetrable problem [30]. For example, it is the only model to explain the full tri-phasic SAR including the power law as an emergent behaviour [28].

The possibility that neutral theory could contribute significantly to the fields of phylogenetics, population dynamics, island biogeography, paleobiology and conservation has been suggested from the start [2], but hardly any work has been done in these areas. It is our hope that neutral models will start to be applied in these other areas of study. There is need for further analytical results [66,90] and simulation methods [68] if we are to achieve this goal.

Classic neutral theory represents the first step towards improved models that might relax the auxiliary assumptions or the neutrality assumption itself. A topical example is 'symmetric models' in which the outcomes of the model are not affected by exchanging species labels. Neutral models are symmetric, but symmetric models need not be neutral and can encompass more biology. For example, symmetric models might allow the birth and death prospects of an individual to be influenced by the density of conspecifics and heterospecifics in a different way [91], even though the identity of the species does not matter.

Our first-hand view of nature shows the importance of niches. However, these niches have dominated our attention and left less obvious, but still important processes forgotten [92]. Perhaps the most important contribution of neutral theory has been to highlight the key roles of dispersal limitation, speciation and ecological drift, by showing how much can be explained by these processes alone without any niches or selection. We hope that neutral theory will ultimately lead to a truly unified theory of biodiversity [2], including the important effects of selection; this would be a tremendous asset to the field.

Acknowledgements

We thank the editor Paul Craze and two anonymous reviewers for their helpful comments and suggestions. We also thank Bart Haegeman, David Alonso, Egbert Leigh, Jr., Evan Economo, Fangliang He, Jonathan Eastman, Luke Harmon, Matt Pennell, Stephen Cornell and members of the Harmon and Rosenblum lab groups for their useful comments and discussions on the paper. JR was funded by the EPSRC (grant EP/F043112/1); RSE acknowledges financial support from NWO.

References

- Gause, G.F. (1934) *The Struggle For Existence*, Williams and Williams
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press
- Kimura, M. and Crow, J.F. (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49, 725–738
- Caswell, H. (1976) Community structure – neutral model analysis. *Ecol. Monogr.* 46, 327–354
- Hubbell, S.P. (1979) Tree dispersion, abundance, and diversity in a tropical dry forest. *Science* 203, 1299–1309
- Leigh, E.G. (1981) The average lifetime of a population in a varying environment. *J. Theor. Biol.* 90, 213–239
- Hubbell, S.P. (1997) A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs* 16, S9–S21
- Bell, G. (2000) The distribution of abundance in neutral communities. *Am. Nat.* 155, 606–617
- Bell, G. (2001) Ecology – neutral macroecology. *Science* 293, 2413–2418
- MacArthur, R.H. and Wilson, E.O. (1963) An equilibrium theory of insular zoogeography. *Evolution* 17, 373–387
- Etienne, R.S. and Olff, H. (2004) A novel genealogical approach to neutral biodiversity theory. *Ecol. Lett.* 7, 170–175
- Etienne, R.S. and Alonso, D. (2005) A dispersal-limited sampling theory for species and alleles. *Ecol. Lett.* 8, 1147–1156
- Etienne, R.S. (2005) A new sampling formula for neutral biodiversity. *Ecol. Lett.* 8, 253–260
- Alonso, D. *et al.* (2006) The merits of neutral theory. *Trends Ecol. Evol.* 21, 451–457
- McGill, B.J. *et al.* (2006) Empirical evaluation of neutral theory. *Ecology* 87, 1411–1423
- He, F. (2005) Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. *Funct. Ecol.* 19, 187–193
- Etienne, R.S. *et al.* (2007) The zero-sum assumption in neutral biodiversity theory. *J. Theor. Biol.* 248, 522–536
- Rosindell, J. *et al.* (2010) Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol. Lett.* 13, 716–727
- Chisholm, R.A. and Pacala, S.W. (2010) Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proc. Natl Acad. Sci. U.S.A.* 107, 15821–15825
- Etienne, R.S. and Haegeman, B. Independent species in independent niches behave neutrally. *Oikos* (in press)
- Azaele, S. *et al.* (2006) Dynamical evolution of ecosystems. *Nature* 444, 926–928
- Hubbell, S.P. (2009) Neutral theory and the theory of island biogeography. In *The Theory of Island Biogeography Revisited* (Losos, J. and Ricklefs, R.E., eds), pp. 264–292, Princeton University Press
- Woodcock, S. *et al.* (2007) Neutral assembly of bacterial communities. *FEMS Microbiol. Ecol.* 62, 171–180
- Horvát, S. *et al.* (2010) A spatially explicit model for tropical tree diversity patterns. *J. Theor. Biol.* 265, 517–523
- Jabot, F. and Chave, J. (2009) Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. *Ecol. Lett.* 12, 239–248
- Rosindell, J. and Cornell, S.J. (2007) Species–area relationships from a spatially explicit neutral model in an infinite landscape. *Ecol. Lett.* 10, 586–595
- Pigolotti, S. and Cencini, M. (2009) Speciation-rate dependence in species–area relationships. *J. Theor. Biol.* 260, 83–89
- Rosindell, J. and Cornell, S.J. (2009) Species–area curves, neutral models and long distance dispersal. *Ecology* 90, 1743–1750
- Nathan, R. (2006) Long-distance dispersal of plants. *Science* 313, 786–788
- Leigh, E.G. (2007) Neutral theory: a historical perspective. *J. Evol. Biol.* 20, 2075–2091
- Haegeman, B. and Etienne, R.S. (2008) Relaxing the zero-sum assumption in neutral biodiversity theory. *J. Theor. Biol.* 252, 288–294
- Conlisk, J. *et al.* (2010) Hubbell's local abundance distribution: insights from a simple colonization rule. *Oikos* 119, 379–383
- Kopp, M. (2010) Speciation and the neutral theory of biodiversity. *Bioessays* 32, 564–570
- Etienne, R.S. and Haegeman, B. (2011) The neutral theory of biodiversity with random fission speciation. *Theor. Ecol.* 4, 87–109
- MacArthur, R.H. (1957) On the relative abundance of bird species. *Proc. Natl Acad. Sci. U.S.A.* 43, 293–295
- Etienne, R.S. *et al.* (2007) Modes of speciation and the neutral theory of biodiversity. *Oikos* 116, 241–258
- Haegeman, B. and Etienne, R.S. (2010) Self-consistent approach for neutral community models with speciation. *Phys. Rev. E* 81, 031911
- Leigh, E.G. *et al.* (2004) Why do some tropical forests have so many species of trees? *Biotropica* 36, 447–473
- Ricklefs, R.E. (2003) A comment on Hubbell's zero-sum ecological drift model. *Oikos* 100, 185–192
- Hubbell, S.P. and Lake, J. (2002) The neutral theory of biodiversity and biogeography, and beyond. In *Macroecology: Patterns and Process* (Blackburn, T. and Gaston, K., eds), pp. 45–63, Blackwell
- Hubbell, S.P. (2003) Modes of speciation and the lifespans of species under neutrality: a response to the comment of Robert E. Ricklefs. *Oikos* 100, 193–199
- Allen, A.P. and Savage, V.M. (2007) Setting the absolute tempo of biodiversity dynamics. *Ecol. Lett.* 10, 637–646

- 43 Ricklefs, R.E. (2006) The unified neutral theory of biodiversity: do the numbers add up? *Ecology* 87, 1424–1431
- 44 Nee, S. (2005) The neutral theory of biodiversity: do the numbers add up? *Funct. Ecol.* 19, 173–176
- 45 Lande, R. *et al.* (2003) *Stochastic Population Dynamics in Ecology and Conservation*, Oxford University Press
- 46 de Aguiar, M.A.M. *et al.* (2009) Global patterns of speciation and diversity. *Nature* 480, 284–387
- 47 Melian, C.J. *et al.* (2010) Frequency dependent selection predicts patterns of radiations and biodiversity. *PLoS Comput. Biol.* 6, e1000892
- 48 Zillio, T. and Condit, R. (2007) The impact of neutrality, niche differentiation and species input on diversity and abundance distributions. *Oikos* 116, 931–940
- 49 Clark, J.S. (2009) Beyond neutral science. *Trends Ecol. Evol.* 24, 8–15
- 50 Chisholm, R.A. and Lichstein, J.W. (2009) Linking dispersal, immigration and scale in the neutral theory of biodiversity. *Ecol. Lett.* 12, 1385–1393
- 51 Perry, G.L.W. *et al.* (2009) Dispersal, edaphic fidelity and speciation in species-rich western Australian shrublands: evaluating a neutral model of biodiversity. *Oikos* 118, 1349–1362
- 52 Etienne, R.S. and Rosindell, J. (2011) The spatial limitations of current neutral models of biodiversity. *PLoS ONE* 6, e14717
- 53 Etienne, R.S. (2007) A neutral sampling formula for multiple samples and an ‘exact’ test of neutrality. *Ecol. Lett.* 10, 608–618
- 54 Munoz, F. *et al.* (2007) Estimating parameters of neutral communities: from one single large to several small samples. *Ecology* 88, 2482–2488
- 55 Etienne, R.S. (2009) Maximum likelihood estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. *J. Theor. Biol.* 257, 510–514
- 56 Volkov, I. *et al.* (2007) Patterns of relative species abundance in rainforests and coral reefs. *Nature* 450, 45–49
- 57 Economo, E.P. and Keitt, T.H. (2008) Species diversity in neutral metacommunities: a network approach. *Ecol. Lett.* 11, 52–62
- 58 Warren, P.B. (2010) Biodiversity on island chains: neutral model simulations. *Phys. Rev. E* 82, 051922
- 59 Vanpeteghem, D. and Haegeman, B. (2010) An analytical approach to spatio-temporal dynamics of neutral community models. *J. Math. Biol.* 61, 323–357
- 60 Muneeppeerakul, R. *et al.* (2008) Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin. *Nature* 453, 220–222
- 61 Holley, R.A. and Liggett, T.M. (1975) Ergodic theorems for weakly interacting systems and the voter model. *Ann. Probab.* 3, 643–663
- 62 Bramson, M. *et al.* (1998) A spatial model for the abundance of species. *Ann. Probab.* 26, 658–709
- 63 Durrett, R. and Levin, S. (1996) Spatial models for species–area curves. *J. Theor. Biol.* 179, 119–127
- 64 Zillio, T. *et al.* (2005) Spatial scaling in model plant communities. *Phys. Rev. Lett.* 95, 098101
- 65 Chave, J. and Leigh, E.G. (2002) A spatially explicit neutral model of beta-diversity in tropical forests. *Theor. Popul. Biol.* 62, 152–168
- 66 O’Dwyer, J.P. and Green, J.L. (2010) Field theory for biogeography: a spatially explicit model for predicting patterns of biodiversity. *Ecol. Lett.* 13, 87–95
- 67 Kingman, J.F.C. (1982) On the genealogy of large populations. *J. Appl. Probab.* 19, 27–43
- 68 Rosindell, J. *et al.* (2008) A coalescence approach to spatial neutral ecology. *Ecol. Inform.* 3, 259–271
- 69 Rosindell, J. and Phillimore, A.B. (2011) A unified model of island biogeography sheds light on the zone of radiation. *Ecol. Lett.* 14, 552–560 DOI: 10.1111/j.1461-0248.2011.01617.x
- 70 Gilbert, B. *et al.* (2006) Can neutral theory predict the responses of Amazonian tree communities to forest fragmentation? *Am. Nat.* 168, 304–317
- 71 Vanpeteghem, D. *et al.* (2008) Dynamics of neutral biodiversity. *Math. Biosci.* 212, 88–98
- 72 Keil, P. *et al.* (2010) Predictions of Taylor’s power law, density dependence and pink noise from a neutrally modeled time series. *J. Theor. Biol.* 265, 78–86
- 73 Adler, P.B. (2004) Neutral models fail to reproduce observed species–area and species–time relationships in Kansas grasslands. *Ecology* 85, 1265–1272
- 74 Hubbell, S.P. (2005) The neutral theory of biodiversity and biogeography and Stephen Jay Gould. *Paleobiology* 31, 122–132
- 75 Tomasovych, A. and Kidwell, S.M. (2010) Predicting the effects of increasing temporal scale on species composition, diversity, and rank-abundance distributions. *Paleobiology* 36, 672–695
- 76 Raup, D.M. *et al.* (1973) Stochastic-models of phylogeny and evolution of diversity. *J. Geol.* 81, 525–542
- 77 Mooers, A.Ø. *et al.* (2007) Some models of phylogenetic tree shape. In *Reconstructing Evolution: New Mathematical and Computational Advances* (Gascuel, O. and Steel, M., eds), pp. 149–170, Oxford University Press
- 78 Davies, T.J. *et al.* (2011) Neutral biodiversity theory can explain the imbalance of phylogenetic trees but not the tempo of their diversification. *Evolution* DOI: 10.1111/j.1558-5646.2011.01265.x
- 79 Economo, E.P. (2011) Biodiversity conservation in metacommunity networks: linking pattern and persistence. *Am. Nat.* 177, doi:10.1086/659946 (in press)
- 80 Halley, J.M. and Iwasa, Y. (2011) Neutral theory as a predictor of avifaunal extinctions after habitat loss. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2316–2321
- 81 Leigh, E.G. *et al.* (1993) The decline of tree diversity on newly isolated tropical islands – a test of a null hypothesis and some implications. *Evol. Ecol.* 7, 76–102
- 82 Hubbell, S.P. *et al.* (2008) How many tree species are there in the Amazon and how many of them will go extinct? *Proc. Natl. Acad. Sci. U.S.A.* 105, 11498–11504
- 83 Levins, R. (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.* 15, 237–240
- 84 Abrams, P.A. (2001) A world without competition. *Nature* 412, 858–859
- 85 Adler, P.B. *et al.* (2007) A niche for neutrality. *Ecol. Lett.* 10, 95–104
- 86 Purves, D.W. and Turnbull, L.A. (2010) Different but equal: the implausible assumption at the heart of neutral theory. *J. Anim. Ecol.* 79, 1215–1225
- 87 Leigh, E.G. *et al.* (2010) Unified neutral theory of biodiversity and biogeography. *Scholarpedia* 5, 8822
- 88 McGill, B.J. (2010) Towards a unification of unified theories of biodiversity. *Ecol. Lett.* 13, 627–642
- 89 Ricklefs, R.E. (2011) Applying a regional community concept to forest birds of eastern North America. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2300–2305
- 90 Allouche, O.O. and Kadmon, R. (2009) A general framework for neutral models of community dynamics. *Ecol. Lett.* 12, 1287–1297
- 91 Volkov, I. *et al.* (2005) Density dependence explains tree species abundance and diversity in tropical forests. *Nature* 438, 658–661
- 92 Vellend, M. (2010) Conceptual synthesis in community ecology. *Q. Rev. Biol.* 85, 183–206
- 93 Etienne, R.S. and Alonso, D. (2007) Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. *J. Stat. Phys.* 128, 485–510
- 94 Volkov, I. *et al.* (2003) Neutral theory and relative species abundance in ecology. *Nature* 424, 1035–1037
- 95 Vallade, M. and Houchmandzadeh, B. (2003) Analytical solution of a neutral model of biodiversity. *Phys. Rev. E* 68, 061902
- 96 Alonso, D. and McKane, A.J. (2004) Sampling Hubbell’s neutral theory of biodiversity. *Ecol. Lett.* 7, 901–910
- 97 Ewens, W.J. (1972) Sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3, 87–112
- 98 Haegeman, B. and Etienne, R.S. (2009) Neutral models with generalised speciation. *Bull. Math. Biol.* 71, 1507–1519
- 99 Chave, J. *et al.* (2002) Comparing classical community models: theoretical consequences for patterns of diversity. *Am. Nat.* 159, 1–23
- 100 Condit, R. *et al.* (2002) Beta-diversity in tropical forest trees. *Science* 295, 666–669