Comparing strategies to preserve evolutionary diversity

Karen Magnuson-Ford [a,b,c], Arne Mooers [c], Sébastien Rioux Paquette [d,e], Mike Steel [b,d],*

[a] Zoology, University of British Columbia, 2212 Main Mall, Vancouver, BC, Canada V6T 1Z4[1]
[b] Biomathematics Research Centre, University of Canterbury, Christchurch, New Zealand
[c] IRMACS and Biological Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6
[d] Allan Wilson Centre for Molecular Ecology and Evolution, New Zealand
[e] Biological Sciences, Victoria University of Wellington, Wellington, New Zealand

ABSTRACT

The likely future extinction of various species will result in a decline of two quantities: species richness and phylogenetic diversity (PD, or 'evolutionary history'). Under a simple stochastic model of extinction, we can estimate the expected loss of these quantities under two conservation strategies: An 'egalitarian' approach, which reduces the extinction risk of all species, and a 'targeted' approach that concentrates conservation effort on the most endangered taxa. For two such strategies that are constrained to experience the same expected loss of species richness, we ask which strategy results in a greater expected loss of PD. Using mathematical analysis and simulation, we describe how the strategy (egalitarian versus targeted) that minimizes the expected loss of PD depends on the distribution of endangered status across the tips of the tree, and the interaction of this status with the branch lengths. For a particular data set consisting of a phylogenetic tree of 62 lemur species, with extinction risks estimated from the IUCN 'Red List', we show that both strategies are virtually equivalent, though randomizing these extinction risks across the tip taxa can cause either strategy to outperform the other. In the second part of the paper, we describe an algorithm to determine how extreme the loss of PD for a given decline in species richness can be. We illustrate the use of this algorithm on the lemur tree.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

An important measure of the biodiversity of a group of organisms is the collective evolutionary breadth that they represent (Purvis et al., 2000; Vane-Wright et al., 1991). At the limit, under conditions of triage (Marris, 2007), it may seem prudent to consider focussing on subsets of species that capture the maximum amount of evolutionary history, quantified as the total length of the evolutionary subtree connecting them to the tree of life (phylogenetic diversity or PD; Faith, 1992). Nee and May (1997) used model trees to suggest that there was little to be gained by focussing on conserving particular species: random subsets of species, or those left after a 'field of bullets' scenario of extinction, captured nearly as much total PD as that captured when an optimal set of species was retained. Subsequent simulation (Heard and Mooers, 2000) and empirical work (Purvis and Hector, 2000; Vamosi and Wilson, 2008; von Euler, 2001) suggested that future predicted losses could greatly exceed the 'field of bullets' scenario, meaning optimum resource allocation might yield increased returns. We return to this question here,

and ask how resource allocation might best be used to decrease the loss of PD through extinction.

Our paper has two parts. In the first, we use a new and flexible approach for describing the probabilities of extinction of lineages through time to compare how much PD can be preserved if conservation resources are applied fairly across all the species in a clade, versus concentrating effort on the most 'at risk' species. In the second part, we describe an algorithm for identifying the subset of species whose collective loss would contribute most to the loss of PD from a focal clade. For both, we offer analytical and simulation results, and then apply the approach to a new and complete tree of a charismatic and highly endangered fauna, the lemurs of Madagascar. We note that our approaches require that all the biodiversity in a clade is enumerated such that the tree is complete; such trees are becoming more common, but do require that the conservation units of biodiversity (e.g. 'species' or 'lineages') have been fairly delimited (see Agapow et al., 2004 for a discussion of this issue).

2. Extinction models and the expected loss of phylogenetic diversity

We first begin by describing a generalization of the 'field of bullets' stochastic model of species extinction (Nee and May,

1997; Raup, 1993). Suppose we have a collection X of species and each species $x \in X$ undergoes extinction independently according to a non-stationary death process, at rate $r_x(t)$. We suppose that the present time is $t=0$, and time is measured forward into the future. Allowing $r_x(t)$ to vary with time allows for: (i) changing environmental conditions (e.g. climate change) and anthropogenic pressures that may alter extinction risk and (ii) changes in the timing or intensity of conservation measures. Let $p_x(t)$ denote the probability that x is extant (i.e. non-extinct) at time $t \geq 0$. This is given by the well-known formula for a non-stationary Poisson process:

$$p_x(t) = \exp\left(-\int_0^t r_x(u)\,du \right). \tag{1}$$

In the special case where $r_x(t) = r_x$, i.e. a constant extinction rate over time, but possibly variable from species to species, one has simply $p_x(t) = e^{-r_x t}$. Notice that $p_x(t)$ decreases monotonically with increasing t, and it converges to zero if and only if $\int_0^t r_x(u)\,du$ diverges (i.e. tends to infinity). This is possible even if the extinction risk reduces towards a limit of 0 with increasing time; for example, if $r_x(t) = (t+1)^{-\gamma}$ for $\gamma \leq 1$ (but not for $\gamma > 1$), then the integral diverges. The divergence of this integral is the condition for 'guaranteed eventual extinction'.

Across all species, this extinction process can leave a system with any subset of species at a future point in time, t. That is, this process induces a continuous-time non-stationary Markov process Y_t on the state space 2^X (the set of all subsets of X) defined by $Y_0 = X$ (with probability 1) and:

$$\mathbb{P}(Y_t = Y) = \prod_{x \in Y} p_x(t) \prod_{x \in X-Y} (1 - p_x(t)),$$

where $p_x(t)$ is given by Eq. (1). This provides further extension of the 'generalized field of bullets model' (g-FOB) from Faller et al. (2008) to allow the extinction probabilities to vary with time, but is still based on independence of extinction events among taxa. So, though we have introduced a generalization, for what follows we use the simpler standard constant extinction rate, $r_x(t) = r_x$.

Suppose we have a rooted phylogenetic X-tree $\mathcal{T} = (V,E)$ with a set of vertices (i.e. nodes) V, edges E and a branch length $l(e)$ assigned to each edge $e \in E$ of $\mathcal{T}$ (we do not necessarily assume the tree is ultrametric). For any subset S of X let $PD(S)$ denote the phylogenetic diversity of S (the sum of the branch lengths of the edges connecting the species in S and the root of the tree—see Fig. 1 for an example). Let $\psi_t = PD(Y_t)$ denote the phylogenetic

diversity of the subset Y_t of species that are extant at future time t. Then, as in Faller et al. (2008), we have

$$\mathbb{E}[\psi_t] = PD(X) - \sum_{e \in E} l(e) \prod_{x \in C(e)} (1 - p_x(t)),$$

where $p_x(t)$ is given by Eq. (1), $C(e)$ is the set of species (i.e. subset of X, the leaf set of $\mathcal{T}$) descendant from e, and $PD(X)$ denotes the total length of the tree (i.e. ψ_0). In the case where $r_x(t) = r$ (constant) then $\mathbb{E}[\psi_t]$ is convex except, possibly, for small values of t (Hartmann and Steel, 2007). The biological significance of this convexity is that we expect most of the phylogenetic diversity to occur earlier rather than later. This might appear to be at odds with a finding from Nee and May (1997) that most loss in expected phylogenetic diversity occurs 'late'; however, there is no contradiction—in Nee and May (1997) the expected phylogenetic diversity was a function of the number of species extinctions, rather than of time, and most species will tend to extinct early under a model in which the extinction rate is time-independent.

An alternative measure of diversity is the 'exclusive molecular phylodiversity' measure (EP) of Lewis and Lewis (2005) that assigns to each subset S of X the value:

$$EP(S) := PD(X) - PD(X-S).$$

That is, $EP(S)$ measures how much phylogenetic diversity would be lost if the species in S were to become extinct. Note that we can write $EP(S)$ as the sum of the lengths of the branches of $\mathcal{T}$ for which all of the species descendent from that branch are in S; thus $EP(S)$ is much more conservative than (and smaller than) $PD(S)$ for which the corresponding branch sum expression replaces the word 'all' by 'at least one'. In Section 6, we give an algorithm for finding sets S of a given size that, if lost, would diminish phylogenetic diversity more than any other set of the same size (i.e. maximizes $EP(S)$). The EP measure is also illustrated in Fig. 1.

Let $\phi_t = EP(Y_t)$ denote the exclusive phylodiversity of the subset Y_t of species that are extant at time t. We have

$$\mathbb{E}[\phi_t] = \sum_{e \in E} l(e) \prod_{x \in C(e)} p_x(t).$$

Notice that if we define the (total) extinction rate of a clade $C(e) \subset X$ as $R_{C(e)}(t) := \sum_{x \in C(e)} r_x(t)$ then:

$$\mathbb{E}[\phi_t] = \sum_{e \in E} l(e) \exp(-R(t)).$$

In particular, if the extinction rates $r_x(t)$ are not time-dependent, then for any tree and any selection of branch lengths, $\mathbb{E}[\phi_t]$ is strictly convex for all $t \geq 0$ on any tree, since it is a convex combination of decaying exponential functions.

3. Two strategies for maximizing expected PD or EP

Suppose our goal is to increase the average probability of survival p_x across all species in a collection, X, by some amount. We partition the species X into two classes: E, a set of endangered species at a higher risk of extinction, and N, a set of non-endangered species which have smaller, but still positive extinction rates. In this section, we address the following somewhat general question: Is the expected future phylogenetic diversity higher if we concentrate all our efforts on just the endangered species, than if we simply try to help all species equally? To make this more precise, consider the following two strategies:

(S_e) *The 'egalitarian' strategy, where we help all species a little*:
Under this strategy, we reduce the extinction rate function $r_x(t)$ of each species x by multiplying $r_x(t)$ by a small positive number $\alpha < 1$. We can denote this as S_e.
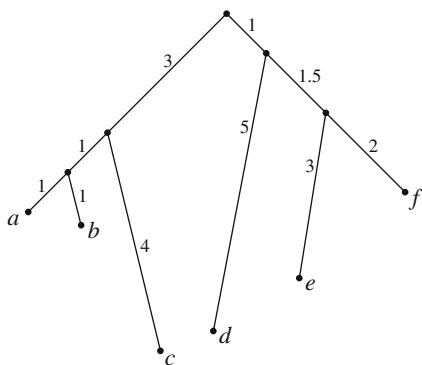

Fig. 1. A simple example to illustrate the measures of phylogenetic diversity (PD) and exclusive molecular phylodiversity (EP) for a species subset. The maximum PD for a subset of three species is 17.5; that is, the total length of the subtree connecting species c, d, and e. The EP measure for this species subset is 12. In this case, no internal branch lengths are included in EP since the descendent species of any one internal branch are not exclusively those in the species subset. The maximum EP for three species, which can be calculated using the algorithm in Section 6, is 12.5, including species d, e, and f.

(S_t) The 'targeted' strategy, where we focus help on endangered species: In this strategy we leave the extinction rate function of each species in N unchanged, but multiply the extinction rate function $r_x(t)$ of each species in E by a small non-negative number β where β is strictly less than α. We can denote this as S_t.

Notice that if we let $p_x = p_x(T)$, the probability that species x is extant at time T, then multiplying $r_x(t)$ by a constant α simply converts p_x to p_x^α since, by (1),

$$\exp\left(-\int_0^T \alpha r_x(u)\,du\right) = \exp\left(-\alpha \int_0^T r_x(u)\,du\right)$$
$$= \left(\exp\left(-\int_0^T r_x(u)\,du\right)\right)^\alpha = p_x(T)^\alpha.$$

Thus we can compactly describe the two strategies by the following transformation table:

	N	E
S_e	$p_x \mapsto p_x^\alpha$	$p_x \mapsto p_x^\alpha$
S_t	$p_x \mapsto p_x$	$p_x \mapsto p_x^\beta$

In order to compare these two strategies, we need to relate β to α. This might be done according to various ways of assessing the cost of decreasing the extinction rate, along the lines of 'Noah's Ark problem' (Hartmann and Steel, 2007; Weitzman, 1998). In this paper, we take a different approach and we determine which strategy leads to higher expected phylogenetic diversity at time T in the future when β is chosen according to the following rule:

The expected total number of species at time T is the same under the two strategies.

While conservationists often choose strategies to maximize the total number of species saved, we chose the above rule to investigate how different strategies conserve PD even when keeping species richness constant between strategies. Our rule means we are comparing strategies that increase the average probability of survival for the collection of Species X by the same amount.

The expected total number of species at time T under S_e is $\sum_{x \in X} p_x^\alpha$, while the expected total number of species at time T under S_t is $\sum_{x \in N} p_x + \sum_{x \in E} p_x^\beta$. Thus β is constrained to satisfy the equation:

$$\sum_{x \in X} p_x^\alpha = \sum_{x \in N} p_x + \sum_{x \in E} p_x^\beta, \tag{2}$$

which has a non-negative solution for β provided that α is not too small. The precise condition on α for Eq. (2) to have a solution for β that is non-negative is that the following inequality holds:

$$\sum_{x \in X} p_x^\alpha \leq |E| + \sum_{x \in N} p_x. \tag{3}$$

To see this, let $s := \sum_{x \in X} p_x^\alpha - \sum_{x \in N} p_x$. When Eq. (2) holds, s is just the right-most term in Eq. (2), the expected number of E species. Now, let $f(\beta) := \sum_{x \in E} p_x^\beta - s$. Notice that f is monotone decreasing as $\beta \geq 0$ increases. If Eq. (2) has a non-negative solution for β then the two terms in $f(\beta)$ are exactly the same, and $f(\beta) = 0$, and so $0 = f(\beta) \leq f(0) = |E| - \sum_{x \in X} p_x^\alpha + \sum_{x \in N} p_x$, which gives (3). Conversely, if (3) holds, then $f(0) = |E| - s \geq 0$ and $f(\alpha) = \sum_{x \in N}(p_x - p_x^\alpha) < 0$. So, by the monotonicity of f, a unique value β between 0 and α exists for which $f(\beta) = 0$; this is the unique solution of Eq. (2).

A particular case of interest is when N and E are both divided into discrete categories, such as in the IUCN categories: 'Least Concern', 'Near Threatened', 'Vulnerable', 'Endangered' and 'Critically Endangered'. Thus E might consist of the last two categories, or the last three, or perhaps just the last one. In each

case N would include the less threatened categories. In the particular case where E consists of just one category (in the IUCN case, this would be 'critically endangered') and all its species are assumed to have the same extinction probability p, then there is an exact explicit formula for β. More precisely, if Eq. (2) has a non-negative solution for β, it is given by

$$\beta = \frac{\log\left(p^\alpha + \frac{1}{|E|}\sum_{x \in N}(p_x^\alpha - p_x)\right)}{\log(p)}. \tag{4}$$

In the case where E consists of the top two (or three) IUCN categories, Eq. (2) leads to a more complex equation, that does not have an explicit solution for β but which still can easily be solved by standard numerical methods. For the examples below, we consider E to contain only the single most threatened category as the most conservative example for comparing the two allocation strategies.

Consider now the expected PD at time T under the egalitarian and targeted strategies, which we write as $\mathbb{E}_e[\psi]$ and $\mathbb{E}_t[\psi]$, respectively. We have

$$\mathbb{E}_e[\psi] = PD(X) - \sum_{e \in E} l(e) \prod_{x \in C(e)} (1 - p_x^\alpha),$$

and

$$\mathbb{E}_t[\psi] = PD(X) - \sum_{e \in E} l(e) \prod_{x \in C(e) \cap N} (1 - p_x) \prod_{x \in C(e) \cap E} (1 - p_x^\beta).$$

We wish to compare $\mathbb{E}_e[\psi]$ and $\mathbb{E}_t[\psi]$ under the assumption that the expected number of species present at time T is the same (i.e. under the constraint (2)). As we might expect, the outcome will depend on the distribution of the species in N and E among the leaves of the phylogenetic tree, and the branch lengths associated with that tree.

Similarly, for expected EP at time T under the two scenarios, which we write as $\mathbb{E}_e[\phi]$ and $\mathbb{E}_t[\phi]$, respectively, we have

$$\mathbb{E}_e[\phi] = \sum_{e \in E} l(e) \prod_{x \in C(e)} p_x^\alpha \quad \text{and} \quad \mathbb{E}_t[\phi] = \sum_{e \in E} l(e) \prod_{x \in C(e) \cap N} p_x \prod_{x \in C(e) \cap E} p_x^\beta.$$

3.1. Two idealized settings

We now consider two very simple scenarios where one can obtain exact equations and we have instructive lower bounds for the difference:

$$\Delta := \mathbb{E}_t[\psi] - \mathbb{E}_e[\psi].$$

In both cases, we adopt the simplifying assumption that p_x takes a constant value q within N and also a constant value $p < q$ within E, and so Eq. (2) reduces to

$$q^\alpha - q = \frac{e}{n}(p^\beta - p^\alpha), \tag{5}$$

where $n = |N|, e = |E|$, and which has a valid solution for β precisely if

$$p^\alpha + \frac{n}{e}(q^\alpha - q) \leq 1.$$

In this case, there is an explicit formula for β given (as a special case of Eq. (4)) by

$$\beta = \frac{\log\left(p^\alpha + \frac{n}{e}(q^\alpha - q)\right)}{\log(p)}. \tag{6}$$

Notice that in this special setting, if $q = 1$ (i.e. non-endangered species have no chance of extinction) then the two scenarios become identical, since $\beta = \alpha$ in this case. Indeed this is always true when the species in N are safe from extinction.

Example 1. Early radiation (and without a molecular clock).

Consider a star phylogeny that has just one interior vertex, its root. Thus all the leaves are adjacent to this root vertex (Fig. 2(a)). Let L_N denote the sum of the lengths of the branches incident with species in N and let L_E denote the sum of the branches incident with species in E:

$$\mathbb{E}_e[\psi] = L_N q^\alpha + L_E p^\alpha;$$

$$\mathbb{E}_t[\psi] = L_N q + L_E p^\beta.$$

Applying (5) and performing elementary algebra shows that:

$$\Delta = e(p^\beta - p^\alpha)\left(\frac{L_E}{e} - \frac{L_N}{n}\right),$$

and since $(p^\beta - p^\alpha) > 0$, we see that Δ is positive (i.e. the targeted strategy is 'better') precisely if the average branch length of the endangered species is greater than that for the non-endangered species.

Example 2. Late radiation with endangered outgroup (and with a molecular clock).

Consider a tree that has one non-root vertex v from which all the non-endangered species have radiated both recently and rapidly (Fig. 2(b); we will assume the sum of the branch lengths in this rapid radiation is equal to zero to simplify calculations, but we can extend this as required). Adjacent to the root we also have a single endangered species, whose branch length is L. Let L also be the length of the path from the root to each leaf in N (i.e. we have a molecular clock).

We have

$$\mathbb{E}_e[\psi] = L(1 - (1 - q^\alpha)^n) + L p^\alpha;$$

$$\mathbb{E}_t[\psi] = L(1 - (1 - q)^n) + L p^\beta;$$

and so:

$$\Delta = L(p^\beta - p^\alpha + x^n - y^n); \tag{7}$$

where $x = (1 - q^\alpha)$ and $y = (1 - q)$. Now, if we apply the algebraic identity:

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + \cdots + y^{n-1})$$

and note that each of the n terms in the second bracket (for our choice of x, y) are positive and less than y^{n-1} (since $x < y$), we have

$$x^n - y^n > -(q^\alpha - q)ny^{n-1},$$

and so, from (7), and the following identity (from (5)):

$$p^\beta - p^\alpha = n(q^\alpha - q) \tag{8}$$

we have

$$\Delta > L(n(q^\alpha - q) - (q^\alpha - q)ny^{n-1}) = Ln(q^\alpha - q)(1 - (1 - q)^{n-1}).$$

Thus, again from (8), we have

$$\Delta > L(p^\beta - p^\alpha)(1 - (1 - q)^{n-1}) \geq Lq(p^\beta - p^\alpha).$$

Therefore, for this scenario, the targeted strategy always yields more expected future PD than the egalitarian approach. The last inequality gives an explicit lower bound on the difference of the expected future PD values.

4. Application to phylogeny of lemurs

In order to compare the egalitarian and targeted strategies in a more realistic setting, we used a complete phylogeny of lemurs. Lemurs are a very diverse and charismatic group and thus are well studied in many aspects of their biology, including molecular systematics. As lemurs face extreme habitat loss (Mittermeier et al., 2008) as well as a recent increase in poaching (Barrett and Ratsimbazafy, 2009), many species have become the recipients of substantial conservation effort. These factors make lemurs an ideal group to compare strategies (S_e) and (S_t).

We constructed a phylogeny based on five mitochondrial genes for nearly all the species of lemurs listed in the field guide *Lemurs of Madagascar* (Mittermeier et al., 2006). We omitted sub-species and those species whose descriptions are based solely on morphological and geographical range data. Thus our phylogeny includes 62 lemur species and two out-groups containing sequences from the *Nycticebus* and *Otolemur* genera. While other recent accounts of lemur diversity may claim a higher number of lemur species (see e.g. Mittermeier et al., 2008), recent changes to lemur taxonomy are largely based on changes in the criteria used to distinguish unique lemur species rather than the discovery of new lemur populations (Tattersall, 2007). As the systematics of lemurs are more fully understood, and if and when conservation efforts are differentiated among more finely delineated lineages, these analyses can be updated.

We used Bayesian analysis to produce a fully resolved phylogeny (Fig. 3; for a complete outline of our methods, see Appendix A). For this phylogeny 47/61 nodes have posterior probabilities equal to 1 and only seven nodes have posterior probabilities less than or equal to 0.92. We used the 50% majority-rule consensus tree for all subsequent analyses. The sum of all branch lengths (i.e. the PD of the tree) equals 5.97 substitutions/site. We note that since we are using an additive tree, PD is
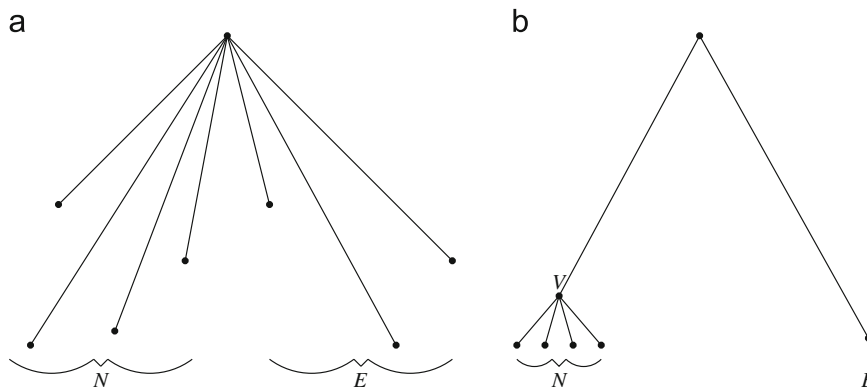

Fig. 2. Two cartoon trees used as examples for how management of endangered (E) and non-endangered (N) species lead to differences in future PD. See main text for details.
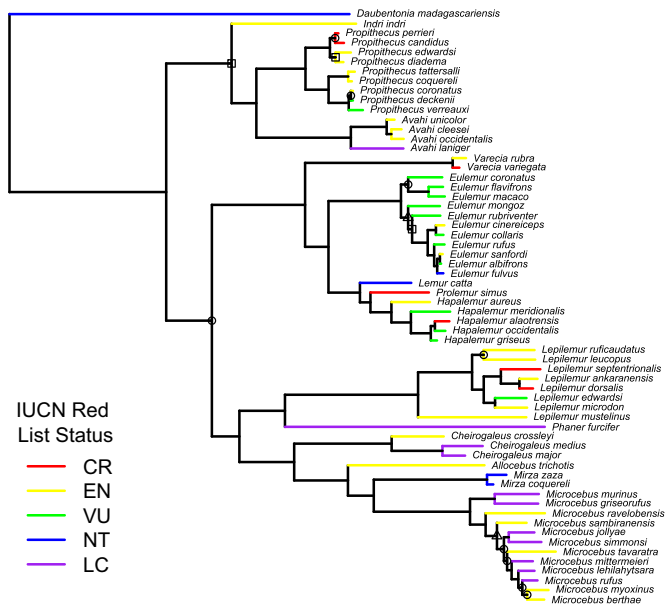
Fig. 3. The lemur phylogeny showing the current IUCN (2009) Red List status for each species. The posterior probabilities for internal nodes are 100% support unless otherwise labelled with a circle: 90–99%, a triangle: 70–89%, or a square: 50–69%.

Table 1
Species' probabilities of extinction based on IUCN categories.

IUCN category	Probability of extinction
Least Concern	0.0001
Near Threatened	0.01
Vulnerable	0.1
Endangered	0.667
Critically Endangered	0.999

measured in units of the number of expected substitutions across the five genes used in the Bayesian analysis. The broad topological outline of this phylogeny is consistent with the recently published primate phylogeny by Fabre et al. (2009), although there are some minor differences within genera.

We used the most recent IUCN (2009) Red List to infer extinction probabilities for each lemur species included in our phylogeny. In the case of *Eulemur flavifrons*, which is recognized as a sub-species of *Eulemur macaco* by the IUCN Red List, we assigned this species the same category as *E. macaco*. Of the 62 species, seven are listed as Critically Endangered (11%), 25 are Endangered (40%), 14 are Vulnerable to extinction (23%), 5 are Near Threatened (8%) and 11 are of Least Concern (18%). We dealt with the species classified as Data Deficient ($n = 15$) by inferring reasonable IUCN categories (see Appendix A for details). We transformed the IUCN Red List status according to the projected probability of extinction in the next 100 years given by the IUCN (2001) and interpolated by Mooers et al. (2008) as presented in Table 1. The extinction risks are also indicated in Fig. 3.

To calculate the expected future PD of the lemur phylogeny under both strategies ($\mathbb{E}_e[\psi]$ and $\mathbb{E}_t[\psi]$), we designated the endangered species set E to be those species listed as Critically Endangered and the non-endangered species set N to include all other species. Unlike the idealized setting above where only two p_x categories were assumed, here we used five: p_x for all species in the E category is set to 0.001, while p_x for species in the N category have the p_x associated with their IUCN status as in Table 1. We set $\alpha = 0.7$ and using Eq. (3), which ensures that the expected number

of remaining species is the same between strategies, we calculated $\beta = 0.091$. Since a low α implies a large change in the probabilities of extinction following an intervention and thus increases the potential for a difference between strategies, we used $\alpha = 0.7$ as it is the lowest value that still yields a meaningful (i.e. positive) value for β. However, even under this low α, the difference between the two strategies for lemurs is only 0.22% of the mean expected future PD (E(PD) is 4.896 under S_e versus 4.885 under S_t; Fig. 4). When the IUCN statuses were shuffled among the tips, the mean expected future PD under S_e is 4.774 and is 4.763 under S_t. While these mean values are very similar, the randomizations of the extinction risks show that there can be considerable variation between strategies. The fact that the actual expected future PD is higher than the mean across randomizations indicates that the Critically Endangered species tend to be spread out evenly across the tree and/or are on short branches. When we assumed the extreme cases that all species classified as Data Deficient are either at zero risk of extinction or classified as Endangered, we observed the same pattern as above, though here the targeted strategy yielded slightly higher expected future PD in both cases (zero risk: E(PD) under S_e = 5.279, under S_t = 5.298; Endangered: E(PD) under S_e = 4.797, under S_t = 4.80).

5. Simulated phylogenies

In addition to considering the lemur tree above, we simulated 1000 trees with 100 tips each, under the Yule model using apTreeshape (Bortolussi et al., 2006) in the statistical package 'R' (R Development Core Team, 2008). The Yule model produces modest variation in tree balance and edge length distributions between the extremes of adaptive radiation (Phillimore and Price, 2008) and equilibrium models (Nee and May, 1997). We randomly assigned an IUCN Red List category to each tip, where the proportion of tips in each category within a given tree corresponds to the overall proportions of all animal species in each category (summarized in Table 2; downloaded from www. iucnredlist.org, November 4, 2009). As before, we assigned CR species to the E category, and all other species to the N category, with all attendant p_x from Table 1.

We calculated the expected future PD under both strategies for every tree, using the value $\alpha = 0.300$, which implies that $\beta = 0.0171$ (Fig. 5(a)). Again, this low value of α was chosen to increase the potential for a difference between strategies. The mean expected future PD under the 'egalitarian' and the 'targeted' strategies was 4655.1 ± 60.21 and 4652.0 ± 53.72, respectively. This constitutes an average reduction in the total PD of 7.80% (under S_e) and 7.86% (S_t). On average, an egalitarian intervention produced an expected future PD value that was higher than that produced under the targeted strategy by only $0.065 \pm 1.34\%$ of the mean expected PD. The differences between the strategies were normally distributed around the mean: 51.9% of the time, the egalitarian intervention conserved more future PD than a targeted intervention; 48.1% of the time, the reverse was true. So, while the mean difference between strategies was negligible, outcomes for individual cases can be quite different, at least with this low value of α (Fig. 5(a)). When we used a value of alpha that was closer to 1, the difference between alpha and the corresponding beta value decreased and thus we observe less variation around the mean (Fig. 5(b)).

6. An algorithm to maximize EP and its application

In this section, we describe and apply a fast algorithm, based on dynamic programming, to find subsets of species of a given
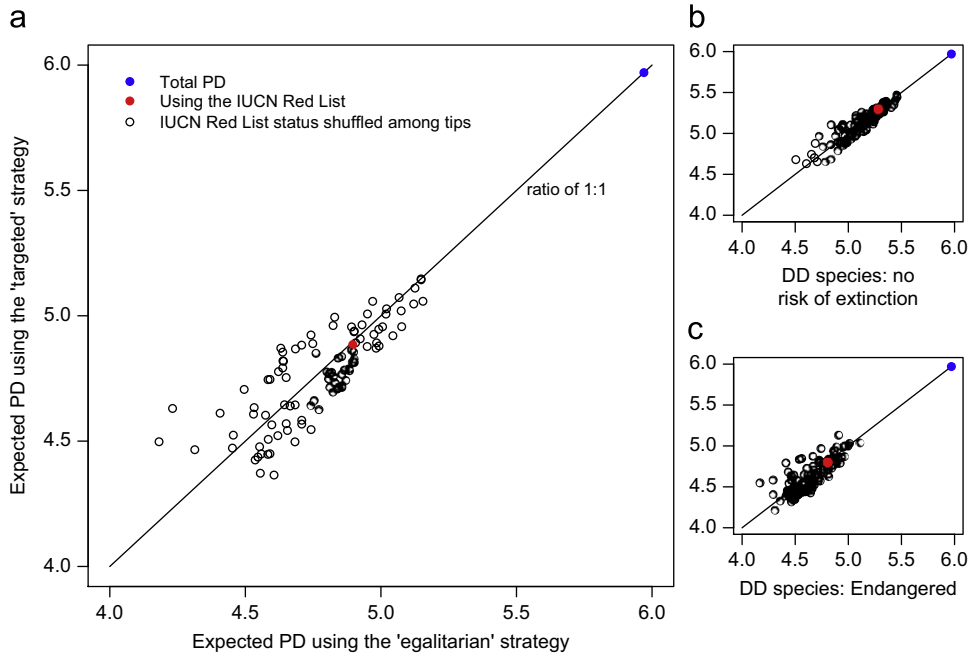
Fig. 4. The expected future PD calculated under both strategies where $\alpha = 0.7$ and $\beta = 0.091$. The species classified as Data Deficient (DD) were treated three ways: (a) their IUCN Red List statuses were inferred ($\beta = 0.091$) and (b) they were considered as having zero risk of extinction ($\beta = 0.12$) or (c) they were considered endangered species ($\beta = 0.038$).

Table 2
Number of animal species within each IUCN category (IUCN, 2009).

IUCN red list category	# species	%
Least Concern	17,535	61
Near Threatened	2574	9
Vulnerable	4467	15
Endangered	2573	9
Critically Endangered	1742	6
Total	28,891	100

size that have a maximal EP score. The fact that dynamic programming can be used to solve this problem was noted in Spillner et al. (2008); here, we provide an explicit description and illustrate its use on the lemur tree.

We first note that maximizing EP is quite a different problem to maximizing phylogenetic diversity, for which there is a simple greedy algorithm (Steel, 2005; Pardi and Goldman, 2005). The problems are related: finding a set of maximal EP score containing a given number k of taxa, selected from a total taxon set of size n, is equivalent to finding a subset of $n-k$ taxa of minimal phylogenetic diversity. However, although maximizing phylogenetic diversity can be solved greedily, minimizing phylogenetic diversity cannot, hence the need for a more sophisticated algorithm.

The algorithm proceeds from the leaves to the root. For each vertex v of the tree that has m species below it, we will compute an $(m+1)$-tuple of pairs $(\varepsilon_0, S_0), (\varepsilon_1, S_1), \ldots, (\varepsilon_m, S_m)$ where ε_i is the maximal EP score possible in the subtree rooted at v if we select a subset of i species from the tips of that subtree, and S_i is a set of such i species that have a maximal EP score. Clearly, $(\varepsilon_0 = 0, S_0) = (0, \emptyset)$ and $(\varepsilon_m, S_m) = (n(v), PD(v))$, where $n(v)$ is the number of tips species below v, and $PD(v)$ is the sum of the lengths of the branches below v.

The base case for the algorithm is a leaf, x for which $m=1$, $\varepsilon_0 = \varepsilon_1 = 0$, $S_0 = \emptyset$, $S_1 = \{x\}$. Now, suppose that we have computed the tuples for the children $w_1, \ldots, w_k$ of a vertex v. Let λ_i

be the length of the branch from v to w_i. We will suppose here that $k=2$ (i.e. a binary tree) though it is possible to describe a more complex algorithm for non-binary trees.

Let (ε_i, S_i) and (ε'_i, S'_i) be the tuples assigned to w_1 and w_2, respectively. For $i = 0, 1, \ldots, n(v)$ let:

$$\varepsilon_i^{(0)} := \max\{\varepsilon_j + \varepsilon'_k : 0 \leq j < n(w_1), 0 \leq k < n(w_2), j+k = i\},$$

and let

$$\varepsilon_i^{(1)} := \begin{cases} PD(w_1) + \lambda_1 + \varepsilon'_{i-n_1} & \text{if } n_1 \leq i, \\ 0 & \text{otherwise}; \end{cases}$$

$$\varepsilon_i^{(2)} := \begin{cases} PD(w_2) + \lambda_2 + \varepsilon_{i-n_2} & \text{if } n_2 \leq i, \\ 0 & \text{otherwise}; \end{cases}$$

and

$$\varepsilon_i^{(1,2)} := \begin{cases} PD(v) & \text{if } i = n_i + n_2, \\ 0 & \text{otherwise}. \end{cases}$$

The following proposition describes how we can easily compute each (ε_i, S_i) value for v from the sequences (ε_j, S_j) and (ε'_k, S'_k) associated to the children (w_1, w_2) of v. The final maximal EP solution is then the one at once we reach the root of the tree.

Proposition 1. For $i = 0, 1, \ldots, n(v)$ we have

$$\varepsilon_i = \max\{\varepsilon_i^{(0)}, \varepsilon_i^{(1)}, \varepsilon_i^{(2)}, \varepsilon_i^{(1,2)}\}.$$

Moreover, if we let L_1, L_2 be the set of leaves of the subtrees rooted at w_1, w_2, respectively, then S_i can be taken to be $S_j \cup S'_k$ (when $\varepsilon_i^{(0)}$ is maximal, and (j, k) provide a maximal pair in the computation of $\varepsilon_i^{(0)}$), or $L_1 \cup S'_{i-n_1}$ (when $\varepsilon_i^{(1)}$ is maximal) or $L_2 \cup S_{i-n_2}$ (when $\varepsilon_i^{(2)}$ is maximal) or $L_1 \cup L_2$ (when $\varepsilon_i^{(1,2)}$ is maximal).

As a simple example of the application of this algorithm, consider the binary tree on six taxa with branch lengths as shown in Fig. 1. The maximal EP values for subsets of species of size k and the set that realises this maximum, for $k = 1, 2, \ldots, 6$ is shown in Table 3. Notice that the sets S_k are not nested in this example.
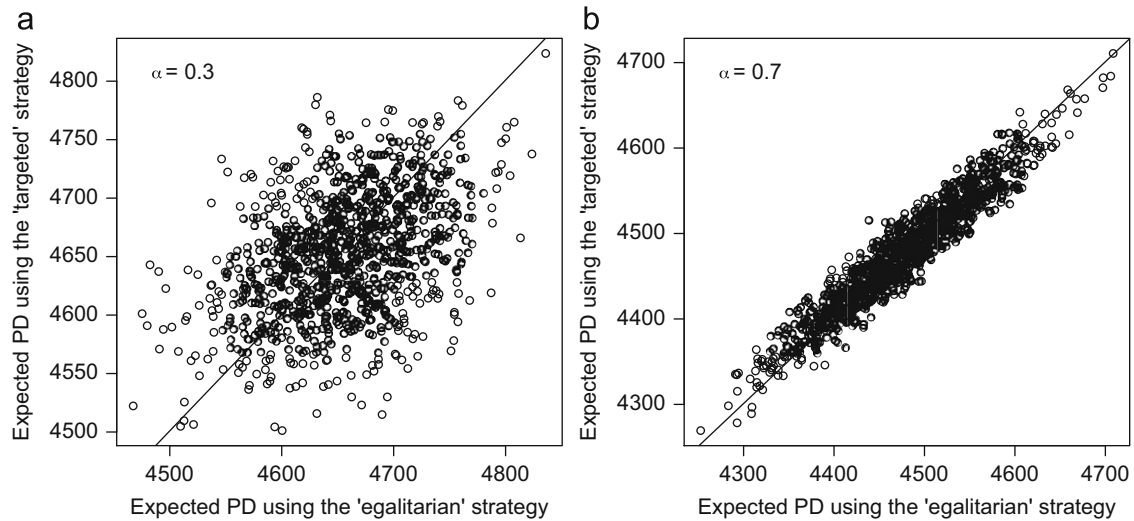
Fig. 5. The expected future PD calculated under both strategies for trees simulated under a Yule model ($n = 1000$), where (a) $\alpha = 0.3$ and (b) $\alpha = 0.7$. The lines show when the strategies are equivalent.

Table 3
The maximal EP values (ε_k) and species set (S_k) for each value of k.

k	ε_k	S_k
1	5.0	d
2	9.0	d, c
3	12.5	d, e, f
4	16.5	c, d, e, f
5	18.0	a, b, c, d, e
6	22.5	a, b, c, d, e, f

6.1. Application

We now apply our algorithm to the lemur tree. Which 10 species, if they were to become extinct, would lead to the greatest loss in phylogenetic diversity compared to any other combination of species? The algorithm identified the following 10 species, with their current IUCN status in brackets: *Allocebus trichotis* (*DD-EN*), *Cheirogaleus major* (*LC*), *C. medius* (*LC*), *C. crossleyi* (*DD-EN*), *Phaner furcifer* (*LC*), *Lepilemur mustelinus* (*DD-EN*), *Varecia variegata* (*CR*), *V. rubra* (EN), *Indri indri* (*EN*), *Daubentonia madagascariensis* (*NT*), as shown in Fig. 6. Losing these 10 species would result in a loss of 32.8% of the total PD (1.96 units of the total 5.97 units), versus an average of 8.9 ± 3.9% when 10 species are lost at random. Most of these species would also rank highest under simple measures of evolutionary distinctiveness (Redding et al., 2008); under the fair proportion measure of distinctiveness, eight of these species are in the top 10, while under the Equal Splits measure, nine of the 10 species listed here are in the top 10. So, not only do evolutionarily distinctive species capture a lot of the total PD in a tree (Redding et al., 2008), their loss would seem to lead to a greater than average loss of PD from that tree. This is because the distinctiveness measures are heavily weighted by the pendant edges (Redding et al., 2008), as is maximum EP.

We also compared the amount of PD lost when we lose those species that are most likely to become extinct (i.e. the seven Critically Endangered species) with the maximum amount of PD that can be lost for the same number of extinctions. We found that it is possible to lose more PD with randomly sampled species because the Critically Endangered species are not on especially
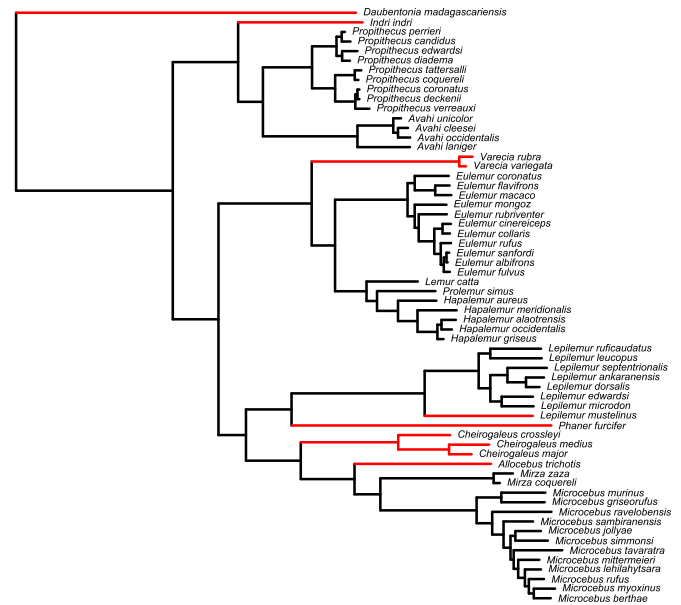

Fig. 6. The 10 lemur species whose extinction would result in the greatest loss of phylogenetic diversity. The amount of PD that would be lost is shown in red (32.8%).

long terminal branches (Fig. 7). Indeed, there is only one species of lemur (*Varecia variegata*) that is both Critically Endangered and belongs to the maximal EP solution set.

7. Concluding comments

We examined expected future PD in both simulated and real trees under two different conservation strategies (egalitarian versus targeted). Despite the fact that total future species richness was the same under both strategies, we found that either strategy could outperform the other (that is, increase the expected future PD) depending on how endangered species are distributed throughout the phylogeny. When endangered species have long pendant edges and are clustered among the tips, a targeted
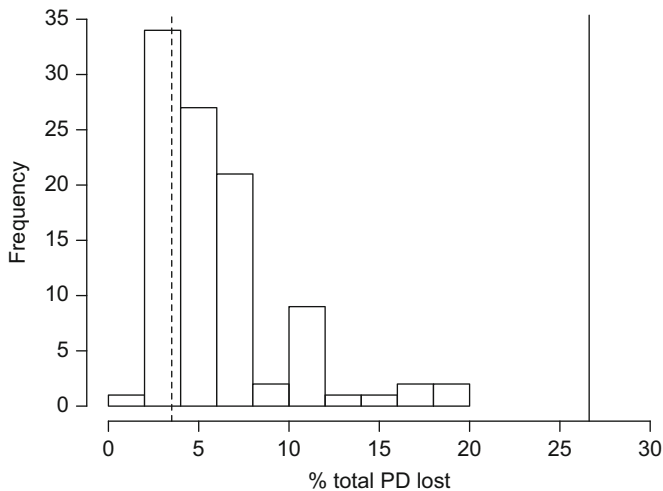
Fig. 7. The distribution of PD loss with the extinction of seven randomly sampled species ($n = 100$). The dashed and solid lines indicate the PD loss with the extinction of the seven critically endangered species (3.5%) and the maximal EP solution set for $k=7$ (26.6%), respectively.

strategy is more effective at preserving future PD; an egalitarian strategy would be preferred when the reverse is true.

In the case of the lemurs of Madagascar, we found the two strategies performed virtually identically. However, when we randomly shuffled the species' conservation statuses among the tips, we found either strategy can outperform the other. We also found that the mean expected future PD of these randomizations was lower than the true expected future PD, under both strategies. This implies that the most endangered species are distributed relatively evenly throughout the tree and/or are not on long terminal branches. Since the pendent edge lengths of the Critically Endangered species are not significantly shorter than the pendent edges of all species (Welch two sample t-test: $t=-0.9638$, degrees of freedom = 8.157, p-value = 0.363), we conclude that it is the even distribution of endangered species throughout the tree that reduces the amount of difference between the two strategies presented here. Indeed we found there was no correlation between species' evolutionary distinctiveness (equal splits measure; Redding and Mooers, 2006) and their survival probability ($r=-0.0719$, $p=0.579$). Further analyses that explicitly consider phylogenetic clustering of survival probability (i.e. 'heritable' survival probabilities) might be illuminating. We expect the conservation strategy chosen in the case of lemurs would focus on other aspects such as economic value or planning feasibility since the difference in PD effects is minor. The algorithm we present here provides a new method by which we may identify the group of species that, if lost, would lead to the greatest loss of PD than the loss of any other group containing the same number of species within a given phylogeny.

In both our simulations and lemur phylogeny analyses, we noted that the average differences between strategies are relatively small as compared with the large differences that were generated under extreme examples (ex. early or late radiations). This may be explained by two factors. First, the number of species that were conserved under each model is constrained to be the same. This decision was made to isolate the effects of allocation strategy on expected PD independent of the number of species conserved. The second was that the degree of extinction risk was spread evenly throughout both the empirical lemur phylogeny and the simulated phylogenies. Since our phylogenies are relatively balanced and the number of pendant edges, which constitute a large proportion of PD, stay the same under both

models, we would expect both strategies to generate similar values. In other clades where there is a correlation between distinctiveness and extinction risk (see e.g. Magnuson-Ford et al., 2009), we might expect a larger difference between strategies.

We note that, while we used α and β to represent 'conservation effort' in these two conservation strategies, the link between increased effort and a corresponding increase in survival probability is less clear in real world situations. If we regard α and β as values directly corresponding to a dollar amount, we make two assumptions. First, increasing the amount of conservation spending will increase a species' probability of survival. While this is widely held to be true, empirical tests of this relationship are only beginning (see Ferraro and Pattanayak, 2006 for discussion). Second, we assume that for a given investment, the survival probability of an endangered species will increase more than that of a less endangered species (for example, when $\alpha=0.5$, an Endangered species' survival probability is increased 24%, from 0.333 to 0.577, whereas for a Near Threatened species, this change is 0.5%, from 0.990 to 0.995). This assumption, however, is more problematic or even backward, since it may be that the cost of conservation increases with the degree of imperilment (see, e.g. Mandel et al., 2010)—for instance, if the cost of habitat preservation increases with its rarity. Thus we would recommend a broad interpretation of α and β as an integrative measure of conservation effort including but not limited to factors such as financial support, public awareness, research and implementation of conservation action plans.

While lemurs provide an excellent example of how the mathematical models and algorithms presented here may be applied to real biological systems, we acknowledge that our model only considers the preservation of future phylogenetic diversity under one metric (PD) and does not include other factors such as cultural significance or ecological diversity. Thus our findings should not be considered as an exclusive recommendation to lemur conservation planning, rather, a starting point for further analyses that explore in more detail different ways we may allocate conservation effort and the effect this has on factors such as PD. We have shown that our methods may be easily extended to consider a variety of more complex factors such as an extinction risk which changes over time or multiple categories of risk within the species set E.

Our methods may be applied at a broader level by incorporating them into existing conservation planning tools. For example, Kremen et al. (2008) have recently identified areas of high conservation priority in Madagascar based on a new spatial conservation algorithm (Zonation; http://www.helsinki.fi/bioscience/consplan/software/Zonation/index.html). This algorithm makes use of species' past and present geographic ranges, abundances, habitat suitability and land cost to identify which locations within current and/or proposed conservation areas, should be of highest conservation priority. Presently, this software can incorporate species 'fractional extinction risk' which is calculated using a species' past change in range size. Another logical extension may be to weight the value of a location by the amount of PD that is conserved (see also Rosauer et al., 2009 for a similar approach). The algorithm we present here could then be used to assign higher value to locations that include those species that, if lost, would lead to the greatest reduction of PD.

Appendix A. Details regarding the lemur phylogeny

A.1. Taxonomy, genetic data, alignment

The entire genetic dataset was partitioned into five genes using Mesquite OSX version 1.1. Species that did not have data for a particular gene(s) were deleted from that gene partition. We used an online model testing site, ModelGenerator and MultiPhyl Online v1.0.6 (http://distributed.cs.nuim.ie/multiphyl.php; Keane et al., 2007) to determine the best model to use for tree inference. We first compared the decisions made by this site with those from MrModeltest2 version 2.2 (MACOSX) for the gene 12S. Both methods gave the same result for the best model to choose (and similar ranking of subsequent models), and so the online model testing programme was used for the other four genes. The accession numbers can be found in the online supplementary material.

The Akaike Information Criterion chose the GTR+I+G model (general time reversible plus invariant plus gamma-distributed rate variation; see, e.g. (Yang, 1997) for a description) for the total gene sequence, and for 12S, Cytb, and PAST separately. The COII gene was best described by the slightly simpler HKY+I+G model and the D-loop gene data was best described by the TVM+I+G model. Since this model is not available in MrBayes 3.1.2, we substituted the closely related GTR+I+G for the D-loop gene. Therefore, all the gene partitions, with the exception of COII, used the GTR model which corresponds to the MrBayes setting nst=6 and the rates= invgamma (gamma-shaped rate variation) setting. COII used the HKY model corresponding to nst=2, again with the rates= invgamma. The parameters of statefreq (stationary nucleotide frequencies), revmat (substitution rates), shape (shape parameter of the gamma distribution of rate variation), and pinvar (proportion of invariable sites) were all set to be unlinked so that each partition had its own set of parameters. Also to allow for different rates for each partition, the rate parameter was set to variable (ratepr=variable). Four unheated chains were run within MrBayes for 10 million generations with the first 25% discarded as burn-in. The sample frequency for the remainder was set to every 1000th generation.

A.2. IUCN red list details

For those species that were classified as Data Deficient (DD), we assumed a reasonable Red List category based on the information given in each species' Red List assessment and recent scientific literature. Recent changes in taxonomy (genus *Lepilemur*) and newly described species (genus *Microcebus*) require further research to know the extent of occurrence of these species according to their newly defined geographic ranges. Proposed classifications are conservative so that if more than one threatened status is tentatively suggested in the justification section of the IUCN Red List entry (www.iucnredlist.org), the more endangered status is chosen. If it is simply stated that 'the species may warrant listing as threatened in future' ('threatened' categories include VU, EN and CR), then EN was chosen as the intermediate of these categories. This is in line with a recent review of the IUCN Red List process which states that 'the precautionary recommendation is that DD species should be afforded the same degree of protection as threatened species, at least until more information is forthcoming' (Mace et al., 2008). In case our assumptions were incorrect, we also carried out the same analyses for two extreme cases, following Purvis and colleagues (Purvis et al., 2000). First we assumed all DD species are classified as Endangered and secondly we assumed they have zero risk of extinction. Table A.1 presents the assignments of IUCN categories for lemur species that are classified as Data Deficient.

Appendix B. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jtbi.2010.06.004.

Table A.1
Assignment of IUCN extinction risk categories to lemur species currently listed as data deficient.

Genus	Species	IUCN category	Justification
Allocebus	*trichotis*	EN	More widely distributed than previously known; however, given known threats, the species may warrant listing as threatened in future (IUCN, 2009)
Avahi	*unicolor*	EN	Given known threats, and assuming a restricted distribution, the species may warrant listing either as Vulnerable or Endangered based on criterion B (IUCN, 2009)
Cheirogaleus	*crossleyi*	EN	Given known threats, the species may warrant listing as threatened in future (IUCN, 2009)
Eulemur	*rufus*	VU	Population trend is decreasing. However, given that threats are no doubt operating within the now more-confined range of E. rufus, with further information this species very likely will require listing as near threatened or vulnerable (IUCN, 2009)
Lepilemur	*dorsalis*	CR	Previously listed as VU on the IUCN red list. It is among the most endangered primates in the world (Mittermeier et al., 2008, p. 22.)
	leucopus	EN	Given known threats and clarity on the distribution range, the species may warrant listing as threatened in future (IUCN, 2009). L. ruficaudatus population trend is decreasing
	microdon	EN	
	mustelinus	EN	
	ruficaudatus	EN	
Microcebus	*jollyae*	LC	Closely related to M. rufus, which is classified as LC. Microcebus spp. are also the most widely distributed and abundant of all the lemurs (Mittermeier et al., 2008)
	lehilahytsara	LC	
	mittermeieri	LC	
	simmonsi	LC	
	myoxinus	EN	Given the likely threat of habitat loss, the species may warrant listing as threatened in future (IUCN, 2009)
Mirza	*zaza*	NT	Very little is known about this species as it has recently been described in 2005, however, it is closely related to M. coquereli (Mittermeier et al., 2008), which is listed as NT

References

Agapow, P., Bininda-Emonds, O., Crandall, K., Gittleman, J., Mace, G., Marshall, J., Purvis, A., 2004. The impact of species concept on biodiversity studies. Q. Rev. Biol. 79, 161–179.

Barrett, M.A., Ratsimbazafy, J., 2009. Luxury bushmeat trade threatens lemur conservation. Nature 461 (7263), 470.

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

Fabre, P.-H., Rodrigues, A., Douzery, E.J.P., 2009. Patterns of macroevolution among primates inferred from a supermatrix of mitochondrial and nuclear DNA. Mol. Phylogen. Evol. 53 (3), 808–825.

Faith, D., 1992. Conservation evaluation and phylogenetic diversity. Biol. Conserv. 61 (1), 1–10.

Faller, B., Pardi, F., Steel, M., 2008. Distribution of phylogenetic diversity under random extinction. J. Theor. Biol. 251 (2), 286–296.

Ferraro, P.J., Pattanayak, S.K., 2006. Money for nothing? A call for empirical evaluation of biodiversity conservation investments. PLoS Biol. 4 (4), e105.

Hartmann, K., Steel, M., 2007. Phylogenetic diversity: from combinatorics to ecology. In: Gascuel, O., Steel, M. (Eds.), Reconstructing Evolution: New Mathematical and Computational Approaches. Oxford University Press, pp. 171–196.

Heard, S., Mooers, A., 2000. Phylogenetically patterned speciation rates and extinction risks change the loss of evolutionary history during extinctions. Proc. R. Soc. Lond. B Biol. 267 (1443), 613–620.

IUCN, 2001. Red List Categories and Criteria. Version 3.1. IUCN Species Survival Commission, Gland, Switzerland and Cambridge, UK, ii+30.

IUCN, 2009. IUCN Red List of Threatened Species. Version 2009.2 ⟨http//www.iucnredlist.org⟩.

Keane, T., Naughton, T., McInerney, J., 2007. Multiphyl: a high throughput phylogenomics webserver using distributed computing. Nucleic Acids Res. 35, W33–W37.

Kremen, C., Cameron, A., Moilanen, A., Phillips, S.J., Thomas, C.D., Beentje, H., Dransfield, J., Fisher, B.L., Glaw, F., Good, T.C., Harper, G.J., Hijmans, R.J., Lees, D.C., Louis, E., Nussbaum, R.A., Raxworthy, C.J., Razafimpahanana, A., Schatz, G.E., Vences, M., Vieites, D.R., Wright, P.C., Zjhra, M.L., 2008. Aligning conservation priorities across taxa in madagascar with high-resolution planning tools. Science 320 (5873), 222–226.

Lewis, L., Lewis, P., 2005. Unearthing the molecular phylodiversity of desert soil green algae (chlorophyta). Syst. Biol. 54 (6), 936–947.

Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Akçakaya, H.R., Leader-Williams, N., Milner-Gulland, E., Stuart, S.N., 2008. Quantification of extinction risk: Iucn's system for classifying threatened species. Conserv. Biol. 22 (6), 1424–1442.

Magnuson-Ford, K., Ingram, T., Redding, D., Mooers, A., 2009. Rockfish (sebastes) that are evolutionarily isolated are also large, morphologically distinctive and vulnerable to overfishing. Biol. Conserv. 142, 1787–1796.

Mandel, J.T., Donlan, C.J., Armstrong, J., 2010. A derivative approach to endangered species conservation. Frontiers Ecol. Environ. 8 (1), 44–49.

Marris, E., 2007. What to let go. Nature 450 (8), 152–155.

Mittermeier, R., Konstant, W., Hawkins, F., EE Louis, J., Langrand, O., Ratsimbafazy, J., Rasoloarison, R., Ganzhorn, J.U., Rajaobelina, S., Tattersall, I., Meyers, D., 2006. Lemurs of Madagascar. Conservation International, Washington, DC.

Mittermeier, R.A., Ganzhorn, J.U., Konstant, W.R., Glander, K., Tattersall, I., Groves, C.P., Rylands, A.B., Hapke, A., Ratsimbazafy, J., Mayor, M.I., Louis, E.E., Rumpler, Y., Schwitzer, C., Rasoloarison, R.M., 2008. Lemur diversity in madagascar. Int. J. Primatol. 29 (6), 1607–1656.

Mooers, A.O., Faith, D.P., Maddison, W.P., 2008. Converting endangered species categories to probabilities of extinction for phylogenetic conservation prioritization. Plos One 3 (11), e3700.

Nee, S., May, R., 1997. Extinction and the loss of evolutionary history. Science 278 (5338), 692–694.

Pardi, F., Goldman, N., 2005. Species choice for comparative genomics: no need for cooperation. PLoS Gen. (1), e71.

Phillimore, A., Price, T., 2008. Density-dependent cladogenesis in birds. PLoS Biol. 6 (3), e71.

Purvis, A., Agapow, P., Gittleman, J., Mace, G., 2000. Nonrandom extinction and the loss of evolutionary history. Science 288 (5464), 328–330.

Purvis, A., Hector, A., 2000. Getting the measure of biodiversity. Nature 405 (6783), 212–219.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Raup, D., 1993. Extinction: Good Genes or Bad Luck. Oxford University Press, Oxford, UK.

Redding, D.W., Hartmann, K., Mirnoto, A., Bokal, D., DeVos, M., Mooers, A.O., 2008. Evolutionarily distinct species often capture more phylogenetic diversity than expected. J. Theor. Biol. 251 (4), 606–615.

Redding, D.W., Mooers, A.O., 2006. Incorporating evolutionary measures into conservation prioritization. Conserv. Biol. 20 (6), 1670–1678.

Rosauer, D., Laffan, S., Crisp, M., Donnellan, S., Cook, L., 2009. Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. Mol. Ecol. 18 (19), 4061–4072.

Spillner, A., Nguyen, B.T., Moulton, V., 2008. Computing phylogenetic diversity for split systems. IEEE ACM Trans. Comput. Biol. 5 (2), 235–244.

Steel, M., 2005. Phylogenetic diversity and the greedy algorithm. Syst. Biol. 54 (1), 527–529.

Tattersall, I., 2007. Madagascar's lemurs: cryptic diversity or taxonomic inflation? Evol. Anthropol. 16 12–23.

Vamosi, J.C., Wilson, J.R.U., 2008. Nonrandom extinction leads to elevated loss of angiosperm evolutionary history. Ecol. Lett. 11 (10), 1047–1053.

Vane-Wright, V.I., Humphries, C.J., Williams, P.H., 1991. What to protect? Systematics and the agony of choice. Biol. Conserv. 55, 235–254.

von Euler, F., 2001. Selective extinction and rapid loss of evolutionary history in the bird fauna. Proc. R. Soc. Lond. B 268, 127–507.

Weitzman, M., 1998. The Noah's Ark problem. Econometrica 66, 1279–1298.

Yang, Z., 1997. Paml: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.