Biological Conservation 144 (2011) 1902–1909

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Endemism and evolutionary history in conflict over Madagascar's freshwater conservation priorities

B. Isambert [a,b,*], J. Bergsten [c], M.T. Monaghan [d], H. Andriamizehy [e], T. Ranarilalatiana [e], M. Ratsimbazafy [e], J.R. Andriniainimanana [e], A.P. Vogler [a,b]

[a] Department of Entomology, Darwin Centre 2, Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom
[b] Department of Biology, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, United Kingdom
[c] Department of Entomology, Swedish Museum of Natural History, Box 50007, SE 10405 Stockholm, Sweden
[d] Leibniz Institute of Freshwater Ecology and Inland Fisheries , Müggelseedamm 301, 12587 Berlin, Germany
[e] Département d'Entomologie, Faculté des Sciences, BP 906, Université d'Antananarivo, Antananarivo 101, Madagascar

ARTICLE INFO

Article history:
Received 17 December 2010
Received in revised form 1 April 2011
Accepted 6 April 2011
Available online 30 April 2011

Keywords:
Phylogenetic diversity
Species richness
Protected areas
Coleoptera
Species delimitation
GMYC model

ABSTRACT

Regional-scale biodiversity indicators provide important criteria for the selection of protected areas in conservation, but their application is often hindered by a lack of taxonomic knowledge. Moreover, different indicators include different types of information, sometimes leading to divergent conservation priorities. Madagascar tops the world list of biodiversity hotspots and much conservation effort has been directed toward its threatened plants and vertebrates. In contrast, its highly diverse freshwater invertebrate fauna has received comparatively little conservation attention. We conducted an inventory of Malagasy adephagan water beetles (Coleoptera, Dytiscidae, Noteridae, Gyrinidae, Haliplidae) using a combined morphological and molecular approach. In total, 2043 beetles from 153 sites were sequenced for cytochrome oxidase subunit I (*cox1*), and species delimitation was carried out using the coalescent-based GMYC model. Phylogenetic relationships of the resulting entities were established using *cox1* combined with partial 16S rRNA and 28S rRNA sequences. Ten national parks were assessed for their species richness, phylogenetic diversity (PD) and endemism. We were particularly interested in the contribution of endemic species to PD. Congruence between molecular and taxonomic identifications was high (91%), with 69% of sampled species endemic to Madagascar. Interestingly, we found that PD at a site was negatively correlated to the proportion of endemic species, most likely because endemics are the result of recent radiations with relatively little branch-length contribution to the measure of PD. This suggests that ranking sites for conservation priority based solely on PD potentially disfavor endemic species by underrating areas where the evolutionary process is most active.

1. Introduction

Madagascar's outstanding levels of endemism and imbalance across taxa are unmatched by any other biodiversity hotspot (Ganzhorn et al., 2001): according to Conservation International, the island is home to 11,600 endemic plant species (89.2% of the total present), 144 endemic mammal species (92.9%), 367 endemic reptile species (95.6%) and 229 endemic amphibian species (99.6%). At the same time, the rate at which habitat and biodiversity disappears on the island is alarming (Allnutt et al., 2008) and the combination makes Madagascar the most threatened and important reservoir of diversity on earth (Ganzhorn et al., 2001; Mittermeier et al., 2005). In 2003, Madagascar counted 21 national parks, five *"réserves naturelles intégrales"* and 21 special reserves, covering more than 1.7 millions ha. However, the same year during the World Parks Congress, the former government announced a target to triple this network of protected areas to 10% of the island by 2012 (Gouvernement Malgache, 2004). Several areas have been under study since then and some were added to the protected network in 2006, raising the critical issue of which biodiversity indicators can best be used to align various conservation priorities.

Species richness is a commonly applied indicator of biodiversity (Gotelli and Colwell, 2001), although it does not take explicitely the genetic and evolutionary dimension of diversity into account. The instrumental value of biodiversity may be better accommodated by accounting for its genetic and taxonomic diversity (Faith, 1994; Moritz, 2002; Polasky et al., 2001), prioritising distantly over closely related taxa (Faith, 1994). Phylogenetic diversity (PD) (Faith, 1992) has been suggested as an appropriate biodiversity

* Corresponding author. Tel.: +33 (0)670142010; fax: +44 (0)2079425229.
E-mail addresses: benjamin.isambert06@imperial.ac.uk (B. Isambert), Johannes.Bergsten@nrm.se (J. Bergsten), Monaghan@igb-berlin.de (M.T. Monaghan), A.Vogler@nhm.ac.uk (A.P. Vogler).

0006-3207/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biocon.2011.04.016

indicator for this purpose (Barker, 2002). Phylogenetic diversity is defined as the sum of the evolutionary pathways connecting a set of taxa (Faith, 1992), and describes 'feature diversity' (Faith, 2002). The concept has been widely adopted and the measure of PD used in conservation practice to identify key regions within hotspots (Forest et al., 2007; Prado et al., 2010).

The general idea behind PD is to give relatively higher priority to those lineages that represent a longer evolutionary time (Krajewski, 1991). Under comparable rates of evolution, longer time translates to more divergent "evolutionary products" and hence maximizing PD would maximize the feature diversity of the protected area that in turn can be seen as a best investment for an unpredictable future (Witting and Loeschcke, 1995). An important consideration is that recent adaptive radiations yield low PD, yet they can be seen as the active motors of speciation as opposed to long branched but species generating evolutionary dead ends (Nee and May, 1997). Consequently, prioritising high PD values may exclude active centres of speciation from the future protected network (Erwin, 1991; Spathelf and Waite, 2007). Utilizing PD values in biodiversity hotspots must therefore be an informed decision, potentially prioritising ancient versus recent radiations (Nee and May, 1997). Thus neither endemism nor PD might be optimal as biodiversity indicators for conservation priority settings. Davis et al. (2008) considered this enigma and identified areas of neoendemism in biodiversity hotspots, by weighting the endemic status of species by their spatial range and their genetic divergence from sister taxa. Similarly, Rosauer et al. (2009) developed a new indicator measuring the spatial restriction of PD, named *phylogenetic endemism*, so as to map endemism of lineages and identify regions where PD is restricted.

In Madagascar, two recent studies suggested the emergence of microendemic species through recent adaptive radiations. Wilmé et al. (2006) investigated the effects of paleoclimatic shifts and river drainage systems on vertebrate dispersal patterns and vicariance within the island based on a comprehensive database of species distribution. In short, Wilmé et al. (2006) outlined a speciation scenario of retracting and expanding forests resulting in isolated patches along altitudinal riversystems during Quaternary climate cycles. As an alternative, several authors underlined the potential role of environmental gradients and the heterogeneous climate of Madagascar in generating local endemism (Dewar and Richard, 2007; Harcourt, 2008). In particular Raxworthy et al. (2007) discussed the role of climatic gradients and the general idea of climate patterns engendering local endemism was referred to as "the current climate hypothesis" by Pearson and Raxworthy (2009). Our intentions are not to test these hypotheses but rather to test the implications on utilizing PD in biodiversity hotspots where recent adaptive radiations and paleoendemics coexist (Yoder and Nowak, 2006).

Adephagan water beetles are predators that occupy most types of water bodies. Larval and adult stages are both fully aquatic. It is a species-rich group, quite easily sampled and ecologically well understood, and it has therefore been used as an indicator group for freshwater biodiversity assessment (Sanchez-Fernandez et al., 2006). Freshwater biodiversity was not considered a conservation priority in Madagascar until recently, although it was shown to be species rich and highly microendemic, with many freshwater species remaining restricted to one or few river basins (Benstead et al., 2003). However this biota is still not given the same attention as flagship vertebrates, while they are equally threatened (Green and Sussman, 1990). We take advantage of the development of high throughput DNA sequencing and operational species delimitation methods, but also evaluate morphological and distributional knowledge to conduct an integrated inventory of the adephagan water beetle fauna of Madagascar. The resulting catalogue of species and regional species pool phylogeny are analyzed to rank ten existing protected areas according to different biodiversity indicators. We also test the hypothesis that if endemic species result from recent radiations their proportional representation would negatively influence measures of PD.

2. Materials and methods

2.1. Sampling

Adephagan water beetles (Coleoptera: Dytiscidae, Noteridae, Haliplidae, Gyrinidae) were collected between 2004 and 2007, following a sampling strategy designed to maximize latitudinal and longitudinal gradients, and habitat and microhabitat diversity of the island. To do so, 153 sampling sites located both inside and outside of 10 protected areas (National Parks and Special reserves) (Table 1) were chosen within each of the seven WWF ecoregions encountered in Madagascar: humid/lowland forests, subhumid forests, mangroves, succulent woodlands, dry deciduous forests, ericoid thickets, and spiny thickets (Fig. 1).

The collecting was conducted with hand nets and sieves, depending on the size of the water bodies and the type of microhabitat (e.g. streams, pools, rivers, water falls, hygropetric rocks). In a few localities 1.5 L bottletraps left overnight were used in addition to active sampling. GPS coordinates were recorded for each sampling locality. The insects were stored in 95% ethanol immediately after collection and returned to the laboratory, where they were identified to species, or, if not possible, sorted to morphospecies. As a sample of the intraspecific genetic variation, up to five individuals per Linnean species/morphospecies per locality were prepared for the molecular analysis. Specimens are deposited in the research collection of the Department of Entomology, Natural History Museum, London.

2.2. DNA extractions and sequencing

In order to allow subsequent taxonomic identification, the DNA was extracted from the insects using non-destructive methods: depending on the size of the beetles, a leg, the whole individual, or the head together with the prothorax were placed overnight in a lysis buffer containing a solution of proteinase K. The WIZARD SV 96 Genomic Kit (Promega) was used, allowing processing 96 individuals at a time. The mitochondrial DNA was amplified for the 3′ region of the Cytochrome Oxidase subunit I (*cox1*) by PCR, using the set of primers Ron Inosine/Ron Dyt and Patty/Pat Dyt (Ron Inosine: GGIGCICCIGATATAGCNTTYC, Ron Dyt: GGAGCACCTGAT ATAGCNTTYC, Patty: GCTTAAATTCATTTGCACTAATCTGC, Pat Dyt: TCATT GCACTAATCTG CCATATTAG). The resulting fragments are 1300 bp. The same PCR master mix was used for every reaction (quantities given per sample): 12 μl of 10% Trehalose, 4.65 μl of DNA grade water, 2.5 μl of Reaction Buffer (10× Bioline), 1.25 μl of 50 mM $MgCl_2$, 0.5 μl of Ron Dyt (or Ron Inosine), 0.5 μl of Pat Dyt (or Patty), 0.5 μl of dNTPs (40 mM), 0.1 μl of *Taq* DNA polymerase (Bioline) and 3 μl of the DNA eluate. The following cycling conditions were used, for both sets of primers: 2 min at 94 °C; 30 s at 94 °C, 1 min at 53 °C, and 2 min and 15 s at 70 °C (repeated over 35 cycles); 10 min at 70 °C. The PCR products were purified and sequenced at the DNA Sequencing Facility of the Natural History Museum London. The primers Jerry (CAACATTTATTTTGATTTTTTGG) and Pat Dyt were used for the sequencing reactions, targeting a fragment of 800 bp within the Ron Dyt–Pat Dyt (or Ron Inosine–Patty) fragment. In order to reconstruct a three gene species phylogeny, two additional markers were targeted: partial 16S rRNA with primers 16Sar and 16Sb2 (Cognato and Vogler, 2001) and 28S rRNA with 28SDD and 28SFF (Monaghan et al., 2007). Within each GMYC cluster, only individuals for which a minimum of two

Table 1
Existing national parks/special reserves considered as potential protected areas for the hydradephagan beetle fauna of Madagascar.

	Area (ha)	Ecor[a]	Elevation (m)	GPS coord[b]	Conservation interest[c]
Andasibe	15,480	LLF	900–1250	18°28′S–48°28′E	14 spp of lemurs, 51 spp of reptiles, 84 spp of amphibians, >100 spp of orchids
Andringitra	31,160	SHF-ET	650–2658	22°07′S-46°47′E	54 spp of mammals, 50 spp of reptiles, 78 spp of amphibians, 1000 spp of plant
Ankarana	18,225	DDF-ET	50–300	12°50′S–49°01′E	11 lemur spp, 50% of Madagascar's bat spp, 60 spp of amphibians and reptiles, 330 plant spp
Isalo	81,540	SHF	514–1268	22°22′S–45°11′E	15 spp of lemurs, 77 spp of birds, >400 plant spp
Marojejy	60,050	ET	60–2132	14°26′S–49°15′E	11 spp of lemurs, 148 spp of amphibians and reptiles, 16 endemic to the park, 33% of all Madagascar's reptiles and amphibians
Masoala	240,520	LHF	0–1224	15°18′S–50°03′E	50% of Madagascar's plant spp, >50% of Madagascar's reptile, amphibian and mammal spp
Montagne d'Ambre	23,010	DDF-SHF	850–1475	12°31′S–49°03′E	1020 spp of plants, 7 spp of lemurs and 24 spp of amphibians
Ranomafana	41,601	SHF	400–1417	21°13′S–47°27′E	98 spp of amphibians, 62 spp of reptiles, 115 spp of birds
Zahamena	64,370	LLF	254–1560	17°30′S–48°41′E	48 spp of mammals, 13 spp of lemurs, 62 spp of amphibians, 46 spp of reptiles, >700 plant spp, average of 1450 trees/ha with a 20 m high canopy
Zombitse	36,308	SW	300–825	22°45′S–44°45′E	47% of Madagascar's endemic bird spp

[a] Ecoregions encountered within the boundaries of the national park (LLF: lowland forest, SHF: subhumid forest, ET: ericoid thickets, DDF: dry deciduous forest, SW: succulent woodland).
[b] GPS coordinates.
[c] Specificities of the area that justified its protected status.

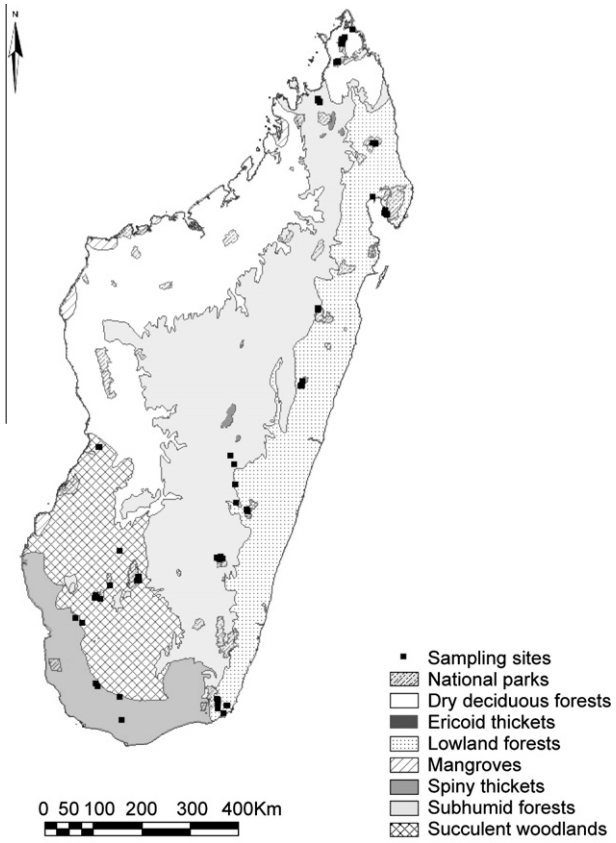

Fig. 1. Distribution of the sampling localities in Madagascar and the seven main ecoregions.

markers were successfully amplified and sequenced were kept for the three gene phylogenetic reconstruction.

Contigs were assembled from the forward and reverse sequences, when both directions were available. Single direction sequences and contigs were edited in Sequencher 4.6 (Gene Codes Corp, Ann Arbor, USA). A total of 2043 sequences (Genbank accessions HQ381640 – HQ383682) were aligned in MAFFT version 6 (http://align.bmr.kyushu-u.ac.jp/mafft/online/server/) following the FFT-NS-1 strategy, recommended for large datasets (i.e. >2000 sequences), with default gap penalties (gap opening penalty = 1.53, offset value = 0). The resulting gap free alignment was trimmed to 740 bp.

2.3. Species delimitation and the GMYC model

A parsimony search was conducted in TNT (Goloboff et al., 2008) on the *cox1* dataset, using the reduced matrix of 1006 unique haplotypes. From this topology, branch lengths were estimated under maximum likelihood using a GTR model and a uniform clock in PAUP*4.0b10 (Swofford, 2002). The General Mixed Yule-Coalescent (GMYC) approach (Pons et al., 2006) was used to conduct the species delimitation using the *splits* package available for R (package at http://r-forge.r-project.org/projects/splits). The number of lineages as a function of time was plotted in order to visualize the transition in branching rate between inter- and intraspecific events. The single threshold option was used on the tree to find the maximum likelihood solution of the GMYC model.

2.4. Multiple alignment and phylogenetic analysis

One representative individual per GMYC group was selected to reconstruct a phylogeny of the regional species pool based on all three markers which was later used to calculate PD across sites and protected parks. A total of 169 *cox*1 sequences, 157 16S sequences and 114 28S sequences constituted the matrix (Genbank accessions: 16S HQ381369 – HQ381525; 28S HQ381526 – HQ381639). The alignment of *cox*1 was unambiguous but different strategies were tested for the length variable 16S and 28S regions: five alignment setting combinations (gap opening penalty: 5, 6.66, 10, 15, 20 and gap extension penalty: 6.66) in ClustalW (Thompson et al., 1994) and three distinct strategies (E-INS-i, G-INS-I and L-INS-i) in MAFFT (gap opening penalty = 1.53, offset value = 0). Congruence between the resulting alignments and the unambiguously aligned *cox*1 region was tested with the ILD test (Farris et al., 1994). Parsimony searches in PAUP*4.0b10 were used to calculate tree lengths (1000 random addition sequence replicates and gaps treated as 5th state). The alignment strategy that gave the lowest ratio ILD/length of combined analysis tree (Wheeler and Hayashi, 1998) was selected (Supplementary material, Table A.1).

For all trees, the single representative of the family Haliplidae was chosen as outgroup. Nucleotide substitution models were

tested on five partitions of the dataset: first, second and third codon positions for *cox*1, 16S and 28S. The selected models from MrModeltest v2 (Nylander, 2004) according to the hierarchical likelihood ratio test and the AIC values were applied to the partitioned dataset and implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Two parallel MCMC runs, each with one cold and three heated chains (temp = 0.2) were run for 10 million generations and sampled every 1000 generations. The mean loglikelihoods and estimated sample sizes (ESS) for the two chains were compared and the burnin was set at 10%, which gave ESS values above 100 for most of the parameters. Trees were summarized using an "all compatible" consensus in MrBayes, storing both branch lengths and posterior probabilities.

2.5. Biodiversity indicators and congruence

Taxon accumulation curves were constructed to assess the comprehensiveness of the sampling at both the species and genotype levels. Presence/absence of each GMYC species and each haplotype were reported for the 153 sites in two community matrices, and used in EstimateS V.8.0 (Gotelli and Colwell, 2001) to build a species and a haplotype accumulation curve with the Mao Tau (given with 95% confidence). Three different indices were used to estimate the total species and haplotype richness: the Incidence based Coverage Estimator (ICE), the Chao 2, and the Michaelis Menten MMRuns. The final estimations were given as an interval between the lowest and the highest estimation among the three indices. Second, to take the evolutionary dimension into account and assess the proportion of total PD captured in the sampling, we evaluated the accumulation curve of PD, defined as the sum of branch lengths of the minimal subtree connecting all taxa in the subset (Faith, 1992). The complete *cox1* dataset (2043 sequences) was formatted in R version 2.8.1 (http://cran.r-project.org/) to construct a PD accumulation curve. Given the large number of sequences, the large number of tree searches during the following randomization process and the consequent computational limitations, neighbor joining trees were preferred to maximum likelihood or parsimony trees in this specific analysis. Following the resampling procedure described by Zhou et al. (2009), a neighbor joining tree was reconstructed using K2P distances and pairwise deletion, and the number of tips (from 1 to 2042) were randomly sampled 500 times. For each randomization step, PD was calculated and its mean value kept for the corresponding tip number.

Incongruence between molecular delimitations and morphological identification was examined in a reciprocal illumination framework (Hennig, 1966). Members of subfamily Copelatinae could not all be identified to the species level at this stage, thus the GMYC species assigned to the genera *Copelatus* and *Madaglymbus* were removed from our calculations of congruence. The GMYC delimitations were used throughout for the downstream diversity analyses. Based on published records and catalogues we categorized species as endemic or non-endemic to the main island of Madagascar (i.e. species reported from the Comoro, Mascarene, Aldabra or Seychelles islands were considered non-endemic). In cases where the GMYC model delimited multiple clusters from what was identified as a single named species, the endemism categorization was judged on a case by case basis: subdivided clusters of widespread African species like *Cybister senegalensis* were all categorized as non-endemic, while subdivided units of endemic species were treated as endemic.

Having defined endemic and non-endemic taxa, four biodiversity indicators were calculated per sampling site and per protected area: the PD, the GMYC species richness (SR), the percentage of endemic species found in each site/protected area (% Endemism) and the percentage of PD represented by the endemic taxa only (% PD End). Phylogenetic diversity values were calculated this time from the 3-gene phylogeny of the regional species pool using the *picante* library (Kembel et al., 2010). PD was calculated as the sum of the branch lengths of the phylogeny connecting all species in the sampling site/area. The Bayesian tree was made ultrametric by applying the nonparametric rate smoothing algorithm described in Sanderson (1997). All subsequent statistical tests based on these biodiversity indicators were conducted in R.

3. Results

3.1. Species delimitation

We analyzed 2043 individuals from 153 collection sites, which produced a total of 1006 mtDNA haplotypes. The maximum likelihood solution of the GMYC model delimited a total of 169 entities (referred to hereafter as GMYC species) (Fig. 2a), 129 as distinct clusters of haplotypes and 40 as singletons (Supplementary material, Figs. A.1 and A.2). The maximum likelihood of the GMYC model was significantly higher than the likelihood of the null model: 9785.035 against 9567.816, $2\Delta L = 434.4380$, χ^2 test: $p < 0.001^{***}$ (Fig. 2b).

3.2. Congruence, endemism and accumulated diversity

Based on the taxonomic identification, more than 91% of the GMYC species were congruent with Linnaean species: 118 cases

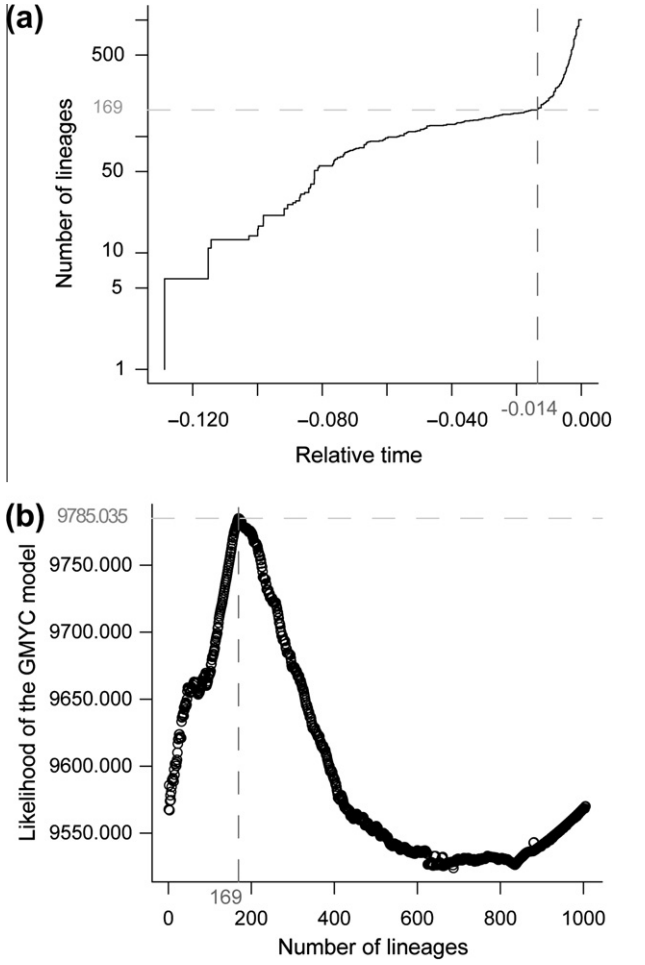

Fig. 2. Delimitation of the species diversity by the GMYC model. (a) Number of lineages through time. The vertical dashed line represents the optimized threshold between speciation events and coalescent events. (b) Likelihood values of the GMYC model obtained as a function of the number of lineages considered.

out of 129 (excluding Copelatinae, see Section 2.5), from which seven were new species and 111 contained only specimens assigned to the same Linnaean name. Eleven GMYC clusters grouped specimens recognized as two different Linnaean species at least, and thus were considered incongruent (Supplementary material, Table A.2). Among these 11 groups, four cases of nested species were identified (*Laccophilus pseustes*/*L. alluaudi*, *Pachynectes costulifer*/*P. sp.n. 4*, *Hydrovatus crassicornis*/*H. madagascariensis*, *Cybister senegalensis*/*C. guignoti*). More than 20 new species were discovered in this integrated inventory (Table A.2), which will be formally described elsewhere. At least ten of these belong to the endemic genus *Madaglymbus* that seems to be the largest endemic radiation of hydradephagan beetles on Madagascar. While the overall level of endemism was conservatively estimated as 69% of the GMYC species (genus *Copelatus* excluded), endemism varied among genera, ranging from 20% (*Hydroglyphus*) to 100% (*Orectogyrus*) (Table A.2). Thirteen out of the 106 Linnaean species identified in our samples were subdivided in several GMYC clusters.

The sample-based rarefaction curves were built from the incidences of the 169 GMYC species and the 1006 mtDNA haplotypes across the 153 sites. The randomized PD accumulation curve was artificially scaled to compare the shape of the curves (Fig. 3). Both the species accumulation curve and the PD accumulation curve approach, but do not reach, an asymptote, whereas the haplotype curve is still curvilinear (Fig. 3). Three different total richness estimators were used: the ICE, the Chao2, and the MMRuns respectively estimated the total richness to be 216, 227 and 210 species, and 4110, 3716 and 8317 haplotypes. Thus the sampling in this study was estimated to represent between 74.4% and 80.5% of the actual fauna at the molecular species level, but covered only 12.1–27.1% of the total haplotype diversity. When comparing the 169 species delimited to the 212 Linnaean species described in Madagascar, we obtain the proportion of 79.7% of the fauna, however only 54% overlap. The PD accumulation curve shows a steeper initial slope compared to the species accumulation curve between 0% and 25% of the number of samples considered, and reaches a total value of 11.298 (before applying the scaling factor).

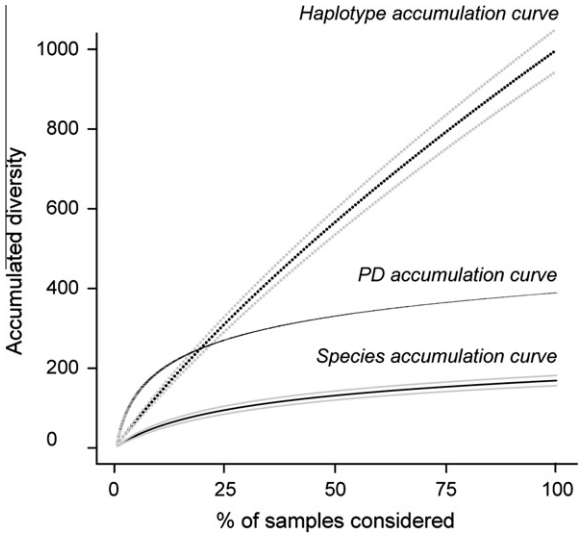

Fig. 3. Randomized accumulation curves for the GMYC species, the haplotypes and for phylogenetic diversity (PD). Species and haplotype accumulation curves were constructed in EstimateS V.8.0, with the 95% confidence interval limits (Mao Tau). The phylogenetic diversity accumulation curve was constructed in R using the *ape* and *CAIC* packages based on the 2043 individuals collected and scaled for comparison.

3.3. Protected areas and biodiversity indicators

PD, SR, % Endemism, and % PD End were calculated for each national park (Fig. 4). PD was highest in Zombitse (8.43), Andringitra (7.81) and Andasibe (7.21) and lowest in Ankarana (3.21). When considering the species richness, Andasibe (45 species) and Zombitse (38 species) still figure among the most diverse parks, together with Isalo (45 species) whereas the percentage of endemism was highest in Zahamena, Montagne d'Ambre, and Marojejy national parks, with 78.6%, 73.3%, and 68.2% of endemic species respectively. Finally, % PD End was greatest in Andasibe (40%), Isalo (34.1%) and Zahamena (38.8%). Out of the three highest ranked areas given by each index, two were common to PD and SR results, as well as to SR and % PD End. However the use of phylogenetic diversity or the percentage of endemic species in the community as biodiversity indicators resulted in two distinct rankings. This discrepancy between PD and % Endemism can have significant implications for conservation priorities. We predicted that endemic elements in the community would decrease the average PD making it an inappropriate indicator for a conservation strategy targeting endemic species. Phylogenetic diversity accumulation curves for endemic and non-endemic taxa showed almost identical shapes (Supplementary material, Fig. A.3). We found that PD and species richness were positively correlated for both endemic and non-endemic groups (Supplementary material, Fig. A.4) (endemic: intercept = 0.454, slope = 0.414, $p < 0.001$; non-endemic: intercept = 0.417, slope = 0.393, $p < 0.001$) but the difference in slopes and intercepts between the two linear models were not significant (Δintercepts = 0.037, $p = 0.639$; Δslopes = 0.021, $p = 0.344$). However, a quadratic regression revealed a significant correlation between PD and % Endemism across the 153 sites: negative for percentages of endemism higher than 46% and positive otherwise (Fig. 5) ($PD = a + b(\arcsin\sqrt{(\%\ \text{Endemism})}) + c(\arcsin\sqrt{(\%\ \text{Endemism})})^2$; $a = 1.2535$, $b = 4.7670$, $c = -3.0335$, $p < 0.001$) (Since the data are not normally distributed, the regression was done using an arcsin-square root transform of the percentage of endemic species).

4. Discussion

Priorities for conservation reserve networks are often set in order to maximize a biodiversity measure, commonly species diversity. Quantifying and defining the boundaries of species therefore remain critical steps (Magurran, 1988). In reality the choice of operational species units is often given little attention in comparison to the choice of the biodiversity indicator itself. Quantifying species until recently required a well worked out taxonomic system of the group and region of interest, which is however missing in many areas of the globe most in need of conservation (Wilson, 2003). The GMYC model overcomes this taxonomic bottleneck and offers operational and quantifiable independently evolving units (Pons et al., 2006). It has recently been used on Malagasy insects (Monaghan et al., 2009) and proved to be a fast and efficient way of delimiting species within poorly described faunas (Pons et al., 2006). However, delimited molecular units as such lack a link to the vast accumulated biological knowledge of species, be it medicinal use, distribution records, phenology or habitat requirements. The morphological identification enabled us to ascribe names to the GMYC units in most cases. This was then used to separate Malagasy endemics from widespread African or Oriental elements.

Lohse (2009) showed that the GMYC method is sensitive to the completeness of the geographic sampling, risking an overestimation of clusters if the proportion of unsampled demes is high. The high congruence between GMYC clusters and Linnaean species

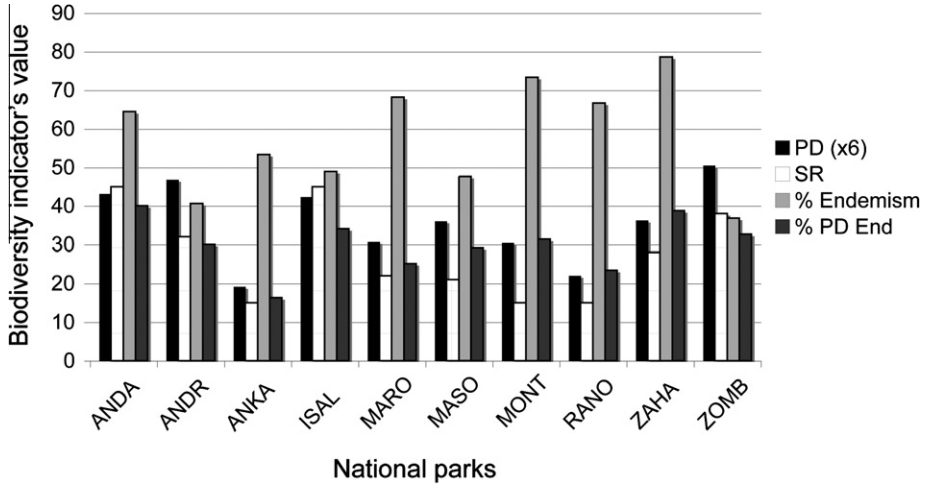

Fig. 4. Scores of the ten existing national parks (ANDA, Andasibe; ANDR, Andringintra; ANKA, Ankarana; ISAL, Isalo; MARO, Marojejy; MASO, Masoala; MONT, Montagne d'Ambre; RANO, Ranomafana; ZAHA, Zahamena; ZOMB, Zombitse) according to the four biodiversity indicators: phylogenetic diversity (PD), GMYC species richness (SR), percentage of endemic species (% Endemism) and proportion of the total phylogenetic diversity represented by the endemic species (% PD End). Values of PD are scaled by a factor of 6.

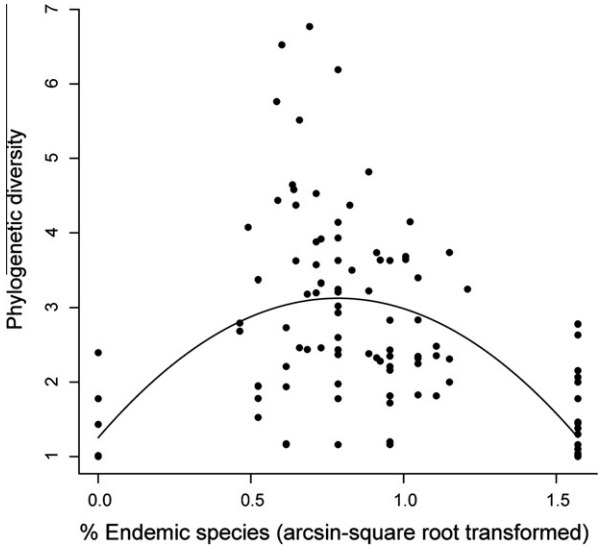

Fig. 5. Correlation between the total phylogenetic diversity and the percentage of endemic species found in each individual sampling site (153 in total). A quadratic regression was fit to the data ($PD = a + b(\arcsin\sqrt{(\%\ Endemism)}) + c(\arcsin\sqrt{(\%\ Endemism)})^2$; $a = 1.2535$, $b = 4.7670$, $c = -3.0335$, $p < 0.001$).

in the present study suggests a low prevalence of this type of sampling related error (also see Papadopoulou et al. (2009)). Nevertheless, a case like the split of the gyrinid *Dineutes proximus* into five, mostly geographically separated, GMYC entities needs to be tested with further sampling of demes. We captured a considerable proportion of the estimated total species richness (up to 80.5%) as well as of the total PD, from more than 2000 individuals and 153 sites. In comparison, Monaghan et al. (2009) found 112 species within Dytiscidae and Hydrophilidae with Chao-estimated richness of 121–166, compared to ca. 210 species in four families here. In addition, the suitability of the GMYC approach could be verified on one hand by the clear peak of the likelihood curve and on the other hand by the level of congruence with the morphological species. More than 90% of the GMYC entities were congruent with the Linnaean species described to date, but considering that several groups are in need of revision (Miller et al., 2009), the congruence will likely increase. The presence of Linnaean species split into more than one GMYC cluster suggests the existence of new species to be evaluated and potentially described. The fact that we found more than 20 new confirmed species out of 169 in total indicates that the hydradephagan beetle diversity is far from completely known. Similarly, the percentage of endemism (69%) calculated from the congruence of the GMYC clusters with endemic Linnaean species is probably below the real level of endemism. This is partly because of the conservative approach we took in linking endemism to the incongruent GMYC species but also because of the remaining endemic species to be discovered. Nevertheless, the figures will likely end up below proportions for other well known groups (Goodman and Benstead, 2005), reflecting the general dispersal capacity of hydradephagan beetles (Ribera and Vogler, 2000).

Despite the considerable scientific effort to understand Madagascar's exceptional biodiversity richness and endemism patterns, not all organisms have been studied to the same degree: very little is known about the invertebrate alpha diversity of the island (see Goodman and Benstead, 2005). Consequently, protected areas in Madagascar are mostly designed for vertebrates and plants. Kremen et al. (2008) proved that multitaxonomic rather than single-taxon approaches were critical for the selection of priorities within the island. By proposing new conservation areas for the 10% target based on expert validated distribution models and rare species records of more than 2000 endemic species, including ants and butterflies, they emphasized the role of invertebrate faunas in the practice of global conservation. Here we evaluated ten of the existing national parks as potential protected areas for aquatic invertebrates and tested four distinct biodiversity indicators. Among the four, many consider PD today as the best measure of priority for biodiversity conservation (Barker, 2002; Forest et al., 2007; Prado et al., 2010). The common practice is to produce DNA sequence data for one representative per predefined species and calculate PD from the resulting phylogeny. This protocol risks omitting the evolutionary history from cryptic, undescribed and lumped/unrevised valid species or species complexes, and the resulting estimations of both taxon richness and PD could mislead the selection of protected areas. A high throughput DNA sequencing approach of whole communities includes all evolutionary history independent of the level of a priori taxonomic knowledge. A second advantage comes with the increased resolution and detail when individuals constitute the terminals instead of species. For example, two sister species, per definition separated at the same time from a common

ancestor, could contribute differently to the total PD by varying in intraspecific genetic variation. This offers the possibility to take into account the genetic variation below the species level often argued to be relevant in conservation biology (Goldstein et al., 2000). Individual-based community phylogenetics is opening new doors not only in conservation biology but also in community ecology where similar measures to PD are used to understand the underlying processes of species coexistence (Webb et al., 2002).

It has been debated whether ancient lineages representing large amounts of evolutionary history should be prioritised over recent radiations that have undergone rapid species diversification recently (Erwin, 1991; Krajewski, 1991; Nee and May, 1997; Spathelf and Waite, 2007). Such prioritising could actually decrease the future potential of species survival, adaptation and speciation, which would run counter to the goal of the conservation priority itself. Our results confirm this possibility as PD is in direct conflict with the percentage of endemic species when the latter reaches considerable levels: the higher the PD, the lower is the proportion of endemic species. This could be seen in Zombitse National Park that had the highest PD but the lowest proportion of endemic species. Our results indirectly align with Wilmé et al.'s (2006) proposition in that large parts of the microendemic patterns could be explained by relatively recent (potentially adaptive) radiations. We conclude that DNA aided inventories are advantageous in order not to omit the contribution made by cryptic species, optimizing both the calculation of PD and the evaluation of endemism in biodiversity hotspots, and thus increasing the efficacy of biodiversity indicators. Finally, we confirm that giving priority to high PD can fail to consider areas rich in recent radiations, and that integrating endemism and PD into one index could contribute to the optimization of evolutionary conservation.

Acknowledgements

This study was part of the HOTSPOTS project funded by the European Commission FP6 EST Marie Curie Actions. JB was supported by Grant 621-2009-3744 from the Swedish Research Council. We thank the Department of Entomology at the University of Antananarivo and the Madagascar Institut pour la Conservation des Ecosystèmes Tropicaux (MICET) for their constant support over the years. We also thank the Madagascar Ministry of the Environment, Water and Forests (MINENV) and the National Association for the Management of Protected Areas (ANGAP) for access to protected areas and for permissions to sample and export.

(Permit numbers: No. 175 MINENVEF/SG/DGEF/DPB/SCBLF; No. 82/06/MINENV.EF/SG/DGEF/DPB/SCBLF/RECH; No. 250/06/MINENV.EF/SG/DGEF/DPB/SCBLF/RECH; No. 354N-EV11/MG04; No. 167N-EA05/MG06; No. 312N-EA12/MG06)

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.biocon.2011.04.016.

References

Allnutt, T.F., Ferrier, S., Manion, G., Powell, G.V.N., Ricketts, T.H., Fisher, B.L., Harper, G.J., Irwin, M.E., Kremen, C., Labat, J.-N., Lees, D.C., Pearce, T.A., Rakotondrainibe, F., 2008. A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. Conservation Letters 1, 173–181.

Barker, G.M., 2002. Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. Biological Journal of the Linnean Society 76, 165–194.

Benstead, J.P., De Rham, P.H., Gattolliat, J.L., Gibon, F.M., Loiselle, P.V., Sartori, M., Sparks, J.S., Stiassny, M.L.J., 2003. Conserving Madagascar's freshwater biodiversity. Bioscience 53, 1101–1111.

Cognato, A.I., Vogler, A.P., 2001. Exploring data interaction and nucleotide alignment in a multiple gene analysis of Ips (Coleoptera : Scolytinae). Systematic Biology 50, 758–780.

Davis, E.B., Koo, M.S., Conroy, C., Patton, J.L., Moritz, C., 2008. The California Hotspots Project: identifying regions of rapid diversification of mammals. Molecular Ecology 17, 120–138.

Dewar, R.E., Richard, A.F., 2007. Evolution in the hypervariable environment of Madagascar. Proceedings of the National Academy of Sciences of the United States of America 104, 13723–13727.

Erwin, T.L., 1991. An evolutionary basis for conservation strategies. Science 253, 750–752.

Faith, D.P., 1992. Systematics and conservation – on predicting the feature diversity of subsets of taxa. Cladistics-the International Journal of the Willi Hennig Society 8, 361–373.

Faith, D.P., 1994. Phylogenetic patterns and the quantification of organismal biodiversity. Philosophical Transactions of the Royal Society of London Series B – Biological Sciences 345, 45–58.

Faith, D.P., 2002. Quantifying biodiversity: a phylogenetic perspective. Conservation Biology 16, 248–252.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics-the International Journal of the Willi Hennig Society 10, 315–319.

Forest, F., Grenyer, R., Rouget, M., Davies, T.J., Cowling, R.M., Faith, D.P., Balmford, A., Manning, J.C., Proches, S., van der Bank, M., Reeves, G., Hedderson, T.A.J., Savolainen, V., 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. Nature 445, 757–760.

Ganzhorn, J.U., Lowry, P.P., Schatz, G.E., Sommer, S., 2001. The biodiversity of Madagascar: one of the world's hottest hotspots on its way out. Oryx 35, 346–348.

Goldstein, P.Z., DeSalle, R., Amato, G., Vogler, A.P., 2000. Conservation genetics at the species boundary. Conservation Biology 14, 120–131.

Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Goodman, S.M., Benstead, J.P., 2005. Updated estimates of biotic diversity and endemism for Madagascar. Oryx 39, 73–77.

Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 4, 379–391.

Gouvernement Malgache, 2004. Journal Officiel de la République Démocratique de Madagascar. Imprimerie Nationale, Antananarivo.

Green, G.M., Sussman, R.W., 1990. Deforestation history of the eastern rain forests of Madagascar from satellite images. Science 248, 212–215.

Harcourt, A.H., 2008. Are lemurs' low basal metabolic rates an adaptation to Madagascar's unpredictable climate? Primates 49, 292–294.

Hennig, W., 1966. Phylogenetic Systematics. University of Illinois Press, Urbana.

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

Krajewski, C., 1991. Phylogeny and diversity. Science 254, 918–919.

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, 222–226.

Lohse, K., 2009. Can mtDNA barcodes be used to delimit species? A response to Pons et al. (2006). Systematic Biology 58, 439–442.

Magurran, A.E., 1988. Ecological Diversity and its Measurement. Princeton University Press, Princeton.

Miller, K.B., Bergsten, J., Whiting, M.F., 2009. Phylogeny and classification of the tribe Hydaticini (Coleoptera: Dytiscidae): partition choice for Bayesian analysis with multiple nuclear and mitochondrial protein-coding genes. Zoologica scripta 38, 591–615.

Mittermeier, R.A., Robles Gil, P., Hoffman, M., Pilgrim, J., Brooks, T., Mittermeier, G.T., Lamoreux, J., da Fonseca, G.A.B., 2005. Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. Chicago University Press, Chicago.

Monaghan, M.T., Inward, D.J.G., Hunt, T., Vogler, A.P., 2007. A molecular phylogenetic analysis of the Scarabaeinae (dung beetles). Molecular Phylogenetics and Evolution 45, 674–692.

Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated species inventory on madagascar using coalescent-based models of species delineation. Systematic Biology 58, 298–311.

Moritz, C., 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology 51, 238–254.

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

Nylander, J.A.A., 2004. MrModeltest v2. Program Distributed by the Author. Uppsala University: Evolutionary Biology Centre.

Papadopoulou, A., Monaghan, M.T., Barraclough, T.G., Vogler, A.P., 2009. Sampling error does not invalidate the Yule-Coalescent model for species delimitation. A response to Lohse. Systematic Biology 58, 442–444.

Pearson, R.G., Raxworthy, C.J., 2009. The evolution of local endemism in Madagascar: watershed versus climatic gradient hypotheses evaluated by null biogeographic models. Evolution 63, 959–967.

Polasky, S., Csuti, B., Vossler, C.A., Meyers, S.M., 2001. A comparison of taxonomic distinctness versus richness as criteria for setting conservation priorities for North American birds. Biological Conservation 97, 99–105.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55, 595–609.

Prado, A., Hawkins, J.A., Yesson, C., Barcenas, R.T., 2010. Multiple diversity measures to identify complementary conservation areas for the Baja California peninsular cacti. Biological Conservation 143, 1510–1520.

Raxworthy, C.J., Ingram, C.M., Rabibisoa, N., Pearson, R.G., 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (Phelsuma) from Madagascar. Systematic Biology 56, 907–923.

Ribera, I., Vogler, A.P., 2000. Habitat type as a determinant of species range sizes: the example of lotic–lentic differences in aquatic Coleoptera. Biological Journal of the Linnean Society 71, 33–52.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rosauer, D., Laffan, S.W., Crisp, M.D., Donnellan, S.C., Cook, L.G., 2009. Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. Molecular Ecology 18, 4061–4072.

Sanchez-Fernandez, D., Abellan, P., Mellado, A., Velasco, J., Millan, A., 2006. Are water beetles good indicators of biodiversity in Mediterranean aquatic ecosystems? The case of the Segura River basin (SE Spain). Biodiversity and Conservation 15, 4507–4520.

Sanderson, M.J., 1997. A nonparametric approach to estimate divergence times in the absence of rate constancy. Molecular Biology and Evolution 14, 1218–1231.

Spathelf, M., Waite, T.A., 2007. Will hotspots conserve extra primate and carnivore evolutionary history? Diversity and Distributions 13, 746–751.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis using Parsimony. Sinauer Associates, Sunderland, MA.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. ClustalW – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

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

Wheeler, W.C., Hayashi, C.Y., 1998. The phylogeny of the extant chelicerate orders. Cladistics-the International Journal of the Willi Hennig Society 14, 173–192.

Wilmé, L., Goodman, S.M., Ganzhorn, J.U., 2006. Biogeographic evolution of Madagascar's microendemic biota. Science 312, 1063–1065.

Wilson, E.O., 2003. The encyclopedia of life. Trends in Ecology and Evolution 18, 77–80.

Witting, L., Loeschcke, V., 1995. The optimization of biodiversity conservation. Biological Conservation 71, 205–207.

Yoder, A.D., Nowak, M.D., 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Annual Review of Ecology Evolution and Systematics 37, 405–431.

Zhou, X., Adamowicz, S.J., Jacobus, L.M., DeWalt, R.E., Hebert, P.D.N., 2009. Towards a comprehensive barcode library for arctic life – Ephemeroptera, Plecoptera, and Trichoptera of Churchill, Manitoba, Canada. Frontiers in Zoology 6.