Molecular Phylogenetics and Evolution 66 (2013) 979–991

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

# Endemism and diversification in freshwater insects of Madagascar revealed by coalescent and phylogenetic analysis of museum and field collections

Laurent Vuataz [a,b,*], Michel Sartori [a], Jean-Luc Gattolliat [a], Michael T. Monaghan [c]

[a] Musée cantonal de zoologie, Palais de Rumine, place de la Riponne 6, 1014 Lausanne, Switzerland
[b] Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland
[c] Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 301, 12587 Berlin, Germany

ARTICLE INFO

*Article history:*
Received 27 February 2012
Revised 26 November 2012
Accepted 5 December 2012
Available online 20 December 2012

*Keywords:*
DNA taxonomy
Phylogeny
GMYC model
Ephemeroptera
Mayfly
Heptageniidae

ABSTRACT

The biodiversity and endemism of Madagascar are among the most extraordinary and endangered in the world. This includes the island's freshwater biodiversity, although detailed knowledge of the diversity, endemism, and biogeographic origin of freshwater invertebrates is lacking. The aquatic immature stages of mayflies (Ephemeroptera) are widely used as bio-indicators and form an important component of Malagasy freshwater biodiversity. Many species are thought to be microendemics, restricted to single river basins in forested areas, making them particularly sensitive to habitat reduction and degradation. The Heptageniidae are a globally diverse family of mayflies (>500 species) but remain practically unknown in Madagascar except for two species described in 1996. The standard approach to understanding their diversity, endemism, and origin would require extensive field sampling on several continents and years of taxonomic work followed by phylogenetic analysis. Here we circumvent this using museum collections and freshly collected individuals in a combined approach of DNA taxonomy and phylogeny. The coalescent-based GMYC analysis of DNA barcode data (mitochondrial COI) revealed 14 putative species on Madagascar, 70% of which were microendemics. A phylogenetic analysis that included African and Asian species and data from two mitochondrial and four nuclear loci indicated the Malagasy Heptageniidae are monophyletic and sister to African species. The genus *Compsoneuria* is shown to be paraphyletic and the genus *Notonurus* is reinstalled for African and Malagasy species previously placed in *Compsoneuria*. A molecular clock excluded a Gondwanan vicariance origin and instead favoured a more recent overseas colonization of Madagascar. The observed monophyly and high microendemism highlight their conservation importance and suggest the DNA-based approach can rapidly provide information on the diversity, endemism, and origin of freshwater biodiversity. Our results underline the important role that museum collections can play in molecular studies, especially in critically endangered biodiversity hotspots like Madagascar where entire species or populations may go extinct very quickly.



## 1. Introduction

The high species richness and endemism found on Madagascar make it one of the world's most important biodiversity hotspots (Myers et al., 2000; Ganzhorn et al., 2001; Goodman and Benstead, 2003). With an estimated 90% of natural habitats having been degraded by human activities (Myers et al., 2000; Goodman and Benstead, 2005), the island's biodiversity is one of the most endangered on Earth (Goodman and Benstead, 2005). Malagasy freshwater biodiversity is particularly high (Groombridge and Jenkins, 2002; Benstead and Pringle, 2004; Isambert et al., 2011) and >95% of freshwater species are endemic to the island (Gibon, 2000; Elouard and Gibon, 2003). Nonetheless, many groups remain poorly known to taxonomy (Goodman and Benstead, 2003). This is the case for most stream insects on Madagascar (Elouard et al., 2001; Benstead et al., 2003), despite their globally recognized importance as bioindicators of water quality.

Mayflies (Ephemeroptera) are a diverse and abundant component of stream ecosystems, particularly in the tropics (Barber-James et al., 2008). Their diversity and endemism on Madagascar is also high, with >100 described species and only one species known from elsewhere (Gattolliat and Rabeantoandro, 2002; Elouard et al., 2003). Most taxonomic work to date has been on the Baetidae, which is the most diverse family of mayflies on Earth (>800 species; Barber-James et al., 2008). In contrast, the Heptageniidae are virtually unknown on Madagascar despite being the third most diverse family globally (>500 species; Barber-James et al., 2008). It is thought that fewer than 10 species of Heptageniidae occur on Madagascar (Elouard et al., 2003) and only two species have been described: *Afronurus matitensis* and *Thalerosphyrus josettae* (Sartori

* Corresponding author at: Musée cantonal de zoologie, Palais de Rumine, place de la Riponne 6, 1014 Lausanne, Switzerland.
*E-mail addresses:* laurent.vuataz@vd.ch, laurent.vuataz@unil.ch (L. Vuataz).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.12.003

and Elouard, 1996). The latter was recombined as *Compsoneuria josettae* (Webb et al., 2006). Both genera have a disjunct distribution in Africa, Madagascar and Southeast Asia (Webb and McCafferty, 2008), and a complex taxonomic status that is still in flux (e.g. Gillies, 1984; Braasch and Soldán, 1986; Wang and McCafferty, 2004; Braasch and Freitag, 2008; Braasch and Boonsoong, 2010). The genera have different ecological requirements, *Afronurus* Lestage, 1924 being a cool adapted genus living on the stony bottom of brooks and streams (Schoonbee, 1973; Harrison and Hynes, 1988; Braasch and Boonsoong, 2010), whereas *Compsoneuria* Eaton, 1881 is a warmer water genus colonizing quieter zones of streams and rivers, in peculiar riparian vegetation and dead wood (Schoonbee, 1967; Gillies, 1984; Elouard et al., 2003; Braasch and Boonsoong, 2010).

The unique composition of the fauna and flora of Madagascar, i.e. the exceptional diversity and endemism of some groups and the absence of others, has intrigued the scientific community for a long time (Yoder and Nowak, 2006). The classical explanation is based on Gondwanan vicariance, implying an origin of the organisms that pre-dates the complete island isolation approximately 80 million years (Myr) ago (Briggs, 2003; de Wit, 2003) after its separation with India. However, it is now widely recognized that an important part of extant Malagasy diversity originated from more recent overseas dispersals followed by adaptive radiations (Yoder and Nowak, 2006). Several molecular studies of insect diversification have found evidence for one or more overseas dispersal events (e.g. Torres et al., 2001; Zakharov et al., 2004; Orsini et al., 2007; Wirta et al., 2008, 2010; Nobre et al., 2010). The only phylogenetic study of mayflies focused on Baetidae (Monaghan et al., 2005) and recovered one large endemic lineage and six other lineages that had African, Asian, or Malagasy origin with one or more oceanic dispersal events in each. This was surprising, in light of the classical view of their limited dispersal capacities (Sartori et al., 2000; Brittain and Sartori, 2003) owing to the very short life of the winged adults and strict habitat fidelity of the larvae.

Most phylogenetic studies of lineage origins and diversification use one individual per species. These are usually described species, named and delineated based on many years of taxonomic research. The time between first collection of new taxa and a phylogenetic understanding of the whole lineage can take many years. Here we combine phylogeny and DNA taxonomy (Vogler and Monaghan, 2007) in what is the first molecular-based study of the virtually unknown Malagasy Heptageniidae. Using multiple mitochondrial (referred to as mt hereafter) and nuclear (n hereafter) markers, amplified from both museum specimens and newly collected individuals, this combined approach allowed us to clarify the evolutionary relationships within the family in Madagascar while obtaining an accurate estimate of species richness of our Malagasy sampling. African and Asian museum specimens were included with the specific aim of testing the relationship of Malagasy species to these faunas in order to understand the depth of endemism and to examine colonization history. Based on the few morphological studies available, we predicted at least two distinct lineages in the Malagasy data set, each forming a monophyletic group with African + Asian *Afronurus* and *Compsoneuria* individuals, respectively. We also expected low species diversity (<10 species) according to earlier estimates (Elouard et al., 2003).

## 2. Materials and methods

### 2.1. Sampling

We used individuals newly collected in the field for this study as well as collections from the Museum of Zoology in Lausanne, Switzerland (MZL). In Madagascar, larval individuals were collected from streams in November and December 2007 from 11 localities in Marojejy, Andringitra and Andohahela National Parks using Surber nets. Adults were caught using hand nets. Individuals were preserved in 100% ethanol in the field, returned to the laboratory, and stored at −20 °C in fresh 100% ethanol. Museum specimens originated from a total of 21 localities in Madagascar. Taken together, sampling covered the main climatic regions of Madagascar except the West coast where Heptageniidae are rare or absent (Elouard et al., 2001). In the laboratory, individuals were separated into *Afronurus* and *Compsoneuria* according to the key provided by Webb and McCafferty (2008). A total of 61 Malagasy individuals were used for analyses (20 *Afronurus* and 41 *Compsoneuria*). Thirty individuals from South Africa (seven localities), Borneo (three localities), Sumatra (eight localities), Java (one locality) Sulawesi (one locality), Sumbawa (one locality) and Luzon (one locality) were also taken from the MZL collections for analyses (Fig. 1). Using the key provided by Webb and McCafferty (2008), 10 South African and five Southeast Asian individuals were assigned to *Afronurus,* whereas one South African and one Southeast Asian individuals were assigned to *Compsoneuria*. Among the remaining individuals, five were attributed to *Thalerosphyrus* Eaton, 1881, four to *Atopopus* Eaton, 1881, one to *Asionurus* Braasch and Soldán, 1986, and three were unidentified at the genus level (hereafter referred to as "Heptageniidae 1, 2 and 3"; see Table S1 for a detailed list of the sampling). All individuals belonged to the subfamily Ecdyonurinae.

Three views (ventral, dorsal and lateral) of each individual were photographed using an Olympus ColorView IIIu camera (Olympus Corporation) connected to a Leica M205 C stereomicroscope (Leica Microsystems). This database is available for later verification of morphological characters. In particular, we aimed to capture coloration that is lost using otherwise non-destructive DNA extraction following Vuataz et al. (2011). Extracted DNA, individuals and photographs are deposited at the MZL.

### 2.2. COI gene tree

We amplified a 658-bp fragment of mt protein-coding cytochrome c oxidase subunit I (COI) using LCO1490 and HCO2198 primers (Folmer et al., 1994) for all 91 individuals of our data set. Polymerase chain reaction (PCR), agarose gel electrophoresis, purification of PCR products and sequencing were conducted as described in Vuataz et al. (2011). Forward and reverse sequencing reads were assembled and edited using CodonCode Aligner 3.7.1.1 (CodonCode Corporation, Dedham, MA).

Sequence alignment, performed using MAFFT (Katoh et al., 2005) as implemented in Jalview 2.6.1 (Clamp et al., 2004), was straightforward as no insertions or deletions were observed in the COI data set. Identical haplotypes were removed from the alignment using Collapse 1.2 (Posada, 2004). The best evolutionary model was selected following the second-order Akaike information criterion (AICc) implemented in MrAIC 1.4.4 (Nylander, 2004; Table 1) under the option using the models implemented in MrBayes (Ronquist and Huelsenbeck, 2003). In order to accommodate different substitution rates among codon positions, we used partitioned models of evolution (e.g. Brandley et al., 2005; Shapiro et al., 2006). Consequently, we examined COI in two partitions, one with first and second codon positions and one with third positions (1 + 2, 3). Bayesian inference (BI) and Maximum Likelihood (ML) tree searches were conducted using MrBayes 3.1.2 and Treefinder v. March 2011 (Jobb et al., 2004), respectively, and performed at the freely available Bioportal (http://www.bioportal.uio.no; Kumar et al., 2009). Two independent analyses of four MCMC chains run for 20 million generations with trees sampled every 1000 generations were used for BI. The stationary

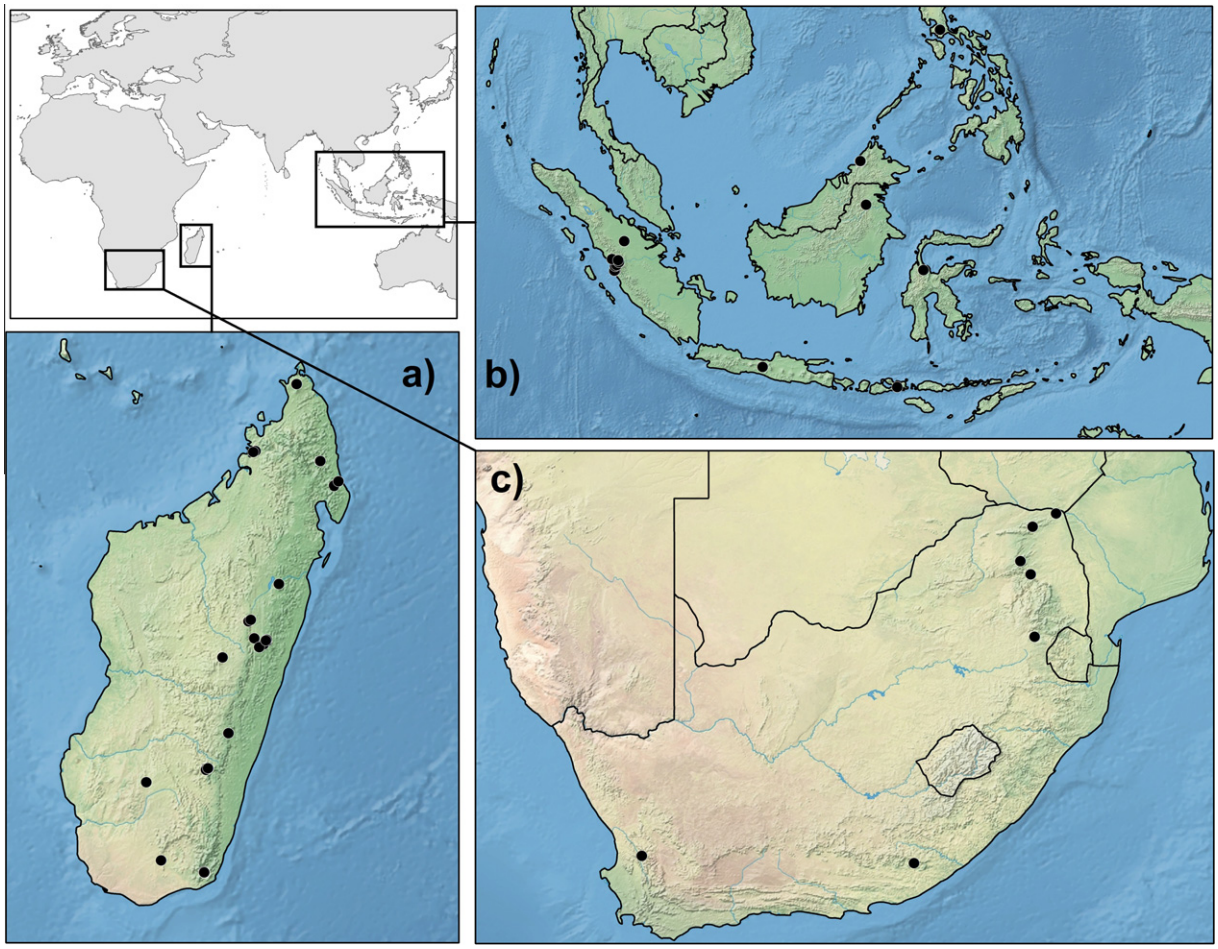


**Fig. 1.** Geographical origins of the Heptageniidae samples used in the study. Sampled localities (filled circles) are shown for (a) Madagascar; (b) Southeast Asia; and (c) South Africa.

**Table 1**
Sequence variation for each sequenced gene region and for the three concatenated matrices (mitochondrial COI + 16S, nuclear wg + EF-1α + H3 + 28S, all six fragments combined). Reduced data set (top) was used for phylogenetic analyses, whereas complete data sets (bottom) were used for gene tree reconstruction (all) and species delimitation (Malagasy).

| Reduced data set | *n* | bp | *K* | *S* | $S_i$ | $\%S_i$ | Model |
|---|---|---|---|---|---|---|---|
| COI | 43 | 658 | 43 | 257 | 247 | 38 | GTR + Γ + I |
| 16S | 36 | 515 | 35 | 188 | 167 | 32 | HKY + Γ + I |
| wg | 43 | 478 | 38 | 121 | 89 | 19 | GTR + Γ |
| EF-1α | 36 | 139 | 22 | 28 | 22 | 16 | K2P + Γ + I |
| H3 | 39 | 328 | 32 | 94 | 75 | 23 | HKY + Γ + I |
| 28S | 34 | 659 | 22 | 85 | 62 | 9 | GTR + Γ + I |
| Mitochondrial | 43 | 1173 | 43 | 445 | 414 | 35 | |
| Nuclear | 43 | 1604 | 41 | 328 | 261 | 17 | |
| Combined | 43 | 2777 | 43 | 773 | 676 | 25 | |
| *Complete data sets* | | | | | | | |
| COI (all) | 91 | 658 | 71 | 257 | 247 | 38 | GTR + Γ + I |
| COI (Malagasy) | 61 | 658 | 42 | 220 | 177 | 27 | HKY + Γ + I |

*n* = Number of sequences, bp = size of aligned matrices, *K* = number of haplotypes (COI, 16S, mitochondrial) or genotypes (wg, EF-1α, H3, 28S, nuclear) or a combination of both (combined), *S* = number of polymorphic sites, $S_i$ = number of parsimony-informative sites, model = best evolutionary model selected following the second-order Akaike information criterion (AICc). Concatenated matrices were examined with partitioned analysis using the best model for each gene fragment (see Materials and methods).

nucleotide frequencies, the alpha shape parameter of the gamma distribution, the relative rate of substitution and the proportion of invariant sites were unlinked across partitions and the ratepr command was set to variable (see Marshall et al., 2006). Two million generations were discarded as a burnin after visually verifying that likelihood curves had flattened-out and that the independent runs converged using Tracer 1.5 (Drummond and Rambaut, 2007). A ML bootstrap analysis of 1000 replicates was conducted with all model parameters set to optimum and all other options set to default. An *Heptagenia sulphurea* (Müller, 1776) individual belonging to the related subfamily Heptageniinae was used as an outgroup. All COI sequences are available from EMBL database (HE651331 – HE651395 and HF536601 – HF536607).

### 2.3. Multiple gene phylogeny

A subset of 43 individuals (21 from Madagascar, five from South Africa and 17 from Southeast Asia) was selected for a multiple gene phylogeny. These individuals were chosen from among each of the main COI gene tree lineages (Fig. S1 and Table S1). We took up to half of the individuals within each clade of closely related COI sequences as well as all singletons that did not obviously cluster with any others. In addition to COI, we amplified five gene fragments for this reduced data set: mt 16S rDNA using the 16Sar (Simon et al., 1994) and 16S2 (Giessler et al., 1999) primers; n 28S rDNA using the 28SFF and 28SDD primers (see Pons et al., 2004); and the n protein-coding genes histone 3 (H3) using the HexAF and HexAR primers (see Ogden and Whiting, 2003), elongation factor 1 alpha (EF-1α) using the primers specified in Takemon et al. (2006), and

wingless (wg) using newly designed wgFlv2 (5′-CTDCCATTATTCCGTGTAGTTGG) and wgRlv498 (5′-GTACATTCACATCTTTCTCTTAC) primers. These were designed from sequences first obtained using Wg578F (Ward and Downie, 2005) and Wg1032R (Abouheif and Wray, 2002) primers. All laboratory procedures and editing of sequences were conducted as described in Section 2.2, but with PCR annealing temperature ranging between 47 °C and 48 °C for 16S and wg, between 47 °C and 52 °C for 28S and EF-1α, and between 47 °C and 58 °C for H3. The heterozygous sites, typically identified as double peaks within the chromatograms, were coded according to the IUBMB code.

Initial alignments of all gene fragments were performed as described in Section 2.2. Alignments of EF-1α and H3 sequences were straightforward as no insertions or deletions were observed in these data sets. Insertions of between three and 30 nucleotides were observed for five Southeast Asian individuals in the wg dataset. In this case, RevTrans 1.4 (Wernersson and Pedersen, 2003), which translates the DNA into peptide sequences before "reverse-translating" the resulting alignment into DNA, was used under the Dialign-T alignment method. The rDNA alignments (16S and 28S) were optimized in MAFFT 6.240 as implemented in GenomeNet (www.genome.jp/tools/mafft/) under the G-INS strategy with a gap opening penalty of 1 and all other parameters set to default. Identical haplotypes (COI, 16S) and genotypes (wg, EF-1α, H3, 28S) were removed, and selection of the best model of evolution for each gene fragment (Table 1) as well as BI and ML tree searches were conducted as described in Section 2.2. Each gene fragment was used as a partition in addition to the (1 + 2, 3) COI partition scheme. As the K2P model (see Table 1) is not implemented in Treefinder, we used the best-fit model of evolution among those available in Treefinder following the AICc for the EF-1α partition, which was a GTR + Γ model. To verify congruence between mt (COI and 16S) and n (wg, EF-1α, H3, 28S) data sets, we first conducted two independent analyses. As no significant incongruence appeared (see Section 3.2), we ran a combined analysis using all six gene fragments. In each analysis, *Heptagenia sulphurea* was used as an outgroup. All sequences are available from EMBL database (16S: HE651396 – HE651431; wg: HE651455 – HE651486 and HF536594 – HF536600; EF-1α: HE651514 – HE651536; H3: HE651487 – HE651513 and HF536587 – HF536593; 28S: HE651432 – HE651454).

### 2.4. Tests of topology

Monophyly of particular lineages (see Section 3.2 and 3.3) was tested by comparing our ML topologies obtained with mt, n or combined data sets with alternative topologies using the Shimodaira-Hasegaewa (SH) test (Shimodaira and Hasegawa, 1999) and the approximately unbiased (AU) test (Shimodaira, 2002) as implemented in Treefinder. We used one million RELL replicates (see Kishino et al., 1990). Alternative tree topologies were reconstructed by manually modifying our ML topologies in TreeView 1.6.6 (Page, 1996).

### 2.5. Species delimitation on Madagascar

Species delimitation was performed on the 61 Malagasy individuals present in the complete COI data set (Table 1) using the general mixed Yule-coalescent (GMYC) approach (Pons et al., 2006; Fontaneto et al., 2007). We selected COI because it was the most variable of our gene fragments (Table 1) and because it is routinely used in species identification (e.g. DNA barcoding) and delimitation. The GMYC model is a likelihood approach that combines equations from coalescent and Yule models to define one or more thresholds representing the species boundary on ultrametric gene trees. More detailed descriptions are available in Pons et al. (2006) and Monaghan et al. (2009).

Identical COI haplotypes were first removed, and ultrametric gene trees were reconstructed under a relaxed molecular clock (uncorrelated lognormal) model in BEAST 1.6.1. (Drummond and Rambaut, 2007). The clock rate was fixed to 1, a UPGMA starting tree was used and a coalescent (constant size) prior was selected because a single coalescent cluster constitutes the GMYC null model (see Monaghan et al., 2009). The best-fit model of evolution (Table 1) was implemented with six gamma categories. To explore the effect of a partitioning scheme on the GMYC outputs, we conducted two separate analyses: one with a single partition (codon positions 1 + 2 + 3) and one with third positions treated as a separate partition (1 + 2, 3). For the partitioned analysis, the substitution rate parameters, the rate heterogeneity model and the base frequencies were unlinked across partitions. All other parameters were set to default. For both analyses, two independent MCMC chains were run for 50 million generations and sampled every 1000 generations, resulting in 50,000 trees for each run. Run convergence was visually verified in Tracer and the independent log and tree files were combined using LogCombiner 1.6.1 (Drummond and Rambaut, 2007) after discarding the first 5,000 trees (10%) from each run as a burnin, resulting in 90,000 trees in the combined analyses. All model parameters of the combined log files reached an estimated sample size (ESS) >600. The maximum clade credibility tree found using TreeAnnotator 1.6.1 (Drummond and Rambaut, 2007) with all options set to default was used as input data for the GMYC model. Single and multiple-threshold GMYC models were optimized using the script available within the SPLITS package (r-forge.r-project.org/projects/splits/) for R (R Development Core Team, 2011).

## 3. Results

### 3.1. Gene fragment characteristics

There were no missing data in the COI matrices and the combined data matrix of six gene fragments (Table 1) was >90% complete. Missing data resulted from seven 16S, seven EF-1α, four H3 and nine 28S sequences that failed to amplify. A total of 93 sites were heterozygous (<0.1% of all data). The length of the 16S and 28S sequences varied between 508 and 511 bp and between 652 and 657 bp, respectively. No introns were detected within the protein-coding gene fragments. Approximately 60% of total variation was found in mtDNA and 40% in nDNA. The percentage of parsimony-informative sites within each gene fragment ranged from 38% (COI) to 9% (28S; Table 1).

### 3.2. Mitochondrial vs. nuclear phylogeny

The mt (COI + 16S) and n (wg, EF-1α, H3, 28S) phylogenies revealed broadly congruent topologies (Figs. 2 and 3), both of which were similar to the gene tree topology of the complete COI data set of all sampled individuals (Fig. S1). In the mt tree, all Malagasy individuals formed a monophyletic lineage with high BI posterior probability (PP) and ML bootstrap support (BS) (PP = 1; BS = 83; Fig. 2) and the African *Compsoneuria* was sister taxon to this Malagasy lineage. In the n tree, fewer nodes were supported overall and the African *Compsoneuria* was recovered within the Malagasy lineage (PP = 1; BS = 92; Fig. 3). To test for the Malagasy lineage monophyly, we compared our mt ML topology with an alternative hypothesis that included the African *Compsoneuria* individual within the Malagasy lineage as a basal, unresolved lineage. Our mt ML topology was significantly better than the alternative topology ($p < 0.05$; Table 2), meaning that the monophyly of the

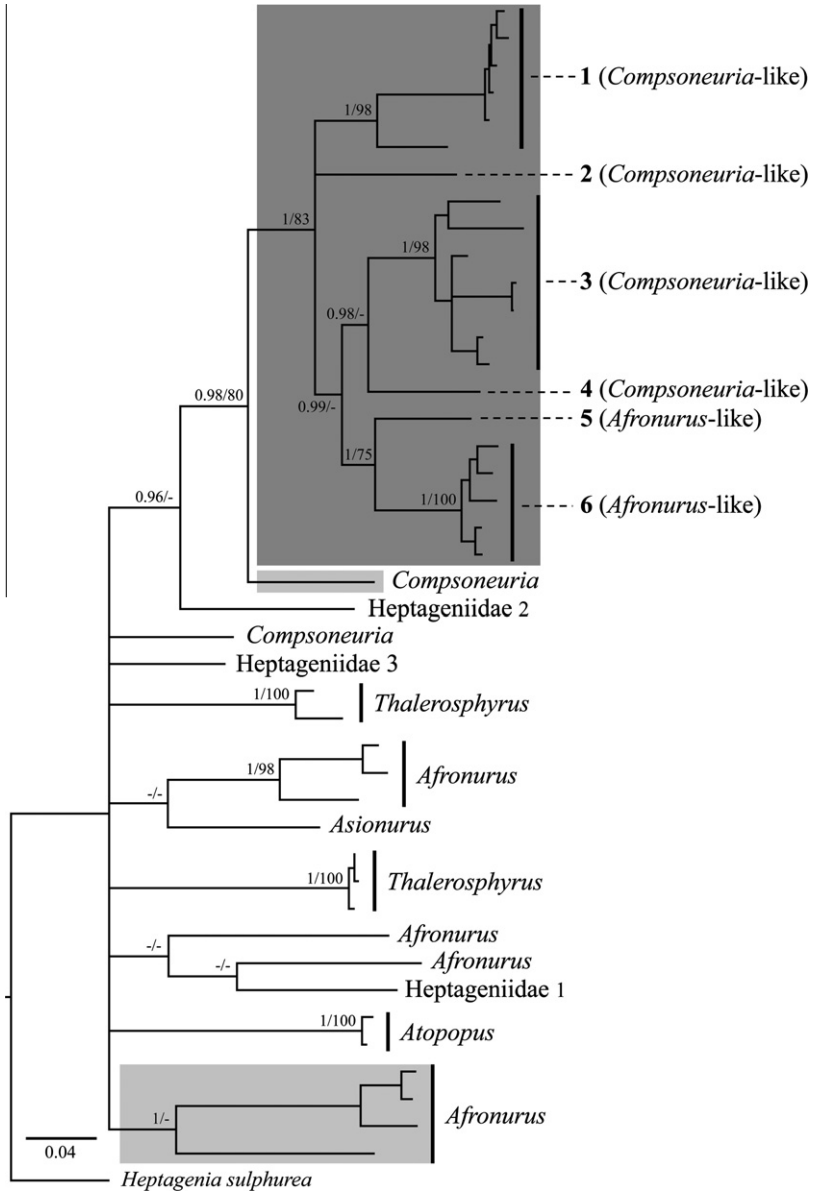


**Fig. 2.** Bayesian majority-rule consensus tree reconstructed from the mitochondrial data set (COI + 16S; 1173 bp). Dark grey shading highlights the Malagasy lineage, grey shading the African lineages, and no shading the Southeast Asian lineages. Malagasy clades and singletons (numbered 1–6) are labeled by their morphological affinities. Deeper nodes are labeled with Bayesian posterior probability (PP)/maximum likelihood bootstrap support (BS) only for PP > 0.8 and BS > 70.

Malagasy lineage was significantly supported by our mt ML reconstruction. We tested monophyly in the n ML topology with an alternative hypothesis placing the African *Compsoneuria* individual as the closest relative to the Malagasy lineage. Our n ML topology was not significantly better than the alternative topology ($p > 0.05$; Table 2), meaning that our n ML reconstruction did not exclude the monophyly of the Malagasy lineage.

In both mt and n phylogenies, three well-supported clades and three singletons (numbered 1–6; Figs. 2 and 3; see also Fig. S1) could be distinguished within the Malagasy lineage. Clades 1 and 3 as well as singletons 2 and 4 were composed of individuals morphologically similar to *Compsoneuria* (labeled "*Compsoneuria*-like"), whereas singleton 5 and clade 6 were composed of individuals morphologically similar to *Afronurus* ("*Afronurus*-like"). The relationships among these lineages varied in different reconstructions: the mt phylogeny recovered clade 3 and singleton 4 as sister to singleton 5 and clade 6 with high PP (0.99) but no BS (<70), whereas the n phylogeny recovered singleton 4 as basal to the others, but without support (PP < 0.8; BS < 70).

In the mt reconstructions, the closest relative to the Malagasy + African *Compsoneuria* lineage was Heptageniidae 2 from Southeast Asia, although only supported by the BI analysis (PP = 0.96). In the n reconstructions, the closest relatives to the Malagasy + African *Compsoneuria* lineage were *Compsoneuria*, Heptageniidae 2 and Heptageniidae 3 from Southeast Asia (PP = 1; BS = 95). The most striking difference between mt and n reconstructions was the better-resolved relationships of deep nodes in the n reconstructions, although most often only supported by the BI analysis. In the n reconstructions, the *Thalerosphyrus* individuals occurred in a single monophyletic lineage (PP = 0.98; BS = 75), as well as the *Afronurus* individuals, although only supported by the BI analysis (PP = 1; BS < 70). In the mt reconstructions, neither genus was monophyletic (unresolved relationships).

### 3.3. Combined phylogeny

The combined phylogeny of the six-gene data set grouped all Malagasy individuals in a single monophyletic lineage, although

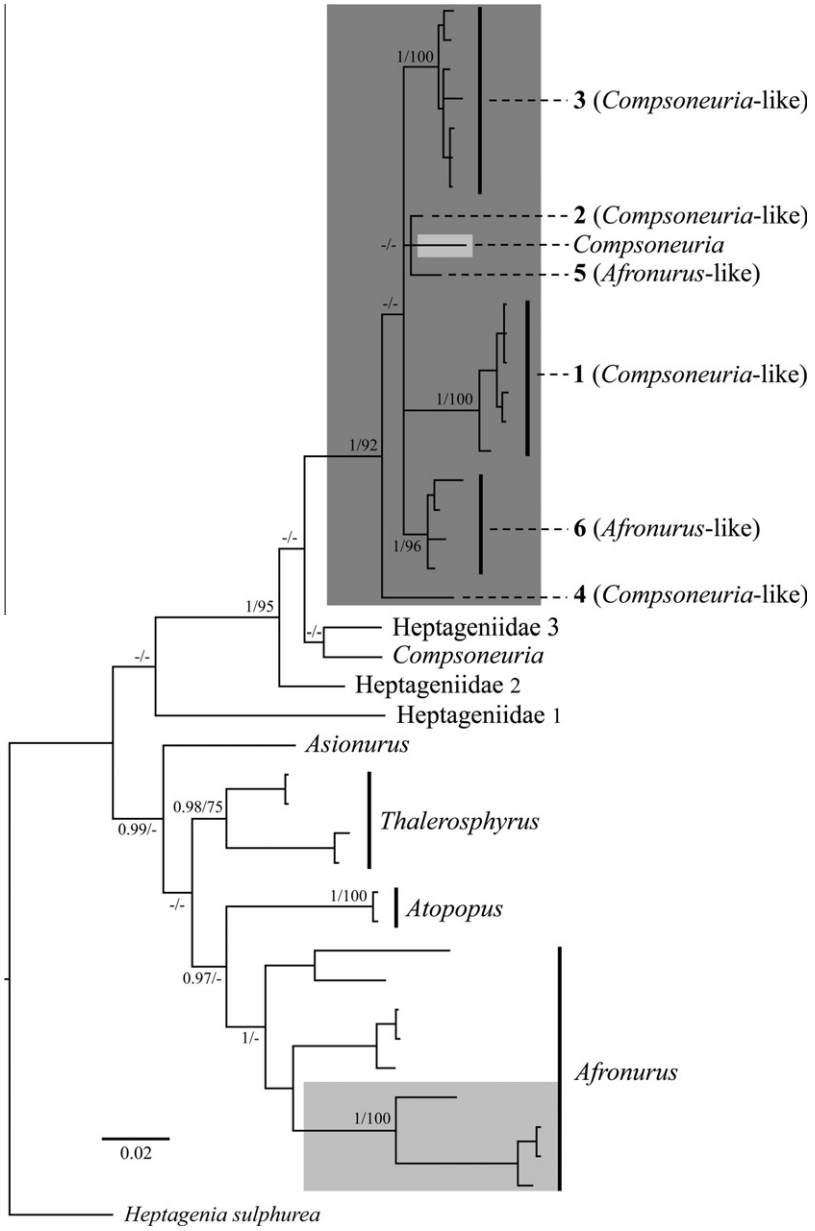


**Fig. 3.** Bayesian majority-rule consensus tree reconstructed from the nuclear data set (wg + EF-1α + H3 + 28S; 1604 bp). Dark grey shading highlights the Malagasy lineage, grey shading the African lineages, and no shading the Southeast Asian lineages. Malagasy clades and singletons (numbered 1–6) are labeled by their morphological affinities. Deeper nodes are labeled with Bayesian posterior probability (PP)/maximum likelihood bootstrap support (BS) only for PP > 0.8 and BS > 70.

**Table 2**
Summary of SH and AU topological test outputs for each alternative hypothesis tested against the best ML topology. The involved data set is specified.

| Data set | Alternative hypothesis | SH | AU |
|---|---|---|---|
| Mitochondrial | African *Compsoneuria* within Malagasy lineage | 0.0404* | 0.0203* |
| Nuclear | African *Compsoneuria* sister to Malagasy lineage | 0.2583 | 0.2486 |
| Combined | African *Compsoneuria* within Malagasy lineage | 0.0381* | 0.0250* |

* Alternative topologies were rejected at $p < 0.05$.

this was weakly supported (PP = 0.89; BS < 70; Fig. 4). African *Compsoneuria* was recovered as the closest relative to the Malagasy lineage (similarly to the mt topology) and the African *Compsoneuria* + Malagasy lineage was highly supported (PP = 1; BS = 93). To test for the Malagasy lineage monophyly, we compared our combined ML topology with an alternative hypothesis that included the African *Compsoneuria* individual within the Malagasy lineage as a basal, unresolved lineage. Our combined ML topology was significantly better than the alternative topology ($p < 0.05$; Table 2), meaning that the monophyly of the Malagasy lineage was significantly supported by our combined ML reconstruction.

The Malagasy clades and singletons 1–6 as defined in Section 3.2 were also recovered and the three clades 1, 3 and 6 were fully supported (PP = 1; BS = 100). Relationships were similar to those

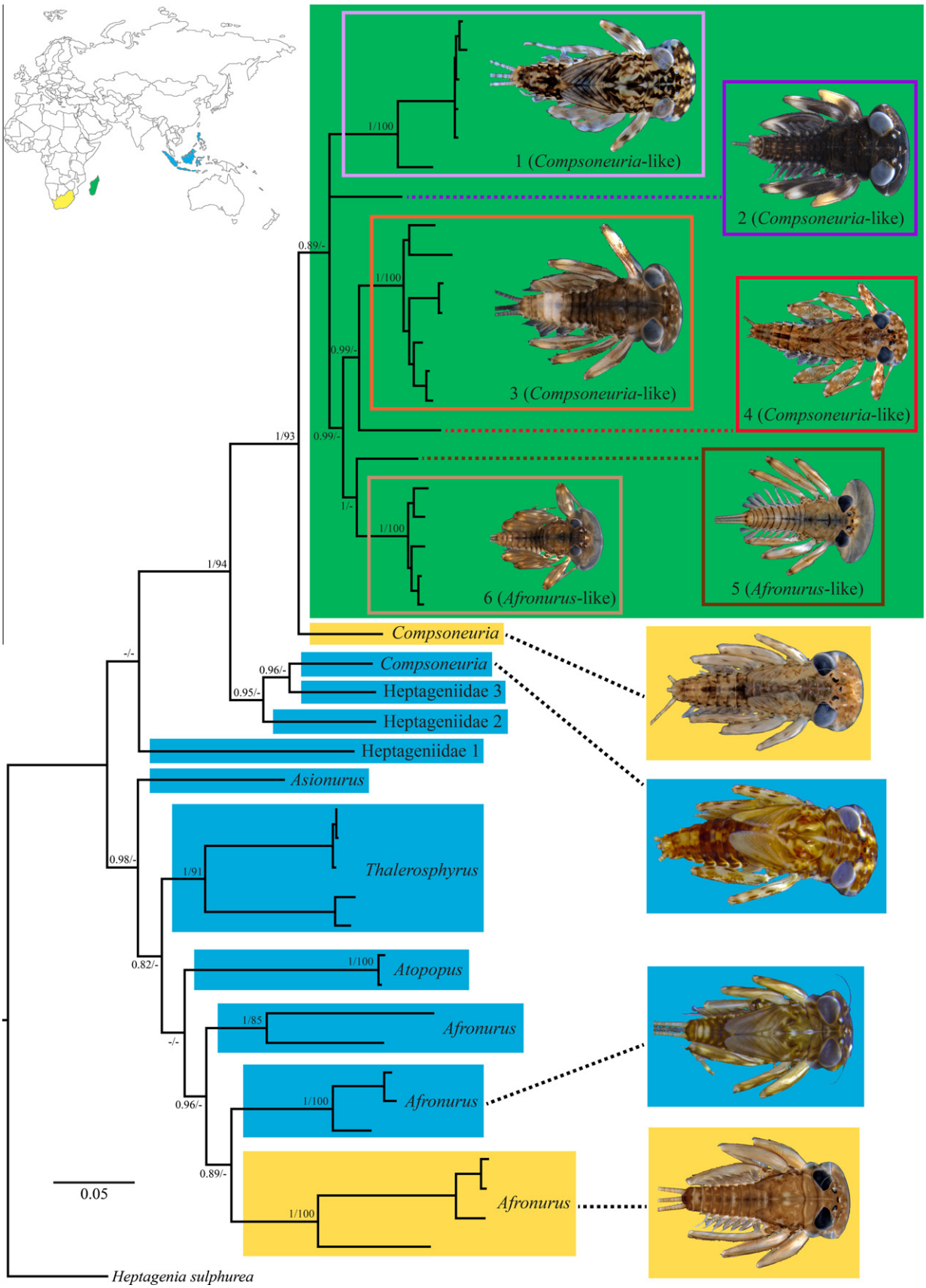


**Fig. 4.** Bayesian majority-rule consensus tree reconstructed from the combined data set (COI + 16S + wg + EF-1α + H3 + 28S; 2777 bp). The Malagasy lineage is highlighted in green, African lineages in yellow, and Southeast Asian lineages in blue. Malagasy clades and singletons (numbered 1–6) are labeled by their morphological affinities. Deeper nodes are labeled with Bayesian posterior probability (PP)/maximum likelihood bootstrap support (BS) only for PP > 0.8 and BS > 70. Photographs depict one individual (nymph) representing each of the six Malagasy lineages, as well as for African and Asian *Afronurus* and *Compsoneuria* lineages.

recovered from the mt reconstruction, with clade 3 and singleton 4 as sister to singleton 5 and clade 6, only supported by BI reconstruction (PP = 0.99; BS < 70). To test whether the combined ML reconstruction significantly supported the six lineages as separate entities, we compared our ML topology with 15 alternative topologies: each of the three clades was moved within the two other ones (six alternative topologies) and each of the three singletons was moved within each three clades (nine alternative topologies). The moved lineages were placed as basal, unresolved lineages within clades. In each case, the alternative topology was significantly rejected with *p*-values <0.0003 (not shown), indicating that the combined ML reconstruction significantly supported the six Malagasy lineages as distinct entities.

As in the n reconstructions, the closest relatives to the Malagasy + African *Compsoneuria* lineage were *Compsoneuria*, Heptageniidae 2 and Heptageniidae 3 (hereafter Asian *Compsoneuria* clade) from Southeast Asia (PP = 1; BS = 94). All the other individuals belonged to a distinct monophyletic lineage well supported by BI reconstruction (PP = 0.98; BS < 70), except Heptageniidae 1 that presented an unsupported position in the topology. As in the n phylogeny, the *Thalerosphyrus* and the *Afronurus* lineages were each monophyletic, although the latter was not supported by ML analysis (BS < 70).

### 3.4. Species delimitation

For each partitioning scheme (see Section 2.5), both single and multiple-threshold GMYC models provided a better fit to the COI ultrametric tree than the null model (likelihood ratio test, $p < 0.005$; Table 3). The single-threshold model delimited 14 putative species composed of six clusters and eight singletons (Fig. 5). The multiple-threshold model delimited 20 putative species (eight clusters and 12 singletons; not shown) but did not fit the data significantly better that the single-threshold model (likelihood ratio test, $p = 0.55$). Consequently, we refer to single-threshold results hereafter.

## 4. Discussion

### 4.1. Phylogeny and colonization

The monophyly of the Malagasy lineage, recovered by the mt and combined analyses and equivocal from the n phylogeny, was surprising and contradicted our expectations based on morphological similarities with *Afronurus* and *Compsoneuria* from Africa and Asia. Our analyses demonstrated that this Malagasy lineage is not closely related to *Afronurus,* but is close to *Compsoneuria*. Interestingly, the African *Compsoneuria* was recovered as sister to the Malagasy lineage, followed by the Southeast Asian *Compsoneuria* clade. The absence of the *Afronurus* lineage in Madagascar and its presence in Africa and Asia casts doubt on the classical hypotheses that have been proposed for its distribution in the Oriental and Afrotropical realms. It is unlikely that the colonization of Southeast Asia occurred through the drifting of the Deccan plate, as documented for instance for the teloganodid mayflies (Sartori et al., 2008).

The Malagasy monophyly and sister relationship to African *Compsoneuria* suggest that Malagasy Heptageniidae resulted from (1) a Gondwanan vicariance, i.e. the split of an ancestral lineage when Africa and Madagascar separated some 165 Myr ago (Rabinowitz et al., 1983); (2) one or more post-Gondwanan overseas colonization events from Africa to Madagascar; or (3) colonization from Madagascar to Africa. Three lines of evidence can be invoked to reject hypothesis (1). Firstly, the oldest recorded Heptageniidae fossil is from the Late Cretaceous, ca. 95 Myr (Sinitshenkova, 2000). All other recorded fossils are from the Eocene or more recent epochs (e.g. Masselot and Nel, 1999; Godunko and Sontag, 2004). While only a minimum age, this oldest Heptageniidae fossil post-dates the separation of Madagascar from Africa by ca. 70 Myr. Secondly, the Heptageniidae comprise >550 species in 33 genera (Barber-James et al., 2010). Its absence from Australia, New Zealand and its very low diversity in Central and South America, combined with its high diversity in the Holarctic realm, suggests an origin in Laurasia (Edmunds, 1979; Barber-James et al., 2008). Three subfamilies are now recognized (Webb and McCafferty, 2008) and all Afrotropical species (*Afronurus, Compsoneuria* and the Malagasy lineage) belong to the subfamily Ecdyonurinae. This includes some species tolerant of low dissolved oxygen, high temperature, and eutrophication (M. Sartori, unpublished data). Together with their ability to move the gills for respiratory movements, a plesiomorphic condition according to Kluge (2004), this could explain their relative success in colonizing Africa. Therefore, there is evidence that the colonization of Africa by a small number of lineages happened thus only after the diversification of the family as a whole, i.e. after the Gondwana break-up. Thirdly, an estimation of the divergence time within the African *Compsoneuria* + Malagasy lineage using molecular dating with a relaxed lognormal molecular clock (data not shown) based on two contrasting COI evolution rates reported for insects (1.5% and 3.5%; see Papadopoulou et al., 2010) suggests a colonization period ranging from 10 to 55 Myr, which is not compatible with a Gondwanan vicariance. Consequently, overseas dispersal is a more likely scenario, as has been documented in a variety of other taxa (e.g. Nagy et al., 2003; Vences et al., 2003; Monaghan et al., 2005; Poux et al., 2005; Kohler and Glaubrecht, 2010).

Our n topology including African *Compsoneuria* within the Malagasy lineage would support hypothesis (3), i.e. overseas colonization from Madagascar to Africa. Hypothesis (3) would also be compatible with all our other topologies (mt and combined topologies) placing the African *Compsoneuria* sister to the Malagasy lineage only if the colonization occurred prior to the Malagasy lineage diversification or if the colonizing lineage went extinct in

**Table 3**
GMYC model outputs using single- and multiple-threshold approaches applied to Malagasy COI complete data set (see Table 1). The two tested partitioning schemes (codon position 1 + 2 + 3 and 1 + 2, 3) are specified.

| Partitions | GMYC | $L_o$ | $L_{GMYC}$ | $T$ | $N_{GMYC}$ | LR |
|---|---|---|---|---|---|---|
| 1 + 2 + 3 | Single | 226.3 | 233.2 | 0.0147 | 14 (8–19) | 13.9** |
| | Multiple | | 234.2 | – | 20 (9–20) | 15.9** |
| 1 + 2, 3 | Single | 228.5 | 235.5 | 0.0142 | 14 (8–19) | 14.0** |
| | Multiple | | 236.5 | – | 20 (10–20) | 16.1** |

Likelihoods are indicated for null ($L_o$) and GMYC ($L_{GMYC}$) models, where null likelihoods are the same for single- and multiple-threshold models. GMYC outputs include the threshold genetic distance from the branch tips where transition occurred ($T$, presented for single-threshold model), and the number of putative species as the sum of sequence clusters and singletons ($N_{GMYC}$), with its corresponding 95% confidence interval (between brackets). Significance of the likelihood ratio (LR) was evaluated using a chi-square test (see Section 2.5 for details of analyses).
** $p < 0.005$.

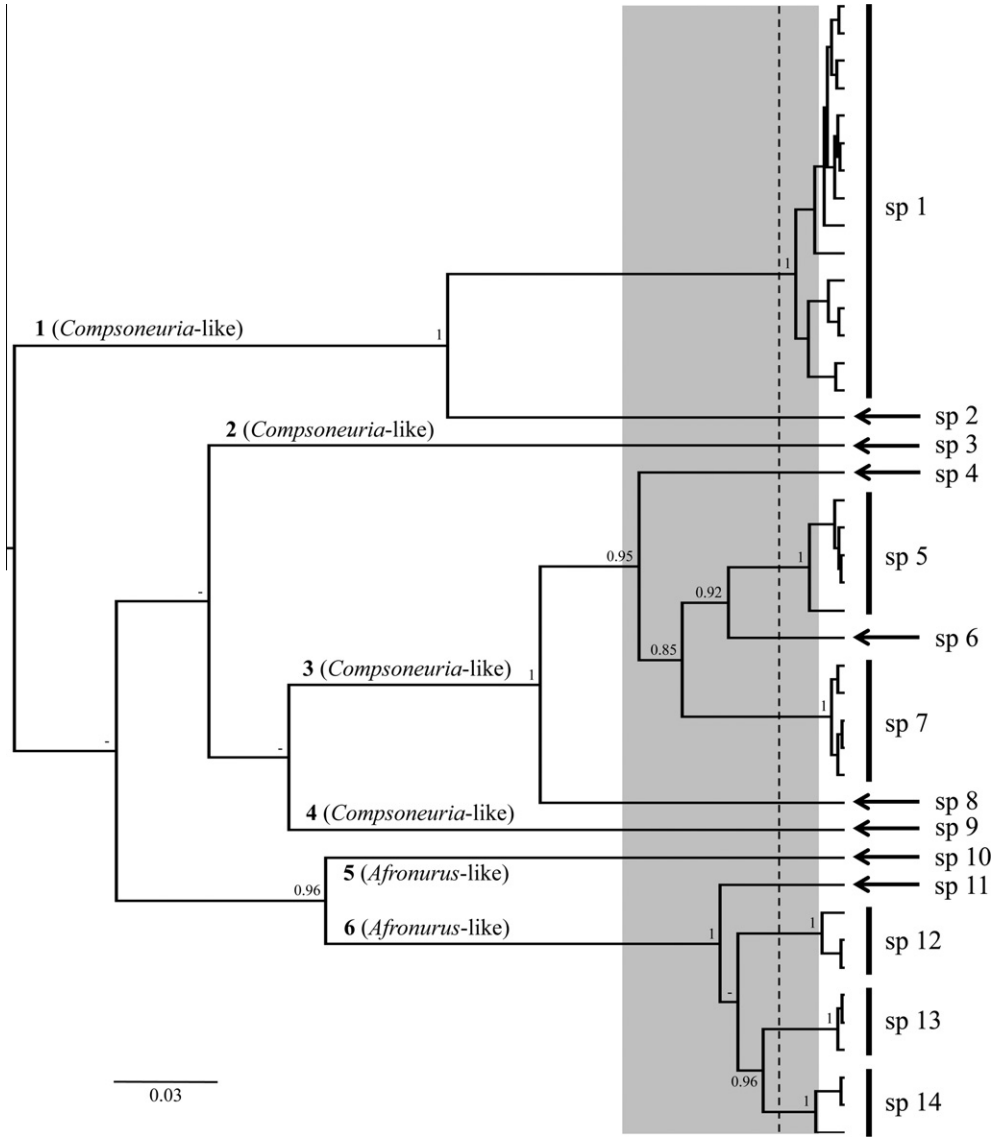


**Fig. 5.** Ultrametric Bayesian maximum clade credibility tree of the complete Malagasy COI data set (658 bp) obtained using a relaxed lognormal clock. The broken vertical line indicates the point of maximum likelihood fit of the single-threshold GMYC model, i.e. the point of transition from interspecies (Yule) to intraspecies (coalescent) branching events. The grey shading corresponds to the confidence interval of the transition point. The bars indicate significant clusters (arrows: significant singletons) that are inferred to be species (numbered sp 1–14). The six well-supported clades and singletons as recovered by mitochondrial, nuclear and combined data sets (Figs. 2–4) are specified above or below corresponding branches. Posterior probabilities (PP) values (if > 0.8) are indicated for nodes up to the GMYC species level.

Madagascar. However, according to all our other topologies and knowing that the n ML topology is compatible with a sister relationship between Malagasy lineage and African *Compsoneuria* (see Section 3.2), the simplest and most probable hypothesis to explain the current Heptageniidae radiation in Madagascar is hypothesis (2), i.e. overseas colonization from Africa to Madagascar. While it is not possible to address the number of distinct colonizations, no evidence of multiple events arose from our data. Including more Asian and African *Compsoneuria* lineages would be necessary for proposing a more accurate scenario.

### 4.2. Diversification hypothesis

Malagasy clade 1 (*Compsoneuria*-like) was composed of sp 1 sampled from 12 localities throughout Madagascar and sp 2 from a single locality (Figs. 5 and 6). Sp 1 was the only widespread putative species of our sampling, and occurred at relatively low elevation (<1000 m) in large rivers with an open canopy (i.e. exposed to sunlight). Interestingly, this is the habitat in which African *Compsoneuria* are usually found (Schoonbee, 1967; authors, pers. obs.). Sp 2 was sampled from one of the sp 1 localities. Clade 3 (*Compsoneuria*-like) included five putative species (spp 4–8). These were found at variable elevations: spp 4 and 6 <400 m; spp 7 and 8 from 850 to 1,300 m; sp 5 > 1750 m), but always in smaller rivers compared with sp 1, typically in tributaries, forest or mountain streams. The *Afronurus*-like lineages (lineages 5 and 6) included five putative species (spp 10–14) throughout Madagascar (sp 14 < 550 m; spp 10–12 from 700 to 1200 m; sp 13 > 1500 m), but always in primary forest streams.

From the above observations, we hypothesized the following evolution of Malagasy Heptageniidae. The African Heptageniidae that colonized Madagascar (see Section 4.1) probably had similar ecological characteristics to present-day African *Compsoneuria*, i.e. larvae inhabiting quiet zones of relatively warm rivers (Schoonbee, 1967; Gillies, 1984). Hence, the best initial environment in Madagascar would have been large and open rivers, like

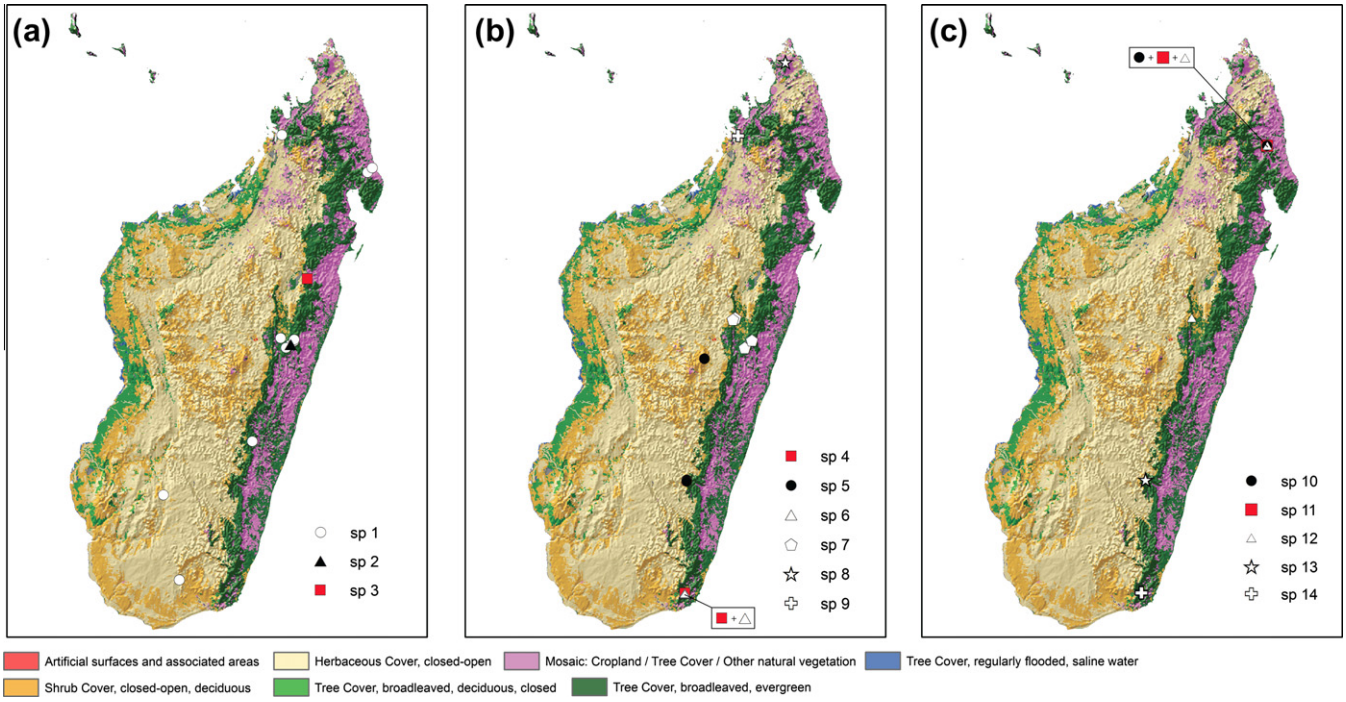


**Fig. 6.** Location of sampling sites (symbols) of Heptageniidae in Madagascar. (a) *Compsoneuria*-like lineages 1 and 2 (sp 1 – sp3); (b) *Compsoneuria*-like lineages 3 and 4 (sp 4–9); (c) *Afronurus*-like lineages 5 and 6 (sp 10–14; see Fig. 5). Land cover GLC2000 (Bartholome and Belward, 2005); ArcGIS$_{TM}$ (ESRI, 2011).

those inhabited today by clade 1 (*Compsoneuria*-like). A part of the earlier Heptageniidae would have adapted later to colder and faster flowing rivers, that is the forest and mountain tributaries that are colonized today by lineages 3 and 4, while another part would have strictly specialized on another environment, which is the forest streams currently inhabited by lineages 5 and 6. It has been documented for several mayfly families that functional morphology is linked to feeding behavior (McShaffrey and McCafferty, 1990), including potential adaptation of mouthparts to contrasting microhabitats in Heptageniidae species (McShaffrey and McCafferty, 1988). The diverging morphology of lineages 5 and 6 (*Afronurus*-like) compared to the other Malagasy Heptageniidae (*Compsoneuria*-like) could thus be the result of a morphological adaptation driven by a shift of feeding strategy due to new environmental conditions (forest streams). Their morphological affinities with *Afronurus*, including mouthparts similarities (see Sartori and Elouard, 1996 for *Afronurus matitensis*) could be a case of convergence due to similar shifts of environmental conditions in both lineages. Both our mt and combined topologies support this diversification hypothesis, with a basal position of clade 1 compared to lineages 3, 4, 5 and 6 (Figs. 2 and 4). Singleton 2 (sp 3) was excluded from these considerations because of its unresolved position in our topologies.

### 4.3. Sampling, species diversity and microendemism

An interesting result of our work was the higher estimated species diversity than expected (Elouard et al., 2003). Nonetheless, we must consider how our sampling strategy could influence the estimated number of species using the GMYC model. Widespread allopatric speciation could lead to an underestimate of species when large-scale sampling uncovers more sister taxa than more restricted sampling (Bergsten et al., 2012). On the other hand, when too few populations (demes) are sampled, an artificial splitting of clades could lead to an overestimation of the species number (Lohse, 2009; Papadopoulou et al., 2009). We considered two lines of evidence that could suggest oversplitting occurred. First, oversplitting should primarily affect sister taxa, and we could examine this using our well-supported topology at the species level. Second, oversplitting should be mainly due to missing intermediate haplotypes resulting from unsampled areas within species' geographical ranges. Consequently, we first identified two cases in which sister GMYC species were sampled from different localities (spp 5 + 6 and spp 13 + 14; see Figs. 5 and 6). We then grouped these as two species and checked whether the sister to these clades also occurred in a different locality. This was the case for spp (5 and 6) + 7 and spp (13 and 14) + sp 12. We noted however that even in the extreme case, where GMYC spp 5–7 and spp 12–14 are only two species, the diversity of our data set would be reduced to ten species (see Fig. 5), a number included in the 95% GMYC confidence interval (Table 3).

Another potential cause of oversplitting is the inadvertent sequencing of divergent, homologous copies of mtDNA that have been transposed into the nuclear genome (NUMTs; Lopez et al., 1994; Bensasson et al., 2001). NUMTs have been coamplified with target mtDNA fragments in a variety of organisms, including insects (e.g. Pons and Vogler, 2005; Martins et al., 2007; Moulton et al., 2010). NUMTs are often nonfunctional pseudogenes that can be detected through examination of sequence characteristics such as indels or stop codons, although they can also be challenging to identify (Song et al., 2008). We found no evidence of NUMT coamplification based on a lack of indels or stop codons, and divergent mtDNA sequences such as singletons 2, 4 and 5 were independently recovered as diverging singletons in the nuclear gene fragment as well, highlighting the benefit of simultaneous study of both mt and n markers (Vuataz et al., 2011).

Most of the delimited species (10 out of 14) were sampled from a single restricted area (Fig. 6). Although possibly due to missing localities, this apparent narrow geographical range size of taxa, or microendemism, is particularly frequent in Madagascar (Wilme et al., 2006; Vences et al., 2009). Moreover, many freshwater species are restricted to single or a few river basins (Benstead et al.,

2003), and it was estimated that 20% of the ca. 100 mayfly species described from Madagascar are microendemic (Benstead et al., 2003). The frequent Malagasy species microendemism could thus explain the observed narrow species range sizes and supports the predominant geographical splitting of clades as recovered by the GMYC analyses. Moreover, taxa including widespread species in open lowland savannah (see Section 4.2) and species restricted to small forested areas have already been found in other mayfly genera including *Proboscidoplocia* Demoulin, 1966 (Elouard and Sartori, 1997), *Xyrodromeus* Lugo-Ortiz and McCafferty, 1997 (Gattolliat, 2002), *Cloeodes* Traver, 1938 (Gattolliat, 2001) and *Dabulamanzia* Lugo-Ortiz and McCafferty, 1996 (Gattolliat and Sartori, 2000), supporting our GMYC species delimitation.

With ca. 70% of the putative species potentially confined to a single river or basin, Heptageniidae may be particularly vulnerable to habitat alteration because they cannot colonize other areas (Wilme et al., 2006). Most of the forest streams have been degraded by intensive deforestation (e.g. Harper et al., 2007), making the five putative forest species (i.e. *Afronurus*-like lineages; see Section 4.2) among the most endangered Malagasy Heptageniidae. Putative species 3 from the *Compsoneuria*-like singleton 2 was sampled from only a single forested locality and may also be critically endangered.

### 4.4. Morphological data and taxonomic status

Our results have important consequences for the taxonomy of the subfamily Ecdyonurinae. The genus *Compsoneuria*, described from Southeast Asia, has as junior synonyms *Compsoneuriella* Ulmer, 1939 from the same area, and *Notonurus* Crass, 1947, from South Africa (Gillies, 1984). We now have strong morphological evidence that the latter synonymy is invalid for at least two reasons. Firstly, the Asian *Compsoneuria* clade (Fig. 4) has fimbriate scattered setae on the galea lacinia, whereas all *Compsoneuria*-related species from Africa and Madagascar have simple setae. Secondly, within the Asian *Compsoneuria* clade, Heptageniidae 2 and 3 lack supracoxal spurs on all legs, thus making the genus *Compsoneuria* paraphyletic. The generic name for African species previously placed in *Compsoneuria* should be *Notonurus* Crass, 1947, and thus the following combinations are reinstalled: *Notonurus bequaerti* (Navás, 1930), *Notonurus njalensis* (Kimmins, 1937), *Notonurus sinuosus* (Navás, 1931), and *Notonurus tortinervis* (Navás, 1930). As South African *Notonurus* and the Malagasy lineage share characters of phylogenetic importance, especially in the setation of the maxilla, and no reliable character was found to support the entire Malagasy lineage, the Malagasy species *Compsoneuria josettae* and *Afronurus matitensis* are also provisionally transferred to *Notonurus* as *Notonurus josettae* (Elouard and Sartori, 1996) comb. nov. and *Notonurus matitensis* (Elouard and Sartori, 1996) comb. nov., respectively.

Whereas the Malagasy lineage cannot be morphologically characterized, the six smaller Malagasy lineages, and especially clades 1, 3 and 6, possess strong and reliable features to identify and separate them. Clade 1 is well characterized by the long and lanceolate bristles covering the whole anterior face of femora and the shape of the tarsal claws, while clades 3 and 6 possess short, generally heart-shaped bristles (Fig. 10 in Sartori and Elouard, 1996). The shape of the fore femora, with a pronounced convex dorsal margin, is characteristic of clade 3. Clade 6 was first considered as belonging to *Afronurus* because of the absence of well developed supracoxal spurs; this clade also possesses a very broad head capsule, very long maxillary palps and peculiar distal margin of abdominal tergites (Figs. 3, 8 and 11 respectively in Sartori and Elouard, 1996). Morphological characterization of lineages 2, 4 and 5 is much more preliminary, as they remain singletons at our stage of knowledge. Singleton 2 possesses very peculiar legs with unusual setation and short tibiae and tarsi.

Putative species within clades 1, 3 and 6 are morphologically relatively well supported. The setation of the legs, especially of the dorsal margin of femora and the presence/absence of feathered setae on the ventral margin of tibia and tarsi, and the spination of the distal margin of abdominal tergites seem to be reliable features to morphologically separate the different putative species. Additional material may challenge the consistency of some characters, especially in the case of morphologically very similar species such as sp. 13 and sp. 14, one possible case of oversplitting as discussed in Section 4.3. The two extant described species of Malagasy Heptageniidae morphologically correspond to GMYC species delimited in this study as sp 5 (*Notonurus josettae*) and sp 13 (*N. matitensis*). Moreover, both species were described from Andringitra National Park (Sartori and Elouard, 1996), where sp 5 and sp 13 were also collected in similar environments and within a comparable altitudinal range. The fact that all Malagasy Heptageniidae are monophyletic and constitute a paleo-endemic clade (see Cronk, 1997; Gillespie and Roderick, 2002) implies that the conservation of this unique genetic pool is of prime importance with regard to biodiversity strategies and responsibilities.

## Acknowledgments

Sampling in Madagascar would not have been possible without the help of H. Andriamizehy, T. Ranarilalatiana, H. Ramandimbiarijaona and D. Ottke. We also thank L. S. Rafarasoa from the University of Antananarivo, and J.-C. Renarason, Samy, L. E. Dodo, M. Madiomanana, J.-L. Randrianandrasana and Jérôme for their assistance in the field, and E. & Y. Sieber and M. & S. Rakotolehibe for their welcome in Madagascar. We also thank collectors of the museum specimens: J.-L. Gattolliat, O. Glaizot, P. Derleth, J.-M. Elouard, E. Doumenq, R. Gerecke, T. Goldschmidt, B. Isambert, M. Whiting, C. Robinson. We are grateful to H. Barber-James, F. de Moor, T. Bellingan, P. Sartori and E. Sartori for their help in the field. We thank the Department of Entomology at the University of Antananarivo and the Madagascar Institut pour la Conservation des Ecosystèmes Tropicaux (MICET) for their support. We are grateful to N. Salamin, R. Studer, C. Bernasconi, B. Isambert and T. Barraclough for helpful discussions and assistance with data analysis, and to O. Broennimann for the sampling maps. The laboratory work would not have been possible without C. Stoffel, C. La Mendola, R. Sermier, N. Remollino-Duvoisin, C. Berney, F. Dolivo, C. Aletti, and C. Peter from University of Lausanne. Finally, we are grateful to L. Fumagalli, D. Cherix, and L. Keller at the University of Lausanne and the team at the Museum of Zoology in Lausanne for their support. Permission to collect in Malagasy National Parks and export specimens for research was granted by the Association Nationale pour la Gestion des Aires Protégées (ANGAP) and the Direction Générale de l'Environnement, des Eaux et Forêts, Direction de la Valorisation des Ressources Naturelles, Service de la Protection des Espèces (autorisation de recherche n° 281/07-MEEFT/SG/DPSAP/SSE du 13.11.07; autorisation de sortie n° 181 N-EA11/MG07). Research was funded by the Swiss National Science Foundation (FNS nr. 31003A-116049; www.snf.ch).

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.12.003.

## References


Abouheif, E., Wray, G.A., 2002. Evolution of the gene network underlying wing polyphenism in ants. Science 297, 249–252.

Barber-James, H.M., Gattolliat, J.L., Sartori, M., Hubbard, M.D., 2008. Global diversity of mayflies (Ephemeroptera, Insecta) in freshwater. Hydrobiologia 595, 339–350.

Barber-James, H.M., Sartori, M., Gattolliat, J.L., Webb, J.M., 2010. World checklist of freshwater Ephemeroptera species. World Wide Web electronic publication. <http://fada.biodiversity.be/group/show/35>. [26.09.12].

Bartholome, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing 26, 1959–1977.

Bensasson, D., Zhang, D., Hartl, D.L., Hewitt, G.M., 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecololgy & Evolution 16, 314–321.

Benstead, J.P., Pringle, C.M., 2004. Deforestation alters the resource base and biomass of endemic stream insects in eastern Madagascar. Freshwater Biology 49, 490–501.

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.

Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M.T., Balke, M., Hendrich, L., Geijer, J., Hermann, J., Foster, G.N., Nilsson, A.N., Barraclough, T.G., Vogler, A.P., 2012. The effect of geographical scale of sampling on DNA barcoding. Systematic Biology 61, 851–869.

Braasch, D., Boonsoong, B., 2010. A contribution to the Heptageniidae (Insecta, Ephemeroptera) of Thailand and Malaysia. Zootaxa, pp. 1–26.

Braasch, D., Freitag, H., 2008. *Palawaneuria*, a new subgenus of *Compsoneuria* and new species of *Compsoneuria* and *Afronurus* (Ephemeroptera, Heptageniidae) from Palawan, Philippines. Deutsche Entomologische Zeitschrift 55, 117–128.

Braasch, D., Soldán, T., 1986. Zur Kenntnis der Gattung *Compsoneuria* Eaton 1881, von den Sunda-Inseln (Ephemeroptera, Heptageniidae). Reichenbachia Staatliches Museum für Tierkunde in Dresden 24, 59–62.

Brandley, M.C., Schmitz, A., Reeder, T.W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Systematic Biology 54, 373–390.

Briggs, J.C., 2003. The biogeographic and tectonic history of India. Journal of Biogeography 30, 381–388.

Brittain, J.E., Sartori, M., 2003. Ephemeroptera. In: Resh, V.H., Cardé, R. (Eds.), Encyclopedia of Insects. Academic Press, New York, pp. 373–380.

Clamp, M., Cuff, J., Searle, S.M., Barton, G.J., 2004. The Jalview Java alignment editor. Bioinformatics 20, 426–427.

Cronk, Q.C.B., 1997. Islands: stability, diversity, conservation. Biodiversity and Conservation 6, 477–493.

de Wit, M.J., 2003. Madagascar: heads it's a continent, tails it's an island. Annual Review of Earth and Planetary Sciences 31, 213–248.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7, 214.

Edmunds, G.F., 1979. Biogeographical relationships of the Oriental and Ethiopian mayflies. In: Pasternak, K., Sowa, R. (Eds.), Proceedings of the Second International Conference on Ephemeroptera. Panstwowe Wydawnictwo Naukowe, Warszawa, pp. 11–14.

Elouard, J.-M., Gibon, F.-M., 2003. Ecology of aquatic invertebrates. In: Goodman, S.M., Benstead, J.P. (Eds.), The Natural History of Madagascar. University of Chicago Press, Chicago, pp. 511–517.

Elouard, J.-M., Sartori, M., 1997. *Proboscidoplocia:* a singular plural (Ephemeroptera: Polymitarcyidae: Euthyplociinae). In: Landolt, P., Sartori, M. (Eds.), Ephemeroptera & Plecoptera Biology-Ecology-Systematics. Mauron + Tinguely & Lachat SA, Fribourg, pp. 439–448.

Elouard, J.-M., Sartori, M., Gattolliat, J.-L., Oliarinony, R., 2001. Ordre des Ephéméroptères. In: Elouard, J.M., Gibon, F.M. (Eds.), Biodiversité et Biotypologie des Eaux Continentales de Madagascar. IRD, Paris, pp. 77–112.

Elouard, J.-M., Gattolliat, J.L., Sartori, M., 2003. Ephemeroptera, mayflies. In: Goodman, S.M., Benstead, J.P. (Eds.), The Natural History of Madagascar. University of Chicago Press, Chicago, pp. 639–645.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3, 294–299.

Fontaneto, D., Herniou, E.A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., Barraclough, T.G., 2007. Independently evolving species in asexual bdelloid rotifers. PLoS Biology 5, 914–921.

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.

Gattolliat, J.-L., 2001. The genus *Cloeodes* (Ephemeroptera, Baetidae) in Madagascar. Revue Suisse de Zoologie 108, 387–402.

Gattolliat, J.-L., 2002. Three new Malagasy species of *Xyrodromeus* (Ephemeroptera: Baetidae) with the first generic description of the adults. Revue Suisse de Zoologie 109, 325–341.

Gattolliat, J.-L., Rabeantoandro, S.Z., 2002. The genus Cloeon (Ephemeroptera, Baetidae) in Madagascar. Mitteilungen der Schweizerischen Entomologischen Gesellschaft 195–209 (2001).

Gattolliat, J.-L., Sartori, M., 2000. Contribution to the systematics of the genus *Dabulamanzia* (Ephemeroptera: Baetidae) in Madagascar. Revue Suisse de Zoologie 107, 561–577.

Gibon, F.-M., 2000. Biologie de la conservation et singularité des cours d'eau. L'exemple des Philopotamidae malgaches (Insecta, Trichoptera). In: Lourenço, W.R., Goodman, S.M. (Eds.), Diversité et Endémisme à Madagascar. Société Française de Biogéographie, Paris, pp. 319–330.

Giessler, S., Mader, E., Schwenk, K., 1999. Morphological evolution and genetic differentiation in *Daphnia* species complexes. Journal of Evolutionary Biology 12, 710–723.

Gillespie, R.G., Roderick, G.K., 2002. Arthropods on islands: colonization, speciation, and conservation. Annual Review of Entomology 47, 595–632.

Gillies, M.T., 1984. On the synonymy of Notonurus Crass with *Compsoneuriella* Ulmer (Heptageniidae). In: Landa, V., Soldán, T., Tonner, M. (Eds.), In: Proceeding of the Fourth International Conference on Ephemeroptera. CSAV, Bechyne, pp. 21–25.

Godunko, R.J., Sontag, E., 2004. *Burshtynogena ferenci* gen. and sp. nov. (Ephemeroptera: Heptageniidae) from Eocene Baltic amber. Annales Zoologici Warszawa 54, 515–518.

Goodman, S.M., Benstead, J.P., 2003. The natural history of Madagascar. University of Chicago Press, Chicago, USA.

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

Groombridge, B., Jenkins, M.-B., 2002. World atlas of biodiversity: earth's living resources in the 21st century. University of California Press, Berkeley, USA.

Harper, G.J., Steininger, M.K., Tucker, C.J., Juhn, D., Hawkins, F., 2007. Fifty years of deforestation and forest fragmentation in Madagascar. Environmental Conservation 34, 325–333.

Harrison, A.D., Hynes, H.B.N., 1988. Benthic fauna of Ethiopian mountain streams and rivers. Archiv für Hydrobiologie Supplement 81, 1–36.

Isambert, B., Bergsten, J., Monaghan, M.T., Andriamizehy, H., Ranarilalatiana, T., Ratsimbazafy, M., Andriniainimanana, J.R., Vogler, A.P., 2011. Endemism and evolutionary history in conflict over Madagascar's freshwater conservation priorities. Biological Conservation 144, 1902–1909.

Jobb, G., von Haeseler, A., Strimmer, K., 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evolutionary Biology 4, 18.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Research 33, 511–518.

Kishino, H., Miyata, T., Hasegawa, M., 1990. Maximum-likelihood inference of protein phylogeny and the origin of chloroplasts. Journal of Molecular Evolution 31, 151–160.

Kluge, N.J., 2004. The phylogenetic system of Ephemeroptera. Kluwer Academic Publishers, Dordrecht.

Kohler, F., Glaubrecht, M., 2010. Uncovering an overlooked radiation: molecular phylogeny and biogeography of Madagascar's endemic river snails (Caenogastropoda: Pachychilidae: *Madagasikara* gen. nov.). Biological Journal of the Linnean Society 99, 867–894.

Kumar, S., Skjaeveland, A., Orr, R.J.S., Enger, P., Ruden, T., Mevik, B.H., Burki, F., Botnen, A., Shalchian-Tabrizi, K., 2009. AIR: a batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. BMC Bioinformatics 10.

Lohse, K., 2009. Can mtDNA barcodes be used to delimit species? A response to Pons et al. (2006). Systematic Biology 58, 439–441.

Lopez, J.V., Yuhki, N., Masuda, R., Modi, W., O'Brien, J., 1994. *Numt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. Journal of Molecular Evolution 39, 174–190.

Marshall, D.C., Simon, C., Buckley, T.R., 2006. Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. Systematic Biology 55, 993–1003.

Martins Jr., J., Solomon, S.E., Mikheyev, A.S., Mueller, U.G., Ortiz, A., Bacci Jr, M., 2007. Nuclear mitochondrial-like sequences in ants: evidence from *Atta cephalotes* (Formicidae: Attini). Insect Molecular Biology 16, 777–784.

Masselot, G., Nel, A., 1999. *Pseudokageronia thomasi* gen. nov., sp. nov. from the Upper Miocene of Murat (France) (Ephemeroptera: Heptageniidae). Ephemera 1, 61–73.

McShaffrey, D., McCafferty, W.P., 1988. Feeding behavior of *Rhithrogena pellucida* (Ephemeroptera: Heptageniidae). Journal of the North American Benthological Society 7, 87–99.

McShaffrey, D., McCafferty, W.P., 1990. Feeding behavior and related functional morphology of the mayfly *Ephemerella needhami* (Ephemeroptera: Ephemerellidae). Journal of Insect Behavior 3, 673–688.

Monaghan, M.T., Gattolliat, J.L., Sartori, M., Elouard, J.M., James, H., Derleth, P., Glaizot, O., de Moor, F., Vogler, A.P., 2005. Trans-oceanic and endemic origins of the small minnow mayflies (Ephemeroptera, Baetidae) of Madagascar. Proceedings of the Royal Society of London Series B: Biological Sciences 272, 1829–1836.

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.

Moulton, M.J., Song, H., Whiting, M.F., 2010. Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from Orthoptera (Arthropoda: Insecta). Molecular Ecology Resources 10, 615–627.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Nagy, Z.T., Joger, U., Wink, M., Glaw, F., Vences, M., 2003. Multiple colonization of Madagascar and Socotra by colubrid snakes: evidence from nuclear and mitochondrial gene phylogenies. Proceedings of the Royal Society of London Series B: Biological Sciences 270, 2613–2621.
Nobre, T., Eggleton, P., Aanen, D.K., 2010. Vertical transmission as the key to the colonization of Madagascar by fungus-growing termites? Proceedings of the Royal Society of London Series B: Biological Sciences 277, 359–365.
Nylander, J.A.A., 2004. MrAIC.pl. Program distributed by the author. Uppsala, Sweden: Evolutionary Biology Centre, Uppsala University.
Ogden, T.H., Whiting, M.F., 2003. The problem with "the Paleoptera Problem:" sense and sensitivity. Cladistics 19, 432–442.
Orsini, L., Koivulehto, H., Hanski, I., 2007. Molecular evolution and radiation of dung beetles in Madagascar. Cladistics 23, 145–168.
Page, R.D.M., 1996. TreeView: an application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12, 357–358.
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 (2009). Systematic Biology 58, 442–444.
Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Molecular Biology and Evolution 27, 1659–1672.
Pons, J., Vogler, A.P., 2005. Complex pattern of coalescence and fast evolution of a mitochondrial rRNA pseudogene in a recent radiation of tiger beetles. Molecular Biology and Evolution. 22, 991–1000.
Pons, J., Bruvo, B., Petitpierre, E., Plohl, M., Ugarkovic, D., Juan, C., 2004. Complex structural features of satellite DNA sequences in the genus *Pimelia* (Coleoptera: Tenebrionidae): random differential amplification from a common 'satellite DNA library'. Heredity 92, 418–427.
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.
Posada, D., 2004. Collapse: describing haplotypes from sequence alignments. Computation Evolutionary Biology Lab, University of Vigo, Vigo, Spain.
Poux, C., Madsen, O., Marquard, E., Vieites, D.R., de Jong, W.W., Vences, M., 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. Systematic Biology 54, 719–730.
R Development Core Team, 2011. R: a language and and environment for statistical computing. R Foundation for Statisitcal Computing, Vienna, Austria.
Rabinowitz, P.D., Coffin, M.F., Falvey, D., 1983. The separation of Madagascar and Africa. Science 220, 67–69.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Sartori, M., Elouard, J.-M., 1996. New Heptageniidae (Insecta: Ephemeroptera) from the Réserve Naturelle Intégrale d'Andringitra, Madagascar. In: Goodman, S. (Ed.), A Floral and Faunal Investory of the Eastern Slopes of the Réserve Naturelle d'Andringitra, Madagascar: with Reference to Elevational Variation. Field Museum of Natural History, Chicago, pp. 121–130.
Sartori, M., Gattolliat, J.-L., Oliarinony, R., Elouard, J.-M., 2000. Biogeography of Malagasy mayflies (Insecta, Ephemeroptera): preliminary results. In: Lourenço, W.R., Goodman, S.M. (Eds.), Diversité et Endémisme à Madagascar. Société Française de Biogéographie, Paris, pp. 307–317.
Sartori, M., Peters, J.G., Hubbard, M.D., 2008. A revision of Oriental Teloganodidae (Insecta, Ephemeroptera, Ephemerelloidea). Zootaxa 1957, 1–51.
Schoonbee, H.J., 1967. A new record of *Compsoneuriella njalensis* (Kimmins) (Ephemeroptera: Heptageniidae) from South Africa. Journal of the Entomological Society of South Africa 29, 151–156.
Schoonbee, H.J., 1973. The role of ecology in the species evaluation of the genus *Afronurus* Lestage (Heptageniidae) in South Africa. In: Peters, W.L., Peters, J.G. (Eds.), Proceedings of the First International Conference on Ephemeroptera. Brill, E.J., Leiden, pp. 88–113.
Shapiro, B., Rambaut, A., Drummond, A.J., 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Molecular Biology and Evolution 23, 7–9.
Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51, 492–508.
Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16, 1114–1116.
Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87, 651–701.
Sinitshenkova, N.D., 2000. New Jersey amber mayflies: the first North American Mesozoic members of the order (Insecta; Ephemeroptera). In: Grimaldi, D. (Ed.), Studies on Fossil in Amber, with Particular Reference to the Cretaceous of New Jersey. Backhuys Publishers, Leiden, pp. 111–125.
Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proceedings of the National Academy of Sciences of the United States of America 105, 13486–13491.
Takemon, Y., Yamamoto, A., Nakashima, M., Tanida, K., Kishi, M., Kato, M., 2006. Isolation of sperm vesicles from adult male mayflies and other insects to prepare high molecular weight genomic DNA samples. Molecular Biology Reports 33, 65–70.
Torres, E., Lees, D.C., Vane-Wright, R.I., Kremen, C., Leonard, J.A., Wayne, R.K., 2001. Examining monophyly in a large radiation of Madagascan butterflies (Lepidoptera: Satyrinae: Mycalesina) based on mitochondrial DNA data. Molecular Phylogenetics and Evolution 20, 460–473.
Vences, M., Vieites, D.R., Glaw, F., Brinkmann, H., Kosuch, J., Veith, M., Meyer, A., 2003. Multiple overseas dispersal in amphibians. Proceedings of the Royal Society of London Series B: Biological Sciences 270, 2435–2442.
Vences, M., Wollenberg, K.C., Vieites, D.R., Lees, D.C., 2009. Madagascar as a model region of species diversification. Trends in Ecology & Evolution 24, 456–465.
Vogler, A.P., Monaghan, M.T., 2007. Recent advances in DNA taxonomy. Journal of Zoological Systematics and Evolutionary Research 45, 1–10.
Vuataz, L., Sartori, M., Wagner, A., Monaghan, M.T., 2011. Toward a DNA taxonomy of Alpine *Rhithrogena* (Ephemeroptera: Heptageniidae) using a mixed Yule-coalescent analysis of mitochondrial and nuclear DNA. PLoS ONE 6, e19728.
Wang, T.Q., McCafferty, W.P., 2004. Heptageniidae (Ephemeroptera) of the world. Part I: phylogenetic higher classification. Transactions of the American Entomological Society 130, 11–45.
Ward, P.S., Downie, D.A., 2005. The ant subfamily Pseudomyrmecinae (Hymenoptera: Formicidae): phylogeny and evolution of big-eyed arboreal ants. Systematic Entomology 30, 310–335.
Webb, J.M., McCafferty, W.P., 2008. Heptageniidae of the world. Part II: key to the Genera. Canadian Journal of Arthropod Identification 7, 1–55.
Webb, J.M., Braasch, D., McCafferty, W.P., 2006. Reevaluation of the genera *Compsoneuria* Eaton and *Trichogenia* Braasch & Soldán (Ephemeroptera: Heptageniidae). Zootaxa 1335, 55–68.
Wernersson, R., Pedersen, A.G., 2003. RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. Nucleic Acids Research 31, 3537–3539.
Wilme, L., Goodman, S.M., Ganzhorn, J.U., 2006. Biogeographic evolution of Madagascar's microendemic biota. Science 312, 1063–1065.
Wirta, H., Orsini, L., Hanski, I., 2008. An old adaptive radiation of forest dung beetles in Madagascar. Molecular Phylogenetics and Evolution 47, 1076–1089.
Wirta, H., Viljanen, H., Orsini, L., Montreuil, O., Hanski, I., 2010. Three parallel radiations of Canthonini dung beetles in Madagascar. Molecular Phylogenetics and Evolution 57, 710–727.
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.
Zakharov, E.V., Smith, C.R., Lees, D.C., Cameron, A., Vane-Wright, R.I., Sperling, F.A.H., 2004. Independent gene phylogenies and morphology demonstrate a malagasy origin for a wide-ranging group of swallowtail butterflies. Evolution 58, 2763–2782.