Evolution and island endemism of morphologically cryptic *Baetis* and *Cloeon* species (Ephemeroptera, Baetidae) on the Canary Islands and Madeira

SEREINA RUTSCHMANN*, JEAN-LUC GATTOLLIAT[†], SAMANTHA J. HUGHES[‡], MARCOS BÁEZ[§],
MICHEL SARTORI[†,¶a] AND MICHAEL T. MONAGHAN*[a]

Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany
[†]*Musée cantonal de zoologie, Palais de Rumine, Lausanne, Switzerland*
[‡]*CITAB, Gabinete H.1.4, Departamento de Ciências Florestais e Arq. Paisagista, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal*
[§]*Departamento de Biología Animal, Universidad de La Laguna, Tenerife, Spain*
[¶]*Biozentrum Grindel und Zoologisches Museum, Universität Hamburg, Hamburg, Germany*

SUMMARY

1. The Canary Islands and Madeira are reportedly home to seven recognised species of baetid mayflies (Ephemeroptera, Baetidae), two of which also occur on the European mainland. Their species status remains unsure, and loss of habitat suggests they are of conservation concern.

2. We applied morphological characters and a general mixed Yule-coalescent (gmyc) model analysis of the cytochrome *c* oxidase subunit 1 (*cox1*) gene to delineate putative species within morphologically cryptic species groups *Baetis* (*Rhodobaetis*) and *Cloeon dipterum* s.l. We used a three-gene mitochondrial data set (1450 base pairs) to infer phylogenetic relationships and a molecular clock calibrated using island geological ages to infer colonisation history.

3. Genetic and morphological evidence indicated the presence of 12 putative species, 11 of which were endemic to the islands. Only *Baetis atlanticus*, on Madeira, also occurs on the European mainland. Two lineages (*B. pseudorhodani* s.l. and *B. canariensis* s.l.) appear to have arisen in the past 15 million years (mya) and diversified in parallel throughout the Canary Islands. Within the canariensis lineage, sister species occur on the island of Gran Canaria and in North Africa.

4. Pronounced island endemism contradicts previous taxonomic work, which reported a depauperate fauna that included several mainland species. Recent diversification among islands and a close link to North Africa suggest a complex evolutionary history. Owing to their small population size and ongoing habitat alteration, several of these island endemics are among the most endangered aquatic insects in Europe.

Keywords: general mixed Yule-coalescent (gmyc) model, island colonisation, Macaronesia, mayfly, mitochondrial phylogeny

Introduction

Oceanic islands are sometimes referred to as 'natural laboratories' because their isolation and distinct biogeographical conditions (e.g. size, age and abiotic factors) can result in the formation of a unique fauna (Emerson, 2008; Losos & Ricklefs, 2009). Gillespie & Roderick (2002) noted that adaptive radiations in arthropods have occurred on the Canary Islands, where a few colonists have diverged into a number of neo-endemics. While much research has been carried out on island evolution and endemism of terrestrial organisms, there has been comparatively little work on aquatic invertebrates (Ribera, Bilton & Vogler, 2003b; Ribera *et al.*, 2003a). This

Correspondence: Sereina Rutschmann, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 301, 12587 Berlin, Germany. E-mail: sereina.rutschmann@igb-berlin.de
[a]These two authors are the senior authors of the paper

is despite the disproportional contribution they make to biodiversity given the relatively small extent of their habitat (Dijkstra, Monaghan & Pauls, 2014).

Genetic studies have identified several colonisation pathways on the Canaries and Madeira (Juan *et al.*, 2000; Emerson, 2002), including a single colonisation event followed by stepping-stone dispersal (Juan, Oromí & Hewitt, 1997; Emerson & Oromí, 2005; Arnedo *et al.*, 2008) or multiple independent colonisation events (Nogales *et al.*, 1998; Ribera *et al.*, 2003a,b). Several studies have proposed an African origin for the fauna, both for the relatively close Canaries as well as the more distant Madeira (Brunton & Hurst, 1998; Weingartner, Wahlberg & Nylin, 2006). Islands are generally considered to be the end points of colonisation events, but in some cases acted as sources rather than sinks of populations (Bellemain & Ricklefs, 2008). Recent studies reported island-to-continent colonisation by land snails (*Theba*, Greve *et al.*, 2010) and birds (*Cyanistes*, Illera *et al.*, 2011) from the Canary Islands to the North African mainland.

The short winged adult stage (often < one day) and strict fidelity as larvae to freshwater habitats of mayflies (Ephemeroptera) suggest fairly limited dispersal abilities, making them good candidates for the study of biogeographical processes (Sartori, 2001). There is phylogenetic evidence for mayfly dispersal between Madagascar and continental Africa (Monaghan *et al.*, 2005), and two mayfly species were recently recorded from the volcanic island of La Réunion, > 700 km from any potential source of colonisation (Gattolliat, 2004). Little is known of their biogeographical history on the

Canary Islands and Madeira. Based on the extensive taxonomic treatment of Canary Island species (Müller-Liebenau, 1971; Alba-Tercedor, Báez & Soldán, 1987) and a recent revision of Madeiran species (Gattolliat *et al.*, 2008), the islands comprise seven recognised species of Baetidae (Table 1).

It was originally thought that two of the species on the islands were identical to the continental species (Eaton, 1885; Navás, 1906; Brinck & Scherer, 1961): *Baetis rhodani* (Pictet, 1843) and *Cloeon dipterum* L. 1761. These are two of the most common and abundant mayfly species in the western Palaearctic (Sowa, 1975; Gattolliat & Sartori, 2008). A number of studies have reported the presence of multiple mitochondrial DNA (mtDNA) lineages within *B. rhodani* that may indicate the presence of cryptic species (Williams, Ormerod & Bruford, 2006; Lucentini *et al.*, 2011). Müller-Liebenau (1971) demonstrated that the supposed presence of *B. rhodani* in the Canary Islands was erroneous and that two new species were present, *B. canariensis* Müller-Liebenau, 1971 and *B. pseudorhodani* Müller-Liebenau, 1971. *Cloeon dipterum* was thought to be the only mayfly species occurring on both archipelagos (Brinck & Scherer, 1961; Müller-Liebenau, 1971; Alba-Tercedor *et al.*, 1987); however, Gattolliat *et al.* (2008) recognised populations on Madeira to be an endemic species, challenging the validity of the species name on any of the islands.

Here, we examined the diversity and evolutionary history of a cryptic group of mayflies found throughout the Canary Islands and Madeira by combining DNA taxonomy and phylogenetic analyses (*sensu* Vuataz *et al.*,

Table 1 Recognised Baetidae species groups on the Canary Islands and Madeira prior to this study, including geographical distribution, primary habitat type (lotic = perennial running water; lentic = permanent or temporary standing water), historical local abundance and number of gmyc species detected in this study with 95% confidence intervals in brackets []

Species group	Distribution	Habitat	Abundance	Gmyc species in this study [95% CI]
BAETIDAE				
Cloeon dipterum (Linnaeus, 1761)*	Azores, Canaries, Palaearctic	Lentic	High	4 [4]
Cloeon peregrinator Gattolliat & Sartori, 2008[†]	Madeira	Lentic	High	
Baetis atlanticus Soldán & Godunko, 2006[‡]	Madeira	Lotic	High	1 [1]
Baetis enigmaticus Gattolliat & Sartori, 2008[§]	Madeira	Lotic	Low	1 [1]
Baetis canariensis Müller-Liebenau, 1971	Canaries	Lotic	High	4 [3–4]
Baetis pseudorhodani Müller-Liebenau, 1971	Canaries	Lotic	Low	3 [3]
Baetis nigrescens Navás, 1932[¶]	Canaries, Iberia, North Africa	Lotic	High	

*Including *Cloeon cognatum* Stephens, 1835 (Alba-Tercedor *et al.*, 1987).
[†]*Cloeon dipterum sensu* Brinck & Scherer (1961).
[‡]*Baetis rhodani* (Pictet, 1843) *sensu* Brinck & Scherer (1961).
[§]*Baetis pseudorhodani sensu* Stauder (1991, 1995); Hughes *et al.* (1998).
[¶]*Baetis nigrescens* was last recorded by Malmqvist *et al.* (1995) and not found in the present study.
[‖]Individal with no *cox1* data.

2013). Putative species were delineated using a gmyc model (Fujisawa & Barraclough, 2013) analysis of *cox1* barcodes. A phylogenetic analysis was then carried out using three mtDNA fragments in order to reconstruct their diversification on the islands. Based on more recent understanding of dispersal and evolution in mayflies, combined with use of new molecular approaches, we predicted higher overall diversity and more pronounced island endemism than previously reported, with important implications for island biodiversity conservation.

Methods

Sampling and material

Baetis and *Cloeon* individuals were collected from larval aquatic habitats at eight sites on four of the Canary Islands and four sites on Madeira between 2006 and 2009 using kick-samples (Fig. 1; see Table S1). The Canary Islands of Fuerteventura, Lanzarote and El Hierro were excluded from our study because they lack perennial running waters necessary for the establishment of *Baetis* populations. Samples were preserved in 99% ethanol in the field and stored at 4 °C until analysis. Individuals were identified to species-group level according to current morphological knowledge (Müller-Liebenau, 1971; Gattolliat & Sartori, 2008; Gattolliat *et al.*, 2008; Bauernfeind & Soldán, 2012). DNA was extracted from a total of 64 individuals for genetic analysis (GenBank accession numbers KF438094, KF438096–KF438103,

KF438105–KF438106, KF438108–KF438123, KF438126–KF438130, KF438133–KF438167, KJ631625–KJ631639; Appendix S1). These were chosen to represent all sampled islands, the broadest possible range of sample sites and morphologies. In addition to the specimens from the Canary Islands and Madeira, we included Corsican *Baetis* (*Rhodobaetis*) *ingridae* Thomas & Soldán, 1987; topotype specimens of *B. rhodani* from Switzerland (Gattolliat & Sartori, 2008), as well as *Baetis rhodani* and *Cloeon dipterum* from central and eastern Europe (Germany, Switzerland, Latvia, Lithuania and Slovakia) and from neighbouring mainland areas (Spain, Morocco and Tunisia).

DNA extraction, amplification and sequencing

Genomic DNA was extracted from whole specimens or from legs using the Wizard SV 96 (Promega, Southampton, U.K.) or NucleoSpin® 96 (Macherey-Nagel, Düren, Germany) tissue kits. We sequenced three fragments of mtDNA: the 5' end of *cox1* commonly used in DNA barcoding, cytochrome oxidase *b* (*cob*) and the large ribosomal subunit *16S* (*rrnL*). The gene fragments were amplified using standard polymerase chain reactions (PCR) protocols with the following primer pairs: LCO1490 + HCO2198 (*cox1*; Folmer *et al.*, 1994), cb3 + cb4 (*cob*; Monaghan *et al.*, 2009) and 16Sar (*rrnL*; Simon *et al.*, 1994) + 16Sb2 (*rrnL*; Giessler, Mader & Schwenk, 1999). All amplification products were purified using the Wizard SV 96 (Promega) standard protocol and used as template for cycle sequencing reactions with the Big-Dye® Terminator v.1.1 or v.3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, U.S.A.). Forward and reverse sequences were analysed with a 3730 or 3500 xL capillary sequencer (Applied Biosystems) and assembled and edited using CODONCODE ALIGNER v.3.5.6 (CodonCode Corporation, Dedham, U.S.A.).

DNA taxonomy and gmyc models

Genetic species delineation was carried out by combining newly sequenced *cox1* data (658 base pairs) with 131 published sequences (Appendix S1). The damselfly *Euphaea formosa* (NC014493) was used as an outgroup. Multiple sequence alignments were made using MAFFT v.7.050b (Katoh & Standley, 2013) and checked for stop codons using MESQUITE v.2.75 (Maddison & Maddison, 2011). The matrix was collapsed by removing identical haplotypes using the PERL script collapsetypes_v4.5.pl (Chesters, 2013). The best-fit models of molecular evolution (Appendix S1) were estimated separately for each
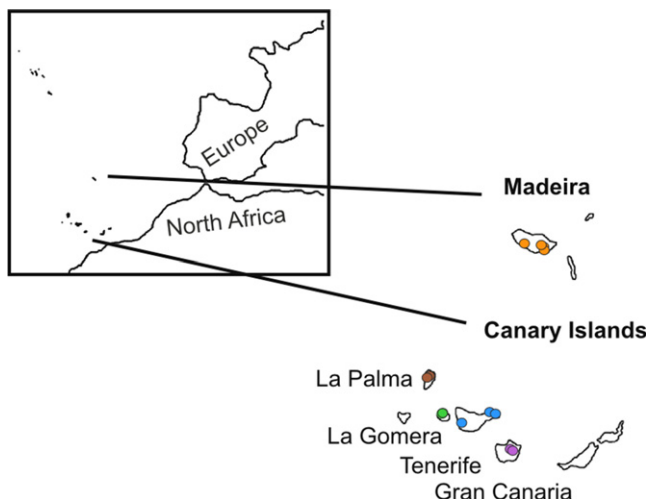

Fig. 1 Map of the sampling sites on the Canary Islands and Madeira. Sites are indicated by filled circles. Each colour corresponds to a different island, and the same colours are used in Figs 2, 3 and Appendix S2. Details of all sampling sites are given in Table S1.

codon position using JMODELTEST v.2.1 (Guindon & Gasc-uel, 2003; Darriba *et al.*, 2012). An ultrametric gene tree was reconstructed using a relaxed (uncorrelated lognor-mal) molecular clock and a coalescent prior *sensu* Mona-ghan *et al.* (2009) in BEAST v.1.8.0 (Drummond *et al.*, 2012). We conducted six independent runs of 20 million generations each. Runs were combined in LOGCOMBINER v.1.8.0 (Drummond *et al.*, 2012), removing 10% of each run as burn-in, whereby all parameters reached effective sample sizes (ESS) > 800. Maximum clade credibility trees were obtained using TREEANNOTATOR v.1.8.0 (Drum-mond *et al.*, 2012). Single-threshold gmyc analysis was conducted using the Splits package (http://r-forge.r-project.org/projects/splits/) for R (R Development Core Team, 2012).

Phylogenetics and molecular dating

Phylogenetic analyses were performed using representa-tives of putative species based on the gmyc analysis. Two species were added (*Baetis rhodani* from Tunisia and *Clo-eon peregrinator* from Madeira) for which *cox1* sequences were unavailable and therefore could not be included in the gmyc analysis. We included only one *Cloeon dipterum* s.l. specimen from the European mainland and added two *cob* sequences for *Baetis rhodani* s.l. and *Cloeon peregrinator* from GenBank (Appendix S1). Sequences were aligned and checked for stop codons as above. In order to accommodate variable substitution rates among different genes and codon positions, we used different evolutionary models for each (Appendix S1) in a partitioned analysis (*sensu* Brandley, Schmitz & Reeder, 2005). Bayesian phylogenetic trees were reconstructed with MRBAYES v.3.2.1 (Ronquist *et al.*, 2012). Nucleotide frequencies, gamma distributions, substitution rates and the proportion of invariant sites were unlinked across partitions. We ran two independent analyses of four MCMC chains, each with ten million generations. The first 25% of each run was discarded as burn-in.

We calculated the age of island lineages using molecu-lar dating of the phylogeny in two different ways. First, we used the age of the oldest Canarian island for which we have samples (Gran Canaria: 14.5 mya; Normal prior; Carracedo *et al.*, 1998) to constrain the roots of the two clades with species on the Canaries and Madeira. This geological calibration analysis used a relaxed (log-normal) molecular clock. In a second analysis, we used a fixed evolutionary rate of 2.69% per million years (Papadopoulou, Anastasiou & Vogler, 2010) and a strict clock. For both analyses, we constrained the separation

of Ephemeroptera and Odonata at 388 mya (fossil cali-bration; Normal prior; 2.5% quantile: 441 mya, 97.5% quantile: 335 mya) following Rehm *et al.* (2011) and used a birth-death process as tree prior (Gernhard, 2008). We performed both calibrations using BEAST 1.8.0 (see above) with seven independent runs of 20 million generations each. Run replicates were analysed and combined as described above, whereby all parameters reached ESS > 800.

Results

Species delineation

There were 184 unique *cox1* haplotypes. The gmyc model was significant (χ^2: 43.5, $P < 0.001$) and delin-eated 33 putative species in the entire *cox1* dataset, com-posed of 21 distinct clusters and 12 singletons (Fig. 2). The 95% confidence interval (CI, defined as 2 log likeli-hood units) ranged from 30 to 35 species. Eleven species were recovered on the Canaries and Madeira, most of which were confined to single islands. Ten of these spe-cies were recognised by the gmyc model even using the most conservative estimate of 30 species in total (based on the lower 95% CI; Table 1).

Four putative gmyc species were delineated in *Cloeon*, including two European mainland species and two spe-cies on Tenerife and Gran Canaria (Fig. 2). *Baetis pseudo-rhodani* s.l. was three gmyc species, one each on La Gomera, Tenerife and Gran Canaria. The morphologi-cally similar *B. enigmaticus* from Madeira was a distinct gmyc species within this clade (Fig. 2). *Baetis canariensis* s.l. clustered into four gmyc species, one each on La Palma, La Gomera, Tenerife and Gran Canaria. All *B. at-lanticus* haplotypes from Madeira were recovered within a larger group composed of *B. rhodani* s.l. from Morocco, Spain and central Europe. One Spanish individual and all specimens of the *B. rhodani* topotype from Switzer-land formed another gmyc species.

Phylogenetics and molecular dating

In the three-gene phylogenetic reconstruction, *Cloeon* was monophyletic, and the species from Gran Canaria was a sister taxon to all others, including Madeiran *C. peregrinator* (Fig. 3, node A). Other relationships within the clade were unresolved. *Rhodobaetis* was monophyletic and composed of two large clades, both of which included Madeiran species. The first clade (Fig. 3, nodes B and C) included *B. ingridae*, the Madeiran *B. at-lanticus* and several gmyc species of *B. rhodani* s.l.,
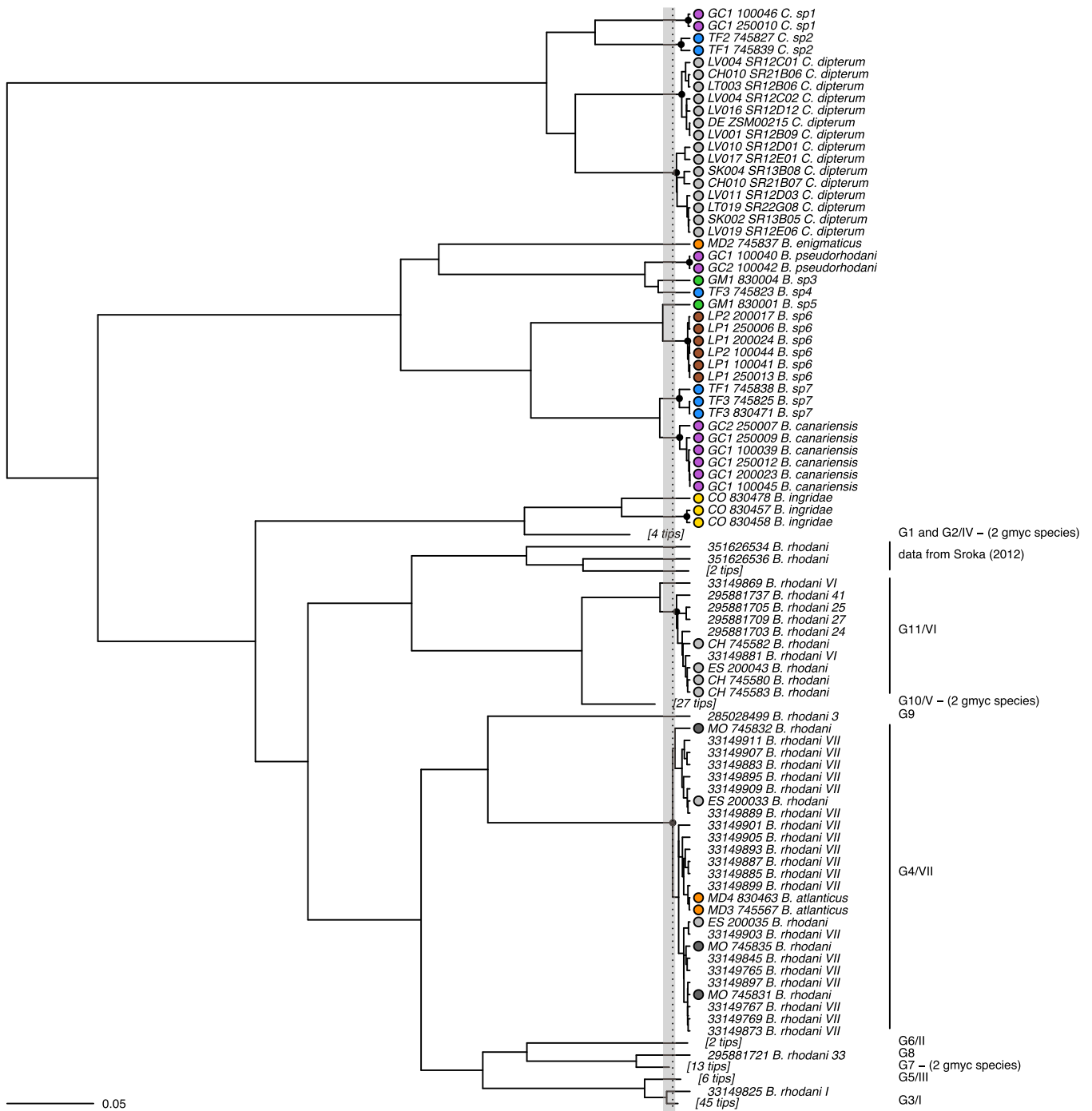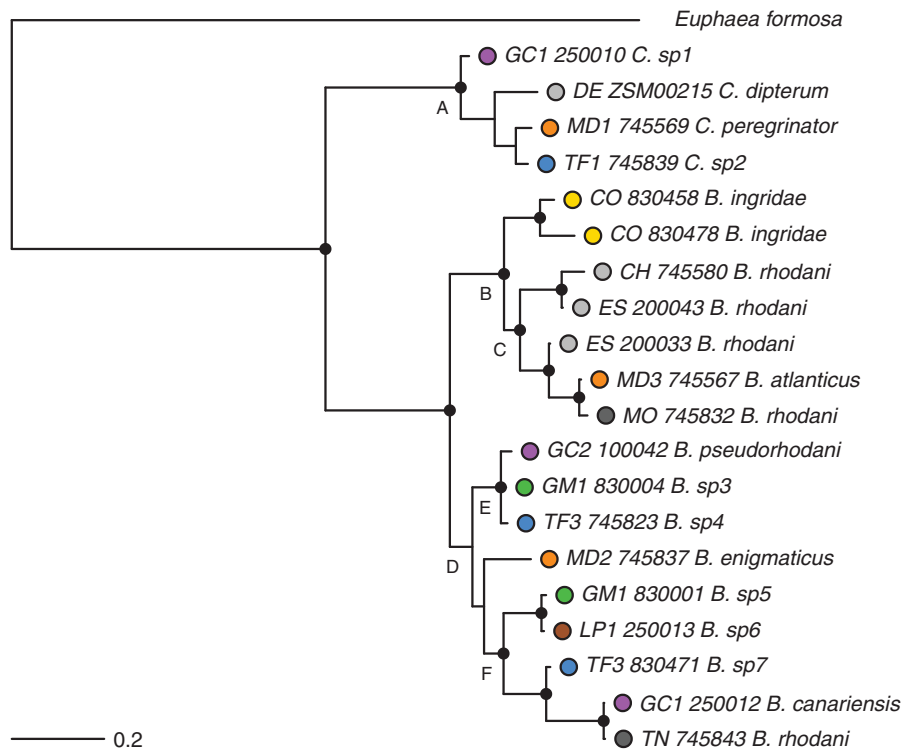
Fig. 2 The *cox1* gene tree used as input for the gmyc model analysis of Canarian, Madeiran, North African and European mayflies (representatives of the *Rhodobaetis* group and *Cloeon dipterum* s.l.). Sequence clusters corresponding to single gmyc species are indicated by open circles at subtending nodes. Where multiple gmyc groups were collapsed for clear display, these are indicated by text in brackets. Terminal labels indicate sampling sites (see Table S1), sample number and morphological assignment. Circles to the left of terminal labels indicate haplotype origin, as Canarian or Madeiran island (coloured according to Fig. 1), European mainland (light grey), North Africa (dark grey) and Corsica (yellow). All other haplotypes are European in origin (see Appendix S1). Putative gmyc species are indicated with *sp1-sp7*. The dotted line indicates the point of maximum-likelihood fit of the single-threshold gmyc model. The grey shading corresponds to the 95% confidence intervals of the gmyc analysis. Scale bar indicates substitutions per site.

Fig. 3 Bayesian inference reconstruction of the phylogenetic relationships among Canarian, Madeiran, North African and European mayflies (representatives of the *Rhodobaetis* group and *Cloeon dipterum* s.l.) based on a concatenated mitochondrial DNA data set (*cob, cox1, rrnL*) using a separate substitution model for each gene. Filled circles indicate well-supported nodes (Bayesian posterior probability ≥ 0.95). Node labels (A–F) are referred to in the text. Putative gmyc species are indicated with *sp1-sp7*. Scale bar indicates substitutions per site. Circles to the left of terminal labels indicate the origin of individuals (see Fig. 2).

including the topotype of *B. rhodani* s.s. from Switzerland. All of the nodes in this clade were well supported (Bayesian posterior probability (BPP) ≥ 0.95). The second clade (Fig. 3, node D) encompassed the three *B. pseudorhodani* s.l. (Fig. 3, node E) gmyc species, *B. enigmaticus* from Madeira, the four *B. canariensis* s.l. gmyc species and *B. rhodani* s.l. from Tunisia (Fig. 3, node F). Monophyly of the three *B. pseudorhodani* s.l. gmyc species from the Canary Islands and the relationships within the *B. canariensis* subclade were all strongly supported, but the position of the Madeiran *B. enigmaticus* remained unclear. The Tunisian *B. rhodani* s.l. was most closely related to the Gran Canaria *B. canariensis* species.

According to the molecular clock analysis using the geological calibration, all lineages from the Canaries and Madeira arose between 15 mya and less than one mya (Appendix S2). The origin of the *B. rhodani* s.l. clade was 23 mya (95% highest posterior density (HPD) interval: 24-4 mya) (Fig. 3, Appendix S2 node C). The separation times inferred using the fixed rate of divergence were somewhat more recent than inferred from the geologically calibrated tree, but fell within the CI. The strict clock recovered a recent origin of the *Cloeon* clade (Fig. 3, Appendix S2, node A: 4 mya; 95% HPD interval: 5-3 mya) and the *B. pseudorhodani* and *B. canariensis* clade (node D: 8 mya; 95% HPD interval: 8-5 mya). The tree resulting from the molecular dating analysis using geographical calibration was identical to the Bayesian

phylogenetic reconstruction. The tree resulting from the fixed divergence rate showed minor differences: *Cloeon dipterum* from the European mainland was basal to other *Cloeon*, and *B. enigmaticus* was a sister taxon to *B. pseudorhodani* (Appendix S2).

Discussion

Species delineation

The delineation of putative species is an important first step in evolutionary biogeographical analysis to ensure adequate sampling of the phylogenetic trees. The gmyc approach using mtDNA has been shown to be an appropriate method of delineating putative species in mayflies as corroborated by nuclear markers and morphological characters (Monaghan *et al.*, 2009; Vuataz *et al.*, 2011, 2013; Pereira-da-Conceicoa *et al.*, 2012). In contrast to DNA-barcoding approaches, it does not require an *a priori* threshold (e.g. 3% difference) or calibration of divergence using congeneric species. Our own morphological data (not shown) also provide strong evidence that the gmyc species here correspond to biological species. Island populations of *B. pseudorhodani* and *B. canariensis* could be reliably separated using the list of 26 morphological characters compiled by Godunko, Prokopov & Soldán (2004). Within the *B. canariensis* clade, the relative length of the gills, the shape of the labial palp and the

ornamentation of the tergites are the most promising characters for separating the different gmyc species of the four Canary Islands (author J.-L.G. & M.S., unpubl. data). Geographical criteria (*sensu* DeSalle, Egan & Siddall, 2005) also corroborate the status of two of the *B. canariensis* gmyc species: one was found only on La Palma, and one was found only on Gran Canaria. Sample sizes for the other gmyc species were too low to use such a criterion.

The recovery of 12 putative species on the Canary Islands and Madeira (11 gmyc species and *C. peregrinator* on Madeira) nearly doubles the mayfly diversity reported previously for these islands (Brinck & Scherer, 1961; Müller-Liebenau, 1971; Alba-Tercedor *et al.*, 1987; Gattolliat *et al.*, 2008). Importantly, most of this diversity occurs as island endemics. These findings are in agreement with the high rate of island endemism in terrestrial invertebrates (e.g. Juan *et al.*, 2000; Emerson & Oromí, 2005). In our analysis, *B. atlanticus* occurred in Spain and Morocco and corresponded to haplogroup VII of Williams *et al.* (2006), which was reported from Ireland, Scotland, England, Denmark and Switzerland. *Baetis atlanticus* is characterised by small but consistent differences from the topotype of *B. rhodani* s.s. (Gattolliat & Sartori, 2008; Gattolliat *et al.*, 2008). The latter matched haplogroup VI of Williams *et al.* (2006). *Baetis pseudorhodani* and *B. canariensis* were described from the Canary Islands, but the type locality of both species is located on Gran Canaria (Müller-Liebenau, 1971). The gmyc species reported here (Figs 2 and 3, *sp1-sp7*) will be formally described elsewhere, and the specific names should be circumscribed to the populations of Gran Canaria, with new names proposed for species on other Canary Islands.

The occurrence of multiple *Cloeon* gmyc species on the Canary Islands was somewhat unexpected. This is because *Cloeon* species are often widespread, and studies from the Indian Ocean report both historical and relatively recent over-ocean dispersal (Gattolliat, 2004; Monaghan *et al.*, 2005). *Cloeon* species live in standing water (e.g. ponds) and appear to fit the repeatedly observed pattern that standing water species are more dispersive than running water species (reviewed in Ribera, 2008). Nonetheless, Gran Canaria and Tenerife were home to two distinct gmyc species, and the *C. dipterum* s.l. species on the islands generally show morphological differences and differentiate slightly from their mainland counterparts (Alba-Tercedor *et al.*, 1987).

The presence of cryptic species within Baetidae was not surprising based on previous studies. Williams *et al.* (2006) first reported the presence of divergent haplogroups in *B. rhodani* s.l., and several more recent studies have detected potentially cryptic species in *Baetis* using molecular markers (Savolainen *et al.*, 2007; Ståhls & Savolainen, 2008; Lucentini *et al.*, 2011; Pereira-da-Conceicoa *et al.*, 2012; Sroka 2012). Here, we add morphological and geographical evidence that undescribed species occur within *Baetis* (Rhodobaetis). Multiple gmyc species within *B. ingridae* and *C. dipterum* may also indicate cryptic species. While we could establish support for species within what was thought to be *B. pseudorhodani* and *B. canariensis*, the status of additional *B. rhodani* s.l. gmyc species on the European mainland requires additional data (nuclear DNA, morphological, ecological or geographical criteria). An important consideration is that single-locus methods may lead to an overestimate of putative species numbers (i.e. 'over-splitting') if too few populations are sampled (Lohse, 2009; Papadopoulou *et al.*, 2009). However, the lack of suitable (freshwater) habitat between the islands and the mainland suggests there are no unsampled intermediate populations. Another concern is the potential amplification of nuclear mitochondrial pseudogenes ('numts', Lopez *et al.*, 1994). Although all sequences were checked for stop codons, we have no knowledge of numts in Baetidae, and thus, an important next step is to corroborate our findings with multiple nuclear markers where possible.

Colonisation history

Our analysis of *Cloeon* and *Baetis* (Rhodobaetis) mayflies revealed a distinct fauna on the Canaries and Madeira that includes members of four well-supported lineages. Three of these lineages contained European or African species, and the phylogenetic reconstruction suggests that the three lineages of *Baetis* (Rhodobaetis) all have different colonisation histories. Madeira appears to have been colonised twice, with closest links to Europe and North Africa (*B. atlanticus*) and to the Canary Islands (*B. enigmaticus*). *Baetis atlanticus* is widespread and abundant in Madeira and is common on the mainland, suggesting a recent colonisation of Madeira from a continental source. We hypothesise that this species arrived when *B. enigmaticus* was already present as the only *Baetis* species in Madeiran running waters. If *B. enigmaticus* is unable to compete with the new immigrant, this could explain its very restricted distribution and extreme rarity in Madeira now. *Baetis enigmaticus* lives in fast-flowing streams and has 'scraping' mouthparts suited for eating diatoms and other algae growing on stony substrata, whereas *B. atlanticus* is a generalist in its habitat and feeding ecology.

Within the *B. canariensis* lineage, well-supported sister relationships were seen between neighbouring islands (La Gomera and La Palma, Tenerife and Gran Canaria), strongly suggesting that speciation has occurred within the archipelago. A close relationship with North Africa was also seen, similarly to Madeiran *B. atlanticus*. The dense sampling of *Rhodobaetis* carried out for the study suggests that there is no sister taxon elsewhere that would disrupt the monophyly of *B. canariensis*. The same was true for *B. pseudorhodani*, with strong support for monophyly of the islands. Interestingly, it appears that the *B. pseudorhodani* lineage never reached La Palma, where only the *B. canariensis* lineage is present. The fact that La Palma is among the youngest islands of the archipelago, dated about two mya (Hughes & Malmqvist, 2005), suggests that it is a younger lineage compared to *B. canariensis*. This is also suggested by a comparison of branch lengths in the two clades and the molecular clock results. The lack of phylogenetic resolution in *Cloeon* lineages makes it difficult to draw any conclusions about their dispersal history, although it was surprising to find strong support for the sister relationship between the Gran Canaria species and all others, including the European *C. dipterum*. A denser sampling of *Cloeon* from a wider range of European populations is now necessary, given the fact here that *C. dipterum* is composed of several species (see also Gattolliat *et al.*, 2008).

Colonisation of the Canaries and Madeira from multiple sources, both mainland and islands, has been reported for several other animals. A colonisation route from the Canary Islands towards Madeira has been suggested for beetles (*Calathus*; Emerson, Oromí & Hewitt, 2000), spiders (*Pholcus*; Dimitrov, Arnedo & Ribera, 2008), birds (*Anthus*; Illera, Emerson & Richardson, 2007) and several plants (e.g. Trusty *et al.*, 2005 and references therein). Clearly, dispersal acts in combination with speciation within and among islands of the archipelagos to determine present-day biodiversity. Based on existing taxonomic data and presumed affinities, mayfly colonisation was thought to be varied, but the degree of endemism, namely the interplay between dispersal and speciation, was unexpected.

Conservation

One implication of our results is that mayfly diversity on these Atlantic island archipelagos has been underestimated, which has considerable impact on the conservation status of these species. In Madeira, *B. atlanticus* is widespread and abundant, found in streams flowing through primary forest, urban and agricultural areas and even artificial channels (Malmqvist, 1988; Soldán & Godunko, 2006). In contrast, *B. enigmaticus* may be on the verge of extinction. Reported from a few streams in the 1990s (Stauder, 1991, 1995), the species was only collected once for this study and is now absent from most of the localities where it was found historically (author S.J.H., unpubl. data), making this species one of the most threatened in Europe. On the Canary Islands, even when *B. canariensis* and *B. pseudorhodani* were considered to occur on several islands of the archipelago, the increasing impact of tourism and agriculture on the natural habitats (Malmqvist *et al.*, 1993; Malmqvist, Nilsson & Báez, 1995; Nilsson *et al.*, 1998) meant they should probably be considered as endangered following IUCN criteria. Our analysis has revealed them to be composed of island endemics. On Gran Canaria, both species were abundant in two streams flowing in agricultural areas where they had been previously reported (Müller-Liebenau, 1971; Nilsson *et al.*, 1998). On Tenerife, the only stream that still possesses the two endemics is the Barranco del Infierno, which is a Special Nature Reserve. The Barranco del Río also has a small population of *B. canariensis*. Three other localities where *Baetis* populations were previously reported (Alba-Tercedor *et al.*, 1987; Malmqvist *et al.*, 1993) were sampled without success for our study (Barranco de Igueste, Barranco de Afur, and Barranco de Masca). In the Río Afur, the absence of *Baetis* probably results from the fact that the stream is periodically dry due to water abstraction (author M.B. & M.S., unpubl. data). On La Gomera, both endemics occur in a pristine stream at El Cedro in Garajonay National Park. Other sampled localities (a stream at Hermingua and one near Agulo) had no *Baetis* populations, clearly because the stream flow regimes are now temporary. Finally, on La Palma, the endemic *B. canariensis* s.l. was abundant in the Río Taburiente in the caldera of the National Park, but the population downstream of the caldera, mentioned by Müller-Liebenau (1971), has vanished because the river is dry for several months each year due to water abstraction (author M.B. & M.S., unpubl. data). In summary, the seven putative species found on the Canary Islands are known from a very restricted number of streams (one or two), making the survival of these species very critical in a near future if some effective regulations are not implemented rapidly. Together with *B. enigmaticus* in Madeira, these seven mayfly species are among the most threatened in Europe. The status of *Baetis nigrescens* on the Canaries is unclear. The species was common on the Canaries until the 1970's, but the last report dates back

to Malmqvist *et al.* (1995). The status of *Cloeon* species is also difficult to assess. They live in standing water habitats which can persist despite degradation of nearby streams. More information regarding the habitat requirements and tolerances of these *Cloeon* species is needed for an accurate assessment.

Acknowledgments

References

Alba-Tercedor J., Báez M. & Soldán T. (1987) New records of mayflies of the Canary Islands (Insecta, Ephemeroptera). *Eos*, **63**, 7–13.

Arnedo M.A., Oromí P., Martín De Abreu S. & Ribera C. (2008) Biogeographical and evolutionary patterns in the Macaronesian shield-backed katydid genus *Calliphona* Krauss, 1892 (Orthoptera: Tettigoniidae) and allies as inferred from phylogenetic analyses of multiple mitochondrial genes. *Systematic Entomology*, **33**, 145–158.

Bauernfeind E. & Soldán T. (2012) *The Mayflies of Europe (Ephemeroptera)*. Apollo Books, Ollerup, Denmark.

Bellemain E. & Ricklefs R.E. (2008) Are islands the end of the colonization road? *Trends in Ecology & Evolution*, **23**, 461–468.

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.

Brinck P. & Scherer E. (1961) On the Ephemeroptera of the Azoreas and Madeira. *Boletim do Museu Municipal do Funchal*, **47**, 55–66.

Brunton C.F.A. & Hurst G.D.D. (1998) Mitochondrial DNA phylogeny of Brimstone butterflies (genus *Gonepteryx*) from the Canary Islands and Madeira. *Biological Journal of the Linnean Society*, **63**, 69–79.

Carracedo J.C., Day S., Guillou H., Rodríguez Badiola E.R., Canas J.A. & Pérez Torrado F.J. (1998) Hotspot volcanism close to a passive continental margin: the Canary Islands. *Geological Magazine*, **135**, 591–604.

Chesters D.. (2013) collapsetypes.pl. Available at: http://sourceforge.net/projects/collapsetypes/.

Darriba D., Taboada G.L., Doallo R. & Posada D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

DeSalle R., Egan M.G. & Siddall M. (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**, 1905–1916.

Dijkstra K.-D.B., Monaghan M.T. & Pauls S.U. (2014) Freshwater biodiversity and aquatic insect diversification. *Annual Review of Entomology*, **59**, 143–163.

Dimitrov D., Arnedo M.A. & Ribera C. (2008) Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: evidence for long-term occupancy yet rapid recent speciation. *Molecular Phylogenetics and Evolution*, **48**, 596–614.

Drummond A.J., Suchard M.A., Xie D. & Rambaut A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Emerson B.C. (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, **11**, 951–966.

Emerson B.C. (2008) Speciation on islands: what are we learning? *Biolological Journal of the Linnean Society*, **95**, 47–52.

Emerson B.C. & Oromí P. (2005) Diversification of the forest beetle genus *Tarphius* on the Canary Islands, and the evolutionary origins of island endemics. *Evolution*, **59**, 586–598.

Emerson B.C., Oromí P. & Hewitt G.M. (2000) Interpreting colonization of the *Calathus* (Coleoptera: Carabidae) on the Canary Islands and Madeira through the application of the parametric bootstrap. *Evolution*, **54**, 2081–2090.

Folmer O., Black M., Hoe 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.

Fujisawa T. & Barraclough T.G. (2013) Delimiting species using single-locus data and the generalized mixed yule coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology*, **62**, 707–724.

Gattolliat J.-L. (2004) First reports of the genus *Nigrobaetis* Novikova & Kluge (Ephemeroptera: Baetidae) from Madagascar and La Réunion with observations on Afrotropical biogeography. *Revue Suisse de Zoologie*, **111**, 657–669.

Gattolliat J.-L., Hughes S.J., Monaghan M.T. & Sartori M. (2008) Revision of Madeiran mayflies (Insecta, Ephemeroptera). *Zootaxa*, **1957**, 52–68.

Gattolliat J.-L. & Sartori M. (2008) What is *Baetis rhodani* (Pictet, 1843) (Insecta, Ephemeroptera, Baetidae)? Designation of a neotype and redescription of the species from its original area. *Zootaxa*, **1957**, 69–80.

Gernhard T. (2008) Using birth-death model on tree. *Journal of Theoretical Biology*, **253**, 769–778.

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.

Godunko R.J., Prokopov G.A. & Soldán T. (2004) Mayflies of the Crimean Peninsula III. The description of *Baetis milani* sp. n. with notes on taxonomy of the subgenus *Rhodobaetis* Jacob, 2003 (Ephemeroptera: Baetidae). *Acta Zoologica Cracoviensia*, **47**, 231–248.

Greve C., Hutterer R., Groh K., Haase M. & Misof B. (2010) Evolutionary diversification of the genus *Theba* (Gastropoda: Helicidae) in space and time: a land snail conquering islands and continents. *Molecular Phylogenetics and Evolution*, **57**, 572–584.

Guindon S. & Gascuel O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Hughes S.J., Furse M.T., Blackburn J.H. & Langton P.H. (1998) A checklist of Madeiran freshwater macroinvertebrates. *Boletim do Museu Municipal do Funchal*, **50**, 5–41.

Hughes S.J. & Malmqvist B. (2005) Atlantic island freshwater ecosystems: challenges and considerations following the EU Water Framework Directive. *Hydrobiologia*, **544**, 289–297.

Illera J.C., Emerson B.C. & Richardson D.S. (2007) Population history of Berthelot's pipit: colonization, gene flow and morphological divergence in Macaronesia. *Molecular Ecology*, **16**, 4599–4612.

Illera J.C., Koivula K., Broggi J., Packert M., Martens J. & Kvist L. (2011) A multi-gene approach reveals a complex evolutionary history in the *Cyanistes* species group. *Molecular Ecology*, **20**, 4123–4139.

Juan C., Oromí P. & Hewitt G.M. (1997) Molecular phylogeny of darkling beetles from the Canary Islands: comparisons of inter island colonization patterns in two genera. *Biochemical Systematics and Ecology*, **25**, 121–130.

Juan I.I., Emerson B.C., Oromí P. & Hewitt G.M. (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology and Evolution*, **15**, 104–109.

Katoh K. & Standley D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

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

Lopez J.V., Yuhki N., Masuda R., Modi W. & O'Brien S.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.

Losos J.B. & Ricklefs R.E. (2009) Adaptation and diversification on islands. *Nature*, **457**, 830–836.

Lucentini L., Rebora M., Puletti M.E., Gigliarelli L., Fontaneto D., Gaino E. *et al.* (2011) Geographical and seasonal evidence of cryptic diversity in the *Baetis rhodani* complex (Ephemeroptera, Baetidae) revealed by means of DNA taxonomy. *Hydrobiologia*, **673**, 215–228.

Maddison W.P. & Maddison D.R.. (2011) MESQUITE: a modular system for evolutionary analysis. Version 2.75. Available at: http://mesquiteproject.org.

Malmqvist B. (1988) Downstream drift in Madeiran levadas: tests of hypotheses relating to the influence of predators on the drift of insects. *Aquatic Insects*, **10**, 141–152.

Malmqvist B., Nilsson A.N. & Báez M. (1995) Tenerifes freshwater macroinvertebrates - status and threats (Canary-Islands, Spain). *Aquatic Conservation: Marine and Freshwater Ecosystems*, **5**, 1–24.

Malmqvist B., Nilsson A.N., Báez M., Armitage P.D. & Blackburn J. (1993) Stream macroinvertebrate communities in the island of Tenerife. *Archiv für Hydrobiologie*, **128**, 209–235.

Monaghan M.T., Gattolliat J.-L., Sartori M., Elouard J.M., James H., Derleth P. *et al.* (2005) Trans-oceanic and endemic origins of the small minnow mayflies (Ephemeropter-

a, Baetidae) of Madagascar. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 1829–1836.

Monaghan M.T., Wild R., Elliot M., Fujisawa T., Balke M., Inward D.J. et al. (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, **58**, 298–311.

Müller-Liebenau I. (1971) Ephemeroptera (Insecta) von den Kanarischen Inseln. *Gewässer und Abwässer*, **50**, 7–40.

Navás S.J. (1906) Catalogo descriptivo de los Insectos Neuropteros de las Islas Canarias. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales*, **4**, 1–24.

Nilsson A.N., Malmqvist B., Báez M., Blackburn J.H. & Armitage P.D. (1998) Stream insects and gastropods in the island of Gran Canaria (Spain). *Annales de Limnologie - International Journal of Limnology*, **34**, 413–435.

Nogales M., López M., Jiménez-Asensio J., Larruga J.M., Hernández M. & González P. (1998) Evolution and biogeography of the genus *Tarentola* (Sauria: Gekkonidae) in the Canary Islands, inferred from mitochondrial DNA sequences. *Journal of Evolutionary Biology*, **11**, 481–494.

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.

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.

Pereira-da-Conceicoa L.L., Price B.W., Barber-James H.M., Barker N.P., de Moor F.C. & Villet M.H. (2012) Cryptic variation in an ecological indicator organism: mitochondrial and nuclear DNA sequence data confirm distinct lineages of *Baetis harrisoni* Barnard (Ephemeroptera: Baetidae) in southern Africa. *BMC Evolutionary Biology*, **12**, 26.

R Development Core Team (2012) *R: A language and environment for statistical computing*. Vienna, Austria. ISBN 3-900051-07-0, Available at: http//www.R-project.org.

Rehm P., Borner J., Meusemann K., von Reumont B.M., Simon S., Hadrys H. et al. (2011) Dating the arthropod tree based on large-scale transcriptome data. *Molecular Phylogenetics and Evolution*, **61**, 880–887.

Ribera I. (2008) Habitat constraints and the generation of diversity in freshwater macroinvertebrates. In *Aquatic Insects: Challenges to Populations* (EdsLancaster J. & Briers R.A.), pp. 289–311. CAB International, Wallingford, U.K.

Ribera I., Bilton D.T., Balke M. & Hendrich L. (2003a) Evolution, mitochondrial DNA phylogeny and systematic position of the Macaronesian endemic *Hydrotarsus* Falkenström (Coleoptera: Dytiscidae). *Systematic Entomology*, **28**, 493–508.

Ribera I., Bilton D.T. & Vogler A.P. (2003b) Mitochondrial DNA phylogeny and population history of *Meladema* diving beetles on the Atlantic Islands and in the Mediterranean basin (Coleoptera, Dytiscidae). *Molecular Ecology*, **12**, 153–167.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Hohna S. et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Sartori M. (2001) Current knowledge of mayfly research in Europe (Ephemeroptera). In *Trends in Research in Ephemeroptera and Plecoptera* (Ed.Dominguez E.), pp. 47–52. Kluwer Academic/Plenum Publishers, New York.

Savolainen E., Drotz M.K., Hoffsten P.-O. & Saura A. (2007) The *Baetis vernus* group (Ephemeroptera: Baetidae) of northernmost Europe: an evidently diverse but poorly understood group of mayflies. *Entomolgica Fennica*, **18**, 160–167.

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.

Soldán T. & Godunko R.J. (2006) *Baetis atlanticus* n.sp., a new species of the subgenus *Rhodobaetis* Jacob, 2003 from Madeira, Portugal (Ephemeroptera: Baetidae). *Genus*, **17**, 5–17.

Sowa R. (1975) What is *Cloeon dipterum* (Linnaeus, 1761)? The nomenclatural and morphological analysis of a group of the European species of *Cloeon* Leach (Ephemerida: Baetidae). *Entomologica Scandinavica*, **6**, 215–233.

Sroka P. (2012) Systematics and phylogeny of the West Palaearctic representatives of subfamily Baetinae (Insecta: Ephemeroptera): combined analysis of mitochondrial DNA sequences and morphology. *Aquatic Insects*, **34**, 23–53.

Ståhls G. & Savolainen E. (2008) MtDNA COI barcodes reveal cryptic diversity in the *Baetis vernus* group (Ephemeroptera, Baetidae). *Molecular Phylogenetetics and Evolution*, **46**, 82–87.

Stauder A. (1991) Water fauna of a Madeiran stream with notes on the zoogeography of the Macaronesian islands. *Boletim do Museu Municipal do Funchal*, **43**, 243–299.

Stauder A. (1995) Survey of the Madeiran limnological fauna and their zoogeographical distribution. *Boletim do Museu Municipal do Funchal*, **4**, 715–723.

Thomas A.G.B. & Solda´n T.. (1987) Baetis ingridae n. sp., Ephe´me´ropte´re nouveau de Corse (Baetidae). *Annales de Limnologie - International Journal of Limnology*, **23**, 23–26.

Trusty J.L., Olmstead R.G., Santos-Guerra A., Sa-Fontinha S. & Francisco-Ortega J. (2005) Molecular phylogenetics of the Macaronesian-endemic genus *Bystropogon* (Lamiaceae): palaeo-islands, ecological shifts and interisland colonizations. *Molecular Ecology*, **14**, 1177–1189.

Vuataz L., Sartori M., Gattolliat J.-L. & Monaghan M.T. (2013) Endemism and diversification in freshwater insects of Madagascar revealed by coalescent and phylogenetic analysis of museum and field collections. *Molecular Phylogenetics and Evolution*, **66**, 979–991.

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.

Weingartner E., Wahlberg N. & Nylin S. (2006) Speciation in Pararge (Satyrinae: Nymphalidae) butterflies - North Africa is the source of ancestral populations of all Pararge species. *Systematic Entomology*, **31**, 621–632.

Williams H.C., Ormerod S.J. & Bruford M.W. (2006) Molecular systematics and phylogeography of the cryptic species complex *Baetis rhodani* (Ephemeroptera, Baetidae). *Molecular Phylogenetics and Evolution*, **40**, 370–382.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 GenBank accession numbers of specimens
Appendix S2 (a) Clock-constrained molecular phylogeny using fossil and geological calibration with a relaxed (lognormal) clock
Table S1 Details of sampling sites on the Canary Islands and Madeira

(Manuscript accepted 13 August 2014)