MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2010) **10**, 264–273 doi: 10.1111/j.1755-0998.2009.02763.x

DNA BARCODING

Mitochondrial DNA barcoding detects some species that are real, and some that are not

KANCHON K. DASMAHAPATRA,* MARIANNE ELIAS,† RYAN I. HILL,‡ JOSEPH I. HOFFMAN§ and JAMES MALLET*¶

*Department of Genetics, Evolution and Environment, University College London, 4 Stephenson Way, London NW1 2HE, UK, †NERC Centre for Population Biology, Imperial College London, Silwood Park Campus, Ascot, Berkshire SL5 7PY, UK; CNRS, UMR 7205, Muséum National d'Histoire Naturelle, 16 Rue Buffon, CP39, 75005 Paris, France, ‡FAS Center for Systems Biology, Harvard University, 52 Oxford St, Northwest Lab Room 458.40-3, Cambridge, MA 02138, USA, §Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK, ¶Wissenschaftskolleg zu Berlin, Wallotstraße 19, 14193 Berlin, Germany

Abstract

Mimicry and extensive geographical subspecies polymorphism combine to make species in the ithomiine butterfly genus *Mechanitis* (Lepidoptera; Nymphalidae) difficult to determine. We use mitochondrial DNA (mtDNA) barcoding, nuclear sequences and amplified fragment length polymorphism (AFLP) genotyping to investigate species limits in this genus. Although earlier biosystematic studies based on morphology described only four species, mtDNA barcoding revealed eight well-differentiated haplogroups, suggesting the presence of four new putative 'cryptic species'. However, AFLP markers supported only one of these four new 'cryptic species' as biologically meaningful. We demonstrate that in this genus, deep genetic divisions expected on the basis of mtDNA barcoding are not always reflected in the nuclear genome, and advocate the use of AFLP markers as a check when mtDNA barcoding gives unexpected results.

Keywords: AFLP, cryptic species, DNA barcoding, ithomiine, *Mechanitis*

Received 6 May 2009; revision received 20 July 2009; accepted 24 July 2009

Introduction

DNA barcoding has recently emerged as a rapid method for species discovery and biodiversity assessment (Hajibabaei *et al*. 2006; Borisenko *et al*. 2008; Stoeckle & Hebert 2008). For animal taxa, the majority of these studies have used a short section of mitochondrial DNA (mtDNA), namely the first ~650 bp of the 5′-end of the *cytochrome oxidase I* gene (*CoI*) (Hebert *et al*. 2003; Elias-Gutierrez *et al*. 2008; Rock *et al*. 2008). DNA barcoding has been argued to revolutionize taxonomy by allowing rapid species identification and discovery without the need for detailed taxonomic expertise with increasing economy (Hajibabaei *et al*. 2007; Stoeckle & Hebert 2008). But the practice of mtDNA barcoding has received much criticism on methodological (Will & Rubinoff 2004), theoretical (Hickerson *et al*. 2006) and empirical grounds (Hurst & Jiggins 2005; Meyer & Paulay 2005; Elias *et al*. 2007; Wiemers & Fiedler 2007). Despite the problems, undoubted successes for mtDNA barcoding have been the discovery of cryptic species overlooked by more traditional taxonomic methods (Smith *et al*. 2006; Burns *et al*. 2007).

Several studies have sought to overcome some of the above problems with mtDNA barcoding by supplementing mtDNA sequences with nuclear sequences (Monaghan *et al*. 2005; Elias *et al*. 2007). However, success with nuclear sequences for DNA taxonomy has been limited, largely because of the difficulty in finding and sequencing nuclear loci that diverge fast enough to distinguish closely related cryptic species (Dasmahapatra & Mallet 2006). A possible alternative for studying the nuclear genome is the analysis of amplified fragment length polymorphisms (AFLPs), which are anonymous dominant nuclear markers, typically fast evolving and readily amplifiable in any organism (Vos *et al*. 1995; Mueller & Wolfenbarger 1999). In this study, we first use mtDNA

Correspondence: Kanchon K. Dasmahapatra, Fax: +44 (0)207 6795052; E-mail: k.dasmahapatra@ucl.ac.uk

sequences to test species limits and affiliations in the difficult ithomiine butterfly genus *Mechanitis* (Lepidoptera; Nymphalidae), and then further investigate these results by utilizing both nuclear gene sequences and AFLP markers.

Four *Mechanitis* species have been described. These species are all abundant locally as well as widely distributed in the neotropics (Brown 1977). Like most other ithomiines, all *Mechanitis* species show multiple geographical subspecies variation and typically have orange, brown, yellow and black wing colouration (Brown 1979; Lamas 2004). The many geographical forms of *Mechanitis*, involved in Müllerian mimicry with various members of the ~360 species of Ithomiinae as well as with *Heliconius*, had led to an extremely problematic species-level taxonomy. However, careful biosystematic studies in the 1970s gave rise to a reasonably stable classification, based on morphology, distribution, hybrid zones and ecology (Brown 1977), that is still generally accepted (Lamas 2004). In this study, we investigate the four putative species comprising this genus.

Materials and methods

A number of subspecies have been described for all the four species of *Mechanitis* (Brown 1977; Lamas 2004), some of which were studied in this work: *Mechanitis mazaeus* (*Mechanitis m. deceptus*, *M. m.* cf. *phasianita*, *M. m. messenoides*, *M. m. pannifera*, *M. m. mazaeus*), *Mechanitis polymnia* (*M. p. proceriformis*, *M. p. casabranca*, *M. p. bolivarensis*, *M. p. eurydice*), *Mechanitis lysimnia* (*M. l. roqueensis*, *M. l. lysimna*, *M. l. solaria*) and *Mechanitis menapis* (*M. m. mantineus*). The first three species have wide distributions (Brown 1979), and our samples were collected from northern and central Peru (eastern San Martin and southern Loreto), Ecuador, northern Venezuela and the Atlantic coast of Brazil (Table 1). *Mechanitis menapis* has a narrower distribution, replacing *M. mazaeus* west of the Andes (Brown 1979) and in Central America; our *M. menapis* specimens were obtained from western Ecuador. We sampled 121 specimens, mainly from Ecuador and Peru, together with five each from Venezuela and eastern Brazil for comparison (Table 1). Details of sampling locations are provided in the Supplementary Material (Table S1).

DNA was extracted from legs and thoraces using QIAamp DNA Micro and DNeasy Blood and Tissue Kits (QIAGEN). Approximately 640 bp of mtDNA comprising the 5′-end of *CoI*, the 'barcoding region', was amplified and sequenced in all specimens. To examine the effect of using more extensive mtDNA sequence data, a further ~1500 bp, comprising the remaining 3′-portion of *CoI*, the *tRNA-leu* gene, and the 5′-end of *CoII*, was also sequenced from 71 specimens representing all the four species. To examine whether the patterns revealed by mtDNA were reflected in the nuclear genome, sequences were also obtained from three nuclear loci: *Tektin* (715 bp, 58 sequences), *Rpl5* (720 bp, 68 sequences) and *Tpi* (1150 bp, 73 sequences) (Mallarino *et al.* 2005; Whinnett *et al.* 2005a). For *Rpl5* and *Tpi*, these sequences included representatives of all the four species, but for *Tektin*, no sequences were obtained for *M. menapis*. Indels in the intronic regions of *Rpl5* and *Tpi* sometimes resulted in the amplification of alleles with different sizes from a single individual. Unless sequence quality was low, sequencing in both directions allowed indels to be readily identified, whereupon each allele was deconvoluted using the information from the double-peak signals following the indel (Flot *et al.* 2006). PCR primers and reaction conditions have been reported previously elsewhere (Whinnett *et al.* 2005a; Dasmahapatra *et al.* 2007; Elias *et al.* 2007); a detailed description is also provided in Table S2. Cycle sequencing was carried out using the Big Dye Terminator 3.1 Cycle Sequencing Kit (Applied Biosystems). All sequences obtained for this study

Table 1 Subspecies and numbers of specimens collected in different countries

	Mechanitis mazaeus		*Mechanitis lysimnia*		*Mechanitis polymnia*		*Mechanitis menapis*	
	Subspecies	*n*	Subspecies	*n*	Subspecies	*n*	Subspecies	*n*
Peru	*deceptus*	8	*roqueensis*	8	*proceriformis*	12	—	
	cf. *phasianita*	7			*eurydice*	1		
	mazaeus	11						
Ecuador	*deceptus*	13	*roqueensis*	7	cf. *proceriformis*	8	*mantineus*	3
	messenoides	16						
	mazaeus	12						
	messenoides × *mazaeus*	4						
	deceptus × *mazaeus*	1						
Venezuela	*pannifera*	2	*solaria*	2	*bolivarensis*	1	—	
Brazil	—		*lysimnia*	2	*casabranca*	3	—	

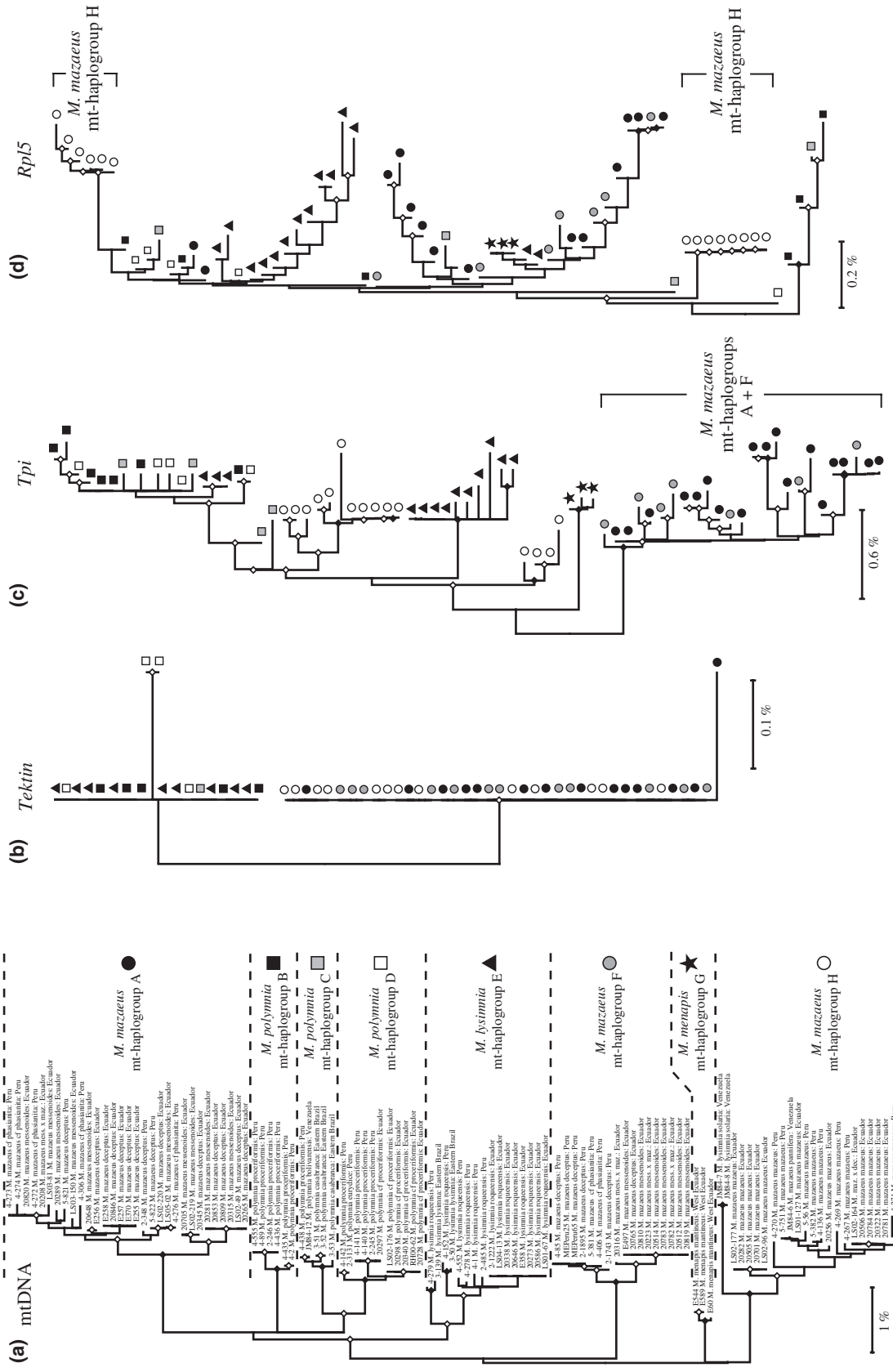

Fig. 1 Neighbour-joining trees constructed from mitochondrial and nuclear sequences showing relationships between the eight *Mechanitis* mitochondrial haplogroups. Nodes marked with filled and open symbols have >90%, and <90%, but >50% bootstrap support respectively. Scale bars represent raw percentage sequence divergence.

have been deposited in GenBank (FJ445856–FJ446152, EU068843–EU068856, EU068966, EU068993–EU068900 and EU069070–EU069075).

Genetic divisions revealed by mitochondrial *CoI* may sometimes not be captured at autosomal nuclear loci. This could be because of the lack of coalescence within species given the ~4× effective population size (N_e) of autosomal loci compared with cytoplasmic mtDNA, and also because typical exons readily sequenced across genera tend to evolve at slower rates than *CoI*. The *Rpl5* and *Tpi* loci we used also span fast-evolving intronic regions. In addition, *Tpi* is a sex-linked locus having only ~3× N_e compared withmitochondrial loci. To obtain even higher resolution nuclear data, we genotyped 84 specimens representing all the four species and all eight mtDNA haplogroups (Fig. 1a) using AFLP markers. The samples were genotyped using four AFLP primer combinations: *Taq*I-CGA + *Eco*RI-ACA; *Taq*I-CAG + *Eco*RI-AGC; *Taq*I-CAG + *Eco*RI-ATG; *Taq*I-CCA + *Eco*RI-ACA. AFLP primers and protocols used are described in Madden *et al.* (2004). AFLP profiles were visualized by autoradiography. Samples with aberrant AFLP profiles were discarded (Bonin *et al.* 2004) and to ensure reliability, the remaining AFLP genotypes were scored by eye; 108 putative loci were polymorphic and could be scored reliably.

For each of the four sequenced loci, bootstrapped neighbour-joining (NJ) trees based on raw sequence divergences were constructed using MEGA 4 (Tamura *et al.* 2007). Haplotype clusters within the mitochondrial tree were defined using a threshold of 1.5% between-cluster raw sequence divergence. The application of such a threshold is somewhat arbitrary (Meyer & Paulay 2005) and is further addressed in the Discussion. Pairwise F_{ST} (Weir & Cockerham 1984) among mtDNA haplogroups (see Results section) was calculated from AFLP genotypes with AFLP-SURV v1.0 (Vekemans 2002), using the approach of Lynch & Milligan (1994). The optimal number of genotypic clusters indicated by the AFLP genotypes was established with the Bayesian program Structure 2.2 (Pritchard *et al.* 2000), using standardized inference criteria (Evanno *et al.* 2005). Following a 100 000 step burn-in period, data were collected over 100 000 Markov chain Monte Carlo repetitions. Structure analysis was carried out on the data set, increasing *K* from 1 to 10. At each value of *K*, the analysis was repeated three times to check between-run consistency.

Results

Topologies of the mtDNA NJ trees based on the 636-bp *CoI* 'barcoding' region and the full 2000 bp are similar and show the same major mtDNA haplogroups. In Fig. 1a we show the former, as sequences are obtained from a larger number of specimens. Eight major nonoverlapping clusters or mtDNA haplogroups are detected on applying a threshold of 1.5% sequence divergence: three each within *Mechanitis mazaeus* (mt-haplogroups A, F and H) and *Mechanitis polymnia* (mt-haplogroups B, C and D), and one each corresponding to *Mechanitis menapis* (mt-haplogroup G) and *Mechanitis lysimnia* (mt-haplogroup E). Not included in these groups are the two specimens of *M. lysimnia solaria* from Venezuela that appear sister to the *M. mazaeus* mt-haplogroup H, rather than lying within main *M. lysimnia* clade; this exception is described further in the Discussion. There is also evidence for another deep division within mt-haplogroup E; however, as one of the clades contains only two specimens, this division may be a consequence of limited sampling.

Pairwise mtDNA distances and AFLP-based F_{ST} between the eight mt-haplogroups are shown in Table 2. The raw average pairwise mtDNA distance between mt-haplogroups is 2.7%, the largest being among *M. mazaeus* mt-haplogroups (≥3.5%) and the smallest between *M. polymnia* mt-haplogroups C and D (1.6%),

Table 2 Average raw percentage pairwise mtDNA distances between mtDNA haplogroups are shown above the diagonal. Intra-haplogroup mtDNA distances are presented along the diagonal. AFLP-based F_{ST} between mtDNA haplogroups are shown below the diagonal in italics

		Mechanitis lysimnia	*Mechanitis menapis*	*Mechanitis polymnia*			*Mechanitis mazaeus*		
Species	mt-haplogroup	E	G	B	C	D	A	F	H
Mechanitis lysimnia	E	0.6	2.3	2.4	2.3	2.0	2.8	2.8	3.1
Mechanitis menapis	G	*0.41*	0.1	2.5	2.5	2.1	2.9	2.1	2.6
Mechanitis polymnia	B	*0.19*	*0.53*	0.2	2.3	2.2	2.8	3.0	3.3
	C	*0.27*	*0.54*	*0.03*	0.2	1.6	2.7	3.3	3.3
	D	*0.24*	*0.55*	*0.03*	*0.00*	0.2	2.7	2.9	3.2
Mechanitis mazaeus	A	*0.41*	*0.47*	*0.49*	*0.49*	*0.50*	0.6	3.6	4.1
	F	*0.39*	*0.42*	*0.47*	*0.47*	*0.48*	*0.00*	0.4	3.5
	H	*0.22*	*0.48*	*0.31*	*0.33*	*0.29*	*0.43*	*0.40*	0.5

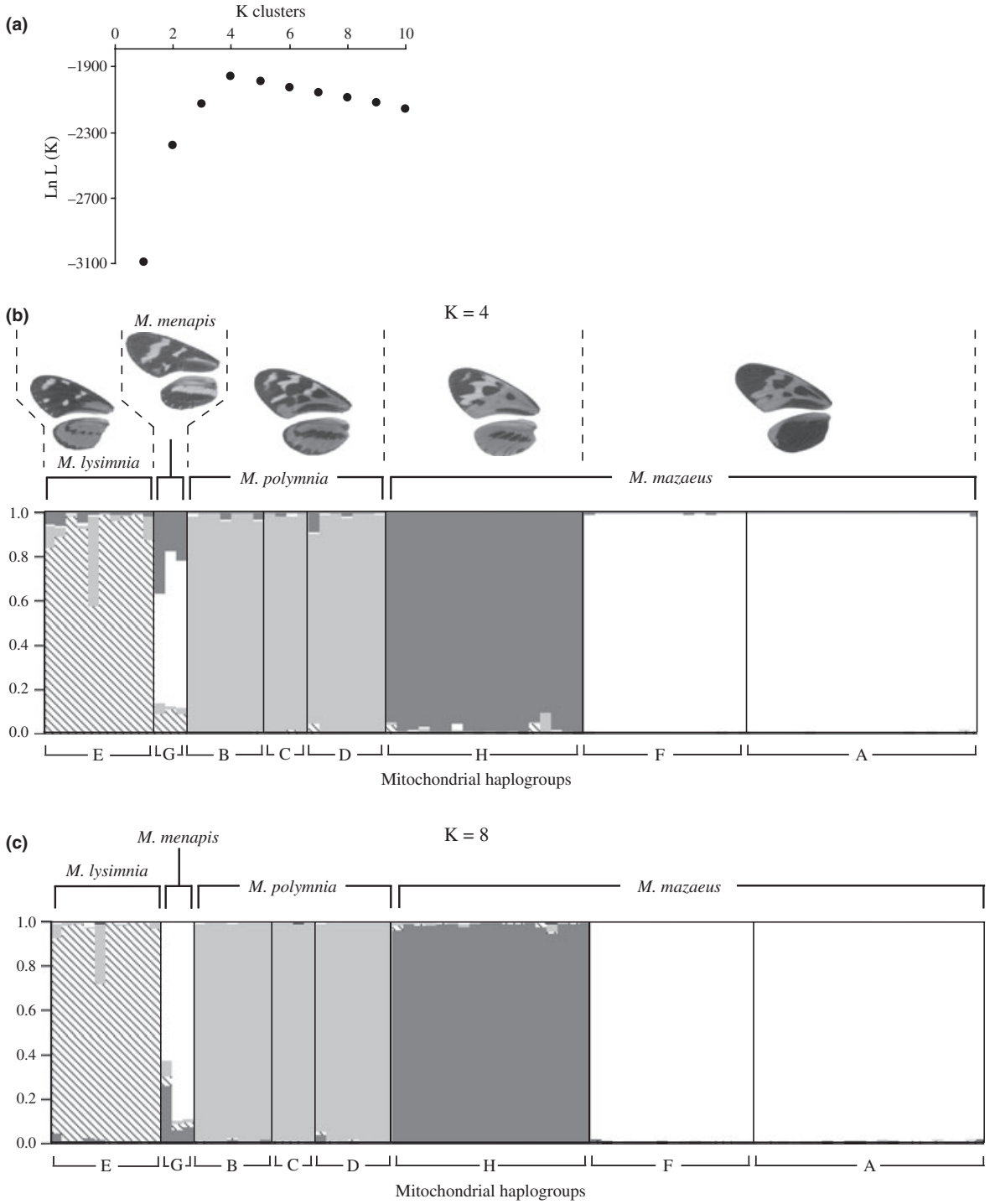

Fig. 2 Structure analysis of AFLP genotypes. (a) Log likelihood of data as a function of *K*, the number of clusters. Likelihoods from each of the three replicate runs at each *K* are indistinguishable and the average likelihood is shown. Highest likelihood is achieved with four clusters, although eight mt-haplogroups are present. Structure results for (b) the optimal number of genotypic clusters, $K = 4$, and (c) the number of mt-haplogroups, $K = 8$. Each of the 85 individuals is represented by a vertical bar broken into *K* shaded segments. The proportion of each colour in the bar indicates the posterior mean probability of ancestry from each genetic cluster. (b) and (c) are virtually identical, indicating the absence of any significant additional genetic structure for more than four clusters. Wing patterns shown are typical of *Mechanitis lysimnia roqueensis*, *Mechanitis menapis mantineus*, *Mechanitis mazaeus deceptus*, *Mechanitis mazaeus mazaeus* and *Mechanitis polymnia proceriformis*.

with all other distances being ≥2.0%. The average intra-haplogroup variation is, in contrast, only 0.3%. AFLP-based F_{ST} between the three *M. polymnia* mt-haplogroups and that between *M. mazaeus* mt-haplogroups A and F are low (≤0.03). In contrast, all other between-mt-haplogroup F_{ST}-values are >0.19.

Hardly any genetic variation was present among *Tektin* sequences. The only clear pattern observable is the single shared *Tektin* haplotype in *M. lysimnia* and *M. polymnia*, suggesting a sister relationship (Fig. 1b). In contrast, high levels of polymorphism are present at both *Tpi* and *Rpl5* (Fig. 1c, d). Neither gene shows distinctions among *M. polymnia* mt-haplogroups B, C and D. In the *Rpl5* genealogy, there is some support for separate evolution of *M. mazaeus* mt-haplogroups H from A and F. Similarly, a monophyletic *M. mazaeus* clade consisting of mt-haplogroups A and F, which is distinct from *M. mazaeus* mt-haplogroup H, is also supported in the *Tpi* genealogy, as is a separate clade corresponding to *M. menapis*. Overall, the nuclear gene genealogies suggest widespread paraphyly of recognized species and mt-haplogroups.

Between-run consistency was high in the Structure analysis of AFLP genotypes: replicate runs at each *K*-value yielded virtually identical likelihoods. Figure 2(a) shows the average likelihood from these replicates. The optimal number of groups was four (Fig. 2a, b): two within *M. mazaeus*, and one each matching *M. lysimnia* and *M. polymnia*. Of the two *M. mazaeus* AFLP clusters, one corresponds exactly to mt-haplogroup H (*M. mazaeus mazaeus*); whereas the other *mazaeus* cluster is a mixture of mt-haplogroups A and F (*M. mazaeus deceptus*, *M. mazaeus* cf. *phasianita* and *M. mazaeus messenoides*), with no evidence of subdivision between them. The *M. lysimnia* and *M. polymnia* AFLP clusters correspond perfectly with morphology-based species designations, except that one specimen identified using morphology as *M. lysimnia* shows genotypic evidence of being a hybrid between the two; and there is no evidence for subdivision of *M. polymnia* along mt-haplogroup lines. There is no single AFLP genotypic cluster associated with *M. menapis*; instead, it appears to share genotypes largely with *M. mazaeus* mt-haplogroups A + F, and to a lesser extent with *M. mazaeus* mt-haplogroup H and *M. lysimnia*. Increasing the number of clusters to eight, the total number of mt-haplogroups did not alter the pattern and Fig. 2b ($K = 4$) is virtually identical to Fig. 2c ($K = 8$), with the four additional clusters making negligible contributions. Additionally, as Structure may be unable to resolve subdivisions among very closely related groups when the data set includes more divergent groups, separate Structure analyses were also carried out on two restricted data sets consisting of only *M. polymnia* individuals ($n = 18$), and only *M. mazaeus* individuals from mt-haplogroups A and F ($n = 36$). Structure was unable to recover any subdivisions within either of these two restricted data sets. This is further evidence for the lack of nuclear divisions corresponding to mt-haplogroups within *M. polymnia* and *M. mazaeus* (mt-haplogroups A and F).

Discussion

Mitochondrial DNA barcoding of the four species in the genus *Mechanitis* revealed deep genetic divisions corresponding to eight mt-haplogroups (Fig. 1a). One mt-haplogroup each corresponded to *Mechanitis lysimnia* and *Mechanitis menapis*. However, three mt-haplogroups were present within each of *Mechanitis mazaeus* and *Mechanitis polymnia*, suggesting the existence of four putative cryptic species in addition to the four species already recognized in morphological and biosystematic work. In contrast, whereas our nuclear sequence data give little resolution, our AFLP data strongly indicate the existence of only four genetic clusters: two within *M. mazaeus* and one each corresponding to *M. lysimnia* and *M. polymnia* (Fig. 2b). Thus, we obtain disparate results using mitochondrial and nuclear markers.

Mitochondrial DNA barcoding relies on intraspecific genetic variation being much less than interspecific genetic variation. When this condition is met, a 'barcoding gap' exists (Meyer & Paulay 2005), and clusters corresponding to genetically more homogenous entities can be discerned. Based on such cases, some proponents of mtDNA barcoding have advocated the use of a threshold of sequence divergence above which genetic clusters may be considered species (Hebert *et al.* 2004). Subsequent careful studies have demonstrated the lack of a barcoding gap in a number of taxa when sampling of species, populations and individuals within the study group is thorough (Meyer & Paulay 2005; Burns *et al.* 2007; Elias *et al.* 2007; Wiemers & Fiedler 2007). In the case of *Mechanitis*, we have sampled all four described species in the genus. Although we have not included all of the many geographical subspecies of these species, we have used specimens collected across a wide geographical area representing opposite ends of the species' geographical distribution in South America. With our current sampling, mitochondrial data show eight major genetic clusters or haplogroups, clearly separated from one another by a large average genetic distance of 2.8%. We found no overlap between intra- and intercluster genetic variation.

Only four of the eight mtDNA divisions are reflected in the nuclear genome. Evidence from nuclear sequence information is weak, probably a result of incomplete lineage sorting because of recency of origin coupled with large effective population sizes within these very widespread and common species. However, even multilocus

AFLP genotyping fails to detect groups corresponding to three of the four novel mt-haplogroups. Although *M. menapis* seems well-separated by mt-haplogroup (Fig. 1a), the *M. menapis* samples do not form their own AFLP genotypic cluster, instead they appear to share most alleles with the two allopatric *M. mazaeus* clusters (Fig. 2b). However, no such patterns are associated with any of the three *M. polymnia* mt-haplogroups or the *M. mazaeus* mt-haplogroups A and F.

Using thresholds to define mitochondrial haplogroups is somewhat arbitrary and changing the threshold clearly impacts the number of mt-haplogroups detected (Meyer & Paulay 2005). For instance, raising the threshold to 2% results in the merger of *M. polymnia* mt-haplogroups C and D, while all other haplogroups remain unchanged. Similarly, reducing the threshold to 1% results in the splitting of *M. lysimnia* mt-haplogroup E into two. Regardless of the method selected to define the mt-haplogroups, evidently we have a situation where not all deep mitochondrial divisions are reflected in the nuclear genome.

Two possibilities might explain the mismatch between mt-haplogroups and AFLP genotypic clusters. First, that some of the mt-haplogroups are not 'real species', which we take to mean taxa that can maintain multilocus genetic, ecological, behavioural and/or morphological differences in sympatry. Second, the AFLP markers may not be sensitive enough to detect small differences between real but closely related species detected by mtDNA barcoding. Several lines of evidence strongly indicate that the former is more likely. Detailed examination of the wing patterns of the specimens in the three *M. polymnia* mt-haplogroups as well as *M. mazaeus* mt-haplogroups A and F failed to reveal any correlation between wing patterning and mt-haplogroup. In contrast, every AFLP genetic cluster detected by STRUCTURE is correlated with wing phenotype (Fig. 2b; discussed below). This suggests that whereas the AFLP clusters are biologically relevant, some of the mitochondrial divisions are not. AFLP markers provide a nuclear, multilocus, genome-wide picture of genetic divergence and as they sample noncoding variation, they have relatively rapid rates of evolution. This sensitivity is routinely exploited to reveal population genetic patterns within single species (Takami *et al*. 2004; Baus *et al.* 2005; Chaput-Bardy *et al.* 2008). The sensitivity of AFLP markers for detecting small genetic differences relative to mtDNA is also demonstrated by results from a parallel study on butterfly species within the ithomiine genus *Melinaea*. In this genus, there are at least six clearly distinguishable morphological species (*Melinaea satevis*, *Melinaea menophilus*, *Melinaea marsaeus*, *Melinaea idae*, *Melinaea mneme* and *Melinaea isocomma*) which cannot be detected using mtDNA barcoding (Whinnett *et al.* 2005b; Elias *et al.* 2007; Dasmahapatra *et al.* in prep.). Yet, where mtDNA barcoding failed to detect distinct groups, the same AFLP primer combinations used in this study were able to confirm morphology-based species divisions (Dasmahapatra *et al.* in prep.). In contrast to the genome-wide picture obtained from AFLP markers, mtDNA reflects evolution only of a single, nonrecombining, maternally inherited mitochondrial genome, which can be affected by factors such as the vagaries of coalescence, interspecific hybridization and effects of selection, such as *via* maternally transmitted endosymbionts like *Wolbachia* (Hurst & Jiggins 2005). As such, genome-wide genetic clustering revealed by multiple AFLP markers is likely to be more generally useful for the discovery of 'real species' than mtDNA barcoding.

Previous detailed morphological and biosystematic work (Brown 1977; Lamas 2004) recognized four species within *Mechanitis* (*M. lysimnia*, *M. polymnia*, *M. mazaeus* and *M. menapis*), each with multiple geographical subspecies. The nuclear data reported in this study strongly support monophyletic *M. lysimnia* and *M. polymnia* clades. However, mitochondrial paraphyly of *M. lysimnia* is suggested by two mtDNA sequences of Venezuelan *M. lysimnia solaria* (Fig. 1a). Unfortunately, it was not possible to genotype these specimens using AFLPs as the DNA was obtained from old dried specimens and was degraded. This split within *M. lysimnia* may be correlated with chromosome numbers (Brown *et al*. 2004), and a group of *M. lysimnia* subspecies, including *M. l. solaria*, might represent a separate species from *M. lysimnia sensu stricto*. Future investigation of phylogenetic relationships within *Mechanitis* should focus on this apparent division within *M. lysimnia*.

Mechanitis mazaeus also exhibited mitochondrial paraphyly, and two genotypic clusters are supported by nuclear data, one comprising mt-haplogroup H and the other combining mt-haplogroups A and F. *M. mazaeus* mt-haplogroups A and F correspond to melanic forms (Fig. 2b) currently considered to represent subspecies *Mechanitis mazaeus messenoides* (Colombia south to Ecuador) and *deceptus* (Ecuador southwards), both of which are generally found at mid elevations on the eastern slopes of the Andes. They are mimetic of melanic sympatric subspecies and races of *Melinaea marsaeus*, *Melinaea isocomma*, *Heliconius numata* and other species (Brown 1977, 1979). Although there is some evidence for intermediate colour patterns (Brown 1977) and Table S1, most specimens in *M. mazaeus* mt-haplogroup H correspond to paler lowland *M. mazaeus* (*sensu stricto*) (Fig. 2b) involved in mimicking the generalized lowland ithomiine and heliconiine tiger patterns. In addition to this differentiation in adult colour pattern, detailed analysis of larval morphology and adult host plant choice also indicate differences between the two *M. mazaeus* nuclear genotypic

clusters, and a lack of differentiation between *M. mazaeus* mt-haplogroups A and F (Hill *et al.* in prep.).

We conclude that forms allied with *M. mazaeus messenoides* and those allied with *M. mazaeus mazaeus* are best considered separate species, as their distributions overlap extensively in the Eastern foothills of the Andes (Brown 1977) with little evidence of hybridization from AFLP loci. Therefore, *M. mazaeus* mt-haplogroups A and F should probably now regain species-level designation as *M. messenoides* (including *M. messenoides messenoides* C. & R. Felder 1865, *M. messenoides* cf. *phasianita* Haensch 1905 and *M. messenoides deceptus* Butler 1873). The nuclear evidence points to the fourth species, *M. menapis*, being a *trans*-Andean form close to *M. mazaeus* and *M. messenoides*, as it shares genotypes with both *cis*-Andean *M. mazaeus* forms described above. *M. menapis* is not sympatric with *M. mazaeus*, and instead replaces it west of the Andes and in Central America.

The deep splits at mtDNA within *M. polymnia* are difficult to explain. There is some evidence that these splits may correspond to geographical areas: *M. polymnia* mt-haplogroups B and C are absent from sites sampled within Ecuador, mt-haplogroup B is found only in Peruvian samples and mt-haplogroup C is dominated in samples from outside Peru and Ecuador (Fig. 1). However, any such patterns are weak as all three mt-haplogroups are found within a small area of Peru. The absence of these haplogroups elsewhere may result from limited sampling outside Peru and Ecuador.

Morphologically based biosystematic work (Brown 1977) has shown itself to be more useful in this genus than mtDNA barcoding. Brown (1977) and Lamas (2004) accepted four species on the basis of morphology, a result largely upheld by our multilocus nuclear analysis. The only major discrepancy is that intermediate colour patterns in Ecuador (Table 1) apparently indicating hybridization between upland *M. mazaeus deceptus/messenoides* and lowland *mazaeus* (*sensu stricto*) led Brown (1977) to lump the highland melanic forms incorrectly as subspecies of *mazaeus*. In comparison, mtDNA barcoding reveals eight nonoverlapping monophyletic mt-haplogroups within *Mechanitis*, three of which are not detected in our nuclear analysis, suggesting that they do not correspond to 'real species'. Mitochondrial DNA barcoding provides a very sensitive technique to find new taxa, but a downside is that such taxa may have no basis in biological reality.

Focussing on the butterfly genus *Mechanitis*, we have used nuclear sequences and sensitive AFLP genotyping to demonstrate how deep genetic divisions in mtDNA are not always reflected by corresponding divisions in the nuclear genome. Such cryptic barcoding clusters may instead represent locally divergent populations that have undergone a bottleneck, or represent retained diversity due to large genetically effective population sizes, rather than having speciated in the normal sense of producing coexisting populations genetically divergent at multiple loci. Putative cryptic species detected by mtDNA barcoding merit closer investigation via analysis of nuclear genetic data, or more in-depth examination of ecology and taxonomy (Smith *et al.* 2006; Burns *et al.* 2007). There are clearly some limitations with using AFLP markers such as their lower per locus information content and higher error rates compared with co-dominant markers such as microsatellites or single nucleotide polymorphisms (Bonin *et al.* 2004; Dasmahapatra *et al.* 2008), as well as size homology of markers (Althoff *et al.* 2007). However, owing to their ease of amplification across taxa, sensitivity to small genetic differences and genome-wide coverage, we advocate the use of AFLP markers in cases where mtDNA barcoding reveals unexpected results, such as a failure to recover known taxonomic divisions or the presence of additional 'cryptic' taxa.

Acknowledgements

We are grateful to Gerardo Lamas of the Museo de Historia Natural, UNMSM, Lima, Peru, and Keith Willmott from the Florida Museum of Natural History, Gainesville, USA, for tireless lepidopteran expertise. Fraser Simpson assisted with many aspects of this project. Collecting permits and support were provided by INRENA (Peru) and the Ministerio del Ambiente and the Museo Ecuatoriano de Ciencias Naturales (Ecuador). Logistic support in Ecuador was provided by the Napo Wildlife Centre and La Selva Jungle Lodge. We thank NERC, DEFRA–Darwin, the Leverhulme Trust and the Margaret C. Walker fund for teaching and research in systematic entomology for funding different aspects of this project.

References

Althoff DM, Gitzendanner MA, Segraves KA (2007) The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. *Systematic Biology*, **56**, 477–484.

Baus E, Darrock DJ, Bruford MW (2005) Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. *Molecular Ecology*, **14**, 3373–3382.

Bonin A, Bellemain E, Eidesen PB *et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Borisenko AV, Lim BK, Ivanova NV, Hanner RH, Hebert PDN (2008) DNA barcoding in surveys of small mammal communities: a field study in Suriname. *Molecular Ecology Resources*, **8**, 471–479.

Brown KS (1977) Geographical patterns of evolution in Neotropical Lepidoptera: differentiation of the species of *Melinaea* and *Mechanitis* (Nymphalidae, Ithomiinae). *Systematic Entomology*, **2**, 161–197.

Brown KS (1979) *Ecologia Geográfica e Evolução nas Florestas Neotropicais*. Universidade Estadual de Campinas, Campinas, Brazil.

Brown KS, von Schoultz B, Suomalainen E (2004) Chromosome evolution in neotropical Danainae and Ithomiinae (Lepidoptera). *Hereditas*, **141**, 216–236.

Burns JM, Janzen DH, Hajibabaei M, Hallwachs W, Hebert PDN (2007) DNA barcodes of closely related (but morphologically and ecologically distinct) species of skipper butterflies (Hesperiidae) can differ by only one to three nucleotides. *Journal of the Lepidopterists' Society*, **61**, 138–153.

Chaput-Bardy A, Lemaire C, Picard D, Secondi J (2008) In-stream and overland dispersal across a river network influences gene flow in a freshwater insect, *Calopteryx splendens*. *Molecular Ecology*, **17**, 3496–3505.

Dasmahapatra KK, Mallet J (2006) DNA barcodes: recent successes and future prospects. *Heredity*, **97**, 254–255.

Dasmahapatra KK, Silva A, Chung J-W, Mallet J (2007) Genetic analysis of a wild-caught hybrid between non-sister *Heliconius* butterfly species. *Biology Letters*, **3**, 360–363.

Dasmahapatra KK, Lacy RC, Amos W (2008) Estimating levels of inbreeding using AFLP markers. *Heredity*, **100**, 286–295.

Elias M, Hill RI, Willmott KR *et al.* (2007) Limited performance of DNA barcoding in a diverse community of tropical butterflies. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **274**, 2881–2889.

Elias-Gutierrez M, Jeronimo FM, Ivanova NV, Valdez-Moreno M, Hebert PDN (2008) DNA barcodes for Cladocera and Copepoda from Mexico and Guatemala, highlights and new discoveries. *Zootaxa*, **1839**, 1–42.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Flot J-F, Tillier A, Samadi S, Tillier S (2006) Phase determination from direct sequencing of length-variable DNA regions. *Molecular Ecology Notes*, **6**, 627–630.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences, USA*, **103**, 968–971.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, **23**, 167–172.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 313–321.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. *PLOS Biology*, **2**, 10.

Hickerson M, Meyer CP, Moritz C (2006) DNA barcoding will often fail to discover new animal species over broad parameter space. *Systematic Biology*, **55**, 729–739.

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic, and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **272**, 1525–1534.

Lamas G, ed. (2004) Ithomiinae. In: *Atlas of Neotropical Butterflies. Checklist: 4A. Hesperioidea–Papilionoidea*, pp. 172–191. Scientific Publishers, Gainesville.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.

Madden JR, Lowe TJ, Fuller HV *et al.* (2004) Neighbouring male spotted bowerbirds are not related, but do maraud each other. *Animal Behaviour*, **68**, 751–758.

Mallarino R, Bermingham E, Willmott KR, Whinnett A, Jiggins CD (2005) Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Molecular Phylogenetics and Evolution*, **34**, 625–644.

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLOS Biology*, **3**, 2229–2238.

Monaghan MT, Balke M, Ryan Gregory T, Vogler AP (2005) DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, **360**, 1925–1933.

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution*, **14**, 389–394.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Rock J, Costa FO, Walker DD *et al.* (2008) DNA barcodes of fish of the Scotia Sea, Antarctica indicate priority groups for taxonomic and systematics focus. *Antarctic Science*, **20**, 253–262.

Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PDN (2006) DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proceedings of the National Academy of Sciences, USA*, **103**, 3657–3662.

Stoeckle MY, Hebert PDN (2008) Bar code of life: DNA tags help classify life. *Scientific American*, **299**, 82–88.

Takami Y, Koshio C, Ishii M *et al.* (2004) Genetic diversity and structure of urban populations of *Pieris* butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology*, **13**, 245–258.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Vekemans X (2002) *AFLP-SURV version 1.0. 2002*. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Whinnett A, Brower AVZ, Lee MM, Willmott KR, Mallet J (2005a) Phylogenetic utility of *Tektin*, a novel region for inferring systematic relationships amongst Lepidoptera. *Annals of the Entomological Society of America*, **98**, 873–886.

Whinnett A, Zimmermann M, Willmott KR *et al.* (2005b) Strikingly variable divergence times inferred across an Amazonian butterfly 'suture zone'. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **272**, 2525–2533.

Wiemers M, Fiedler K (2007) Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology*, **4**, doi:10.1186/1742-9994-4-8.

Will KW, Rubinoff D (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics*, **20**, 47–55.

Supporting information

Additional Supporting information may be found in the online version of this article.

Table S1 Details of samples and sampling localities

Table S2 PCR conditions for the amplicons used in this study

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.