MOLECULAR ECOLOGY

Molecular Ecology (2010) doi: 10.1111/j.1365-294X.2010.04802.x

The anatomy of a 'suture zone' in Amazonian butterflies: a coalescent-based test for vicariant geographic divergence and speciation

KANCHON K. DASMAHAPATRA,* GERARDO LAMAS,† FRASER SIMPSON* and JAMES MALLET*‡

*Department of Genetics, Evolution and Environment, University College London, 4 Stephenson Way, London NW1 2HE, UK, †Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Av. Arenales 1256, Apartado 14-0434, Lima-14, Peru; ‡Radcliffe Institute for Advanced Study, Byerly Hall, Harvard University, Cambridge, MA 02138, USA

Abstract

Attempts by biogeographers to understand biotic diversification in the Amazon have often employed contemporary species distribution patterns to support particular theories, such as Pleistocene rainforest refugia, rather than to test among alternative hypotheses. Suture zones, narrow regions where multiple contact zones and hybrid zones between taxa cluster, have been seen as evidence for past expansion of whole biotas that have undergone allopatric divergence in vicariant refuges. We used coalescent analysis of mutilocus sequence data to examine population split times in 22 pairs of geminate taxa in ithomiine and heliconiine butterflies. We test a hypothesis of simultaneous divergence across a suture zone in NE Peru. Our results reveal a scattered time course of diversification in this suture zone, rather than a tight cluster of split times. Additionally, we find rapid diversification within some lineages such as *Melinaea* contrasting with older divergence within lineages such as the Oleriina (*Hyposcada* and *Oleria*). These results strongly reject simple vicariance as a cause of the suture zone. At the same time, observed lineage effects are incompatible with a series of geographically coincident vicariant events which should affect all lineages similarly. Our results suggest that Pleistocene climatic forcing cannot readily explain this Peruvian suture zone. Lineage-specific biological traits, such as characteristic distances of gene flow or varying rates of parapatric divergence, may be of greater importance.

Keywords: coalescent theory, *Heliconius*, Ithomiini, phylogeography, Pleistocene refuges, speciation

Received 7 April 2010; revision received 15 July 2010; accepted 19 July 2010

Introduction

The highest terrestrial diversity on the planet is found in the Amazon basin, yet processes responsible for its evolution are unclear. There is a widely held view that most speciation occurs in geographic isolation, or 'allopatry' (Mayr 1963; Coyne & Orr 2004). However, the Amazon is a large, continuously forested region with few obvious geographic barriers. If allopatric speciation is the predominant mode of speciation, the simplest prediction would be that the Amazon should have low, rather than high, species diversity. Many theories have therefore been put forward to explain spatial patterns of diversification in the Amazon and other tropical regions, reviewed in (Haffer 1997, 2008), such as riverine barrier hypotheses (Patton & Da Silva 1998; Gascon *et al.* 2000), Pleistocene refuge hypotheses (Haffer 1969; Simpson & Haffer 1978; Haffer & Prance 2001), ecological (parapatric) hypotheses (Benson 1982; Endler 1982b), peripheral speciation (Fjeldså 1994), centrifugal speciation (Brown 1957) as well as vicariance hypotheses based on pre-Pleistocene geographic barriers (Cracraft & Prum 1988; Patton & Da Silva 1998; Hall & Harvey

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

2002). Among these, Pleistocene refuge hypotheses have received much attention from both supporters and critics (Endler 1982a; Haffer 1997; Moritz *et al.* 2000; Colinvaux *et al.* 2001; Carnaval *et al.* 2009).

The Pleistocene (1 800 000–10 000 years ago) is the most recent of a number of geological periods with cyclical climate changes during which there were repeated glaciations in the temperate zone with a periodicity of ~100 000 years. According to the Pleistocene refuge hypothesis (Haffer 1969), ice ages in the temperate zone led to periods of aridity in the tropics. Formerly continuous forest was hypothesized to become fragmented into refuges separated by expanses of grass-dominated savannah or desert. The forest biota was supposed then to have diverged during this period of confinement. After wetter conditions returned, the forest biota expanded from the refuges and came back into contact with formerly identical conspecifics. If reproductively isolated, these taxa could overlap as separate species. Alternatively if reproductive barriers were weak, the divergent taxa either fused or remained in contact at narrow hybrid zones.

The ability of the refuge hypothesis to give rise to allopatric speciation in an area where allopatry seemed at first sight unlikely, together with apparent correspondence of likely refuge areas with the contemporary patchwork biodiversity pattern in both temperate (Hewitt 2000) and tropical regions (Haffer 1969; Prance 1973; Brown 1987a), have contributed to its popularity. A Pleistocene refuge theory for divergence and diversification is widely and generally supported for the temperate zone. For example, refuges undoubtedly existed in Southern Europe during repeated glaciations, while Central and Northern Europe was scoured by a thick layer of ice (Hewitt 2000; Stewart *et al.* 2010). In contrast, applications of refuge theory to explain patterns of neotropical biodiversity remain contentious. Pleistocene refuges have been invoked to explain neotropical species distributions in birds (Haffer 1969), plants (Prance 1973, 1982) and butterflies (Brown 1977, 1987a). Yet, for a number of taxa, factors other than isolation in past refuges might also explain contemporary species distributions (Moritz *et al.* 2000). For example, current geographic distributions of Amazonian birds can be explained via a null model of random placement of species ranges better rather than via refuges (Beven *et al.* 1984), and putative centres of endemism for Amazonian plants correlate strongly with collecting density, rather than providing a strong biogeographic signal (Nelson *et al.* 1990).

Geological data can be used as evidence for existence of past refuges, just as they have to establish patterns and movement of ice in the temperate zone. However, geological data are harder to come by in the Amazon than almost anywhere else, and the palaeoecological implications are disputed. Pollen core data are commonly used to reconstruct past vegetation changes, but palynological records from Amazonia are scarce. Some studies claim evidence for great expansions of grass-dominated savannah during the Pleistocene based on fragmentary palynological and geomorphological data (Haffer & Prance 2001), but others argue that the forest overall remained continuous, rather than patchy as predicted by the refuge hypothesis. At the same time, changes in Amazonian pollen flora do indeed indicate changes in forest tree composition during temperature fluctuations correlated with temperate zone glaciation (Colinvaux *et al.* 2000, 2001).

Given that direct evidence from geology is still unclear, current patterns of distribution of organisms remain among the most powerful data available to test refuge theory in the Amazon. One class of biogeographic evidence that can suggest vicariant divergence (of which refuge theory is an example) is the existence of 'suture zones.' Suture zones are defined as narrow regions with unusual concentrations of contact zones and hybrid zones. They have been argued to represent the meeting places where whole biotas, having emerged after divergence from a pair of vicariant refuges, have begun to interact after climate amelioration (Remington 1968). Pleistocene climatic cycles have been implicated in the formation of a number of possible terrestrial suture zones, such as the Great Plains suture zone in North America where numerous avian hybrid zones cluster (Remington 1968; Moore & Price 1993; Klicka & Zink 1997), the clustered suture zones in the Alps and Central Europe (Hewitt 1996, 2000) and the Australian tropics (Moritz *et al.* 2009).

If divergence occurred during refuge formation, the multiple pairs of sister taxa across a suture zone might be expected to have split at a similar time, corresponding to the timing of the vicariant event itself (Coyne & Orr 2004). Divergence owing to known or probable single vicariant events has been studied mainly in marine systems, for example in species pairs between the Atlantic and Gulf coasts of USA (Avise 1994, 2000) and particularly across the Isthmus of Panama 2.7–3.5 Ma (Lessios 1979; Vawter *et al.* 1980; Bermingham & Lessios 1993; Knowlton *et al.* 1993; Knowlton & Weigt 1998; Hurt *et al.* 2009). Distribution patterns and biogeographic features such as suture zones by themselves may be insufficient to address biogeographic patterns with complex underlying evolutionary histories such as species diversification in the Amazon. Incorporation of molecular sequence data can help uncover useful temporal information about taxa which shed light on these complex histories (Elias *et al.* 2009). At the same time, phylogeographical analyses based on single genes and simple estimates of divergence such as sequence

difference may also be inadequate (Hurt *et al.* 2009; Hickerson *et al.* 2010). Importantly, even in the case of the Isthmus of Panama, where a single vicariant event is undisputed, recent analyses reveal that divergence of a number of species pairs predates the vicariant event (Hurt *et al.* 2009).

We here focus on a well-defined suture zone near Tarapoto in NE Peru (Fig. 1) between two areas of endemism in the lowland rainforest: the Río Mayo/upper Río Huallaga valley systems on the one hand (Departments of San Martín in the North and Huánuco further South), and the lower Río Huallaga and Río Ucayali regions on the other (Departments of San Martín (Eastern part), Loreto (Western part) in the North and Ucayali in the South). These two regions are separated by a low (mostly < 1500 m altitude) easternmost extension of the Andes, known as the Cordillera Escalera/Cordillera Azul. Around 40 forest-inhabiting butterfly species from tribe Ithomiini and subtribe Heliconiina (Nymphalidae) show morphologically differentiated geminate pairs of taxa across this suture zone (Fig. 1). The taxa consist both of subspecies which hybridize freely, or 'semi-species,' which do not (hereafter we refer to both collectively as 'subspecies' for convenience). Suture zones of this kind are argued to form the meeting place for separate biotas recently expanded from refuges, here from the putative 'Huallaga' (Western) and 'Ucayali' (Eastern) Pleistocene refuges (Brown *et al.* 1974; Brown 1979; Lamas 1982; Whinnett *et al.* 2005b). The Ithomini and Heliconiina contain about 350 and 70 species overall, respectively (Lamas *et al.* 2004, 2004), each of which consists of up to 30 morphologically recognizable subspecies distributed across the Neotropics (Brown 1979; Lamas *et al.* 2004). Distributions of heliconiine and ithomiine subspecies have previously been used to support and indeed to refine substantially Haffer's Pleistocene refuge model (Brown *et al.* 1974; Brown 1979, 1987b). Characteristically, each subspecies is a member of a different geographic Müllerian mimicry ring, varying in step with other similarly geographically differentiated taxa occurring in the same areas (Brown 1979).

Ithomiine subspecies pairs across this suture zone have previously been found to show variable levels of

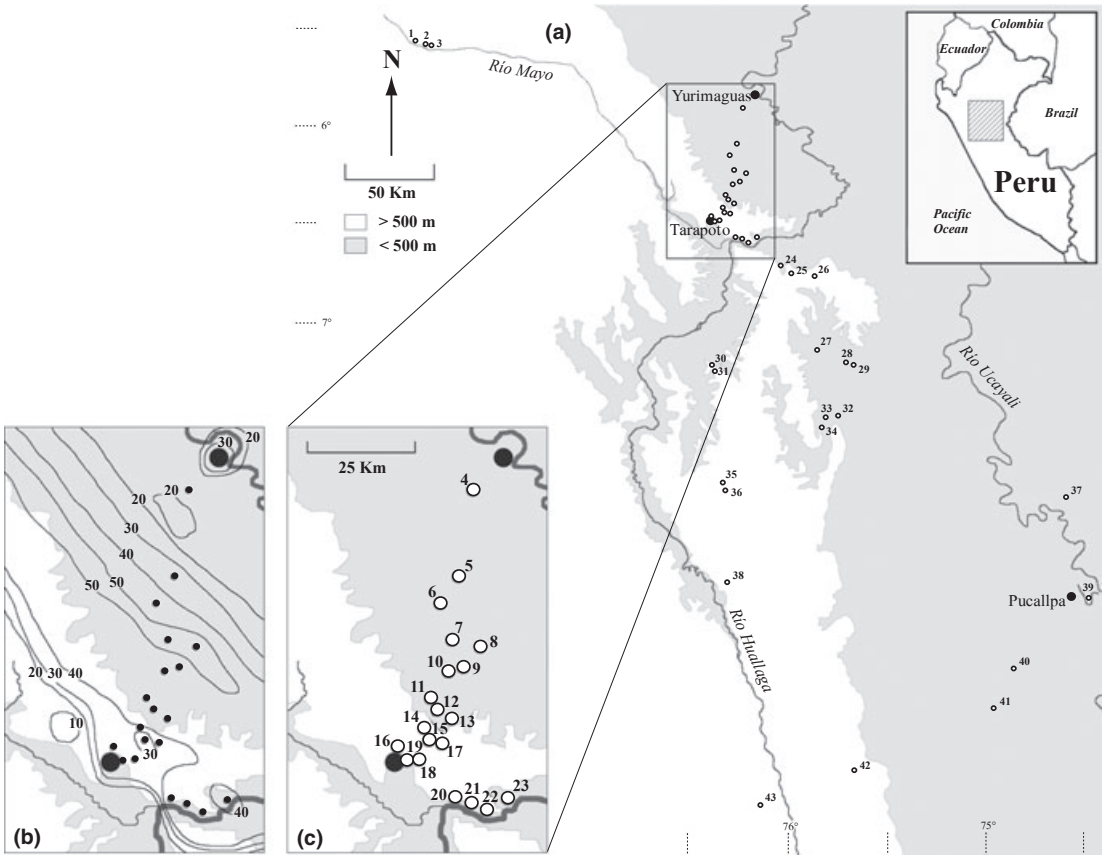

Fig. 1 (a) Map of sampling locations across the suture zone. Sampling location numbers refer to the site numbers in Table 2. (b) Contours of percentage of taxa hybridizing in the suture zone between the Río Mayo/upper Río Huallaga and the lower Río Huallaga/Río Ucayali regions of endemism. (c) Detail of collections made in the region of more intensive sampling.

mitochondrial (mtDNA) divergence, suggesting highly variable split times among different pairs of subspecies (Whinnett *et al.* 2005b). If variable nucleotide divergence is typical of the butterflies' genomes overall, a single recent forest vicariance event may not be responsible for divergence. However, several potential problems prevent drawing this conclusion from the previous work: (i) Mitochondrial DNA may evolve differently from the rest of the genome. For example, selective sweeps could be important in the nonrecombining mitochondrial genome, especially via indirect selection from cytoplasmic endosymbionts (Hurst & Jiggins 2005). (ii) Whinnett *et al.* (2005b) estimated split time via a simple mtDNA molecular clock. However, given that these taxa diverged during the last few million years, coalescence times of molecular markers may be far older than the time when the populations themselves split (Hudson & Turelli 2003). The overall DNA divergence between taxa includes both stochastic coalescence time within ancestral populations before the split and the time since the populations split (Hudson & Turelli 2003). The expected time to coalescence within a population of effective size N is $2N$ generations. As many Amazonian ithomiine and *Heliconius* species probably have large population sizes, the degree to which gene coalescence predates population divergence might be substantial. Coalescent-based methods must therefore be employed to account for variance in coalescence times.

In this study, we investigated multilocus DNA sequence divergence across the Tarapoto suture zone to test a null hypothesis of simultaneous vicariance using coalescent-based methods. As an alternative, cyclical climatic changes of the Pleistocene could have resulted in repeated forest expansions and contractions centred on the same refuges, and a more complex pattern might be observed. These and other hypotheses are dealt with in the discussion. We improve on the previous data and analyses (Whinnett *et al.* 2005b) as follows:

1 We obtain molecular sequence information from 22 pairs of ithomiine and heliconiine taxa differing morphologically at species or subspecies level across the suture zone, doubling the number of taxon pairs studied previously.
2 We double the mitochondrial sequence data per specimen to ~2100 bp
3 We add three nuclear loci: sex-linked *Tpi* (~800 bp), as well as autosomal *Tektin* (~700 bp) and *Rpl5* (~650 bp).
4 We obtain sequence data from larger numbers of individuals per subspecies.
5 We exploit more rigorous coalescent-based analyses to estimate relative split time, while allowing for variation in effective ancestral population size (θ) and stochastic coalescence times, instead of using clock-based simple nucleotide divergence as estimators of split time

Methods

Ithomiine and heliconiine samples of 22 species, each with morphologically divergent subspecies across the suture zone (Table 1), were collected (Fig. 1 and Table S1). These comprised 17 species from nine ithomiine genera and a further five *Heliconius* species. Collections were made in 2002, 2004 and 2005. All specimens were identified by GL and JM. Wings were removed from specimens and kept in transparent envelopes for identification, and remaining tissue was preserved in salt-saturated DMSO and stored at −20 °C. Both wings and tissue are held at UCL. Whenever possible, at least three individuals from each subspecies were sampled (Table S1) although for a few rarer subspecies, such as *Scada reckia junina* and *Melinaea satevis tarapotensis*, we obtained only 1–2 specimens. Representative outgroup taxa were also used (Table 1).

DNA was extracted from one-third of the thorax of each specimen using the DNeasy Blood and Tissue Kit (QIAGEN), and DNA extracts stored at −20 °C. We amplified ~2130 bp of mtDNA comprising *cytochrome oxidase I* (*CoI*), *tRNA-leu* and the 5'-end of *cytochrome oxidase II* (*CoII*), as well as three nuclear loci. The nuclear loci were *Tektin* (700–740 bp), *ribosomal protein L5* (*Rpl5*) (400–900 bp) and *triose phosphate isomerase* (*Tpi*) (690–770 bp) and were chosen as they have been found to have relatively rapid rates of evolution and are readily amplifiable in both the Ithomiini and *Heliconius* species (Mallarino *et al.* 2005; Whinnett *et al.* 2005a; Dasmahapatra *et al.* 2007, 2010; Elias *et al.* 2009; de-Silva *et al.* 2010). Details of PCR primers and reaction conditions are provided in Tables S2 and S3. PCR products were cleaned and cycle sequenced from both ends using the Big Dye Terminator 3.1 Cycle Sequencing Kit (Applied Biosystems). Cycle sequenced products were analysed on a 3730xl Genetic Analyzer (Applied Biosystems). Chromatograms were checked and edited with ChromasPro 1.41 (Technelysium Pty Ltd). Sequence alignment was carried out using ClustalX and checked manually. Both *Rpl5* and *Tpi* spanned several introns and varied in length owing to indels. When indel heterozygosity resulted in variable-sized products from the same individual, indels could often be identified, and the two alleles 'deconvoluted' and haplotyped using information from the double peak signal reads beyond the indel after bidirectional sequencing (Flot *et al.* 2006; Dasmahapatra *et al.* 2007). Sequence data were lodged with Genbank (HM051677–HM052795,

Table 1 The number of sequences of the four loci used in each of the pairs of subspecies for the 22 species across the suture zone

Species	Subspecies pair		mtDNA		*Tektin*		*Rpl5*		*Tpi*		Outgroup taxon used
			ssp1	ssp2	ssp1	ssp2	ssp1	ssp2	ssp1	ssp2	
Heliconius erato	*favorinus*	*emma*	4	6	4	5	4	7	10	7	*Heliconius demeter ucalayensis*
Heliconius melpomene	*amaryllis*	*aglaope*	13	4	4	5	5	3	7	8	*Heliconius ethilla aerotome*
Heliconius pardalinus	*sergestus*	*dilatus*	4	4	3	3	4	2	5	3	*Heliconius ethilla aerotome*
Heliconius numata *	*tarapotensis* + *bicoloratus*	*aurora* + *silvana*	4	7	4	6	4	6	2	5	*Heliconius ethilla aerotome*
Heliconius demeter	*ssp. nov.*	*ucayalensis*	7	13	3	5	8	11	7	13	*Heliconius sara sara*
Melinaea marsaeus	*mothone*	*phasiana*	7	22	10	20	9	13	8	16	*Melinaea ludovica ludovica*
Melinaea satevis	*tarapotensis*	*cydon*	2	7	1	5	2	6	2	6	*Melinaea ludovica ludovica*
Melinaea menophilus	*ssp. nov.*	*hicetus*	22	5	22	5	12	3	17	4	*Melinaea ludovica ludovica*
Oleria onega	*ssp. nov.*	*janarilla*	27	64	11	9	5	3	10	8	*Oleria gunilla serdolis*
Oleria gunilla	*serdolis*	*lota*	11	5	7	3	6	7	10	7	*Oleria onega janarilla*
Hyposcada kena	*ssp. nov.*	*flexibilis*	3	5	3	5	4	5	4	5	*Hyposcada anchiala mendaz*
Hyposcada anchiala	*mendax*	*kezia*	5	4	5	5	9	7	4	6	*Hyposcada kena flexibilis*
Hyposcada illinissa	*ssp. nov.*	*margarita*	6	8	6	7	6	7	5	4	*Hyposcada anchiala mendax*
Scada zibia	*quotidiana*	*batesi*	12	4	6	7	5	4	7	6	*Scada reckia junina*
Scada reckia	*ethica*	*junina*	1	4	1	4	1	4	1	3	*Scada zibia quotidiana*
Brevioleria aelia	*quadrona*	*pachiteae*	12	7	10	7	10	6	12	6	*Brevioleria arzalia arzalia*
Brevioleria arzalia	*ssp. nov.*	*arzalia*	3	3	3	4	2	4	4	4	*Brevioleria aelia*
Napeogenes pharo	*lamia*	*pharo*	7	8	4	6	9	8	2	6	*Napeogenes larina otaxes*
Napeogenes sylphis	*rindgei*	*corena*	11	6	8	3	11	9	8	6	*Napeogenes inachia pozziana*
Ithomia salapia	*derasa*	*aquinia*	7	8	6	10	7	12	6	10	*Ithomia lagusa peruana*
Pseudoscada florula	*gracilis*	*aureola*	4	3	9	6	7	5	7	6	*Pseudoscada timna timna*
Mechanitis mazaeus	*deceptus* + *cf. phasianita*	*fallax*	13	8	10	7	12	12	16	7	*Scada zibia quotidiana*†
Total number of sequences			390		277		286		300		

**Heliconius numata* is locally polymorphic (Joron *et al.* 2001).
†An appropriate outgroup within *Mechanitis* is not available because of uncertain relationships among *Mechanitis* species (Dasmahapatra *et al.* 2010).

Table S1). In addition, pre-existing data from Genbank were also used (Table S1).

Approximate Bayesian computational methods have been developed to test for simultaneous divergence of taxa taking into account variance in coalescence times (Hickerson *et al.* 2006). However, currently msBayes (Hickerson *et al.* 2007) does not allow for variation in substitution rates between taxa, a feature of our data set. Instead, we employed the coalescent-based Bayesian program MCMCcoal 1.2 (Yang 2002; Rannala & Yang 2003) to estimate parameters relevant to neutral coalescence for each subspecies pair, followed by the scaling of parameter estimates to account for this variation in substitution rates. The program allows for the estimation of split times for a pair of taxa ($\tau = t\mu$, the product of split time in generations and substitution rate per generation) as well as of ancestral and current effective population sizes ($\theta = 4N\mu$, the product of genetically effective population size and substitution rate). Both split time and effective population size parameters are scaled by substitution rate μ, which under the neutral theory is expected to equal the mutation rate. To correct for among-lineage variation in substitution rate, we arranged our taxa into groups of three consisting of a

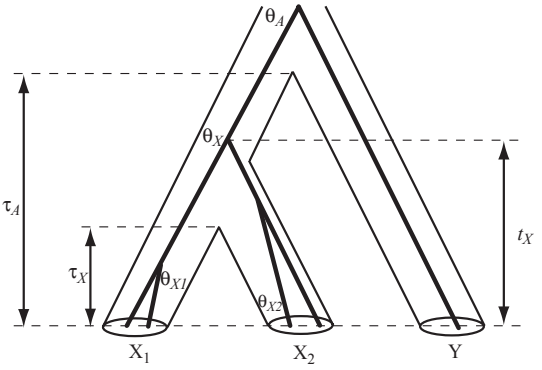

Fig. 2 A group of three taxa comprising a subspecies pair, X_1 and X_2, and an outgroup taxon, Y, used in the MCMCcoal analysis. Six parameters were estimated from each subspecies pair: θ_{X_1}, θ_{X_2}, θ_X, θ_A, τ_X and τ_A. For recently diverged pairs of taxa, the coalescence time, t_X, may be considerably longer than the split time τ_X.

Table 2 Details of sampled locations shown in Fig. 1

Site no.	Site Name	Latitude	Longitude
1	Jorge Chávez, San Martín	5°39′43″S	77°43′58″W
2	Puente Serranoyacu, San Martín	5°40′31″S	77°40′28″W
3	Puente Aguas Verdes, San Martín	5°41′16″S	77°38′00″W
4	Km 103.1 Tarapoto-Yurimaguas, San Martín	5°58′18″S	76°13′55″W
5	Km 78 Tarapoto-Yurimaguas, San Martín	6°08′58″S	76°16′00″W
6	Santa Rosa de Davidcillo, San Martín	6°14′57″S	76°16′30″W
7	Convento, San Martín	6°16′22″S	76°17′21″ W
8	Km 7.2 Pongo-Barranquita, San Martín	6°17′12″S	76°13′54″W
9	Río Shucshuyacu, Pongo del Cainarachi, San Martín	6°20′07″S	76°15′15″W
10	Km 43–46 Tarapoto-Yurimaguas, San Martín	6°22′36″S	76°17′01″W
11	Km 22 Tarapoto - Yurimaguas, San Martín	6°24′30″S	76°19′30″W
12	Km 24 Tarapoto-Yurimaguas, San Martín	6°24′58″S	76°18′24″W
13	Km 35 Tarapoto-Yurimaguas, San Martín	6°25′29″S	76°15′02″W
14	Km 21 Tarapoto-Yurimaguas, San Martín	6°24′23″S	76°18′07″W
15	La Antena, San Martín	6°27′18″S	76°17′54″W
16	Bocatoma Río Shilcayo, San Martín	6°27′27″S	76°20′58″W
17	Túnel/Biodiversidad, San Martín	6°27′30″S	76°17′15″W
18	Km 8 Tarapoto-Yurimaguas, San Martín	6°27′43″S	76°19′23″W
19	Urahuasha, San Martín	6°27′60″S	76°20′05″W
20	Shapaja, San Martín	6°34′47″S	76°15′46″W
21	Chumía, San Martín	6°36′57″S	76°11′10″W
22	Río Pucayaquillo	6°35′10″S	76°13′05″W
23	Chasuta, San Martín	6°34′28″S	76°08′12″W
24	Robashca, PNCAZ, San Martín	6°42′38″S	76°01′60″W
25a	Camp 1 Robashca-Quebrada Yanayacu, PNCAZ, Loreto	6°44′27″S	75°58′54″W
25b	Camp 2 Quebrada Yanayacu, PNCAZ, Loreto	6°44′55″S	75°56′23″W
26	Laguna del Mundo Perdido, San Martín	6°45′05″S	75°52′08″W
27	Boca del Ushpayacu, Loreto	7° 05′ 40″S	75°51′30″W
28	Río Pauya/ Río Cushabatay confluence, PNCAZ, Loreto	7°09′18″ S	75°43′24″W
29	Quebrada Paco, Loreto	7°09′49″S	75°41′21″W
30	Km 22, Nuevo Lima - La Perla del Ponacillo, San Martín	7°10′04″S	76°20′42″W
31	Km 22, Nuevo Lima - Selva Andina, San Martín	7°11′36″S	76°20′06″W
32	Quebrada Huacanqui, Loreto	7°24′13″S	75°46′11″W
33	Pongo del Río Pauya, Loreto	7°24′19″S	75°49′11″W
34	Cerro Mira Culo, Loreto	7°27′12″S	75°50′16″W
35	Cachatigre, Río Biabo, PNCAZ	7°43′05″S	76°20′51″W
36	Caño Negro, Río Biabo, PNCAZ	7°45′10″S	76°20′03″W
37	Quebrada Tunuya, San Jeronimo, Loreto	7°46′32″S	74°41′20″W
38	Cachiyacu, San Martín	6°26′00″S	76°23′00″W
39a	Caño Tushmo, Lago Yarinacocha, Ucayali	8°20′32″S	74°35′31″W
39b	El Pandisho, Lago Yarinacocha, Ucayali	8°19′02″S	74°35′28″W
40	IVITA, Ucayali	8°38′28″S	74°57′13″W
41	Bosque von Humboldt, Ucayali	8°49′48″S	75°03′28″W
42	8 km west of Boquerón del Padre Abad, Ucayali	9°06′13″S	75°44′28″W
43	Villa Jennifer, Tingo María, Huánuco	9°16′11″S	76°00′29″W

subspecies pair (X_1, X_2) belonging to species X, together with a single representative of a closely related outgroup species Y (Fig. 2 and Table 1). Using this model, MCMCcoal was set to estimate six parameters: scaled population sizes for the subspecies pairs, θ_{X_1}, θ_{X_2}, and of their ancestors, θ_X, θ_A, as well as scaled split times τ_X and τ_A; Fig. 2). In each analysis, the major parameter of interest was τ_X, the population split time for the focal pair of subspecies.

Heredity multipliers (inheritance scalars) of 0.25, 0.75, 1.00 and 1.00 were used in MCMCcoal for mtDNA, *Tpi*, *Tektin* and *Rpl5* loci, respectively, as correction factors for θ in mtDNA and sex-linked *T*pi, relative to autosomal loci (Yang 2007a). Relative substitution rates (μ) of the four loci were estimated by comparing maximum likelihood ithomiine and *Heliconius* tree lengths for each locus using PAML (Yang 2007b). Tree topologies used for relative rate estimation were based on (Brower *et al.*

2006) for ithomiines and (Beltrán *et al.* 2007) for *Heliconius*; taxa used for this estimation are shown in Fig. S1. Relative substitution rates of mtDNA, *Tpi*, *Tektin* and *Rpl5* were estimated as 1.00:1.79:0.80:1.23 (ithomiines) and 1.00:1.02:0.31:0.57 (*Heliconius*). By setting the mitochondrial substitution rates to one, the rate of substitution at mtDNA is assumed to be the same in the ithomiine and heliconiine lineages. This allowed the comparison of parameter estimates among the groups. For each of the six estimated parameters (Fig. 2), MCMCcoal requires gamma distributed priors specified by parameters α and β, which determine the mean (α/β) and variance (α/β^2) of the prior distributions. In these analyses, we employed permissive priors for all estimated parameters to minimize any strong influence of the prior on the posterior distribution. Some subspecies pairs show few fixed sequence differences, suggesting recent divergence. Therefore, for both τ estimates we set priors with $\alpha < 1$, which allow $\tau = 0$. The α and β values for the two τ estimates were chosen to reflect the fact that $\tau_X < \tau_A$ (Fig. 2). Thus, α and β were set to 1.1 and 28 respectively for all four θ priors, 0.7 and 3.5 for τ_X, and 0.7 and 35 for τ_A. To avoid bias, identically permissive prior distributions were specified for all 22 subspecies pairs.

MCMCcoal analyses were conducted using a burn-in of 20 000 iterations, following which 20 000 samples were taken from the posterior distribution with a sample interval of 10. 'Fine-tune parameters' were optimized for each subspecies pair to maintain acceptance proportions with the interval of (0.15, 0.7) as suggested in the manual (Yang 2007a). Output stabilization was examined to ensure the reliability of parameter estimates. Each analysis was run twice with different random seeds to check consistency of parameter estimates and assess convergence. In all cases, parameter estimates converged on the same values in different runs.

The trees and sequences used to estimate relative substitution rates μ of each locus (Fig. S1) were also used to test for variation in evolutionary rate among lineages, or deviation from molecular clock, using PAUP* 4.0b10 (Swofford 2003). Deviation from a molecular clock was found within both Ithomiini (mtDNA: $2\Delta\ln L = 81$, $P < 10^{-5}$; *Tektin*: $2\Delta\ln L = 86$, $P = 10^{-6}$ with d.f. = 32) and *Heliconius* (mtDNA: $2\Delta\ln L = 43$, $P = 10^{-4}$; *Tektin*: $2\Delta\ln L = 26$, $P = 0.007$; *Tpi*: $2\Delta\ln L = 56$, $P = 10^{-7}$ with d.f. = 11). Thus, it was desirable to correct the estimates of τ obtained from MCMCcoal to reflect relative absolute time among lineages.

The 'local clock' model implemented in PAML (Yoder & Yang 2000; Yang & Yoder 2003) was used to estimate relative substitution rates for each of the four loci in different branches of the ithomiine and heliconiine trees. Within the Ithomiini, separate relative substitution rates were calculated simultaneously for each of the nine genera using sequence data from 35 species and subspecies (Fig. S1). Within *Heliconius*, sequences from 13 species and subspecies were used to calculate separate substitution rates for each locus for *melpomene-wallacei-hierax* and *erato-sara* groups (Fig. S4). In each lineage involving a subspecies pair, rates for the four loci were averaged to provide local rates of substitution which were then used to scale local estimates of τ obtained with MCMCcoal. Lineage correction factors are shown in Table 3. The estimated substitution rates varied rather strongly with a standard deviation of 25%.

Estimates of split time between subspecies pairs, τ_X, obtained from MCMCcoal were compared with uncorrected average pairwise sequence divergence calculated in MEGA 4.0 (Tamura *et al.* 2007) between each subspecies in a pair. Values of τ_X from MCMCcoal were also compared with those from IMa, another coalescent-based Bayesian analysis program (Nielsen & Wakeley 2001; Hey & Nielsen 2007). For the purposes of this investigation, migration rates were set to zero in IMa. As we regularly found some sites with three states in these analyses, we used the Hasegawa-Kishino-Yano (HKY) (Hasegawa *et al.* 1985) substitution model for all loci rather than the infinite sites model. (MCMCcoal employs a similar Jukes-Cantor substitution model which allows multiple changes at a single site). Inheritance scalars and relative substitution rates for each locus were set as in MCMCcoal. The truncated uniform priors are specified differently from those in MCMCcoal. Following trial runs, the upper bound for the priors on θ_{X_1}, θ_{X_2}, θ_X were set at 15, while the upper bound for τ was set at 10. Metropolis-coupling was implemented using five chains with a two step heating scheme (Hey 2009). A burn-in of a million steps was used, and effective sample size values were examined to assess convergence, following which the posterior distribution was sampled for an hour with a sample interval of 10 (the number of steps sampled varied between 170 000 and 670 000). Each data set was analysed twice to ensure consistent parameter estimates.

To investigate differential patterns of divergence among lineages further, gene genealogies of example genera *Hyposcada* and *Melinaea* were estimated. Each of these two genera includes three species showing morphological subspecies differences across the Huallaga suture zone. After hierarchical likelihood ratio tests implemented within MODELTEST 3.7 (Posada & Crandall 1998) were used to select the most likely model of sequence evolution, maximum likelihood trees were constructed using PAUP* (Swofford 2003) from concatenated sequences of all four loci using the GTR+I+G model. In addition, for each genus, separate

Table 3 Estimates of subspecies split times (τ_X) and effective population size (θ) estimates derived from MCMCcoal. All estimates are scaled by the neutral mutation rate, μ, calculated relative to the per site mutation rate of mtDNA. Figures in parentheses are 95% credibility intervals. The lineage correction factors are the multipliers used to correct τ_X for among-lineage variation in substitution rate

Species	τ_X	θ_{X_1}	θ_{X_2}	θ_X	Lineage correction	Corrected τ_X
Heliconius demeter	0.003983 (0.002228–0.006391)	0.014496 (0.007755–0.027739)	0.021873 (0.013729–0.035694)	0.041461 (0.023856–0.073746)	1.2	0.004780 (0.002674–0.007669)
Oleria onega	0.003683 (0.002152–0.005274)	0.056224 (0.040405–0.078309)	0.076784 (0.052647–0.118445)	0.041189 (0.021564–0.070165)	1.0	0.003683 (0.002152–0.005274)
Heliconius pardalinus	0.003250 (0.001310–0.006747)	0.005568 (0.001858–0.019430)	0.009163 (0.004195–0.022157)	0.025293 (0.009029–0.073215)	1.1	0.003575 (0.001441–0.007422)
Heliconius erato	0.002856 (0.001512–0.004178)	0.086824 (0.040903–0.187583)	0.090473 (0.041586–0.200876)	0.053195 (0.031076–0.087631)	1.2	0.003427 (0.001814–0.005014)
Napeogenes pharo	0.002310 (0.000913–0.003864)	0.029507 (0.015746–0.073117)	0.027792 (0.013313–0.062165)	0.030085 (0.018087–0.018087)	1.3	0.003003 (0.001187–0.005023)
Scada reckia	0.002648 (0.001062–0.005696)	0.004932 (0.002143–0.012130)	–	0.018092 (0.004333–0.056524)	1.1	0.002913 (0.001168–0.006266)
Hyposcada anchiala	0.002527 (0.001238–0.003805)	0.002486 (0.001179–0.005338)	0.002757 (0.001353–0.005901)	0.002478 (0.000151–0.013348)	1.0	0.002527 (0.001238–0.003805)
Napeogenes sylphis	0.001401 (0.000562–0.002162)	0.009124 (0.004297–0.020502)	0.024823 (0.013378–0.050994)	0.010292 (0.004769–0.019603)	1.3	0.001821 (0.000731–0.002811)
Pseudoscada florula	0.002540 (0.001422–0.003649)	0.020931 (0.009993–0.051900)	0.018260 (0.009491–0.038680)	0.030907 (0.011878–0.065928)	0.7	0.001778 (0.000995–0.002554)
Hyposcada kena	0.001773 (0.000740–0.003246)	0.002135 (0.000915–0.004884)	0.000413 (0.000055–0.001690)	0.023967 (0.009023–0.074433)	1.0	0.001773 (0.000740–0.003246)
Mechanitis mazaeus	0.001163 (0.000700–0.001923)	0.007591 (0.004171–0.014740)	0.022701 (0.012891–0.049628)	0.015580 (0.009927–0.025424)	1.2	0.0001396 (0.000840–0.002308)
Heliconius numata	0.001215 (0.000651–0.001929)	0.079231 (0.029334–0.209320)	0.052320 (0.014840–0.166982)	0.014919 (0.008365–0.026411)	1.1	0.001337 (0.000716–0.002122)
Scada zibia	0.001165 (0.000535–0.001900)	0.033939 (0.011897–0.122042)	0.017957 (0.008800–0.037615)	0.015727 (0.008849–0.026913)	1.1	0.001282 (0.000589–0.002090)
Oleria gunilla	0.001177 (0.000611–0.001980)	0.004542 (0.002086–0.010743)	0.007541 (0.003952–0.015641)	0.004193 (0.001750–0.009067)	1.0	0.001177 (0.000611–0.001980)
Ithomia salapia	0.000753 (0.000393–0.001388)	0.008974 (0.004600–0.021947)	0.003587 (0.001687–0.008238)	0.010504 (0.006053–0.018711)	1.3	0.000979 (0.000511–0.001804)
Heliconius melpomene	0.000425 (0.000030–0.000974)	0.060535 (0.014381–0.178497)	0.018014 (0.006128–0.085410)	0.029296 (0.018691–0.045681)	1.1	0.000468 (0.000033–0.001071)
Hyposcada illinissa	0.000378 (0.000174–0.000703)	0.029416 (0.005346–0.135927)	0.015179 (0.003276–0.099033)	0.003877 (0.002126–0.006908)	1.0	0.000378 (0.000174–0.000703)
Melinaea marsaeus	0.000678 (0.000401–0.001068)	0.017304 (0.010347–0.028736)	0.005319 (0.002843–0.010072)	0.052688 (0.033262–0.089017)	0.5	0.000339 (0.000201–0.000534)
Melinaea satevis	0.000554 (0.000044–0.001833)	0.021756 (0.007108–0.102169)	0.006909 (0.000870–0.081087)	0.022020 (0.013147–0.039588)	0.5	0.000277 (0.000022–0.000917)
Brevoleria arzalia	0.000294 (0.000010–0.000995)	0.028373 (0.004151–0.137638)	0.023330 (0.002771–0.124083)	0.008382 (0.002987–0.017189)	0.8	0.000235 (0.000008–0.000796)
Melinaea menophilus	0.000229 (0.000009–0.000620)	0.030204 (0.004407–0.137586)	0.028671 (0.009503–0.093340)	0.027408 (0.018571–0.040049)	0.5	0.000115 (0.000005–0.000310)
Brevoleria aelia	0.000038 (0.000001–0.000143)	0.033757 (0.003731–0.144522)	0.039463 (0.005265–0.152316)	0.004094 (0.002346–0.006654)	0.8	0.000030 (0.000000–0.000114)

bootstrapped neighbour-joining trees for each locus were constructed using MEGA 4.0 (Tamura *et al.* 2007) to investigate any locus-specific effects.

Results

A total of 1282 mitochondrial and nuclear sequences (390 mtDNA, 277 *Tektin*, 286 *Rpl5*, 300 *Tpi*) comprising ~1.5 Mb total bases from 399 specimens within the 30 species (Tables 1 and S1) were used for the analyses. The average coverage was seven haplotypes sequenced per locus for each subspecies (Table 1).

As explained previously, MCMCcoal estimates of split time for the 22 subspecies pairs, τ_X, were corrected for variation among lineages in substitution rate. The resultant estimates of τ_X varied more or less continuously from near zero (*Brevioleria aelia* and *Melinaea menophilus*) up to 0.005 (*Heliconius demeter* and *Oleria onega*) with little tendency towards clustering (Fig. 3). A very similar pattern of scattered split times is found even for uncorrected τ_X values, suggesting problems with correction are not the cause (Table 3). If the suture zone formed as a result of contact between biotas separated by a simple vicariant event, MCMCcoal divergence estimates for the 22 subspecies pairs are expected to be clustered around a single value. The 95% credibility intervals (i.e. the interval containing 95% of the posterior probability) of all 22 divergence estimates should overlap. In this study, 95% credibility intervals for τ_X were wide, a consequence of using only four loci; however, a significant fraction is nonoverlapping (Fig. 3), strongly indicating that not all pairs of taxa split at the same time, thereby implicating a problem with a single split time.

In addition to the assessment based on credibility intervals, we carried out two *ad hoc* tests of significance to gauge statistically whether our estimates could have resulted from a single vicariance event. In the first *ad hoc* test, we repeated the initial MCMCcoal analyses but constrained τ_X by specifying a very narrow prior centred about the same average τ_X ($\tau_X = 0.001695$) for all 22 subspecies pairs (by setting priors on τ_X: $\alpha = 20000$, $\beta = 11900000$, giving a very low prior variance of 2×10^{-10}). Priors for all other parameters were left unchanged. For each taxon pair, the difference in maximum likelihood obtained in MCMCcoal runs between constrained and unconstrained τ_X was used to carry out likelihood ratio tests. As expected, in most (17 of the 22) comparisons the unconstrained calculations showed a higher likelihood than the constrained calculations, seven of which were significantly higher likelihoods after Bonferroni correction ($-2\Delta\ln L > 9.32$, d.f. = 1, $P < 0.0023$; Table S4). Alternatively, summing likelihoods from unconstrained and constrained runs across all 22 species indicates strongly ($-2\Delta\ln L = 263$, d.f. = 22, $P \approx 10^{-42}$) that the same split time does not apply to all species.

The second *ad hoc* test was carried out in two stages. First, using MCMCoal a single τ_X for all 22 taxon pairs was estimated combining sequence data from all species in a single analysis: effectively calculating a single τ_X of 0.000922 across the suture zone from 88 loci (four loci from 22 species). Then, individual analyses for each taxon pair were repeated with τ_X constrained by a narrow prior to be centred about the single divergence value obtained in the first stage (the prior on τ_X was forced to have a median of 0.000922 using $\alpha = 20\,000$, $\beta = 22\,000\,000$, giving a very low prior variance of

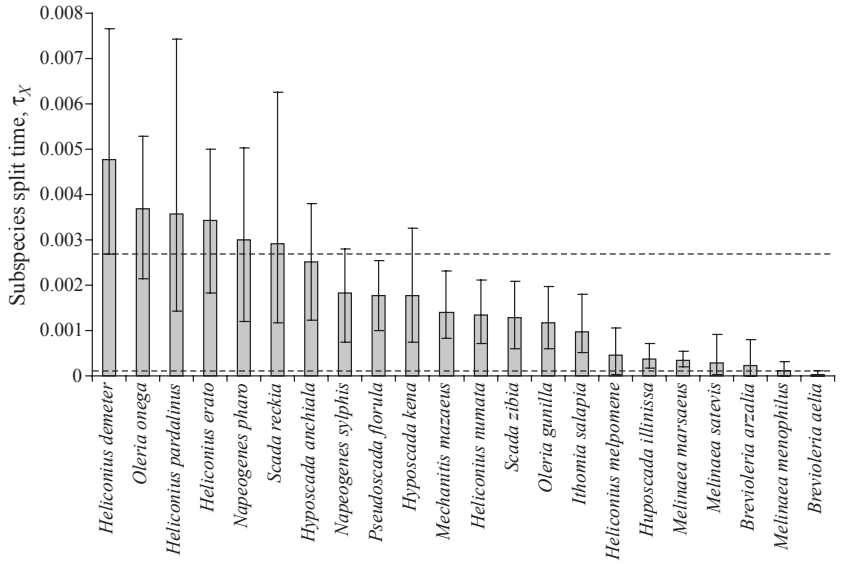

Fig. 3 MCMCcoal divergence estimates (τ_X, the product of absolute split time and substitution rate) for the 22 subspecies pairs across the suture zone. Split times are scaled to allow for lineage-specific substitution rate variation. For each species, the median τ_X is shown with 95% credibility intervals. The upper and lower dashed lines correspond to the lower credibility interval for the largest value of τ_X (*Heliconius demeter*) and the higher credibility interval for the smallest τ_X (*Brevioleria aelia*). These highlight the lack of overlap among estimates of split time among subspecies pairs.

4×10^{-11}), followed by likelihood ratio testing between constrained and unconstrained τ_X as before. No outgroup species were used in this second *ad hoc* test. After Bonferroni correction, in three of the 22 comparisons the unconstrained calculations showed a significantly higher likelihood than the constrained calculations (Table S4). However, summing likelihoods from unconstrained and constrained runs across all 22 species again indicated strongly ($-2\Delta\ln L = 85.6$, d.f. = 22, $P \approx 10^{-8}$) that a single split time cannot be applied across all species pairs.

It should be noted that these *ad hoc* tests are not strictly valid as the MCMC procedure is not intended to generate maximum likelihood estimates but to explore the region around the likelihood peak in proportion to posterior probability. In the likelihood ratio tests described, we employed the value showing the maximum likelihood from the 20 000 iterations sampled: this will rarely be the overall maximum likelihood value. This explains why not all likelihoods are greater in unconstrained than constrained analyses (although unconstrained likelihoods were never significantly less than constrained likelihoods, Table S4). The visual inspection of 95% credibility provides no quantified statistical test but is in theory more appropriate and intuitive. Nonetheless, although the tests performed here are somewhat informal, our *ad hoc* approach is today used widely in similar genealogical problems, for example to test the null hypothesis of zero gene flow among populations (Hey & Nielsen 2007).

The estimates of τ_X for the taxon pairs using the two different coalescent programs, MCMCcoal and IMa, were strongly correlated (Fig. 4a, $r^2 = 0.86$, $P = 10^{-10}$, 20 d.f.), demonstrating that the programs recovered similar information from the data in spite of different prior specification. Split time τ_X (from MCMCcoal) is also correlated with raw net sequence divergence among subspecies pairs at mtDNA (Fig. 4b, $r^2 = 0.56$, $P = 6 \times 10^{-5}$, 20 d.f.), and also, more weakly, at *Tpi* (Fig. 4c, $r^2 = 0.20$, $P = 0.04$, 20 d.f.). We found no significant correlations between τ_X and sequence divergence at either *Tektin* or *Rpl5* (Fig. 4d, e). The weaker correlation of MCMCcoal τ_X with sequence divergence than with IMa τ_X suggests that similar additional information about split time is recovered by both coalescent-based analysis programs and that this information is lacking in raw sequence divergence. There were no clear differences in the estimated effective population sizes of pairs of subspecies (θ_{X_1} and θ_{X_2}): 95% credibility intervals overlapped strongly in all subspecies pairs (Table 3). θ_{X_1} and θ_{X_2} estimates were strongly correlated with one another ($r^2 = 0.67$, $P = 1 \times 10^{-5}$, 19 d.f.) showing that estimated effective population sizes of pairs of subspecies are probably similar to each other (Table 3).

A comparison of *Hyposcada* and *Melinaea* maximum likelihood trees (Fig. 5) and neighbour-joining trees at each locus (Fig. 6) demonstrates the effect of more recent species and subspecies diversification within the latter. Within *Hyposcada,* both conspecifics and subspecies form monophyletic clusters separated by relatively long branches (Figs 5 and 6). In contrast, within *Melinaea* the branches are shorter, and species (*M. marsaeus* and *M. satevis*) and subspecies paraphyly is widespread (Figs 5 and 6). These patterns are present across all loci (Fig. 6). This is probably a result of incomplete lineage sorting owing to the recency of diversification and larger ancestral population sizes in *Melinaea* (θ_X for *M. satevis* = 0.022, *M. menophilus* = 0.027, *M. marsaeus* = 0.053) compared to in *Hyposcada* (θ_X for *H. anchiala* = 0.0025, *H. illinissa* = 0.0039, *H. kena* = 0.024) (Table 3).

Discussion

Suture zones form the boundaries between centres of endemism under vicariance theories of diversification, with centres of endemism interpreted as likely sites for former Pleistocene refuges in heliconiine and ithomiine butterflies (Brown *et al.* 1974; Brown 1979, 1987b). Here, we use coalescent-based analysis of mutilocus sequence data from 22 pairs of taxa across an Amazonian butterfly suture zone to test a hypothesis of simultaneous split times. Our analysis clearly reveals a scattered time course of diversification, rather than a tight cluster of split times as predicted by a simple vicariant event. These results suggest that diversification is more complex than that suggested by the simplistic, single-split vicariance model.

With molecular technology and improved methods of analysis rapidly advancing, we now have tools to investigate the effects of coalescence time on overall divergence among pairs of related taxa. The time to coalescence, t_X, of two gene sequences now present in two descendant species is always longer than the time back to the split of the two species, τ_X (Fig. 2), because coalescence must always occur in the ancestor some time before species divergence. As the species split time increases, the discrepancy between the two diminishes as the fraction of the total time to coalescence within the ancestral population becomes smaller relative to the time within the descendants. For recently diverged populations with large effective population sizes, as is likely for many Amazonian butterflies, the discrepancy can be substantial (Edwards & Beerli 2000; Nichols 2001; Beaumont *et al.* 2010).

A simple measure of nucleotide divergence will therefore give an inadequate estimate of split time. In such cases, coalescent-based analyses such as MCMCcoal and IMa are useful to estimate split times and

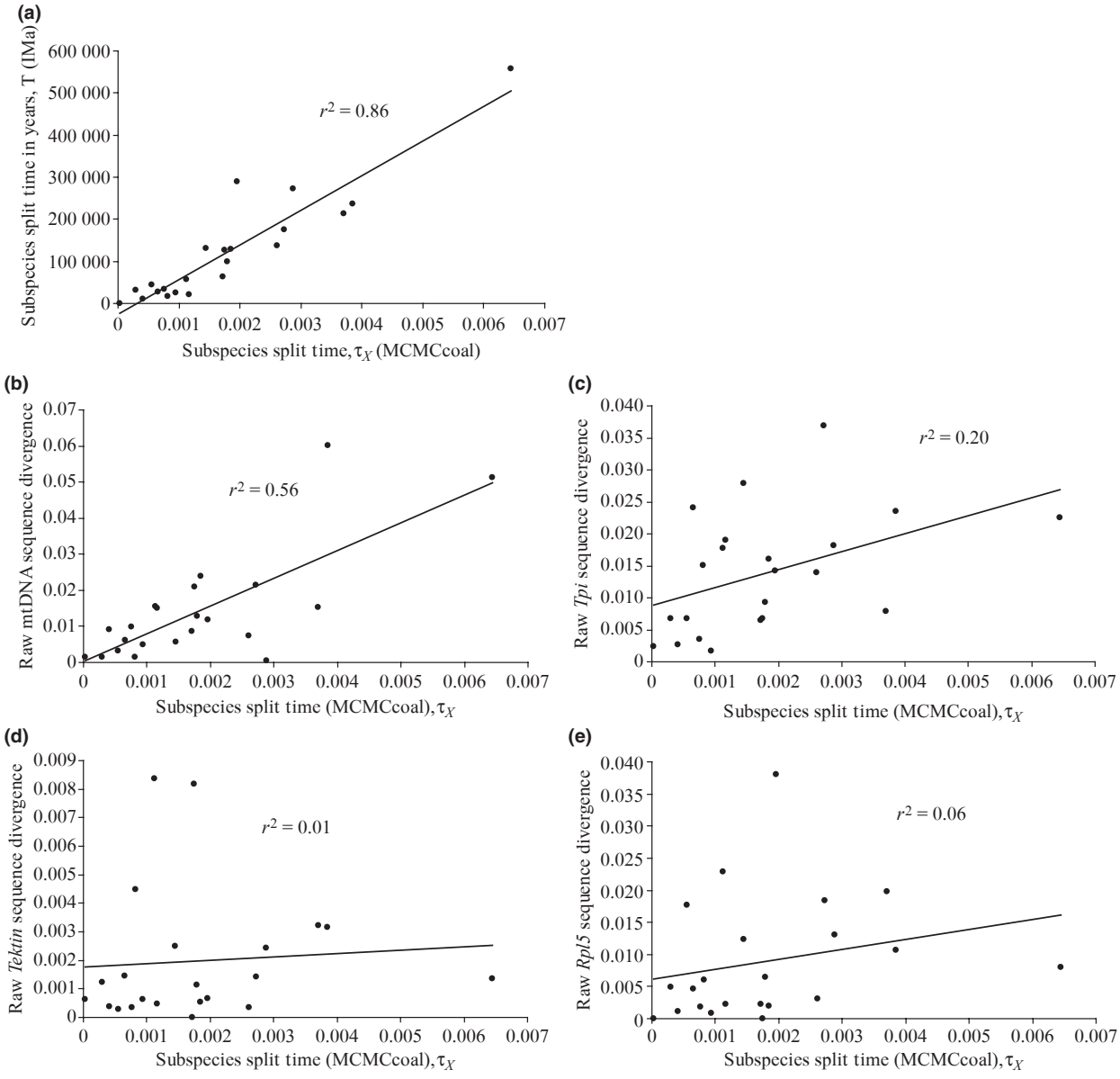

Fig. 4 Relationship between τ_X per nucleotide estimated using MCMCcoal and (a) split time in years estimated using IMa (assuming a substitution rate of 2.3% per million years at mtDNA), (b) raw percentage sequence divergence between subspecies at mtDNA, (c) *Tpi*, (d) *Tektin*, (e) *Rpl5*.

effective population sizes of ancestral and descendant populations using mutlilocus data sets, while accounting for coalescence in the ancestral population. We report a strong correlation ($r^2 = 0.56$) between τ_X and mitochondrial sequence divergence. The correlation of τ_X with sequence divergence at *Tpi* is weaker, and correlations were not significant with sequence divergence at *Rpl5* and *Tektin*. This decline in correlation may partly reflect the larger relative effective population sizes of the nuclear genes (n.b., mtDNA, *Tpi*, *Rpl5* and *Tektin* 'heredity multipliers' of population size are expected to be 0.25, 0.75, 1, 1, see Methods), although it is also likely to be due to smaller sequence read lengths at each nuclear gene and increased stochasticity owing to indel and noncoding DNA variation among lineages. Similar discrepancies between coalescent-based divergence estimates and sequence divergence were reported by Hurt *et al.* (2009) in their analysis of divergences between species pairs across the Isthmus of Panama.

While we report a strong correlation between τ_X and mitochondrial sequence divergence, there is substantial unexplained variance (Fig. 4b). For example, mtDNA divergence within *H. demeter* is ~6%, which would suggest a split time of 6–7 Myr from applying the

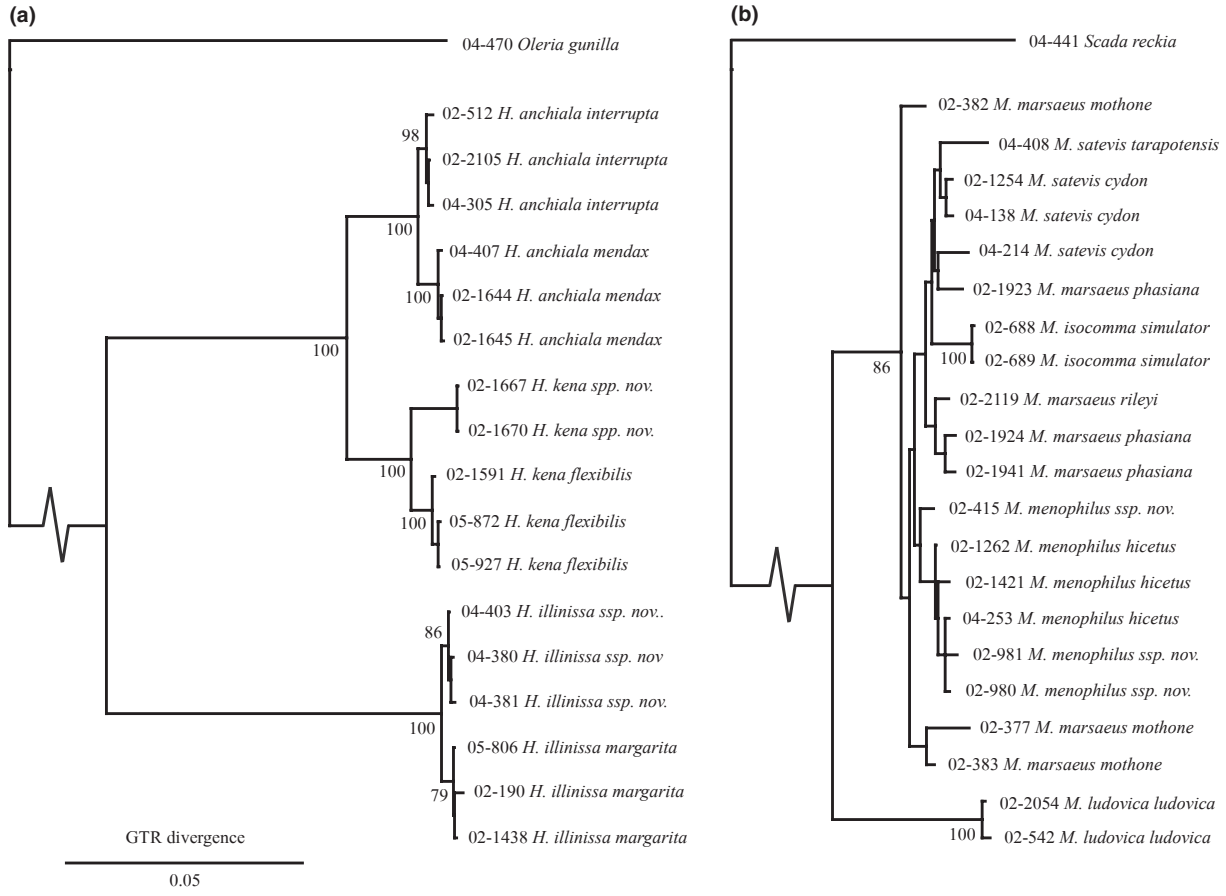

Fig. 5 Maximum likelihood trees based on concatenated mitochondrial and nuclear sequences of (a) *Hyposcada* and (b) *Melinaea* species found on either side of the suture zone. Nodes with 50% or greater bootstrap support are labelled.

traditionally accepted mtDNA substitution rate of 2.3% per branch per million years for arthropods (Brower 1994). Our analyses indicate that in this case, the split time is in fact less than half of this (Fig. 4b, Table 3). The opposite effect is seen in *H. erato*, where the net mtDNA sequence divergence between subspecies is only 0.05%, yet split times from both coalescent programs are considerably higher (Fig. 4B, Table 3), suggesting that the subspecies are not of as recent origin as that expected based from sequence divergence alone. Mean coalescence time in the ancestor is $2N_e$ generations, and these apparent inconsistencies are presumably owing to variation in ancestral and current population sizes as well as stochastic variation in coalescence time in ancestral populations: population size estimates, ancestral (θ_X) and particularly current population sizes (θ_{X_2} and θ_{X_1}), in *H. erato* are higher than in the other species (Table 3). Such discrepancies demonstrate the problem with simple application of sequence divergence from single loci for dating split times and the importance of approaches that simultaneously estimate split times as well as effective population sizes.

Although we correct for variation in overall substitution rate among lineages within both *Heliconius* and Ithomiini, we have assumed that the mean rate of substitution at mtDNA is the same in both groups to make comparisons between *Heliconius* and ithomiine subspecies pairs. There are no recent fossil calibrations for nymphalid butterflies (Wahlberg *et al.* 2009), and saturation of mtDNA changes among nymphalid groups makes it difficult to assess relative substitution rates for ithomiines and heliconiines. Even if average substitution rates are not equal in the two groups, it is unlikely that the conclusions reached here will be altered. In the Ithomiini, some taxa have high τ_X (*Oleria onega* and *Scada reckia*), some show intermediate values (*Pseudoscada florula* and *Oleria gunilla*), and others have virtually no divergence at all (*Brevioleria aelia* and *Melinaea menophilus*) (Fig. 3). Similarly, within *Heliconius*, estimates of τ_X vary from high (*H. demeter* and *H. pardalinus*) to near zero (*H. melpomene*). The same pattern of variable split times across the suture zone is found within each group.

There are many potential pitfalls with a coalescent-based approach. For example, all substitutions are

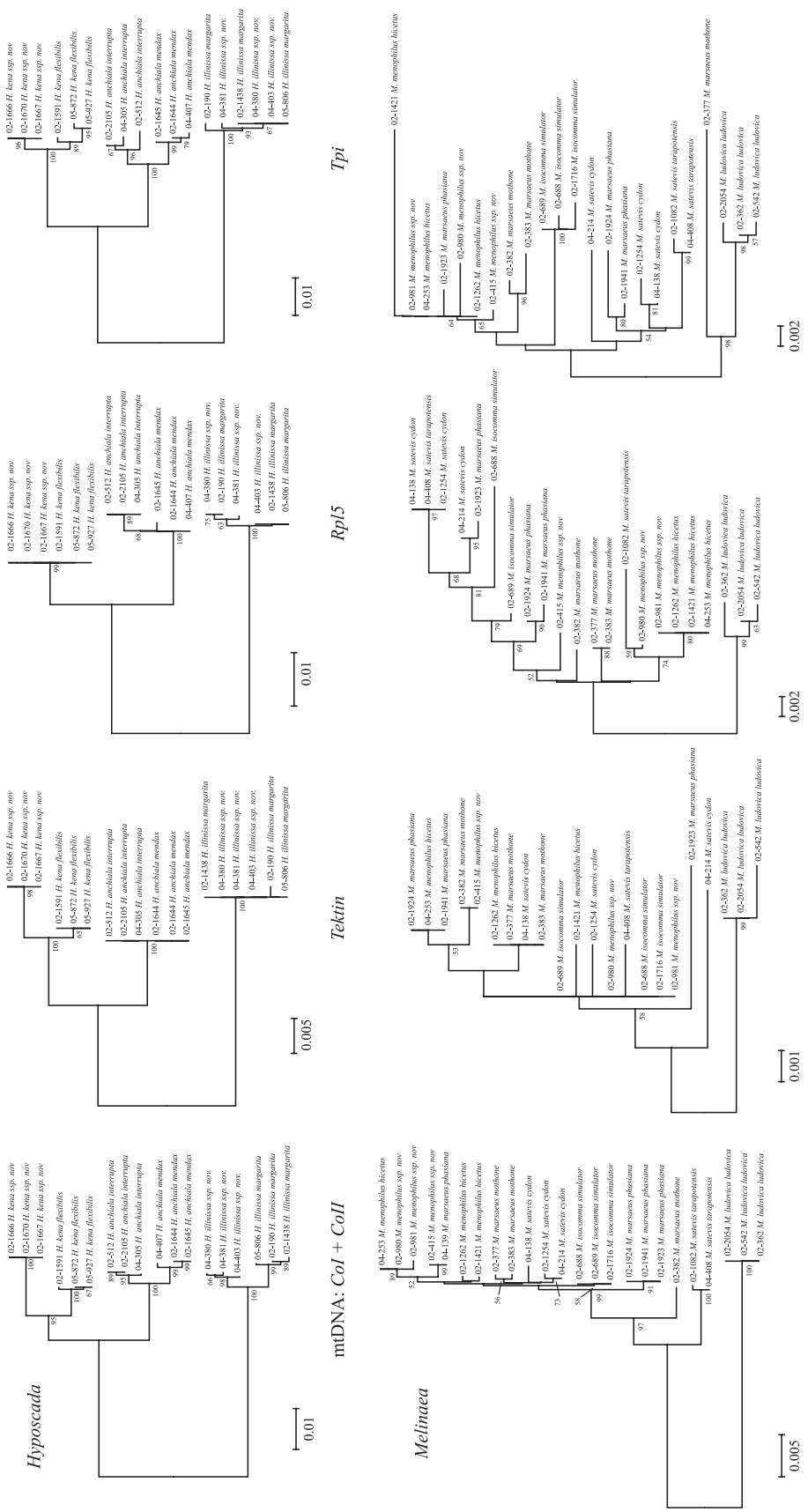

Fig. 6 Neighbour-joining trees of mitochondrial *Col* and *ColII*, *Tektin*, *Rpl5* and *Tpi* for the ithomiine genera *Melinaea* and *Hyposcada*. Scale bar indicates K2P distances. Nodes with 50% or greater bootstrap support are labelled.

assumed neutral. This will not be strictly true for coding regions such as those we have selected. However, given that the majority of divergence is synonymous and that we have corrected for relative substitution rates among loci (see Methods), we have minimized the problem. A further simplifying assumption is the absence of gene flow between subspecies since separation. Narrow hybrid zones between subspecies of many of our species are known to occur naturally in this region (Mallet 1993; Mallet & Lamas 2002) and have been well studied in *H. erato*, *H. melpomene* and *H. numata* (Mallet 1989; Mallet *et al.* 1990; Joron *et al.* 2001); in these zones, hybrid genotypes show unimodal distributions as expected when rates of hybridization in the zone are high (Jiggins & Mallet 2000). In contrast, contact zones between the highly divergent *Oleria onega* subspecies used in this study (de Silva *et al.* in prep.) and *Mechanitis mazaeus* (Dasmahapatra *et al.* 2010) show bimodal distributions of genotypes, implying low rates of hybridization and gene flow within the zone. The frequency of hybridization among sympatric taxa decreases with increasing genetic divergence and follows an approximately log-linear distribution in *Heliconius* (Mallet *et al.* 2007), and the parapatric taxa studied here are scattered across this spectrum. However, unlike sympatric species, many of the subspecies studied here are largely isolated by distance, with the possibility of limited gene exchange only within the narrow contact zone: it is likely that gene flow has little effect on divergence of each subspecies as a whole. Hybrid zones usually ensure a strong barrier to gene flow, as evidenced by strong frequency differences or fixed differences maintained across the zone, delaying neutral allele diffusion by many thousands of generations (Barton & Hewitt 1983, 1985). We have not taken the potential for long-range gene flow among subspecies into account, so that our estimates of split time could be somewhat underestimated. If compatibility does follow an exponential decline, it is possible that fractional underestimation would be greater for more recently diverged subspecies, and the appropriate correction could have the effect of reducing the variance in estimated split times across the suture zone. In general however, because split time is defined as the time since gene flow was completely abolished, there will inevitably be conflation of gene flow since split time and recency of actual split time.

We have successfully compared relative split times τ_X without a reliable estimate of absolute substitution rate, μ. An estimate of the absolute time of diversification would require calibration, for example using fossils, but these are not available for such recent events in butterflies. Average base divergence rate estimates for mitochondrial *CoI* and *CoII* vary from 2.3% per million years estimated in arthropods (Brower 1994) to 0.78–1.02% per million years in swallowtail butterflies (Zakharov *et al.* 2004). These rates suggest that divergence across the suture zone began 1500–200 000 or 4000–500 000 years ago, depending on which calibration is used. As substitution rates within *CoI* and *CoII* are known to vary widely in insects (Zakharov *et al.* 2004), these estimates are hardly reliable. Nonetheless, it is clear that most taxa across this suture zone are recent, probably diverging in the Pleistocene and Holocene.

The highly variable split times lead us to rule out a single vicariant event as the cause of the suture zone. Of course, our null hypothesis is simplistic, so several alternative hypotheses can be examined. If the Amazon basin went through glaciation-associated arid periods during the Pleistocene, climatic oscillations could have resulted in repeated forest contraction and expansion centred on the same forest refuges. Variable split times clearly do not rule out the possibility that different pairs of subspecies were formed during each forest contraction/expansion episode. However, by rejecting a simple 'climatic forcing' argument, we are left with a model of 'climatic assistance' that generates nonreplicable patterns across multiple species under identical environmental cycles. A simple, parsimonious null model becomes hard to reject: that the pattern of divergences we observe is similar to that resulting from a simple stochastic model of speciation, independent of climatic oscillations, such as a constant low probability of species/subspecies divergence and extinction per unit time, perhaps approximately Poisson-distributed (cf. Fig. 3). A 'climatic assistance' explanation involving multiple split times but geographically coincident vicariant events would be expected to affect all lineages similarly: variable split times across the suture zone should be found within all lineages. The same would be true for simple stochastic null models of divergence. However, in the butterfly data presented here, there appear to be strong lineage effects that are unexpected under either hypothesis.

Within the genus *Melinaea*, all the taxa studied are of very recent divergence (Fig. 3). Furthermore, four fully sympatric species (*M. menophilus*, *M. marsaeus*, *M. satevis* and *M. isocomma*), three of which show subspecies differentiation across the zone, are themselves virtually indistinguishable based on the loci sequenced here: they have polyphyletic genealogies at each of these loci (Figs 5 and 6). The species themselves are extremely young (perhaps <100 000 years old), so that their subspecies must be even younger (perhaps ~10 000 years or less). These results are in stark contrast to those from subtribe Oleriina (*Hyposcada* and *Oleria*) which show considerably greater genetic divergence (Figs 5 and 6) and longer split times among pairs of

subspecies (Fig. 3), greater even than *Melinaea* show among species (Figs 5 and 6). Thus, both the random speciation model and the climatic assistance model seem inadequate. Even if climatic assistance 'explains' divergence of subspecies and species, lineage-specific biological factors seem to have greater importance.

While ecological factors such as larval host plant adaptation (Willmott & Freitas 2006) and changes in pheromone signalling (Schulz *et al.* 2004; Estrada *et al.* 2010) may affect diversification in the Ithomiini and Heliconiina, there is increasing evidence that changes in wing colour patterns are important. Mimetic shifts coupled with wing colour–mediated mate preference are known to be important reproductive isolation factors between several recently diverged *Heliconius* sister-species as well as colour pattern races within the same species (Jiggins *et al.* 2001; Chamberlain *et al.* 2009). In the Ithomini, the diversity of colour patterns found varies from genera to genera. Within *Ithomia*, wing pattern changes are associated with diversification above the species level (Jiggins *et al.* 2006). By contrast, in the mimetically more homogenous Oleriina, local geographic isolation is thought to be more important in diversification (de-Silva *et al.* 2010), demonstrating how biological factors vary among lineages.

We should emphasize that we here use pairs of morphologically distinct taxa only to test the null hypothesis of simultaneous divergence. More than 40 other species in these same butterfly groups, which are often as abundant and widespread as the divergent taxa, are also distributed across the suture zone and yet show no obvious morphological or subspecies-level divergence. Under the vicariant model, these nondivergent taxa would have had to have gone through the single or multiple range splits without diverging at all.

Generalizing to other taxa, we might expect that forest-inhabiting birds would also diversify in rainforest refuges with the butterflies, given a perhaps naïve application of the Haffer and Brown models (correlations of centres of endemism among birds, butterflies and other taxa have often been claimed). Yet although there are endemic bird taxa in the Eastern Andes, there are virtually no 'Huallaga' vs. 'Ucayali' geminate subspecies or species pairs similar to those in the butterflies studied here (Haffer 1987): most lowland birds in this region have gone through the hypothesized cycles of geographic isolation without diverging at all. In contrast, poison frogs of the genus *Dendrobates* and *Epipedobates*, which themselves form mimicry rings (Symula *et al.* 2001), show very narrow patterns of endemism and genetic differentiation in the same region of Peru, with multiple different forms on opposite banks of the Río Mayo, as well as up and down the Mayo valley (Symula *et al.* 2001; Roberts *et al.* 2006, 2007). These patterns of endemism are much narrower than those shown by ithomiine and heliconiine butterflies. Similar, highly local patterns of differentiation in *Epipedobates* are seen in Ecuador (Graham *et al.* 2004).

If the suture zone near Tarapoto, Peru was not produced as a result of simple contact between two recently re-expanded forest refuge biotas, what caused the well-defined pattern of biotic contacts (Fig. 1)? An alternative explanation for the existence of this suture zone is required. Hybrid zones and clines are potentially mobile, and their position is maintained by a balance between selection and dispersal (Barton & Hewitt 1985; Barton & Hewitt 1989). In neotropical butterflies, there is empirical evidence that hybrid and contact zones can move rapidly (Blum 2002; Dasmahapatra *et al.* 2002). Differential selection for or against particular genotypes across an ecotone will cause a moving cline to settle on boundaries between environments, or in regions of low population density which act as sinks for migration (Barton 1979; Barton *et al.* 1985; Mallet 2010, 1993). Therefore, the suture zone could be the result of multiple contact and hybrid zones moving independently of one another and becoming trapped on a common ecological hiatus. In some cases, the transitions are between geminate mid-elevation-adapted and lowland forms (e.g. *Hyposcada kena*, *Napeogenes pharo* and *Ithomia salapia*); in others, the transitions are among geminate lowland rainforest taxa (e.g. *Oleria onega*, *Heliconius erato*, *H. melpomene* and *H. pardalinus*). The Tarapoto suture zone corresponds approximately with the eastern base (rather than the ridge top) of the Cordillera Escalera range (Fig. 1) east of which lies the virtually unbroken Amazon basin. The eastern slopes of this ridge attract high levels of orogenic rainfall (Pongo del Cainarachi, site 8 on Fig. 1, has 3637 mm rain per year) compared with further out into the Amazon basin (e.g. Yurimaguas, 2279 mm), and with the rain shadow Southwest of the Cordillera Escalera [e.g. Tarapoto, 1004 mm, see Fig. 1 for locations (Mallet 1993)]. Given that butterflies fly little during wet weather, reducing their potential rates of increase, the wetter eastern slopes could therefore act as a population sink which traps clines and hybrid zones.

Thus, we see a range of levels of local geographic differentiation and diversification for forest taxa, from weak (birds, some butterflies), to strong (the geminate taxa of butterflies studied here), to very strong (dendrobatid frogs). There are clear lineage-specific effects on this broader phylogenetic scale, just as found on a finer scale for genetic divergence and split time among the geminate subspecies of butterflies. The lineage-specific rates of divergence argue that climate forcing or climatic assistance vicariance models provide a very incomplete explanation for diversification. Climate-

related, geographical isolation may be one of the drivers of evolution in the Amazon rainforest but fails as a general explanation. If they existed, past refuges apparently affected each group of taxa in different ways. Alternatively, the divergence could have taken place in parapatry and was more constrained by ecological factors than geographic isolation (Benson 1982; Endler 1982b, Moritz *et al.* 2000; Mallet 1993, 2010). Lineage-specific traits of the organisms themselves, such as dispersal distance and vagility, or the tendency to local ecological adaptation and adaptive radiation appears to have greater importance than vicariance events in the diversification of the Amazon fauna.

Acknowledgements

We thank INRENA providing collection permits (096-2004-INRENA-IFFS-DCB, 021C/C-2005-INRENA-IANP) and CIMA for logistic assistance in the field. In particular, we thank Lily Rodríguez, Bertha Alvarado, Luis Benites, Wacho Aguirre, Tatiana Pequeño, Alvaro del Campo, Raúl Enrique Torrico Zavaleta, Jabin Grández Shapiama, Agústin Macedon Flores and the other park guards of Parque Nacional Cordillera Azul. We also thank NERC for funding this research and Ziheng Yang for guidance with implementing analyses in MCMCcoal and PAML and for suggesting the *ad hoc* likelihood ratio tests. Assistance in the field was also provided by Jorge Carillo, Lisa de Silva, Moisés Abanto, Douglas Cotrina, Laura Ferguson and Armando Silva-Vásquez.

References

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman and Hall, London.

Avise JC (2000) *Phylogeography*. Harvard University Press, Boston, Massachusetts.

Barton NH (1979) The dynamics of hybrid zones. *Heredity*, **43**, 341–359.

Barton NH, Hewitt GM (1983) Hybrid zones as barriers to gene flow. In: *Protein Polymorphism: Adaptive and Taxonomic Significance* (ed. Oxford GS), pp. 341–359. Academic Press, London & New York.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

Beaumont MA, Nielsen R, Robert C *et al.* (2010) In defence of model-based inference in phylogeography. *Molecular Ecology*, **19**, 436–446.

Beltrán M, Jiggins CD, Brower AVZ, Bermingham E, Mallet J (2007) Do pollen feeding and pupal-mating have a single origin in *Heliconius*? Inferences from multilocus sequence data *Biological Journal of the Linnean Society*, **92**, 221–239.

Benson WW (1982) Alternative models for infrageneric diversification in the humid tropics: tests with passion vine butterflies. In: *Biological Diversification in the Tropics* (ed. Prance GT), pp. 608–640. Columbia Univ. Press, New York, NY.

Bermingham E, Lessios HA (1993) Rate variation of protein and mtDNA evolution as revealed by sea urchins separated by the Isthmus of Panama. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 2434–2738.

Beven S, Connor EF, Beven K (1984) Avian biogeography in the Amazon basin and the biological model of diversification. *Journal of Biogeography*, **11**, 383–399.

Blum MJ (2002) Rapid movement of a *Heliconius* hybrid zone: evidence for phase III of Wright's shifting balance theory? *Evolution*, **56**, 1992–1998.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.

Brower AVZ, Freitas AVL, Lee MM *et al.* (2006) Phylogenetic relationships among the Ithomiini (Lepidoptera: Nymphalidae) inferred from one mitochondrial and two nuclear gene regions. *Systematic Entomology*, **31**, 288–301.

Brown WL (1957) Centrifugal speciation. *Quarterly Review of Biology*, **32**, 247–277.

Brown KS (1977) Geographical patterns of evolution in Neotropical forest Lepidoptera (Nymphalidae: Ithomiinae and Nymphalidae-Heliconiini). In: *Biogeographie et Evolution en Amérique Tropicale* (ed. Descimon H), pp. 118–160. Publications du Laboratoire de Zoologie de l'Ecole Normale Superieure No. 9, Paris.

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

Brown KS (1987a) Areas where humid tropical forest probably persisted. In:*Biogeography and Quaternary History in Tropical America* (ed. Whitmore TC), p. 45. Oxford University Press, Oxford, U.K.

Brown KS (1987b) Biogeography and evolution of neotropical butterflies. In: *Biogeography and Quaternary History in Tropical America* (ed. Whitmore TC), pp. 66–104. Oxford University Press, Oxford, U.K.

Brown KS, Sheppard PM, Turner JRG (1974) Quaternary refugia in tropical America: evidence from race formation in *Heliconius* butterflies. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **187**, 369–378.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic Forest Hotspot. *Science*, **323**, 785–789.

Chamberlain NL, Hill RI, Kapan DD, Gilbert LE, Kronforst MR (2009) Polymorphic butterfly reveals the missing link in ecological speciation. *Science*, **326**, 847–850.

Colinvaux PA, De Oliviera PE, Bush MB (2000) Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. *Quaternary Science Reviews*, **19**, 141–169.

Colinvaux PA, Irion G, Räsänen ME, Bush MB, Nunes de Mello JAS (2001) A paradigm to be discarded: geological and paleoecological data falsify the HAFFER & PRANCE refuge hypothesis of Amazonian speciation. *Amazoniana*, **16**, 609–646.

Coyne JA, Orr HA (2004) *Speciation*, Sinauer Associates, Sunderland, Massachusetts.

Cracraft J, Prum RO (1988) Patterns and processes of diversification: speciation and historical congruence in some neotropical birds. *Evolution*, **42**, 603–620.

Dasmahapatra KK, Blum MJ, Aiello A *et al.* (2002) Inferences from a rapidly moving hybrid zone. *Evolution*, **56**, 741–753.

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, Elias M, Hill RI, Hoffman JI, Mallet J (2010) Mitochondrial DNA barcoding detects some species that are real, and some that are not. *Molecular Ecology Resources*, **10**, 264–273.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Elias M, Joron M, Willmott K *et al.* (2009) Out of the Andes: patterns of diversification in clearwing butterflies. *Molecular Ecology*, **18**, 1716–1729.

Endler JA (1982a) Comment on John Turner's paper. In: *Biological Diversification in the Tropics* (ed. Prance GT). pp. 331–332. Columbia University Press, New York.

Endler JA (1982b) Problems in distinguishing historical from ecological factors in biogeography. *American Zoologist*, **22**, 441–452.

Estrada C, Yildizhan S, Schulz S, Gilbert LE (2010) Sex-specific chemical cues from immatures facilitate the evolution of mate guarding in *Heliconius* butterflies. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **277**, 407–413.

Fjeldså J (1994) Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity and Conservation*, **3**, 207–226.

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.

Gascon C, Malcolm JR, Patton JL *et al.* (2000) Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences, USA*, **97**, 13672–13677.

Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781–1793.

Haffer J (1969) Speciation in amazonian forest birds. *Science*, **165**, 131–137.

Haffer J (1987) Biogeography of Neotropical birds. In: *Biogeography and Quaternary History in Tropical America* (ed. Whitmore TC), pp. 105–150. Oxford University Press, Oxford, U.K.

Haffer J (1997) Alternative models of vertebrate speciation in Amazonia: a review. *Biodiversity and Conservation*, **6**, 451–476.

Haffer J (2008) Hypotheses to explain the origins of species in Amazonia. *Brazilian Journal of Biology*, **68**, 917–947.

Haffer J, Prance GT (2001) Climatic forcing of evolution in Amazonia during the Cenozoic: on the refuge theory of biotic differentiation. *Amazoniana*, **16**, 579–607.

Hall JPW, Harvey DJ (2002) The phylogeography of Amazonia revisited: new evidence from riodinid butterflies. *Evolution*, **56**, 1489–1497.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hey J (2009) Using the IMa program. User Documentation

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences USA*, **104**, 2785–2790.

Hickerson MJ, Stahl E, Lessios HA (2006) Test for simultaneous divergence using approximate Bayesian computation. *Evolution*, **60**, 2435–2453.

Hickerson MJ, Stahl E, Takebayashi N (2007) msBayes: pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics*, **8**, 268.

Hickerson MJ, Carstens BC, Cavender-Bares J *et al.* (2010) Phylogeography's past, present, and future: 10 years after Avise 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.

Hudson RR, Turelli M (2003) Stochasticity overrules the ''three-times rule'': genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**, 182–190.

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.

Hurt C, Anker A, Knowlton N (2009) A multilocus test of simultaneous divergence across the Isthmus of Panama using snapping shrimps in the genus *Alpheus*. *Evolution*, **63**, 514–530.

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution*, **15**, 250–255.

Jiggins CD, Naisbit RE, Coe RL, Mallet J (2001) Reproductive isolation caused by colour pattern mimicry. *Nature*, **411**, 302–305.

Jiggins CD, Mallarino R, Willmott KR, Bermingham E (2006) The phylogenetic pattern of speciation and wing pattern change in Neotropical *Ithomia* butterflies (Lepidoptera: Nymphalidae). *Evolution*, **60**, 1454–1466.

Joron M, Wynne IR, Lamas G, Mallet J (2001) Variable selection and the coexistence of multiple mimetic forms of the butterfly *Heliconius numata*. *Evolutionary Ecology*, **13**, 721–754.

Klicka J, Zink RM (1997) The importance of recent ice ages in speciation: a failed paradigm. *Science*, **277**, 1666–1669.

Knowlton N, Weigt L (1998) New dates and new rates for the divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **265**, 2257–2265.

Knowlton N, Weigt LA, Solórzano LA, Mills DK, Bermingham E (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science*, **260**, 1629–1632.

Lamas G (1982) A preliminary zoogeographical division of Peru based on butterfly distributions (Lepidoptera, Papilionoidea).

In: *Biological Diversification in the Tropics* (ed. Prance GT), pp. 336–357. Columbia University Press, New York.

Lamas G, Callaghan C, Casagrande MM *et al.* (2004) *Atlas of Neotropical Lepidoptera. Checklist: Part 4A. Hesperioidea—Papilionoidea*. Association for Tropical Lepidoptera/Scientific Publishers, Gainesville, Florida.

Lessios HA (1979) Use of Panamanian sea urchins to test the molecular clock. *Nature*, **280**, 599–601.

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.

Mallet J (1989) The genetics of warning colour in Peruvian hybrid zones of *Heliconius erato* and *H. melpomene*. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **236**, 163–185.

Mallet J (1993) Speciation, raciation, and color pattern evolution in *Heliconius* butterflies: evidence from hybrid zones. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 226–260. Oxford University Press, New York.

Mallet J (2010) Shift happens! Shifting balance and the evolution of diversity in warning colour and mimicry. *Ecological Entomology*, **35**, 90–104.

Mallet J, Lamas G (2002) Heliconiina and Ithomiinae of Río Mayo and lower Río Huallaga, Peru. http://www.ucl.ac.uk/taxome/ith/tarapoto/sutzon.html

Mallet J, Barton N, Lamas G *et al.* (1990) Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics*, **124**, 921–936.

Mallet J, Beltrán M, Neukirchen W, Linares M (2007) Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology*, **7**, 28.

Mayr E (1963) *Animal Species and Evolution*, Harvard University Press, Cambridge, Massachusetts.

Moore WS, Price JT (1993) Nature of selection in the northern flicker hybrid zone and its implications for speciation theory. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 196–225. Oxford University Press, New York.

Moritz C, Patton JL, Schneider CJ, Smith TB (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533–563.

Moritz C, Hoskin CJ, Mackenzie JB *et al.* (2009) Identification and dynamics of a cryptic suture zone in tropical rainforest. *Proceedings of the Royal Society of London Series B—Biological Sciences*, **276**, 1235–1244.

Nelson BW, Ferreira CAC, da Silva MF, Kawasaki ML (1990) Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature*, **345**, 714–716.

Nichols R (2001) Gene trees and species trees are not the same. *Trends in Ecology and Evolution*, **16**, 358–364.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation. A Markov-chain Monte-Carlo approach. *Genetics*, **158**, 885–896.

Patton JL, Da Silva MNF (1998) Rivers, refuges and ridges. The geography of speciation of Amazonian mammals. In: *Endless Forms. Species and Speciation* (ed. Howard DJ), pp. 202–216. Oxford University Press, New York.

Posada NM, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Prance GT (1973) Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon basin, based on evidence from distribution patterns in Caryocaraceae, Chrysobalanaceae, Dichapetalaceae and Lecythidaceae. *Acta Amazonica*, **3**, 5–28.

Prance GT (1982) Forest refuges: evidence from woody angiosperms. In: *Biological Diversification in the Tropics* (ed. Prance GT), pp. 137–156. Columbia University Press, New York.

Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, **164**, 1645–1656.

Remington CL (1968) Suture-zones of hybrid interaction between recently joined biotas. *Evolutionary Biology*, **1**, 321–428.

Roberts JL, Brown JL, von May R *et al.* (2006) Genetic divergence and speciation in lowland and montane Peruvian poison frogs. *Molecular Phylogenetics and Evolution*, **41**, 149–164.

Roberts JL, Brown JL, Schulte R, Arizabal W, Summers K (2007) Rapid diversification of colouration among populations of a poison frog isolated on sky peninsulas in the central cordilleras of Peru. *Journal of Biogeography*, **34**, 417–426.

Schulz S, Beccaloni G, Brown KS *et al.* (2004) Semiochemicals derived from pyrrolizidine alkaloids in male ithomiine butterflies (Lepidoptera : Nymphalidae : Ithomiinae). *Biochemical Systematics and Ecology*, **32**, 699–713.

de-Silva DL, Day JJ, Elias M *et al.* (2010) Molecular phylogenetics of the neotropical butterfly subtribe Oleriina (Nymphalidae: Danainae: Ithomiini). *Molecular Phylogenetics and Evolution*, **55**, 1032–1041.

Simpson BB, Haffer J (1978) Speciation patterns in the Amazonian forest biota. *Annual Review of Ecology and Systematics*, **9**, 497–518.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **277**, 661–671.

Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4*, Sinauer Associates, Sunderland, Massachusetts.

Symula R, Schulte R, Summers K (2001) Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **268**, 2415–2421.

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.

Vawter AT, Rosenblatt R, Gorman GC (1980) Genetic divergence among fishes of the Eastern Pacific on the Caribbean: support for the molecular clock. *Evolution*, **34**, 705–711.

Wahlberg N, Leneveu U, Kodandaramaiah U *et al.* (2009) Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. *Proceedings of the Royal Society B-Biological Sciences*, **276**, 4295–4302.

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 B*, **272**, 2525–2533.

Willmott KR, Freitas AVL (2006) Higher-level phylogeny of the Ithomiinae (Lepidoptera: Nymphalidae), with a review of classification, patterns of larval hostplant colonisation and diversification. *Cladistics*, **22**, 297–368.

Yang Z (2002) Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics*, **162**, 1811–1823.

Yang Z (2007a) MCMCcoal Markov Chain Monte Carlo Coalescent Program Version 1.2. User Documentation.

Yang Z (2007b) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.

Yang Z, Yoder AD (2003) Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. *Systematic Biology*, **52**, 705–716.

Yoder AD, Yang Z (2000) Estimation of primate speciation dates using local molecular clocks. *Molecular Biology and Evolution*, **17**, 1081–1090.

Zakharov EV, Caterino MS, Sperling FAH (2004) Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). *Systematic Biology*, **53**, 193–215.

Kanchon Dasmahapatra's main research interests lie in using molecular markers to investigate species biology and evolutionary processes, in particular species boundaries in neotropical butterflies. Gerardo Lamas works on the evolutionary systematics and biogeography of butterflies. Fraser Simpson has a broad interest in natural history, and particularly in ornithology. James Mallet is interested in ecological adaptation, speciation and the nature of species, with a special emphasis on neotropical butterflies.

Supporting information

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

Fig. S1 A) Ithomiine and B) *Heliconius* tree topologies used to estimate relative substitution rates of the four loci (mtDNA, *Tpi*, *Tektin*, *Rpl5*) by comparing maximum likelihood tree lengths for each locus. The trees were also used to test for variation in evolutionary rate among lineages and to estimate the relative substitution rates in each of the labelled clades. Trees based on phylogenies described in Brower et al. (2006) and Beltrán et al. (2007).

Table S1 Details of samples used in this study together with Genbank accession numbers of sequences used

Table S2 PCR conditions for the amplicons used in this study

Table S3 Additional primers required for some specimens in a few genera

Table S4 Summary of likelihood ratio tests used in the two *ad hoc* tests comparing MCMCcoal results with unconstrained and constrained split times (τ_X)

Please note: Wiley-Blackwell are 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.