REPEATED GENETIC PATTERNS ACROSS A FAUNAL "SUTURE ZONE"

Introduction: underlying rationale

Since the earliest times, biogeographic patterns across the planet have inspired theories to explain the origins of biodiversity. However, because of the complexity of the patterns, even recent biogeographers have tended to use distribution patterns to confirm preferred theories, rather than making genuine attempts to test between hypotheses. Recently, the ease with which DNA sequences can be read has given hope for more objective tests, an approach generally referred to as "comparative phylogeography". Here, I advocate the use of a new and original replicated approach to distinguish between sets of simple alternative hypotheses as a route towards testing the major theories.

There are a large number of theories for diversification in the world's most diverse biome, the Amazon rainforest. Despite a need to understand diversity processes in order to conserve Amazonian biota, a general theory has yet to emerge. Below, I briefly summarise conflicting biogeographic hypotheses for the rainforests of the Amazon basin. I then outline a molecular genetic approach to search across multiple contact zones of heliconiine and ithomiine butterflies within a small region of Peru at the edge of the Amazon basin, and show how the approach can exploit the high species richness to ensure replicated sampling of general biogeographic patterns. Finally, I discuss potential significance of such work in terms of NERC's mission statement.

Theories of diversification in the Neotropical rainforest. The Neotropical rainforest, and the Amazon basin in particular, is the most diverse terrestrial biome on the planet. Many biogeographic theories have been put forward to explain its diversification (Haffer 1997), but it is safe to say that opinions still vary widely. Among the major theories are: (1) *riverine barrier hypotheses*, whereby major rivers act as dispersal barriers, today largely discounted (Capparella 1988, Patton et al. 1998, Gascon et al. 2000); (2) *refugium hypotheses*, in which drier conditions during the Pleistocene led to allopatric speciation in forest refuges (Simpson & Haffer 1978, Whitmore & Prance 1987, Haffer 1997). These hypotheses are particularly relevant to the current study, because the Heliconiina and Ithomiinae are the two best-studied groups among those that led to the development of refugium theory (Turner 1971, Brown 1987, Turner & Mallet 1996). Today, this theory is under strong attack (Endler 1982, Beven et al. 1984, Mallet 1993, Mallet & Turner 1998, Willis & Whittaker 2000, Bush 2003), but modified versions termed (3) *vicariance hypotheses* are gaining support (Cracraft & Prum 1988, Brumfield & Capparella 1996, Hall & Harvey 2002). It is now clear from molecular work that the originally envisaged time-scale for diversification in Pleistocene refuges of tens of thousands of years is almost certainly incorrect. Instead, molecular divergence suggests speciation and divergence took hundreds of thousands to millions of years, often occurring long before the Pleistocene, at least in the temperate zone where most molecular studies have been performed (Zink & Slowinski 1995, Zink 1996, Avise et al. 1998, Hewitt 2000). Palynologists have also found little evidence for Pleistocene restriction of rainforest in the Amazon ("all claims for arid land processes shown to be in error" – Colinvaux et al. 2001; see also Bush 2003), although others disagree (Haffer 1997). (4) *Ecological hypotheses* argue that speciation took place via local adaptation to environmental conditions that probably pertained, and may still pertain, even in the absence of geographic isolation (Endler 1982). (5) *Peripheral speciation/Amazonian "museum" hypotheses* propose that small local areas at the edge of the neotropical forest produced evolutionary novelties and new species, which then variably spread in towards the centre of the Amazon basin (Mayr 1954, Carson & Templeton 1984, Fjeldså 1994, Fjeldså & Rahbek 1998). In Fjeldså's (1994) analysis, DNA-DNA hybridization data from Sibley & Ahlquist (1990) were used to show that Andean species were on average younger than Amazonian species; the Amazon was seen as a "museum" of older diversity. (6) In *centrifugal divergence hypotheses* (W.L. Brown 1957), the centre of the rainforest range, such as the Amazon basin, produced novelties which then variably spread out, leaving peripheries of the range with relictual taxa, for instance in the Andes (Mallet 1993, Mallet & Turner 1998) perhaps as a result of rapid movement of hybrid zones (Parsons et al. 1993, Dasmahapatra et al. 2002, Blum 2002). W.L. Brown's (1957) hypothesis was originally devised as a testable alternative to Mayr's (1954) peripheral founder model for the drongos of New Guinea and other examples. Many unexplained "leap-frog" biogeographic patterns in Andean and other extra-Amazonian taxa are very likely due to secondary disjunction after centrifugal spread (Remsen

1984, Mallet & Turner 1998). Also, Fjeldså's (1994) results (above) could be explained if the taxa he regards as "young species" in the Andes are actually "old subspecies", and if Amazonian species consist of even younger subspecies not so prone to being split by taxonomists into separate species. The Andes would then become a repository of relictual, older forms of varying ages (older subspecies, as well as older species; Fjeldså also found ancient, endemic species in the Andes whose distributions correlated with his "young species"). These patterns would be explained if diversification and speciation successively spread centrifugally after originating in Amazonia until they lapped up against the Andes, leaving remnants of earlier spread clinging to the periphery of their former range; the precise opposite of Fjeldså's own "museum hypothesis" of Amazonian biodiversity.

Synthesis. Many or all of the conflicting theories could apply to particular groups (Haffer 1997), but this result would be an unsatisfactory result because it lacks generality. (Multiple modes of diversification are likely, of course, but clear biogeographic patterns such as "suture zones" – see below – argue for general causes). Even if there are multiple causes of diversification, it would be helpful to identify the features that predispose particular taxa to particular modes of speciation. In practice, even the most basic tenets of these theories, such as the existence of correlated centres of endemism required by refugium and vicariance theories, are poorly demonstrated (Beven et al. 1984, Nelson et al. 1990, Mallet 1993). In part this is because of the great complexity of analysing spatial data. Another reason is that, though temperate glacial refuges are well studied (Hewitt 2000), genetic data from the vast species storehouse of the tropics are still scarce (but see Brumfield and Capparella 1996, Fjeldså 1994, Gascon et al. 2000).

Mechanisms of diversification. Major advances in understanding speciation have taken place in the last 20 years (e.g. Coyne & Orr 1997; Turelli et al. 2001). Studies of hybrid zones have contributed strongly to knowledge of the genetics of speciation (e.g. Barton & Hewitt 1989; Jiggins & Mallet 2000). Recently, the ecology of speciation has been explored as more studies have been carried out in the field (see examples in Howard & Berlocher 1998), with Darwin's finches (Grant 1986), sticklebacks (Schluter 2000) and *Rhagoletis* fruit flies (Feder et al. 1998) providing some of the best understood cases. In my own group, we have been investigating speciation in *Heliconius* in this vein (e.g. McMillan et al. 1997, Mallet et al. 1998, Jiggins et al. 2001a,b, Naisbit et al. 2001, 2002, 2003, Beltrán et al. 2002, Bull 2003).

Understanding individual speciation events, while satisfying in detail, gives little hint of general patterns. Biogeography can extend and generalise these recent advances in genetic and ecological knowledge of speciation by analysing multiple taxa simultaneously. Phylogeography (Zink & Slowinski 1995, Zink 1996, Hewitt 1996, Avise et al. 1998) and inferences about geographic modes of speciation depending on range geography (Barraclough & Vogler 2000) are good examples of a trend towards a more comparative approach. This proposal builds on my laboratory's expertise with speciation in Lepidoptera (particularly *Heliconius*) by using a deliberately simplified but highly replicated form of molecular "comparative phylogeography" (Zink 1996) to test hypotheses in a small region of the neotropical rainforest.

Rationale for genetic tests. The complexity of Amazon biogeography and high species diversity in the region makes it hard to know where to start. Nearly 400 species of Heliconiina and Ithomiinae exist in neotropical rainforests, each with up to 30 morphologically defined subspecies, separated by hybrid or contact zones. To obtain sequences from all of these would require a massive program grant lasting at least 5 years (sequencing technology is becoming very rapid, but the fieldwork and collecting alone will slow the process). Instead, this proposal builds toward the collaborative goal of a complete "DNA taxonomy" for Neotropical Heliconiina + Ithomiinae by focussing on a pair of adjacent, well-defined zones of endemism, implicated in the Pleistocene refugium theory as the Huallaga and Ucayali refuges (K.S. Brown, e.g. 1987), and containing among the highest species richness in the Amazon.

Specific objectives: hypotheses to be tested

Some important corollaries of alternative theories of Amazonian diversification are highly amenable to testing, even though overall ideas may be hard to put into a testable framework. While temperate zone genetic studies are well advanced, work has hardly begun in the tropics. Simple and robust predictions such as those made here should also be more easily testable in the tropics because of the extremely high

species diversity, allowing greater replication. Testing these hypotheses in the tropics also has great relevance to global conservation, where development is a pressing threat to the world's largest concentration of biodiversity. Among possible testable predictions are:

1) *Clustered time course of differentiation.* Under vicariance/refuge explanations, we in effect claim that relatively few biogeographic events produced divergence in many taxa simultaneously. The seemingly unlikely alternative is that a number of separate events have combined at different times and places to produce the spatially correlated patterns of differentiation we observe today. A clustered time-course for divergence would support vicariance, while a scattered time course would suggest multiple causes of diversification. In the temperate zone, it is now clear that sister species pairs show an extremely scattered time course of speciation throughout the Pliocene and Pleistocene, and a simple hypothesis of simultaneous vicariance for all taxa is thereby disproved. Allopatric divergence may have been involved, but if so, it was piecemeal and irregular, rather than clustered in time for large fractions of the biota (Avise et al. 1998, Hewitt 2000). To date, most evidence has come from the temperate zone, and hardly any data on the time course of geographic diversification is available from the tropics. The fossil record for recent taxa is virtually non-existent in the Amazon basin, so we must rely even more strongly on molecular divergence data as a surrogate for time than in the temperate zone.

2) *Tropical species younger than temperate species*. The relatively ancient divergence of many temperate species (>2My) may well be due to high rates of extinction during Pleistocene climatic vicissitudes (Hewitt 2000, Avise et al. 1998). Studies across mainly higher latitude hybrid zones showed on average 20% of allozyme loci with fixed differences, again suggesting ancient divergence even between geographic races of the same species (Barton & Hewitt 1983). However, races of the tropical *Heliconius erato* and *H. melpomene*, which form part of this proposal, show little or no molecular differentiation (Barton & Hewitt 1983, Mallet et al. 1998; Brower 1994,1996, Bull 2003, Flanagan et al. 2003). If speciation in temperate and tropical zones was affected in a similar way by climate change throughout the Pliocene and Pleistocene, sister species/races in tropical and temperate zones should be of roughly the same age. If on the other hand, speciation rates were roughly equivalent globally, but tropical species diversity is now higher due to lower rates of extinction (Hewitt 2000), or if speciation occurs more rapidly in the tropics, species and subspecies could be much younger in the tropics than in the temperate zone, as already hinted by lack of differentiation between *Heliconius* races, and as already known for African cichlids in Lake Victoria (hundreds of speciation events in < 100,000 yrs; Verheyen et al. 2003).

3) *Peripheral speciation should lead to younger Andean species than Amazonian species*. If either Fjeldså's (1994) hypothesis for an "Andean species pump", or Mayr's peripheral founder speciation model (Mayr 1954) were correct, the ages of taxa in the periphery of the range, for instance in Andean valleys, should be younger than those of Amazonian "museum" taxa; they should have smaller branch lengths to their nearest close relative. If on the other hand, centrifugal diversification were more common, Amazonian taxa should be younger than their relatives in Andean valleys. These two clear alternatives can again be tested by means of genetic data. In addition, founder effect speciation (Mayr 1954) should lead to the younger Andean species having lower levels of polymorphism than ancestral, central forms. Replicated patterns demonstrating older Amazonian forms, coupled with correlated ages of divergence, would also provide additional evidence for a general, vicariance explanation (prediction 1).

Methodology and approach

The system: the suture zone for Ithomiinae and Heliconiina in the Río Mayo and adjacent Amazonia. The region under study forms a well-defined "suture zone" (region of correlated contact zones between many pairs of taxa – Remington 1968, Hewitt 2000) only 20 km across near the easternmost cordilleras of the Andes. The area hasbecome a focus for individual studies of speciation in butterflies (*Heliconius*) and frogs (*Dendrobates*) (Mallet & Barton 1989, Mallet et al. 1990, Joron et al. 2001, 2002, Symula et al. 2001, Joron 2002, Mallet & Lamas 2002); in fact we are the first to demonstrate the existence of this striking biogeographic anomaly. We have found 140 species of Heliconiina and Ithomiinae butterflies in this region over the last 20 yrs.

The suture zone (Fig. 1) lies between two lowland regions of endemism below 1000m altitude: the Río Mayo/upper Río Huallaga valley floor (hereafter "Andean") and the Río Ucayali/lower Río Huallaga ("Amazonian") floodplains. The key feature of the system is the repeated pattern of differentiation across the suture zone: 38 pairs of "contact zone" taxa show contact or hybrid zones between Andean and Amazonian geographic "subspecies" across the suture zone (some of these are in fact good species that hybridize rarely or not at all). Another 37 "widespread" species occur in both Andean and Amazonian regions with no obvious morphological differentiation. A further 41 species have only Andean, and 20 have only Amazonian distributions.

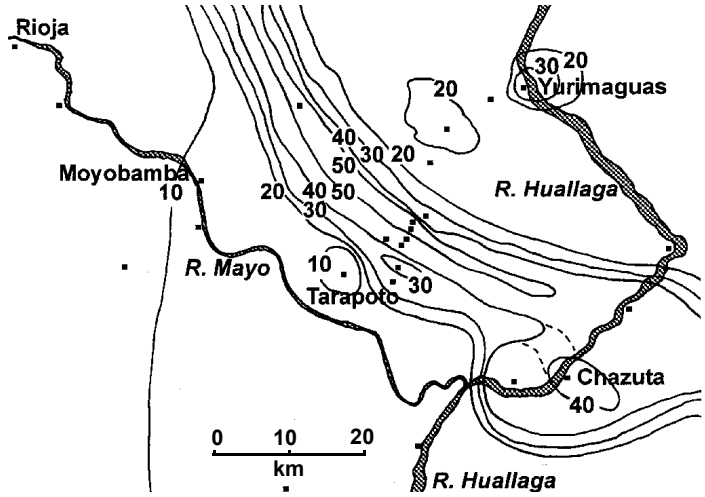

Fig. 1. Contours of % of taxa hybridising in a butterfly suture zone between Amazonian (NE) and Andean (SW) faunas near Tarapoto, Peru. The two zones of endemism were interpreted as the *Ucayali* and *Huallaga* Pleistocene Refuges (Brown 1987). The suture zone runs along the eastern slopes of the eastern Andes, SW of Yurimaguas.
Sources: Joron (2002), Mallet & Lamas (2002).

Samples. It is proposed to test these biogeographic predictions by obtaining sequences from a sample of the above species. For each species or geographic race used, we will extract DNA from a total of 3 representatives from at least two sites. We will sample both Andean and Amazonian races in each of 15 "contact zone" species (i.e. 15 x 2 x 3 = 90 inds.; species names are omitted here to save space; see Joron 2002, Mallet & Lamas 2002 for species names). We will also sample 10 purely "Andean", and 10 purely "Amazonian" species (totalling an additional 10 x 3 + 10 x 3 = 60 inds.), totalling 150 individuals overall.

Molecular markers. For each individual, DNA will be extracted and amplified to obtain COI & COII mtDNA coding sequences, nuclear coding sequences *wingless* and *elongation factor-1α* (*EF-1α*), and nuclear introns *triose phosphate isomerase* (*Tpi*) and *cubitus interruptus* (*Ci*) (Table 1) using PCR and sequencing methods already well-tested in our lab (Cho et al. 1995, Brower 1996, Brower & Egan 1997, Beltrán et al. 2002, Bull 2003).

Table 1. Molecular markers.

Locus	*Size (bp)*
COI+COII (mt, coding)	1600
wingless (coding)	400
EF-1α (coding)	1150
Tpi (intron)	550
Ci (intron)	850-1100

This study is embedded within an international collaborative program to study ithomiine and heliconiine genetics, phylogeny and speciation. By 2004-5, expressed sequence tag and microsatellite (W.O. McMillan lab., Puerto Rico; J. Mavárez, UCL/STRI), as well as 5x coverage BAC (https://www.fastlane.nsf.gov/servlet/showaward?award=0208388, M. Goldsmith) libraries will be available for *Heliconius*. In addition, detailed α-taxonomy and phylogeny is being revised for the whole group (G. Lamas, Peru; A. Freitas, Brazil; K. Willmott, UCL/NHM), while species-level molecular phylogenetic and genomic mapping studies in *Heliconius* and other heliconiines (M. Beltrán, UCL/STRI-Panama; W.O. McMillan, Puerto Rico; M. Joron, UCL), *Oleria* and *Hyposcada* (A.Whinnett, UCL), *Melinaea* and *Hypothyris* (M. Zimmermann, UCL), and *Ithomia* (Chris Jiggins, Edinburgh) are well underway under other funding arrangements. Andrew Brower (Oregon) has an NSF grant to obtain a generic phylogeny of the Ithomiinae based on sequence data (particularly COI/COII, *EF-1α* and *wingless*, and we are actively exchanging material with his group. The current project will both contribute to and benefit from the overall program, which will relatively soon obtain sequence information for all approx. 390 ithomiine and heliconiine species, including many of their subspecies. Thus we will be able to exploit a common pool of sequence information from many laboratories around the world to place the species and subspecies studied in an overall phylogeny. The current study will itself also generate many data useful in these other collaborative projects. At the very least, we will soon possess much more useful data than that from the Sibley-Ahlquist bird DNA/DNA hybridization studies, the only comparable large-scale molecular data from South America (Fjeldså 1994).

Data analysis. Sequence data will be analysed using a combination of phylogenetic methods and coalescent-based population genetics approaches. Sequence distance will be estimated for each species pair at each locus using software packages such as PAUP* and PHYLIP (see Beltrán et al. 2002, Bull 2003). To test for locus-specific effects, we will perform analyses with five loci for each taxon pair. As a simple analysis, the estimated distances can then be subjected to analysis of variance, to estimate and test for locus- and lineage-specific effects. Locus effects account for variation in evolutionary rates among loci, while lineage effects account for differences between species pairs. Interactions between the two effects will be due to variation of individual locus rates between lineages, and to variable times of stochastic coalescence. The null hypothesis of a common divergence time in all pairs of species implies no lineage effect, and will be used to test prediction 1 using a standard ANOVA F-test. This analysis is robust, but lacks power since it does not account for sampling errors or ancestral polymorphisms when estimating pairwise distances.

A more sophisticated analysis will employ recent major advances in implementing Monte-Carlo Markov-chain algorithms for analyzing sequence data of closely related species (Nielsen and Wakeley 2001; Yang 2002; Rannala and Yang 2003). These methods use coalescent models to account for ancestral polymorphism, and can estimate species divergence times (scaled by the mutation rate) while accounting for polymorphisms in ancestral populations. The method of Nielsen and Wakeley (2001) is limited to two species, but can accommodate gene flow after divergence of the two species. The method of Yang (2002) and Rannala and Yang (2003) is applicable for species trees of any size but does not account for gene flow after species divergence. This is currently an area of active statistical research, so I have arranged to collaborate with my UCL colleague Ziheng Yang, who is already collaborating on a similar study of diversification of Madagascar primates with Dr. Ann Yoder. We will use the approach of Yang (2002) and Rannala and Yang (2003) to estimate scaled species divergence times and to perform likelihood ratio tests of the null hypothesis that divergence dates are identical among species pairs (prediction 1). These phylogenetic methods are also required to estimate relative age of species via branch lengths (peripheral isolation vs. centrifugal speciation, tropical vs. temperate; predictions 2-3): species with longer estimated branch lengths to their nearest relative are "older"; species with shorter branch lengths are "younger" (c.f. Fjeldså 1994).

Although this work will be conducted using molecular divergence as a surrogate for actual time of divergence, the accuracy of the molecular clock, and variability in the rates of molecular evolution is another area under active investigation. Lineage-specific rate variation is found between major groups of mammals, but is not strong within groups (within primates, for example; Yoder & Yang 2000), so we should not expect much rate variation within Ithomiinae or Heliconiina. The majority of dates from insect mtDNA depend on the clock calibration by Brower (1994). We will check the accuracy of this calibration by examining data emerging from other recent projects on the same and related taxa. The timing of the orogeny of the Andes provides a useful upper bound of about 10-20 My before present for time of divergence between sister taxa across the Andes in this region (e.g. sister taxa in W. Ecuador and E. Ecuador, for example).

This proposal is an invited resubmission. This is an update of an earlier proposal, submitted in Nov 2002, and given "invited resubmission" status. I here provide more detail on statistical analysis, answering the major criticism of the committee. In particular, I discuss in detail how we will deal with variation in rate of evolution among loci and among lineages. I would like to emphasise that my proposal employs a simplified, highly replicated design, and does not (and never did) attempt to generate a complete biogeography for the whole of South America at this stage; I concentrate on answering clear alternative hypotheses using a reductionist and highly replicated analysis of a single suture zone, consisting of multiple pairs of differentiated taxa. The suggestion of the committee that I use "Brooks Parsimony Analysis", mainly relevant to finding "centres of origin," has therefore not been taken up. Some reviewers argued we lack expertise in molecular genetics of Lepidoptera; here, I refute this, and have chosen genes which in our experience with *Heliconius*, *Melinaea*, *Oleria*, *Ithomia* and *Hypothyris* provide a good balance of both deep and shallow phylogenetic information for taxa that speciated in the last 5 million years. Furthermore, I name members of the international Lepidoptera genetics community

with whom we are sharing information. A suggestion was also made that I collaborate with a "comparative biologist", but the novel design of this study demands more work on statistical analysis of pairwise molecular divergence and coalescence theory than on comparative methods. Thus I have instead formalised a collaboration with my UCL colleague Ziheng Yang, a world leader in bioinformatics and analysis of molecular evolution.

Plan of research

The postdoc position will be advertised and I will hire as soon as possible. Starting in November 2003, the equipment will be purchased, and we will continue to extract and sequence material already collected (August 2002). In Jan-Mar 2004 (when air-fares are "off-peak"), we will carry out the remaining field work in the Tarapoto region. The remainder of the grant period will involve curation of collections, DNA extraction, amplification, sequencing and data analysis.

Justification of resources sought

A *postdoc*, experienced in efficient, high-throughput PCR and sequencing and well-versed in molecular evolution theory, is essential for this project. He/she will have had at least two years' postdoctoral experience. Support for my *technician*, Fraser Simpson, an expert in molecular techniques, is required for lab work. The taxonomic expertise of our *visiting researcher* from Peru, Gerardo Lamas, to be paid from non-NERC sources, strongly underpins this project. *Travel.* We must attend meetings to present results and to keep up-to-date in a rapidly moving field. *Fieldwork.* We have obtained about 75% of the samples during a recent field trip to the region (August 2002), but remaining species, which are rarer, will be more difficult. In particular, the Amazonian side is extensively deforested, and suitable sites for lowland races and lowland species can only be reached via day-long marches into forest trails or *lancha* river trips. Nonetheless, we should be able to obtain the samples with one further two-month field trip. *Consumables* are calculated for 200 samples (even allowing for difficulties, this should easily give 150 individuals with the complete information required to complete the study; in the absence of any problems, we will run more species). We will use the Dept. Biology ABI 3700 sequencers for a total of 4,000 sequencing reactions. The high cost of so many sequences is justified by the need to replicate sampling across as many species and genes as possible for 3 individuals per subspecies or species; much of the sequence information will also serve multiple duty for the lab's other phylogenetic and DNA taxonomy projects, as well as the specific hypotheses outlined here. Freezers, gel tanks, field equipment and general consumables are required to replace worn-out equipment and consumables or, in the case of the ultra-cold freezer, to add to our storage capacity for this project.

Management of project, training, and career development opportunities. The Departmental and College financial support systems will regulate accounting. All personnel on the grant will be kept up-to-date with safety and other skills via training programmes frequently run by our Human Resources Department. Former students and postdocs have successfully found careers as internationally respected academics (e.g. W.O. McMillan, C. Jiggins, R. Naisbit), as agricultural scientists (e.g. Igor Emelianov) and as freelancers (e.g. Martin Brookes). As well as providing world-recognised results in the service of the understanding of global biodiversity and climate change, this project will train scientists in molecular expertise, and give a broad background in molecular evolution and field work. These skills open up wide opportunities in productive employment. For example, a former PhD student, Vanessa Bull, has just started as a molecular biologist in the Home Office Forensic Science Service.

Strategic value and anticipated results

Science underpinning sustainable biodiversity and climate change. The extraordinary rate of forest destruction and rapid climate change suggests that human activities will soon cause many extinctions, as well as exacerbating climate change (Myers et al. 2000). Conservation strategies are of two broad types. An obvious simple strategy is to maximize the numbers of species (or diversity measured some other way) in reserves, as in "hotspot" or "complementarity" approaches (e.g. Vane-Wright et al. 1991, Myers et al. 2000). Many existing conservation areas are set up in regions perceived to have high value because they have high rates of endemism. A second approach argues that we should site reserves in climatically stable areas important in past diversification, even if local species diversity is not currently high (Lovejoy

1982, K.S. Brown 1987, Erwin 1991, Fjeldså 1994, Fjeldså & Rahbek 1998). Siting reserves in areas stable during Pleistocene climate change seems sensible, because they may also be buffered against human-induced climate change. In effect, this second approach trades current diversity for future diversity. Now that the Pleistocene refuge hypothesis is largely discredited, at least in its original form, the causes of the Amazon biodiversity hotspot are again open to question. Possibly, conservation based on untested biogeographic theories is better than nothing, but this seems dangerous. It is therefore urgent to fully understand the systems that we value for conservation before it is too late. In my view, understanding neotropical biogeography is crucial for developing conservation strategies.

Anticipated results. The molecular markers proposed here give a range of levels of molecular divergence between related species, and also because they can be amplified reliably, without cloning (except in the case of *Ci*, which is autosomal and has abundant heterozygosity and length variation). COI/COII and the intron at sex-linked *Tpi* are ideal because they evolve rapidly and can be readily amplified. In heterogametic butterfly females, *Tpi* is hemizygous (COI/COII are of course haploid) and therefore requires no cloning to obtain readable sequence (Beltrán et al. 2002, Bull 2003). *EF-1α* and *wingless* coding regions are autosomal in Lepidoptera and are therefore always diploid. They are reliably amplified but evolve slowly, and so are less informative at subspecies level than *Tpi*, but are highly informative at deeper levels. Nonetheless, we have found that both COI and *EF-1α* are highly informative between species in the genus *Hypothyris*, yet neither marker is strongly differentiated even between good species of *Melinaea* that co-occur in sympatry (M. Zimmermann, in prep.; both these ithomiine genera are to be included in the current study). The latter result may indicate that *Melinaea* species, which drive the evolution of major Müllerian mimicry rings in a huge array of other heliconiines and ithomiines (Brown 1987, Joron et al. 2001), form a recent explosive radiation. However, we need more sequence information to verify that these results hold true across the genome.

If the suture zone is due to a historical split between faunas as in refuge/vicariance theories, similar genetic divergences should be found across the zone in many species (prediction 1). It has already been shown that the narrow morphological hybrid zones in *Heliconius erato* and *H. melpomene* have little differentiation at allozymes or mitochondrial DNA (Mallet et al. 1990, 1998, Brower 1996, Bull 2003, Flanagan et al. 2003). Other contact zones studied here give strong deficits of hybrids; in these cases, greater levels of differentiation are likely, as in *H. himera* x *H. erato* hybrid zones in Ecuador (Jiggins et al. 1997). Previous studies of all of the loci used here in a few key species of *Heliconius* and Ithomiinae demonstrate that the range of divergences required for these tests will be available in the other species to be analysed in this study.

To assess divergence accurately, multiple gene regions (we will use 5), and multiple individuals per subspecies (we will use 3) are needed (see also Avise et al. 1998, Edwards & Beerli 2000, Marko 2002). Marine species pairs (Knowlton et al. 1993, Marko 2002) found across the Isthmus of Panama, which divided the Pacific from the Caribbean 3My ago, give an excellent example of the kinds of clustered divergence due to a known vicariance event. If our early results from prediction 1, showing very strong variation in levels of differentiation, consistent across multiple loci, are proved correct, we should find a much more scattered time course of speciation than in these marine taxa, as found in the temperate zone terrestrial biota (Hewitt 2000). With respect to the other two predictions, I believe that many differentiated taxa should on average be much younger in the tropics (prediction 3), because they are not so prone to extinction via ice coverage (Hewitt 2000). I currently have no basis for a guess as to whether Andean species are on average older or younger than Amazonian taxa (prediction 2).

These results will be the first large-scale "DNA taxonomy" survey of a rapidly radiating group of terrestrial species in the tropics. At minimum, this work will provide useful comparisons with benchmarks from many temperate zone studies (Avise et al. 1998, Hewitt 1996), as well as of classic marine studies across the Isthmus of Panama (Knowlton et al. 1993, Marko 2002). If successful, we will provide tests of (1) whether the suture zone of Fig. 1 is caused by one or a few vicariant events, of similar or younger age to those in Panama, or whether it is due to continued evolutionary change throughout the Pliocene and Pleistocene, as observed in temperate regions (Avise et al. 1998, Hewitt 1996), with an

additional environmental effect causing clustered contact zones; (2) whether tropical diversification was more rapid than in temperate regions recently (Hewitt 2000), or whether the high diversity was due to ancestral species diversity before the Pleistocene; and (3) whether Andean species are younger than Amazonian species, as suggested by theories of Fjeldså (1994, Fjeldså & Rahbek 1998).

Publications, long-term stewardship of datasets and outreach. Where possible, interim results will be published, and meetings will be attended. I envisage final results to be published in major papers in top-ranking international journals late in the grant period. I believe we have a good track record in this respect. We will also be write up other papers on molecular evolution, mimicry evolution, Lepidoptera systematics and biogeography. Preprints and additional data will continue to be made available on our website. All sequence data will be stored and made available to the global community online at GENBANK/EMBL. Tissue and DNA extracts will be kept at -80C, and made available as necessary to internationally recognized researchers interested in using the collections for purposes other than those envisaged here. We are on good terms with many Latin American organizations and persons, and have permits for current work in Peru and Panama. We share our results in international talks, and have been recently been involved in such outreach events in Peru, Colombia, Ecuador and Panama, as well as in this country.

References

Avise, J. C. 1994. Molecular Markers. Chapman and Hall, London.

Avise, J. C., D. Walker, and G. C. Johns. 1998. Proc. Roy. Soc. Lond. B 265:1707-1712.

Barraclough, T. G., and A. P. Vogler. 2000. Amer. Nat. 155:419-434.

Barton, N. H., and G. M. Hewitt. 1983. Pp. 341-359 *in* G. S. Oxford and D. Rollinson, eds. Protein Polymorphism. Academic Press, London & New York.

Barton, N. H., and G. M. Hewitt. 1989. Nature 341:497-503.

Beltrán, M. S., C. D. Jiggins, V. Bull, M. Linares, J. Mallet, W. O. McMillan, and E. Bermingham. 2002. Molec. Biol. Evol. 19:2176-2190.

Beven, S., E. F. Connor, and K. Beven. 1984. J. Biogeog. 11:383-399.

Blum, M. J. 2002. Evolution 56:1992-1998.

Brower, A. V. Z. 1994. Proc. Natl. Acad. Sci., USA 91:6491-6495.

Brower, A. V. Z. 1996. Evolution 50:195-221.

Brower, A. V. Z., and M. G. Egan. 1997. Proc. Roy. Soc. Lond. B 264:969-977.

Brown, K. S. 1987. *in* T. C. Whitmore and G. T. Prance, eds. Biogeography and Quaternary History in Tropical America. Oxford Univ. Press, Oxford, U.K.

Brown, W. L. 1957. Centrifugal speciation. Q. Rev. Biol. 32:247-277.

Brumfield, R. T., and A. P. Capparella. 1996. Evolution 50:1607-1624.

Bull, V. J. 2003. Genealogy and speciation in *Heliconius cydno* and *H.melpomene*. PhD Thesis, University of London, submitted June 2003.

Bush, M.B. 2003. The rise and fall of the refugial hypothesis of Amazonian speciation. http://www.fit.edu/biology/bushlab/refugia.html

Capparella, A. P. 1988. Acta Congr. Int. Ornith. 19:1658-1664.

Carson, H. L., and A. R. Templeton. 1984. Ann. Rev. Ecol. Syst. 15:97-131.

Cho, S., A. et al. 1995. Molec. Biol. Evol. 12:650-656.

Colinvaux, P. A. et al. 2001. Amazoniana. 16:609-646.

Coyne, J. A., and H. A. Orr. 1989. Pp. 180-207 *in* D. Otte and J. A. Endler, eds. Speciation and its Consequences. Sinauer Associates, Sunderland, Mass.

Coyne, J. A., and H. A. Orr. 1997. Evolution 51:295-303.

Cracraft, J., and R. O. Prum. 1988. Evolution 42:603-620.

Dasmahapatra, K. K., M. Blum, A. Aiello, S. Hackwell, N. Davies, E. P. Bermingham, and J. Mallet. 2002. Evolution 56:741-753.

Edwards, S. V., and P. Beerli. 2000. Evolution 54:1839-1854.

Endler, J. A. 1982. Amer. Zoologist 22:441-452.

Erwin, T. L. 1991. Science 253:750-752.

Feder, J. L., S. H. Berlocher, and S. B. Opp. 1998. *in* S. Mopper and S. Y. Strauss, eds. Genetic Structure and Local Adaptation. Chapman & Hall, New York.

Fjeldså, J. 1994. Biodiv. Conserv. 3:207-226.

Fjeldså, J., and C. Rahbek. 1998. Pp. 139-160 *in* G. M. Mace, et. al., eds. Conservation in a Changing World Cambridge Univ. Press, Cambridge.

Gascon, C., et al. 2000. Proc. Natl. Acad. Sci., USA 97:13672-13677.

Grant, P. R. 1986. Ecology and Evolution of Darwin's Finches. Princeton, NJ.

Haffer, J. 1997. Biodiv. Conserv. 6: 451-476.

Hall, J. P. W., and D. J. Harvey. 2002. Evolution 56:1489-1497.

Hewitt, G. M. 1996. Biol. J. Linn. Soc. 58:247-276.

Hewitt, G. 2000. Nature 405:907-913.

Howard, D. J., and S. H. Berlocher, eds. 1998. Endless Forms. Oxf Univ. Press, NY.

Jiggins, C. D., M. Linares, J. Mallet, R. E. Naisbit, C. Salazar, and Z. Yang. 2001. Sex-linked hybrid sterility in a butterfly. Evolution 55:1631-1638.

Jiggins, C. D., and J. Mallet. 2000. Trends Ecol. Evol. 15:250-255.

Jiggins, C. D., W. O. McMillan, P. King, and J. Mallet. 1997. Heredity 79:495-505.

Jiggins, C. D., R. E. Naisbit, R. L. Coe, and J. Mallet. 2001. Nature 411:302-305.

Joron, M., I. R. Wynne, G. Lamas, and J. Mallet. 2001. Evol. Ecol. 13:721-754.

Knowlton, N., et al. 1993. Science 260:1629-1632.

Joron, M. 2002. Ithomiine butterflies of Tarapoto, Peru. http://abacus.gene.ucl.ac.uk/joron/mjitho.html

Lovejoy, T. E. 1982. *in* G. T. Prance, ed. Biological Diversification in the Tropics.

Mallet, J. 1993. Pp. 226-260 *in* R. G. Harrison, ed. Hybrid Zones. OUP, New York.

Mallet, J., and N. H. Barton. 1989. Evolution 43:421-431.

Mallet, J., et al. 1990. Genetics 124:921-936.

Mallet, J. & Lamas, G. 2002. Heliconiina and Ithomiinae of Río Mayo & lower Río Huallaga, Peru. http://www.ucl.ac.uk/taxome/ith/tarapoto/sutzontab.pdf

Mallet, J., W. O. McMillan, and C. D. Jiggins. 1998. Pp. 390-403 *in* D. J. Howard and S. H. Berlocher, eds. Endless Forms. Oxford Univ. Press, New York.

Mallet, J., and J. R. G. Turner. 1998. Pp. 262-280 *in* P. R. Grant, ed. Evolution on Islands. Oxford University Press, Oxford.

Marko, P. B. 2002. Molec. Biol. Evol. 19:2005-2021.

Mayr, E. 1954. Pp. 157-180 *in* J. Huxley, A. C. Hardy, and E. B. Ford, eds. Evolution as a Process. Allen and Unwin, London.

Mayr, E. 1970. Populations, Species, and Evolution. Harvard Univ. Press, Mass.

McMillan, W. O., C. D. Jiggins, and J. Mallet. 1997. PNAS, USA 94:8628-8633.

Myers, N., et al. 2000. Nature 403:853-858.

Naisbit, R. E., C. D. Jiggins, and J. Mallet. 2001. Proc. Roy. Soc. B 268:1849-1854.

Naisbit, R. E., C. D. Jiggins, and J. Mallet. 2002. Genetics 161:1517-1526.

Naisbit, R. E., C. D. Jiggins, and J. Mallet. 2003. Evol. Devel. 5:269-280.

Nielsen, R. and J. Wakeley (2001). Genetics 158(2): 885-96.

Nelson, B. W., et al. 1990. Nature 345:714-716.

Parsons, T. J., S. L. Olson, and M. J. Braun. 1993. Science 260:1643-1646.

Patton, J. L., and M. N. F. da Silva. 1998. Pp. 202-216 *in* D. J. Howard and S. H. Berlocher, eds. Endless Forms. Species and Speciation. OUP, New York.

Presgraves, D. C. 2002. Evolution 56:1168-1183.

Price, T. D., and M. M. Bouvier. 2002. Evolution 56:2083-2089.

Rannala, B. and Z. Yang (2003). Genetics: in press. (manuscript and program available at ftp://abacus.gene.ucl.ac.uk/pub/MCMCcoal/)

Remington, C. L. 1968. Evol. Biol. 1:321-428.

Remsen, J. V. 1984. Science 224:171-173.

Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale Univ. Press, New Haven.

Simpson, B. B., and J. Haffer. 1978. Ann. Rev. Ecol. Syst. 9:497-518.

Symula, R., et al. 2001. Proc. Roy. Soc. Lond. B 268:2415-2421.

Taberlet, P., and R. Cheddadi. 2002. Science 297:2009-2010.

Turner, J. R. G. 1971. Pp. 224-260 *in* E.R. Creed, ed. Ecological Genetics and Evolution. Blackwell, Oxford.

Turner, J. R. G., and J. L. B. Mallet. 1996. Phil. Tr. Roy. Soc. Lond. B 351:835-845.

Vane-Wright, R. I., et al. 1991. Biol. Conserv. 55:235-254.

Whitmore, T. C., and G. T. Prance, eds. 1987. Biogeography and Quaternary History in Tropical America. Oxford University Press, Oxford.

Willis, K., and R. J. Whittaker. 2000. The refugial debate. Science 287:1406-1407.

Yang, Z. (2002). Genetics 162: 1811-1823.

Yoder, A. D., and Z. Yang. 2000. Estimation of primate speciation dates. Molec. Biol. Evol. 17:1081-1090.

Zink, R. M. 1996. Evolution 50:308-317.

Zink, R. M., and J. B. Slowinski. 1995. Proc. Natl. Acad. Sci., USA 92:5832-5835.