The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2011, **102**, 391–403. With 5 figures

Three fishes in one: cryptic species in an Amazonian floodplain forest specialist

MAXINE P. PIGGOTT[1,2], NING L. CHAO[3] and LUCIANO B. BEHEREGARAY[1,4]*

[1]*Molecular Ecology Laboratory, Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia*
[2]*The Fenner School of Environment and Society, Australian National University, Canberra, ACT 0200, Australia*
[3]*Departamento de Ciências Pesqueiras, Universidade Federal do Amazonas, Manaus, AM 69700-000, Brazil*
[4]*Molecular Ecology Laboratory, School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia*

Received 18 May 2010; accepted for publication 19 August 2010

Accurately describing biodiversity in tropical regions such as Amazonia is difficult because of insufficient morphological inventories and the lack of studies on the distribution of genetic diversity. Aquatic organisms from Amazonian flooded forests are generally expected to move laterally along the forests during the annual inundation cycle, a behaviour that should promote admixture of populations and reduce within-drainage speciation. We used an unprecedented fine-scale sampling effort and multiple DNA markers to quantify region-wide population differentiation in an Amazonian floodplain forest specialist, the black-wing hatchet fish *Carnegiella marthae* (Myers, 1927). Our study revealed three previously unsuspected and ancient cryptic species of black-wing hatchet fish in the Rio Negro floodplain (RNF), in central Amazonia. Two species produce occasional first-generation hybrids. The third and rarer species, although found in extreme sympatry with another species, appears to be reproductively isolated, and also differs in external morphology and dentition. Our findings have important implications for guiding conservation management because *C. marthae* is harvested commercially in the RNF ornamental fishery. They also suggest that the diversity of Amazonian ichthyofauna is vastly underestimated, including that found in landscapes lacking contemporary barriers to account for population divergence and speciation. *Biological Journal of the Linnean Society*, 2011, **102**, 391–403.

ADDITIONAL KEYWORDS: Amazon – biodiversity – conservation genetics – microsatellites – phylogeography – Rio Negro – tropical rainforest.

INTRODUCTION

Detecting and measuring biodiversity are fundamental tasks in ecological and evolutionary studies, but can be hampered by cryptic biodiversity. Cryptic species are discrete species that are difficult or sometimes impossible to distinguish morphologically, despite reproductive isolation and sometimes extensive genetic divergence (Colborn *et al*., 2001; Beheregaray & Caccone, 2007; Bickford *et al*., 2007). In addition to biodiversity and conservation issues (Knowlton, 1993; Schönrogge *et al*., 2002; Beheregaray & Caccone, 2007), cryptic species also present some interesting evolutionary questions, including how phenotypic similarity persists in the face of extensive molecular evolution (Taylor, Finston & Hebert, 1998; Colborn *et al*., 2001; Mayer & von Helversen, 2001), and understanding the mechanisms of stasis or convergence promoting morphological similarity (Knowlton, 1993; Bickford *et al*., 2007).

Problems in identifying cryptic species using phenotypic differences can be overcome by the use of molecular techniques and analyses, which have

*Corresponding author.
E-mail: luciano.beheregaray@flinders.edu.au

proved useful for discovering unexpected species diversity (e.g. Taylor *et al*., 1998; Bickford *et al*., 2007; Pfenninger & Schwenk, 2007). This is particularly true in tropical rainforests, where recent reports of diversification imply that species richness in these regions has been under-documented by morphological inventories (Parra-Olea & Wake, 2001; Beheregaray & Caccone, 2007; Sistrom, Chao & Beheregaray, 2009). Although a recent review suggests that the proportion of cryptic species in nature is similar across different biogeographic regions (Pfenninger & Schwenk, 2007), intensive and systematic population sampling is needed in groups of tropical rainforest organisms before such conclusions are made. One such biological group is the ichthyofauna of the Amazon Basin.

The rivers draining Amazonia host the most diverse freshwater fish fauna of the world (Reis, Kullander & Ferraris, 2004), and are likely to harbour many cryptic species (e.g. Fernandes *et al*., 2005; Sistrom *et al*., 2009). The ecological and biogeographic history of Neotropical fishes, and their levels of biodiversity, especially intraspecific genetic diversity, is largely unknown (Beheregaray, 2008). Nonetheless, the increase of anthropogenic perturbations of both terrestrial and aquatic ecosystems has made Amazonia a priority for conservation and biodiversity analysis (Myers *et al*., 2000; Laurance, Powell & Hansen, 2002; Ferraz *et al*., 2003). An important source of freshwater fish diversity in the region is represented by small-sized species inhabiting flooded forests and streams, such as those found in the Rio Negro floodplain (RNF) (Gery, 1984; Smith, 1985; Chao, 2001). The Rio Negro is one of the largest tributaries of the Amazon, extending 1700 km to its headwaters. The river level oscillates annually by up to 16 m, inundating an extensive floodplain that covers an area of approximately 0.75 million km^2 of largely undisturbed primary forest (Filizola, 1999; Latrubesse & Franzinelli, 2005). This hot spot of fish diversity is the seat of a thriving ornamental fishery that provides about 60% of the income of local riverine people (Chao, 2001). One ornamental species used in this fishery is the blackwing hatchet fish *Carnegiella marthae* (Myers, 1927). This small characin species is found exclusively in black-water streams and flooded forests of the Negro and upper Orinoco River basins (Géry, 1977; Weitzman & Palmer, 2003). This strangely shaped species is popular in the aquarium trade and represents a valuable resource for ornamental fishermen from middle Rio Negro, in Brazil (Chao, 2001).

Being a floodplain forest specialist, *C. marthae* is expected to move laterally in the RNF during the annual inundation cycle (Latrubesse & Franzinelli, 2005; Marshall, Forsberg & Thome-Souza, 2008), a behaviour that should promote admixture of populations at a regional level and reduce the likelihood of allopatric speciation. In this study, we quantify region-wide population differentiation and test for cryptic diversity in this floodplain forest specialist using DNA data derived from nuclear microsatellite markers, mitochondrial DNA (mtDNA) sequences, and morphological information. Testing for cryptic lineages relies on a comprehensive spatial assessment of reproductive isolation. We used a fine-scale sampling regime covering tributaries along the entire known distribution of the species in the Rio Negro, including its remote headwaters region. Whenever possible, population samples were obtained from tributaries equidistantly separated, and also located on opposite river margins. Unexpectedly, we present evidence for highly differentiated and divergent population groups representing three cryptic species of black-wing hatchet fish in the Rio Negro basin. We discuss the implications of our findings for the conservation management of *C. marthae* populations, and the role of contemporary drainage structure as a predictor of cryptic biodiversity in Amazonian floodplain-dependent organisms.

MATERIAL AND METHODS

Sampling *C. marthae* populations for genetic analysis

Carnegiella marthae is found in forest streams (*igarapés*) and flooded forest habitat (*igapós*) in the middle, upper, and headwater regions of the Rio Negro basin (Chao, 2001). A total of 403 *C. marthae* individuals were collected from 21 tributary populations in the Rio Negro basin (central Amazonia, Brazil) in 2002 and 2004 (Fig. 1; Table 1). This sampling effort essentially covers the entire distribution of the species in the Rio Negro basin. This taxon is also reported for the Orinoco basin (Géry, 1977), based on specimens collected in the confluence between the Negro and Orinoco basins, only ~40 km from one of our sampling sites in the Rio Negro's headwaters. Fish were collected in the flooded forest (*igapó*) with hand nets. Importantly, all individuals representing each tributary sample were caught at the very same site in the *igapó* (e.g. they were schooling together at the time of sampling). All fishes were collected by the senior author (L.B.B.), and no samples from the ornamental fishing trade were used. A small piece of muscle tissue was taken from behind the dorsal fin of each fish and preserved in 95% ethanol. In addition, six specimens of the only other congeneric species reported to occur in the RNF, the marbled hatchet fish *Carnegiella strigata* (Günther,

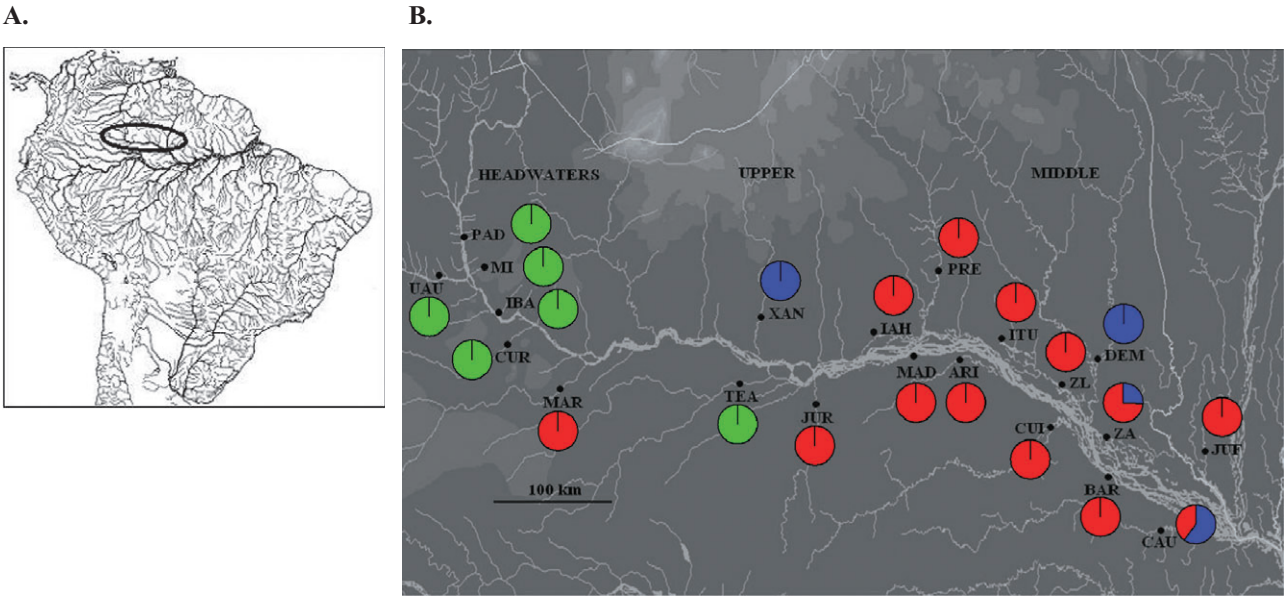

Figure 1. A, location of the Rio Negro in northern South America. B, sampled localities in the Rio Negro, central Amazonia, Brazil. Tributary abbreviations are as listed in Table 1. Also shown is the proportion of each *Carnegiella marthae* population sample assigned to groups A, B, or C, based on analyses of microsatellite data (see text for details). Group A is represented by blue (dark grey in print version) , B by green (pale grey), and C by red (white).

Table 1. Sampled localities for *Carnegiella marthae*, including population abbreviation, sample size, geographic co-ordinates, margin of the Rio Negro where collected (facing downstream), and genetic group based on STRUCTURE and principle components analysis

Locality	Population abbreviation	*N*	Geographic coordinates	Region/margin	Genetic group
Jufari	JUF	20	00.59.40 °S, 62.06.10 °W	Midwaters/left	C
Caurés	CAU	20	01.19.01 °S, 62.24.54 °W	Midwaters/right	A, C
Zamula	ZA	20	00.51.57 °S, 62.46.22 °W	Midwaters/left	A, C
Baruri	BAR	20	00.53.35 °S, 63.03.38 °W	Midwaters/right	C
Demini	DEM	20	00.23.40 °S, 62.51.17 °W	Midwaters/left	A
Zalala	ZL	20	00.39.59 °S, 63.00.32 °W	Midwaters/left	C
Cuiuni	CUI	20	00.46.09 °S, 63.10.40 °W	Midwaters/right	C
Arirahá	ARI	20	00.26.54 °S, 63.41.43 °W	Midwaters/right	C
Itu	ITU	20	00.26.00 °S, 63.07.00 °W	Midwaters/left	C
Preto	PRE	20	00.06.40 °S, 64.05.03 °W	Midwaters/left	C
Iahá	IAH	20	00.23.47 °S, 64.36.26 °W	Midwaters/left	C
Madiquié	MAD	20	00.25.48 °S, 62.24.04 °W	Midwaters/right	C
Jurabaxi	JUR	20	00.33.07 °S, 64.48.06 °W	Midwaters/right	C
Arixana	XAN	20	00.21.50 °S, 62.11.51 °W	Midwaters/left	A
Tea	TEA	20	00.32.59 °S, 65.15.13 °W	Midwaters/right	B
Niaua-mirim	MI	20	00.08.24 °S, 66.54.91 °W	Headwaters/left	B
Curicuriari	CUR	3	00.13.35 °S, 66.24.58 °W	Headwaters/right	B
Marié	MAR	21	00.26.32 °S, 66.24.58 °W	Headwaters/right	C
Uaupés	UAU	20	00.04.47N, 67.24.13 °W	Headwaters/right	B
Paduá	PAD	20	00.12.23N, 67.19.23 °W	Headwaters/left	B
Igarapé Ibará	IBA	20	00.20.15 °S, 66.35.04 °W	Headwaters/left	B

1864), were collected from the tributary Marié (MAR), as a comparison in the analyses.

Microsatellite genotyping and mitochondrial DNA sequencing

DNA extraction of tissue samples was carried out using a modified salting-out method (Sunnucks & Hales, 1996). All samples were genotyped at four microsatellite loci (*Cm8*, *Cm6*, *Cm10*, and *Cm20*; Beheregaray *et al*., 2006). These markers were the only loci out of a panel of seven that amplified well for all samples. Polymerase chain reactions (PCR) contained 1 μL of DNA (~50–100 ng), 200 μM of dGTP, dTTP, and dCTP, 20 μM dATP, 2.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5 pmol of each primer, 0.1% Triton X-100, 0.05 μl [α-^{33}P]dATP at 1000 Ci mmol^{-1}, and 0.5 U of *Taq* polymerase in a total volume of 10 μL. Annealing temperatures are given in Beheregaray *et al*. (2006). PCR products were electrophoresed through a 6% polyacrylamide sequencing gel and visualized by autoradiography. Allele sizes were scored against an A- or T-terminating M13 control sequencing reaction size marker.

We also obtained sequence data from one or two individuals per tributary for two mtDNA genes: the *adenosine triphosphatase subunits 6* and *8* (*ATPase6* and *ATPase8*). Data from the mitochondrial genome were obtained to assess evolutionary distinctiveness between any population groups depicted in the analyses of microsatellite data. *ATPase6* and *ATPase8* fragments were amplified with primers ATP8.2 and CO3.2 (Martin & Bermingham, 1998) using PCR conditions described in Corrigan *et al*. (2008). All PCR products were purified using an ULTRA CLEAN purification kit (Mo Bio Laboratories, Carlsbad, CA, USA) and sequenced in both directions on an ABI377 sequencer, following the manufacturer's instructions.

Species distinction

Prior to carrying out analyses to clarify the number of species collected, multilocus genotypes were tested for linkage disequilibrium (option 2) using GENEPOP v3.4 (Raymond & Rousset, 1995). Genotypes for each pair of loci were tested for independence in each population using a contingency test, and significance values were calculated using Fisher's exact test. We used our microsatellite data to carry out two different analyses based on individual assignment and allelic similarities. The Bayesian clustering method in STRUCTURE v2.1 (Pritchard, Stephens & Donnelly, 2000) was used to test the assignment of each individual and determine the number of genetically distinct groups of *C. marthae* in the Rio Negro. We set most parameters to their default values, but specifically chose the admixture model and correlated allele frequencies to allow for individuals with mixed ancestry. The length of the initial burn-in period was set at 50 000 iterations, followed by a run of 1 000 000 Markov chain Monte Carlo repetitions with five runs to determine the level of variation of the likelihood for each K value. The range of possible K values that we tested was from 1 to 15. We used the method described by Evanno, Regnaut & Goudet (2005) to determine the estimated K value for our data, as using the maximal value of $L(K)$ to identify the true number of populations (K) may not provide a correct estimation of the number of clusters. We also looked at the degree of similarity of all individuals based on allele frequencies using a principal components analysis (PCA) in GENALEX v6 (Peakall & Smouse, 2006). For the latter analysis we also included genotypes from the six *C. strigata* samples for comparison.

Secondly, we conducted a phylogenetic analysis to assess evolutionary distinctiveness among mtDNA lineages. The *ATPase6* and *ATPase8* sequences were aligned using SEQUENCHER v4.1 (Gene Codes Corporation, Ann Arbor, MI, USA). Maximum parsimony (MP) and maximum likelihood (ML) phylogenetic analyses were conducted in PAUP* v4.0b10 (Swofford, 1998). The MP and ML trees were obtained with the branch-and-bound search, with 100 and ten random-addition sequence replicates, respectively, and tree–bisection–reconnection branch swapping. As the model of evolution, we used HKY85 (Hasegawa, Kishino & Yano, 1985), based on the results of the Akaike information criterion, as implemented in MODELTEST v3.07 (Posada & Crandall, 1998). Node support for different reconstructions was assessed by generating 10 000 bootstrap replications (Felsenstein, 1985). The out-group taxon used in all reconstructions was the marbled hatchetfish *C. strigata*.

Genetic differentiation between putative species

To determine the relative level of genetic differentiation within and between the putative genetic groups determined by STRUCTURE and PCA, we carried out a hierarchical analysis of molecular variance (AMOVA) using ARLEQUIN v2.0 (Schneider, Roessli & Excoffier, 2000). The degree of genetic differentiation was calculated using F_{ST} statistics, based on the estimator θ (Weir & Cockerham, 1984), and R_{ST} statistics (Slatkin, 1995), based on differences in allele size. We used R_{ST} in addition to F_{ST}, as it may be more appropriate for studying the levels of genetic variation under the stepwise mutation models thought to apply to microsatellites. To assess whether R_{ST} performs better than F_{ST} on this data set, we carried out an allele permutation test (Hardy *et al*., 2003) using

SPAGEDI v1.2 (Hardy & Vekemans, 2002). SPAGEDI v1.2 uses random permutations to reassign allele size information among alleles. Significantly smaller permuted (pR_{ST}) than observed R_{ST} values suggest that at least a partial stepwise mutation process has contributed more to the observed differentiation than genetic drift alone (Hardy *et al.*, 2003). One-tailed tests were used to assess whether R_{ST} was equal to pR_{ST}.

Morphological analysis

A morphological examination was conducted in 100 *C. marthae* specimens collected during a previous expedition of one of the authors (N.L.C.), and deposited at the fish collection of the Universidade Federal do Amazonas (Manaus, Brazil). These specimens were sampled on several occasions from the same site sampled for our genetic analysis in the Igarapé Zamula (ZA) (Fig. 1). The examination revealed two colour morphs: one dark morph and one light-coloured morph, herein referred to as the white morph (Fig. 2). We compared morphometric and meristic characters between these morphs using standard protocols (Hubbs & Lagler, 1958) based on 50 dark and 50 white morph specimens. Additional osteological comparisons on dentition were made with five fishes per morph. The jaw bones were removed from the fishes, and then cleared and stained for observation, following the method described by Taylor & Van Dyke (1985). This analysis was conducted to test for differences in dentition between the two morphs potentially related to foraging specializations.

RESULTS

Three species in one taxon

Microsatellite variability ranged from 6 to 49 alleles per locus (Appendix S1). Although four locus pairs showed significant evidence of linkage disequilibrium in the 126 tests performed, this was not consistent for any particular locus or population, which indicates

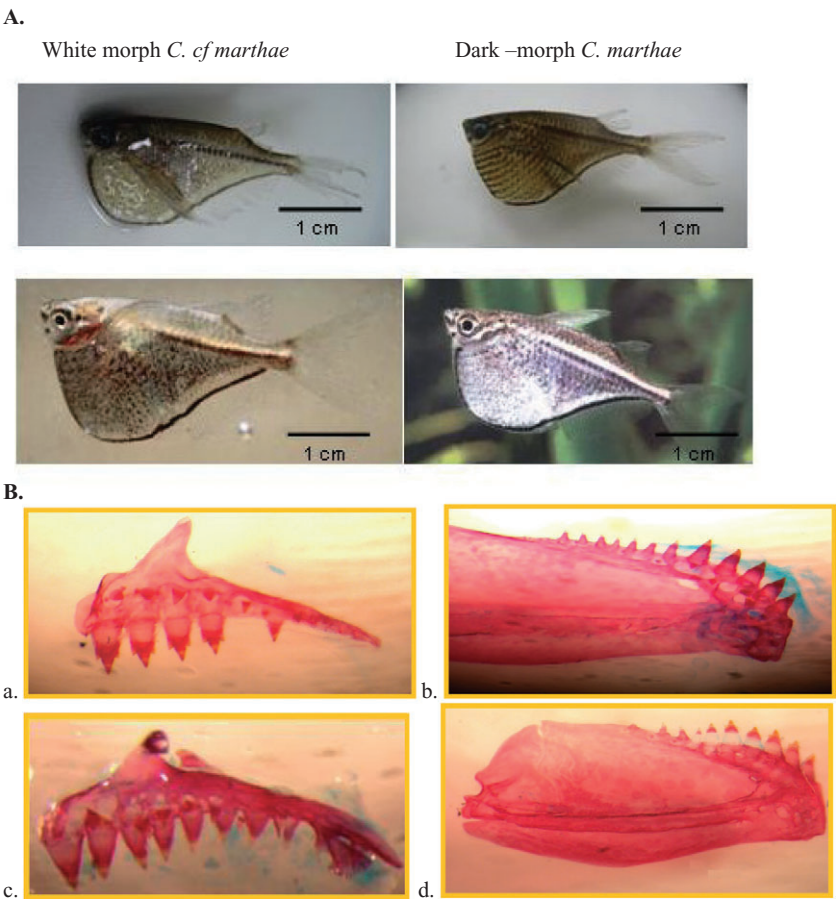

Figure 2. A, live and preserved morphs of *Carnegiella marthae* in the Rio Negro: 'white' (i.e. light-coloured) morph and 'dark' morph. B, clear and stained specimens showing differences in dentition between the two morphs of *C. marthae*: a and b are the premaxillary and the dentary of the dark morph; c and d are the premaxillary and the dentary of the white morph.

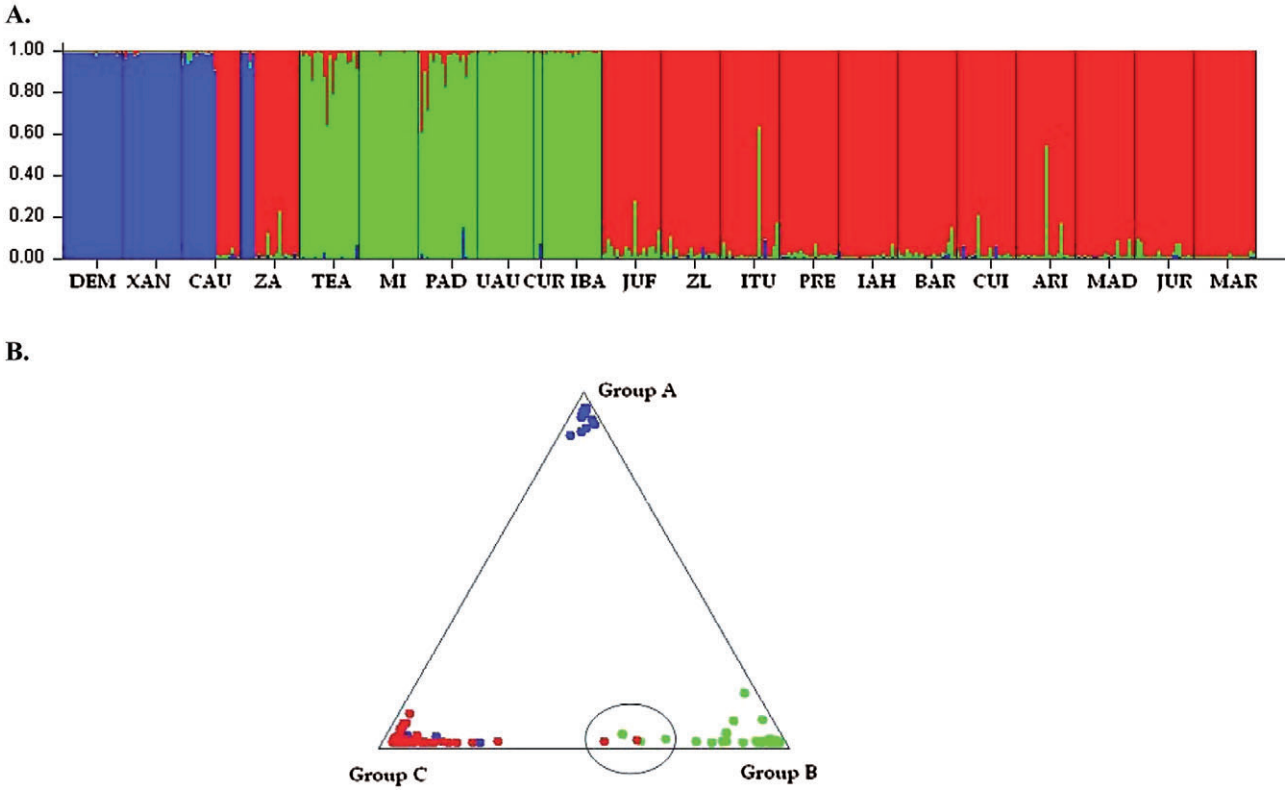

Figure 3. A, a STRUCTURE (Pritchard *et al*., 2000) bar plot of population assignment showing three distinct genetic clusters from the 21 populations of *Carnegiella marthae* sampled in the Rio Negro. B, STRUCTURE triangle plot showing three distinct genetic clusters and hybrid individuals between groups B and C.

that they can be considered as independent markers. The Bayesian clustering method using microsatellite data identified three distinct genetic clusters in *C. marthae* (Fig. 3A, B). Overall, the majority of individuals assigned to one of the three genetic clusters in STRUCTURE had a very high probability of assignment (group A, mean $q = 0.985$, range 0.939–0.992; group B, mean $q = 0.960$, range 0.554–0.992; group C, mean $q = 0.969$, range 0.578–0.991). The six *C. strigata* samples were excluded from the STRUCTURE analysis, as the most likely number of genetic groups could not be resolved with their inclusion. This is most likely as a result of the failure of two loci to amplify for the *C. strigata* samples: *Cm20*, which has diagnostic alleles for each *C. marthae* genetic group (see below), and *Cm10*. PCA also revealed three genetic groups that matched the individuals assigned to the three STRUCTURE clusters, and discriminated one group represented exclusively by the six *C. strigata* samples (Fig. 4).

Each *C. marthae* genetic group identified by the Bayesian clustering method and the PCA had diagnostic microsatellite alleles for locus *Cm20*: alleles 127 and 129 were diagnostic for group A, alleles 125, 169, and 170 were diagnostic for group B, and allele 123 was diagnostic for group C (Appendix S1; Table 2). This pattern was observed when comparing populations across tributaries and, importantly, also in sympatric comparisons. Six individuals (found in populations ZA, ITU, JUR, MAD, and ARI) were heterozygous for the diagnostic locus *Cm20* with an allele from group B and one from group C (e.g. 123/125, 123/169; Fig. 3B; Table B), and also grouped

Table 2. Microsatellite alleles (in base pairs) and allele frequencies for locus *Cm20* in *Carnegiella marthae*, separated into genetic groups A, B, and C, and putative first-generation hybrid individuals based on results from STRUCTURE and principle components analysis

Alleles	Group A	Group B	Group C	Hybrid
n	57	102	238	6
123			1.00	0.500
125		0.505		0.333
127	0.958			
129	0.042			
169		0.490		0.167
170		0.005		

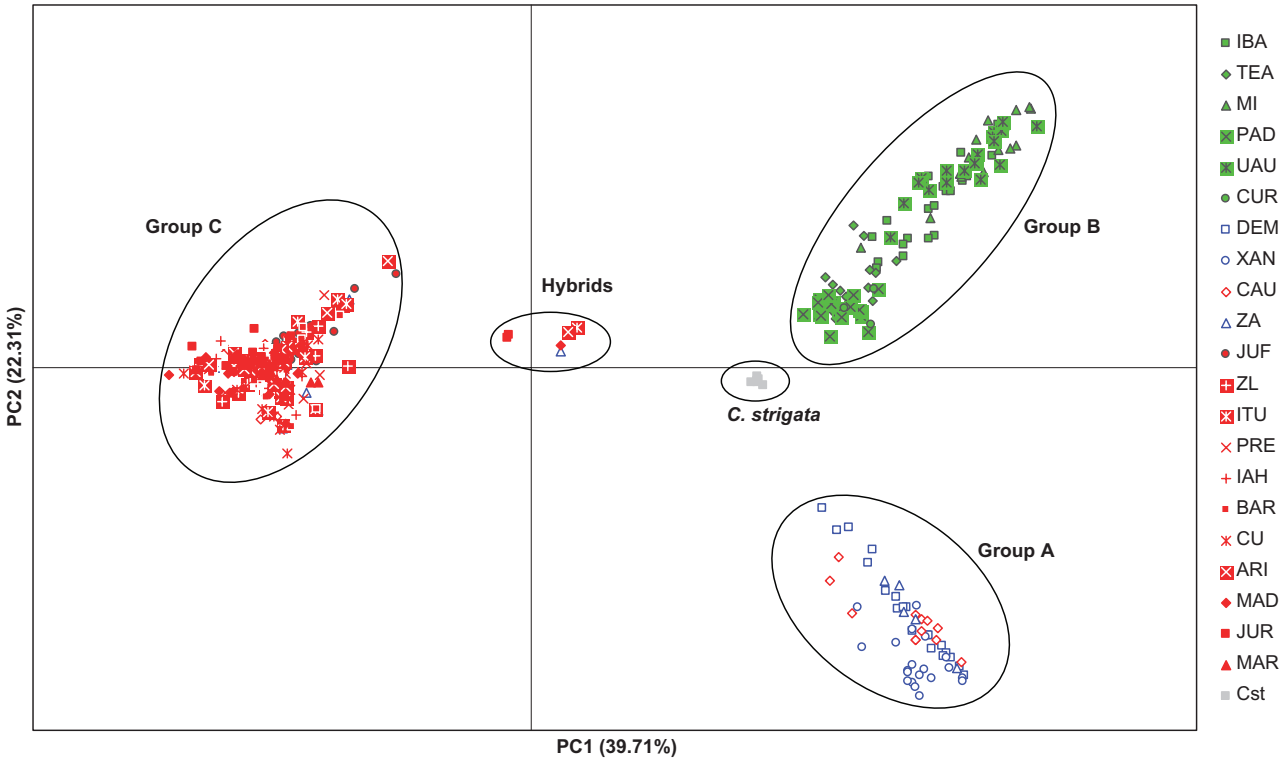

Figure 4. Principal components analysis based on the genetic distances of 403 *Carnegiella marthae* individuals from 21 populations and six *Carnegiella strigata* individuals from one population in the Rio Negro. Also shown are the hybrids between groups B and C.

Table 3. Hierarchical analysis of molecular variance (AMOVA) using F_{ST} and R_{ST} within and between the three *Carnegiella marthae* genetic groups determined by STRUCTURE

Statistic	Source of variation	d.f.	Sum of squares	Variance	% variation	*P*
F_{ST}	Among groups	2	215.320	0.45494	23.51	< 0.001
	Among pops. within groups	20	133.058	0.15337	7.93	< 0.001
	Within pops.	783	1 038.485	1.32629	68.56	< 0.001
R_{ST}	Among groups	2	399 183.227	860.41072	52.17	< 0.001
	Among pops. within groups	20	173 070.495	233.16190	14.14	< 0.001
	Within pops.	783	435 182.082	555.78810	33.70	< 0.001

F_{ST} fixation indices: F_{ST}, 0.31442; F_{SC}, 0.10365; F_{CT}, 0.23514.
R_{ST} fixation indices: R_{ST}, 0.66303; R_{SC}, 0.29553; R_{CT}, 0.52166.

together separately from the other three genetic groups in the PCA (Fig. 4). This suggests that there may be the potential for hybridisation between groups B and C. There was no evidence of possible hybridization between group-A fishes with fishes from either groups B or C.

For two of the three groups marked divergence was detected in sympatry (see below), which strongly indicates reproductive isolation between them. In addition, the remaining three microsatellite loci also showed marked differences in allele frequency distributions between the three groups (Appendix S1). As expected based on the results from previous analyses, there was highly significant genetic differentiation between the three genetic groups using *F* and *R* statistics (Table 3). There was a large difference between the F_{ST} and R_{ST} values, with the R_{ST} value (52.17%) being much higher than the F_{ST} value (23.51%) (Table 3). Permutation tests of allele sizes were significant over all loci ($R_{ST} = 0.476$, $P < 0.001$), and individually for three loci, *Cm8* ($R_{ST} = 0.417$, $P < 0.001$), *Cm10* ($R_{ST} = 0.368$, $P < 0.001$), and *Cm6*

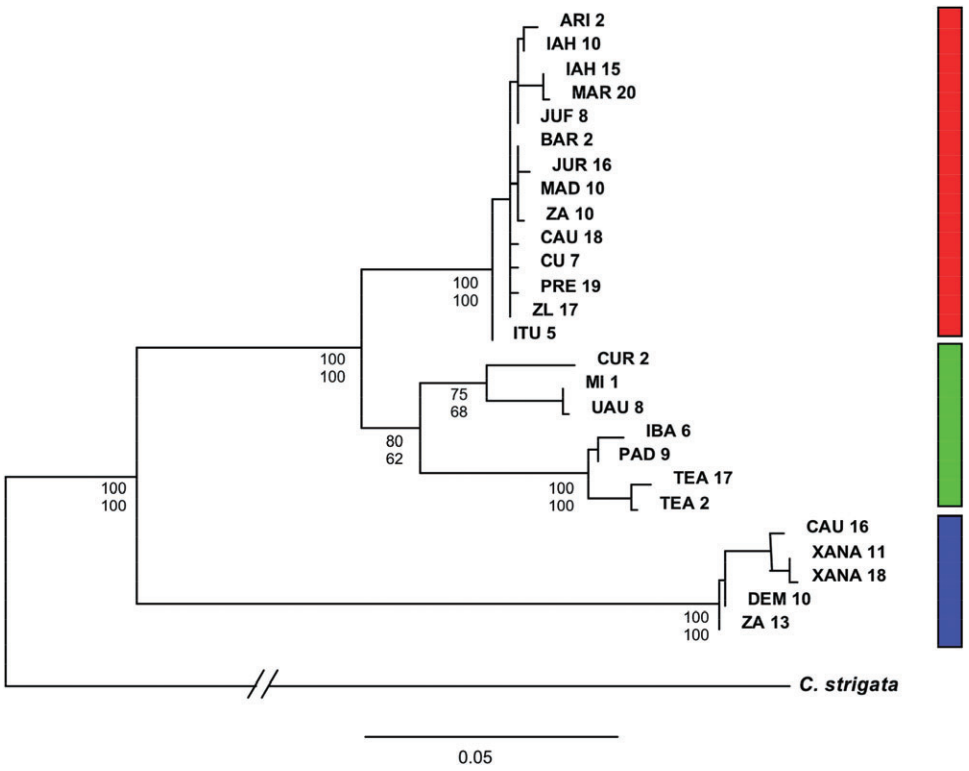

Figure 5. Maximum-likelihood tree based on 803 bp of the mitochondrial DNA *ATPase6* and *ATPase8* genes of *Carnegiella marthae*. Maximum-likelihood and maximum-parsimony bootstrap support values for the major clades are given below branches. Tributary abbreviations are as listed in Table 1, and numbers are the specimens used for sequencing. The genetic groups A, B, and C, identified based on analyses of microsatellite data, are also shown. The marbled hatchet fish *Carnegiella strigata* is the out-group.

($R_{ST} = 0.537$, $P < 0.001$). Locus *Cm20* was not significant, which probably reflects its low allelic diversity compared with the other three loci (Appendix S1).

The geographic distribution of the three genetic groups in the Rio Negro is provided in Figure 1B and Table 1. Group A is composed of individuals from four populations located in the middle and upper RNF, whereas individuals from six populations were assigned to group B, which was found predominantly in the headwaters of the basin (Fig. 1B). The most widespread group (C) had individuals assigned from 13 populations and spanned mostly the middle of the RNF, but also had tributary populations from the upper section and headwaters. Generally, a population was composed of individuals from one genetic group only, except for populations ZA and CAU, which had individuals assigned to either group A or C (Fig. 1B).

The mtDNA phylogenetic reconstructions provided strong support for the three genetic groups identified in population analyses using nuclear DNA data. Here, 26 *ATPase6* and *ATPase8* haplotypes with a combined length of 783 bp were characterized for 30 individuals representing the 21 tributaries. This included six ZA individuals: three from the dark morph and three from the white morph. Both phylogenetic methods resulted in trees with identical topologies that characterize three divergent clades (A, B, and C; Fig. 5). The mtDNA clade membership for each of the 30 individuals matched exactly their assignments to the genetic groups A, B, and C identified with microsatellite data (Fig. 5; Table 1). Clades A and C were supported by very high bootstrap values (100% for both ML and MP), whereas bootstrap support for clade B was lower (80 and 62%, respectively). As expected from the topology of the tree, very deep levels of divergence were observed between haplotypes from different clades (9–19.5% HKY85 sequence divergence), and lower levels of divergence were observed between haplotypes from the same clade (0.13–6.8%). In addition to the three major groups, the mtDNA phylogenetic analysis also suggests that group-B tributary populations are further subdivided into two subclades (Fig. 5). The possibility of genealogical subdivision within group B needs to be verified with additional genetic data, and with an analytical framework that combines both population genetic and phylogeographic approaches.

Table 4. Number of tricuspid and conical teeth on premaxillary and dentary bones of dark and white forms of *Carnegiella marthae* ($N = 5$ per taxa)

	No. of dentary teeth number			No. of premaxillary teeth		
Taxa	Total	Tricuspid	Conical	Total	Tricuspid	Conical
C. marthae Dark	13–14	6	7–8	8–10	1–3	5–8
C. marthae White	12–14	0–4	8–14	8–10	0–4	4–10

Morphological divergence

Our analysis of the two sympatric colour morphs identified in ZA disclosed a pattern of morphological divergence highly consistent with the separation of ZA individuals in genetic groups A and C. Several diagnostic morphological characters are apparent in these two morphs. The dark morph, which is the commonly found *C. marthae* in the RNF, has a dozen dark oblique lines on the side of the body, whereas the white morph has diffused dotted lines (Fig. 2). The white morph can be readily distinguished from the dark morph by having 27–30 ($N = 50$) soft anal fin rays versus 23 or 24 ($N = 50$) in the dark morph. In terms of divergence in foraging traits, the number of tricuspid teeth on the dentary bone is six for the dark morph, but is less than four or they are absent in the white morph (Fig. 2B; Table 4). Several morphometric measurements (data not shown), such as the eye diameter, distance between nares, and dorsal fin base length, are proportionally greater in the white morph, whereas the least distances between dorsal fin base and upper caudal fin base is greater in the dark morph ($P < 0.05$).

For the majority of individuals only the tissue sample, rather than the entire fish, was brought to the laboratory. Nonetheless, we deposited 30 voucher specimens (including three dark and three white morphs from ZA) in the fish collection of the Universidade Federal do Amazonas. The nuclear genotypes and the mtDNA sequences of these ZA individuals were assigned to two genetic groups that correspond to the two morphs: all genetic group-A fishes were from the white morph, whereas all group-C fishes were from the dark morph. Voucher specimens belonging to genetic group B ($N = 5$) are morphologically similar to dark morph fishes, but a better sample is needed to test for morphological divergence between group B and the other two groups.

DISCUSSION

This study used an unprecedented fine-scale sampling effort in central Amazonia to report on three previously unsuspected cryptic species in the black-wing hatchet fish *C. marthae*, a small floodplain forest specialist from the Rio Negro basin. Below we discuss the nature of our molecular and morphological evidence supporting species delimitation, the putative age of these species, and the conservation implications of our findings. We also discuss the result of cryptic speciation within a well-connected floodplain system that lacks contemporary barriers for dispersal, an intriguing discovery with ramifications for guiding biodiversity inventories in Amazonia.

Reproductive isolation, and phylogenetic and morphological divergence

The evidence presented here satisfies a number of different properties used by biologists to delineate species. As a result, the evidence also satisfies the operational criteria used for empirical applications of several species concepts, including the biological species concept (Wright, 1940; Mayr, 1942; Dobzhansky, 1950) and different versions of the phylogenetic species concept (Nelson & Platnick, 1981; Baum & Shaw, 1995), and is also consistent with the unification of species concepts recently proposed by de Queiroz (2007). For instance, analyses based on both nuclear and mitochondrial genomes revealed three groups of separately evolving metapopulation units in our sample (groups A, B, and C). These groups of populations can be readily distinguished based on individual assignment of multilocus genotypes (STRUCTURE; Fig. 3), differences in nuclear allele frequencies (PCA; Fig. 4), and evolutionary divergence between maternal lineages (phylogenetic tree; Fig. 5). Molecular diagnosability is also evident in these three groups. They are diagnosed based on fixed nuclear microsatellite alleles (Table 2) and mtDNA substitutions. In addition, morphological analysis of specimens from groups A and C revealed diagnostic differences in the number of teeth and soft anal fin rays between sympatric samples (Table 4).

The marked phylogenetic distinctiveness of the three groups (e.g. 9–19.5% corrected mtDNA sequence divergence) suggests an old history of speciation in the black-wing hatchet fish. In the absence of obvious external calibration points for our study system (e.g. Teske & Beheregaray, 2009), we assumed a tentative molecular clock for the mtDNA *ATPase* data set of 1.3%

per million years. This *ATPase* clock was proposed based on analyses of groups of geminate fishes found across the Panama Isthmus (Bermingham, McCafferty & Martin, 1997). This clock suggests split events of Miocene age in the black-wing hatchet fish, ranging from around 7 Mya (between groups B and C) to 15 Mya (between groups A and C). This suggestion of ancient diversification is corroborated by the large difference between R_{ST} and F_{ST} values observed in the microsatellite data set of the three groups (Table 3). It can be difficult to discriminate between the effects of genetic drift and new mutations that might inflate these fixation indices, especially in the absence of sequence data to test for allele length homoplasy. However, the high R_{ST} divergence observed between cryptic species and the significantly smaller permuted (pR_{ST}) than observed R_{ST} values for three loci, and over all loci, suggests that at least a partial stepwise mutation process has contributed more to the observed differentiation than genetic drift alone (Hardy *et al*., 2003). As mutation accumulation depends on time, the observed pattern would not be expected if cryptic species are of recent (i.e. Pleistocene) age.

Complete reproductive isolation was suggested when comparing group-A fishes with those from the other two groups. This included intrinsic reproductive isolation (i.e. not based on geographic barriers), as individuals from groups A and C were collected in extreme sympatry (in the same net) in two tributaries (see below). Nonetheless, even in this scenario, no evidence of putative hybrids was detected. Furthermore, for the tributary sample for which we also collected morphological data (ZA), we found a perfect match between the genetic and morphological assignment of fishes into two groups. Partial reproductive isolation was identified between groups B and C, for which we identified a small number ($N = 6$) of putative hybrids in five tributaries. Thus, although genetically divergent from one another, groups B and C appear to have the potential to mate and produce the occasional first-generation hybrid. Interestingly, reproductive isolation appears to be in agreement with the evolutionary time since the separation of maternal lineages: the species that hybridize are more closely related to each other phylogenetically than either is to the species that showed complete reproductive isolation (Fig. 5). As discussed in Vähä & Primmer (2006), efficient detection of F_1 hybrid individuals requires a large number of loci (12–24) and pairwise F_{ST} between parental populations of 0.21 or 0.12. Although there were very high levels of genetic divergence between the groups and high allelic diversity in the loci used in this study, more loci would be required to effectively investigate hybridization between groups B and C.

Tributary populations generally contained a single genetic group, except for midwater tributaries ZA and CAU, which contained fishes from groups A and C (Fig. 1). This may reflect their natural overlapping distribution, or perhaps presents a case of sympatric or peripatric speciation. Another possible explanation is the unintentional mixing of the two species by ornamental fishermen. The ornamental fishery operates in several tributaries of the RNF (including ZA and CAU), and the catch is transported to the city of Barcelos (near ZA and CAU), from where fishes are later shipped to the capital of Manaus (Chao, 2001). In Barcelos, fishes deemed of low commercial value (e.g. of small size) or that exceeded market demands are released into the Rio Negro's main channel. Thus, it is possible that released group-A fishes collected elsewhere in the RNF might have entered tributaries near Barcelos and established populations in the flooded forest habitat.

Implications for the management of fishery resources

To date, black-winged hatchet fish from the Rio Negro basin have been sold as a single species (*C. marthae*) to the national and international aquarium trade (Chao, 2001). Our findings have important implications for conservation management. In this study, only 14% of all sampled individuals were assigned to group A. These were found in only four of the 21 sampled tributaries, and thus appear to be more rare than the other two putative species. The highly divergent group-A fishes also have lower nuclear genetic diversity than fishes from groups B and C (Appendix S1), suggesting it may be more isolated and perhaps composed of smaller populations. Thus, the potential for overfishing this putative species needs to be addressed by the Rio Negro ornamental fishery association and fishery managers. In addition, it is critically important that the three species reported here should be managed independently. As shown by Daugherty *et al*. (1990), an assumption of monotypy and failure to identify diversity within a species, such as the New Zealand tuatara *Sphenodon*, can result in the extinction of populations and subspecies.

Cryptic biodiversity within a well-connected floodplain system

Findings of cryptic biodiversity in tropical rainforests are not completely unexpected. Amazonia has very high freshwater fish diversity (Reis *et al*., 2004; Fernandes *et al*., 2005), and is largely understudied compared with densely sampled and well-characterized temperate systems (Willig, Kaufman & Stevens, 2003; Beheregaray, 2008). However, what is surprising in our study is the finding of cryptic speciation within a well-connected system of flooded forests that lacks

contemporary barriers that could account for population divergence and speciation. In the RNF, water level can increase up to 15 m annually because of the inundation cycle, creating a vast lateral floodplain forest that provides feeding, reproduction, and refuge for small fishes (Goulding, Carvalho & Ferreira, 1988; Chao, 2001; Marshall *et al*., 2008). The black-winged hatchet fish is expected to behave as other small forest fishes found in the same habitat, such as the cardinal tetra *Paracheirodon axelrodi* (e.g. Cooke, Chao & Beheregaray, 2009). During the inundation cycle these small fishes move upstream from the shallow areas of Rio Negro's tributaries, and then laterally into the flooded forest, returning to the tributaries during the low-water season. Thus, the flooded forest environment provides a route for connectivity of forest fish populations between streams and adjacent wetlands (Winemiller, 1993; Marshall *et al*., 2008), reducing the possibility of within-drainage speciation. Geomorphological evidence indicates that the vast anabranching system that constitutes the RNF is very recent: of Holocene age (Latrubesse & Franzinelli, 2005). However, this basin was affected by numerous well-characterized tectonic events during the Miocene that did not leave obvious geographic features in the contemporary landscape (Lundberg, 1998; Latrubesse & Franzinelli, 2005). The Miocene speciation scenario for black-winged hatchet fish within the Rio Negro agrees not only with geomorphology, but also with outcomes from our ongoing project on the comparative phylogeography and speciation of Rio Negro's forest fishes. These include evidence for cryptic speciation of Miocene age (for one-lined pencilfish, *Nannostomus unifasciatus*; Sistrom *et al*., 2009), and for historical isolation followed by recent population expansions linked to the establishment of the RNF (for *P. axelrodi*; Cooke *et al*., 2009). Additional sequence data being collected for *C. marthae* and other species will be used in a multispecies comparative study (Beheregaray, unpubl. data) to clarify the chronology of diversification events in RNF forest fishes, and to identify geological and environmental processes driving speciation.

Our study suggests that contemporary drainage structures in Amazonia might be a poor predictor of cryptic biodiversity in rainforest-dependent aquatic organisms.

It also shows the importance of employing a fine-scale sampling regime associated with multiple unlinked DNA markers to detect region-wide cryptic biodiversity. To date, the approach employed here is unique for genetic studies of Amazonian forest fishes. Sampling on a larger scale, such as between basins, is unlikely to have uncovered these multiple species (e.g. species B was confined to the headwaters and species A is comparatively rare).

Amazonian biodiversity is likely to be severely under-documented based on morphological inventories alone (Parra-Olea & Wake, 2001; Beheregaray & Caccone, 2007). Cryptic speciation is expected to be abundant in tropical rainforests because they are a species-rich habitat, and because many organisms are involved in specialized interspecific interactions (Willig *et al*., 2003; Hebert *et al*., 2004). Although some have argued that the proportion of cryptic species is likely to be similar across different biogeographic regions (Pfenninger & Schwenk, 2007), we believe a much more intensive and systematic population sampling approach, as carried out here, is needed for tropical rainforest organisms. There is also the possibility of morphological stasis in these regions because of extreme environmental conditions, which may reduce or eliminate morphological change that can accompany speciation (Bickford *et al*., 2007). As discussed in Bickford *et al*. (2007), DNA methods are the best approach to identifying cryptic species, and should be employed by biologists who are discovering and describing new species. However, morphology-based taxonomy is still very important for species identification, for suggesting functional differences among cryptic taxa, and for assigning taxonomically valid names, and is essential for species descriptions (Schlick-Steiner *et al*., 2007). Once cryptic species are discovered, they can usually be identified by external physical characters, especially if morphometric statistics are employed (Seifert, 2002; Saez & Lozano, 2005; Schlick-Steiner *et al*., 2007). Our findings support the expectation of high cryptic diversity in Amazonia, and provide perspectives for future studies aimed at documenting biodiversity in the region.

ACKNOWLEDGEMENTS

This study was funded by the Discovery programme of the Australian Research Council (ARC grant DP0556496 to L.B.B.) and by the ECOSAVE programme of the Yale Institute of Biospheric Studies (YIBS), Yale University. Logistics and local arrangements in Amazonia were supported in part by Project Piaba, based at the Universidade Federal do Amazonas, and partially funded by the Brazilian National Council of Research and Technology (CNPq grant 46.6098/2001-4), and Bio-Amazonia Conservation International. Our collection permit was issued under IBAMA number 12273-1 to N.L.C. We also acknowledge Lilian Gonsalvez for assistance with the collection of morphological data.

REFERENCES

Baum DA, Shaw KL. 1995. Genealogical perspectives on the species problem. In: Hoch C, Stephenson AG, eds.

Experimental and molecular approaches to plant biosystematics (P.). St. Louis: Missouri Botanical Garden, 298–303.

Beheregaray LB. 2008. Twenty years of Phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology* **17:** 3754–3774.

Beheregaray LB, Caccone A. 2007. Cryptic biodiversity in a changing world. *Journal of Biology* **6:** 1–5.

Beheregaray LB, Piggott M, Chao NL, Caccone A. 2006. Development and characterization of microsatellite markers for the Amazonian blackwing hatchetfish, *Carnegiella marthae* (Teleostei, Gasteropelecidae). *Molecular Ecology* **6:** 787–788.

Bermingham E, McCafferty SS, Martin AP. 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian isthmus. In: Kocher TD, Stepien CA, eds. *Molecular systematics of fishes*. San Diego, CA: Academic Press, 113–126.

Bickford D, Lohman DJ, Sohdi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22:** 148–155.

Chao NL. 2001. The fishery, diversity and conservation of ornamental fishes in the Rio Negro basin, Brazil – A review of Project Piaba (1989–99). In: Chao NL, Petry P, Sonneschien PGL, Tlusty M, eds. *Conservation and management of ornamental fish resources of the Rio Negro Basin, Amazonia, Brazil – Project Piaba*. Manaus: Editora da Universidade do Amazonas, 161–204.

Colborn J, Crabtree RE, Shaklee JB, Pfeiler E, Bowen BW. 2001. The evolutionary enigma of the bonefishes (Albula spp.): cryptic species and ancient separations in a globally distributed shorefish. *Evolution* **55:** 807–820.

Cooke GM, Chao NL, Beheregaray LB. 2009. Phylogeography of a flooded forest specialist fish from central Amazonia based on intron DNA: the cardinal tetra Paracheirodon axelrodi. *Freshwater Biology* **54:** 1216–1232.

Corrigan S, Huveneers C, Schwartz TS, Harcourt RG, Beheregaray LB. 2008. Genetic and reproductive evidence for two species of ornate wobbegong shark *Orectolobus* spp. on the Australian east coast. *Journal of Fish Biology* **73:** 1662–1675.

Daugherty CH, Cree A, Hay JM, Thompson MB. 1990. Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature* **347:** 177–179.

De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56:** 879–886.

Dobzhansky T. 1950. Mendelian populations and their evolution. *American Naturalist* **84:** 401–418.

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

Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* **125:** 1.

Fernandes FMC, Albert JS, Daniel-Silva MDZ, Lopes CE, Crampton WGR, Almeida-Toledo LF. 2005. A new Gymnotus (Teleostei: Gymnotiformes: Gymnotidae) from the Pantanal Matogrossense of Brazil and adjacent drainages: continued documentation of a cryptic fauna. *Zootaxa* **933:** 1–14.

Ferraz G, Russell GJ, Stouffer PC, Bierregaard RO, Pimm SL, Lovejoy TE. 2003. Rates of species loss from Amazonian forest fragments. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 14069–14073.

Filizola NP. 1999. *O fluxo de sedimentos em suspensão nos rios da bacia Amazônica Brasileira*. Brasilia: ANEEL.

Gery J. 1984. The fishes of Amazonia. In: Sioli H, ed. *The Amazon: limnology and landscape ecology of a mighty tropical river and its basin*. Dordrecht: Dr W. Junk Publishers, 353–370.

Géry J. 1977. *Characoids of the world*. Neptune City, NJ: TFH Publications.

Goulding M, Carvalho ML, Ferreira EG. 1988. *Rio Negro, rich life in poor water*. The Hague: SPB Academic Publishing.

Günther A. 1864. Catalogue of the fishes in the British Museum. Catalogue of the Physostomi, containing the families Siluridae, Characinidae, Haplochitonidae, Sternoptychidae, Scopelidae, Stomiatidae in the collection of the British Museum. *Catalogue of the Fishes in the British Museum* **5:** 1-411–1-455.

Hardy OJ, Vekemans X. 2002. SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2:** 618–620.

Hardy OJ, Charbonnel N, Freville H, Heuertz M. 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics* **163:** 467–1482.

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

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 14812–14817.

Hubbs CL, Lagler KF. 1958. *Fishes of the Great Lakes region*. Michigan: University of Michigan Press.

Knowlton N. 1993. Sibling species in the sea. *Annual Review of Ecology and Systematics* **24:** 189–216.

Latrubesse EM, Franzinelli E. 2005. The late Quaternary evolution of the Negro River, Amazon, Brazil: implications for island and floodplain formation in large anabranching tropical systems. *Geomorphology* **70:** 372–397.

Laurance WF, Powell G, Hansen L. 2002. A precarious future for Amazonia. *Trends in Ecology and Evolution* **17:** 251–252.

Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabarba MCSL, Wesselingh F. 1998. The stage for neotropical fish diversification: a history of tropical South American rivers. In: Malabarba LR, Reis RE, Vari RP, Lucena ZM, Lucena CAS, eds. *Phylogeny & Classification of Neotropical Fishes*. Porto Alegre: Edipucrs.

Marshall BG, Forsberg BR, Thome-Souza MJF. 2008. Autotrophic energy sources for *Paracheirodon axelrodi* (Osteichthyes, Characidae) in the middle Negro River, Central Amazon, Brazil. *Hydrobiologica* **596:** 95–103.

Martin AP, Bermingham E. 1998. Systematics and evolution of lower Central American cichlids inferred from

analysis of cytochrome b gene sequences. *Molecular Phylogenetics and Evolution* **9:** 192–203.

Mayer F, von Helversen O. 2001. Cryptic diversity in European bats. *Proceedings of the Royal Society of London Series B* **268:** 1825–1832.

Mayr E. 1942. *Systematics and the origin of species*. New York: Columbia University Press.

Myers GS. 1927. Descriptions of new South American freshwater fishes collected by Dr. Carl Ternetz. *Bulletin of the Museum of Comparative Zoology* **68:** 107–135.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403:** 853–858.

Nelson G, Platnick NI. 1981. *Systematics and biogeography*. New York: Columbia University Press.

Parra-Olea G, Wake DB. 2001. Extreme morphological and ecological homoplasy in tropical salamanders. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 7888–7891.

Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:** 288–295.

Pfenninger M, Schwenk K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology* **7:** 121–126.

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

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 9197–9201.

Raymond M, Rousset F. 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86:** 248–249.

Reis RE, Kullander SO, Ferraris CJ. 2004. *Checklist of freshwater fishes of South and Central America*. Porto Alegre: Edipucrs.

Saez AG, Lozano E. 2005. Body doubles. *Nature* **433:** 111.

Schlick-Steiner BC, Seifert B, Stauffer C, Christian E, Crozier RH, Steiner FM. 2007. Without morphology cryptic species stay in taxonomic crypsis following discovery. *Trends in Ecology and Evolution* **22:** 391–392.

Schneider S, Roessli D, Excoffier L. 2000. *Arlequin: a software for population genetic data analysis*. Geneva: Genetics and Biometry Laboratory, University of Geneva.

Schönrogge K, Barr B, Wardlaw JC, Napper E, Gardner MG, Breen J, Elmes GW, Thomas JA. 2002. When rare species become endangered: cryptic speciation in myrmecophilous hoverflies. *Biological Journal of the Linnean Society* **75:** 291–300.

Seifert B. 2002. How to distinguish most similar insect species – improving the stereomicroscopic and mathematical evaluation of external characters by example of ants. *Journal of Applied Entomology* **126:** 1–9.

Sistrom MJ, Chao NL, Beheregaray LB. 2009. Population history of the Amazonian one-lined pencilfish based on intron DNA data. *Journal of Zoology* **278:** 287–298.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139:** 457–462.

Smith N. 1985. The impact of cultural and ecological change on Amazonian fisheries. *Biological Conservation* **32:** 355–373.

Sunnucks P, Hales DF. 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus Sitobion (Hemiptera: Aphidae). *Molecular Biology and Evolution* **13:** 510–524.

Swofford DL. 1998. PAUP – A computer-program for phylogenetic inference using maximum parsimony. *Journal of General Physiology* **102:** A9.

Taylor DJ, Finston TL, Hebert PDN. 1998. Biogeography of a widespread freshwater crustacean: pseudocongruence and cryptic endemism in the North American *Daphnia laevis* Complex. *Evolution* **52:** 1648–1670.

Taylor WR, Van Dyke GC. 1985. Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage. *Cybium* **9:** 107–119.

Teske PR, Beheregaray LB. 2009. Evolution of seahorses' upright posture was linked to Oligocene seagrass expansion. *Biology Letters* **5:** 521–523.

Vähä JP, Primmer CR. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15:** 63–72.

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

Weitzman SH, Palmer L. 2003. Gasteropelecidae. In: Reis RE, Kullander SO, Ferraris CJ, eds. *Checklist of freshwater fishes of South and Central America*. Porto Alegre: Edipucrs, 101–103.

Willig MR, Kaufman DM, Stevens RD. 2003. Latitudinal gradients of biodiversity: pattern process, scale, and synthesis. *Annual Review of Ecology, Evolution and Systematics* **34:** 273–309.

Winemiller KO. 1993. Seasonality of reproduction by livebearing fishes in tropical rainforest streams. *Oecologia* **95:** 266–276.

Wright S. 1940. The statistical consequences of Mendelian heredity in relation to speciation. In: Huxley J, ed. *The new systematics*. London: Oxford University Press, 161–183.

SUPPORTING INFORMATION

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

Appendix S1. Allele frequency distributions for polymorphic loci in genetic groups A (blue bars), B (green bars), and C (red bars) in *Carnegiella marthae*.

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