doi:10.1111/evo.12410

RIVERSCAPE GENETICS IDENTIFIES REPLICATED ECOLOGICAL DIVERGENCE ACROSS AN AMAZONIAN ECOTONE

Georgina M. Cooke,[1,2] Erin L. Landguth,[3] and Luciano B. Beheregaray[1,4,5]

[1]*Molecular Ecology Lab, Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia*

[2]*The Australian Museum, The Australian Museum Research Institute, Sydney, New South Wales 2010, Australia*

[3]*Division of Biological Sciences, University of Montana, Missoula, Montana 59812*

[4]*Molecular Ecology Lab, School of Biological Sciences, Flinders University, Adelaide, South Australia 5001, Australia*

[5]*E-mail: luciano.beheregaray@flinders.edu.au*

Received May 17, 2013
Accepted March 4, 2014

Ecological speciation involves the evolution of reproductive isolation and niche divergence in the absence of a physical barrier to gene flow. The process is one of the most controversial topics of the speciation debate, particularly in tropical regions. Here, we investigate ecologically based divergence across an Amazonian ecotone in the electric fish, *Steatogenys elegans*. We combine phylogenetics, genome scans, and population genetics with a recently developed individual-based evolutionary landscape genetics approach that incorporates selection. This framework is used to assess the relative contributions of geography and divergent natural selection between environments as biodiversity drivers. We report on two closely related and sympatric lineages that exemplify how divergent selection across a major Amazonian aquatic ecotone (i.e., between rivers with markedly different hydrochemical properties) may result in replicated ecologically mediated speciation. The results link selection across an ecological gradient with reproductive isolation and we propose that assortative mating based on water color may be driving the divergence. Divergence resulting from ecologically driven selection highlights the importance of considering environmental heterogeneity in studies of speciation in tropical regions. Furthermore, we show that framing ecological speciation in a spatially explicit evolutionary landscape genetics framework provides an important first step in exploring a wide range of the potential effects of spatial dependence in natural selection.

KEY WORDS: **Adaptive divergence, Amazon Basin, CDPOP, ecological genomics, evolutionary landscape genetics, isolation by environment.**

Studying the evolution of reproductive isolation and niche divergence in the absence of a physical barrier to gene flow is an important endeavor in speciation research. Ecological speciation results from divergent natural selection acting on adaptive traits responsible for post- and prezygotic reproductive isolation along a continuum from adaptive variation within panmictic populations to complete reproductive isolation between species (Coyne 1992; Schluter 2000, 2009; Rundle and Nosil 2005; Hendry 2009; Hendry et al. 2009; Nosil et al. 2009a). Yet, despite the growing acceptance that divergent selection has generated much of life's diversity (Schluter 2000, 2001; Coyne and Orr 2004; Nielsen 2005; Rundle and Nosil 2005; Nosil et al. 2009b; Schluter 2009), our understanding at a molecular level of how environmental heterogeneity influences complex evolutionary processes, such as adaptation and gene flow, is still limited (but see "geographic mosaic hypothesis"; Thompson 2005). This deficiency can be partly explained by the historical reliance of population genetic surveys on information from putatively neutral genetic markers (Holderegger and Wagner 2008; Storfer et al. 2010). Nowadays, there is a growing capacity to gain information from functionally

EVOLUTION INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

important genes or genomic regions targeted by selection (e.g., via genome scans: Beaumont and Balding 2004) and as a result, divergent selection can now be examined within a spatial framework that includes environmental heterogeneity in the underlying speciation processes (Feder et al. 2012). Moreover, the emerging field of landscape genetics, which explores the degree to which complex landscape facilitates the movement of organisms (Manel et al. 2003; Storfer et al. 2007, 2010), is also beginning to include selection-driven loci in spatial selection environments providing a powerful approach for studying ecological speciation systems (e.g., Landguth et al. 2012).

However, the space–ecology relationship is complex and there are several known spatial and selective pathways that may result in divergent natural selection in the face of gene flow. These range from complete sympatric speciation through to secondary contact after allopatric isolation (Endler 1977; Coyne and Orr 2004). Indeed, numerous environmental factors as well as geographical distance and geomorphological history may be correlated with a signature of adaptive divergence. These effects must be controlled for to detect ecologically mediated adaptive divergence (Lee and Mitchell-Olds 2011). If selection is driving divergence between populations, gene flow should be reduced between selective environments. In turn, if geography is driving divergence, geographic distance and physical barriers should reduce gene flow. Yet, both these processes may be occurring simultaneously. As such, many studies of adaptive divergence and ecological speciation must consider the contribution of both space and selective environment (e.g., Crispo et al. 2006; Nosil 2008; Thorpe et al. 2008; Labonne and Hendry 2010; Lee and Mitchell-Olds 2011; Cooke et al. 2012b).

In this study, we explore a system in which two codistributed and cryptic species of Amazonian electric fish seem to suggest genetic divergence in parallel across the same environmental gradient. This scenario offers an opportunity to use "genome scans" of selection with empirical data and spatially explicit individual-based simulations to build evidence for the process of divergent natural selection in the wild across space and selective environments. Indeed, the comparison of landscape patterns of selected loci ("outlier loci" that exhibit exceptionally high genetic differentiation) to those of neutral loci can assist in the dissection of the space–ecology relationship. Furthermore, an even greater understanding of how the environment influences ecological speciation is expected from studies that use "natural replicates of the ecological speciation in progress" (sensu Rosenblum and Harmon 2011). Studies of replicated speciation events have included comparisons of closely related lineages in geographically separate but similar environments (e.g., whitefish [Lu and Bernatchez 1999]; sticklebacks [Berner et al. 2009]) and of distantly related lineages in a shared environment (e.g., lizards [Rosenblum and Harmon 2011]). Both scenarios are expected to assist with the discovery of general rules about divergent natural selection that may result in ecological speciation.

The South American electric fish, also known as knife fish or weakly electric fish (Teleostei: Gymnotiformes), are nocturnal organisms with very poor vision that inhabit deep river channels and floating floodplain meadows (Albert 2001). They have evolved elaborate electrosensory systems capable of emitting and decoding electrical discharges (Bass 1986). The system is used for both electrolocation and communication, and certain aspects of the rhythmic electric organ discharge (EOD) are not only species specific, but may also be sexually dimorphic (Kramer et al. 1981; Hagedorn and Carr 1985; Crampton 1998; Stoddard 1999). It is thought that predation avoidance initiated the evolution of EOD complexity in Gymnotiformes. However, sexual selection exploited this complexity, resulting in further signal elaboration co-opted for mate attraction (Stoddard 1999).

Our study organism is the barred electric fish, *Steatogenys elegans* (Hypopomidae), a taxon distributed throughout the lowland freshwaters of the Amazon and Orinoco Basins and the Guyana Shield (Albert 2001). Like other electric fish, EOD sexual dimorphism has been noted in many species of this family, and is likely involved in prezygotic isolation (Stoddard 1999). We examine the relationship between genetic variation and selection within *S. elegans* in three major river systems of the Amazon Basin: the Amazon River and its two largest tributaries, the Negro and Madeira Rivers (Fig. 1). The study area encompasses a putatively strong selection gradient represented by an aquatic ecotone with dramatic differences in hydrochemical properties, sediment composition, and optical characteristics. These differences are best illustrated in terms of "water color." Black water (i.e., Negro River), although translucent, is stained dark by tannins and humic acids leached from vegetation, has an acidic pH (∼5 or lower), and is nutrient-poor. White water (i.e., Amazon and Madeira Rivers), by contrast, has an Andean origin, is very turbid with high amounts of dissolved solids and nutrients, and has a neutral pH (∼7) (Sioli 1984). Within our study transect, the black waters of the Negro River coalesce into the white waters of the Amazon River generating a steep ecological gradient—this is clearly illustrated in the transition region known as the "meeting of waters" (Fig. 1A). The transect also allows controlling for genetic structure geographically associated with the confluence of a river by comparing systems with the same water type (i.e., the Madeira and Amazon Rivers).

Here, we investigate divergent natural selection in the electric fish *S. elegans* across an Amazonian aquatic ecotone by combining methods in phylogenetics, genome scans, and population genetics with a recently developed individual-based evolutionary landscape genetics approach that incorporates selection. The individual-based, spatially explicit, and environmentally driven landscape genetics approach has yet to be applied to addressing

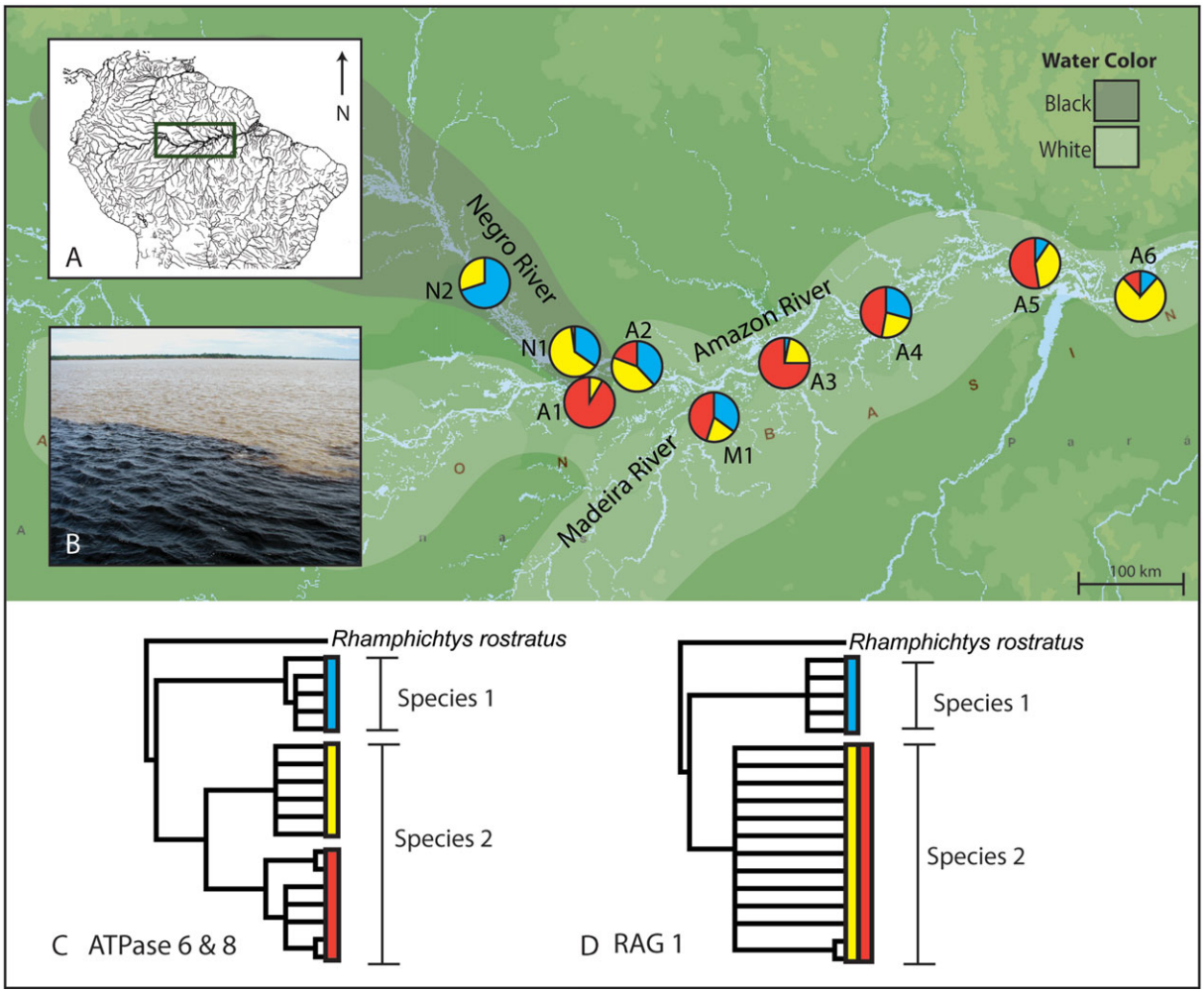

Figure 1. **The nine sampling localities of *S. elegans* in the Amazon Basin together with summarized phylogenetic results. Pie charts show the proportion of individuals in each sampled site assigned to mitochondrial phylogroup A, B, or C (blue, A; yellow, B; red, C). Insets show (A) the sampling area within northern South America; (B) the meeting of the black waters of the Negro River and white waters of the Amazon River (near our site A2); (C) maximum likelihood tree based on mitochondrial ATPase 6 and 8 sequences (bootstrap support values are show in Fig. S1, i); (D) maximum likelihood tree based on nuclear RAG1 sequences (bootstrap support values are shown in Fig. S1, ii).**

questions in divergent natural selection and ecological speciation in the wild, making this study a first of its kind. Our initial aims are (1) to assess the relative contributions of geography and divergent selection between environments as biodiversity drivers in Amazonian electric fish using phylogenetics, genome scans, and population genetics, and (2) to evaluate the sensitivity of simulations of riverscape genetics under a range of spatial selection scenarios by comparing them with the empirical riverscape genetic results. Unexpectedly, we report on two cryptic and divergent species within *S. elegans* that are sympatrically distributed across our vast study region. The two species show similar intraspecific patterns of divergence in relation to the same selective environment, a finding suggestive of progress toward replicated incipient ecological speciation. Our landscape genetics results provide support to this hypothesis, indicating that computational approaches that explicitly evaluate the interactions between gene flow and selection can be extrapolated to empirical studies in complex scenarios. To the best of our knowledge, this represents one the first replicated cases of incipient ecological speciation in Amazonia and has implications to understand how environmental heterogeneity influences the distribution and evolution of biodiversity in this complex and species-rich ecosystem.

Methods

SAMPLING AND GENETIC MARKERS

We sampled *S. elegans* at nine sites ($N = 233$ individuals) from the three major river systems of the Amazon Basin; the Negro (black water), Amazon (white water), and Madeira (white water) Rivers (Fig. 1) in 2005 and 2008 (see Supporting Information

Methods for details about sampling). Geographic coordinates and hydrochemical variables were recorded at each site (Table S1). Muscle tissue was dissected from each fish and ethanol-preserved DNA was extracted using a salting-out protocol (Sunnucks and Hales 1996).

Sequence data were obtained from two mitochondrial DNA (mtDNA) regions, the adenosine triphosphatase subunits 6 and 8 (ATPase 6 and 8), for all samples and two outgroups (*Sternopygus macrurus* and *Rhamphichtys rostratus*). Results from the mtDNA phylogenetic analysis (described below; Figs. S1, S2) were used to select a subset of samples ($N = 18$) representing each monophyletic group to be sequenced for a single-copy nuclear DNA fragment of the recombination-activating gene 1 (RAG1), plus the two outgroup taxa. Primers and amplification conditions used for the polymerase chain reactions (PCRs) for the mtDNA and nuclear genes and sequencing protocol are detailed in Cooke et al. (2012c).

Amplified fragment length polymorphisms (AFLPs) were generated for all samples following a modified protocol (Zenger et al. 2006). Details about the PCR procedure and scoring of AFLP profiles are found in Cooke et al. (2012a). AFLPSCORE version 1.4 (Whitlock et al. 2008) was employed to calculate the mismatch error rate for our data using ~10% of our samples as replicates for each primer combination.

PHYLOGENETIC ANALYSES AND DIVERGENCE DATING

Phylogenetic relationships based on mtDNA ATPase 6 and 8 and nuclear RAG1 sequences were investigated using maximum parsimony, neighbor-joining, maximum likelihood, and Bayesian inference (BI) (see Supporting Information Methods for details). Divergence times between the main lineages recovered based on mtDNA and nuclear datasets (see Results) were then estimated as time to most recent common ancestor (T_{mrca}; details in Supporting Information Methods).

GENOME SCANS OF SELECTION

Phylogenetic analyses provided strong support for two divergent clades that appear further separated in lineages A, B, and C (see Results). We searched for positive selection within each clade using two distinct analytical approaches. First, DFDIST (Beaumont and Nichols 1996) was used to identify outlier loci, which under a scenario of divergent selection should display significantly higher F_{ST} values than the majority of neutral loci in a sample. We chose not to specifically test for loci with unusually low F_{ST} (i.e., expected to be under balancing selection) because divergence-based methods are known to have little power to detect stabilizing selection (Beaumont and Balding 2004). Outlier loci were identified by comparing their specific F_{ST} values with a null F_{ST} distribution estimated from the empirical data using the method of Zhivotovsky (1999). Using the null distribution, a "trimmed" mean F_{ST} was calculated by removing 30% of the highest and lowest F_{ST} values (Beaumont and Balding 2004). The "trimmed" mean represents the "neutral" F_{ST} value, uninfluenced by loci under stabilizing or directional selection. The null distribution was generated using 50,000 realizations, and analyses were performed excluding loci with allele frequencies higher than 0.98 and using a smoothing parameter of 0.04. DFDIST was used to search for directional selection in pairwise comparisons between locations within each clade. In an attempt to exclude type I errors outlier loci detected using DFDIST, loci were classified into three groups (after Nosil et al. 2008): (1) putatively neutral loci, (2) nonrepeat outlier loci, outliers detected in only one pairwise comparison, and (3) repeat outlier loci, outliers detected in multiple comparisons. Additionally, we noted the geographic and ecological setting of these outliers to make inferences about the possible forces driving outlier behavior. Nonrepeat outlier loci were considered potential type I errors, irrespective of their *P*-value. Thus, only repeat outlier loci were considered as loci under selection and removed from the AFLP dataset for subsequent analyses.

Additionally, we used the Bayesian method implemented in BAYESCAN (Foll and Gaggiotti 2008) to estimate the probability that each locus is subject to selection. BAYESCAN defines two alternative models: (1) the effect of selection included and (2) the effect of selection excluded. Here, model choice is based on the posterior model probabilities (Bayes factors) and evidence for selection is based on Jeffreys (1961) scale of evidence. The sample size was set to 10,000 and the thinning interval was set to 20, and all loci were ranked according to their estimated posterior probability. Three independent runs were performed to check the consistency between detected outliers. BAYESCAN was run separately for each lineage.

EMPIRICAL RIVERSCAPE GENETICS

For the empirical dataset, Nei's (1972) standard genetic dissimilarity coefficient (D_{N72}) and population structure (F_{ST} and Θ^B) based on AFLP was assessed within each lineage (A and BC) according to Lynch and Milligan (1994). Population structure was calculated for both total and putatively neutral datasets. Genetic subdivision was investigated using the putatively neutral AFLP data in STRUCTURE version 2.3.1 (Falush et al. 2003). We use the admixture model to determine the number of populations (K) by following the ΔK method of Evanno et al. (2005) and also by comparing the log-likelihood ratios across multiple independent runs by varying the assumed number of K. Each run consisted of a burn-in phase of 100,000 iterations, followed by 1,000,000 iterations.

To test for an association between genetic (θ_{ST}) and geographic distance measured as riverine distance (isolation by distance; Wright 1943), we used Mantel permutation tests (Mantel

1967) in ARLEQUIN 3.5.1.2 (Excoffier et al. 2005) for each mtDNA lineage (i.e., A, B, and C), and in GENALEX 6.1 (Peakall and Smouse 2006) for each cryptic species (i.e., 1 and 2) using the putatively neutral AFLP dataset. Associations between population genetic structure and hydrochemistry (i.e., a test of isolation by environment) were explored with an analysis of molecular variance (AMOVA) using neutral AFLP data (Excoffier et al. 1992). Two AMOVAs were conducted for each cryptic species within each dataset: (1) white water populations (M1, A1–A6) versus black water populations (N1 and N2), and (2) white water Amazonas sites (A1–A6) versus white water Rio Madeira (M1). The latter tests for genetic differentiation associated with drainage structure (in this case, the tributary confluence of the Madeira River) by comparing populations from different rivers that have the same water color (i.e., type I error rates).

SIMULATING RIVERSCAPE GENETICS

We simulated individual genetic exchange over 100 nonoverlapping generations as a function of individual-based movement, mating, dispersal, and selection, with 100 individuals spatially located at each of the nine populations in our study system using the individual-based landscape genetics program CDPOP version 1.2 (Landguth and Cushman 2010; Landguth et al. 2012; see Supporting Information Methods for more details).

We simulated one locus under spatial selection and 19 neutral loci with 0.0005 *k*th-allele model mutation rate, free recombination, no initial linkage disequilibrium, and random distribution of initial starting alleles (i.e., simulating a panmictic initial population with maximum allelic diversity). Following similar spatial selection simulation studies (e.g., Thibert-Plante and Hendry 2010; Landguth and Balkenhol 2012) but expanding to a spatially explicit environmental gradient in a riverscape setting, we altered selection pressures due to "water color" between populations by considering three spatially explicit relative fitness surface scenarios. (1) No spatial selection gradient (uniform): In this scenario, the three genotypes (AA, Aa, and aa) were being selected against, but uniformly across the "water color" riverscape scenario, thus having no spatial dependency and allowing us to test for type I statistical errors. (2) Gentle spatial selection gradient (gentle): Here, we used a "gentle" spatial selection gradient corresponding to the three river color locations. For AA, we used the relative fitness coefficients of 0.4, 0.3, and 0.2 for black, mixed, and white waters, respectively. For aa, we implemented an opposite spatial selection gradient (relative fitness coefficients of 0.2, 0.3, and 0.4 for black, mixed, and white waters, respectively). (3) Steep spatial selection gradient (steep): For this scenario, stronger spatial selection gradients were assigned to each genotype, with the relative fitness coefficients for AA of 1.0, 0.6, and 0.2 for black, mixed, and white waters, respectively. An opposite spatial selection gradient was implemented for aa (0.2, 0.6, and 1.0 for black, mixed, and white waters, respectively). Aa received a uniform selection gradient of 0.2 in all three scenarios.

For each simulated dataset, we ran a population genetics approach to compare to the empirical riverscapes genetic dataset observed population structure and a landscape genetics approach to compare with the genetic subdivision run in STRUCTURE (see empirical riverscapes genetics methods above). For the population genetics, we quantified genetic differentiation at each generation through the estimator D_{est} (Jost 2008) using only neutral loci, only selection-driven loci, and total loci (neutral plus selection-driven). For the landscape genetics approach, we implemented the distance-based approach of partial Mantel tests (Smouse et al. 1986; Legendre and Fortin 2010) to test for correlation between the spatial genetic signature observed at each generation and the environmental variable of "water color," partialling out the effect of riverine distance (for more details see Supporting Information Methods). We ran 10 Monte Carlo replicates on all scenarios to assess variability in results, and plotted the average D_{est} and average partial Mantel *r* obtained with the different data types (selection-driven, selection-driven plus neutral, and neutral). We evaluated the temporal development of genetic differentiation under the different spatial selection gradients ("uniform," "gentle," and "steep"). We also compared the spatial selection gradient simulations with simulations of a model of secondary contact (see Supporting Information Methods).

Results

GENETIC VARIATION

We obtained and aligned 820 base pairs (bp) of the mtDNA ATPase 6 and 8 genes that characterized 117 unique haplotypes. These were composed of 189 variable characters of which 129 were parsimony informative. For RAG1, we aligned 1494 bp of the gene and identified five unique sequences. These were composed of nine variable characters of which seven were parsimony informative. Genbank accession numbers for all sequences are KJ526476–KJ526610. AFLP profiles were resolved for the 233 individuals with 310 polymorphic loci. Mismatch error rates as calculated by AFLPScore were on average 7.5% per primer combination (6.5–11.14%), which is within the acceptable error rate for AFLPs (Bonin et al. 2007).

CRYPTIC SPECIES AND MOLECULAR DATING

Phylogenetic analyses of mtDNA sequences resolved three monophyletic and well-supported lineages herein referred to as "phylogroups A, B, and C." For each phylogenetic method, phylogroup A appeared as ancestral (Figs. 1C, D, S1, S2). These three phylogroups appear largely sympatrically distributed in our study

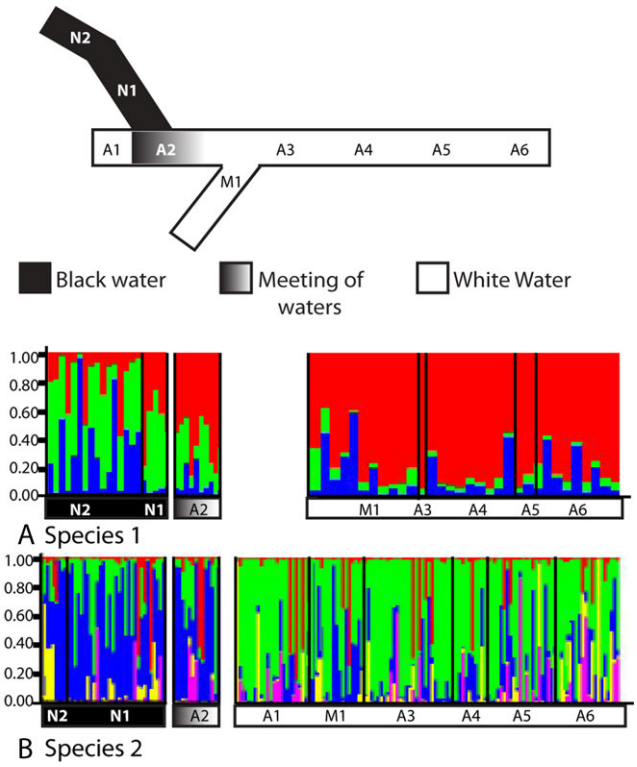

Figure 2. STRUCTURE **results for (A) cryptic species 1 ($n = 62$) and for (B) cryptic species 2 ($n = 171$). Individuals are grouped by sampling location, and each individual is represented by one vertical column. Sample sites and water colors as listed on** STRUCTURE **graphs correspond to the sampling locations as shown in the simplified map above the graphs.**

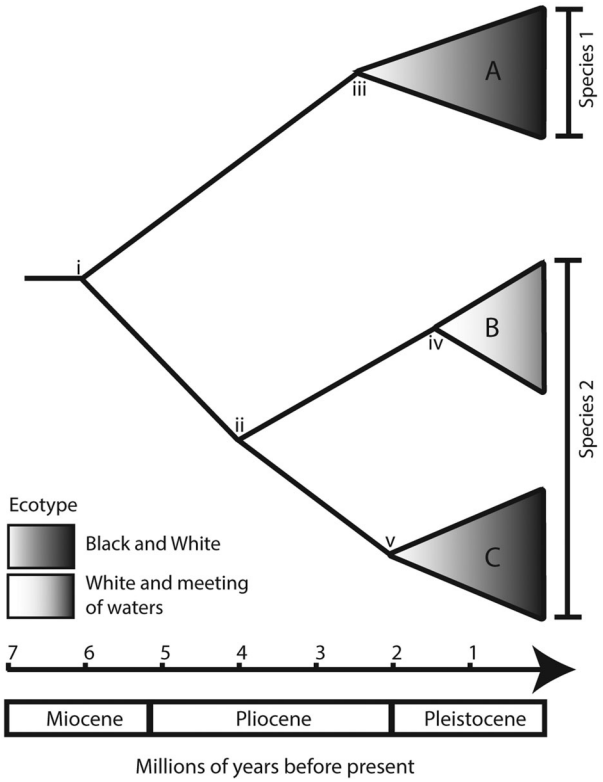

Figure 3. Chronogram of the *Steatogenys elegans* species complex showing divergence time estimates between phylogroups A, B, and C (nodes i–v: Table 2) based on mtDNA ATPase 6, 8, and RAG1 sequence data. Branch lengths represent the estimated ages of each lineage with respect to the geological time scale. The ecotype typical of each phylogroup, based on where the fish were sampled, is shown.

region, except for phylogroup A, which was not sampled west of the confluence of the Negro and Amazon Rivers, and phylogroup C, which was not sampled in the Negro River (Fig. 1).

The results based on nuclear RAG1 sequences and 310 AFLP loci strongly support only two major clades (Figs. 1, S2, S3). In agreement with the mtDNA data, individuals from lineage A appear divergent and ancestral to a second group composed by mtDNA phylogroups B and C. The latter result is probably due to the overall slower rate of nuclear to mitochondrial evolution, and the faster lineage sorting of the mitochondrial genome (Figs. 1C, 2A). Herein, we refer to phylogroup A as "cryptic species 1" (sp. 1) and the combined phylogroups B and C as "cryptic species 2" (sp. 2). Based on our sample, cryptic sp. 1 and sp. 2 are found in sympatry throughout the black waters of the Negro River and the white waters of the Amazon and Madeira Rivers. Consistent with presumed lineage ages, genetic diversity was generally higher in sp. 1 than sp. 2 (Table S2). The AFLP dataset suggest low levels of introgression from sp. 1 to sp. 2 (Fig. S3), a pattern that could not be detected in our phylogenetic analyses, which indicate reciprocal monophyly based on both mtDNA and nuclear sequences (Figs. 1C, S1).

Divergence estimates (T_{mrca}) were consistent for both the nuclear and mtDNA datasets, with speciation between sp. 1 and sp. 2 starting around 6 million years ago (Ma) during the late Miocene (Fig. 3, Table 1). Sp. 1 appears older and based on estimates of divergence (Table 1), diverged in the early Pliocene. More recently, probably in the middle Pliocene (~4 Ma), the mitochondrial lineages B and C split (Table 1). For these lineages, reproductive isolation within the nuclear genome has not yet ensued.

OUTLIER LOCI AND THEIR ASSOCIATION TO WATER COLOR

For *S. elegans* sp. 1, genome scans of population pairwise comparisons conducted using DFDIST identified less than 1% of outlier loci deviating from neutral expectations (Table 2). Of these, none were repeatedly identified within sp. 1. BAYESCAN performed on all sp. 1 populations identified 88 loci with positive α_i values, none of which included the outliers identified using DFDIST. Of these, only one locus had decisive evidence for divergent selection with a $\log_{10}$ (Bayes factor) >2, and just one had

Table 1. Time to most recent common ancestor (T_{mrca}) estimates and 95% lower and upper highest probability densities (HPD) for ATPase 6 and 8 and RAG 1 based on mutation rates after Bermingham et al. (1997) and Quenouille et al. (2004), respectively, and calculated using the relaxed clock method drawn from a lognormal distribution in BEAST version 1.4.6. The nodes correspond to Figure 3.

Node	Gene	T_{mrca}	95% HPD lower	95% HPD upper
I	ATPase 6 and 8	5.9640	3.0763	9.8560
	RAG 1	5.4120	0.8374	12.0296
II	ATPase 6 and 8	3.5995	1.6716	6.0239
	RAG 1	3.6040	0.4454	8.5188
III	ATPase 6 and 8	2.2930	0.6296	4.6542
	RAG 1	2.5100	0.1810	6.7369
IV	ATPase 6 and 8	1.1711	0.2765	2.5577
V	ATPase 6 and 8	1.8133	0.5702	3.4747

Table 2. Number of outlier and repeat-outlier AFLP loci found in each pairwise comparison using Dfdist where repeat outliers refer to those identified in multiple pairwise comparisons.

Species 1 (Lineage A)				Species 2 (Lineages B and C)				
Comparison	99% outliers	Repeat outliers	Total F_{ST}	Comparison	99% outliers	Repeat outliers	Total F_{ST}	Neutral F_{ST}
Black versus white	0	–	–	Black versus white	16	8	–	–
N2N1	0	–	0.0857*	N2N1	2	1	0.0314	0.0312
A1N1	NA	–	NA	A1N1	9	8	0.0717*	0.0437*
N1A2	0	–	0.0373	N1A2	1	0	0.0199	0.0173
A1A2	0	–	NA	A1A2	0	–	0.0497*	0.0466*
A2A3	0	–	0.0500	A2A3	2	1	0.0432*	0.0411*
A2M1	NA	–	NA	A2M1	1	0	0.0335	0.0335
M1A3	1	0	NA	M1A3	1	1	0.0227	0.0221
A3A4	0	–	0.1407*	A3A4	5	4	0.0130	0.0141
A4A5	0	–	0.0488	A4A5	6	3	0.0334	0.0386
A5A6	1	0	0.0416	A5A6	0	–	0.0379*	0.0383
Total	2	0		Total	29	12		

Total F_{ST}, and the presumed "neutral" F_{ST} value (after Lynch and Milligan 1994) are also shown.
*$P \leq 0.05$.
NA indicates pairwise comparisons for which there were insufficient data.

substantial evidence for selection with $\log_{10}$ (Bayes factor) >0.5 based on Jeffreys (1961) scale of evidence (Table S4).

For *S. elegans* sp. 2, genome scans conducted using DFDIST between populations identified 9% of AFLP loci as outliers deviating from neutral expectations. Of these 2.5% were repeatedly identified within sp. 2. Loci repeatedly identified as outliers between pairwise comparisons are unlikely to represent type I errors (Campbell and Bernatchez 2004). We could be confident that these loci were not due to random chance because the proportion of repeat outliers was significantly greater than the proportion of nonrepeat outliers detected over the 310 loci ($P < 0.001$, $\chi^2 = 16.3$, df = 1). Thus, we conservatively identified at least 2.5% of homologous loci within sp. 2 that may be directly subject to selection or tightly linked to selected genes via "hitchhiking" (Jensen et al. 2007). Within sp. 2, there was a significantly greater proportion of outlier loci detected in genome scans between black and white water populations than within white water populations ($P < 0.001$, $\chi^2 = 15.9$, df = 1), including those between the Amazon and Madeira Rivers (Table 2). BAYESCAN performed on all sp. 2 populations identified 102 loci with positive α_i values. Of the outliers identified using DFDIST, 83% of these were also identified using BAYESCAN. Applying Jeffreys (1961) scale of evidence, only four loci had substantial evidence of selection with $\log_{10}$ (Bayes factor) >0.5, whereas the remaining loci with positive α_i values had low posterior probabilities. Of the outliers repeatedly identified using both methods ($n = 12$), eight were identified in pairwise comparisons between black and white water sites (Table S4).

Table 3. Population differentiation within and between black and white water regions calculated from AFLP data using an analysis of molecular variance (AMOVA).

	Black versus white water				White versus white water			
	Source of variation	Variation (%)	FI	*P*	Source of variation	Variation (%)	FI	*P*
Species 1	Among regions	6	Φ_{RT}: 0.060	0.001*	Among regions	2	Φ_{RT}: 0.017	0.298
	Among populations	1	Φ_{PR}: 0.009	0.272	Among populations	4	Φ_{PR}: 0.042	0.174
	Among individuals	93	Φ_{PT}: 0.069	0.002*	Among individuals	94	Φ_{PT}: 0.058	0.024*
Species 2	Among regions	2	Φ_{RT}: 0.020	0.050*	Among regions	0	Φ_{RT}: –0.003	0.494
	Among populations	1	Φ_{PR}: 0.010	0.090	Among populations	1	Φ_{PR}: 0.011	0.145
	Among individuals	97	Φ_{PT}: 0.030	0.020*	Among individuals	99	Φ_{PT}: 0.008	0.192

In the black versus white water AMOVA, regions include: (1) black water populations (N1 and N2), and (2) white water populations (M1 and A2–A6). In white versus white water AMOVA, regions include: (1) the Madeira River (M1), and (2) white water Amazon River sites A2–A6.
FI = fixation index.
***Significant results.**

EMPIRICAL RIVERSCAPE GENETICS AND ECOLOGICAL SPECIATION

Analyses of population structure within *S. elegans* sp. 1 suggest a strong correlation between "water color" and genotype (Fig. 2A, Table 3). Overall, there was little mtDNA differentiation between white water sites, indicative of high connectivity within this selective environment (e.g., Φ_{ST} A2 vs. A3 = 0.12, $P > 0.05$; Table S3), whereas there was strong genetic differentiation between black and white water populations (e.g., Φ_{ST} N1 vs. A2 = 0.40, $P \leq 0.05$; Table S3). Further, mtDNA Mantel tests provided no statistical support for associations between genetic (Φ_{ST}) and riverine distance ($P = 0.595$). Results based on AFLP data also support the population boundary between the ecologically distinct white and black water sites with no correlation between genetic and geographic distance ($P = 0.454$). The STRUCTURE analysis (Fig. 2A) shows a distinct cline where the black and white waters meet at site A2 (Fig. 2A). Here, both mean L(K) and ΔK inferred three populations that correlate with white, black, and meeting of water habitats (Fig. 2A). Likewise, the AMOVA assessing population differentiation within and between water color habitats also supported the hypothesis that the ecotone between black and white water is a significant barrier to gene flow (Table 3). Finally, by applying the overall AFLP divergence rate of $D_{N72} = 0.0370$ (SD = 0.0406) per 10,000 years (after Kropf et al. 2009) to our data, it appears that AFLP divergence between black and white water populations in sp. 1 is recent (~8378 generations).

Analyses of population structure within *S. elegans* sp. 2 also suggest a marked population boundary between the black and white water habitats with mixing and gene flow at the meeting of waters (A2) (Fig. 2B, Table 2). For the STRUCTURE analysis, ΔK inferred two populations and the mean L(K) plateaued at $K = 5$ at which point the black and white water ecotypes become visible in the STRUCTURE output (shown in Fig. 2B), a finding typical of systems with hierarchical structure. Furthermore, the mtDNA phylogroup C was not sampled in black waters, except at the meeting of waters (A2); whereas phylogroup B was sampled at every site (Fig. 1). In both mtDNA phylogroups, there was no correlation between geographic and genetic distance (B, $P = 0.488$; C, $P = 0.093$), whereas there was a weak yet significant correlation in the AFLP data ($R_{xy} = 0.0243$, $R^2 = 0.0107$, $P = 0.006$). Similarly to sp. 1, the AMOVAs also supported the significant barrier to gene flow represented by the black and white water ecotone, the absence of a population barrier between the white water Amazon and Madeira Rivers (Table 3), and divergence timing estimates between the ecotypes were also recent (~3946 generations). Interestingly, the removal of outlier loci from the dataset resulted in a substantial reduction of population differentiation between black and white water populations (e.g., cryptic sp. 2, N1 vs. A1; total $F_{ST} = 0.0717$, $P \leq 0.05$, neutral $F_{ST} = 0.0437$, $P \leq 0.05$, Table 2). This reduction was not observed in any other pairwise comparison after the removal of outlier loci, suggesting that the contribution of those loci under selection to the genetic structure observed across the ecotone is relatively high.

To summarize, intraspecific divergence within both sp. 1 and sp. 2 appears to be recent and a barrier to gene flow exists between black and white water whereas no barrier to gene flow was identified at the confluence of the white waters of the Madeira River into the white waters of the Amazon River. This suggests that geographically driven population structure generated by the confluence of major tributaries is unlikely in our study system.

SIMULATED RIVERSCAPE GENETICS AND ECOLOGICAL SPECIATION

Simulations with three scenarios of relative selection pressures due to "water color" between populations were first conducted to assess the population structure in the simulated dataset and the relative contribution of selection driven versus neutral genetic differentiation. The difference in neutral and selection-driven

genetic differentiation was clearly influenced by the spatial selection gradient (Fig. S4 and Supporting Information Results for more details). In addition, when a spatial selection gradient exists, we show using partial Mantel tests that the environmental signature of "water color" can be discerned and increased in magnitude from the "gentle" to "steep" selection-driven scenarios ($r = 0.30$ (0.237,0.368) and $r = 0.75$ (0.730,0.776), respectively, at generation 100; Fig. S4iv, vi; see Supporting Information Results). Simulations of a scenario of secondary contact revealed strong population structure between the Negro and Amazon and the Amazon and Madeira Rivers when assuming low migration (see Table S4).

Discussion

SEEING DOUBLE: TWO CRYPTIC AND CODISTRIBUTED SPECIES OF AMAZONIAN ELECTRIC FISH

Our molecular analyses provided evidence for two cryptic species within *S. elegans*. These include reciprocal monophyly of the two lineages based on mtDNA and the conserved RAG1 sequences, and strong genetic structure based on 310 AFLP loci. Clade membership of all 233 individuals matched their assignments to the two groups identified with AFLP data. Reproductive isolation was apparent between lineages because they were found in extreme sympatry (i.e., sampled during the same round of dragging) in all three river systems surveyed (Fig. 1); albeit some level of introgression was evident from sp. 1 into sp. 2 (Fig. S3). Molecular dating indicates deep evolutionary separation between these codistributed lineages. Although caution should be taken when interpreting our divergence estimates given the absence of *Steatogenys* molecular clocks, similar estimates were independently obtained for the nuclear and mtDNA datasets (~5.4 and 6 Ma, respectively). This strengthens the notion of a long history of isolation between lineages. Forthcoming phenotypic studies of sp. 1 and sp. 2 are expected to inform on diagnostic morphological characters for species description and on identification of traits under selection within each lineage.

Although the Amazon Basin sustains the world's richest freshwater fish fauna (Reis et al. 2003), the growing number of cryptic species of Amazonian fish detected with molecular techniques (Littmann et al. 2001; Nakayama et al. 2001; Hubert et al. 2007; dos Santos Silva et al. 2008; Sistrom et al. 2009; Nagamachi et al. 2010; Piggot et al. 2011; Cooke et al. 2012d) suggests that species richness in this group is vastly underestimated. A recent comprehensive assessment of cryptic diversity in Amazonian frogs also pointed to a similar conclusion (Funk et al. 2012). The evolution of aquatic biodiversity in Amazonia appears to be intrinsically linked to complex and relatively old geomorphological events that have impacted its riverscape (e.g., uplifts, erosions, and changes in sediment supplies from the Andes) and to major climatic and sea-level changes during the Miocene (see Fig. S4; Hoorn et al. 2010). Accordingly, there are several examples of vicariant biogeographic events driving population divergence and speciation in Amazonian fish and amphibians (e.g., Lynch and Duellman 1997; Lovejoy et al. 1998; Sivasundar et al. 2001; Hubert and Renno 2006; Beheregaray and Caccone 2007; Hubert et al. 2007; Cooke et al. 2009; Sistrom et al. 2009; Piggott et al. 2011). Yet, there has been little recognition for the role of ecological speciation in the generation of Amazonian and tropical diversity alike, with spatially defined models of speciation dominating the literature (Moritz et al. 2000; Hoorn et al. 2010; Turchetto-Zolet et al. 2013). Our study is not aimed at assessing biogeographic scenarios underpinning the split between the two cryptic species of *S. elegans*. We have instead explored the progress toward ecologically based divergent natural selection within each cryptic species (discussed below), and show how environmental heterogeneity influences biodiversity in the complex and species-rich Amazon Basin.

ECOLOGICAL SPECIATION AND EMPIRICAL AND SIMULATED RIVERSCAPE SIGNAL

During ecological speciation, divergent selection will act on populations utilizing different environments. This may result directly or indirectly in speciation (Schluter 2001). Indeed, reproductive isolation usually arises from resource acquisition and competition, mate attraction, and predator avoidance (Schluter 2001; Rundle and Nosil 2005). Here, in this replicated *S. elegans* system, we find evidence for recent divergence linked to a major hydrochemical gradient within each cryptic species using F_{ST}-based genome scans and population genetic analyses that may eventuate in ecological speciation. We further corroborate these findings by conducting individual-based, evolutionary landscape genetics simulations. These show that neutral data can give a low population differentiation signal (similar to the empirical neutral data findings) and selection-driven loci can respond with high population differentiation to the water color ecotone (similar to the empirical outlier loci findings). Furthermore, our empirical and simulated landscape genetics analysis explicitly links selection-driven population genetic structure to the water color ecotone.

The two sympatric cryptic species of *S. elegans* show a relatively old history of divergence (~6 Ma) that is likely a combination of geomorphological history and natural selection. However, intraspecific population level interactions of cryptic sp. 1 and sp. 2 are most informative in identifying divergent selection involved in the progress toward ecological speciation. This is because divergent selection occurring between dissimilar ecotypes that do not yet exhibit complete reproductive isolation reveals insights into processes of ecological speciation that may not be apparent long after speciation is complete (Beheregaray and Sunnucks

2001; Hendry 2009; Via 2009). During early stages of ecological speciation, genomic divergence is likely to be heterogeneous. Genetic differentiation is generally thought to accumulate in some regions (genomic islands) that affect ecologically important traits before others, whereas gene flow continues throughout the rest of the genome (Schluter 2000; Nosil et al. 2009a; Via 2009). With time, however, divergent selection will promote reproductive isolation, further facilitating genome-wide neutral divergence via genetic drift or selection for different traits (Schluter 2000; Rundle and Nosil 2005; Nosil et al. 2009a). Thus, by examining recently isolated or diverging ecologically dissimilar populations, genetic changes that may contribute to speciation can be identified before these become confounded by changes taking place once speciation is complete (Schluter 2000; Via 2009).

Our genome scans within cryptic sp. 2 identified 2.5% of loci repeatedly deviating from neutral expectations. Although uncertainty still remains regarding the role selection may play over these loci, repeat outliers are unlikely to be type I errors (Cooper 2000; Campbell and Bernatchez 2004). Instead, it is probable that these loci are directly subject to selection or tightly linked to selected genes via "hitchhiking" (Jensen et al. 2007). Here, a major barrier to gene flow was identified between black and white water sites. On the other hand, no barrier was identified at the confluence of the Madeira and Amazon Rivers (Tables 2, 3), with the low population structure between white water Amazon sites partially explained by isolation by riverine distance. In cryptic sp. 1, genome scans identified less than 1% of outlier loci and no repeat outliers (Table 3). Likely reasons include the small sample size in many pairwise comparisons (less than 10 individuals per population; Beaumont and Balding 2004), different selection pressures compared to sp. 1, and/or similar selection pressures with a different underlying genetic architecture. Nevertheless, as observed with sp. 2, a barrier was also identified between black and white water sites, whereas no barriers were detected within the same selective environment or geographically associated with the confluence of a major tributary (Tables 2, 3). The above provides evidence that divergent selection is acting within each cryptic species between the black and white water ecotypes.

A key factor in identifying the presence of adaptive divergence is the association of outlier loci to contrasting environments (Nosil et al. 2009a). In sp. 2, there was a significantly greater proportion of outlier loci detected in genome scans between water colors than within (Table 2). Also, removing outlier loci substantially reduced population differentiation between black and white water populations (Table 2). Such reduction was not observed in any other pairwise comparison. Thus, the majority of loci identified that exhibit higher levels of genetic divergence than expected under neutrality were found in comparisons between sites characterized by different hydrochemical properties. This finding was corroborated by the spatially explicit riverscape simulations that showed that neutral versus selection-driven loci tied to an environmental variable can be differentiated using population genetics and correlated spatially via landscape genetics (Fig. S4). Based on the association of genotype and water color within sp. 1 and sp. 2, on the identification and spatial association of "outlier loci" to an ecological gradient, and on our landscape genetics results, we have some evidence for divergent selection that may eventuate in replicated ecological speciation within the *S. elegans* species complex.

Nonetheless, it is well recognized that distinguishing between secondary contact zones and ongoing adaptive divergence of parapatrically isolated forms is extremely difficult (Endler 1977). Indeed, spatial isolation and secondary contact has been implicated in the adaptive radiation of cichlid fish in the Great African Lakes (Schwarzer et al. 2012), as well as speciation in terrestrial Amazonian vertebrates, particularly in birds (Haffer 1969, 1997; Sedano and Burns 2010). As such, this alternative hypothesis warrants exploration here. Generally, it is accepted that the west to east transcontinental flow of the Amazon River and its major tributaries (including Negro and Madeira Rivers) had formed by the late Miocene (Hoorn et al. 1995; Lundberg et al. 1998) with the final establishment of the modern Amazon River drainage system being ~2.5 Ma following the breach of the Madre de Dios formation (Campbell et al. 2006; Fig. S5). Based on our molecular dating results for mtDNA and nuDNA sequence data (Table 1, Fig. 3) in sp. 2, phylogroups B (predominantly white water) and C (predominantly black water) diverged ~3.6 Ma. Although there is not sufficient data to obtain a similar date estimate for the white and black water ecotypes in sp. 1, this result is interesting as it coincides with formation of the Amazon River and its largest tributaries. Prior to this time, if fish inhabited the major tributaries such as the Madeira and Negro they would have been isolated from the extensive freshwater rivers and lakes system in the western Amazon Basin. Following the formation of the Amazon River however, these tributary populations would come in contact with an Amazon River population. Under a scenario of secondary contact following the formation of the Amazon River, we would expect to see equal population subdivision associated with the presence of the Madeira and Negro Rivers, irrespective of water color. The latter was the pattern detected in our landscape genetic simulations of a secondary contact scenario (Table S5). However, these patterns were not observed in the empirical data in either sp. 1 or sp. 2. Rather, our results show that population subdivision is associated with water color more than the geomorphological history or riverine distance. Thus, adaptive divergence or progress toward ecological speciation may be the most parsimonious explanation for our findings.

THE GENERALITY OF THE WATER COLOR ECOTONE AND THE FATE OF INCIPIENT SPECIES

Information about how distantly related species respond to a shared environment are also particularly important in identifying factors that promote or inhibit ecological speciation (Rosenblum and Harmon 2011). Our hypothesis of ecological speciation in two sister species is corroborated by recent studies of two unrelated taxa sampled from the very same sites as *S. elegans*; the Amazonian puffer *Colomesus asellus* (Cooke et al. 2012b), and the characin *Triportheus albus* (Cooke et al. 2012a). These studies combined genome scans and population genetics to disclose heightened divergent selection at the interface of water types, providing strong independent evolutionary replicates that strengthen the generalities of our findings.

Yet, there is no certainty that adaptively diverging lineages will result in reproductively isolated species (Futuyma 1987; Coyne and Orr 2004; Hendry 2009). Indeed, the link between adaptive divergence and speciation within closely related species is often unclear, simply because the process of adaptive divergence itself drives lineages apart (Reznick and Ricklefs 2009). Our data consist of samples and populations along the divergence spectrum providing us with the opportunity to identify patterns of divergence hitchhiking around loci potentially involved with ecological speciation. Outliers detected within each cryptic species might be the genetic signature of divergence hitchhiking associated with ecologically important traits (Via 2009). Importantly however, our AFLP scans are based on anonymous loci, limiting the investigation about putative ecological selective traits (Stinchcombe and Hoekstra 2008). This deficiency is expected to be overcome by functional studies that combine quantitative genomics, transcriptomics, and candidate gene analysis to identify genomic signatures associated with phenotypic traits under selection.

During ecological speciation, genes under divergent selection cause reproductive isolation pleiotropically via divergence hitchhiking (Rundle and Nosil 2005; Via 2009). Under divergence hitchhiking, combinations of genes that cause assortative mating can accumulate and be protected from recombination, because traits that drive resource use also affect mate choice (Schluter 2001; Via 2009). Thus, ecological speciation can be simply the direct consequence of behavioral isolation whereby individuals mate in their preferred habitat (Johnson et al. 1996; Rundle and Nosil 2005). In this way, sexual isolation can evolve as a consequence of the ecologically driven adaptive divergence of mating cues such as communication systems (Boughman 2002).

In weakly electric fish, the precise synchronization of external fertilization must be achieved via EOD communication, in which courtship signaling involves conspicuous and diagnostic EODs (Silva et al. 2008). During the breeding season, many gymnotiform species produce sexually dimorphic signals enabling greater distinction between conspecifics, heterospecifics, and gender (e.g., Stoddard 1999). However, electrical current requires the movement of ions. Thus, pH, dissolved minerals, dissolved oxygen, and temperature should affect the transmission of EODs between individuals within chemically different white and black water habitats. Because EODs carry information that is of both a communicative and social value, it is likely that weakly electric fish are also sensitive to changes in water conductivity. In fact, such changes have been shown to trigger breeding in tropical gymnotiformes (Kirschbaum 1995; Silva et al. 2008). We therefore speculate that conductivity or "water color" may be an ecologically dependent mechanism of behavioral isolation, driving divergence within the *S. elegans* cryptic species complex across this ecotone. This is consistent with the proposal that EODs in African electric fish are drivers of sympatric speciation (Feulner et al. 2006), which is the most extreme form of ecological speciation.

We have described a case of two closely related lineages that exemplify how divergent selection across an aquatic ecotone in Amazonia may eventuate in replicated ecologically mediated speciation. Our findings highlight the importance of considering environmental heterogeneity in studies of speciation in Amazonia and other species-rich tropical regions.

ACKNOWLEDGMENTS

We thank C. Moritz and A. Hendry for their helpful comments on an earlier version of this manuscript, N. Chao for assistance with fieldwork and logistics, and M. Ashcroft for help with GIS. This study was funded by the Discovery Program of the Australian Research Council (ARC DP0556496 to LBB) and by Macquarie University through a postgraduate travel grant and student award to GMC. Local arrangements were supported in part through the Brazilian National Council of Research and Technology (CNPq-SEAP No. 408782/2006–4 to N. Chao). Collection permit is under IBAMA #1920550, and ethical approval under Macquarie University #2007/033. The authors have no conflict of interest to declare.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.7g2h4.

LITERATURE CITED

Albert, J. S. 2001. Species diversity and phylogenetic systematics of American knifefishes (Gymnotiformes, Teleostei). Misc. Publ. Mus. Zool. Univ. Mich. 190:1–140.

Bass, A. H. 1986. Electric organs revisited. Pp. 13–70 *in* T. H. Bullock and W. Heiligenberg, eds. Electroreception. Wiley-Interscience, New York.

Beaumont, M. A., and D. J. Balding. 2004. Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol. 13:969–980.

Beaumont, M. A., and R. A. Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. Proc. R. Soc. Lond. Ser. B Biol. Sci. 263:1619–1626.

Beheregaray, L. B., and A. Caccone. 2007. Cryptic biodiversity in a changing world. J. Biol. 6:1–5.

Beheregaray, L. B., and P. Sunnucks. 2001. Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. Mol. Ecol. 10:2849–2866.

Berner, D., A.-C. Grandchamp, and A. P. Hendry. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. Evolution 63:1740–1753.

Bermingham, E., S. S. McCafferty, and A. P. Martin. 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. Pp. 113–128 *in* T. D. Kocher and C. A. Stepien, eds. Molecular systematics of fishes. Academic Press, San Diego, CA.

Bonin, A., D. Ehrich, and S. Manel. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. Mol. Ecol. 16:3737–3758.

Boughman, J. W. 2002. How sensory drive can promote speciation. Trends Ecol. Evol. 17:571–577.

Campbell, D., and L. Bernatchez. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Mol. Biol. Evol. 21:945–956.

Campbell, K. E., C. D. Frailey, and L. Romero-Pittman. 2006. The Pan-Amazonian Ucayali Peneplain, late Neogene sedimentation in Amazonia, and the birth of the modern Amazon River system. Palaeogeogr. Palaeoclimatol. Palaeoecol. 239:166–219.

Cooke, G. M., N. L. Chao, and L. B. Beheregaray. 2009. Phylogeography of a flooded forest specialist fish from central Amazonia based on intron DNA: the cardinal tetra *Paracheirodon axelrodi*. Freshwater Biol. 54:1216–1232.

———. 2012a. Divergent natural selection with gene flow along major environmental gradients in Amazonia: insights from genome scans, population genetics and phylogeography of the characin fish *Triportheus albus*. Mol. Ecol. 21:2410–2427.

———. 2012b. Natural selection in the water: freshwater invasion and isolation by adaptation to water colour in the Amazonian pufferfish. J. Evol. Biol. 25:1305–1320.

———. 2012c. Marine incursions, cryptic species and ecological diversification in Amazonia: the biogeographic history of the croaker genus *Plagioscion* (Sciaenidae). J. Biogeogr. 39:724–738.

———. 2012d. Five cryptic species in the Amazonian catfish *Centromochlus existimatus* identified based on biogeographic predictions and genetic data. PloS One 7:e48800.

Cooper, M. L. 2000. Random amplified polymorphic DNA analysis of southern brown bandicoot (*Isoodon obesulus*) populations in Western Australia reveals genetic differentiation related to environmental variables. Mol. Ecol. 9:469–479.

Coyne, J. A. 1992. Genetics and speciation. Nature 355:511–515.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Crampton, W. G. R. 1998. Electric signal design and habitat preferences in a species rich assemblage of gymnotiform fishes from the Upper Amazon Basin. An. Acad. Bras. Cienc. 70:805–847.

Crispo, E., P. Bentzen, D. N. Reznick, M. T. Kinnison, and A. P. Hendry. 2006. The relative influence of natural selection and geography on gene flow in guppies. Mol. Ecol. 15:49–62.

dos Santos Silva, D., S. S. R. Milhomem, A. C. Paes de Souza, J. C. Pieczarka, and C. Y. Nagamachi. 2008. A conserved karyotype of *Sternopygus macrurus* (Sternopygidae, Gymnotiformes) in the Amazon region: differences from other hydrographic basins suggest cryptic speciation. Micron 39:1251–1254.

Endler, J. A. 1977. Geographic variation, speciation, and clines. Princeton Univ. Press, Princeton, NJ.

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

Excoffier, L. G., Laval G., and S. Schneider. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol. Bioinfom. Online. 1:47–50.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics 164:1567–1587.

Feder, J. L., S. P. Egan, and P. Nosil. 2012. The genomics of speciation-with-gene-flow. Trends Genet. 28:342–350.

Feulner, P. G., F. Kirschbaum, C. Schugardt, V. Ketmaier, and R. Tiedemann. 2006. Electrophysiological and molecular genetic evidence for sympatrically occuring cryptic species in African weakly electric fishes (Teleostei: Mormyridae: Campylomormyrus). Mol. Phylogenet. Evol. 39:198–208.

Foll, M., and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977–993.

Funk, W. C., M. Caminer, and S. R. Ron. 2012. High levels of cryptic species diversity uncovered in Amazonian frogs. Proc. R. Soc. B. 279:1806–1814.

Futuyma, D. J. 1987. On the role of species in anagenesis. Am. Nat. 130:465–473.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165:131–137.

———. 1997. Alternative models of vertebrate speciation in Amazonia: an overview. Biodiv. Conserv. 6:451–476.

Hagedorn, M., and C. Carr. 1985. Single electrocytes produce a sexually dimorphic signal in South-American electric fish, *Hypopomus occidentalis* (Gymnotiformes, Hypopomidae). J. Comp. Physiol. A 156:511–523.

Hendry, A. P. 2009. Ecological speciation! Or the lack thereof? Canadian. J. Fish. Aquat. Sci. 66:1383–1398.

Hendry, A. P., D. I. Bolnick, D. Berner, and C. L. Peichel. 2009. Along the speciation continuum in sticklebacks. J. Fish Biol. 75:2000–2036.

Holderegger, R., and H. H. Wagner. 2008. Landscape genetics. BioScience 58:199–207.

Hoorn, C., J. Guerrero, G. A. Sarmiento, and M. A. Lorente. 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. Geology 23:237–240.

Hoorn, C., F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartín, A. Sanchez-Meseguer, C. L. Anderson, J. P. Figueiredo, et al. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330:927–931.

Hubert, N., and J.-F. Renno. 2006. Historical biogeography of South American freshwater fishes. J. Biogeogr. 33:1414–1436.

Hubert, N., F. Duponchelle, J. Nunez, C. Garcia-Davila, D. Paugy, and J. F. Renno. 2007. Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of the *Neotropical ichthyofauna*. Mol. Ecol. 16:2115–2136.

Jeffreys, H. 1961. Theory of probability. Clarenden Press, Oxford, U.K.

Jensen, J. D., A. Wong, and C. F. Aquadro. 2007. Approaches for identifying targets of positive selection. Trends Genet. 23:568–577.

Johnson, P. A., F. C. Hoppensteadt, J. J. Smith, and G. L. Bush. 1996. Conditions for sympatric speciation: a diploid model incorporating habitat fidelity and non-habitat assortative mating. Evol. Ecol. 10:187–205.

Jost, L. 2008. G(ST) and its relatives do not measure differentiation. Mol. Ecol. 17:4015–4026.

Kirschbaum, F. 1995. Taxonomy, zoogeography and general ecology of South American knifefishes (Gymnotiformes). Pp. 446–464 *in* P. Moller, ed. Behavior, history, electric fishes. Chapman and Hall, Lond.

Kramer, B., F. Kirschbaum, and H. Mark. 1981. Species specificity of electric organ discharges in a sympatric group of gymnotoid fish from Manaus (Amazonas). Pp. 195–219 *in* T. Szabo and G. Czeh, eds. Advances in physiological sciences. Pergamon Press, Oxford, U.K.

Kropf, M., H. P. Comes, and J. W. Kadereit. 2009. An AFLP clock for the absolute dating of shallow-time evolutionary history based on the intraspecific divergence of southwestern European alpine plant species. Mol. Ecol. 18:697–708.

Labonne, J., and A. P. Hendry. 2010. Natural and sexual selection giveth and taketh away reproductive barriers: models of population divergence in guppies. Am. Nat. 176:26–39.

Landguth, E. L., and N. Balkenhol. 2012. Relative sensitivity of neutral versus adaptive genetic data for assessing population differentiation. Conserv. Genet. 13:1421–1426.

Landguth, E. L., and S. A. Cushman. 2010. CDPOP: a spatially explicit cost distance population genetics program. Mol. Ecol. Res. 10:156–161.

Landguth, E. L., B. K. Hand, J. Glassy, S. A. Cushman, and M. A. Sawaya. 2012. UNICOR: a species connectivity and corridor network simulator. Ecography 35:9–14.

Lee, C. R., and T. Mitchell-Olds. 2011. Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. Mol. Ecol. 20:4631–4642.

Legendre, P., and Fortin M. J. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol. Ecol. Res. 10:831–844.

Littmann, M. W., B. M. Burr, and U. A. Buitrago-Suarez. 2001. A new cryptic species of *Sorubim* Cuvier (Siluriformes: Pimelodidae) from the upper and middle Amazon Basin. Proc. Acad. Nat. Sci. Phila. 151:87–93.

Lovejoy, N. R., E. Bermingham, and A. P. Martin. 1998. Marine incursion into South America. Nature 396:421–422.

Lu, G. Q., and L. Bernatchez. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (Coregonus clupeaformis): support for the ecological speciation hypothesis. Evolution 53:1491–1505.

Lundberg, J. C., L. G. Marshall, J. Guerrero, B. Horton, L. R. Malabarba, and F. Wesselingh. 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. Pp. 13–48 *in* L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena, eds. Phylogeny and classification of neotropical fishes. Edipucrs, Porto Alegre.

Lynch, J. D., and W. E. Duellman. 1997. Frogs of the genus Eleutherodactylus (Leptodactylidae) in Western Ecuador: systematics, ecology, and biogeography. Univ. Kans. Publ., Mus. Nat. Hist. 23:1–236.

Lynch, M., and B. G. Milligan. 1994. Analysis of population genetic structure with rapid markers. Mol. Ecol. 3:91–99.

Manel, S., M. K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. 18:189–197.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Can. Res. 27:209–220.

Moritz, C., J. L. Patton, C. J. Schneider, and T. B. Smith. 2000. Diversification of rainforest faunas: an integrated molecular approach. Ann. Rev. Ecol. Syst. 31:533–563.

Nagamachi, C. Y., J. C. Pieczarka, S. S. R. Milhomem, P. C. M. O'Brien, A. C. P. de Souza, and M. A. Ferguson-Smith. 2010. Multiple rearrangements in cryptic species of electric knifefish, *Gymnotus carapo* (Gymnotidae, Gymnotiformes) revealed by chromosome painting. BMC Genet. 11: 1–9.

Nakayama, C., M. Jegu, J. I. R. Porto, and E. Feldberg. 2001. Karyological evidence for a cryptic species of piranha within *Serrasalmus rhombeus* (Characidae, Serrasalminae) in the Amazon. Copeia 2001: 866–869.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283–292

Nielsen, R. 2005. Molecular signatures of natural selection. Ann. Rev. Genet. 39:197–218.

Nosil, P. 2008. Ernst Mayr and the integration of geographic and ecological factors in speciation. Biol. J. Linn. Soc. 95:26–46.

Nosil, P., S. P. Egan, and D. J. Funk. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. Evolution 62:316–336.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009a. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

Nosil, P., L. J. Harmon, and O. Seehausen. 2009b. Ecological explanations for (incomplete) speciation. Trends Ecol. Evol. 24:145–156.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes. 6:288–295.

Piggott, M. P., N. L. Chao, and L. B. Beheregaray. 2011. Three fishes in one: cryptic species in an Amazonian floodplain forest specialist. Biol. J. Linn. Soc. 102:391–403.

Quenouille, B., E. Bermingham, and S. Planes. 2004. Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. Mol. Phy. Evol. 31:66–88.

Reis, R. E., S. O. Kullander, and C. J. Ferraris. 2003. Checklist of freshwater fishes of South and Central America. Edipucrs, Porto Alegre.

Reznick, D. N., and R. E. Ricklefs. 2009. Darwin's bridge between microevolution and macroevolution. Nature 457:837–842.

Rosenblum, E. B., and L. J. Harmon. 2011. "Same same but different": replicated ecological speciation in White Sands. Evolution 65:946–960.

Rundle, H. D., and P. Nosil. 2005. Ecological speciation. Ecol. Lett. 8:336–352.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.

———. 2001. Ecology and the origin of species. Trends Ecol. Evol. 16:372–380.

———. 2009. Evidence for ecological speciation and its alternative. Science 323:737–741.

Schwarzer, J., E. R. Swartz, E. Vreven, J. Snoeks, F. P. D. Cotterill, B. Misof, and U. K. Schliewen. 2012. Repeated trans-watershed hybridization among haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution. Proc. R. Soc. B Biol. Sci. 279:4389–4398.

Sedano, R. E., and K. J. Burns. 2010. Are the Northern Andes a species pump for Neotropical birds? Phylogenetics and biogeography of a clade of Neotropical tanagers (Aves: Thraupini). J. Biogeogr. 37:325–343.

Silva, A., L. Quintana, R. Perrone, and F. Sierra. 2008. Sexual and seasonal plasticity in the emission of social electric signals. Behavioral approach and neural bases. J. Physiol. Paris 102:272–278.

Sioli, H. 1984. The Amazon limnology and landscape ecology of a mighty tropical river and its basin. Dr. Junk Publisher, Dordrecht.

Sistrom, M. J., N. L. Chao, and L. B. Beheregaray. 2009. Population history of the Amazonian one-lined pencilfish based on intron DNA data. J. Zoo. 278:287–298.

Sivasundar, A., E. Bermingham, and G. Orti. 2001. Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. Mol. Ecol. 10:407–417.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple-regression and correlation extensions of the mantel test of matrix correspondence. Syst. Zoo. 35:627–632.

Stinchcombe, J. R., and H. E. Hoekstra. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158–170.

Stoddard, P. K. 1999. Predation enhances complexity in the evolution of electric fish signals. Nature 400:254–256.

Storfer, A., M. A. Murphy, J. S. Evans, C. S. Goldberg, S. Robinson, S. F. Spear, R. Dezzani, E. Delmelle, L. Vierling, and L. P. Waits. 2007. Putting the 'landscape' in landscape genetics. Heredity 98:128–142.

Storfer, A., M. A. Murphy, S. F. Spear, R. Holderegger, and L. P. Waits. 2010. Landscape genetics: where are we now? Mol. Ecol. 19:3496–3514.

Sunnucks, P., and D. F. Hales. 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol. Biol. Evol. 13:510–524.

Thibert-Plante, X., and A. P. Hendry. 2010. When can ecological speciation be detected with neutral loci? Mol. Ecol. 19:2301–2314

Thompson, J. N. 2005. The geographic mosaic of coevolution. University of Chicago Press, Chicago, IL.

Thorpe, R. S., Y. Surget-Groba, and H. Johansson. 2008. The relative importance of ecology and geographic isolation for speciation in anoles. Philos. Trans. R. Soc. B Biol. Sci. 363:3071–3081.

Turchetto-Zolet, A. C., F. Pinheiro, F. Salgueiro, and C. Palma-Silva. 2013. Phylogeographical patterns shed light on evolutionary process in South America. Mol. Ecol. 22:1193–1213

Via, S. 2009. Natural selection in action during speciation. Proc. Nat. Acad. Sci. USA 106:9939–9946.

Whitlock, R., H. Hipperson, M. Mannarelli, R. K. Butlin, and T. Burke. 2008. An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. Mol. Ecol. Res. 8:725–735.

———. 1943. Isolation by distance. Genetics 28:114–138.

Zenger, K. R., A. J. Stow, V. Peddemors, D. A. Briscoe, and R. G. Harcourt. 2006. Widespread utility of highly informative AFLP molecular markers across divergent shark species. J. Hered. 97:607–611.

Zhivotovsky, L. A. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. Mol. Ecol. 8:907–913.

Associate Editor: A. Hendry

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Maximum likelihood phylogenetic tree showing the relationships between cryptic species based on the mitochondrial ATPase 6 and 8 (Fig. S1a) and the nuclear RAG1 genes (Fig. S1b).

Figure S2. Neighbor-joining tree showing the relationships between cryptic species and their sampling location based on the entire ATPase 6 and 8 dataset

Figure S3. STRUCTURE results of the total AFLP dataset using putatively neutral loci ($n = 289$).

Figure S4. Results of simulations showing genetic differentiation for selection-driven loci (dashed line), neutral loci (dash-dotted line), and combined selection-driven and neutral loci (solid line).

Figure S5. Geomorphological history of South American (1) rivers, lakes, and wet lands largely confined to a sedimentary basin in western Amazonia, (2) the formation of the modern trans-continental west-to-east flow of the Amazon River, and (3) the modern Amazon Basin with water "color" catchments shown.

Table S1. Sampling locations, sample size (n), and hydrochemical variables for *Steatogenys elegans* in the Amazon Basin (temperature, ºC; pH; turbidity, cm; dissolved oxygen (mg / L), OD; oxygen saturation, O_2 %).

Table S2. Population estimates of genetic diversity for mtDNA and AFLP data for each phylogroup.

Table S3. mtDNA Φ_{ST} value.

Table S4. BAYESCAN results following Jeffreys (1961) scale of evidence.

Table S5. Pairwise G_{ST} values for nine sites and low and high migration simulation scenarios.