Biological Journal of the Linnean Society, 2009, **97**, 259–274. With 5 figures

Population structure in the catfish *Trichogenes longipinnis*: drift offset by asymmetrical migration in a tiny geographic range

KELLY R. ZAMUDIO[1]*, JEANNE M. ROBERTSON[1,2], LAUREN M. CHAN[1,3] and IVAN SAZIMA[4]

[1]*Cornell University, Department of Ecology and Evolutionary Biology, Ithaca, NY 14853-2701, USA*
[2]*University of Idaho, Department of Biological Sciences, Moscow, ID 83844-3051, USA*
[3]*Duke University, Department of Biology, Durham, NC 27708, USA*
[4]*Universidade Estadual de Campinas, Museu de Zoologia, 13083-970 Campinas, São Paulo, Brazil*

Received 16 July 2008; accepted for publication 22 November 2008

Based on population genetic theory and empirical studies of small populations, we expect that species with very small ranges (narrow endemics) will exhibit reduced genetic diversity, increasing their susceptibility to the negative effects of genetic homogeneity. Although this pattern of reduced diversity applies to most narrow endemics, conservation biologists have yet to identify a general pattern for the degree of spatial population genetic structure expected in species with very small ranges. In part, this is because the degree of population structure within narrow endemics will be highly variable depending on the equilibrium between the homogenizing effects of dispersal and the diversifying effects of drift and local selection in small populations, thus precluding general predictions about the relative importance of small range, small population sizes, and habitat patchiness for maintaining genetic diversity in narrowly-distributed species. We document a striking example of high population structure in the tiny geographic range of a stream-dwelling catfish, *Trichogenes longipinnis*, endemic to the Atlantic Forest of Brazil. The maintenance of this diversity results from a combination of asymmetrical and limited dispersal, and drift in small populations. Our results highlight the need to understand population structure, and not only overall genetic diversity, of narrowly-distributed species for their conservation planning. *Biological Journal of the Linnean Society*, 2009, **97**, 259–274.

ADDITIONAL KEYWORDS: dispersal – genetic drift – microgeographic differentiation – rarity – Siluriformes.

INTRODUCTION

Plants and animals with highly restricted geographic ranges (narrow endemics) are considered rare and often are protected by international conservation laws (IUCN, 2004) based on the assumption that narrowly-distributed taxa are precariously avoiding extinction due to stochastic events and the detrimental effects of genetic uniformity (Brown, 1995; Lande, 1999). Genetic erosion and the negative effects of lowered adaptive genetic variation have been well documented in small populations (Purvis *et al.*, 2000; Spielman, Brook & Frankham, 2004), yet the assumption that narrow endemics are particularly susceptible to these negative effects is based on the simple observation that commonness (measured as abundance and/or range size) is positively correlated with genetic diversity (Frankham, 1996; Cole, 2003). Although reduced genetic diversity is evident in many narrow endemics, this pattern is not universal. Indeed, because of the interactions among population sizes, habitat patchiness, and population connectivity, it is difficult to derive generalized predictions of rangewide genetic diversity in species with exceedingly small ranges (Young & Brown, 1996; Cole, 2003).

Conservationists are particularly concerned with microevolutionary processes that cause genetic

*Corresponding author. E-mail: kelly.zamudio@cornell.edu

erosion in small populations (Ellstrand & Elam, 1993; Frankham, 1996). Genetic drift can be especially important and, in extreme cases, can lead to accumulation and fixation of deleterious mutations (Lynch & Gabriel, 1990; Lynch, Conery & Bürger, 1995) and the overall reduction of both neutral and adaptive genetic variation (Lande, 1988; Willi *et al.*, 2007). The rate at which genetic diversity is lost is mediated by other processes, such as population subdivision in patches of habitat, local adaptation in independent populations, and patterns of gene flow among subpopulations (Manier & Arnold, 2006; Willi *et al.*, 2007). Inbreeding and loss of evolutionary potential can be more extreme in narrow endemics (Cole, 2003) than in widespread species (Whitlock, 1992; Jehle *et al.*, 2005). If subpopulations of narrow endemics become isolated due to fragmented or patchy habitats and migration is curtailed, this can exacerbate overall decreases in diversity within populations and potentially result in a drastic loss of rangewide genetic diversity (Whitlock & Barton, 1997). Alternatively, if sufficient genetic diversity remains across the range of the species, even low levels of migration among subpopulations can slow the rate at which genetic diversity is lost (Slatkin, 1973). Therefore, in narrow endemics, the extent and direction of gene flow in subdivided populations are critical parameters for understanding the maintenance of rangewide genetic variation.

In the present study, we investigated the population genetics of *Trichogenes longipinnis*, a narrowly-distributed freshwater species endemic to the Atlantic Coastal Forest of southeastern Brazil, and a member of the pencil and parasitic catfish Family Trichomycteridae (de Pinna & Wosiacki, 2003). Our aim was to quantify microevolutionary processes that typically lead to population structure in subdivided populations, and to ask how they impact the spatial distribution of genetic diversity in a species with an extremely narrow distribution and small overall population size. *Trichogenes longipinnis* occurs only in a few steep streams along approximately 10 km of coast in São Paulo and Rio de Janeiro states (Sazima, 2004) (Fig. 1); however, despite its highly restricted range, this species exhibits substantial population-level variation in morphology (Sazima, 2004). Given this combination of a tiny geographic distribution and potentially high population differentiation and structure, this species is an ideal candidate for examining the relative impact of dispersal, population connectivity, and small population sizes on genetic variation among populations.

We sampled streams throughout the small range of this species and assessed within and among-population variation at nuclear microsatellite loci. If drift is sufficiently strong to reduce overall genetic differentiation rangewide, we expect all populations to exhibit lowered genetic diversity, even in the face of migration among subdivided populations. Alternatively, if drift acts on independent populations, but is mitigated by gene flow among populations, we expect to find variance among populations in levels of genetic diversity that are correlated with patterns of dispersal and population connectivity. Specifically, we expect: (1) higher genetic diversity in historically larger populations that have suffered less genetic erosion by drift; (2) diverse and larger populations serving as sources for recently founded or re-colonized populations; and (3) asymmetry in historical migration with disproportionately higher migration out of these larger source populations. These population dynamics are well known for species with patchy but widespread distributions (Francisco, Galetti & Galetti, 2006; Hanfling & Weetman, 2006; Koizumi, Yamamoto & Maekawa, 2006); however, it is less clear if and how they contribute to maintenance of rangewide population variation in species with very narrow ranges. Given the potentially larger effects of drift in narrow endemics, and the potential for rangewide loss of genetic variability, it is especially important to understand interactions among microevolutionary forces shaping the distribution of genetic variation in these species.

MATERIAL AND METHODS

Study species and population sampling

The genus *Trichogenes* is monotypic (Britski & Ortega, 1983) and diverged early in the evolution of the Trichomycteridae (de Pinna, 1992, 1998; Ribeiro, 2006). *Trichogenes longipinnis* has been recorded from approximately ten coastal streams or streamlets in three independent drainages near the border between São Paulo and Rio de Janeiro states (São Thiago, 1990; Sazima, 2004; F. C. T. Lima, pers. comm.) and is a strict habitat specialist, living exclusively in waterfall or riffle-fed pools of streams on steep hills with rocky or sandy bottoms (Sazima, 2004). It occurs from 149–656 m a.s.l. and is not found in more gently sloping terrain or lowland reaches of the streams it inhabits (Fig. 1).

We sampled populations of *T. longipinnis* from five streams in the municipalities of Ubatuba, São Paulo state and Parati, Rio de Janeiro state, Brazil (see Appendix). The two streams in Rio de Janeiro state (RJ599 and Buracão) belong to the Parati–Mirim watershed and drain from two opposite escarpments into a common river system (Fig. 1). The three streams in São Paulo (Cachoeira do Amor, Rio Camburi, and Paralelo Camburi) belong to two drainages that are isolated topographically from each other and from the

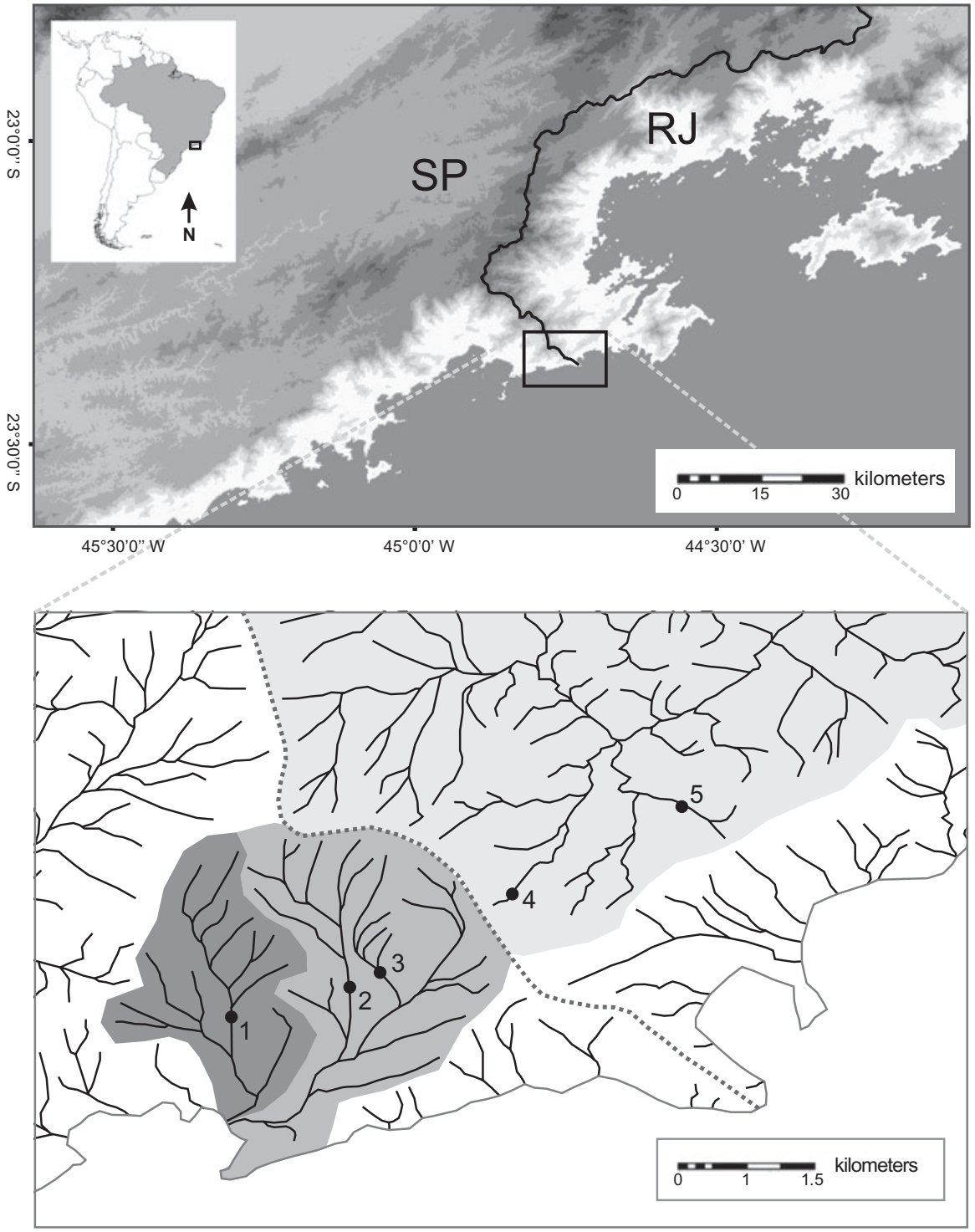

Figure 1. Topographic map of northern São Paulo and southern Rio de Janeiro with collection localities for *Trichogenes longipinnis*. The political state boundary coincides with a topographical divide that separates the range of this catfish species into the Rio Parati–Mirim watershed (Rio de Janeiro State) to the northeast and two independent and isolated drainages in São Paulo state to the southwest. The five sampled populations are: 1, Cachoeira do Amor; 2, Rio Camburi; 3, Paralelo Camburi; 4, RJ599; and 5, Buracão.

Parati–Mirim watershed, and drain independently into the sea (Fig. 1). The five streams span most of the known range of the species (São Thiago, 1990; Sazima, 2004). Two additional populations are known from two streamlets immediately south of Cachoeira do Amor (hereafter referred to as Amor), where *Trichogenes* has low (1–4 individuals m^{-2}) yet widely fluctuating densities (I. Sazima, pers. observ.); a few other populations are known from the Parati–Mirim watershed in Rio de Janeiro (F. C. T. Lima, pers. comm.).

At each sampling locality, we captured juveniles and adults with aquarium nets and baited funnel traps. We clipped fins from each fish with sterile scissors and preserved them in absolute ethanol; all sampled individuals were released at the site of capture. Samples were collected on three days over the course of two months (a period fin clips remain evident) to avoid duplicate collections. Population densities vary among streams (Sazima, 2004); thus, we aimed to collect 30 individuals from most sites, and collected as many individuals as possible from other localities. Our sample sizes per stream vary from 15 (Amor) to 36 (Rio Camburi).

Despite its extremely restricted range and habitat specialization, *T. longipinnis* can be fairly common in some pools: moving visual censuses of a 20 m diameter pool at the Amor site and stationary censuses of a smaller (4 m diameter) pool at Rio Camburi yielded estimates of 3–12 individuals m^{-2} and 18–25 individuals m^{-2}, respectively. Aggregations of fish attracted to fish food pellets in multiple smaller (< 4 m diameter) pools at Rio Camburi, Paralelo Camburi, and Buracão streams resulted in estimates of 15–18, 9–10, and 4–7 individuals m^{-2}, respectively. Therefore, the demographic and spatial distribution that characterizes this species appears to be moderate densities in each pool, but few isolated pools per stream. The populations in different streams have quantifiable differences in spot patterns (Sazima, 2004). Based on spot pattern similarity (I. Sazima, pers. observ.) we can infer three groups of streams with discernible phenotypes: (1) Amor and Paralelo Camburi; (2) Rio Camburi; and (3) RJ599 and Buracão (the two populations in Rio de Janeiro; Fig. 1).

Marker characterization and genotyping

We cloned microsatellite loci from an enriched partial genomic library prepared with liver tissue from one adult *T. longipinnis* from the RJ599 site. The voucher specimen was deposited in the ichthyological collection at the Museu de Zoologia, Universidade de Campinas, Brazil (accession number ZUEC 6226). Genomic DNA was extracted using a Qiagen DNeasy kit, digested with *Alu*I/*Hae*III (New England Biolabs), size selected for 500–700-bp fragments, and ligated to SNX linkers. Linked fragments were enriched for microsatellites with biotinylated dimer, trimer, and tetramer probes bound to streptavidin-coated magnetic beads (Dynabeads, Dynal Biotech). DNA fragments containing microsatellites were captured magnetically and amplified via polymerase chain reaction (PCR) with linker-specific primers. Amplification products were digested with *Nhe*I, cloned into pUC19 vector, and transformed using DH5α competent cells (Invitrogen). Colonies were grown on X-Gal/IPTG-coated agar plates and transferred to Magna Lift nylon membranes (Osmonics Inc.) that were later probed with the same series of di-, tri-, and tetra-nucleotide radio-labelled repeats. We cultured all positive clones and extracted plasmid DNA with Qiagen miniprep columns. Template DNA was sequenced directly with vector-specific primers (M13 F and R) using dGTP BigDye terminator cycle sequencing components on an ABI 3100 Genetic Analyzer (Applied Biosystems). We designed PCR primers in the flanking regions of 33 loci using Seqman (DNA STAR, version 5.05). Following optimization, seven loci consistently yielded specific PCR product of good concentration and were polymorphic in a pilot study (Table 1).

Fin clips were digested in lysis buffer with Proteinase K, and genomic DNA was purified using phenol–chloroform extraction (Sambrook & Russell, 2001). Concentrations of genomic DNA were measured and diluted (100 ng μL^{-1}) for use as template in amplification reactions. Each PCR reaction consisted of 100 ng of template, 0.05 μL of *Taq* polymerase (5 U μL^{-1}), 1.0 μL of 10 × PCR buffer (100 mM Tris-HCl, 15 mM $MgCl_2$, 500 mM KCl), 0.1 μL of dNTPs (40 mM), and 0.2 μL of each primer (10 μM), in a total volume of 10 μL. Forward primers were 5′-labelled with a fluorescent dye. Loci were amplified in a thermal cycler under the conditions: 5 min of initial denaturation at 94 °C; 35 cycles of 1 min of denaturation at 94 °C, 1 min of annealing at primer-specific temperatures (Table 1), 1 min of extension at 72 °C; and a final extension of 72 °C for 30 s. Amplified products with different labels or non-overlapping size ranges were multiplexed and electrophoresed on an ABI 3100 Genetic Analyzer. Fragment sizes were determined with a LIZ-500 standard using GENEMAPPER, version 3.5. We genotyped 117 individuals from the five sites sampled throughout the range of the species.

Population structure analysis

We calculated global and population genetic diversity indices. Allelic and private allelic richness within each population were estimated using rarefaction

Table 1. Microsatellite primers, motif, and optimal annealing temperatures for *Trichogenes longipinnis*

Locus*	Primer sequence (5′- to 3′)	Motif†	T_a (°C)
Tld33	F: TTA GGT TGC ATC ATG GGA CA R: GGC TGG CAT GCA TTT TCT TA	$(AC)_{24}$	61
Tlt16	F: CTC CAG CAT CTA GCA TCC AAC AT R: AAG AAA GGC TGA GCA GGT GAA AAT	$(CA)_{16}GA(CA)_3$	61
Tld15	F: AGA GAA AAC TAG GAC GCA GAA GG R: CGA GGG GAG ACG GCA TTA C	$(TG)_3CT(TG)_3CTGCTCG(TG)_{10}$	58
Tld19	F: AAA CAA TGT AAA AGC CCA GTA AT R: TGA AGC ATG TAA GGC AAG AGG TTT	$(CA)_6CTCACG(CA)_6$	64
Tld27	F: TGG TTG GCT TGT CTC AGG GTT TCT R: GTG CAA ATC AAT CTC CAA CAG C	$(CA)_3(CG)_3(TA)_2(CA)_8$	61
Tld17	F: TGT GAG TTA GCA GGG CAA GTT R: GCC TGG TCA TGA ATA T	$(CT)_8CG(CT)_6(CA)_2(CT)_8$	53
Tlt35	F: TTA CCG CTG CAT TAC TTG R: CTC ATG AAC CTC CAG GAT A	$(TG)_{17}$	54.4

*Forward primers were 5′ labelled with a fluorescent tag for automated fragment size detection.
†Original clone sequences have been submitted to GenBank (FJ489252-FJ489258).

implemented using the software HP-RARE to account for variance in samples size across sites (Kalinowski, 2004, 2005). Heterozygosity averaged over all loci in each population and the proportion of polymorphic loci in each population were estimated in GENALEX, version 6 (Peakall & Smouse, 2006). Observed (H_O) and expected (H_E) numbers of heterozygotes, and statistical deviations from Hardy–Weinberg (H–W) equilibrium were estimated in GENEPOP, version 3.1 (Raymond & Rousset, 1995). We used a Monte Carlo approximation of Fisher's exact test (Guo & Thompson, 1992) and a Bonferroni correction to test for H–W proportions in genotypic frequencies. The Markov chain algorithm was run for 100 000 steps following 10 000 dememorization steps. We also estimated pairwise probabilities of linkage disequilibrium using a Fisher's exact test implemented in GENEPOP, version 3.1, with a 10 000-step dememorization, 1000 batches, and 10 000 iterations per batch.

We performed a global test of overall population differentiation (not assuming H–W equilibrium within populations) using FSTAT, version 2.9.3 (Goudet, 1995). This test permutes genotypes among populations to create a null distribution for comparison with observed levels of population differentiation (Goudet *et al.*, 1996). We estimated population divergence using F_{ST} (using the estimator θ of Weir & Cockerham, 1984) and performed pairwise significance tests for F_{ST} (Goudet *et al.*, 1996) by permutation and resampling of multilocus genotypes among pairs of populations. A table-wide significance at the 5% nominal level after standard Bonferroni corrections (adjusted P-value = 0.005) was reached after 200 randomizations. We estimated R_{ST} (Slatkin, 1995) and F-statistics for hypothetical combinations of populations, with confidence intervals inferred by bootstrapping over loci. We grouped populations according to three a priori hypotheses of population division: assuming all five populations independently, assuming isolation among populations from streams in three different watersheds (Fig. 1), and assuming isolation among the three groups of populations that show similar patterns in morphology (Sazima, 2004; I. Sazima, pers. observ.). For each hypothetical grouping, we tested for significance of overall population differentiation (not assuming H–W equilibrium within populations) and the significance of F_{IS} within each of the pooled population groups.

Bayesian clustering analyses

The number of genetic demes in our sample, K, was estimated using Bayesian assignment implemented in STRUCTURE, version 2.1 (Pritchard, Stephens & Donnelly, 2000). STRUCTURE assumes K genetic clusters, each characterized by a set of allele frequencies, and the admixture model probabilistically estimates the proportion of individuals with ancestry in each cluster (Pritchard *et al.*, 2000). We clustered samples excluding information on population of origin, assuming independence among loci, and non-informative priors. We estimated $\Pr(X|K)$, where X represents the data, for K between 1 (one breeding deme) and 6 (the number of populations plus one).

Mean and variance of log likelihoods for each K were calculated from 20 independent runs of 1 000 000 iterations (following 100 000 iterations burn-in). We used ΔK (Evanno, Regnaut & Goudet, 2005) as our criterion to estimate the number of demes in our sample and DISTRUCT, version 1.0 (Rosenberg, 2004) to graph individual membership coefficients to inferred demes.

Contemporary and historical demographic parameters

We estimated migration at two temporal scales. First, we used GENECLASS2, version 2 (Piry *et al.*, 2004) to infer potential first generation immigrants and individual assignment probabilities of individuals to their source population. We used a Bayesian classification method (Rannala & Mountain, 1997) and a Monte Carlo simulation algorithm (Paetkau *et al.*, 2004) with 10 000 simulated individuals and $\alpha = 0.05$. For assignment probabilities, we used the Bayesian assignment and the Monte Carlo simulation method of Rannala & Mountain (1997) with 10 000 simulated individuals and both $\alpha = 0.001$ and $\alpha = 0.05$.

Second, to estimate historical population sizes and patterns of migration, we implemented Bayesian estimates of historical demographic parameters under the isolation with migration model (Hey & Nielsen, 2004; Hey, 2005) in the program IM. We used the six-parameter model, assuming constant population sizes for all pairs of populations (Hey, 2005); we estimated current and ancestral populations sizes (θ_1, θ_2, θ_A), time subsequent to divergence (t), and asymmetric migration rates (m_1 and m_2). We included in these analyses all pairs of genetic demes detected in the previous STRUCTURE analyses (i.e. populations RJ599 and Buracão pooled as a single panmictic population). Demographic estimates were based on a Markov chain Monte Carlo of at least 5 000 000 steps, following a 100 000 step burn-in with sampling every ten steps. We assumed a stepwise mutation model and used a geometric heating scheme with 20 coupled chains and heating parameters set to $g_1 = 0.8$ and $g_2 = 0.9$; for the Amor and RJ599/Buracão comparison, we used ten chains with the same heating parameters. We evaluated the efficacy of chain coupling and convergence by examining swapping rates among chains and extending runs until the effective sample size (ESS) values of parameter correlations along the chain were greater than 50. We ran each analysis three times with a different random seed number to evaluate congruency of results. Posterior probabilities (HiSmth) and 95% credible intervals were estimated from the distribution for each parameter. Asymmetry in population migration rates ($2N_1m_1$ and $2N_2m_2$), differences in population sizes between each pair of populations, and differences between the ancestral population sizes and each current population size were evaluated statistically by comparing the proportion of times one parameter was larger than the other over the course of the run.

RESULTS

Population structure analysis

Populations were generally in H–W equilibrium, however two populations (Amor and Paralelo Camburi) showed overall significant deviations from equilibrium after a correction for multiple comparisons (Table 2). Coincidentally, these are the least and most genetically diverse populations, respectively. Disequilibrium is due to a deficit of heterozygotes at a single locus (Tld33); this locus shows deviation from equilibrium in three of the five populations, possibly due to a null allele. We detected no significant associations (linkage disequilibrium) among loci using a Fisher's exact test across all populations (range of P-values = 0.056–0.948).

Our population with the smallest genotypic sample at a single locus included 11 individuals; therefore, we assumed a maximum of 22 genes when estimating allelic richness in HP-RARE (Kalinowski, 2005). Population genetic diversity shows a clear geographic pattern (Figs 2, 3). Our range edge populations (Amor and Buracão) are less genetically diverse than those in the center of the species' distribution (Fig. 2, Table 2). The edge populations have fewer alleles, fewer common and private alleles, and lower heterozygosity. By contrast, Rio Camburi and Paralelo Camburi, the two populations in the middle of the species' range, are more variable. This genetic diversity may be related to population sizes; field censuses indicate that Rio Camburi and Paralelo Camburi harbour the largest populations (Sazima, 2004).

We found high population differentiation in overall randomization tests and F_{ST} (Tables 3, 4). A randomization test for overall population differentiation (Goudet *et al.*, 1996) showed significant deviations at each locus individually, and over all loci combined ($P < 0.001$). Likewise, pairwise F_{ST} values were high, ranging from 0.1438 (Buracão/RJ599) to 0.5851 (Amor/Buracão, the two most distant populations). All F_{ST} estimates were significantly different from zero (Table 3). The lowest F_{ST} values were estimated for the sites Rio Camburi/Paralelo Camburi and for RJ599/Buracão, indicating genetic connectivity between neighbouring streams within each of those watersheds.

We estimated F-statistics for two hypothetical population groups reflecting a priori expectations of connectivity based on stream membership in watersheds and the geographic distribution of spot patterns

Table 2. Genetic variation in five sampled populations of *Trichogenes longipinnis*

Locus	Population	Amor	Rio Camburi	Paralelo Camburi	RJ599	Buracão
N		15	36	30	18	18
PPL	%	57.14	100	100	85.71	71.43
Tld33	*A*	6	4	9	5	5
	H_O	2	11	8	7	7
	H_E	8.428	21.508	22.818	10.226	11.935
	H–W *P*-value	**0.0000**	**0.0002**	**0.0000**	0.0603	0.0138
Tlt16	*A*	3	4	5	3	2
	H_O	10	11	13	2	1
	H_E	7.476	9.862	11.655	1.960	1.000
	H–W *P*-value	0.0469	1	0.7494	1	–
Tld15	*A*	1	3	4	3	2
	H_O	–	10	16	5	5
	H_E	–	9.030	14.736	4.600	4.429
	H–W *P*-value	–	1	0.6982	1	1
Tld19	*A*	2	2	2	1	1
	H_O	7	8	13	–	–
	H_E	6.5172	7.050	13.000	–	–
	H–W *P*-value	1	1	1	–	–
Tld27	*A*	1	6	7	4	2
	H_O	–	19	20	3	4
	H_E	–	19.522	20.947	3.788	7.097
	H–W *P*-value	–	0.9187	0.4668	0.1780	0.1066
Tld17	*A*	2	2	3	2	1
	H_O	3	23	17	8	–
	H_E	2.778	17.232	15.070	8.000	–
	H–W *P*-value	1	0.0783	0.6376	1	–
Tlt35	*A*	1	5	3	4	4
	H_O	–	21	9	4	4
	H_E	–	18.123	9.305	3.774	3.774
	H–W *P*-value	–	0.3943	1	1	0.6069
Overall	χ square (d.f.)	26.14 (8)	24.72 (14)	∞ (14)	9.07 (12)	14.05 (8)
	P-value	**0.001**	0.037	**0.000**	0.697	0.081

N, number of fish genotyped; PPL, proportion of polymorphic loci; A, number of alleles, H_O, observed number of heterozygotes, H_E, expected number of heterozygotes.
Significant deviations from Hardy–Weinberg (H–W) equilibrium (after Bonferroni correction for multiple comparisons) are shown in bold.

Table 3. Pairwise F_{ST} (below diagonal) and straight-line geographic distances (km; above diagonal) for all pairs of sampled populations of *Trichogenes longipinnis*

	Amor	Rio Camburi	Paralelo Camburi	RJ599	Buracão
Amor	–	1.459	1.728	3.541	5.708
Rio Camburi	0.4467	–	0.273	2.089	4.267
Paralelo Camburi	0.4158	0.2219	–	1.816	3.994
RJ599	0.5551	0.5064	0.3875	–	2.184
Buracão	0.5851	0.4949	0.4210	0.1438	–

All pairwise F_{ST} comparisons are statistically significant at $P < 0.005$, the tablewide nominal level with corrections for multiple comparisons.

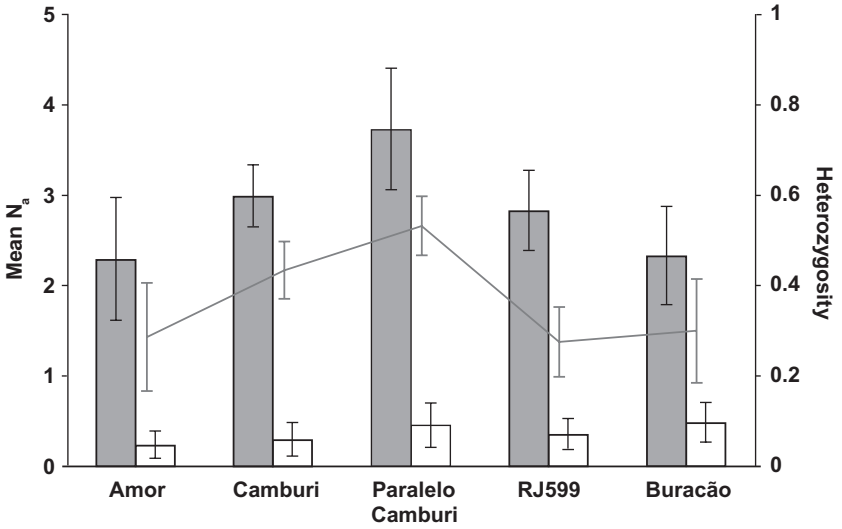

Figure 2. Patterns of genetic diversity among populations of *Trichogenes longipinnis*. Bars represent the mean ± SD of allelic richness (grey bars) and private allelic richness (white bars). The grey line plot represents mean heterozygosity (and associated SE) for each population (estimated across all loci). Populations are arranged according to their geographic position along the coast.

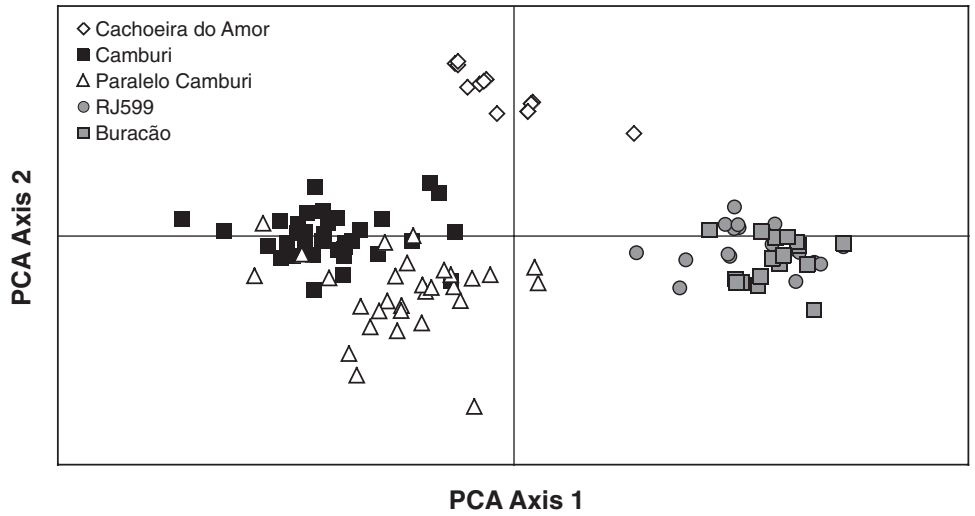

Figure 3. Scattergram of first two axes of a principal component analysis (PCA) of genetic variation in *Trichogenes longipinnis*. Black and white symbols are populations from the isolated streams in São Paulo; gray symbols represent populations in the Parati–Mirim watershed in Rio de Janeiro. Clustering of individuals corroborates the genetic similarity between the stream pairs Rio Camburi/Paralelo Camburi and RJ599/Buracão, and underscores the distinct nature of the population from Cachoeira do Amor.

(Table 4). In both cases, pooling samples resulted in reduced, but still significant, estimates of F_{ST} compared to the individual population analyses, suggesting that a significant amount, but not all, of the genetic variation is still explained within these groupings. By contrast, F_{IS} estimated for the pooled samples were higher and showed significant deviations from equilibrium, likely due to Wahlund effects as a result of combining samples from significantly distinct groups (Table 4).

Principal component analysis (PCA) corroborates genetic similarity between some pairs of neighbouring streams (Fig. 3). The first two PCA axes explain 44.13% and 21.06% of the sampled variation, respectively; the first three axes explain 80.03% of total genetic variation in our sample. Pairs of neighbouring

Table 4. Population genetic structure for various hypothetical clusters of *Trichogenes longipinnis* populations

Grouping†	Statistic*	F_{ST}	F_{IS}	F_{IT}	R_{ST}
Five populations	Mean	**0.415**	0.088	0.466	0.383
	Lower CI	0.274	–0.133	0.360	
	Upper CI	0.529	0.316	0.543	
	P-value	< 0.001			
Watersheds (three clusters)	Mean	**0.407**	**0.184**	0.516	0.290
	Lower CI	0.253	–0.018	0.404	
	Upper CI	0.535	0.369	0.588	
	P-value	< 0.001			
Morphology (three clusters)	Mean	**0.339**	**0.215**	0.481	0.193
	Lower CI	0.213	0.003	0.383	
	Upper CI	0.460	0.388	0.555	
	P-value	< 0.001			

*Estimates of F_{ST}, F_{IS}, F_{IT}, and R_{ST}, confidence intervals (CI) and F_{ST} *P*-values were calculated in FSTAT (Goudet, 1995).
†In the watershed-based analysis, populations were grouped according to membership in the three watersheds occupied by the species (Amor/Camburi + Paralelo Camburi/RJ599 + Buracão; Fig. 1). In the phenotype-based analysis, populations were grouped by morphological similarity (Amor + Paralelo Camburi/Camburi/RJ599 + Buracão).
Significant *F*-statistics are shown in bold.

streams within watersheds are clustered: Paralelo Camburi and Rio Camburi are somewhat distinct but overlap in the PCA scattergram. By contrast, sites RJ599 and Buracão, the two streams in the Parati–Mirim watershed in Rio de Janeiro are not distinguishable. The PCA analysis underscores the large genetic differentiation of the Amor population from other streams.

Bayesian clustering analysis

Bayesian clustering strongly supports four genetic demes (Fig. 4). The three populations in each of the São Paulo streams (Amor, Rio Camburi, and Paralelo Camburi) are independent genetic demes; the final deme is composed of the Parati–Mirim watershed populations in Rio de Janeiro (Buracão and RJ599). Partitioning of individual genotypes into more or less than four clusters was not supported: ΔK was 619.22 for $K = 4$ and less than 24.07 for other possible values of K. For $K = 4$, almost all individuals had high membership coefficients (mean $q = 0.9153 \pm 0.122$) in the cluster from which they were sampled. The isolation of the Amor site and the single deme in Rio de Janeiro state are consistent with membership of those streams to independent drainages (Fig. 1); Camburi and Paralelo Camburi clearly belong to independent demes, despite their close proximity within the same drainage. A total of 15 individuals showed $q < 0.8$ (three from Rio Camburi, seven from Paralelo Camburi, four from RJ599 and Buracão; Fig. 4). Closer examination of their membership coefficients to other demes indicates the highest shared ancestry between Paralelo Camburi and Rio Camburi, and also some mixing between those two streams and the Parati–Mirim watershed (Fig. 4).

Contemporary and historical demographic parameters

Tests for recent immigrants show that contemporary gene flow is extremely limited among *T. longipinnis* populations. A single individual met the statistical threshold for assignment as a first generation immigrant; this individual was collected at population Paralelo Camburi and its likely source population was the neighbouring stream Rio Camburi, a stream in the same watershed. Assignment tests corroborate low levels of gene flow; 96.6% of our samples were correctly assigned to the source populations (results identical for both levels of α). Four individuals were incorrectly assigned: one individual from Camburi was assigned genetically to Paralelo Camburi, and three individuals from RJ599 were assigned to Paralelo Camburi, a neighbouring stream, but in another watershed. In the four cases of incorrect assignments, probabilities met the tablewide threshold probability of $P > 0.01$, but were not exceedingly high (range 0.124–0.344). In all four cases, a smaller, but nontrivial assignment probability was also estimated for the actual source population. Assignment probabilities should be interpreted with caution because their power varies depending on number of loci and degree of differentiation among subpopulations. Cornuet *et al.* (1999) demonstrated that high assignment success could be obtained using a Baye-

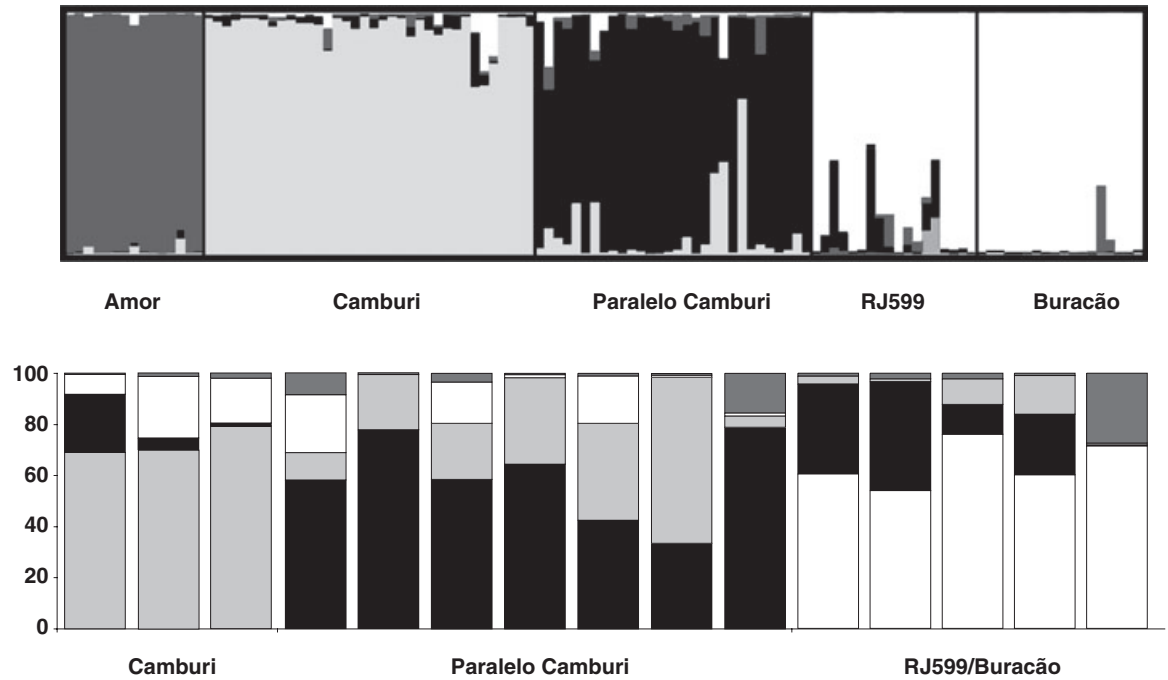

Figure 4. Population structure inferred by Bayesian assignment (implemented in STRUCTURE, version 2.1). *Trichogenes longipinnis* samples can be assigned to four geographic genetic demes (top), each represented by a single population or a pair of neighbouring populations. One deme (populations RJ599 and Buracão) is restricted to the Rio Parati–Mirim watershed, whereas the other three demes are from isolated streams in two drainages in São Paulo state. Fifteen individuals in our sample showed evidence of mixed ancestry (defined as less than 80% membership coefficient to a single deme). A closer look at the membership coefficients of these mixed individuals indicates a geographic pattern to shared ancestry (bottom). Most of admixed individuals from Paralelo Camburi and Rio Camburi share membership coefficient between those two demes. By contrast, mixed individuals from the two Rio de Janeiro populations show coancestry with the Paralelo Camburi population. None of the individuals from Amor show significant mixed ancestry, corroborating isolation of that population.

sian approach (Rannala & Mountain, 1997) in simulations with ten loci and 30–50 individuals per population, neither of which we have in the present study. However, a recent study of the performance of assignment tests in structured populations (Waples & Gaggiotti, 2006) showed that assignment success increases sharply with restricted gene flow, a condition that is met in the present study.

Historical migration rates between pairs of neighbouring populations (m_1 and m_2; Table 5) were low, with the lowest rates between Amor and the three other genetic demes. Migration rates between non-neighbouring populations were also low, in the range of 0.135–2.115. Populations Camburi and Paralelo Camburi showed the highest migration rates, followed by migration rates between those two populations and the genetic deme composed of RJ599 and Buracão (Table 5). We found evidence of asymmetrical migration rates for two population pairs: (1) immigration rates into Rio Camburi from Paralelo Camburi are higher than the reciprocal migration rate and (2) immigration rates into Paralelo Camburi from RJ599/Buracão are higher than the reciprocal. These patterns indicate that populations in this highly restricted range are primarily connected by migration between neighbouring demes (stepping stone pattern), with an overall pattern of asymmetric migration from larger populations such as Paralelo Camburi and the RJ599/Buracão deme (Table 5).

Migration rates need to be interpreted in light of population sizes because these determine effective migration between populations. Our IM estimates of population sizes (θ_1 and θ_2) are generally consistent with previous field estimates (Sazima, 2004). The overall pattern is one of increasing population sizes as one moves from southwest to northeast across the species' range. Amor, the most isolated stream also has the smallest estimated population size, whereas the easternmost demes Paralelo Camburi and RJ599/Buracão have the largest population sizes (Table 5). The ancestral population sizes (θ_A) are similar for all populations (Table 5) and estimates have broadly

Table 5. Historical demographic parameter inferred from pairwise population comparisons under a model of isolation with migration

		Parameter estimates*						Posterior probabilities†			
Population 1	Population 2	θ_1	θ_2	θ_A	t	m_1	m_2	$\theta_1 > \theta_2$	$\theta_1 > \theta_A$	$\theta_2 > \theta_A$	$\frac{2N_1m_1}{N_2m_2} > 2$
Amor	Camburi	0.051 (0.051–0.659)	0.090 (0.030–0.687)	12.359 (4.930–42.074)‡	0.025 (0.015–0.205)	0.008 (0.008–9.428)	1.058 (0.008–7.238)‡	0.0173	0	0	0.2596
Camburi	Paralelo Camburi	0.090 (0.030–0.568)	0.278 (0.119–1.072)	12.543 (0.053–42.181)‡	0.045 (0.015–0.275)	10.313 (0.938–40.988)‡	0.660 (0.020–17.140)‡	0.0380	0.0009	0.0022	0.5518
Paralelo Camburi	RJ599 & Buracão	0.238 (0.238–1.508)	0.227 (0.076–1.285)	15.798 (6.430–45.489)	0.095 (0.025–0.325)	2.530 (0.170–8.570)	1.050 (0.070–5.390)	0.7713	0	0	0.8824
Amor	Paralelo Camburi	0.051 (0.051–0.760)	0.595 (0.199–1.548)	13.845 (3.985–94.075)‡	0.155 (0.015–1.495)‡	0.735 (0.015–6.765)	0.585 (0.015–2.775)	0.0011	0.0001	0.0003	0.1403
Camburi	RJ599 & Buracão	0.090 (0.090–0.628)	0.113 (0.038–0.718)	13.450 (5.491–40.753)‡	0.045 (0.015–0.195)	2.115 (0.225–9.645)‡	1.215 (0.045–8.265)‡	0.3647	0	0	0.5878
Amor	RJ599 & Buracão	0.051 (0.051–0.760)	0.340 (0.113–1.020)	24.245 (8.577–85.971)	0.145 (0.055–0.465)	0.135 (0.015–5.505)	0.165 (0.015–1.755)	0.0048	0	0	0.4283

*Historical demographic parameters are scaled by mutation rate, μ, where $\theta = 4N_e\mu$ is the effective size of each population (θ_1, θ_2) and the ancestral population (θ_A), $t = t\mu$ is the coalescence time for populations 1 and 2, and $m = m/\mu$ is the migration rate into population 1 (m_1) and into population 2 (m_2). Parameter estimates include the peak of the posterior probability and the associated 90% credible interval (CI).
†Posterior probabilities for comparisons between parameters for effective sizes (θ) and the effective migration rates ($2Nm$).
‡Approximations to the 90% CI for instances of posterior probability densities with nonzero tails to the distribution.

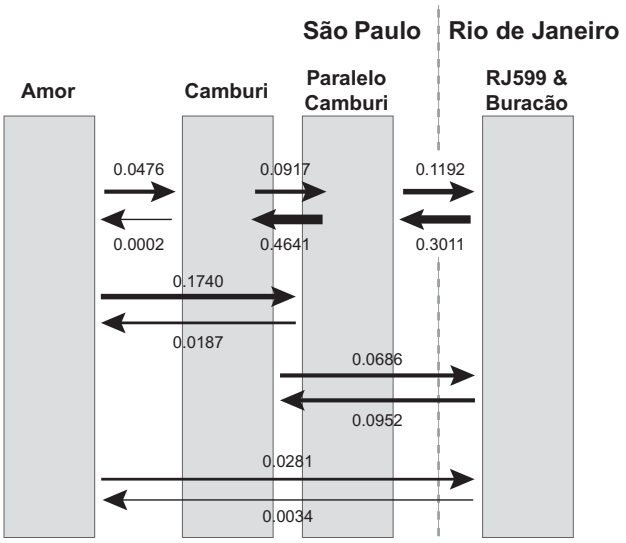

Figure 5. Effective migration rates among five stream populations of *Trichogenes longipinnis* under a model of isolation with migration. Populations are represented by vertical grey bars, organized in their relative positions along the coast. Values above and below each arrow are the directional effective migration rates between population pairs calculated as $(\theta_i \times m_i)/2$ using the peak estimate from the posterior probability densities for θ and m. The relative magnitude of directional migration is indicated by the thickness of each arrow.

overlapping 95% intervals. We found no evidence of a large population expansion for any of our five populations by comparing current and ancestral population sizes (Table 5) and the times subsequent to divergence for all pairs of populations are statistically indistinguishable. Taken in combination, the IM analyses indicate that our five populations currently have very different demographies, despite temporally coincident and similar demographic histories.

Adjusting migration rates to account for the large differences in effective population size corroborates the asymmetry in migration from larger source populations (RJ599/Buracão and Paralelo Camburi) to the smaller populations in the southwestern part of the range (Fig. 5).

DISCUSSION

Comparisons of broadly- and narrowly-distributed species show a significant pattern of decrease in genetic diversity with decreasing range size (Cole, 2003). This pattern is usually attributed to a correlation between narrow distributions and small population sizes and, despite large variation among taxa in population densities, it is evident in most taxa that have been examined to date (Hamrick & Godt, 1989;

Gaston, 1994; Gitzendanner & Soltis, 2000). An alternative explanation for overall reduced genetic diversity in narrow endemics involves the interaction between the age of taxa and the stability of habitat where they are found. Taxa basal within their lineages are often depauperate in species richness when compared to their sister groups, and can show extremely restricted geographical distributions (Stiassny & de Pinna, 1994). This pattern was recently demonstrated for several freshwater fish lineages from coastal drainages of Brazil, including Trichomycteridae (Ribeiro, 2006). These old, narrowly-endemic taxa therefore should have lower genetic variability due to long isolation and adaptation to stable environments; this hypothesis has been proposed in cases of extreme habitat specialization such as in the case of cave-dwelling taxa (Poulson & White, 1969) but the same reasoning can be applied to most long-branch taxa. The hypothesis posits that basal, low diversity taxa have survived over long periods because they are adapted to relatively stable environments; if they inhabited highly stochastic and unpredictable environments, they would have become extinct (Young & Brown, 1996). Directional selection and specialization to a narrow range of environments further reduces genetic diversity within species (Poulson & White, 1969; Willi *et al.*, 2007). *Trichogenes longipinnis* belongs to the large and diverse Trichomycteridae, a monophyletic family composed of approximately 170 species (de Pinna & Wosiacki, 2003). *Trichogenes longipinnis* (Trichogeninae) and five species of the genus *Copionodon* (Copiodontinae) diverged early from the ancestor of all other trichomycterids (de Pinna, 1998; de Pinna & Wosiacki, 2003) and the cladogenic event separating them from the remaining Trichomycteridae is very old (Ribeiro, 2006). Therefore, *T. longipinnis* is both a narrow-endemic and a long-branch taxon; however, our data show surprising levels of genetic diversity within and among populations (Fig. 2) and we can easily reject the hypothesis that this species as a whole is genetically uniform.

The geographic distribution of genetic diversity in *T. longipinnis* indicates two competing scenarios for historical population dynamics. The first is that genetic diversity is preserved in the larger northeastern populations that are not subject to high genetic drift. These large populations may have acted historically as colonization sources for range-edge populations, with overall low rates of immigration, resulting in founding of small edge populations with lower diversity. The alternative hypothesis is that the southwestern populations (Amor, Camburi, or both) are relicts of more widespread lineages of *Trichogenes* that were genetically differentiated, and upon secondary contact, the lineages admixed in the center of the species' current range. This introgression between two distinct lineages could have resulted in higher genetic diversity in central populations because of the combination of divergent alleles from the two source populations (Gantenbein & Largiadèr, 2002), the selective fixation or increase in alleles that were rare in parent populations (Woodruff, 1989) or increased mutation rates in hybrid populations (Thompson & Woodruff, 1978). Combined, our data on population genetic diversity, assignment probabilities, and migration support the hypothesis that Paralelo Camburi and the Parati–Mirim streams in Rio de Janeiro have been source populations in this metapopulation system with limited and primarily historic dispersal. Our coalescent migration analyses indicate that range-edge streams at Amor and Camburi are primarily recipients of immigrants, and likely serve as sinks in this population system with renewal of genetic diversity and population numbers from the easternmost populations in the species' range.

Narrow distributions are not always correlated with high genetic structure among subpopulations (Cole, 2003). An inverse relationship between range size and population structure is predicted based on the assumption that narrow endemics have small populations, and thus lower gene flow, the product of population size and migration rates (Nm) (Ellstrand & Elam, 1993). Therefore, the pattern of genetic structure depends both on dynamics within populations (drift, selection, and the decrease in genetic diversity), as well as dispersal among them. The balance between drift and migration in small populations will vary depending on the amount of standing variation, the degree of isolation, and the potential colonization–extinction or source-sink dynamics that may exist among subpopulations (Whitlock & Barton, 1997). The more isolated and subdivided subpopulations become, the more susceptible to increased loss of diversity through drift, especially in cases of narrow endemics that already have low genetic diversity (Lande, 1999; Cole, 2003). The degree of dispersal among subpopulations is highly variable and species-dependent (Ellstrand & Elam, 1993; Murray *et al.*, 2002; Rundle *et al.*, 2002), limiting generalizations about population structure in narrow endemics. Nonetheless, the pattern that we uncovered in the present study system may be representative of other aquatic taxa or terrestrial species with similarly restricted movement. Combined, our analyses reveal a pattern of historical isolation, some small populations with reduced diversity, and asymmetry in migration from larger source populations (Fig. 5). Patterns of current and historical connectivity emphasize the isolation and lowered genetic diversity of smaller populations (Amor), presumably due to genetic drift. We also found asymmetry in historical

migration between the larger northeastern populations and those on the southwestern edge of the range (Fig. 5). Paralelo Camburi and the Parati–Mirim streams are the likely source for migrants to other populations; exchange may have been associated with colonization of those streams, with repeated genetic rescue events after local extinctions or with drastic reductions in genetic diversity in the small edge populations.

The results obtained in the present study emphasize the distinct nature of edge populations, even in species with exceedingly small ranges. Both Buracão and Amor sites are range limit populations and have the lowest population densities of all populations sampled (Sazima, 2004), yet they differ in degree of isolation and connectivity with source populations. Buracão is part of the more extensive and interconnected Parati–Mirim watershed and therefore shows little differentiation from RJ599, the other stream sampled in that drainage, suggesting they are members of a larger interbreeding group of populations in the Rio de Janeiro drainage. By contrast, the Amor stream runs independently and approximately parallel to other streams along the Atlantic Coastal Forest escarpment before draining into the sea, resulting in larger genetic differentiation at the southeastern limit of the species' range.

Morphological variation (e.g. size and number of spots) among populations of *T. longipinnis* is not clinal (Sazima, 2004); however, the distribution of phenotypes is also unrepresentative of historical connections among populations. The degree of phenotypic differentiation is striking, considering that streams in the same drainage (Camburi and Paralelo Camburi) are highly morphologically divergent. This might reflect genetic drift and the fixation of different spot patterns in certain streams (e.g. Rio Camburi and Amor, the most differentiated populations) due to stochastic events during the founding of populations or historical reductions in population sizes. It is also possible that this pattern is due to localized selection for specific phenotypes in different streams (Crispo *et al.*, 2006). In fragmented or patchy landscapes, directional selection can reduce gene flow and genetic diversity within small peripheral populations because maladapted immigrants and hybrids should have reduced survival (Willi *et al.*, 2007). Selective advantages of spot patterns (e.g. imposed by visual predators) in streams with different physical characteristics would be an interesting avenue for future research. We cannot differentiate between drift and selection as causes for the morphological diversity among *Trichogenes* populations, but this phenotypic diversity appears to be independent of genetic connectivity among populations.

Our study of *Trichogenes longipinnis* has two practical implications: one for systematics and the other for conservation. *Trichogenes* is monotypic (Britski & Ortega, 1983; de Pinna & Wosiacki, 2003) but the present study indicates that it may contain more species, corresponding to our four differentiated demes (Amor, Paralelo Camburi, Rio Camburi, and RJ599/Buracão; Fig. 4). Of these, the populations of Rio Camburi and Amor are the most distinctive (Sazima, 2004) and also the most genetically isolated. The populations from streams across the Parati–Mirim watershed (RJ599, Buracão and others that remain unsampled) most likely constitute a single species. Additional genetic studies with sampling of other isolated streams across the tiny range of this catfish may reveal an even more complex scenario of genetic and morphological differentiation. Sazima (2004) showed consistent differences in phenotype among populations and suggested that, if significant genetic differences were to be found, each of these morphologically differentiated populations (Amor and Rio Camburi) should receive the status of endangered, EN B1+2bcd (IUCN, 2004). Our results, however, would qualify these two populations and that from Paralelo Camburi as critically endangered, CR B1ab (iii, iv) (IUCN, 2004). A complicating factor in the conservation of these populations is that Paralelo Camburi, the most genetically diverse population, is being depleted by subsistence over fishing by local people. The opposite case is observed in the population of Rio Camburi, where pools are used for recreation and the fish receive anthropogenic food supplement from local people and tourists (I. Sazima, pers. observ.). On the other hand, populations from the Parati–Mirim watershed in Rio de Janeiro should receive the status of vulnerable, VU B1ab (iii, iv) (IUCN, 2004) because they have higher population connectivity and are more widely distributed. Conservation strategies for this species should not be concerned about low rangewide population genetic variation, unless all populations become significantly reduced. However, to maintain high genetic variation overall, it will be important to conserve all differentiated populations from within the species' range.

The general assumption underlying the view of endangerment in narrow endemics is that there is a correlation between small range and small population size (Gaston, 1994; Purvis *et al.*, 2000). Both range size and population density are good predictors of vulnerability to extinction (Purvis *et al.*, 2000), suggesting a link between those two characteristics; however, empirical evaluation of narrow endemics indicates that this is not always the case (Trajano, 2001; Bichuette & Trajano, 2005; Buhay & Crandall, 2005). Given the genetic and conservation implications of these two traits, it would be useful to explore

the taxonomic distribution and ecological correlates of taxa with this combination of demography and distribution. Many narrow endemics are protected based on the criterion of small range size alone (IUCN, 2004). In addition, basal clades of widespread lineages are commonly species-poor and have restricted distribution (Stiassny & de Pinna, 1994) and thus deserve special attention from the conservation viewpoint. Most narrow endemics have reduced genetic diversity (Cole, 2003) compared to their broad-ranged sister taxa; likewise, threatened or endangered species in general show reduced genetic diversity (Spielman *et al.*, 2004). As our example shows, some extreme narrow endemics can be highly structured over very short geographic distances, exacerbating the difficulties of conserving genetic diversity in those rare taxa. Conservationists should not assume that small geographic ranges signify low genetic diversity; understanding genetic structure and the interplay between inter- and intrapopulation dynamics in maintaining this genetic diversity may be critical to guarantee adaptive potential to different selective environments and avoid deleterious effects of inbreeding in narrow endemics.

ACKNOWLEDGEMENTS

We thank C. A. Brasileiro, C. F. B. Haddad, M. R. C. Martins, R. J. Sawaya, and M. Sazima for company on field trips to Picinguaba; H. W. Greene for comments on the manuscript; L. Davis and A. Talaba for help with genotyping; and J. G. R. Giovanelli and C. G. Becker for help with distribution maps. Samples were collected under IBAMA permit number 69/2002; tissues were exported under permit number 0116729BR issued by the Brazilian Ministry of the Environment (MMA/BR). K.Z. thanks Célio Haddad, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Fulbright for support during a sabbatical at Universidade Estadual Paulista, Rio Claro. Field and laboratory expenses were funded by research grants from the Conselho Nacional de Desenvolvimento Científico e Technológico (CNPq) and FAPESP (to I.S.), and awards from the National Geographic Society, and the National Science Foundation (NSF) (to K.Z.). Analyses benefited from computing resources at the Computational Biology Service Unit at Cornell University, a facility partially funded by Microsoft Corporation.

REFERENCES

Bichuette ME, Trajano E. 2005. A new cave species of *Rhamdia* (Silurifomes: Heptapteridae) from Serra do Ramalho, northeastern Brazil, with notes on ecology and behavior. *Neotropical Ichthyology* **3:** 587–595.

Britski HA, Ortega H. 1983. *Trichogenes longipinnis*, novo gênero e espécie de Trichomycterinae do sudeste do Brasil (Pisces, Siluriformes). *Revista Brasileira de Zoologia* **1:** 211–216.

Brown JH. 1995. *Macroecology*. Chicago, IL: Chicago University Press.

Buhay JE, Crandall KA. 2005. Subterranean phylogeography of freshwater crayfishes shows extensive gene flow and surprisingly large population sizes. *Molecular Ecology* **14:** 4259–4273.

Cole TC. 2003. Genetic variation in rare and common plants. *Annual Review of Ecology and Systematics* **34:** 213–237.

Cornuet J-M, Piry S, Luikart G, Estoup A, Solignac M. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153:** 1989–2000.

Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP. 2006. The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology* **15:** 49–62.

Ellstrand NC, Elam DR. 1993. Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics* **24:** 217–242.

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.

Francisco M, Galetti M, Galetti PM Jr. 2006. Atlantic forest fragmentation and genetic diversity of an isolated population of the Blue-manakin, *Chiroxiphia caudata* (Pipridae), assessed by microsatellite analyses. *Revista Brasileira de Ornitologia* **14:** 21–28.

Frankham R. 1996. Relationship of genetic variation to population size in wildlife. *Conservation Biology* **10:** 1500–1508.

Gantenbein B, Largiadèr CR. 2002. *Mesobuthus gibbosus* (Scorpiones: Buthidae) on the island of Rhodes – hybridization between Ulysses' stowaways and native scorpions? *Molecular Ecology* **11:** 925–938.

Gaston KJ. 1994. *Rarity*. London: Chapman & Hall.

Gitzendanner MA, Soltis PS. 2000. Patterns of variation in rare and widespread plant congeners. *American Journal of Botany* **87:** 783–792.

Goudet J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86:** 485–486.

Goudet J, Raymond M, DeMeeus T, Rousset F. 1996. Testing differentiation in diploid populations. *Genetics* **144:** 1933–1940.

Guo SW, Thompson EA. 1992. Performing the exact tests of Hardy–Weinberg proportions for multiple alleles. *Biometrics* **48:** 361–372.

Hamrick JL, Godt MJW. 1989. Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS, eds. *Plant population genetics. Breeding and genetic resources*. Sunderland, MA: Sinauer Associates, 43–63.

Hanfling B, Weetman D. 2006. Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the River Sculpin, *Cottus gobio*. *Genetics* **173:** 1487–1501.

Hey J. 2005. On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology* **3:** 965–975.

Hey J, Nielsen R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167:** 747–760.

IUCN. 2004. *IUCN red list categories and criteria: version 3.1*. Gland: IUCN Species Survival Commission.

Jehle R, Wilson GA, Arntzen JW, Burke T. 2005. Contemporary gene flow and the spatio-temporal genetic structure of subdivided newt populations (*Triturus cristatus*, *T. marmoratus*). *Journal of Evolutionary Biology* **18:** 619–628.

Kalinowski ST. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics* **5:** 539–543.

Kalinowski ST. 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5:** 187–189.

Koizumi I, Yamamoto S, Maekawa K. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden charr. *Molecular Ecology* **15:** 3175–3189.

Lande R. 1988. Genetics and demography in biological conservation. *Science* **241:** 1455–1460.

Lande R. 1999. Extinction risk from anthropogenic, ecological and genetic factors. In: Landweber LF, Dobson AP, eds. *Genetics and the extinction of species*. Princeton, NJ: Princeton University Press, 1–22.

Lynch M, Conery J, Bürger R. 1995. Mutation accumulation and the extinction of small populations. *American Naturalist* **146:** 489–518.

Lynch M, Gabriel W. 1990. Mutation load and the survival of small populations. *Evolution* **44:** 1725–1737.

Manier MK, Arnold SJ. 2006. Ecological correlates of population genetic structure: a comparative approach using a vertebrate metacommunity. *Proceedings of the Royal Society Biological Sciences Series B, Biological Sciences* **273:** 3001–3009.

Murray BR, Thrall PH, Gill AM, Nicotra AB. 2002. How plant life-history and ecological traits relate to species rarity and commonness at varying spatial scales. *Austral Ecology* **27:** 291–310.

Paetkau D, Slade R, Burden M, Estoup A. 2004. Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13:** 55–65.

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

de Pinna MCC. 1992. A new subfamily of Trichomycteridae (Teleostei: Siluriformes), lower loricarioid relationships and a discussion on the impact of additional taxa for phylogenetic analysis. *Zoological Journal of the Linnean Society* **106:** 175–229.

de Pinna MCC. 1998. Phylogenetic relationships of Neotropical Siluriformes (Teleostei: Ostariophysi); historical overview and synthesis of hypotheses. In: Malabarba LR, Reis RE, Vari RP, Lucena ZMS, Lucena CAS, eds. *Phylogeny and classification of neotropical fishes*. Porto Alegre: EdiPUCRS, 279–330.

de Pinna MCC, Wosiacki W. 2003. Family Trichomycteridae (Pencil or parasitic catfishes). In: Reis RE, Kullander SO, Ferraris CJ Jr, eds. *Check list of the freshwater fishes of South and Central America*. Porto Alegre: EdiPUCRS 270–290.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A. 2004. GENECLASS2: a software for genetic assignment and first generation migrant detection. *Journal of Heredity* **95:** 536–539.

Poulson TL, White WB. 1969. The cave environment. *Science* **165:** 971–981.

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

Purvis A, Gittleman JL, Cowlishaw G, Mace GM. 2000. Predicting extinction risk in declining species. *Proceedings of the Royal Society Biological Sciences Series B* **267:** 1947–1952.

Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Science of the United States of America* **94:** 9197–9201.

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

Ribeiro AC. 2006. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. *Neotropical Ichthyology* **4:** 225–246.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4:** 137–138.

Rundle SD, Foggo A, Choiseul V, Bilton DT. 2002. Are distribution patterns linked to dispersal mechanism? An investigation using pond invertebrate assemblages. *Freshwater Biology* **47:** 1571–1581.

Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

São Thiago H. 1990. *Composição e distribuição longitudinal da ictiofauna do Rio Parati-Mirim (RJ) e período reprodutivo das principais espécies*. MSc Thesis, Museu Nacional.

Sazima I. 2004. Natural history of *Trichogenes longipinnis*, a threatened trichomycterid catfish endemic to Atlantic forest streams in southeast Brazil. *Ichthyological Exploration of Freshwaters* **15:** 49–60.

Slatkin M. 1973. Gene flow and selection in a cline. *Genetics* **75:** 733–756.

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

Spielman D, Brook BW, Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Science of the United States of America* **101:** 15261–15264.

Stiassny M, de Pinna MCC. 1994. Basal taxa and the role of cladistic patterns in the evaluation of conservation priorities: a view from freshwater. In: Forey PL, Humphries CJ, Vane-Wright RI, eds. *Systematics and conservation evaluation*. Oxford: Clarendon Press, 235–249.

Thompson JN, Woodruff RC. 1978. Mutator genes – pacemakers of evolution. *Nature* **274:** 317–321.

Trajano E. 2001. Habitat and population data of troglobitic armoured cave catfishes, *Ancistrus cryptophthalmus* Reis 1987, from Central Brazil (Siluriformes: Locariidae). *Environmental Biology of Fishes* **62:** 195–200.

Waples RS, Gaggiotti O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15:** 1419–1439.

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

Whitlock MC. 1992. Non-equilibrium population structure in forked fungus beetles: extinction, colonisation, and the genetic variance among populations. *American Naturalist* **139:** 952–970.

Whitlock MC, Barton NH. 1997. The effective size of a subdivided population. *Genetics* **146:** 427–441.

Willi Y, Van Buskirk J, Schmid B, Fischer M. 2007. Genetic isolation of fragmented populations is exacerbated by drift and selection. *Journal of Evolutionary Biology* **20:** 534–542.

Woodruff DS. 1989. Genetic anomalies associated with Cerion hybrid zones: the origin and maintenance of new electromorphic variants called hybridzymes. *Biological Journal of the Linnean Society* **36:** 281–294.

Young AG, Brown AHD. 1996. Comparative population genetic structure of the rare woodland shrub *Daviesia suaveolens* and its common congener *D. mimosoides*. *Conservation Biology* **10:** 1220–1228.

APPENDIX

Coordinates for five *Trichogenes* sampling localities in southeastern Brazil. Population numbers correspond to those in Figure 1.

Number	Locality	Stream name	State	Municipality	Latitude	Longitude
1	Amor	Cachoeira do Amor	São Paulo	Ubatuba	–23.36006667	–44.78330000
2	Camburi	Rio Camburi	São Paulo	Ubatuba	–23.35571667	–44.76981667
3	Paralelo Camburi	None	São Paulo	Ubatuba	–23.35451667	–44.76748333
4	RJ599	None	Rio de Janeiro	Parati	–23.34728333	–44.75153333
5	Buracão	None	Rio de Janeiro	Parati	–23.33670000	–44.73351667