Molecular Ecology (2006) **15**, 49–62 doi: 10.1111/j.1365-294X.2005.02764.x

The relative influence of natural selection and geography on gene flow in guppies

ERIKA CRISPO,* PAUL BENTZEN,† DAVID N. REZNICK,‡ MICHAEL T. KINNISON§ and ANDREW P. HENDRY*

**Redpath Museum and Department of Biology, McGill University, Montréal, QC H3A 2K6, Canada, †Department of Biology, Dalhousie University, Halifax, NS B3H 4J1, Canada, ‡Department of Biology, University of California, Riverside, CA 92521, USA, §Department of Biological Sciences, University of Maine, Orono, ME 04469-5751, USA*

Abstract

Two general processes may influence gene flow among populations. One involves divergent selection, wherein the maladaptation of immigrants and hybrids impedes gene flow between ecological environments (i.e. ecological speciation). The other involves geographic features that limit dispersal. We determined the relative influence of these two processes in natural populations of Trinidadian guppies (*Poecilia reticulata*). If selection is important, gene flow should be reduced between different selective environments. If geography is important, gene flow should be impeded by geographic distance and physical barriers. We examined how genetic divergence, long-term gene flow, and contemporary dispersal within a watershed were influenced by waterfalls, geographic distance, predation, and habitat features. We found that waterfalls and geographic distance increased genetic divergence and reduced dispersal and long-term gene flow. Differences in predation or habitat features did not influence genetic divergence or gene flow. In contrast, differences in predation did appear to reduce contemporary dispersal. We suggest that the standard predictions of ecological speciation may be heavily nuanced by the mating behaviour and life history strategies of guppies.

Keywords: migration, ecological speciation, F_{ST}, isolation by distance, parapatry, reproductive isolation

Received 27 April 2005; revision received 5 August 2005; accepted 1 September 2005

Introduction

Adaptive divergence can occur when populations in different ecological environments experience divergent patterns of selection (Schluter 2000). These same populations, however, may be linked by the physical dispersal of individuals (Slatkin 1987). If dispersal results in high gene flow, adaptive divergence may be constrained. The resulting balance between diversifying selection and homogenizing gene flow has been investigated in a number of theoretical models (e.g. Hendry *et al*. 2001; Lenormand 2002) and empirical studies (e.g. Riechert 1993; King & Lawson 1995; Hendry *et al*. 2002; Hendry & Taylor 2004; Nosil & Crespi 2004). Less studied is the opposite causal pathway: adaptive divergence may reduce gene flow because individuals that disperse between ecological environments are maladapted and thus may have lower fitness than adapted residents (i.e. 'ecological speciation'; reviews: Schluter 2000; Rundle & Nosil 2005). This reduction in gene flow may occur if dispersers are more likely to die before reproducing or have lower fecundity or mating success than residents, or produce hybrid offspring with reduced fitness (e.g. Via 1999; Via *et al*. 2000; Rundle 2002; Hendry 2004; Nosil 2004; Nosil *et al*. 2005). Under these conditions, realized gene flow may be substantially lower than dispersal, leading to at least partial reproductive isolation between populations adapted to different ecological environments.

A number of empirical studies have demonstrated that divergent selection can reduce gene flow in nature (e.g. Lu & Bernatchez 1999; Cooper 2000; Ogden & Thorpe 2002; Rocha *et al*. 2005). But what is the relative importance of this ecological effect within the overall context of nonecological, geographic features that can also influence gene flow? For example, dispersal should be negatively impacted by

Correspondence: Erika Crispo, Fax: (514)398-3185; E-mail: erika.crispo@mail.mcgill.ca

vicariance owing to physical barriers (review: Coyne & Orr 2004, pp. 86–105) and by increasing physical distance (i.e. isolation by distance; Wright 1943). If these geographic features explain most of the variation in gene flow, then perhaps divergent selection plays only a limited additional role. Indeed, some studies have found that gene flow can be high even in the presence of divergent selection (e.g. Smith *et al*. 2001; Emelianov *et al*. 2004; Hendry & Taylor 2004; Smith *et al*. 2005). More studies are thus needed to determine the relative importance of selection and geography on gene flow in nature. We address this question by examining genetic differentiation and gene flow among natural populations of Trinidadian guppies (*Poecilia reticulata*), a system conveniently characterized by adaptive divergence and the potential for substantial, but not unlimited, gene flow.

Adaptive divergence in guppies

Guppies are a live-bearing fish, common throughout Trinidad. A variety of selective factors influence adaptive divergence in guppies, with predation having been studied in greatest detail (reviews: Endler 1995; Houde 1997; Magurran 1998). In particular, populations below waterfalls in the lower reaches of streams typically coexist with a suite of piscivorous fishes ('high predation'), whereas those above waterfalls in the upper reaches of streams typically coexist with only a few weakly piscivorous fishes ('low predation'). Reflecting this selective dichotomy, high- and low-predation guppies differ in a broad suite of phenotypic traits. With respect to external appearance, males in low-predation sites tend to be more colourful than those in high-predation sites (Endler 1978, 1980), because bright male colour is favoured by females (review: Houde 1997) but selected against by piscivorous fishes (Endler 1980, 1983; Godin & McDonough 2003). With respect to behaviour, guppies in high-predation sites show increased schooling (Seghers 1974; Breden *et al*. 1987; Magurran *et al*. 1992, 1995), more careful predator inspection (Magurran & Seghers 1990; Magurran *et al*. 1992; Kelley & Magurran 2003), and increased escape ability (O'Steen *et al*. 2002; Ghalambor *et al*. 2004). With respect to life history, females in high-predation sites mature earlier, have higher reproductive effort, and have more (but smaller) offspring (Reznick & Bryga 1987; Reznick *et al*. 1996a, 2002).

Predation is clearly an important selective factor for guppies, but physical habitat features are also important. Canopy openness influences primary productivity and food availability, which in turn influence life history traits (Grether *et al*. 2001a; Reznick *et al*. 2001) and male colour (Kodric-Brown 1989; Grether *et al*. 2001b). Substrate size influences colour spot size through selection for background matching (Endler 1978, 1980). Water velocity selects for aspects of swimming ability and male courtship behaviour (Nicoletto 1996; Nicoletto & Kodric-Brown 1999). Stream size (e.g. width and depth) might influence resource use and foraging strategies, although this has yet to be examined. In short, divergence among guppy populations reflects adaptation to both predation and physical habitat features. The theory of ecological speciation therefore predicts that gene flow should decrease between sites with increasing differences in these ecological factors.

Gene flow in guppies

Guppies show potential for gene flow between selective environments. First, heterozygosity and allelic diversity are greater downstream than upstream (Shaw *et al*. 1991, 1992, 1994), as expected with downstream gene flow. Second, genetic markers in guppies introduced to new locations introgress into populations downstream of the introduction site (Haskins *et al*. 1961; Shaw *et al*. 1992; Becher & Magurran 2000). Third, several aspects of guppy mating behaviour may increase gene flow. For example, females from high-predation sites mate willingly with low-predation males (Endler & Houde 1995), females often prefer novel male phenotypes (Hughes *et al*. 1999; Eakley & Houde 2004), and males may use 'sneaky' copulations to circumvent female choice (review: Magurran 1998).

The guppy system also shows considerable potential for geographic features to reduce contemporary dispersal (i.e. movement of individuals) and long-term gene flow (i.e. long-term integration of gene pools). First, the dispersal of individual guppies appears limited by distance (Reznick *et al*. 1996a; D. Weese, M.T. Kinnison, A.P. Hendry, unpublished). We might therefore expect to see a signature of isolation by distance (IBD) within watersheds. Second, waterfalls should severely limit upstream dispersal (Becher & Magurran 2000). Third, water flow should facilitate downstream dispersal and hinder upstream dispersal. Substantial variation in selective environments and geographic features in this system thus affords an opportunity to determine the relative importance of various factors influencing gene flow.

Our goal was to assess the relative impacts of selection and geography on guppy gene flow. One potential problem, however, is that waterfalls limit upstream movement by both guppies and predators. Waterfalls thus create covariance between geography (waterfall barriers) and differences in selection (predation), making it difficult to disentangle their relative effects. Fortunately, multiple waterfalls are found on some tributaries, such that some waterfalls separate sites in high- vs. low-predation environments, whereas other waterfalls separate sites that are both in low-predation environments. In addition, predators in our study streams are sometimes absent from tributary sites not blocked by waterfalls, such that guppies can be found in adjacent low- and high-predation sites that are not separated by a major barrier. These types of site comparisons break the otherwise strong covariance between selection and

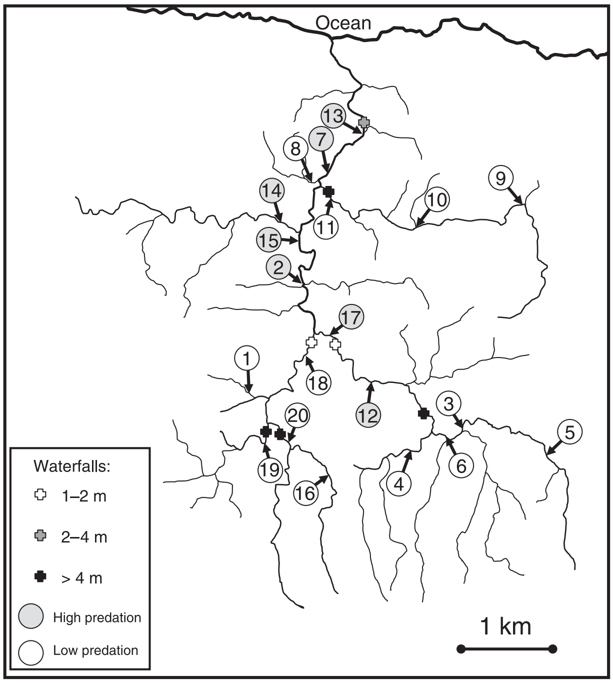

Fig. 1 Sites in the Marianne watershed from which guppies were sampled. Sites are indicated as high predation (shaded circles) or low predation (open circles). Site numbers reflect the chronological order of sampling. Waterfalls of different heights are indicated with crosses. Additional waterfalls occur in the watershed but these are not shown because they are not located between our sampling sites (i.e. they are located on tributaries we did not sample).

geography, allowing a determination of the relative effects of waterfalls and predation on gene flow.

Methods

We studied the Marianne River, located on the north slope of the Northern Range mountains of Trinidad (Fig. 1). The Marianne is characterized by substantial variation in several factors that might influence gene flow. First, a number of waterfalls are found on various tributaries (Fig. 1). Second, predation intensity varies spatially, with high-predation sites generally found in the main stem of the river and low-predation sites generally found in the tributaries (Fig. 1). Third, physical habitat features, such as canopy openness and stream size, also vary spatially.

Genetic data

In March 2002, we collected 20 male and 20 female guppies from each of 20 sites (Fig. 1). Minor exceptions to these sample sizes included sites 11, 13, and 18 (19 males each) and site 8 (19 females and 18 males). We killed the fish with an overdose of tricaine methanesulphonate and then preserved them in 95% ethanol. We used DNeasy Tissue Kits (QIAGEN) to extract genomic DNA from the tail fins of all sampled fish ($n = 794$). For genotyping, we used two previously published dinucleotide microsatellite loci (*Pre21* and *Pre92*; Becher *et al*. 2002) and five new tetranucleotide microsatellite loci (*Pre7*, *Pre8*, *Pre13*, *Pre15*, and *Pre18*; Paterson *et al*. 2005). Specific primers and polymerase chain reaction (PCR) conditions are given in Becher *et al*. (2002) and Paterson *et al*. (2005). Site-specific heterozygosities and total allele numbers are shown in Table 1. Locus-specific heterozygosities and allele numbers are provided in Crispo (2005; Appendix 2).

We used GENEPOP (version 3.4; Raymond & Rousset 1995) to test for site-specific departures from Hardy–Weinberg equilibrium (HWE; 100 batches and 1000 iterations per batch) and linkage equilibrium (10 000 batches and 10 000 iterations per batch). The only locus that greatly deviated from HWE was *Pre7*, which showed significant ($P < 0.05$) heterozygote deficits at 9 of the 17 sites for which the test was computable ($L < 0.001$ using the binomial likelihood function; Chapman *et al*. 1999). Owing to the possibility of null alleles at this locus, all subsequent analyses were performed with and without *Pre7*. These two sets of analyses yielded nearly identical results, and we take the conservative approach of reporting only those without *Pre7*. An exception was the partial Mantel tests (see below), for which we report results from two tests including *Pre7*. We also calculated null allele frequencies at this locus for the sites (1, 2, 7, 14–19) at which null alleles were suspected, and we then adjusted genotypes accordingly, using the correction algorithm of van Oosterhout (MICROCHECKER version 2.2.3; van Oosterhout *et al*. 2004). The adjusted genotypes for *Pre7* were included in a partial Mantel test (see below). The only hint of linkage disequilibrium was between *Pre15* and *Pre18* ($P < 0.05$ for 7 of 18 sites; $L < 0.001$ using the binomial likelihood function; Chapman *et al*. 1999). These loci are not strongly linked, however, or linkage would have been observed at more sites, and we therefore include both loci in our analyses. Summaries of site- and locus-specific deviations from HWE and linkage equilibrium are provided in Crispo (2005; Appendices 3 & 4).

Geographic features, predation, habitat

We recorded the presence and approximate height of all waterfalls located between our sampling sites (Fig. 1). We considered waterfalls to be potential barriers to gene flow only if they were > 2 m in height, because very small waterfalls were likely absent during parts of the wet season (D. N. Reznick, personal observation). We used maps to measure geographic distances along the length of the stream between sites (Trinidad and Tobago Government, 1978, Edition 2, Sheets 4 and 14, 1:25 000). We categorized the predation regime at each site as 'high' or

Table 1 Summary values for each collection site: expected heterozygosity based on seven loci (unbiased; TFPGA), number of alleles summed over all seven loci, geographic distance from the ocean along the stream, and site means for habitat variables

Site	Heterozygosity	Number of alleles	Distance from ocean (km)	% canopy openness	Water velocity (m/s)	Substrate size (mm)	Stream width (cm)	Stream depth (cm)
1	0.61	57	6.29	7.07	0.15	34.5	169.9	5.6
2	0.69	70	4.23	18.50	0.23	20.8	613.3	25.3
3	0.49	33	7.98	10.53	0.00	38.9	122.9	17.7
4	0.26	25	7.86	79.01	0.08	7.5	66.5	6.1
5	0.33	20	9.48	52.30	0.03	6.6	160.4	4.3
6	0.47	37	7.73	24.83	0.21	33.4	380.0	11.8
7	0.69	70	2.22	22.62	0.09	13.5	1037.9	41.6
8	0.56	48	2.47	13.21	0.08	22.8	164.0	11.3
9	0.46	48	5.44	16.24	0.22	38.6	325.8	9.1
10	0.44	47	3.94	11.04	0.12	21.1	394.0	14.2
11	0.46	51	2.67	15.62	0.11	18.9	317.6	20.3
12	0.40	36	6.26	32.43	0.22	37.6	490.5	14.9
13	0.70	67	1.44	35.07	0.12	20.5	995.1	61.1
14	0.63	70	3.43	22.26	0.05	22.3	347.5	22.8
15	0.68	72	3.30	25.38	0.16	23.4	855.2	23.9
16	0.52	37	7.66	20.12	0.17	19.8	115.8	6.3
17	0.69	69	5.33	28.24	0.17	22.5	447.3	28.1
18	0.66	63	5.73	16.48	0.22	36.6	396.7	18.9
19	0.37	37	7.36	n/a	n/a	n/a	n/a	n/a
20	0.46	37	6.91	10.50	0.19	35.1	315.0	15.9

'low' based on our own visual surveys (February–April, 2002–2005) and on previous surveys (Endler 1983; Reznick & Bryga 1996; Reznick *et al*. 1996b). In the Marianne and other north-slope streams of Trinidad, high-predation sites generally contain one or more species of goby (*Dormitator maculatus*, *Eleotris pisonis*, and *Gobiomorus dormitor*) and the mountain mullet (*Agonostomus monticola*). The presence of any of these species indicates accessibility of the site to predatory fishes, and led to our designation of the site as 'high-predation.' Both high- and low-predation sites also contain killifish (*Rivulus hartii*) and freshwater prawns (*Macrobrachium crenulatum* and *Macrobrachium faustinum*) (Liley & Seghers 1975; Reznick & Bryga 1996; Reznick *et al*. 1996b).

Near the end of the dry season in March 2003, we measured habitat features that might influence gene flow and divergent selection (see Introduction). We measured these features at all sites, except for site 19, which was excluded owing to limited accessibility and time constraints. Stream width was measured as the wetted width of the channel at 6–11 transects spaced every 2.5–10 m along the stream. Spacing between transects was constant within each site, but varied among sites so as to reflect the area from which guppies were collected. Stream depth, water velocity, and substrate size were then measured at three equidistant points along each transect (i.e. 18–33 points per site). Water velocity was measured using a flow meter (Swoffer model 2100) with the impeller positioned approximately 60% of the distance from the substrate to the surface. Substrate size was measured by randomly selecting a rock at each point and measuring its median diameter (modified from Kondolf & Li 1992). We ignored rocks larger than 200 mm in diameter, because these are probably too large to influence the evolution of male colour spot size (Endler 1978, 1980). If the substrate was not rock, we excluded the observation from the analyses. Canopy openness was quantified with a concave spherical densiometer (Lemmon 1957), which yields estimates comparable to those obtained using hemispherical photography (Englund *et al*. 1999). Four densiometer readings, one facing each cardinal direction, were taken in the middle of the stream at five equidistant points between the first and last transect at each site. Site means were calculated for each habitat feature (Table 1), and these were used in all subsequent analyses (for details see Millar *et al*. in press).

Site means for several habitat features were highly correlated with each other, specifically stream depth and width ($r = 0.831$, $P < 0.001$) and percent canopy openness and substrate size ($r = -0.682$, $P = 0.003$). We therefore used principle components (SPSS version 11.0.1) based on site means to develop composite variables for each of these two pairs of habitat features. These principal components explained 90.52% (depth and width) and 80.22% (canopy and substrate) of the variation. Subsequent analyses used these two sets of site-specific principal component scores rather than the original four habitat variables.

Data analysis

Our first step was to characterize guppy population structure. We used STRUCTURE (version 2.1; Pritchard *et al*.

2000) to group genotypes into clusters according to HWE and linkage equilibrium. Three iterations were run for each value of k (the putative number of clusters) using the admixture model, with burn-in and Monte Carlo Markov chain values of 100 000 each, and values of $k = 1$ through 30. We also used ARLEQUIN (version 1.0; Schneider *et al*. 1997) to partition the total molecular variance among different hierarchical groups (AMOVAS). Three separate AMOVAS were run, each grouping sites in different ways. The first ('waterfalls only') grouped sites together that were not separated by waterfalls (four groups). The second ('predation only') grouped sites together that had the same predation regime (two groups). The third ('adjusted predation') grouped sites together that were in the same predation regime and were not separated by a different predation regime (four groups). The latter grouping split low-predation sites into three different sets (sites 3–6; sites 9–11; and sites 1, 16, and 18–20) that were separated by high-predation environments, acknowledging that dispersers moving through a different selective regime might suffer decreased fitness. Groups with only one site could not be analysed, and so sites 19 and 8 were excluded from the 'waterfalls only' and 'adjusted predation' comparisons, respectively.

Our second step was to examine how within-site genetic variation was related to predation, waterfalls, and geographic distance from the ocean. We used TOOLS FOR POPULATION GENETIC ANALYSIS (TFPGA; version 1.3; Miller 1997) to calculate site-specific numbers of alleles and heterozygosities (unbiased) across loci (Table 1). We then used ANCOVAS (SPSS) to examine how site-specific number of alleles or heterozygosity (dependant variables) were related to distance from the ocean (covariate) and the presence or absence of waterfalls separating sites from the main stem (fixed factor). These ANCOVAS were then repeated with the presence or absence of predators (fixed factor) replacing the presence or absence of waterfalls.

Our third step was to quantify genetic divergence, long-term gene flow, and contemporary dispersal. We estimated global genotypic differentiation (GENEPOP) and F_{ST} (Weir & Cockerham 1984) and R_{ST} (Rousset 1996) (FSTAT version 2.9.3; Goudet 2001), as well as pairwise F_{ST} (Weir & Cockerham 1984) and R_{ST} (Michalakis & Excoffier 1996) (GENEPOP). We then evaluated the relative suitability of F_{ST} vs. R_{ST} through allele size permutation tests (Hardy *et al*. 2003) implemented in SPAGEDI (version 1.1; 20 000 permutations; Hardy & Vekemans 2002). We did not estimate gene flow ($N_e m$) from F_{ST} because the necessary assumptions (Whitlock & McCauley 1999) were violated by our study system. Instead, we estimated gene flow using the maximum-likelihood coalescent approach implemented in MIGRATE (version 1.6.9; microsatellite model, threshold value of 10, default values for all other parameters; Beerli & Felsenstein 1999, 2001). Our sample set was too large for MIGRATE to estimate gene flow when specifying a matrix of migration between adjacent sites. We therefore estimated gene flow independently between all possible adjacent upstream–downstream site pairs. We estimated contemporary dispersal using GENECLASS (version 2.0; Piry *et al*. 2004) to detect 'first-generation immigrants'. We here used the L_home likelihood computation, the Bayesian (Rannala & Mountain 1997) method of classification, and the Monte Carlo simulation algorithm (Paetkau *et al*. 2004) with 10 000 simulated individuals and an α value of 0.05. All sites were included in a single analysis, from which we identified immigrants at our sampling sites and the site from which each immigrant had the highest likelihood of dispersing from. If an individual was detected as an immigrant, but the sampling site was chosen as the most likely site of origin, we excluded it from further analyses.

Our fourth step was to evaluate factors influencing genetic divergence, long-term gene flow, and contemporary dispersal. We here used partial Mantel tests to compare matrices of pairwise differences between sites (Mantel 1967; Smouse *et al*. 1986; 20 000 randomizations; implemented in FSTAT). Predictor matrices were based on habitat features (water velocity, stream depth/width PC, canopy openness/substrate size PC), predation regime (0 = same regime, 1 = different regimes), geographic distance along the stream, and the number of waterfalls separating sites (Fig. 1). Paired differences in geographic distance, water velocity, and principle component scores were $\log_{10}$ transformed to improve normality and homoscedasticity. Response matrices were based on pairwise F_{ST} and R_{ST} (with and without *Pre7*, and with adjusted genotypes for *Pre7*), or on the number of first-generation immigrants (i.e. 'dispersers'; with and without *Pre7*; GENECLASS). We also evaluated Rousset's (1997) linearized F_{ST} [i.e. $F_{ST}/(1 - F_{ST})$], but this transformation proved less suitable because F_{ST} values were already normally distributed and gave stronger correlations with distance than did the transformed values. We pooled the number of dispersers (from GENECLASS), rather than analysing upstream and downstream dispersal separately, because there exists a greater potential for individuals to be classified as upstream dispersers than as downstream dispersers (see Results).

Nuances of our study system, and results of the above partial Mantel tests, precipitated additional analyses. First, we performed tests where predation regime was replaced by 'adjusted predation' (as described above). Second, although we normally only considered waterfalls > 2 m high, we also performed partial Mantel tests including waterfalls in the 1–2 m category (Fig. 1). Third, because waterfalls had a dominant influence on genetic divergence and gene flow (see Results), we also performed partial Mantel tests on two different subsets of sites: pairs separated by waterfalls and pairs not separated by waterfalls. This allowed us to determine the effects of habitat features, predation, and distance, independent of the effects of waterfalls. Fourth, predation had a strong influence on contemporary dispersal (see Results), and so we also performed these partial Mantel tests on two different subsets of sites: pairs experiencing the same predation and pairs experiencing different predation.

Partial Mantel tests may inflate type I errors relative to non-partial (i.e. one predictor and one response matrix) Mantel tests (Raufaste & Rousset 2001; Rousset 2002; but see Castellano & Balletto 2002). We therefore also used non-partial Mantel tests to confirm the influence of waterfalls on F_{ST} and dispersal. We did not conduct non-partial tests for any of the other predictor variables because of the overwhelming influence of waterfalls (see Results). Also, strongly correlated predictor matrices may inflate type II errors (Castellano & Balletto 2002). Some of the predictor variables in the partial Mantel tests were correlated with each other: waterfalls and predation ($r = -0.185$, $P = 0.011$), waterfalls and distance ($r = 0.570$, $P < 0.001$), predation and width/depth PC ($r = 0.339$, $P < 0.001$), distance and canopy openness/substrate size PC ($r = 0.227$, $P = 0.003$), and width/depth PC and canopy openness/substrate size PC ($r = 0.284$, $P < 0.001$). To ensure that these correlations were not influencing our results, we also conducted partial Mantel tests on all possible groups of uncorrelated variables. These included (i) waterfalls, water velocity, and width/depth PC, (ii) predation, distance, and water velocity, (iii) waterfalls, water velocity, and canopy openness/substrate size PC, (iv) predation, distance, and canopy openness/substrate size PC, and (v) distance, width/depth PC, and water velocity.

Table 2 Results from STRUCTURE based on the most likely number of clusters ($k = 7$, as obtained from the average of three iterations)

	Inferred cluster						
Site	1	2	3	4	5	6	7
1	0.019	0.005	0.257	**0.686**	0.007	0.014	0.012
2	**0.645**	0.026	0.017	0.063	0.018	0.200	0.031
3	0.005	**0.926**	0.005	0.006	0.038	0.005	0.015
4	0.005	0.019	0.006	0.003	**0.952**	0.006	0.009
5	0.006	**0.930**	0.004	0.003	0.039	0.006	0.012
6	0.007	**0.789**	0.014	0.005	0.155	0.008	0.022
7	0.364	0.021	0.009	0.013	0.022	**0.540**	0.031
8	0.102	0.007	0.012	0.022	0.007	**0.825**	0.025
9	0.011	0.011	0.006	0.005	0.020	0.011	**0.936**
10	0.013	0.010	0.005	0.005	0.013	0.010	**0.944**
11	0.012	0.014	0.008	0.007	0.012	0.013	**0.933**
12	0.006	0.124	0.007	0.003	**0.842**	0.006	0.012
13	0.293	0.015	0.014	0.014	0.013	**0.580**	0.071
14	0.442	0.008	0.010	0.022	0.008	**0.476**	0.034
15	0.415	0.011	0.047	0.068	0.016	**0.418**	0.024
16	0.010	0.014	**0.928**	0.006	0.022	0.010	0.009
17	**0.692**	0.019	0.041	0.027	0.013	0.169	0.038
18	0.069	0.009	0.334	**0.542**	0.010	0.025	0.011
19	0.008	0.003	0.006	**0.965**	0.004	0.011	0.004
20	0.008	0.015	**0.930**	0.013	0.018	0.006	0.010

Proportions shown are from the iteration with the highest value of ln P [$k = 7$]. Proportions of genotypes from each sampling site within each of the seven inferred clusters are shown. Bold values indicate the cluster with the highest proportion of individuals from each site.

Results

Genetic structure

The most probable number of genetic clusters, averaged over three iterations using STRUCTURE, was seven [ln P($k = 7$) = −13 503.4, −13 493.8, and −13 502.3], with the next most probable being six [ln P($k = 6$) = −13 756.2, −13 771.3, and −13 619.7] or eight [ln P($k = 8$) = −13 757.5, −13 805.1, and −13 454.5]. The degree to which guppies from a given site were assigned together to a given cluster varied widely. In discussing these clusters, we refer to the assignment of 'guppies', but note that the admixture model can divide an individual's assignment between multiple sites (Pritchard *et al*. 2000). Guppies from tributaries isolated from the main stem by waterfalls usually formed coherent clusters that included guppies from the sites in that tributary (Table 2). In particular, 79–93% of guppies from sites 3, 5, and 6 were clustered together, 93% from sites 16 and 20 were clustered together, and 93–94% from sites 9, 10, and 11 were clustered together. Guppies from sites above and below waterfalls in the same tributary also tended to cluster together and separately from other sites. In particular, 84–95% from sites 4 and 12 were clustered together, 97% from site 19 were clustered with individuals from sites 1 and 18, and 54–69% from sites 1 and 18 were clustered with individuals from site 19. In contrast, guppies from the main stem (sites 2, 7, 13, 15, 17) and tributaries not separated by waterfalls (sites 8 and 14) were assigned much more diffusely to several different clusters that spanned larger geographic distances (Table 2). In fact, only 42–69% of the guppies from any one of these sites were clustered together. These general patterns held across all three iterations, and similar trends were observed for $k = 6$ and $k = 8$. In short, waterfalls played a strong role in isolating tributary populations, whereas sites below waterfalls in the main stem comprised widely spread and more diffuse groupings.

AMOVAS complemented the results from STRUCTURE in revealing that the proportion of genetic variation was much greater among sites separated by waterfalls than among sites not separated by waterfalls (Table 3). In contrast, the proportion of genetic variation was much lower among sites from different predation regimes than among sites within predation regimes, presumably because the 'low predation' grouping included sites from several

Table 3 Results from AMOVAS (ARLEQUIN)

Source of variation	Percentage of variation	*P* value
Waterfalls only		
Among groups	22.50	< 0.001
Within groups, among sites	11.41	< 0.001
Within sites	66.09	< 0.001
Predation only		
Among groups	3.20	0.123
Within groups, among sites	29.91	< 0.001
Within sites	66.89	< 0.001
Adjusted predation		
Among groups	22.54	< 0.001
Within groups, among sites	12.45	< 0.001
Within sites	65.01	< 0.001

Sites for the 'waterfalls only' comparison were grouped so that sites within each grouping were not separated by a waterfall (four groups). Sites for the 'predation only' comparison were grouped according to predation (high-predation and low-predation; two groups). Sites for the 'adjusted predation' comparison were grouped so that high-predation sites were grouped together and low-predation sites were divided into groups that were not separated by high-predation environments (four groups). Groups with only one site could not be analysed, and sites 19 and 8 were therefore excluded from the 'waterfalls only' and 'adjusted predation' comparisons, respectively.

isolated tributaries. When low-predation sites were divided such that grouped sites were not separated by high-predation environments (i.e. 'adjusted predation'), the proportion of genetic variation among groups was similar to that based on separation by waterfalls (Table 3). This similarity was not surprising because 'adjusted predation' groupings were largely defined by waterfalls.

Within-site genetic variation

In ANCOVAS, distance from the ocean was negatively correlated with the total number of alleles ($P = 0.004$) but not with heterozygosity ($P = 0.115$), sites above waterfalls had fewer total alleles ($P = 0.005$) and lower heterozygosity ($P = 0.003$), and distance and waterfalls did not interact in their influence on total alleles ($P = 0.477$) or heterozygosity ($P = 0.642$). Considering predation instead of waterfalls, distance from the ocean was negatively correlated with the total number of alleles ($P = 0.002$) and with heterozygosity ($P = 0.041$), sites with predators had more alleles ($P = 0.026$) and marginally higher heterozygosity ($P = 0.067$), and distance and predation did not interact in their influence on total alleles ($P = 0.770$) or heterozygosity ($P = 0.425$). These patterns suggest some combination of (i) founder effects associated with upstream colonization, and (ii) downstream-biased gene flow, particularly across waterfalls and predation regimes.

Genetic divergence, gene flow, dispersal

Considering all sites (i.e. globally), the Marianne was characterized by significant genotypic differentiation ($P < 0.001$) and high genetic divergence ($F_{ST} = 0.302$, 95% CI = 0.164–0.485; weighted $R_{ST} = 0.408$). Considering site pairs, F_{ST} and R_{ST} were highly correlated across all pairs ($r = 0.763$, $P < 0.001$), across pairs not separated by waterfalls ($r = 0.914$, $P < 0.001$), and across pairs separated by waterfalls ($r = 0.584$, $P < 0.001$). As well, SPAGEDI revealed that actual R_{ST} was not significantly greater than R_{ST} estimated after allele size permutation (i.e. pR_{ST}; across all sites and loci; $P = 0.192$). When individual loci were tested (across all sites), R_{ST} was greater than pR_{ST} for only *Pre13* ($P = 0.003$). When site pairs were considered individually (across all loci), R_{ST} was significantly greater than pR_{ST} for only 24 of 190 pairs (none after implementing the sequential Bonferroni technique; Rice 1989). F_{ST} and R_{ST} thus yielded similar information, but F_{ST} proved to be more appropriate for subsequent analyses: *Pre18* manifested some non-standard alleles that appeared to increase in increments of two rather than the expected four base pairs (Crispo 2005), inconsistent with the stepwise-mutation model for which R_{ST} is designed. We therefore restrict our presentation of results to those based on F_{ST}.

We base our interpretation of MIGRATE results on visual trends, while noting that statistical inference cannot be made owing to the non-independence of pairs of sites. MIGRATE revealed that gene flow across waterfalls was usually lower in the upstream direction than in the downstream direction (Fig. 2). For the five pairs of adjacent sites separated by waterfalls, only one showed (marginally) higher gene flow in the upstream direction (sites 6 and 12). A similar downstream bias in gene flow was evident for 10 of the 14 site pairs not separated by waterfalls. Differences in predation did not have an obvious effect on overall gene flow, nor on the directional bias in gene flow (Fig. 2). Although shifts in predation were highly correlated with the presence of waterfalls (Fig. 1), occasional decoupling of these two effects confirmed that predation did not reduce gene flow. First, when considering only sites not separated by waterfalls, gene flow between sites in different predation regimes fell within the range of gene flow between sites in the same regime (Fig. 2). Second, when considering only sites separated by waterfalls, gene flow between sites in the same predation regime fell within the range of gene flow between sites in different regimes (Fig. 2).

Contemporary dispersal (estimated by GENECLASS) differed from long-term gene flow (estimated by MIGRATE) in that a higher number of fish appeared to disperse upstream than downstream, including over waterfalls (Table 4). However, this upstream vs. downstream comparison is biased, because the number of downstream sites to which immigrants could be randomly assigned was

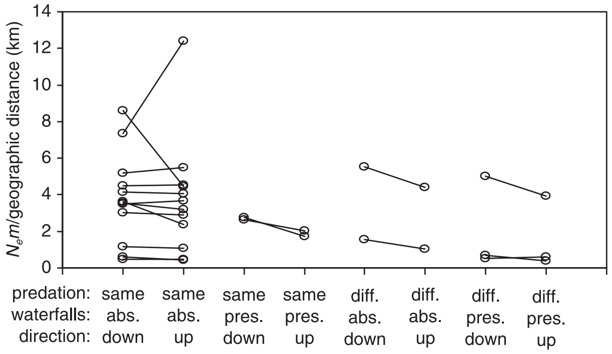

Fig. 2 Directional gene flow between adjacent pairs of sites (estimated by MIGRATE), divided by the geographic distance between sites. (We divided by geographic distance to remove the confounding effects of distance between sites). Site pairs are categorized by whether or not the two sites experience the same ('same') or different ('diff.') predation regimes and by whether waterfalls between them are present ('pres.') or absent ('abs.'). The upstream ('up') and downstream ('down') estimates for each pair are connected by a line.

Table 4 Contemporary dispersal quantified as the number first-generation immigrants (GENECLASS)

Sampling site	Dispersal up	Dispersal down	Dispersal across waterfalls up	Dispersal across waterfalls down	Donor sites
1	4				**2**,18
2	3				7,15
3	2				6
4	3		1		6,**12**
5	1				6
6		2			3
7	1	6			2,13,15,17
8		1			**17**
9	2				10
10	2	1			9,11
11		4			9,10
12		2		2	**3,6**
13		4			2,15,17
14	3	1			2,13,15
15	1	1			13,14
16	3		1		18,20
17	2				**8**,15
18	2				**7**,**17**
20	1	2	1		16,18
Total	30	24	3	2	

Shown are the estimated numbers of immigrants that dispersed primarily upstream or primarily downstream of the sampling site, and the number of immigrants that dispersed upstream or downstream over a waterfall. Includes the sites that individuals dispersed from (i.e. donor sites). Donor sites that are in a different predation regime than the sampling site are indicated in bold.

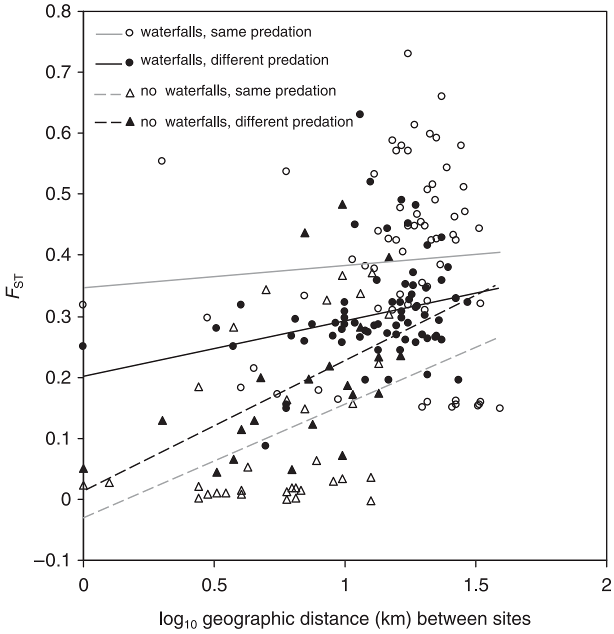

Fig. 3 Pairwise F_{ST} values between all possible site pairs in relation to $\log_{10}$ transformed geographic distance (km) between sites along the length of the stream. Circles and solid lines represent site pairs separated by one or more waterfalls, triangles and dashed lines represent site pairs not separated by a waterfall. Open points and grey lines represent site pairs experiencing the same predation regime; closed points and black lines represent site pairs experiencing different predation regimes.

greater than the number of upstream sites. For example, only one downstream site (13) had no sampled sites further downstream, whereas eight upstream sites (1, 4, 5, 8, 9, 14, 16, and 19) had no sampled sites further upstream. The potential for an immigrant to be assigned to a downstream site was therefore much greater than to an upstream site. We thus do not attempt to further interpret this directional pattern of dispersal, and all analyses are based on pooled upstream and downstream dispersers between sites. It is also important to note that results based on the identification of immigrants should be interpreted with caution. First, as genetic differentiation decreases, the ability to correctly identify immigrants decreases, and we observed low F_{ST} values for several site pairs (Fig. 3; Manel *et al.* 2005). Second, several putative immigrants had *smaller* –log(L) values for their sampling site than did putative residents from those same sites. Third, several identified immigrants had multiple sites with –log(L) values that were lower than that for the sampling site, indicting that an immigrant could potentially have arisen from one of several sites [we chose the site with the lowest –log(L)].

Partial Mantel tests explained up to 60% of the variance in pairwise F_{ST} (Table 5). The strongest effect was that F_{ST} increased with the number of waterfalls between sites

Table 5 Correlation coefficients from partial Mantel tests (FSTAT) comparing matrices of pairwise F_{ST} or the number of first-generation immigrants (i.e. 'dispersers'; GENECLASS) to matrices of predation (sites in the same or a different predation regime), waterfalls (number of waterfalls separating sites), $\log_{10}$ transformed geographic distance along the stream, and $\log_{10}$ transformed differences in habitat variables or habitat variable principle components

	n	r^2	Waterfall	Distance	Predation	Canopy openness/ substrate size	Stream width/depth	Water velocity
F_{ST}								
All data points	190	0.422	0.606*	0.169	0.056	0.120	−0.012	0.092
No waterfalls	54	0.291	n/a	0.419*	0.276	0.116	−0.137	−0.085
Waterfalls only	136	0.201	n/a	0.126	−0.333*	0.121	0.099	0.224*
Pre7 included	190	0.392	0.565*	0.207*	0.111	0.101	−0.001	0.088
Adj. genotypes	190	0.394	0.569*	0.209*	0.101	0.099	−0.001	0.086
Adj. predation	190	0.438	0.606*	0.128	0.169	0.116	−0.008	0.113
Extra waterfalls	190	0.599	0.764*	−0.049	0.056	0.063	0.049	0.061
Dispersal								
All data points	190	0.314	−0.422*	−0.175*	−0.318*	0.001	0.040	0.051
No waterfalls	54	0.264	n/a	−0.303*	−0.384*	0.095	0.068	0.106
Waterfalls only	136	0.072	n/a	−0.261*	−0.011	−0.047	−0.010	−0.045
Same predation	99	0.330	−0.528*	−0.218*	n/a	0.057	0.023	−0.001
Diff. predation	91	0.100	−0.257	−0.139	n/a	−0.015	0.010	0.117
Pre7 included	190	0.401	−0.435*	−0.360*	−0.285*	−0.015	0.005	0.026
Adj. predation	190	0.318	−0.422*	−0.210	−0.309*	0.007	−0.017	0.017
Extra waterfalls	190	0.341	−0.484*	−0.114	−0.297*	0.044	0.014	0.048

Results are shown for all data points, for only sites not separated by waterfalls, for only sites separated by waterfalls, for data with locus *Pre7* included, for 'adjusted genotypes' for *Pre7* (note that GENECLASS could not estimate dispersal using adjusted genotypes), for 'adjusted predation' (see text), and for waterfalls < 2 m high included (i.e. 'extra waterfalls'). In addition, results for tests on dispersal are also shown for only site pairs experiencing the same predation regime, and only site pairs experiencing different predation regimes. Also shown are the number of site pairs (n) and the amount of variance explained by the models (r^2).
$*P < 0.05$.

(Fig. 3), and this effect was found in every test that included variation in waterfalls between sites (Table 5). The effect of waterfalls was also highly significant ($r = 0.606$, $P < 0.001$) in the non-partial test excluding all factors other than waterfalls. F_{ST} also increased with increasing geographic distance between sites when sites were not separated by waterfalls (Fig. 3; Table 5), possibly due to the otherwise overwhelming effect of waterfalls. Distance was also a significant variable when including *Pre7 (*both with and without the adjusted genotypes), possibly due to increased power. In contrast, F_{ST} was not higher between sites with different predation regimes. In fact, the only significant result for predation was a *decrease* in F_{ST} when considering only sites separated by waterfalls (Fig. 3; Table 5). Differences in habitat features also had little effect. The only significant trend was a slight increase in F_{ST} with increasing differences in water flow, but only for sites separated by waterfalls (Table 5). These results were maintained in partial Mantel tests incorporating only uncorrelated variables: F_{ST} increased with the number of waterfalls separating sites ($P < 0.001$ for both tests) and with increasing geographic distance between sites ($P < 0.001$ for all three tests). Predation and habitat variables were not significant in the tests on uncorrelated variables after implementing the sequential Bonferroni technique (Rice 1989).

Partial Mantel tests explained considerably less of the variance in dispersal than in F_{ST} (Table 5), perhaps because only a small proportion of fish were immigrants (Table 4). Waterfalls and distance between sites were important in reducing dispersal, but here differences in predation were also important (Table 5). The only nonsignificant influence of predation involved only sites separated by waterfalls, but the overall r^2 value for the test was low ($r^2 = 0.072$; Table 5), probably due to a low number of dispersers over waterfalls (Table 4). No habitat features appeared to influence dispersal. Non-partial Mantel tests, incorporating only the effect of waterfalls, again revealed their strong influence ($r = -0.422$, $P < 0.001$). Similar results to the above were obtained for partial Mantel tests incorporating only uncorrelated variables: dispersal was lower between predation regimes ($P < 0.001$ for both tests), decreased with the number of waterfalls ($P < 0.001$ for both tests) and distance ($P < 0.001$ for all three tests) between sites, and habitat variables were nonsignificant.

The earlier-noted association between waterfalls and shifts in predation regime was unlikely to hamper our analyses because a substantial number of site pairs broke this covariance. In particular, 66 of 190 pairwise comparisons, involving 13 sites, had the same predation regime but were separated by one or more waterfalls (Fig. 3; 73 comparisons involving all 20 sites if including waterfalls < 2 m high). Similarly, 21 of 190 comparisons, involving 10 sites, had different predation regimes but no waterfall separation (Fig. 3; six comparisons involving seven sites if including waterfalls < 2 m high). We can thus place confidence in our estimates of the effects of predation and waterfalls on genetic divergence and dispersal.

Discussion

Effects of geography

Geography had substantial effects on guppy genetic structure. Principal among these effects was that waterfalls substantially reduced gene flow. This conclusion was supported by (i) coherent genetic clusters above waterfalls (Tables 2 and 3), (ii) lower genetic diversity above waterfalls, (iii) less upstream than downstream gene flow across waterfalls (Fig. 2), and (iv) greater genetic divergence, lower gene flow, and lower dispersal across waterfalls (Figs 2 and 3; Tables 4 and 5). These results are intuitively satisfying given that upstream movement over waterfalls should be very difficult for guppies, perhaps accomplished only occasionally and by few individuals, such as during floods or when pregnant females are dropped by piscivorous birds. Gene flow across waterfalls should be much easier in the downstream direction, an expectation consistent with the higher genetic diversity below waterfalls and the downstream-biased estimates of long-term gene flow (Fig. 2). Instead, it was surprising that estimates of gene flow across waterfalls were not *more* downstream biased (Fig. 2). One factor increasing upstream gene flow may be that dispersers from downstream sites, even if rare, could have a high reproductive output when they reach upstream sites. This suggestion arises because high-predation guppies, which are typical of downstream sites, have an earlier age of reproduction and a greater reproductive effort than do their low-predation counterparts, which are typical of upstream sites (Reznick & Bryga 1987; Reznick *et al*. 2001, 2002). This suggestion remains hypothetical, however, because the lower resource availability at low-predation sites may place residents (adapted for the competitive conditions) at a selective advantage over immigrants (Grether *et al*. 2001a; Reznick *et al*. 2001, 2002; Bronikowski *et al*. 2002). One factor decreasing downstream gene flow may be that guppies dispersing from low-predation sites survive at low rates when they move to high-predation environments, owing to their lack of antipredator adaptations (see Introduction). A combination of these two scenarios may explain why the downstream bias in gene flow across waterfalls is not stronger than might be expected.

Geography also influenced guppy genetic structure through the isolating effects of geographic distance (i.e. IBD). This effect of distance was weaker than that of waterfalls, and was most evident when considering sites not separated by waterfalls (Fig. 3; Table 5). This observed signature of IBD nevertheless provides some indication that guppies in the Marianne are at, or are approaching, genetic equilibrium (Slatkin 1993). Similar results were observed in bullfrogs (*Rana catesbeiana*), in which IBD was only evident at small spatial scales (Austin *et al*. 2004), and in brook charr (*Salvelinus fontinalis*), in which IBD was strongest when considering sites not historically separated by barriers (Poissant *et al*. 2005). The proximate effects of distance on gene flow probably relate to dispersal distances. Indeed, we also found that increasing geographic distance generally reduced contemporary dispersal, as identified by genetic assignment tests (Table 5). These results are further consistent with evidence of decreasing dispersal with increasing distance, as estimated by mark–recapture experiments in the Marianne (D. Weese, M. T. Kinnison, & A. P. Hendry, unpublished). We conclude that the distance between sites does indeed influence dispersal and gene flow, but that these effects become difficult to detect when gene flow is restricted by waterfalls.

The direction of water flow is another physical factor that might influence guppy dispersal, even in the absence of waterfalls. It seems intuitive that water flow would bias the movement of guppies in the downstream direction, since upstream dispersal should be more difficult. However, we found relatively minor effects of the direction of gene flow when not considering waterfalls (Fig. 2). Moreover, mark–recapture experiments at various locations in the Marianne have found no evidence for downstream-biased dispersal (D. Weese, M. T. Kinnison, & A. P. Hendry, unpublished). The direction of water flow seems to be of only minor importance in our system, implying a tendency for guppies to hold their position in a given location, or perhaps to move upstream.

Effects of natural selection

We found no evidence that divergent natural selection increased genetic divergence or reduced long-term gene flow (Figs 2 and 3; Table 5), a result that conflicts with the standard predictions of ecological speciation (e.g. Lu & Bernatchez 1999; Ogden & Thorpe 2002; Rocha *et al*. 2005; reviews: Schluter 2000; Rundle & Nosil 2005). Yet, we might have made these predictions for guppies because (i) males moving from low-predation to high-predation sites should have greater predator-induced mortality (Endler 1978, 1980; Godin & McDonough 2003), and (ii) males moving from high-predation sites to low-predation sites should be less likely to procure mates owing to their reduced coloration (review: Houde 1997). Instead, our results are congruent with observations that neutral gene flow can be high even in the presence of substantial adaptive divergence (e.g. Smith *et al*. 2001; Emelianov *et al*. 2004; Hendry & Taylor 2004; Smith *et al*. 2005). Guppies thus provide an interesting test case for how the standard predictions of ecological speciation can be heavily nuanced by the specifics of natural systems.

Sexual selection may be one reason why shifts in predation did not reduce gene flow in guppies. First, low-predation males that disperse into high-predation sites are more colourful than residents, and may therefore be favoured by females (review: Houde 1997). Second, males can adopt 'sneaky mating' tactics that undermine female choice (review: Magurran 1998). High-predation males are more likely to use these tactics, which may increase their mating success in low-predation environments (Endler 1987; Luyten & Liley 1991). Third, females sometimes prefer 'novel' male phenotypes (Hughes *et al*. 1999; Eakley & Houde 2004), which might increase the success of dispersers relative to residents. Fourth, low-predation females are large (Reznick & Bryga 1987; Endler 1995) and may therefore be favoured by males at high-predation sites (Herdman *et al*. 2004). Thus, although dispersers from low- to high-predation sites may be at a selective disadvantage with respect to predation, they may have increased reproductive success due to mate choice. This hypothesis is consistent with our observation (Table 5) that differences in predation reduce contemporary dispersal (i.e. selection against migrants; Hendry 2004; Nosil *et al*. 2005), but do not reduce long-term gene flow. Thus, these two aspects of selection (natural and sexual) may act in opposition to one another.

Natural selection may, paradoxically, be another reason why shifts in predation did not reduce gene flow. Specifically, high-predation females that colonize low-predation sites may have a reproductive advantage owing to their earlier maturity and more rapid production of offspring (Reznick & Bryga 1987; Reznick *et al*. 2001, 2002). Indeed, a similar phenomenon has been observed in snow buttercups (*Ranunculus adoneus*), where gene flow was higher between different selective regimes due to the movement of higher quality seeds into an area occupied by lower quality seeds (Stanton *et al*. 1997). (Of course, guppies dispersing upstream may be at a selective disadvantage, as upstream residents may be better adapted to the competitive environments of upstream, low-predation habitats, as noted above.) In addition, heterosis may increase the fitness of immigrants (Ingvarsson & Whitlock 2000; Ebert *et al*. 2002) because inbreeding has negative effects on guppy mating ability (van Oosterhout *et al*. 2003) and survival (Shikano & Taniguchi 2002). Thus, although upstream dispersal may be low, immigrants could potentially have greater reproductive success than residents, thereby increasing gene flow relative to dispersal.

When considering selection on guppies, we have emphasized the role of predation, which is undoubtedly predominant (reviews: Endler 1995; Houde 1997; Magurran 1998), but other selective factors can also influence guppy evolution. These factors include physical habitat features such as canopy openness, water velocity, and substrate size (Endler 1980; Kodric-Brown 1989; Nicoletto 1996; Nicoletto & Kodric-Brown 1999; Grether *et al*. 2001a). Despite the documented contributions of these features to guppy evolution, we also failed to find evidence that sites with greater differences in these habitat features experience lower gene flow, except perhaps for a very weak effect of water flow (Table 5). We recognize, however, that a definitive correlative test for ecological speciation would include more watersheds, more sites within watersheds, more variables that may contribute to selection, and a further decoupling of waterfalls from shifts in predation.

Conclusions

Our findings are consistent to varying degrees with previous population genetic work on guppies in other watersheds. First, Shaw *et al*. (1994) found that genetic variability decreased with increasing distance from the ocean, but was not influenced by predation. This same study, however, failed to find an influence of waterfalls. Second, artificial introductions have shown that gene flow is common, including from areas of low to high predation (Haskins *et al*. 1961; Shaw *et al*. 1992; Becher & Magurran 2000). Third, Carvalho *et al*. (1991) found that vicariance (physical isolation of watersheds) played a strong role in genetic differentiation, but failed to find evidence of IBD. Our use of hypervariable microsatellites, and recently developed analytical methods, allowed us to confirm and extend these findings. In particular, we conclude that north-slope guppy populations are characterized by low gene flow across waterfalls, moderate IBD, a moderate downstream gene flow bias, and little effect of divergent selection on gene flow.

Our study suggests that divergent natural selection is much less important than geography in determining patterns of genetic isolation in a natural system. This conclusion is likely robust because our study system was characterized by differences in selection that promote adaptive divergence, and because we focused on a single watershed (i.e. geography would be even more important across watersheds). We close by noting that a consideration of guppy biology suggests the intriguing possibility that divergent natural and sexual selection might both increase and decrease gene flow in a variety of interacting ways. The standard predictions of ecological speciation may thus obscure a more interesting and complicated dynamic between selection and gene flow. Further research should concentrate on these interactions, as well as the effect of gene flow on adaptation and mate choice evolution. Experimental manipulations of gene flow would be particularly useful for empirically evaluating the extent and mechanisms by which gene flow constrains adaptation and vice versa.

Acknowledgements

Michael A. Hendry, Meghan L. Kelly, Nathan P. Millar, Amy K. Schwartz, and Dylan Weese assisted with the field work. Kristy Brohman and Ian G. Paterson performed the genotyping. Special thanks to Scott A. Pavey for conducting the AMOVAs. Eldredge Bermingham, Felix Breden, and three anonymous reviewers provided insightful comments. Funding was provided by a Natural Sciences and Engineering Research Council of Canada Discovery Grant (Andrew P. Hendry), a National Science Foundation grant (Michael T. Kinnison and David N. Reznick), and the Maine Agricultural and Forest Experiment Station (Michael T. Kinnison).

References

Austin JD, Lougheed SC, Boag PT (2004) Controlling for the effects of history and nonequilibrium conditions in northern bullfrog (*Rana catesbeiana*) populations. *Genetics*, **168**, 1491–1506.

Becher SA, Magurran AE (2000) Gene flow in Trinidadian guppies. *Journal of Fish Biology*, **56**, 241–249.

Becher SA, Russell ST, Magurran AE (2002) Isolation and characterization of polymorphic microsatellites in the Trinidadian guppy (*Poecilia reticulata*). *Molecular Ecology Notes*, **2**, 456–458.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763–773.

Beerli P, Felsenstein J (2001) Maximum-likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA*, **98**, 4563–4568.

Breden F, Scott M, Michel E (1987) Genetic differentiation for antipredator behaviour in the Trinidadian guppy, *Poecilia reticulata*. *Animal Behaviour*, **35**, 618–620.

Bronikowski AM, Clark ME, Rodd FH, Reznick DN (2002) Population-dynamic consequences of predator-induced life history in the guppy (*Poecilia reticulata*). *Ecology*, **83**, 2194–2204.

Carvalho GR, Shaw PW, Magurran AE, Seghers BH (1991) Marked genetic divergence revealed by allozymes among populations of the guppy *Poecilia reticulata* (Poeciliidae), in Trinidad. *Biological Journal of the Linnean Society*, **42**, 389–405.

Castellano S, Balletto E (2002) Is the partial Mantel test inadequate? *Evolution*, **56**, 1871–1873.

Chapman RW, Sedberry GR, Koenig CC, Eleby BM (1999) Stock identification of gag, *Mycteroperca microlepis*, along the southeast coast of the United States. *Marine Biotechnology*, **1**, 137–146.

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

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

Crispo E (2005) *Factors influencing gene flow in guppies*. MSc Thesis, McGill University, Montréal, Québec.

Eakley AL, Houde AE (2004) Possible role of female discrimination against 'redundant' males in the evolution of colour pattern polymorphism in guppies. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, S299–S301.

Ebert D, Haag C, Kirkpatrick M, Riek M, Hottinger JW, Pajunen VI (2002) A selective advantage to immigrant genes in a *Daphnia* metapopulation. *Science*, **295**, 485–488.

Emelianov I, Marec F, Mallet J (2004) Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 97–105.

Endler JA (1978) A predator's view of animal color patterns. *Evolutionary Biology*, **11**, 319–364.

Endler JA (1980) Natural selection on color patterns in *Poecilia reticulata*. *Evolution*, **34**, 76–91.

Endler JA (1983) Natural and sexual selection on color patterns in poeciliid fishes. *Environmental Biology of Fishes*, **9**, 173–190.

Endler JA (1987) Predation, light intensity, and courtship behaviour in *Poecilia reticulata*. *Animal Behaviour*, **35**, 1376–1385.

Endler JA (1995) Multiple-trait coevolution and environmental gradients in guppies. *Trends in Ecology & Evolution*, **10**, 22–29.

Endler JA, Houde AE (1995) Geographic variation in female preference for male traits in *Poecilia reticulata*. *Evolution*, **49**, 456–468.

Englund SR, O'Brien JJ, Clark DB (1999) Evaluation of digital and film hemispherical photography and spherical densiometry for measuring forest light environments. *Canadian Journal of Forest Research*, **30**, 1999–2005.

Ghalambor CK, Reznick DN, Walker JA (2004) Constraints on adaptive evolution: the functional trade-off between reproduction and fast-start swimming performance in the Trinidadian guppy (*Poecilia reticulata*). *American Naturalist*, **164**, 38–50.

Godin J-GJ, McDonough HE (2003) Predator preference for brightly colored males in the guppy: a viability cost for a sexually selected trait. *Behavioral Ecology*, **14**, 194–200.

Goudet J (2001) *FSTAT (version 2.9.3): A program to estimate and test gene diversities and fixation indices*. Available from http://www.unil.ch/izea/softwares/fstat.html.

Grether GF, Millie DF, Bryant MJ, Reznick DN, Mayea W (2001a) Rain forest canopy cover, resource availability, and life history evolution in guppies. *Ecology*, **82**, 1546–1559.

Grether GF, Hudson J, Endler JA (2001b) Carotenoid scarcity, synthetic pteridine pigments and the evolution of sexual colouration in guppies (*Poecilia reticulata*). *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 1245–1253.

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

Hardy OJ, Charbonnel N, Fréville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Haskins CP, Haskins EF, McLaughlin JJA, Hewitt RE (1961) Polymorphism and population structure in *Lebistes reticulatus*, an ecological study. In: *Vertebrate Speciation* (ed. Blair WF), pp. 320–395. University of Texas Press, Austin, Texas.

Hendry AP (2004) Selection against migrants contributes to the rapid-evolution of reproductive isolation. *Evolutionary Ecology Research*, **6**, 1219–1236.

Hendry AP, Taylor EB (2004) How much of the variation in adaptive divergence can be explained by gene flow? An evaluation use lake-stream stickleback pairs. *Evolution*, **58**, 2319–2331.

Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of quantitative characters in discrete populations: a theoretical framework for empirical tests. *Evolution*, **55**, 459–466.

Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution*, **56**, 1199–1216.

Herdman EJE, Kelly CD, Godin J-GJ (2004) Male mate choice in the guppy (*Poecilia reticulata*): do males prefer larger females as mates? *Ethology*, **110**, 97–111.

Houde AE (1997) *Sex, Color and Mate Choice in Guppies*. Princeton University Press, Princeton, New Jersey.

Hughes KA, Du L, Rodd FH, Reznick DN (1999) Familiarity leads to female mate preference for novel males in the guppy, *Poecilia reticulata*. *Animal Behaviour*, **58**, 907–916.

Ingvarsson PK, Whitlock MC (2000) Heterosis increases the effective migration rate. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 1321–1326.

Kelley JL, Magurran AE (2003) Effects of relaxed predation pressure on visual predator recognition in the guppy. *Behavioral Ecology and Sociobiology*, **54**, 225–232.

King RB, Lawson R (1995) Color-pattern variation in Lake Erie water snakes: the role of gene flow. *Evolution*, **49**, 885–896.

Kodric-Brown A (1989) Dietary carotenoids and male mating success in the guppy: an environmental component to female choice. *Behavioral Ecology and Sociobiology*, **25**, 393–401.

Kondolf GM, Li S (1992) The pebble count technique for quantifying surface bed material size in instream flow studies. *Rivers*, **3**, 80–87.

Lemmon PE (1957) A new instrument for measuring forest overstory density. *Journal of Forestry*, **55**, 667–668.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Liley NR, Seghers BH (1975) Factors influencing the morphology and behaviour of guppies in Trinidad. In: *Function and Evolution in Behaviour* (eds Baerends G, Beer C, Manning A), pp. 92–118. Clarendon Press, Oxford.

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

Luyten PH, Liley NR (1991) Sexual selection and competitive mating success of male guppies (*Poecilia reticulata*) from four Trinidad populations. *Behavioral Ecology and Sociobiology*, **28**, 329–336.

Magurran AE (1998) Population differentiation without speciation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **353**, 275–286.

Magurran AE, Seghers BH (1990) Population differences in predator recognition and attack cone avoidance in the guppy *Poecilia reticulata*. *Animal Behaviour*, **40**, 443–452.

Magurran AE, Seghers BH, Carvalho GR, Shaw PW (1992) Behavioural consequences of an artificial introduction of guppies (*Poecilia reticulata*) in N. Trinidad: evidence for the evolution of anti-predator behaviour in the wild. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **248**, 117–122.

Magurran AE, Seghers BH, Shaw PW, Carvalho GR (1995) The behavioural diversity and the evolution of guppy, *Poecilia reticulata*, populations in Trinidad. *Advances in the Study of Behaviour*, **24**, 155–202.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology & Evolution*, **20**, 136–142.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Michalakis Y, Excoffier L (1996) Genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, **142**, 1061–1064.

Millar NP, Reznick DN, Kinnison MT, Hendry AP (2006) Disentangling the selective factors that act on male colour in wild guppies. *Oikas*, in press.

Miller MP (1997) *Tools for Population Genetic Analysis (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data*. Computer software distributed by author.

Nicoletto PF (1996) The influence of water velocity on the display behavior of male guppies, *Poecilia reticulata*. *Behavioral Ecology*, **7**, 272–278.

Nicoletto PF, Kodric-Brown A (1999) The relationship among swimming performance, courtship behavior, and carotenoid pigmentation of guppies in four rivers in Trinidad. *Environmental Biology of Fishes*, **55**, 227–235.

Nosil P (2004) Reproductive isolation caused by visual predation on migrants between divergent environments. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **271**, 1521–1528.

Nosil P, Crespi BJ (2004) Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. *Evolution*, **58**, 102–112.

Nosil P, Vines TH, Funk DJ (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, **59**, 705–719.

O'Steen S, Cullum AJ, Bennett AF (2002) Rapid evolution of escape ability in Trinidadian guppies (*Poecilia reticulata*). *Evolution*, **56**, 776–784.

Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical habitats. *Proceedings of the National Academy of Sciences, USA*, **99**, 13612–13615.

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

Paterson IG, Crispo E, Kinnison MT, Hendry AP, Bentzen P (2005) Characterization of tetranucleotide microsatellite markers in guppy (*Poecilia reticulata*). *Molecular Ecology Notes*, **5**, 269–271.

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.

Poissant J, Knight TW, Ferguson MM (2005) Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscape on the genetic structure of a stream-dwelling fish. *Molecular Ecology*, **14**, 1321–1331.

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

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences, USA*, **94**, 9197–9221.

Raufaste N, Rousset F (2001) Are partial Mantel tests adequate? *Evolution*, **55**, 1703–1705.

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

Reznick DN, Bryga HA (1987) Life-history evolution in guppies (*Poecilia reticulata*). 1. Phenotypic and genetic changes in an introduction experiment. *Evolution*, **41**, 1370–1385.

Reznick DN, Bryga HA (1996) Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). V. Genetic basis of parallelism in life histories. *American Naturalist*, **147**, 339–359.

Reznick DN, Butler IVMJ, Rodd FH, Ross P (1996a) Life-history evolution in guppies (*Poecilia reticulata*). 6. Differential mortality as a mechanism for natural selection. *Evolution*, **50**, 1651–1660.

Reznick DN, Rodd FH, Cardenas M (1996b) Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. *American Naturalist*, **147**, 319–338.

Reznick D, Butler MJ, Rodd FH (2001) Life-history evolution in guppies. VII. The comparative ecology of high- and low-predation environments. *American Naturalist*, **157**, 126–138.

Reznick D, Bryant MJ, Bashey F (2002) *r*- and *K*-selection revisited: the role of population regulation in life-history evolution. *Ecology*, **83**, 1509–1520.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Riechert SE (1993) Investigation of potential gene flow limitation of behavioral adaptation in an aridlands spider. *Behavioral Ecology and Sociobiology*, **32**, 355–363.

Rocha LA, Robertson DR, Roman J, Bowen BW (2005) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **272**, 573–579.

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142**, 1357–1362.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rousset F (2002) Partial Mantel tests: reply to Castellano and Balletto. *Evolution*, **56**, 1874–1875.

Rundle HD (2002) A test of ecologically dependent postmating isolation between sympatric sticklebacks. *Evolution*, **56**, 322–329.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.

Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

Schneider S, Kueffer JJ-M, Roessli D, Excoffier L (1997) *ARLEQUIN: An exploratory population genetics software environment*. Computer software distributed by authors.

Seghers BH (1974) Schooling behavior in the guppy (*Poecilia reticulata*): an evolutionary response to predation. *Evolution*, **28**, 486–489.

Shaw PW, Carvalho GR, Magurran AE, Seghers BH (1991) Population differentiation in Trinidadian guppies (*Poecilia reticulata*): patterns and problems. *Journal of Fish Biology*, **39**, S203–S209.

Shaw PW, Carvalho GR, Seghers BH, Magurran AE (1992) Genetic consequences of an artificial introduction of guppies (*Poecilia reticulata*) in N. Trinidad. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **248**, 111–116.

Shaw PW, Carvalho GR, Magurran AE, Seghers BH (1994) Factors affecting the distribution of genetic variability in the guppy, *Poecilia reticulata*. *Journal of Fish Biology*, **45**, 875–888.

Shikano T, Taniguchi N (2002) Heterosis for neonatal survival in the guppy. *Journal of Fish Biology*, **60**, 715–725.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Smith TB, Schneider CJ, Holder K (2001) Refugial isolation versus ecological gradients. *Genetica*, **112–113**, 383–398.

Smith TB, Calsbeek R, Wayne RK, Holder KH, Pires D, Bardeleben C (2005) Testing alternative mechanisms of evolutionary divergence in an African rain forest passerine bird. *Journal of Evolutionary Biology*, **18**, 257–268.

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627–632.

Stanton ML, Galen C, Shore J (1997) Population structure along a steep environmental gradient: consequences of flowering time and habitat variation in the snow buttercup, *Ranunculus adoneus*. *Evolution*, **51**, 71–94.

van Oosterhout C, Trigg RE, Carvalho GR *et al.* (2003) Inbreeding depression and genetic load of sexually selected traits: how the guppy lost its spots. *Journal of Evolutionary Biology*, **16**, 273–281.

van Oosterhout C, Hutchison WF, Wills DPM, Shipley P (2004) MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, **53**, 1446–1457.

Via S, Bouck AC, Skillman S (2000) Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution*, **54**, 1626–1637.

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

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity*, **82**, 117–125.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

This work was completed as part of E. Crispo's MSc research at McGill University, under the supervision of A.P. Hendry. Her primary interests include speciation and the evolution of biological diversity. P. Bentzen is a professor at Dalhousie University who studies fisheries biology and conservation genetics. D.N. Reznick is a professor at the University of California who specializes in evolutionary ecology of guppies. M.T. Kinnison is an assistant professor at the University of Maine who studies microevolution, aquatic ecology, and conservation genetics. A.P. Hendry is an assistant professor at McGill University who studies the evolution of biological diversity.