Generic Scan Using AFLP Markers as a Means to Assess the Role of Directional Selection in the Divergence of Sympatric Whitefish Ecotypes

David Campbell and Louis Bernatchez

Département de Biologie, Université Laval, Ste-Foy, Québec, Canada

Under the ecological theory of adaptive radiation, adaptation and reproductive isolation are thought to evolve as a result of divergent natural selection. Accordingly, elucidating the genetic basis of these processes is essential toward understanding the role of selection in shaping biological diversity. In this respect, the number of genes that evolved by selection remains contentious. To address this issue, the pattern of genetic differentiation obtained using 440 AFLP loci was compared with that expected under neutrality in four sympatric pairs of lake whitefish ecotypes that evolved adaptive phenotypic differences associated with the exploitation of distinct ecological niches. On average, 14 loci showed restricted gene flow relative to neutral expectation, suggesting a role of directional selection on their divergence. Among all loci that are most likely under directional selection, six exhibited parallel patterns of divergence, which provided further support for the role of selection in driving their divergence. Overall, these results indicate that only a small proportion of scored AFLP loci (between 1.4% and 3.2%) might be linked to genes implicated in the adaptive radiation of lake whitefish.

Introduction

Elucidating the processes underlying population divergence and ultimately speciation is central towards understanding the origin of biodiversity (Howard 1998). Ecological shifts in habitat use accompanied by phenotypic differentiation represents a specific case of divergence where deterministic forces, rather than random processes, are likely to be implicated (Orr and Smith 1998). Under the ecological theory of adaptive radiation, such shifts are the result of divergent natural selection stemming from environment heterogeneity and competitive interactions (Schluter 2000). For speciation to occur, reproductive isolation must build up between divergent populations (Mayr 1963). Within the process of ecological speciation (Schluter 1996*a*, 1996*b*), reproductive isolation is often thought to evolve as a by-product of divergent natural selection because traits responsible for extrinsic postzygotic and prezygotic reproductive isolation are usually adaptive (Coyne 1992; Turelli, Barton, and Coyne 2001; Presgraves et al. 2003). Thus, reproductive isolation could share a common genetic basis with adaptation through pleiotropic interactions or physical linkage (Hawthorne and Via 2001).

The number and effect of genes fixed during adaptation and reproductive isolation is still a contentious issue, given empirical support for either Fisher's (Fisher 1930) geometric model (Barton and Hewitt 1981; Szymura and Barton 1991; Rieseberg, Baird, and Gardner 2000) or Kimura's (Kimura 1983) oligogenic model (Tanksley 1993; Jiggins et al. 1996; Bradshaw et al. 1998; Sucena and Stern 2000). Alternatively, it has been proposed that a model including both types of mutations might better reflect the genetic basis of adaptation (Orr 1998). These contrasting views implicate the necessity of further investigation in this field, especially for nonmodel organisms (Rogers et al. 2001) in search of a global understanding of the genetic basis underlying the processes of adaptation and reproductive isolation in wild populations.

Hybrid zones have long been viewed as natural laboratories, within which all components of gene flow barriers are represented, therefore, providing a means to estimate the number of genetic factors contributing to adaptation and reproductive isolation (Barton and Hewitt 1989; Harrison and Rand 1989). After hybridization, recombinant genotypes with maladaptive combinations of loci will be removed, resulting in a differential pattern of introgression. Theory, thus, predicts that loci incurring a "disadvantage" (or linked loci) in recombinant genotypes will also be characterized by a greater genetic differentiation relative to "neutral" or "selectively" advantageous loci (Barton and Hewitt 1981). Thus, identification of loci linked to genes implicated in differential adaptation and/or reproductive isolation can be achieved by quantifying patterns of genetic differentiation between populations that have recently or are still exchanging genes (Bowcock et al. 1991; Rieseberg, Linder, and Seiler 1995; Beaumont and Nichols 1996; Rieseberg, Whitton, and Gardner 1999; Rogers et al. 2001; Wilding, Butlin, and Grahame 2001; Akey et al. 2002; Schlötterer 2002*a*, 2002*b*; Kayser, Brauer, and Stoneking 2003). This, however, requires the assay of a large number of loci to accurately estimate the expected level of genetic differentiation under "neutrality," and the proportion of loci linked to genes implicated in adaptation and reproductive isolation (Schlötterer 2003). In model organisms for which important genomic resources are available, the high-throughput use of single-nucleotide polymorphism (SNP) allows the efficient screening of many thousands of loci (Akey et al. 2002; Brumfield et al. 2003). Until similar resources become accessible for studying nonmodel organisms, the analysis of amplified fragment length polymorphism (AFLP) (Vos et al. 1995) currently represents the best alternative to SNPs for efficiently screening large number of loci. However, to our knowledge, the use of AFLP in the context of assessing the role of selection in shaping patterns of differentiation among wild populations has been limited to a single study (Wilding, Butlin, and Grahame 2001).

Key words: *Coregonus clupeaformis*, directional selection, speciation, hybridization, reproductive isolation, adaptation.

E-mail: louis.bernatchez@bio.ulaval.ca.

Mol. Biol. Evol. 21(5):945–956. 2004
DOI:10.1093/molbev/msh101
Advance Access publication March 10, 2004

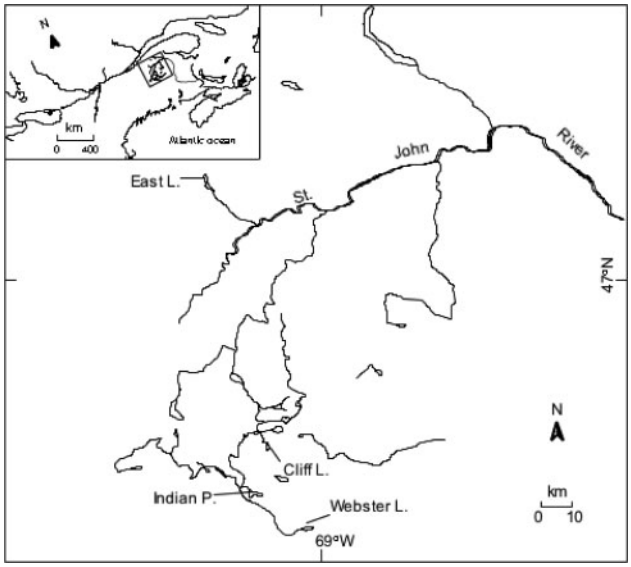

FIG. 1.—Map of study site with location of sampled lake whitefish populations.

After the Pleistocene glaciations, groups of north-temperate freshwater fish populations colonized new habitats that favored the widespread sympatric divergence of populations under conditions of high ecological opportunity (Bernatchez and Wilson 1998; Taylor and McPhail 1999; Robinson and Schluter 2000). The lake whitefish complex, composed of North American (*Coregonus clupeaformis*) and Eurasian (*Coregonus lavaretus*) populations represent such a group (Bernatchez 2003). Within the St. John River drainage, located in southeastern Québec to northern Maine, sympatric ecotypes (dwarf and normal) of lake whitefish (*Coregonus clupeaformis*) are observed in multiple lakes (Fenderson 1964; Chouinard, Pigeon, and Bernatchez 1996; Pigeon, Chouinard, and Bernatchez 1997). These have evolved adaptive phenotypic differences in life history and in behavioral, physiological, and morphological characters for the differential exploitation of limnetic (dwarf) and benthic (normal) ecological niches (Bodaly 1979; Bernatchez, Chouinard, and Lu 1999; Rogers, Gagnon, and Bernatchez 2002). The facts that the dwarf ecotype is never found in allopatry and that lakes harboring sympatric ecotypes have been isolated after the last glacial retreats (Bernatchez, Chouinard, and Lu 1999) indicate that whitefish ecotypes have evolved in parallel more than once (Pigeon, Chouinard, and Bernatchez 1997). Furthermore, the observed correlation between the extent of genetic differentiation and trophic specialization supported the hypothesis that these forms diverged through the process of ecological speciation (Lu and Bernatchez 1999). Whitefish, therefore, represents a very pertinent system for investigating the genetic basis of adaptation and reproductive isolation in wild populations of nonmodel organisms.

The first objective of this study was to document the differential pattern of genetic differentiation at AFLP markers between members of sympatric pairs of whitefish from four lakes (Indian Pond, Webster Lake, Cliff Lake, and East Lake) to estimate the proportion of loci potentially linked to genes implicated in their adaptive divergence and/or reproductive isolation. These lakes have a different history of colonization. In Indian Pond, Webster Lake, and Cliff Lake, dwarf and normal ecotypes are derived from two distinct evolutionary lineages (Acadian and Atlantic, respectively) that came into secondary contact after the Wisconsinian glaciation, whereas in East Lake, they are derived from only one evolutionary lineage (Acadian) (Bernatchez, Chouinard, and Lu 1999). Because dwarf and normal ecotypes differ in many adaptive traits likely under polygenic control (e.g., growth and swimming behavior), we predicted that a large proportion of loci would be potentially linked to genes under the effect of directional selection. The evolution of similar adaptations in independent environments posit natural selection as the main cause of diversification (Schluter 2000). Therefore, parallel trends in the divergence of genes (or linked markers) underlying these adaptations would further support the role of selection in shaping population divergence (Gravilets 1997). Thus, a second objective was to determine whether trends in the divergence of loci potentially linked to genes implicated in the adaptive divergence and/or reproductive isolation of lake whitefish were indicative of the action of parallel selection across the four sympatric pairs.

Materials and Methods

Sampling

Pairs of lake whitefish sympatric dwarf and normal ecotypes were sampled in four lakes of the St. John River drainage using gillnets (fig. 1 and table 1). Liver or white muscle tissues were collected from each specimen and preserved in 95% ethanol for subsequent total DNA extraction using phenol-chloroform (Bernatchez, Guyomard, and Bonhomme 1992).

AFLP Analysis

The AFLP plant mapping kit (Applied Biosystems, Inc.) was used according to the protocol of Vos et al. (1995). After a preselective amplification, seven selective amplification primer combinations were used to generate AFLP markers (table 1). The resulting PCR products were multiplexed (two duplexes [CAAG-GGTG and CTAG-CGTC] and one triplex [CATA-AGTT-CCTC]) and migrated on 6% long-ranger polyacrylamide gels for 3.5 h using an ABI 377 automated sequencer (Applied Biosystems, Inc.). The multiplex strategy allowed us to run the complete set of selective amplifications for 10 individuals on one gel for a total of 24 gels. Fragment analysis was completed using Genescan software (Applied Biosystems, Inc.) with a fluorescent signal detection threshold set to 100 units to avoid background noise. Loci were scored using the Binthere software developed by N. Garnhart (University of New Hampshire) and available at http://tilapia.unh.edu/AFLP/BinThere.html. Binthere generates spreadsheets for AFLP data whereby all scored amplified fragments are automatically placed within 1-bp, size-specific bins (from 50.5 to 499.5 bp; below 50 bp, there are to many fragments to permit an unambiguous scoring), allowing an objective scoring of presence/absence from the bins after correction of fragments alignment by size.

Table 1
Genetic Diversity at Seven AFLP Primer Combinations Within Sympatric Lake Whitefish Ecotypes

Lake	Ecotype		CAAG	GGTG	CTAG	CGTC	CATA	AGTT	CCTC	NB_T	H_A^{primer}	H_M	H_A^{primer}
		Dye	FAM	JOE	FAM	JOE	FAM	JOE	NED				
		NB	103	61	55	48	68	38	67	440			
		S_R	50–248	50–231	52–243	50–219	51–249	50–235	50–195				
Indian	Dwarf	***N***	28	30	29	26	30	30	30				
		H_E	0.34	0.29	0.32	0.29	0.32	0.23	0.30		0.30	0.59	
		P	0.91	0.92	0.73	0.52	0.71	0.84	0.99				0.80
	Normal	***N***	28	34	32	34	34	34	34				
		H_E	0.30	0.27	0.33	0.28	0.33	0.27	0.35		0.30	0.49	
		P	0.89	0.97	0.76	0.63	0.97	0.79	0.99				0.86
East	Dwarf	***N***	28	29	28	24	29	24	29				
		H_E	0.30	0.29	0.32	0.27	0.31	0.25	0.35		0.30	0.61	
		P	0.34	0.30	0.75	0.44	0.71	0.76	0.25				0.51
	Normal	***N***	29	29	29	29	29	29	29				
		H_E	0.31	0.37	0.31	0.29	0.31	0.29	0.33		0.32	0.59	
		P	0.77	0.44	0.84	0.83	0.75	0.79	0.42				0.69
Webster	Dwarf	***N***	30	29	30	30	30	30	30				
		H_E	0.35	0.27	0.22	0.29	0.36	0.16	0.25		0.27	0.60	
		P	0.94	0.75	0.95	0.63	0.71	0.47	0.84				0.75
	Normal	***N***	30	26	29	30	30	30	30				
		H_E	0.25	0.34	0.29	0.35	0.28	0.20	0.30		0.29	0.59	
		P	0.86	0.98	0.85	0.92	0.82	0.87	0.66				0.85
Cliff	Dwarf	***N***	24	24	24	24	24	24	24				
		H_E	0.32	0.38	0.31	0.31	0.35	0.26	0.29		0.32	0.57	
		P	0.83	0.69	0.55	0.48	0.78	0.89	0.60				0.69
	Normal	***N***	22	22	23	22	22	23	23				
		H_E	0.41	0.33	0.37	0.32	0.29	0.22	0.24		0.31	0.50	
		P	0.83	0.77	0.64	0.81	0.57	0.42	0.52				0.65
		H_A^{pop}	0.32	0.32	0.31	0.30	0.32	0.24	0.30				
		H_A^{pop}	0.80	0.73	0.76	0.66	0.75	0.73	0.66				

NOTE.—Primer combinations are denoted by the *Eco*RI Axx and *Mse*I Cxx selective nucleotides, respectively. *Dye*, dye used to label each *Eco*RI primer; *NB*, number of bands scored per primer combination; NB_T, total number of bands scored; S_R, range of band size in base pairs; *N*, number of samples successfully scored; H_E, mean expected heterozygosity; H_A^{primer} mean expected heterozygosity averaged over primer combinations; H_A^{pop}, mean expected heterozygosity averaged over populations; *P*, proportion of polymorphic loci; P_A^{primer}, proportion of polymorphic loci averaged over primer combinations; and P_A^{pop}, proportion of polymorphic loci averaged over populations. All statistics were obtained assuming HWE. Mean heterozygosity at six microsatellite loci (H_M) are reported from Lu and Bernatchez (1999) for comparison.

Genetic Diversity and Population Differentiation

Genetic diversity was quantified in terms of mean expected heterozygosity at each primer combination (H_E), mean expected heterozygosity averaged over primer combinations (H_A^{primer}), mean expected heterozygosity averaged over populations (H_A^{pop}), proportion of polymorphic loci at each primer combination (*P*), proportion of polymorphic loci averaged over primer combinations (P_A^{primer}), and proportion of polymorphic loci averaged over populations (P_A^{pop}). All statistics were obtained assuming Hardy-Weinberg equilibrium (HWE), as Lu and Bernatchez (1999) showed that all populations used in this study were not deviating from HWE using microsatellites. The extent of genetic differentiation between ecotypes in each lake was estimated using the *F*-statistic θ^B of Holsinger, Lewis, and Dey (2002) by running the f = 0 model (assuming HWE) in Hickory version 0.8 (Holsinger and Lewis 2003). Estimates of mean expected heterozygosity (H_M) and genetic differentiation (θ) (Weir and Cockerham 1984) based on six microsatelitte loci are reported from Lu and Bernatchez (1999) for comparison with AFLP data.

Patterns of Divergence Versus Neutral Distribution Inferred from Simulations

To target loci for which genetic differentiation could not be explained solely by interaction between mutation, drift, and migration, the level of differentiation at individual loci was compared with the distribution of differentiation across loci under neutral expectation. Such loci should correspond to those with most restricted gene flow because of the direct or indirect action of differential selection (Barton and Hewitt 1981). The expected distribution was established by running simulations under a neutral model of evolution, with parameters set such that simulations would return the mean observed level of differentiation between populations (Bowcock et al. 1991; Beaumont and Nichols 1996; Wilding, Butlin, and Grahame 2001). Differentiation was estimated using the F_{ST} of Nei (1977) after Nei and Chesser (1983) for the observed data because this estimator was implemented in the simulator. Allelic frequencies were estimated assuming Hardy-Weinberg Equilibrium (HWE) from the equation $f_p = 1 - \sqrt{(N-P)/N}$, where f_p is the frequency of the dominant allele, *N* is the sample size. and *P* is the number of individuals with the band.

Simulations were performed using the algorithm developed by Wilding, Butlin, and Grahame (2001) for dominant loci, which assumes a finite island model of migration between two populations of diploid individuals with a given average effective population size (N_e), mutation rate (μ), and migration rate (*m*) per generation. Such a model mimics a closed system consisting of two lacustrine fish populations exchanging migrants with each other

while being completely isolated from any other population outside the lake. N_e was approximated for each population using the equilibrium relationship of Ohta and Kimura (1973), $H_E = 1 - 1/\sqrt{1 + 8N_e\mu}$, where H_E is gene diversity estimated from microsatellite data (Lu and Bernatchez 1999), and μ was taken as 5×10^{-4} mutations/locus/generation, a microsatellite mutation rate commonly applied for salmonids (Estoup and Angers 1998). This yielded comparable N_e estimates in each population (mean = 1,123; range = 711 to 1,394) that were also in the order of those reported for other salmonids (Miller and Kapuscinski 1997; Hansen et al. 2002). Thus, N_e was set to 1,000 in all simulations. The AFLP mutation rate is unknown. However, preliminary data from laboratory crosses suggest a value of the order of 10^{-4} for whitefish (S. Rogers, personal communication), which was therefore used in all simulations. Finally, migration rate was approximated for each sympatric pair from the mean F_{ST} observed for AFLP data using the formula of Crow and Aoki (1984) for a finite island model of migration with two populations ($F_{ST} = 1/[1 + 16N_em + 16N_e\mu]$).

Bowcock et al. (1991) demonstrated that the distribution of F_{ST} under such a model of neutral evolution was affected by the initial allele frequency used to run the simulations. Thus, they suggested use of multiple simulations with different initial allele frequencies for comparing observed F_{ST} with the appropriate distribution. In doing so, however, they related two different allele frequencies: the initial allele frequencies used in the simulation and the allele frequencies observed in the data. Because there is no reason for observed allele frequencies to reflect initial allele frequencies, mainly because of drift, allele frequencies for 500 loci were randomly drawn from a uniform distribution between 0 and 1, equal in the two populations, and then allowed to drift for 10 N_e generations. In doing so, the observed F_{ST} values were compared with a single simulated distribution reflecting a large subsample of all possible distributions obtained with different initial allele frequencies. Samples of 30 individuals were taken from each simulated population to estimate F_{ST} for each locus. One hundred simulations were performed for each sympatric pair, providing a total of 50,000 F_{ST} values (minus those loci [13.3% on average] that were monomorphic in the simulated samples; see Wilding, Butlin, and Grahame [2001]) that were used for building the expected distribution of differentiation under neutrality.

In the simulations, by setting the initial allele frequencies equal in the two populations, it was assumed that dwarf and normal ecotypes were derived from only one population. However, this is not true for Indian Pond, Webster Lake, and Cliff Lake, where dwarf and normal ecotypes are derived from two distinct evolutionary lineages (Acadian and Atlantic, respectively) that came into secondary contact after the Wisconsinian glaciation (Bernatchez, Chouinard, and Lu 1999). Therefore, to ensure that observed distributions in these lakes could be adequately compared with expected distributions obtained by running simulations, we investigated whether equilibrium had been approached both in the natural systems and in the simulations. Equilibrium in Indian Pond, Webster Lake, and Cliff Lake must be evaluated considering colonization by two different lineages. Thus, we used Whitlock's (1992) equation that approximate the time required for genetic variance to decay half way to equilibrium: $t_{1/2} = \ln(1/2)/\ln[1 - m)^2(1 - 1/2N_e)]$, where N_e is the average effective population size obtained from microsatellite data and m is the average migration rate obtained from observed mean F_{ST} based on AFLP data and from average N_e. From this, the time required to reach equilibrium was approximately 572 generations in Indian Pond ($m/2 = 0.000955$, $N_e = 974$), 618 generations in East Lake ($m/2 = 0.000930$, $N_e = 1316$), 1,706 generations in Webster Lake ($m/2 = 0.000210$, $N_e = 1275$), and 1,490 generations in Cliff Lake ($m/2 = 0.000195$, $N_e = 926$). Taking a generation time of 3 years for whitefish (Chouinard, Pigeon, and Bernatchez 1996), approximately 4,000 generations elapsed since the two lineages colonized the system (approximately 12,000 years ago). Thus, according to the above estimates, equilibrium has likely been achieved in these lakes. Although this model cannot strictly be applied to East Lake populations, it is plausible they also reached equilibrium, given that their N_e estimates were comparable to others, that gene flow was more pronounced than for other lakes, and that time to reach equilibrium is a function of $\sqrt{2N_em}$ (Slatkin 1993).

The approach to equilibrium within the simulations was evaluated analytically using equation (5) of Chakraborty and Jin (1992). This equation predicts F_{ST} in the absence of gene flow as a function of time (10 N_e generations for the simulations), N_e (average in each simulation = 1,000), the number of populations (two per sympatric pair), and the heterozygosity (average across AFLP loci and populations in each simulation: $H_E^{Indian} = 0.33$, $H_E^{East} = 0.32$, $H_E^{Webster} = 0.31$, $H_E^{Cliff} = 0.30$). Using this equation, the amount of divergence was estimated at equilibrium (F_{ST}^*) and after 10,000 generations (F_{ST}^{10000}) using the estimates obtained from the simulations (Indian Pond: $F_{ST}^* = 0.50$, $F_{ST}^{10000} = 0.48$; East Lake: $F_{ST}^* = 0.52$, $F_{ST}^{10000} = 0.49$; Webster Lake: $F_{ST}^* = 0.53$, $F_{ST}^{10000} = 0.50$; Cliff Lake: $F_{ST}^* = 0.54$, $F_{ST}^{10000} = 0.51$). From these results, equilibrium values would have almost been reached after 10,000 generations in the absence of gene flow. However, because simulations were run for 10,000 generations in presence of gene flow, and that equilibrium values are lower and reached more quickly under such conditions, these results indicate that equilibrium values have been attained in the simulations.

Assuming that equilibrium has been reached both in the simulations and in the natural systems as supported above, observed and simulated distributions of F_{ST} can be compared as a mean to identify loci potentially linked to genes implicated in the adaptive divergence and/or reproductive isolation of lake whitefish. Moreover, mean F_{ST} estimated from the simulated samples were not significantly different from values expected from theory in all but one sympatric pair. In Indian Pond, where it was significant, divergence was greatest in the simulation, making the test for selection more conservative (see *Results*).

Observed and simulated F_{ST} distributions were divided into 11 categories of F_{ST} (from -0.1 to $+1.0$) to test whether they were significantly different in each

Table 2
Estimates of Genetic Differentiation Between Dwarf and Normal Whitefish Ecotypes Based on 440 AFLP Loci and Six Microsatellite Loci

	Genetic Differentiation	
Lake	AFLP	Microsatellite
Indian	0.0420 (0.0331–0.0526)	0.0840 (0.0120–0.1750)
East	0.1140 (0.0935–0.1355)	0.0580 (0.0130–0.1100)
Webster	0.1723 (0.1513–0.1948)	0.1400 (0.0200–0.3020)
Cliff	0.2198 (0.1942–0.2479)	0.2560 (0.1470–0.3390)

NOTE.—Estimates for AFLP are θ^B with 95% credible intervals, and those for microsatellites are θ with 95% confidence intervals. Estimates for microsatellites are reported from Lu and Bernatchez 1999.

category for each of the four lakes. Differences between observed and simulated distribution in each category were considered significant when simulated and empirical 95% confidence intervals did not overlap. Confidence intervals for the simulated distributions were obtained based on 100 simulations separately, while they were obtained by bootstrapping (100 bootstrap replicates) over loci for the observed distributions. No bootstrapping procedure was used to estimate confidence intervals for the simulations. Instead 100 simulations were performed. Loci with F_{ST} greater than 95% quantile of the expected distribution of differentiation under neutrality (hereafter, S_L) were considered the most likely to be linked to genes under the effect of directional selection in comparison with those with F_{ST} less than 95% quantile (hereafter, NS_L) (Bowcock et al. 1991; Beaumont and Nichols 1996).

Test of Parallel Pattern of Divergence

We then determined whether trends in the divergence of S_L markers among the four sympatric pairs could reveal the effect of parallel selection. Parallel trends of divergence corresponded to loci for which the frequency of presence (presence of an AFLP band) was greatest in the same ecotype (either dwarf or normal) in all four sympatric pairs. A chi-square test was used to test the hypothesis that parallel trends were not observed more or less often than expected under neutrality for both NS_L and S_L markers. The prediction was that there should be no significant difference between observed and expected numbers of parallel trends among NS_L markers, whereas a significant excess of observed parallel trends should be observed among S_L markers. Expected numbers of loci showing parallel and nonparallel trends among NS_L and S_L markers were calculated from expected probability under neutrality (see details in *Results* section).

Neighbor-Joining population trees were then constructed to compare the topologies obtained using NS_L versus S_L markers exhibiting a parallel trend of divergence. Because the number of NS_L markers largely exceeded the number of S_L, 100 sets of NS_L markers were generated by randomly sampling a number of loci equal to the number of S_L markers exhibiting a parallel trend. Nei's genetic distances were then estimated by running 100 bootstrap replicates using AFLP-SURV version 1.0 (Vekemans 2002) for each of the 100 sets of NS_L and for S_L exhibiting a parallel trend. Consensus trees were constructed using the distance matrices obtained by bootstrapping with the procedures NEIGHBOR and CONSENSE from the PHYLIP version 3.6 software package (Felsentstein 1993).

Results

Genetic Diversity and Population Differentiation

The seven primer combinations used generated a total of 440 AFLP loci and the number of fragments obtained for each of them varied from 38 to 103 (table 1). Moderate to high levels of genetic diversity were observed within each sample, with mean expected heterozygosity varying from 0.16 to 0.41 and mean proportion of polymorphic loci varying from 0.25 to 0.99 for individual primer combination (table 1). As previously observed for microsatellites (Lu and Bernatchez 1999), mean expected heterozygosity values averaged over all loci were generally similar across populations. Estimates of genetic differentiation based on AFLP loci (θ^B) and microsatellite loci (θ) did not differ significantly as their confidence (microsatellites) and credible (AFLP) intervals always overlapped (table 2). However, the extent of genetic differentiation between dwarf and normal ecotypes differed among sympatric pairs, varying from 0.0420 to 0.2198 with AFLP and from 0.0580 to 0.2560 with microsatellites (table 2).

Patterns of Divergence Versus Neutral Distribution Inferred from Simulations

The percentage of observed and simulated polymorphic loci were significantly different in all but one

Table 3
Comparison Between Observed and Simulated Percentage of Polymorphic Loci (*P*) and Genetic Differentiation Based on Nei's Estimator of F_{ST} for Each Sympatric Pair of Whitefish Ecotypes

		P		F_{ST}	
Lake	*m*	Observed	Simulated	Observed	Simulated
Indian	0.0019 (0.0015–0.0023)	91.36 (88.86–94.09)	85.34 (82.40–88.60)	0.0309 (0.0258–0.0365)	0.0477 (0.0415–0.0535)
East	0.0011 (0.0009–0.0012)	72.27 (69.09–75.91)	86.50 (83.40–89.00)	0.0532 (0.0445–0.0595)	0.0670 (0.0572–0.0751)
Webster	0.0006 (0.0005–0.0007)	95.23 (93.86–97.50)	87.56 (84.60–90.00)	0.0868 (0.0746–0.0989)	0.0994 (0.0875–0.1079)
Cliff	0.0004 (0.0003–0.0004)	86.59 (84.09–88.86)	87.31 (84.40–89.60)	0.1202 (0.1053–0.1355)	0.1281 (0.1150–0.1390)
Mean		86.36	86.68	0.0727	0.0855

NOTE.—Estimates of migration rate (m) between ecotypes obtained from observed F_{ST} ($F_{ST} = 1/[1 + 16Nm + 16N\mu]$) and used for simulations are also indicated. Values of simulated P and F_{ST} are means with 95% confidence intervals obtained from 100 simulations. The 95% confidence intervals of m, observed P, and observed F_{ST} were obtained by bootstrapping 100 times over loci.

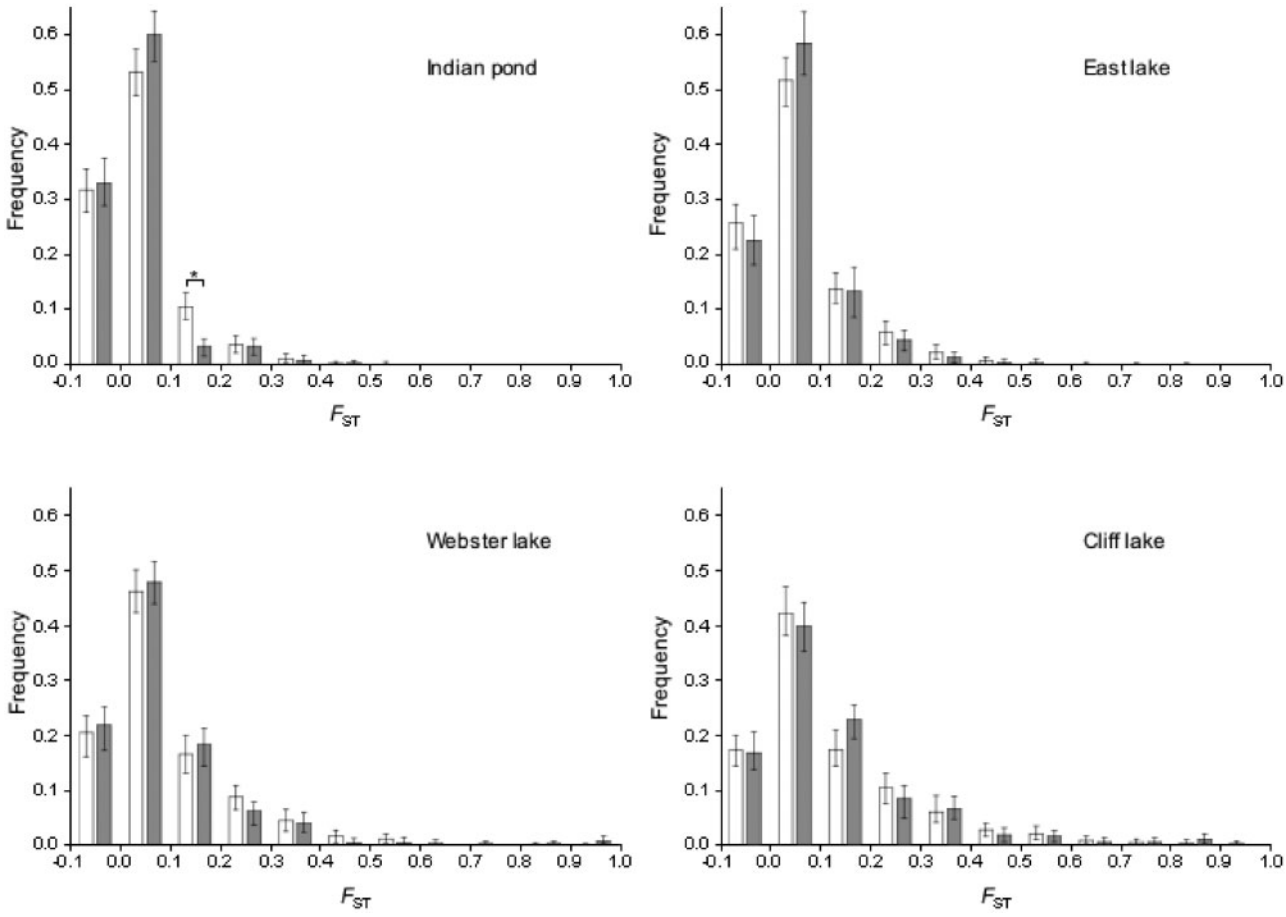

FIG. 2.—Comparison of simulated (open bars) and observed (filled bars) distributions of F_{ST} for each sympatric pair of whitefish ecotypes. Frequencies of simulated bars are means with 95% confidence intervals obtained from 100 simulations. Error bars on the empirical distributions are 95% confidence intervals obtained by bootstrapping 100 times over loci. Differences between simulated and empirical distributions are considered significant where simulated and empirical 95% confidence intervals do not overlap (indicated by an asterisk).

sympatric pair, as the 95% confidence intervals never overlapped except for Cliff Lake (table 3). Yet, the observed differences were relatively weak such that the mean percentage of polymorphic loci for observed and simulated loci were almost identical (observed, $P = 86.36$; simulated, $P = 86.68$) (table 3). Mean simulated F_{ST} values were slightly greater than mean observed F_{ST} values, but the difference was only significant in Indian Pond (table 3).

The overall comparison of observed and simulated patterns of divergence revealed no significant difference in East Lake, Webster Lake, and Cliff Lake. A significant differentiation was observed in Indian Pond. This difference was restricted to a single category of small F_{ST} values (from 0.1 to 0.2) for which fewer loci than expected under the neutral model were observed. Although no significant difference was observed in Webster Lake and Cliff Lake, a small excess of loci with high F_{ST} values were observed relative to the simulated data (fig. 2). In Webster Lake, three loci (CI: 0 to 6) with F_{ST} values between 0.9 and 1.0 were observed compared with no simulated loci (CI: 0 to 1), and in Cliff Lake, four loci (CI: 1 to 8) with F_{ST} values between 0.8 and 0.9 were observed compared with one simulated loci (CI: 0 to 3).

In Indian Pond, East Lake, Webster Lake, and Cliff Lake, there were, respectively, 19, 9, 11, and 16 loci (mean = 14 S_L or 3.2% of screened loci) exhibiting F_{ST} values greater than the 95% quantile of their respective simulated distribution of differentiation under neutrality. Overall, there were a total of 48 S_L, which is not equal to the sum of S_L over all sympatric pairs, because some loci exhibited F_{ST} values greater than the threshold in more than one sympatric pair. The proportion of S_L among the loci generated by each primer combination varied importantly. Namely, the mean percentage of S_L among loci from primer combination CATA (6.30%) was more than twice as important as those observed for any other primer combination, except for primer combination CCTC. This translated into a variation of the contribution of each primer combination to the pool of S_L, with contribution of primer combination CATA being, on average, twice as important as those observed for any other primer combination, except for primer combination CCTC. In contrast, not a single locus of the AGTT primer combination exceeded the 95% quantile threshold (table 4).

Test of Parallel Pattern of Divergence

A chi-square test was used to assess whether parallel trends in the divergence of loci were observed more or less often than expected under neutrality for both NS_L (392

Table 4
Proportion of Loci Potentially Linked to Genes Under Differential Selection (S_L)

Lake	Q_{95}	Number S_L	CAAG	GGTG	CTAG	CGTC	CATA	AGTT	CCTC
Indian	0.1957	19	0.00	10.5	15.8	10.5	63.2	0.00	0.00
			0.00	3.30	5.50	4.20	17.6	0.00	0.00
East	0.2562	9	55.6	11.1	11.1	0.00	22.2	0.00	0.00
			4.90	1.60	1.80	0.00	2.90	0.00	0.00
Webster	0.3624	11	9.10	0.00	0.00	9.10	18.2	0.00	63.6
			1.00	0.00	0.00	2.10	2.90	0.00	10.4
Cliff	0.4556	16	6.20	31.3	6.20	6.20	31.3	0.00	18.8
			1.00	8.20	1.80	2.10	7.40	0.00	4.50
Mean		14	17.7	13.2	8.30	6.50	33.7	0.00	20.6
			1.50	2.90	2.30	1.60	6.30	0.00	3.70

NOTE.—Included in the table are the 95% quantile (Q_{95}) of the expected distribution of F_{ST} under neutrality in each sympatric pair and the number of loci with F_{ST} greater than Q_{95} (Number S_L). Two proportions are given for each primer combination. On the first line is the proportion of S_L from each primer combination, and on the second line is the proportion of S_L among the loci generated by each primer combination.

loci) and S_L (48 loci). Probabilities of expected parallel and nonparallel trends under neutrality were obtained based on NS_L. From these loci, the expected proportion of parallel trends in which the frequency of presence is always greatest in the dwarf ecotype was calculated as the products of the proportions, from each sympatric pair, of time where the frequency of presence was greatest for that ecotype. In Indian Pond, East Lake, Webster Lake, and Cliff Lake, these proportions were equal to 0.5281, 0.2066, 0.5281, and 0.4719, respectively, such that the expected proportion of parallel trends in which the frequency of presence is greatest in the dwarf ecotype equals 0.0272. The expected proportion of parallel trends in which the frequency of presence is greatest in the normal ecotype was obtained in the same manner. In Indian Pond, East Lake, Webster Lake, and Cliff Lake, the proportions of time where the frequency of presence was greatest in the normal ecotype were equal to 0.3648, 0.5128, 0.3954, and 0.4133, respectively, such that the expected proportion of parallel trends in which the frequency of presence was greatest in the normal ecotype equals 0.0306. The expected probability of parallel trends equals the sum of the expected proportions of both possible parallel trends ($0.0272 + 0.0306 = 0.0578$), and the expected probability of nonparallel trends, thus, equals 0.9422. Expected numbers of parallel and nonparallel trends within NS_L and S_L were obtained by multiplying these probabilities by the absolute number of loci in each category (table 5).

Among NS_L, the number of observed parallel trends was not significantly different from the expected number of parallel trends under neutrality (23 versus 22.64; $P = 0.9381$), whereas with S_L, parallel trends were observed significantly more often than expected under neutrality (6 versus 2.77; $P = 0.0458$) (table 5). Among S_L exhibiting a parallel trend in their divergence, three loci were found to be potentially linked to genes under selection in more than one sympatric pair (table 6). The primer combination CATA, which showed the greatest proportion of S_L, also predominated within S_L exhibiting a parallel trend, representing 67% of these loci but representing only 15% of all loci screened (table 6).

The neighbor-joining population tree constructed from the S_L exhibiting a parallel trend clearly reflected the divergence between dwarf and normal ecotypes as it grouped both forms at opposing end of the tree with strong bootstrap values (fig. 3*B*). Such topology of dichotomic grouping by ecotype was reproduced in only one of the 100 consensus trees constructed with 100 sets of six loci that were randomly selected among 392 NS_L. In fact, populations in these trees tended to group more by lake rather than by form (fig. 3*A*).

Table 5
Chi-Square Test to Assess Whether Parallel Trends in the Divergence of Loci Were Observed More or Less Often Than Expected Under Neutrality for Both NS_L (392 Loci) and S_L (48 Loci)

		Observed	Expected	*P*-value
NS_L	Parallel trend	23	22.6	
	No parallel trend	369	369.4	0.94
S_L	Parallel trend	6	2.8	
	No parallel trend	42	45.2	0.046

NOTE.—For explanations on how expected numbers of parallel and nonparallel trends among NS_L and S_L were obtained, see *Results*.

Discussion

Patterns of Divergence Versus Neutral Expectation

The first objective of this study was to estimate the proportion of loci potentially linked to genes implicated in

Table 6
Genetic Differentiation Based on Nei's Estimator of F_{ST} at Six S_L Exhibiting a Parallel Trend in Their Divergence

Locus ID	Indian Pond	East Lake	Webster Lake	Cliff Lake
CATA-67	0.0212	0.3254*	0.2945	0.3435
CATA-73*	0.2584*	0.1454	0.3947*	0.4613*
CATA-104*	0.0324	0.4493*	0.3689*	0.6174*
CATA-143	0.2711*	0.0003	0.2248	0.1835
CGTC-60*	0.2290*	0.0450	0.0436	0.8652*
CTAG-77	0.2605*	< 0.0001	0.0006	0.4100

NOTE.—F_{ST} values greater than the 95% quantile of the expected distribution of differentiation under neutrality are indicated by an asterisk. Locus ID followed by an asterisk indicates that the locus was found to be potentially linked to genes under directional selection in more than one sympatric pair. Locus IDs are made up of the name of the primer combination from which the loci are derived and of the size of AFLP fragments in base pairs.

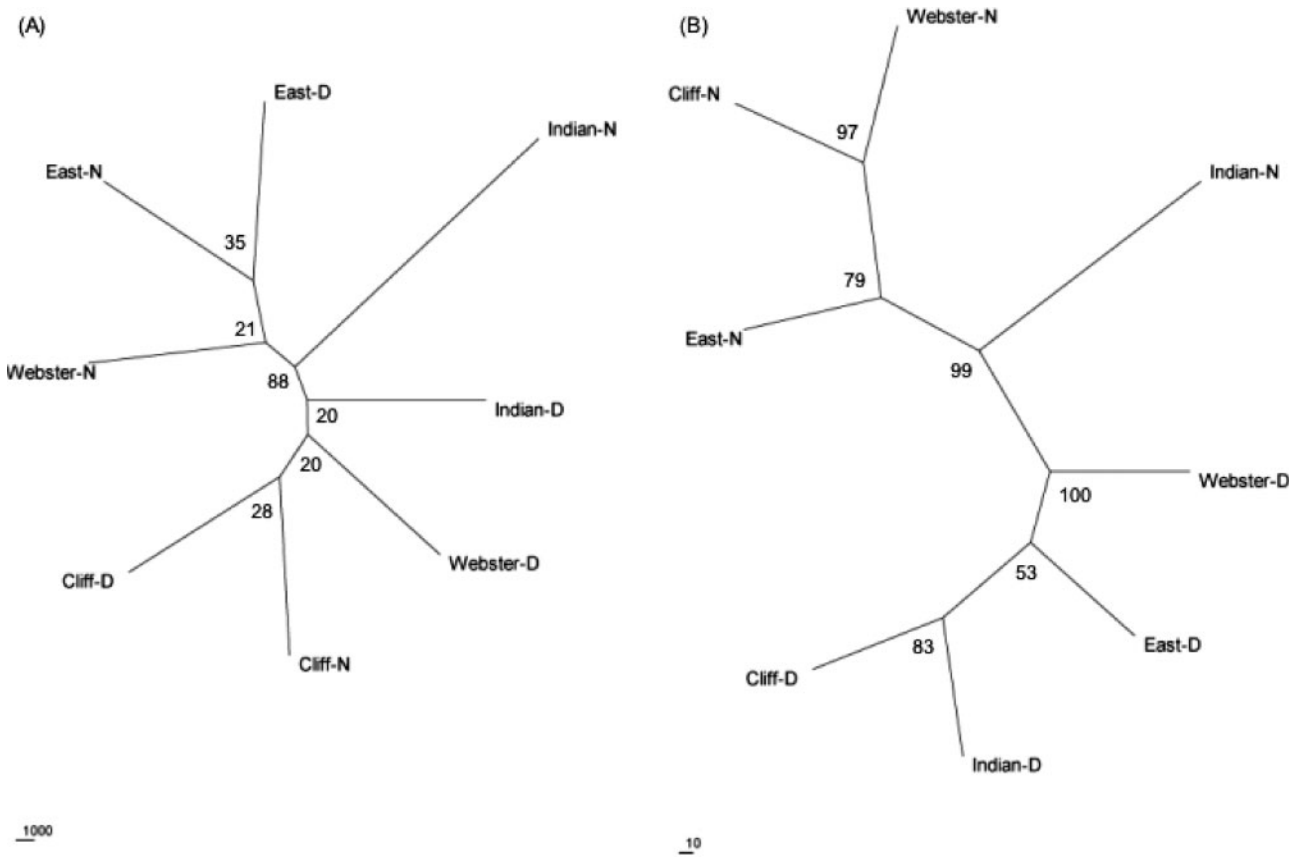

FIG. 3. Neighbour-Joining population trees obtained using Nei's genetic distance. Numbers at nodes represent bootstrap estimates (in percent). Tree distances were estimated (*A*) by pooling 100 sets of six loci randomly selected among 392 NS_L, and (*B*) by using six S_L exhibiting a parallel trend in their divergence. Dwarf and normal ecotypes are identified by D and N, respectively.

adaptation and/or reproductive isolation during population divergence of lake whitefish. To achieve this, the pattern of genetic differentiation of 440 AFLP loci among four sympatric pairs was compared with that expected under neutrality.

Overall, the comparison of observed and simulated patterns of divergence did not provide clear signs of the signature of directional selection, as no significant difference was observed in the tail of observed and simulated distributions of F_{ST} in all sympatric pairs. Still, a small excess of loci with high F_{ST} values were observed in the empirical data relative to the simulated data for Webster Lake and Cliff Lake, which is indicative of the action of directional selection (Bowcock et al. 1991; Akey et al. 2002). On average, our results suggest that within each sympatric pair, gene flow has likely been constrained by selection against less "fit" hybrid recombinant genotypes at 3.2% of scored loci. These represent the most likely loci to be linked to genes implicated in adaptation and/or reproductive isolation of lake whitefish ecotypes (Barton and Hewitt 1981). Assuming that the 440 AFLP loci scored represent an unbiased sampling of the genome (but see below), these results suggest that only a small proportion of genes might have been under the effect of directional selection during population divergence of lake whitefish. This finding might apply more generally to population divergence as other studies on various organisms consistently reported a small proportion of loci potentially under the effect of directional selection. For example, using 306 AFLP loci, Wilding, Butlin, and Grahame (2001) demonstrated that differentiation between parapatric morphs of *Littorina saxatilis* was likely maintained by directional selection acting on 5% of the genome. Using 501 microsatellites, Vigouroux et al. (2002) demonstrated that approximately 2% of the genome had likely been under selection during maize domestication. Finally, Akey et al. (2002) conducted an extensive screening through the human genome using 26,530 SNPs evenly distributed over the 23 chromosomes in three populations and demonstrated that only 0.6% of screened loci had likely been under directional selection.

Test of Parallel Pattern of Divergence

The second objective of this study was to determine whether parallel selection could explain the observed patterns of divergence at S_L markers among the four sympatric pairs. This was supported because parallel trends were observed more often than expected under neutrality among S_L markers (six loci against three expected), which was not the case for the NS_L markers. This was further exemplified by the contrasting tree topologies obtained with both groups of markers. Indeed, the tree constructed from the S_L markers exhibiting a parallel trend clearly reflected the

divergence between dwarf and normal ecotypes as it grouped both forms at opposing ends of the tree with strong bootstrap values. In contrast, populations tended to group more by lake rather than by ecotype with NS_L markers.

Observed parallel trends among loci potentially linked to genes underlying phenotypic differentiation across similar habitats strongly suggest that their divergence was shaped by selection (Harvey and Pagel 1991). Still, few loci among S_L markers exhibited a parallel trend of divergence (six out of 48 loci). Many factors may explain the lack of parallel trends among the remaining 42 S_L markers. First, the hypothesis that divergence at these loci could simply be the result of a balance between mutation, drift, and migration cannot strictly be ruled out. In such a case, an even smaller proportion of genes (1.4%; 6 loci out of 440) might have been under the effect of directional selection during the process of population divergence in whitefish. Second, local rather than parallel selective factors could have contributed to their diversification (Taylor 1991). Indeed, a number of ecological factors may promote divergence between dwarf and normal lake whitefish in different ways between lakes. Third, because natural selection acts indirectly on genes through the phenotypes, if there exist different ways to evolve a particular phenotype, the parallel evolution of a given phenotype will not necessarily be accompanied by the parallel evolution of a given genotype. For instance, Hoekstra and Nachman (2003) discovered that the parallel evolution of adaptive melanism in *Chaetodipus intermedius* has involved different genetic bases. Thus, the similarity of phenotypic adaptations among sympatric pairs of lake whitefish might be in part the result of different genetic bases, in which case. parallel trends in the divergence of S_L markers would not be expected.

Effect of Primer Combinations

Interestingly, loci that are most likely linked to genes under the influence of directional selection were not randomly distributed among primer combinations. In particular, the proportion of S_L from primer combination CATA and the proportion of S_L among the loci generated using the primer combination CATA were at least twice as important as those from all but one of the other primer combinations. Furthermore, two thirds of S_L exhibiting a parallel trend were from the primer combination CATA as well. Although little attention has been paid to this issue thus far in the literature, there is some evidence that AFLP loci from the same primer combination may not be distributed randomly among chromosomes. For instance, Peng et al. (2000) discovered that AFLP markers derived from two primer combinations disproportionately mapped to nine out of 14 linkage groups in wild hemmer wheat. In lake whitefish, four linkage groups out of 29 from a preliminary AFLP linkage map constructed using 12 primer combinations have segregating markers derived from a single primer combination, which, interestingly, happens to be CATA (Rogers et al. 2001). These observations raise the hypothesis that a small proportion of scored loci not only are linked to genes under the influence of directional selection but are also distributed on few linkage groups. However, further investigation into this must await the completion of a more detailed genetic map for homology search of fragments scored within natural populations and laboratory crosses. The apparent nonrandom distribution of AFLP markers derived from the same primer combination over chromosomes also stresses the importance of using many primer combinations for increasing the number of independant loci in studies of population genetic structure and obtaining an unbiased sample of the genome when screening for the signature of selection.

A Hypothetical Scenario for the Genetic Basis of Adaptation and Reproductive Isolation in Lake Whitefish

Although this study indicated that only a small proportion of the genome, in terms of both proportion of loci and number of genomic regions, is likely under directional selection, they do not rule out the possibility that life history and morphological and behavioral traits implicated in the divergence of lake whitefish are under polygenic control. Indeed, genes underlying a given trait may be interacting with one another at the molecular level such that changes in the level and/or timing of expression of only one gene (especially a regulatory gene) could affect the whole network and result in a novel phenotype (White 2001).

Under this scenario, most of the genome would be permeable to introgression such that the maintenance of phenotypic differentiation, despite the persistence of gene flow, would require strong selective pressures (Barton 1983; Barton and Bengtsson 1986; Wu 2001). In lake whitefish, Q_{ST}-F_{ST} analyses indicated that adaptive traits of dwarf and normal ecotypes, such as gill-raker counts and swimming behavior, are indeed under strong directional selection (Rogers, Gagnon, and Bernatchez 2002; Bernatchez 2003). Furthermore, as sympatric pairs of lake whitefish exhibited varying degrees of genetic and morphological differentiation (Lu and Bernatchez 1999, it suggests that the intensity of directional selection acting on each sympatric pair has been different, perhaps because of differences in the extent of ecological niches segregation in their respective habitat (Bernatchez, Chouinard, and Lu 1999). AFLP estimates indicated that effective migration rate between sympatric ecotypes was most pronounced in Indian Pond ($m = 0.0019$) followed by East Lake ($m = 0.0011$), Webster Lake ($m = 0.0006$), and Cliff Lake ($m = 0.0004$), whereas morphological differentiation was greatest in Cliff Lake followed by Indian Pond, Webster Lake, and East Lake (Lu and Bernatchez 1999). Thus, in Indian Pond, morphological differentiation was more pronounced than in East Lake and Webster Lake, despite a higher constraint imposed by gene flow on adaptive divergence (Hendry, Day, and Taylor 2001). This suggests that the intensity of directional selection might have been strongest in Indian Pond relative to East Lake and Webster Lake.

Conclusion

This study provided evidence that only a small proportion (between 1.4% and 3.2%) of scored AFLP loci

might have been linked to genes implicated in the adaptive radiation of lake whitefish which, together with earlier investigations (Tanksley 1993; Bradshaw et al. 1998), support the view that few genes may be underlying processes of adaptation and/or reproductive isolation. However, this conclusion has to be interpreted cautiously, as identification of S_L has relied on the use of simulations with many underlying assumptions. Therefore, if differentiation under neutrality has been underestimated or overestimated as a result of improper assumptions on the demographic history of lake whitefish populations, than the estimated proportion of loci that might have been linked to genes implicated in adaptation and/or reproductive isolation might be biased upwardly or downwardly.

Consequently, before S_L identified in this study can be safely regarded as under the effect of directional selection, concordant results from an independent source are required. Such results may be obtained from QTL mapping. Indeed, if S_L fall among QTL for adaptive traits and/or reproductive isolation, it would strongly support the role of selection on their differentiation. Moreover, QTL mapping is likely to provide other insights of importance to the genetic architecture of adaptation, including establishing the magnitude of effect of loci on adaptive traits, localizing their position on a linkage map and determining the interactions of loci controlling the expression of traits involved in the divergence of ecotypes. This research is currently underway (Bernatchez 2003).

Acknowledgments

We are grateful to Vicky Albert for technical assistance in the laboratory, as well as to Sean Rogers, Dylan Fraser, Jay F. Storz, and two anonymous reviewers for their constructive comments on the manuscript. We also thank R. K. Butlin for kindly providing us with the simulator. This work was supported by the Canadian Research Chair in conservation genetics of aquatic organisms and a research grant from the Natural Science and Engineering Council (NSERC, Canada) to LB. D.C. is supported by an NSERC Postgraduate Scholarship.

Literature Cited

Akey, J. M., G. Zhang, K. Zhang, L. Jin, and M. D. Shriver. 2002. Interrogating a high-density SNP map for signatures of natural selection. Genome Res. **12**:1805–1814.

Barton, N. H. 1983. Multilocus clines. Evolution **37**:454–471.

Barton, N. H., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity **57**:357–376.

Barton, N. H., and G. M. Hewitt. 1981. The genetic basis of hybrid inviability in the grasshopper *Podisma pedestris*. Heredity **47**:367–383.

———. 1989. Adaptation, speciation and hybrid zones. Nature **341**:497–503.

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

Bernatchez, L. 2004. Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (Salmoniformes). Pp. 175–207 *in* A. P. Hendry and S. Stearns, eds. Salmonids and evolution. Oxford University Press, New York.

Bernatchez, L., A. Chouinard, and G. Lu. 1999. Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. Biol. J. Linn. Soc. **68**:173–194.

Bernatchez, L., R. Guyomard, and F. Bonhomme. 1992. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. Mol. Ecol. **1**:161–173.

Bernatchez, L., and C. Wilson. 1998. Comparative phylogeography of nearctic and palearctic fishes. Mol. Ecol. **7**:431–452.

Bodaly, R. A. 1979. Morphological and ecological divergence within the whitefish (*Coregonus clupeaformis*) species complex in Yukon Territory, Canada. Can J. Fish. Aquat. Sci. **36**:1214–1222.

Bowcock, A. M., J. R. Kidd, J. L. Mountain, J. M. Hebert, L. Carotenuto, K. K. Kidd, and L. L. Cavalli-Sforza. 1991. Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. Proc. Natl. Acad. Sci. USA **88**:839–843.

Bradshaw, H., K. Otto, B. Frewen, J. McKay, and D. Schemske. 1998. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). Genetics **149**:367–382.

Brumfield, R. T., P. Beerli, D. A. Nickerson, and S. V. Edwards. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol. Evol. **18**:249–256.

Chakraborty, R., and L. Jin. 1992. Heterozygote deficiency, population substructure and their implications in DNA fingerprinting. Hum. Genet. **88**:267–272.

Chouinard, A., D. Pigeon, and L. Bernatchez. 1996. Lack of specialization in trophicmorphology between genetically differentiated dwarf and normal forms of lake whitefish (*Coregonus clupeaformis*, Mitchill) in Lac de l'Est, Québec. Can. J. Zool. **74**:1989–1998.

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

Crow, J. F., and K. Aoki. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proc. Natl. Acad. Sci. USA **81**:6073–6077.

Estoup, A., and B. Angers. 1998. Microsatellites and minisatellites for molecular ecology: theoritical and experimental considerations. Pp. 55–86 *in* G. R. Carvhalo, ed. Advances in molecular ecology. IOS Press, Amsterdam, The Netherlands.

Felsentstein, J. 1993. PHYLIP (phylogeny inference package). Version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

Fenderson, O. C. 1964. Evidence of subpopulations of lake whitefish, *Coregonus clupeaformis*, involving a dwarf form. Trans. Am. Fish. Soc. **93**:77–94.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford University Press, Oxford, UK.

Gravilets, S. 1997. Hybrid zones with Dobzhansky-type epistatic selection. Evolution **51**:1027–1035.

Hansen, M. M., D. E. Ruzzante, E. E. Nielsen, D. Bekkevold, and K. L. D. Mensberg. 2002. Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. Mol. Ecol. **11**:2523–2535.

Harrison, R. G., and D. M. Rand. 1989. Mosaic hybrid zones and the nature of species boundaries. Pp. 111–133 *in* D. Otte, and J. A. Endler, eds. Speciation and its consequences. Sinauer Associates, Sunderland, Mass.

Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford University Press, New York.

Hawthorne, D. J., and S. Via. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature **412**:904–907.

Hendry, A. P., T. Day, and E. B. Taylor. 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoritical framework for empirical tests. Evolution **55**:459–466.

Hoekstra, H. E., and M. W. Nachman. 2003. Different genes underlie adaptive mechanism in different populations of rock pocket mice. Mol. Ecol. **12**:1185–1194.

Holsinger, K. E., and P. O. Lewis. 2003. Hickory: a package for the analysis of population genetic data. Version 0.8. Distributed by the authors, Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Conn.

Holsinger, K. E., P. O. Lewis, and D. K. Dey. 2002. A Bayesian approach to inferring population structure from dominant markers. Mol. Ecol. **11**:1157–1164.

Howard, D. J. 1998. Unanswered questions and future directions in the study of speciation. Pp. 439–448 *in* D. J. Howard and S. H. Berlocher, eds. Endless forms: species and speciation. Oxford University Press, New York.

Jiggins, C. D., W. O. McMillan, W. Neukirchen, and J. Mallet. 1996. What can hybrid zones tell us about speciation? The case of *Heliconius erato* and *H. himera* (Lepidoptera: Nymphalidae). Biol. J. Linn. Soc. **59**:221–242.

Kayser, M., S. Brauer, and M. Stoneking. 2003. A genome scan to detect candidate regions influenced by local natural selection in human populations. Mol. Biol. Evol. **20**:893–900.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK.

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

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, Mass.

Miller, L. M., and A. R. Kapuscinski. 1997. Historical analysis of genetic variation reveals low effective population size in a northern pike (*Esox lucius*) population. Genetics **147**:1249–1258.

Nei, M. 1977. F-statistics and analysis of gene diversity in subdivided populations. Ann. Hum. Genet.. **41**:225–233.

Nei, M., and R. K. Chesser. 1983. Estimation of fixation indices and gene diversities. Ann. Hum. Genet. **47**:253–259.

Ohta, T., and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet. Res. **22**:201–204.

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution **52**:935–949.

Orr, M. R., and T. B. Smith. 1998. Ecology and speciation. Trends Ecol. Evol. **13**:502–506.

Peng, J., A. B. Korol, T. Fahima, M. S. Röder, Y. I. Ronin, Y. C. Li, and E. Nevo. 2000. Molecular genetic maps in wild Emmer wheat, *Triticum dicoccoides*: genome-wide coverage, massive negative interference, and putative quasi-linkage. Genome Res. **10**:1509–1531.

Pigeon, D., A. Chouinard, and L. Bernatchez. 1997. Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). Evolution **51**:196–205.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature **423**:715–719.

Rieseberg, L. H., S. J. Baird, and K. A. Gardner. 2000. Hybridization, introgression, and linkage evolution. Plant Mol. Biol. **42**:205–224.

Rieseberg, L. H., C. R. Linder, and G. J. Seiler. 1995. Chromosomal and genic barriers to introgression in *Helianthus*. Genetics **141**:1163–1171.

Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics **152**:713–727.

Robinson, B. W., and D. Schluter. 2000. Natural selection and the evolution of adaptive genetic variation in northern freshwater fishes *in* T. A. Mousseau, B. Sinervo, and J. A. Endler, eds. Adaptive genetic variation in the wild. Oxford University Press, New York.

Rogers, S. M., D. Campbell, S. J. E. Baird, R. G. Danzmann, and L. Bernatchez. 2001. Combining the analyses of introgressive hybridisation and linkage mapping to investigate the genetic architecture of population divergence in the lake whitefish (*Coregonus clupeaformis*, Mitchill). Genetica **111**:25–41.

Rogers, S. M., V. Gagnon, and L. Bernatchez. 2002. Genetically based phenotype-environment association for swimming behavior in lake whitefish ecotypes (*Coregonus Clupeaformis*, MITCHILL). Evolution **56**:2322–2329.

Schlötterer, C. 2002*a*. A microsatellite-based multilocus screen for the identification of local selective sweeps. Genetics **160**:753–763.

———. 2002*b*. Towards a molecular characterization of adaptation in local populations. Curr. Opin Genet. Dev. **12**:683–687.

———. 2003. Hitchiking mapping—functional genomics from the population genetics perspective. Trends Genet. **19**:32–38.

Schluter, D. 1996*a*. Ecological speciation in postglacial fishes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **351**:807–814.

———. 1996*b*. Ecological causes of adaptive radiation. Am. Nat. **148**:S40–S64.

———. 2000. The ecology of adaptive radiation. Oxford University Press, New York.

Slatkin, M. 1993. Isolation by distance in equilibrium and nonequilibrium populations. Evolution **47**:264–279.

Sucena, E., and D. L. Stern. 2000. Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven baby*. Proc. Natl. Acad. Sci. USA **97**:4530–4534.

Szymura, J. M., and N. H. Barton. 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *Bombina variegata*: comparisons between transects and between loci. Evolution **45**:237–261.

Tanksley, S. D. 1993. Mapping polygenes. Annu Rev. Genet. **27**:205–233.

Taylor, E. B. 1991. A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture **98**:185–207.

Taylor, E. B., and J. D. McPhail. 1999. Evolutionary history of an adaptive radiation in species pairs of threespine stickelbacks (*Gasterosteus*): insights from mitochondrial DNA. Biol. J. Linn. Soc. **66**:271–291.

Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. Trends Ecol. Evol. **16**:330–343.

Vekemans, X. 2002. AFLP-SURV. Version 1.0. Distributed by the author, Laboratoire Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

Vigouroux, Y., M. McMullen, C. T. Hittinger, K. Houchins, L. Schulz, S. Kresovich, Y. Matsuoka, and J. Doebley. 2002. Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc. Natl. Acad. Sci. USA **99**:9650–9655.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M.

Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23**:4407–4414.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution **38**:1358–1370.

White, K. P. 2001. Functional genomics and the study of development, variation and evolution. Nature Rev. Genet. **2**:528–537.

Whitlock, M. C. 1992. Temporal fluctuations in demographic parameters and the genetic variance among populations. Evolution **46**:608–615.

Wilding, C. S., R. K. Butlin, and J. Grahame. 2001. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. J Evol. Biol. **14**: 611–619.

Wu, C.-I. 2001. The genic view of the process of speciation. J. Evol. Biol. **14**:851–865.

Diethard Tautz, Associate Editor

Accepted Accepted January 9, 2004