Genetica **111:** 25–41, 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)

S.M. Rogers[1], D. Campbell[1], S.J.E. Baird[2], R.G. Danzmann[3] & L. Bernatchez[1,*]
[1]*Groupe interuniversitaire de recherches océanographiques du Québec (GIROQ), Département de Biologie, Université Laval, Ste-Foy, Québec, G1K 7P4, Canada;* [2]*Department of Zoology and Entomology, The University of Queensland, Qld, 4072, Australia;* [3]*Department of Zoology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada;* **Author for correspondence (Phone: 418-656-3402; Fax: 418-656-2339; E-mail: Louis.Bernatchez@bio.ulaval.ca)*

Key words: adaptive radiation, AFLP, *Coregonus clupeaformis*, hybrid index, introgression, linkage mapping, maximum likelihood, speciation

Abstract

Adaptation and reproductive isolation, the engines of biological diversity, are still elusive when discussing the genetic bases of speciation. Namely, the number of genes and magnitude of selection acting positively or negatively on genomic traits implicated in speciation is contentious. Here, we describe the first steps of an ongoing research program aimed at understanding the genetic bases of population divergence and reproductive isolation in the lake whitefish (*Coregonus clupeaformis*). A preliminary linkage map originating from a hybrid cross between dwarf and normal ecotypes is presented, whereby some of the segregating AFLP markers were found to be conserved among natural populations. Maximum-likelihood was used to estimate hybrid indices from non-diagnostic markers at 998 AFLP loci. This allowed identification of the most likely candidate loci that have been under the influence of selection during the natural hybridisation of whitefish originating from different glacial races. As some of these loci could be identified on the linkage map, the possibility that selection of traits in natural populations may eventually be correlated to specific chromosomal regions was demonstrated. The future prospects and potential of these approaches to elucidate the genetic bases of adaptation and reproductive isolation among sympatric ecotypes of lake whitefish is discussed.

Introduction

Ecological shifts, resulting in extremely rapid rates of divergence, stem from the process of adaptive radiation (Orr & Smith, 1998). The ecological theory of adaptive radiation purports that resource-based divergent natural selection is the underlying cause of diversification (Huxley, 1942; Mayr, 1942; Schluter, 2000), and that reproductive isolation may evolve as a byproduct of divergent selection. In such cases, lineages may display rapid speciation under conditions of high ecological opportunity, lending to the term 'ecological speciation' to describe the evolution of reproductive isolation originating from ecological forces (Schluter, 1996a, b, 1998).

Overall, very little is known of the genetic basis of adaptation and reproductive isolation such that the conflict over the number and magnitude of genetic changes responsible for speciation continues today (Barton, 1998; Orr, 1998; Schemske & Bradshaw, 1999). The existing paradigm implies that adaptation and reproductive isolation are caused by an infinite number of genes each with a small individual effect. While Darwin (1859) was the first to suggest the 'accumulation of innumerable slight variation, each good for the individual possessor, Fisher's (1930) geometric model of gradualism appeared to mathematically confirm this idea. However, conflicting evidence now suggests that a small number of genes invoking large adaptive phenotypic effects may be more

descriptive of the genetic basis of adaptation (Kimura, 1983; Barton, 1998; Orr, 1998). The empirical evidence is contradictory, with some studies implicating a large number of loci each with a relatively small effect, that is, infinitesimal (Kim & Rieseberg, 1999; Sawamura, Davis & Wu, 2000). Other studies have found evidence of small number of loci each with a large phenotypic effect, that is, oligogenic (Tanksley, 1993; Bradshaw et al., 1995, 1998).

Linkage mapping techniques offer a promising avenue to resolve this debate through mapping the genetic architecture of reproductive barriers, that is, the number, effects, interactions, and location of genetic factors contributing to reproductive isolation (Rieseberg, 1998a; Rieseberg, Whitton & Gardner, 1999). Moreover, utilisation of a hybrid cross between evolutionary lineages where prezygotic reproductive isolation is incomplete is highly amenable to the construction of an informative linkage map (Rieseberg & Linder, 1999). Controlled hybridisation experiments between different species that exhibit almost complete reproductive isolation are typically restricted to only one or a few generations of recombination. In contrast, hybrid crosses exhibiting incomplete reproductive isolation essentially liberate a wide array of genotypes historically resulting from many generations of recombination that would have otherwise been 'hidden' by natural selection in natural populations. In such hybrid crosses all components of 'interspecific' gene flow barriers are represented (Rieseberg, 1998b). Only two studies to date have successfully utilised hybrid zones (in Sunflower species, *Helianthus* spp.) to investigate the genetic architecture of barriers to interspecific gene flow and correlate the results to available linkage maps (Rieseberg, Whitton & Gardner, 1999; Carney, Gardner & Rieseberg, 2000). The utility of this approach has never been tested in natural animal populations.

Because of their young age, northern temperate fish populations provide an effective window within which critical speciation issues can be addressed directly (Bernatchez & Wilson, 1998). Numerous species consisting of sympatric forms undergoing rapid phenotypic divergence and reproductive isolation have been discovered within this zone (reviewed in Taylor & McPhail, 1999; Robinson & Schluter, 2000 and see also Turgeon, Estoup & Bernatchez, 1999). One of the best studied groups of such sympatric forms is the whitefish complex composed of North American and Eurasian populations (traditionally referred to *Coregonus clupeaformis* (lake whitefish) and *C. lavaretus*, (European whitefish) respectively; but see Bernatchez and Dodson (1994) for inconsistencies between nomenclature and phylogenetic relationships). Phylogeographic studies have revealed the existence of five glacial lineages that diverged during inhabitation of isolated Pleistocene glacial refugia, 18,000 to approximately 500,000 ya (reviewed in Bernatchez, Chouinard & Lu, 1999). Secondary contact, accompanied by introgressive hybridisation has subsequently occurred between glacial races, typically in lakes of the Allagash basin (St. John river drainage) located in south-eastern Québec (Canada) to northern Maine (USA) (Fenderson 1964; Bernatchez & Dodson, 1990; Lu, Basley & Bernatchez, 2001). Whitefish of this area also constitute a species complex given the existence of several sympatric pairs, whereby dwarf and normal ecotypes co-inhabit in at least 10 lakes (Fenderson, 1964; Scott & Crossman, 1973; Bernatchez, Chouinard & Lu, 1999). These ecotypes display adaptive trait phenotypic differences in life-history, behavioural, and morphological characters associated with the use of distinct trophic resources (Fenderson, 1964; Bernatchez, Chouinard & Lu, 1999). Although not firmly demonstrated by the experimental approach, there is little doubt that ecological processes are contributing to maintaining phenotypic divergence, and possibly reproductive isolation between whitefish ecotypes (Bernatchez et al., 1996; Pigeon, Chouinard & Bernatchez, 1997; Lu & Bernatchez, 1999). In addition, intrinsic genetic processes may also elicit the formation of reproductive isolation between populations issued from different glacial races, following evidence of hybrid inviability under experimental conditions (Lu & Bernatchez, 1998).

Overall, whitefish offer many characteristics of an ideal system within which critical issues of relevance for understanding the genetic basis of population divergence and reproductive isolation can be addressed. Namely, the genomic regions identified by mapping of whitefish hybrid forms are more likely to represent genes actually involved in speciation rather than genes that have diverged following speciation. In this paper, we present the first steps of an ongoing research program on whitefish in order to illustrate the potential of both linkage mapping and introgressive hybridisation analyses to investigate the genetic architecture of population divergence and reproductive isolation. First, we tested the amenability of 250 AFLP markers towards creating a preliminary linkage map using a hybrid dwarf/normal lake whitefish backcross. Second, we genotyped whitefish individuals

from a naturally introgressed and putative parental populations at 1171 AFLP loci to illustrate the utility of maximum-likelihood method in estimating hybrid indices from non-diagnostic, dominant markers. We then show how this information can be used to identify candidate loci for which introgression has likely been constrained (or favoured) by selection in the naturally introgressed population. In light of these results, we discuss the potential of the combined use of genomic/QTL mapping and comparative analysis of hybrid indices for testing future hypotheses that can detail the genetic bases of adaptation and reproductive isolation among sympatric ecotypes of lake whitefish.

Materials and methods

Generation of hybrid mapping families

The parental generation of dwarf whitefish were collected in the fall of 1996 from Témiscouata Lake (47°36′N, 68°45′W), whereas the parental generation of normal whitefish were collected from Aylmer Lake (45°50′N, 71°26′W) (Figure 1). The F1 generation consisted of reciprocal intraspecific pure dwarf (DD), pure normal (NN), and hybrid dwarf/normal (ND) crosses as described in Lu and Bernatchez (1998). The pure dwarf F1 were generated with a complete diallel cross using 20 males and 20 females. The notation H1 will be used to define a normal female crossed with a dwarf male, that is, ND.

The backcross (BC1) generation was crossed on November 1, 1999, using the same methodology as Lu and Bernatchez (1998). Adipose fin tissue from each parent was also collected for subsequent DNA extractions. The mapping family was generated by crossing an F1 hybrid female with a pure F1 dwarf male (H1D). After hatching, 60 progeny were collected and used for construction of the preliminary linkage map. Other hybrid backcross families were also generated for future comparative mapping studies but are not discussed here.

Pure and introgressed glacial lineage samples

Lake whitefish were sampled from the secondary contact zone of the St. John river basin of north-eastern North America (Figure 1). Representatives of Atlantic and Acadian glacial lineages were sampled from three different lake populations, respectively. The 'purity' of lineage samples was assessed through a previous factorial correspondence analysis (FCA) that described glacial lineage associations among populations based on six microsatellite loci (Lu, Basley & Bernatchez, 2001). A total sample population of 18 fish originating from the three most representative Atlantic lineage populations (Cliff Lake-normal ecotype ($n = 6$); Ross Lake ($n = 6$); and Clear Lake ($n = 6$)) and 17 fish from the three most representive Acadian lineage populations (Cliff Lake-dwarf ecotype ($n = 6$); East Lake-normal ($n = 6$); and Crescent Pond ($n = 5$) were sampled (see details in Lu et al., 2001). The introgressed sample consisted of 36 individuals of the Témiscouata Lake dwarf ecotype. The estimate of admixture proportion at six microsatellites obtained by the weighted least squares method of Long (1991) was 63% versus 37% for Acadian and Atlantic lineages (Lu, Basley & Bernatchez, 2001). Tissues of liver or muscle were collected from each specimen and preserved in 95% ethanol for subsequent total DNA extraction.

Characterization of AFLP markers for linkage mapping and analysis of hybrid indices

The AFLP plant mapping kit (Applied Biosystems, Inc.) was used according to the protocol of Vos et al. (1995) to generate informative linkage markers from the whitefish mapping family and for the quantification of hybrid indices. We initially tested 64 selective amplification primer possibilities to discern which selective amplification primer pairs yielded the most polymorphic bands for linkage map construction and introgression analysis. For the linkage map, 12 primer combinations that yielded the most consistent and reproducible fragments were chosen while five primer combinations were chosen for the quantification of hybrid indices (Table 1).

Electrophoresis of selective amplification products was performed using an ABI 377 automated sequencer (Applied Biosystems, Inc.) on 5% long-ranger polyacrylamide gels for a migration period of 3.5 h. Analysis of fragments was completed using Genescan software (Applied Biosystems, Inc.) with a fluorescent signal detection threshold set to 100 fluorescent units. The scoring of loci was performed using the software Binthere, developed by N. Garnhart (University of New Hampshire), and available on the website of T. Kocher laboratory (http://www.tilapia.unh.edu). Binthere generates spreadsheets for AFLP data whereby all scored amplified fragments are automatically placed within 1 bp size-specific bins ranging from 50.5 to 499.5 bp

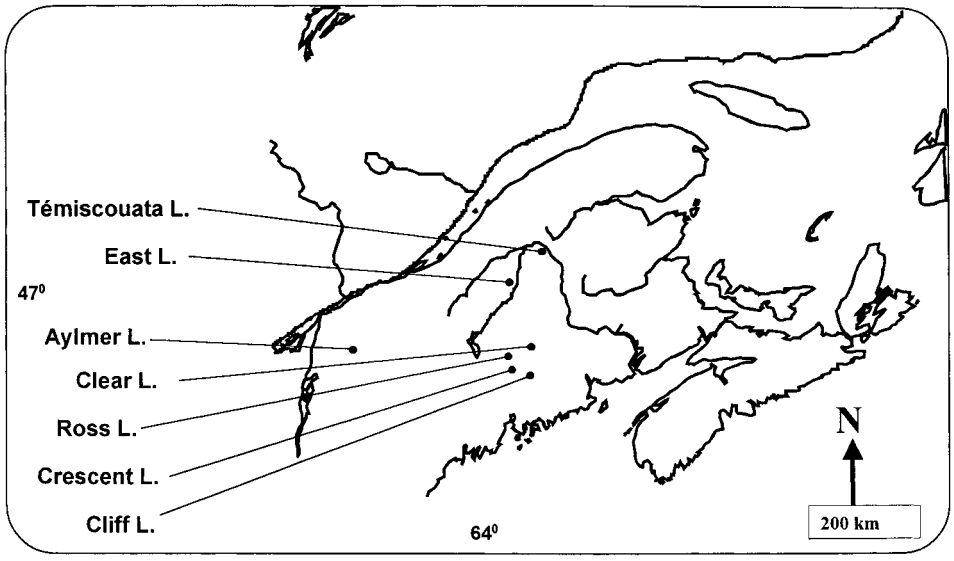

Figure 1. Location map of sampled whitefish populations for this study. (see Lu, Basley & Bernatchez, 2001 for details).

permitting an objective scoring of presence/absence from the bins.

Linkage map construction

All polymorphic AFLP fragments were tested for Mendelian 1:1 segregation ratios with a log-likelihood *G*-test using the software package LinkMFex (R.G. Danzmann, www.uoguelph.ca/~rdanzman/software/LINKMFEX). The significance of distorted segregation ratios was corrected for multiple comparisons using the sequential Bonferonni correction (Rice, 1989), while all markers with table-wide significant deviations were monitored throughout linkage analysis in efforts to guard against the formation of pseudolinkages (Cloutier, Cappadocia & Landry, 1997). As this study represents the first mapping effort of the *Coregonus* genome, linkage groups had to be first designated in LINKGRP (R.G. Danzmann, www.uoguelph.ca/~rdanzmann/software/LINKMFEX) and remaining markers were subsequently linked on the basis of the 'assign' command output in MAPMAKER 3.0 (LOD threshold of 3.0). Genotypic data were entered in both linkage phases to satisfy the requirements of the software (Lander et al., 1987). For linkage groups with nine or fewer markers, the exhaustive analysis using the 'compare' command was performed. For linkage groups with more than 10 markers, a three-point analysis (LOD > 3.0) followed by the 'order' command determined the sequence of markers. Final map distances (cM) were calculated using the Kosambi function (Ott, 1991). All linkage groups were generated under the assumption that sex-specific recombination, a phenomenon implicated in numerous salmonid species, is likely occurring in whitefish (Sakamoto et al., 2000).

Maximum likelihood estimates of hybrid indices using non-diagnostic markers

Hybrid indices are used for both description of and inference about hybrid populations. In the absence of information about the exact nature of the crossing history which has produced a hybrid population, the hybrid indices of individuals from the population allow us to visualise the relative contribution of source populations, and the variation in that contribution among individuals. When hybrid indices are calculated over diagnostic alleles the components of the variance in contributions can be dissected allowing rapid calculation of the pairwise association between alleles (Barton & Gale, 1993). When an allele is non-diagnostic there is doubt about its source population. Therefore, a hybrid index which includes information for any non-diagnostic allele must be estimated rather than calculated. This issue was first addressed by Rieseberg, Baird and Desrochers (1998), and the estimation procedure used in that study has also been used by Rieseberg, Whitton and Gardner (1999). The hybrid index estimates presented here were calculated using the method laid out in the Appendix of Rieseberg,

Table 1. Summary of AFLP locus data for the H1D mapping family and the analysis of hybrid indices

	Selective AFLP primers												Total
	AGAC	CAAG	CTAG	CATA	AGTT	CAAT	ACTA	ACTC	CCTC	CGTC	CTTC	GGTG	
No. of markers characterized													
In H1D mapping family	4	47	13	38	13	14	3	9	6	53	14	36	250
In natural populations	104	185	397	248	243								1177
No. of markers conserved between													
H1D and natural populations	4	33	11	30	12								90
No. of markers successfully mapped	4	21	5	21	2	9	2	1	2	17	9	26	119
No. of mapped markers conserved													
in natural populations	4	10	4	17	2								37
No. of mapped markers													
deviating from 1:1 segregation ratio	3	7	0	6	0	3	0	0	1	5	1	2	28

AFLP primers are denoted by the *Eco*RI Axx & *Mse*I Cxx selective nucleotide combinations, respectively.

Baird and Desrochers (1998). We briefly summarise the formal principles of the approach below.

The approach is largely analogous to the analysis of variance in hybrid index for diagnostic traits described by Barton and Gale (1993). When considering diagnostic traits, the variance in hybrid indices over all individuals can be split into two components: that due to the variance within each trait across all individuals, and that due to the covariance between traits across all individuals. If the traits are discrete Mendelian markers, then the covariance among traits are due to statistical associations between the markers, that is linkage disequilibria. The variance of the hybrid index over all individuals and the variance of marker states at each locus are calculated, and their difference is the variance due to linkage disequilibria. As we know the number of pairs of loci involved, we can therefore rapidly calculate the average pairwise linkage disequilibrium that would explain this variance. We can then make inferences about the various processes which generate linkage disequilibrium: migration, epistatic selection and drift, those which maintains linkage disequilibrium, for example, selection against hybrids, and those which degrade linkage disequilibrium, for example, segregation and recombination.

Hybrid indices over non-diagnostic loci were estimated using the likelihood framework (Fisher, 1925; Edwards, 1972; Hilborn & Mangel, 1997). Likelihood estimation requires a model concerning observations. The model for hybrid populations is based on inheritance: we assume that two source populations that were separated have both contributed to a mixed population. This means that each allele in the mixed population must trace its copy-by-copy ancestry either through one source or the other. Likewise all the parent-by-parent ancestral lineages of an individual must at some point trace through one source or the other. We thus define the *hybrid index for a locus* as the proportion of alleles at that locus which trace their copy-by-copy ancestry through one of the source populations. We then define the *hybrid index for an individual* as the proportion of its parent-by-parent ancestral lineages that trace through one of the source populations, in the generation before the first mixed mating in its ancestry. These definitions are consistent with the calculation of hybrid indices over diagnostic loci (Baird, in preparation).

To estimate hybrid indices we assume to know the frequencies of allele states in the source populations. In the present study the source populations are the putative Acadian and Atlantic glacial lineages. We calculated the likelihood of the source of an allele using the comparative frequencies of alleles of the same state in the two source populations. If the allele's copy-by-copy ancestry traces either through one source or the other, then the more likely source is the one in which alleles of the same state are more common. The likelihood of the hybrid index for a locus is simply the product of the likelihoods of the alleles at that locus. The likelihood of the hybrid index (HI) of an individual is calculated in two stages. First, the frequencies of allele states over all the individual's ancestral lineages are calculated given that a proportion HI are from one source and a proportion (1-HI) are from the other. Second, the likelihood of each of the individual's alleles being observed is calculated, given the allele state frequencies in its ancestral lineages. The likelihood for the hybrid index of the individual is the product of these likelihoods calculated for its alleles.

We can then make inferences about the evolutionary process by comparing hybrid indices estimated over loci and hybrid indices estimated over individuals in an approach analogous to the comparison of the within-locus and across-individual variance components of diagnostic hybrid indices (Barton & Gale, 1993). The nature of these inferences is clear, given the definitions of per locus and per individual hybrid indices: we are contrasting the modes of inheritance of alleles and individuals. The joint estimate of the hybrid index over all individuals is the best estimate of the entire individual-based ancestral contribution from one source. This will be a function of the proportion of individuals which have entered the mixed population from each source, and the relative fitnesses of the source populations in the context of the mixed population (the tendency of one source to give rise to more child-by-child descendants than the other in the mixed population). This can be contrasted with the effect of fitness differences at the per locus level. Although alleles from the two sources will be mixed necessarily at the same ratio as individuals, their subsequent fate need not follow that determined by the population-level fitness of the sources. When hybrid indices within loci are compared to the prediction made across all individuals, we can identify candidates for markers (or linked loci) under strong selection in the context of the introgressed population. Thus, loci with hybrid indices departing the most from the joint estimate of the hybrid index over all individuals are the most likely candidates for directional selection during the process of hybridisation.

The drop in likelihood when a number of estimates are replaced by a single estimate over the same data can be thought of an analogous to the variance among those estimates. These likelihood drops can be used in the same way that variance among individuals and variance within loci can be used to calculate the variance due to linkage disequilibrium. All computations for estimating linkage disequilibrium, hybrid index indices and associated error (2 units of support) were performed using Mathematica 4.0 (Wolfram, 1992).

Results

Characterization of AFLP markers

The selective amplification primer combinations used revealed a high number of markers potentially amenable to linkage map construction and for the introgression analysis. Overall, the 12 primer combinations deemed useful for linkage mapping in the H1D family generated 250 potential polymorphic fragments (Table 1). We also scored the presence/absence of 1177 AFLP loci in the Atlantic and Acadian reference populations that could be used for the maximum likelihood estimates of hybrid indices in the introgressed Témiscouta dwarf population. We found that 37 of these loci were conserved within the mapping families (31% of all mapped markers), suggesting that conservation of AFLP markers between these samples may be sufficient to eventually link traits from laboratory crosses to natural populations, and vice versa (Table 1, but see Discussion).

Linkage analysis and genomic map

The 250 potentially informative AFLP markers for linkage mapping were scored in 60 progeny. Three individuals were missing a substantial number of genotypes and were rejected from the data set. We succeeded in mapping 119 markers to 29 sex-specific linkage groups (Figure 2). The total map distance of the *Coregonus clupeaformis* H1D map was 1462 cM, with an average distance of 16.2 cM between markers. The size of the linkage groups ranged from 12.7 cM to 200.9 cM (average = 50.4 cM). The number of markers per linkage group ranged from 2 to 17.

Of the mapped markers, 28 loci exhibited significant (tablewide $\alpha = 0.05$) deviations from 1:1 Mendelian segregation ratios, indicating that 24% of the total number of mapped markers had distorted segregation ratios (Table 1, Figure 2). High levels of segregation distortion in dominant loci are not uncommon and may even be indicative of linked sequences or loci under positive or negative selection (Rieseberg et al., 1993). Given our main interest in assessing the homology of potentially selected loci between mapped families and natural populations, distorted loci were retained for building the map. Of the distorted AFLP loci, dominant markers were under and overrepresented at 19 and 9 loci, respectively. The distribution of these did not appear randomly distributed, as several linkage groups had no or very few such loci (e.g., Lg1, Lg2, Lg3, Lg11), while others were mainly composed of such loci (e.g., Lg7, Lg20, Lg29).

Several loci could not be mapped into the H1D family as we found 92 loci that were grossly underrepresented in the genotypes. Further investigations revealed that these loci were very close in signal intensity to the threshold of 100 fluorescent units. As such, many of these genotypes went undetected during scoring of the raw data. Because of their instability, they were rejected from preliminary mapping efforts. A remaining 39 of 250 loci were also rejected as they resulted in either substantial linkage conflicts within the preliminary map possibly due to tetraploidy ancestry or sex-specific recombination (see Discussion), or did not show any linkage. Loose linkage relationships observed indicate that proposed linkage groups and accompanying map distances should be considered as 'best' estimates until additional loci can be added.

Maximum likelihood estimates of per individual and per locus hybrid indices.

Out of 1177 AFLP loci in the pure Atlantic and Acadian reference populations, 998 (85%) were useful for calculating estimates of per individual and per locus hybrid indices in the Témiscouata dwarf population. None of these were diagnostic, exhibiting dominant marker frequency differentials between the parental Atlantic and Acadian glacial lineages ranging from of 2 to 44%. Loci that were rejected from the data set included markers that either displayed novel dominants in the Témiscouata population or that did not contribute to inferences of introgressive hybridisation (equal frequencies).

Our approach was designed to show not just the best estimates for hybrid indices, but what the estimates would have been if our dataset was limited to less than $\sim$1000 loci, as is most often the case

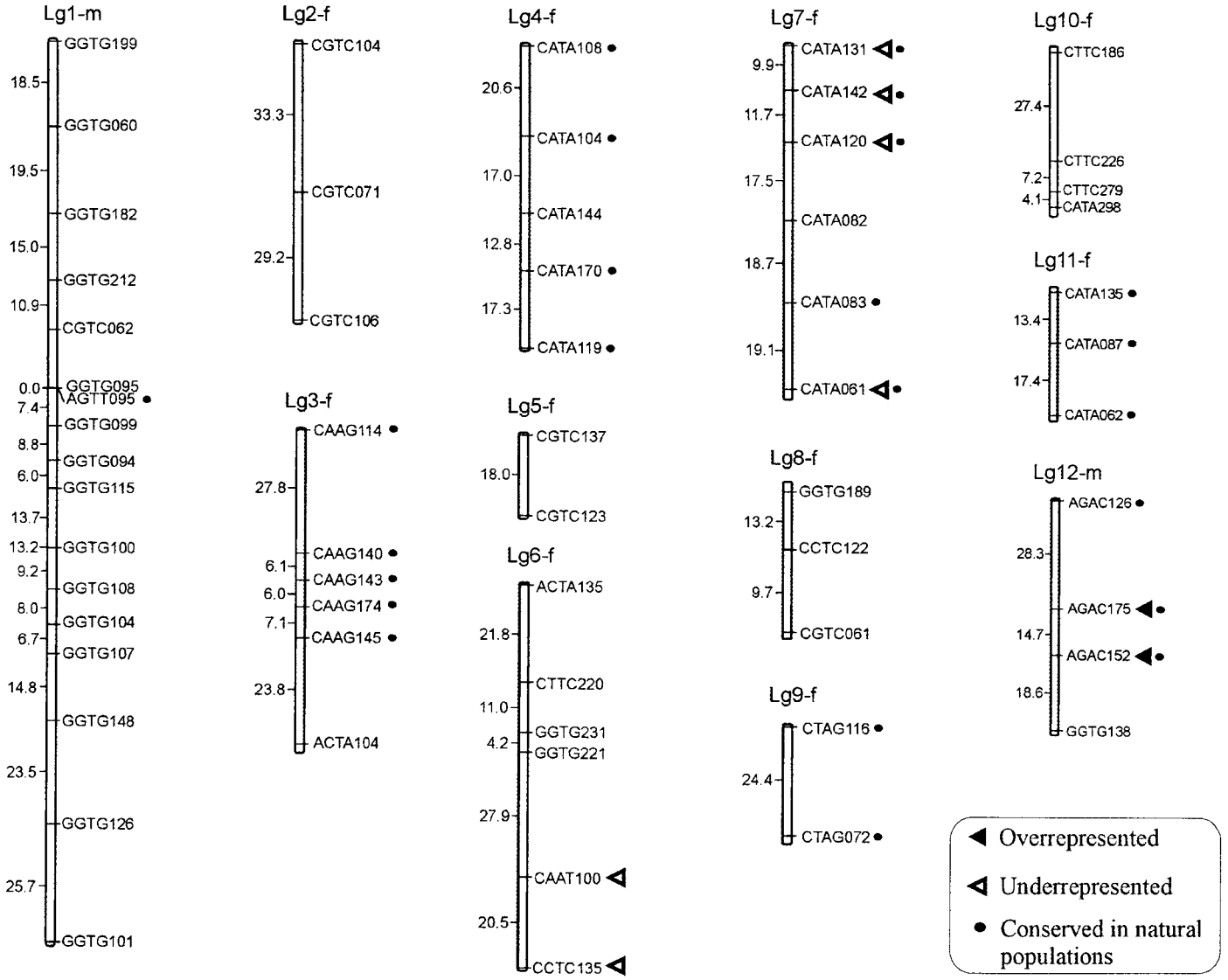

Figure 2. A preliminary sex−specific linkage map of whitefish detailing the segregation of 119 AFLP loci over 29 linkage groups. The linkage groups were given numerical and female/male designations while locus names correspond to the selective amplification primer combo used and locus size (bp). Legend provides details about loci that were significantly under- and over-represented as well as loci conserved in the natural populations.

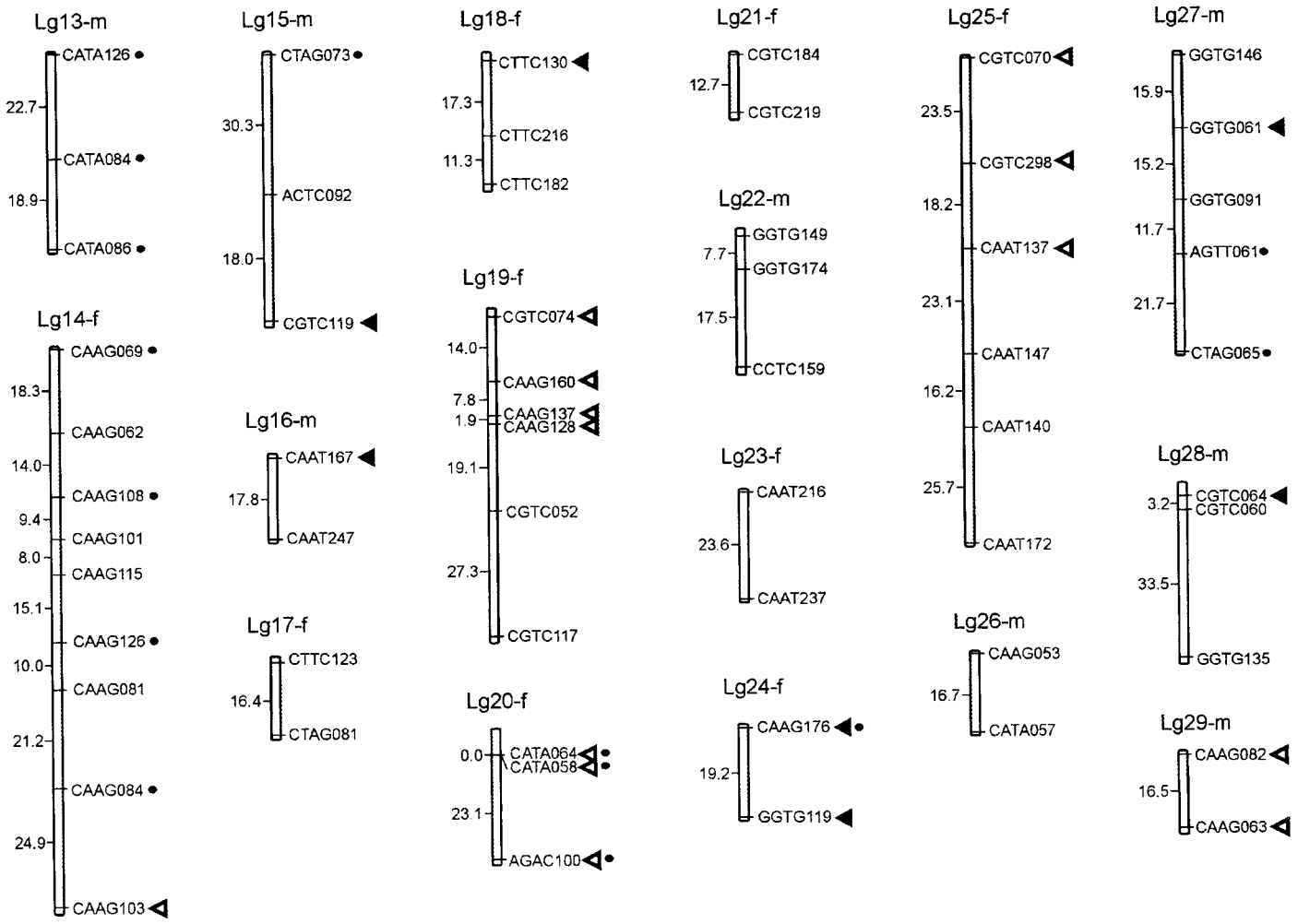

Figure 2. (*continued*)

in such studies. We performed this analysis, first as if we had observed only two of the loci, then as if we had observed four, then six etc. This approach served three purposes: as a form of power analysis of our own dataset, as an indication of the power of inference available to studies in general which use non-diagnostic loci, and as a first stage in differentiating the effects during hybridisation on loci which have differentiated most during isolation of source populations, versus loci which have differentiated least.

The joint maximum-likelihood estimate of per individual hybrid index over all Témiscouata individuals and when all 998 loci were considered was 0.56 (2 units of support: 0.54 − 0.59, Figure 3). This value revealed that the Témiscouata dwarf population was highly introgressed, whereby the genotypic composition of individuals were overall more likely to be of Atlantic than Acadian lineage origin. Note however, that the hybrid index estimate over individuals varied substantially as a function of the number and type (based on frequencies of dominant allele) over which it was made (Figures 4).

The maximum-likelihood estimate of per locus hybrid index over all 998 loci was 0.79 (2 units of support: 0.78 − 0.81), (extreme right, Figure 4). The discrepancy (0.79− 0.56 = 0.23) between both hybrid indices estimates provided a first indication of the effects of linkage disequilibrium, and of the possible role of selective effects in shaping the observed pattern of introgression in the Témiscouata dwarf population. Further indication of this was illustrated by comparing the pattern of both hybrid indices as a function of loci included in the analysis, as well as that of the associated estimate of pairwise linkage disequilibrium. When up to the 10 most divergent loci (5 most characteristic of each lineage) were included in the analysis, both indices were relatively similar, although characterized by large errors around estimates and linkage disequilibrium was near zero (left of Figure 4). When the 10 following most divergent loci were included (going from 10 to 20 loci), a strong discrepancy developed between estimates of both indices, characterized by a decrease of per individual hybrid index strongly biased towards the Acadian lineage. This was also accompanied by a sharp negative increase of the pairwise estimate of linkage disequilibrium. This effect slowly eroded as the number of loci included in the analysis was increased. When more than approximately 100 loci were included, both estimates essentially behaved in the same manner (similar curve shape), although the per individual hybrid index remained approximately 20% lower than that of per locus. In other words, the overall discrepancy between both hybrid indices was caused by a relatively limited number of associated loci. This indicated that these loci are either held in association in a part of the genome with very low recombination relative to other regions, or that their association is maintained by selection.

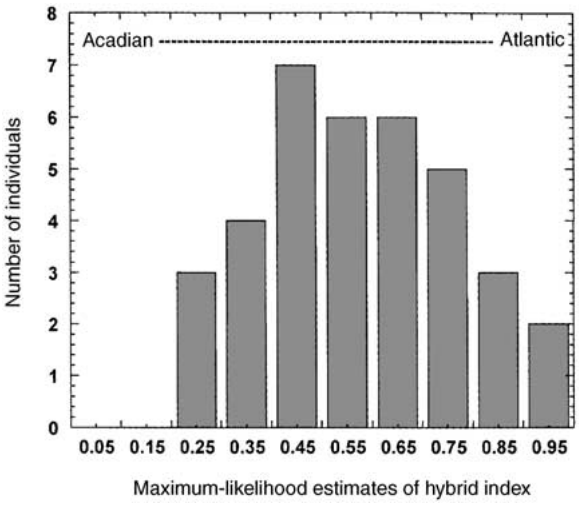

Figure 3. The distribution of per individual hybrid index among individuals over 36 whitefish from the dwarf population of Témiscouata Lake when all 998 loci are considered. The per individual estimates are categorised without reference to their support bounds. The joint hybrid index estimate over all individuals is 0.56 (2 units of support: 0.54–0.59).

Figure 5 which illustrates the deviation of observed marker frequencies for the 998 individual AFLP loci from their expected frequencies given the estimated hybrid index over all individuals of 0.56. A majority of loci ($n = 585$) did not significantly deviate from expectation (their frequency falling within the two units of support of the estimate of expected frequency). On the other hand, a total of 199 and 214 loci were more likely characterized by an allelic composition of Acadian lineage origin (left) or of Atlantic lineage origin (right), respectively. The extremes of this distribution comprise the most likely candidate markers for directional selection during hybridisation. If we set a threshold of 10% deviation outside the two units of support of HI expectation in order to pinpoint the loci significantly most deviating from expectation, a total of 14 loci with an overrepresentation of Atlantic lineage and 12 loci with an overrepresentation of Acadian lineage were retained.

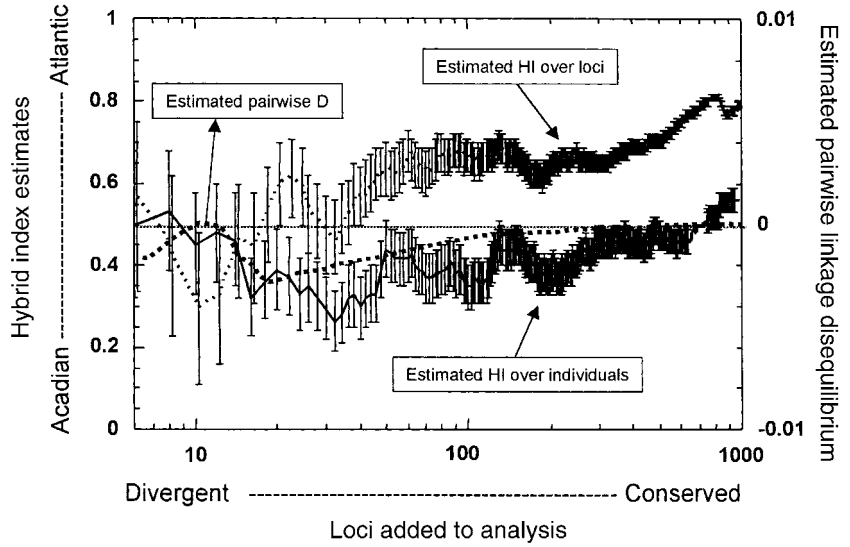

Figure 4. Change in estimates of hybrid index as increasing numbers of non-diagnostic loci are included in the analysis. Hybrid indices over loci and individuals are defined in the text. The number of loci is shown on a log scale. Loci were ordered with the highest frequencies of dominant markers in the Atlantic source population, relative to the Acadian source population, and also with the highest frequencies of dominant markers in the Acadian source population, relative to the Atlantic source population. Loci were then introduced pairwise into the analysis. The first such pair consists of the locus with the highest frequency of dominant markers in the Atlantic lineage relative to the Acadian lineage, together with the locus with the highest frequency of dominant markers in the Acadian lineage, relative to the Atlantic source population.

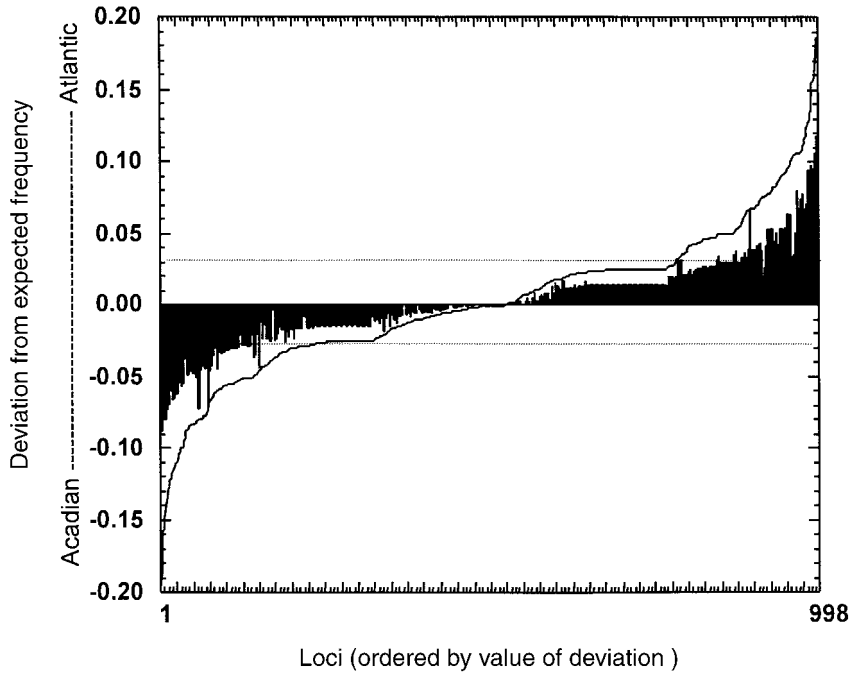

Figure 5. The deviation of observed locus marker frequencies for 998 individual AFLP loci from their expected frequencies given the estimated hybrid index over all individuals (0.56 with two units of support given by horizontal dotted lines). The sigmoid curve shows that some loci are more likely characterized by an allelic composition of Acadian lineage origin (left) or of Atlantic lineage origin (right). The extremes of this graph therefore comprise the most likely candidate loci for directional selection. The histograms in black indicate the minimum deviation with 2 units of support from the expectation. Deviation from expectation is more certain for some loci (those with higher sharp black peaks in the histogram) than others (those with white dents in the histograms).

Table 2. Summary of AFLP loci mapped in the H1D mapping family and genotyped in the introgressed Témiscouata Lake dwarf whitefish population

	Locus number in hybrid index	Locus ID	Linkage group	Locus deviation from average hybrid index (0.56)	Locus differences from zero signed
Acadian	5	CAAG108	14	−0.158	−0.046
	17	CAAG 114	3	−0.119	−0.030
	34	AGAC100	20	−0.100	−0.055
	47	CAAG084	14	−0.084	−0.042
	62	CAAG126	14	−0.079	−0.007
	63	CAAG174	3	−0.079	−0.038
	76	AGTT061	27	−0.068	−0.047
	155	CAAG140	3	−0.044	−0.025
	389	CATA086	13	−0.011	−0.007
	394	CAAG143	3	−0.010	−0.004
	415	CAAG145	3	−0.008	−0.003
	443	CTAG116	9	−0.005	0
	459	CTAG065	27	−0.004	−0.003
	529	CATA087	11	0	0
	612	CAAG176	24	0.018	0.016
	617	CATA135	11	0.020	0.011
	618	CATA142	7	0.020	0.009
	632	CAAG069	14	0.022	0.011
	776	CATA064	20	0.027	0.020
	831	CATA062	11	0.046	0.026
	845	CATA120	7	0.049	0.024
	874	CATA084	13	0.053	0.038
	879	CTAG072	9	0.055	0
	916	CATA170	4	0.075	0.020
	929	CATA061	7	0.080	0.050
	935	CTAG073	15	0.084	0
	947	AGAC152	12	0.092	0.045
	951	AGAC175	12	0.095	0.027
	955	CATA126	13	0.097	0.039
	968	CATA119	4	0.106	0.067
	969	CATA104	4	0.106	0.067
	981	AGAC126	12	0.123	0.054
	982	CATA108	4	0.124	0.089
	995	CATA058	20	0.179	0.114
	996	CATA131	7	0.183	0.098
Atlantic	998	CATA083	7	0.187	0.094

Loci are presented in ascending order concordant with the deviation from expected frequency (last column) presented in Figure 5.

A total of 37 markers identified in the analysis of hybrid indices were conserved in the mapping family and segregated onto the linkage map (Table 2). Some of the strongest candidate loci for directional selection, (e.g., CAAG108, CATA058, CATA083, CATA131) were located on the map. In several instances, we also observed a comparable level of deviation from expected frequency for several loci located on the same linkage group, suggesting either very low recombination in the introgressed population or similar selective effects on loci in the context of hybridisation. For instance, high positive deviations were generally observed for markers of linkage groups Lg4, Lg7 and Lg12, while high negative deviations were observed

for Lg14 loci and small deviations from expectation were generally observed for Lg3 and Lg11 loci. These qualitative observations also provided evidence of linkage group conservation between mapped and natural population loci, and support the possibility that traits elucidated with linkage mapping may be inferred in natural populations. Further qualitative comparisons between the mapping family and the natural population also supported possible selective effects on several linkage groups. Namely, it was noteworthy that several linkage groups, such as Lg7 and Lg12, that were characterized by a majority of loci that significantly deviated from expected Mendelian proportions in the mapping family also showed strong deviations from expectation of neutral introgression in the naturally introgressed population (Table 2, Figure 5).

Discussion

The main objectives of this paper were to demonstrate the amenability of linkage mapping and the comparative analysis of likelihood estimates of hybrid indices using non-diagnostic markers for direct studies into the genetic bases of population divergence and reproductive isolation in the lake whitefish. The generation of a first linkage map constructed from AFLP markers genotyped from a reciprocal hybrid dwarf/normal whitefish mapping family was the first step in achieving this goal. It also represents the first genomic mapping effort for *Coregonus clupeaformis*. Furthermore, the differential behaviour of the maximum-likelihood estimates of per locus and per individual hybrid indices provided indication for the possible role of selection in shaping the observed pattern of introgressive hybridisation in the Témiscouata Lake dwarf population. Notably, this allowed identification of likely candidate markers (or linked loci) for which introgression has likely been constrained (or favoured) by selection in a naturally introgressed whitefish population. This study also revealed that many loci and several linkage groups identified in the mapping family were conserved in the natural populations. Finally, several linkage groups characterized by a majority of loci that significantly deviated from expected genotypic proportions in the mapping family also showed strong deviations from the neutral expectation in the natural population. These first results therefore illustrate the potential of the combined use of genomic/QTL mapping and the comparative analysis of per locus and per individual hybrid indices derived from non-diagnostic markers for testing future hypotheses that can detail the genetic basis of population divergence and reproductive isolation in whitefish.

Linkage mapping

Amplified fragment length polymorphisms (AFLP) were first described by Vos et al. (1995) as a genetic marker system capable of generating DNA fingerprints for any organism regardless of the complexity of the genome (Van der Lee et al., 1997). Since then, AFLPs are emerging as the markers of choice for genetic mapping of organisms for which few codominant markers (e.g., microsatellites) have been developed (Mueller & Wolfenbarger, 1999). The utility of AFLP in generating high density linkage maps has been demonstrated numerous times (Kocher et al., 1998; Young et al., 1998; Kim & Rieseberg 1999). Overall, the choice of genetic marker for mapping is correlated to the amenability the marker has to not only generating a linkage map, but to efficiency and cost of development. With relatively little effort compared to similar development of codominant markers, 250 polymorphic AFLP markers were generated with only 12 selective amplification primer combinations. Furthermore, the multiplexing potential of AFLP allowed routine migrations of all primer pairs for multiple individuals within a single gel.

Past karyotypic investigations of lake whitefish show that the genome is comprised of 40 linkage groups (Allendorf & Thorgaard, 1984). We have thus far successfully identified 29 sex-specific linkage groups covering 1462 cM of the whitefish genome. Some of the AFLP markers were found to deviate significantly from Mendelian segregation ratios. This may reflect a real departure caused by segregation distortion underlying certain genomic traits not yet identified, a limitation of the sample size in the preliminary map, or inherent genomic characteristics of the whitefish genome. Although dominant markers are typically easier to generate for a pedigree, they may have limited usefulness across families unless the marker is cloned and sequenced to establish homology (Knapik et al., 1998). Yet, the apparent conservation of many loci and several linkage groups between the mapping family and the natural populations indicates that this problem can at least partly be circumvented. Future mapping efforts of the whitefish genome will necessarily imply a combined approach

using dominant and co-dominant markers such that the map efforts are amenable across different families and increases the likelihood that overall marker distribution will be evenly spaced (Kocher et al., 1998). Moreover, markers that exhibit Mendelian inheritance by co-dominant segregation will have both parents contributing to the allelic state in offspring, increasing the flexibility of detecting marker segregation in the progeny.

It was not possible to assess quadrivalent formations that appear to be prevalent in male salmonids and that are likely responsible for the reduced recombination contributing to sex-specific differences (Sakamoto et al., 2000). Mutivalent pairings are presumably higher in whitefish as a result of the inherent tetraploid ancestry, but are also likely confined to males (Allendorf & Thorgaard, 1984). Such structural constraints in males implicated in recombination within multivalent pairings may be possible to visualise in future studies through comparative family mapping techniques and gene-centromere mapping approaches. Future mapping efforts in whitefish should therefore continue to incorporate the use of female-based hybrid backcross families, decreasing observable sex-specific segregational differences. With reference to tetraploidy, in order to elucidate the degree of tetraploid genomic regions that persist in the *Coregonus* genome, co-dominant markers may reveal duplicated genes, possibly allowing comparisons with the signal intensity of AFLP loci that flank the duplicated genes to identify duplicated AFLP loci. Further investigations into the tetraploidy ancestry of this species is required as polyploidy may have played an important role in the evolution of whitefish (Allendorf & Thorgaard, 1984).

Many AFLP markers identified in the whitefish linkage map were also homologous to markers originating from natural populations in the introgression analysis. Altogether, when only considering primer pairs used in the analysis of hybrid indices of the Témiscouata Lake population, 90 of 115 markers (78%) were conserved between the mapping family and the wild populations. Reciprocally, 31% of all mapped markers were conserved in the wild populations. These results therefore demonstrate the potential to building comparative links between experimental mapping families and wild populations. The potential for finding even more loci of this nature will obviously grow as mapping efforts develop, and sample sizes are increased in the screening of natural populations.

Comparative analysis of hybrid indices

The comparative analyses of per locus and per individual hybrid indices estimated from non-diagnostic markers using maximum-likelihood provided important insights into the nature of the pattern of introgression observed in the Témiscouata Lake dwarf population. This first showed that the overall genetic composition of that population is the result of pronounced introgression between the Acadian and Atlantic glacial lineages. It also revealed, however, that there was substantial variance in estimates among individuals, with the genetic composition of some being more likely to be strongly influenced by the Acadian lineage (e.g., those with HI of 0.25 on Figure 3), and that of others being more likely to be of almost entirely Atlantic lineage origin (e.g., those with HI of 0.95). This result is in sharp contrast with previous interpretations based on mitochondrial DNA which suggested that the Témiscouata dwarf population was of pure Acadian origin (Pigeon, Chouinard & Bernatchez, 1997). The results of our study were however more comparable to those recently reported from the analysis of six microsatellite loci. Based on the admixture proportions of different lineages computed by means of the weighted least squares method (Long, 1991), Lu, Basley and Bernatchez (2001) estimated that the genetic composition of the Témiscouta Lake population was 63% versus 37% of Acadian and Atlantic lineage origin, respectively. While also indicating pronounced introgression, this latter result is opposite to the hybrid indices estimated in this study over all 998 loci. There are several possible explanations for such discrepancies, the most obvious being the limited number of loci used in the previous analysis. This study clearly showed that the extent of introgression may vary substantially among specific loci (Figure 5) and consequently, that the maximum-likelihood estimate of hybrid index may vary depending on the actual markers used, as well as their numbers. This was best illustrated by the results of Figure 4. When only few of the most divergent loci were used (up to 10), both HI estimates were relatively similar and suggested a more pronounced level of introgression towards the Acadian lineage. Errors around estimates were however quite large for both estimates. When adding the next 10 most divergent loci, a pronounced discrepancy developed between both indices due to the contribution of an apparently limited number of loci to linkage disequilibrium. Only when a very large number of loci were included in the analysis did both estimates re-

spectively stabilise at a given value with narrow errors around them. This clearly shows that any interpretation of overall level of introgression derived from a limited number of markers (either diagnostic or not) should be interpreted cautiously since patterns of introgression are expected to vary substantially across individual loci for deterministic causes, including selection. As such, this study illustrates the benefits of AFLP analysis for studying patterns of introgression. Clearly, the potential for analysing hundreds of loci far outweighed the potential drawbacks associated with the use of dominant markers in estimating hybrid index and inferring the possible causes of the observed pattern of introgression. Admitedly, however, the small sample sizes used in this study, namely for the source populations, also creates uncertainity regarding the estimates we obtained. Clearly, future efforts will necessarily imply increased sample sizes to improve the representativity of the allelic composition of the source populations.

The comparative analysis of the joint estimate of the hybrid index over all individuals with the variance observed among individual loci also allowed us to identify possible candidates for loci under strong selection in the introgressed Témiscouta Lake population. Thus, specific loci and even conserved linkage groups deviated significantly from neutral expectations at the genomic level. In the lake whitefish system, it may be hypothesised that the genetic bases of adaptation and reproductive isolation will implicate two main groups under selection. First, genetic admixture at the loci implicated in existent genetic incompatibility between the glacial lineages (Lu & Bernatchez, 1998) may be counter selected while in contrast, ecological divergence between ecotypes may implicate a selection for certain loci towards the divergence of adaptive niches. Indeed, in the lake whitefish system under a selective hypothesis of neutral expectations for introgressed markers, introgression of loci (and linked markers) contributing to reproductive isolation is expected to be retarded, whereas neutral or positively selected chromosomal segments should introgress at higher frequencies (Rieseberg, Whitton & Gardner, 1999). If the markers have been genetically mapped, as was the case for some markers here, the observed patterns of introgression should make it possible to locate potential chromosomal segments contributing to isolation. Indeed, some markers implicated in the Témiscouata Lake dwarf population that were contenders for selection were also segregated on the linkage map, and in some cases these loci showed pronounced distortion from the expected segregation ratio in the mapping family.

Future directions

Schemske and Bradshaw (1999) proposed two experiments for elucidating the genetic basis of adaptation. First, the genetic basis of adaptive phenotypic traits and intrinsic genomic incompatibilities contributing to speciation must be determined. Second, the 'response' of wild populations to these localised traits must be determined. To date, there have been no studies of this nature on natural animal populations. We have demonstrated that both experiments necessary for determining the genetic basis of adaptation and reproductive isolation can potentially be achieved using a gene mapping approach coupled with comparative analyses of hybrid indices in the lake whitefish system.

Future studies should, therefore, focus on the two experimental domains proposed by Schemke and Bradshaw (1999) in an attempt to determine the genetic architecture of loci affecting these traits. The overall hypothesis to be tested will be that the genetic basis of reproductive isolation in diverging populations can be understood as the genetics of behavioural, morphological, physiological traits directed at trophic resources (Bush & Smith, 1998; Diekmann & Doebeli, 1999; Kondrashov & Kondrashov, 1999). The number and effect of quantitative trait loci examined will reveal whether adaptations and genomic incompatibility are always based on a large number of mutations, each with a small effect, or if small mutations have a large phenotypic effect (Coyne, 1995). It is likely that polygenic adaptations contributing to adaptive traits in our hybrid dwarf/normal crosses will become disassociated by recombination, thus revealing variability in traits that may otherwise not be seen due to natural selection (Rieseberg, 1998b; Jiggens & Mallet, 2000). Therefore, in the hybrid crosses it is likely that the number of genes and their magnitude of effect on genomic regions implicated in intrinsic postzygotic isolation or QTLs associated with selection can be quantified.

The second major hypothesis that requires investigation is if the patterns of introgression in a natural whitefish hybrid zone are determined by the intensity of natural selection acting on genomic traits responsible for phenotypic differentiation and reproductive isolation as measured under laboratory conditions. It may be predicted that the levels of introgression at genetic markers associated with implicated genomic

regions will be lower than that of chromosomal regions not assocciated to phenotypic traits and genetic incompatibility between laboratory whitefish hybrid crosses.

Finally, once the patterns of introgression observed between ecotypes have been measured in the laboratory it will be possible to compare these measures with that among natural populations inhabiting different lakes. A repeatability and correlation of patterns among lakes would offer a very strong demonstration of selection at the genomic level (Rieseberg, Whitton & Gardner, 1999). By sampling individuals from lakes where sympatric whitefish ecotypes coexist, the same maximum likelihood estimates of hybrid indices described here could be used to examine marker loci that are asociated with genomic incompatibility and critical QTL loci in the hybrid mapping families. We are currently implementing such research initiatives with the ultimate goal of elucidating the role of selection in shaping the genetic basis of phenotypic divergence and reproductive isolation in natural populations of whitefish.

Acknowledgements

We would like to thank the whitefish rearing and laboratory technical assistance of S. Higgins, E.-J. Arsenault, P.-P. Dupont, and L. Papillon. We would also like to acknowledge our intellectual debt to G. Lu for his insightful assistance into the development of this research. We are grateful for the constructive comments of two anonymous reviewers. This work was supported by a research grant from the Natural Science and Engineering Research Council (NSERC, Canada) to L.B. and is a contribution to the programme of the Groupe interuniversitaire de recherches Océanographiques du Québec (GIROQ)

References

Allendorf, F.W. & G.H. Thorgaard, 1984. Tetraploidy and the evolution of salmonid fishes, in Evolutionary Genetics of Fishes, edited by B.J. Turner. Plenum Press, New York.

Barton, N.H., 1998. The geometry of adaptation. Nature 395: 751–752.

Barton, N.H. & K.S. Gale, 1993. Genetic analysis of hybrid zones, pp. 13–45 in Hybrid Zones and the Evolutionary Process, edited by R.G. Harrison. Oxford University Press, Oxford.

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

Bernatchez, L. & J.J. Dodson, 1990. Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial DNA restriction analysis. Evolution 44: 1263–1271.

Bernatchez, L. & J.J. Dodson, 1994. Phylogenetic relationships among paleartic and neartic whitefish (*Coregonus* sp.) populations as revealed by mitochondrial DNA variation. Can. J. Fish. Aquat. Sci. 51: 240–251.

Bernatchez, L., J.A. Vuorinen, R.A. Bodaly & J.J. Dodson, 1996. Genetic evidence for reproductive isolation and multiple origins of sympatric trophic ecotypes of whitefish (*Coregonus*). Evolution 50: 624–635.

Bernatchez, L. & C.C. Wilson, 1998. Comparative phylogeography of Nearctic and Palearctic fishes. Mol. Ecol. 7: 431–452.

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

Bradshaw, H.D., S.M. Wilbert, K.G. Otto & D.W. Schemske, 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Minulus*). Nature 376: 762–765.

Bush, G.L. & J. Smith, 1998. The genetics and ecology of sympatric speciation: a case study. Res. Populat. Ecol. 40: 175–187.

Carney, S.E., K.A. Gardner & L.H. Rieseberg, 2000. Evolutionary changes over the fifty-year history of a hybrid population of sunflowers (*Helianthus*). Evolution 54: 462–474.

Cloutier, S., M. Cappadocia & B.S. Landry, 1997. Analysis of RFLP mapping inaccuracy in *Brassica napus* L. Theor. Appl. Genet. 95: 83–91.

Coyne, J.A., 1995. Speciation in monkeyflowers. Nature 376: 726–727.

Darwin, C., 1859. On the Origin of Species By Means of Natural Selection. Random House, USA, Modern Library edn. (1998).

Diekmann, U. & M. Doebeli, 1999. On the origin of species by sympatric speciation. Nature 400: 354–357.

Edwards, A.W.F., 1972. Likelihood. Cambridge University Press, Cambridge.

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., 1925. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Oxford University Press, Oxford.

Hilborn, R. & M. Mangel, 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton.

Huxley, J., 1942. Evolution, the Modern Synthesis. Allen and Unwin, London.

Jiggens, C.D. & J. Mallet, 2000. Bimodal hybrid zones and speciation. Trends Ecol. Evol. 15: 250–255.

Kim, S.-C. & L.H. Rieseberg, 1999. Genetic architecture of species differences in annual sunflowers: implications for adaptive trait introgression. Genetics 153: 965–977.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge.

Knapik, E.W., A. Goodman, M. Ekker, M. Chevrette, J. Delgado, S. Neuhaus, N. Shimoda, W. Driever, M.C. Fishman & H.J. Jacob, 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). Nature Genet. 18: 338–343.

Kocher, T., W.-J. Lee, H. Sobolewska, D. Penman & B. McAndrew, 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). Genetics 148: 1225–1232.

Kondrashov, A.S. & F.A. Kondrashov, 1999. Interactions among quantitative traits in the course of sympatric speciation. Nature 400: 351–354.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln & L. Newburg, 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181.

Long, J.C., 1991. The genetic structure of admixed populations. Genetics 127: 417–428.

Lu, G., D. Basley & L. Bernatchez, 2001. Complex evolutionary history of lake whitefish (*Coregonus clupeaformis*) in northeastern North America assessed with mtDNA and microsatellite markers. Mol. Ecol. 10: 965–985.

Lu, G. & L. Bernatchez, 1998. Experimental evidence for reduced hybrid viability between dwarf and normal ecotypes of lake whitefish (*Coregonus clupeaformis*, Mitchill). Proc. Roy. Soc. London B 265: 1025–1030.

Lu, G. & 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., 1942. Systematics and the Origin of Species from the Viewpoint of a Zoologist. Columbia University Press, New York.

Mueller, U. & L. Wolfenbarger, 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. 14: 389–394.

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

Orr, H.A. & T.B. Smith, 1998. Ecology and speciation. Trends Ecol. Evol. 13: 502–506.

Ott, J., 1991. Analysis of Human Genetic Linkage. The Johns Hopkins University Press, Baltimore, Revised edn.

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

Rice, W.R., 1989. Analysing tables of statistical tests. Evolution 43: 223–225.

Rieseberg, L.H., 1998a. Genetic mapping as a tool for studying speciation, pp. 459–487 in Molecular Systematics of Plants, edited by D.E. Soltis, P.S. Soltis & J.J. Doyle. Chapman and Hall, New York, 2nd edn.

Rieseberg, L.H., 1998b. Molecular ecology of hybridization, pp. 243–265 in Advances in Molecular Ecology, edited by G.R. Carvalho. IOS Press.

Rieseberg, L.H., S.J.E. Baird & A.M. Desrochers, 1998. Patterns of mating in wild sunflower hybrid zones. Evolution 52: 713–726.

Rieseberg, L.H., H. Choi, R. Chan & C. Spore, 1993. Genomic map of a diploid hybrid species. Heredity 70: 285–293.

Rieseberg, L.H. & C.R. Linder, 1999. Hybrid classification: insights from genetic map-based studies experimental hybrids. Ecology 80: 351–370.

Rieseberg, L.H., J. Whitton & K.A. 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. & D. Schluter, 2000. Natural selection and the evolution of adaptive genetic variation in northern freshwater fishes, pp. 65–94 in Adaptive Genetic Variation in the Wild, edited by T.A. Mousseau, B. Sinervo & J.A. Endler. Oxford University Press, New York.

Sakamoto, T., R.G. Danzmann, K. Gharbi, P. Howard, A. Ozaki, S.-K. Khoo, R.A. Woram, N. Okamoto, M.M. Ferguson, L.-E. Holm, R. Guyomard & B. Hoyheim, 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*). Genetics 155: 1331–1345.

Sawamura, K., A. Davis & C.-L. Wu, 2000. Genetic analysis of speciation by means of introgression into *Drosophila melanogaster*. Proc. Nat. Acad. Sci. 97: 2652–2655.

Schemske, D.W. & H.D. Bradshaw, 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). Proc. Nat. Acad. Sci. 96: 11910–11915.

Schluter, D., 1996a. Ecological speciation in postglacial fishes. Philos. Trans. Roy. Soc. London B 351: 807–814.

Schluter, D., 1996b. Ecological causes of adaptive radiation. Am. Naturalist 148: 540–564.

Schluter, D., 1998. Ecological causes of speciation, pp. 114–129 in Endless Forms: Species and Speciation, edited by D.J. Howard & S.H. Berlocher. Oxford University Press, New York.

Schluter, D., 2000. Ecological character displacement in adaptive radiation. Am. Naturalist 156: S4–S16.

Scott, W. & E. Crossman, 1973. Freshwater fishes of Canada. Fish. Res. Board Can. 184: 269–277.

Tanksley, S., 1993. Mapping polygenes. Ann. Rev. Genet. 27: 205–233.

Taylor, E. & J. McPhail, 1999. Evolutionary history of an adaptive radiation in species pairs of three spine sticklebacks (*Gasterosteus*): insights from mitochondrial DNA. Biol. J. Linnean Soc. 66: 271–291.

Turgeon, J., A. Estoup & L. Bernatchez, 1999. Species flock in the North American Great Lakes: molecular ecology of Lake Nipigon ciscoes (Teleostei: Coregonidae: Coregonus). Evolution 53: 1857–1871.

Van der Lee, T., I. De Witte, A. Drenth, C. Alfonso & F. Govers, 1997. AFLP linkage map of the oomycete Phytophthora infestans. Fungal Genet. Biol. 21: 278–291.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper & M. Zabeau, 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. 23: 4407–4414.

Wolfram, S., 1992. Mathematica: A System for Doing Mathematics by Computer. Addison-Wesley Publishing Company, Redwood City.

Young, W., P. Wheeler, V. Coryell, P. Keim & G. Thorgaard, 1998. A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148: 839–850.