Journal of Fish Biology (2003) **62,** 517–533
doi:10.1046/j.0022-1112.2003.00032.x, available online at http://www.blackwell-synergy.com

Microgeographic population structure of brook charr: a comparison of microsatellite and mark-recapture data

B. K. ADAMS* AND J. A. HUTCHINGS

Department of Biology, Dalhousie University, Halifax, Nova Scotia, B3H 4J1, Canada

(Received 22 September 2001, Accepted 20 November 2002)

Polymorphism at five microsatellite genetic markers (genotyped $n = 496$) and mark-recapture tagging data (tagged $n = 9813$) were used to define the population structure of brook charr, *Salvelinus fontinalis* from the Indian Bay watershed, Newfoundland, Canada. Despite the absence of physical barriers to migration among lakes, both genetic and tagging data suggest that brook charr in each lake represent reproductively isolated populations. Exact tests comparing allele frequencies, θ (global value = 0·063), *R*st (global value = 0·052), individual assignment tests, and Nei's genetic distance provided congruent estimates of population subdivision in agreement with the tagging data (only 2·2% of recaptures were lake-to-lake). The genetic structure of the brook charr populations corresponded with the geographic structure of the drainage basin on a qualitative level, although linear distance over water was not significantly correlated with the tagging data or the genetic distance measures. The agreement between the tagging and the genetic data suggest that microsatellite markers can be useful tools for defining real biological units. The results also suggest that brook charr exhibit microgeographic population structure at the watershed scale, and that this is the scale at which conservation and management of this salmonid might best be implemented.

Key words: brook charr; genetic distance; genetic structure; microsatellite; migration rate; tagging.

INTRODUCTION

To identify the fundamental principles governing plant and animal distribution and abundance, reliable information on population structure is required. Studies of population differentiation on fishes have included research on migration patterns (Groot *et al.*, 1989), identification and distribution of separate stocks (Lester *et al.*, 1988; McConnell *et al.*, 1995; Bentzen *et al.*, 1996), and the differentiation of populations of wild and domestic origin (Danzmann *et al.*, 1991). Various techniques and approaches have been used to address these questions, *e.g.* parasite biotags (Groot *et al.*, 1989), artificial tagging (Healey, 1978), morphometrics and meristics (Wilder, 1952) and genetic markers (Danzmann *et al.*, 1991; Bentzen *et al.*, 1996; Angers & Bernatchez, 1998). These techniques have allowed researchers to identify and to assess the likelihood of movement among natural

*Author to whom correspondence should be addressed. Tel.: +1 709 256 6133; fax: +1 902 494 3736; email: bkadams@dal.ca

populations, thus providing information critical to the understanding of the fundamental biological principles underlying patterns of distribution and abundance.

Salmonidae comprise one of the most interesting groups of animals for which population structure has been examined. For example, Pacific salmon *Oncorhynchus* spp. have been shown to form reproductively isolated populations among rivers (Brykov *et al.*, 1996), within tributaries of the same catchment (Small *et al.*, 1998), and even between odd and even years within the same tributary [*e.g.* the two-year life cycle of pink salmon *Oncorhynchus gorbuscha* (Walbaum)] (Beacham *et al.*, 1996). Atlantic salmon *Salmo salar* L. and brown trout *Salmo trutta* L. populations also exhibit geographical and temporal subdivisions within rivers (Hansen & Loeschcke, 1996; O'Connell *et al.*, 1996).

Brook charr *Salvelinus fontinalis* (Mitchill) are ubiquitous on the island of Newfoundland. Despite their abundance, little is known about the population structuring of this salmonid in Newfoundland watersheds. Hutchings (1993, 1994, 1996), Adams (1999), and van Zyll de Jong *et al.* (1999) have greatly expanded the data available on life-history variation in Newfoundland brook charr. A comparison of the populations described by Hutchings (1993) and Adams (1999) demonstrates the substantial amount of life-history variation that occurs both within and among watersheds. It is probable that a proportion of this life-history variation can be attributed to local adaptation, suggesting that brook charr may exhibit microgeographical population structuring. A substantial body of work in Québec, Canada, suggests that brook charr exhibit strong microgeographical population structuring in the watersheds of southern Québec (Angers *et al.*, 1995, 1999; Angers & Bernatchez, 1998). Microsatellite markers and mark-recapture programmes can provide the data required to define the scale of brook charr population structure in Newfoundland.

Microsatellite genetic markers are commonly used in contemporary fine-scale salmonid population structure analyses. (Ferguson & Danzmann, 1998). During the past decade, several statistical tools have been adapted (Nei *et al.*, 1983; Weir & Cockerhan, 1984) or created (Goldstein *et al.*, 1995; Slatkin, 1995) to analyse microsatellite data. The dependability of these metrics has been assessed with empirical data (Paetkau *et al.*, 1997; Ruzzante, 1998) and with analytical or simulated approaches (Goldstein *et al.*, 1995; Shriver *et al.*, 1995; Slatkin, 1995; Takezaki & Nei, 1996; Gagiotti *et al.*, 1999). Comparative studies of population structure described by both microsatellite data and tagging data, however, are uncommon.

In addition to discriminating among population units, both microsatellites and tagging data have also been used to estimate number of migrants (N_m) among populations. For example, Gagiotti *et al.* (1999) compared the performance of θ and *R*st in estimating N_m with microsatellite data. They concluded that the stepwise mutation model (SMM)-based *R*st performed best under 'ideal' conditions (*i.e.* large sample sizes and high number of loci scored), while the infinite alleles model (IAM)-based θ performed best under 'realistic' conditions (*i.e.* small sample size and low number of loci scored).

Irrespective of the mutation model underlying indirect estimates of gene flow based on θ and *R*st, these metrics include implicit assumptions about the nature of the populations sampled: (1) selection is absent (*i.e.* genetic markers are neutral); (2) mutation rates are low; (3) there is equilibrium between migration

and genetic drift; (4) population sizes are static. Whitlock & McCauley (1999) describe the effect of these assumptions, or violation thereof, on the reliability of estimates of N_m based on θ. They concluded that this metric was reliable only for broad-scale applications (*e.g.* comparing gene flow among species). Despite this caveat, genetic-based estimates of N_m continue to provide a basis for fisheries and conservation management decisions.

Tagging programmes are also an effective tool for discerning population structure and individual movement patterns (Hilborn, 1990; Myers *et al.*, 1997). Tagging data cannot provide estimates of reproductive relationships among potentially distinct populations. Interpretation of tagging data is limited to the time frame during which the data were collected, whereas genetic data are a product of historical events and relationships. There are also assumptions implicit in the use of tagging data to quantify migration rates among populations: (1) all animals have an equal chance of capture, (2) tagging does not affect catchability or fish behaviour (*e.g.* induce migration), and (3) animals do not lose tags or experience tagging-induced mortality during the study period. Tagging data have a long history of use for fisheries management and conservation, even though at least one of these assumptions is often violated (Krebs, 1998). A study including both methodologies can thus provide a valuable comparison of the strengths and weaknesses of these commonly used techniques.

The primary objective of the present study was to compare indirect estimates of gene flow and migration, using microsatellites and mark-recapture data. As far as is known, such a comparison has not been made in brook charr. A secondary objective was to describe the population structure of wild brook charr populations in a large postglacial watershed. Specifically, whether or not brook charr inhabiting lakes within a single watershed can represent distinct, reproductively isolated populations was examined.

MATERIALS AND METHODS

Right pelvic finclips were collected from brook charr captured by randomly set nets in the littoral zone of eight lakes in the Indian Bay watershed (Fig. 1) and preserved in 95% ethanol. Samples were collected during June and July and during September and October in 1997 and during September and October in 1998 (Table I). The bulk of the sampling was undertaken during September and October when brook charr are spawning and populations were expected to be most discrete. Genetic sampling was independent of the tagging study (*i.e.* none of the individuals tagged were sampled for the genetic analysis).

GENETIC METHODS

Sfo-18 and Sfo-23 (Angers *et al.*, 1995), Omy-38A, Omy-38B and Omy-105 (unpubl. data) were chosen as microsatellite markers, based on their variability and the technical convenience of their use, to quantify the amount of genetic structuring of Indian Bay brook charr populations. Extraction of DNA and PCR amplification of the alleles at each locus were carried out using standard techniques (McConnell *et al.*, 1995).

Genetic variation was quantified in terms of observed and expected heterozygosity, allele numbers per locus and allele frequencies. Bonferroni corrections for table-wide levels of significance were made where appropriate (Rice, 1989). Heterozygosity at each locus was calculated using the descriptive statistics option in TFPGA 1.3 (Tools For Population Genetics Analysis; Miller, 1997). Fisher's exact tests were used to test for departures from

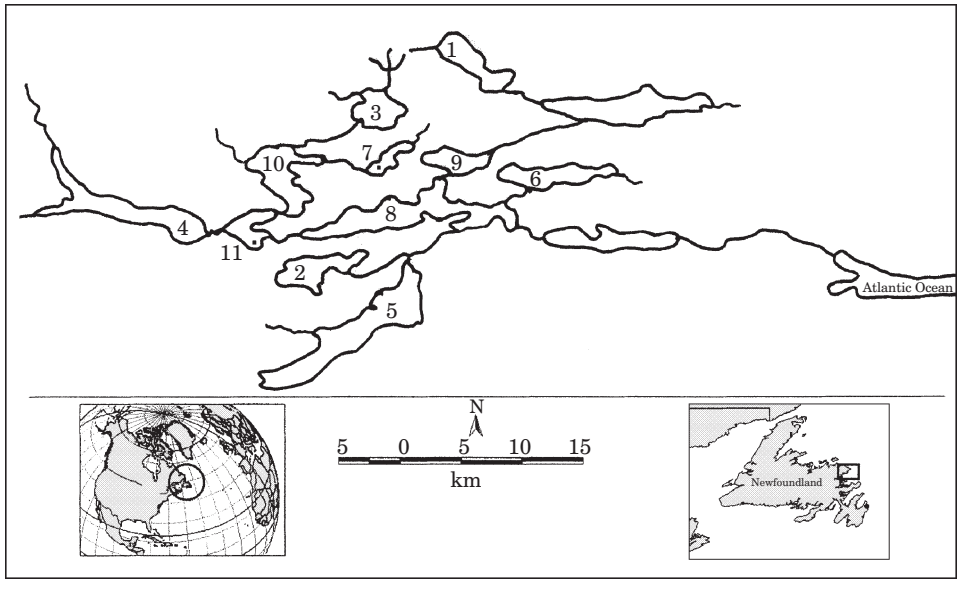

FIG. 1. Genetic and tagging sampling sites for brook charr within the Indian Bay watershed, Newfoundland. (1) Moccasin Pond, (2) Southern Pond, (3) Alleys Pond, (4) Fourth Pond, (5) Wings Pond, (6) Fourmile Pond, (7) Skippers Pond, (8) Second Pond, (9) Little Bear Cave, (10) Back-up and (11) Third Pond.

Hardy–Weinberg (H–W) proportions (Raymond & Rousset, 1995). TFPGA 1.3 employs a Markov chain method (Guo & Thompson, 1992) with 1000 iterations to test for excess or deficiency of heterozygotes. Arlequin 1.1 (Schneider *et al.*, 1998) was used to test for linkage disequilibrium.

The level of population differentiation was estimated using exact tests for pair-wise comparisons of all spatial and temporal sample sets at each locus and across all loci. Pair-wise exact tests for population differentiation were implemented by TFPGA 1.3 as described above for tests of H–W equilibrium. Spatial and temporal pair-wise comparisons were also made using θ (Weir & Cockerhan, 1984), a metric analogous with *F*st. TFPGA 1.3 was

TABLE I. Collection dates and sample sizes of brook charr from each site within the Indian Bay watershed

		September to October	
Sampling site	June to July 1997	1997	1998
Alleys Pond	—	61	53
Fourmile Pond	—	—	52
Fourth Pond	—	50	—
Moccasin Pond	—	30	—
Second Pond	—	30	—
Skippers Pond	—	31	—
Southern Pond	45	44	52
Wings Pond	—	48	—

used to calculate θ within and among loci, individuals and populations. CI were estimated by a bootstrap procedure with 1000 iterations. *R*st was calculated for each pair-wise comparison using *R*st-Calc (Goodman, 1997). θ, *R*st and Slatkin's private allele method (via Genepop) (Slatkin, 1985) were used to estimate N_m among populations and for the watershed overall. Individual assignment tests were carried out using the Geneclass software (Cornuet *et al.*, 1999). The 'leave one out' procedure along with the Bayesian likelihood estimation was used to assign individuals directly to the most probable population of origin. The simulation method in Geneclass was used to test the probability of rejecting individuals from each population. The number of simulated individuals was set at 10000 for each population with a 0·05 probability of rejection. Nei's genetic distance (D_A) (Nei *et al.*, 1983) was calculated among populations. Paetkau *et al.* (1997) suggested D_A maybe the most appropriate and conservative approach when drift is the dominant force underlying population differentiation (Takezaki & Nei, 1996). Multi-dimensional scaling plots based on θ, *R*st and D_A were used to describe the relationship among populations. A Mantel test was used to test for correlations between the matrices of genetic metrics, geographical distance (linear distance by water) and tagging data. Multi-year samples from Southern Pond and Alleys Pond were pooled within lake for the aforementioned analysis.

TAGGING DATA

Trap nets were used to sample brook charr for tagging. The Indian Bay Ecosystem Corporation (IBEC) surveys 12–14 of the watershed's 16 lakes each year. The mark-recapture survey focuses on estimating population abundance, thus the sample sizes and effort were dictated by an effort to reduce the error in population size estimates creating a substantial disparity in effort among lakes. Data from 11 of these lakes are presented in the present study. Trap net mark-recapture surveys in individual lakes are based on eight to 140, 24 h trap net sets. The order in which lakes were sampled each year was random, as were the lakes that were chosen for simultaneous sampling. At least two lakes were sampled simultaneously at all times during the tagging study. Differential selectivity does not appear to affect the catch rates of brook charr from ages 2–7 years (Adams, 1999). Deployment sites were set 100 m apart along the perimeter of each lake. Daily trap net sets represented a randomly selected subset of these permanent deployment sites. Trap net set duration was *c.* 24 h. For each net set, the catch was enumerated, and all brook charr were anaesthetized, measured (fork length, L_F, to the nearest 0·1 mm), weighed (nearest 0·1 g), tagged with an individually numbered Floy tag anterior to the dorsal fin, and released. Individual brook charr handling times were usually <1 min.

The percentage of recaptures that were not in the initial tagging lake was calculated for each lake surveyed. Given the uneven distribution of effort and relatively low rate of lake-to-lake recaptures, quantitative estimates of N_m among lakes based on the tagging data could not be differentiated from zero.

RESULTS

GENETIC VARIABILITY IN INDIAN BAY BROOK CHARR

Number of alleles per locus averaged 11, ranging from two at Omy-105 to 25 at Sfo-23 (Table II), and varied by as much as three-fold among sites. Significant departures from Hardy–Weinberg proportions at an 0·05 α-level were detected in only two of a possible 55 cases, significantly less than expected by chance alone ($P = 0{\cdot}001$). Expected heterozygosity averaged 0·69 across all sites and loci, ranging from 0·27 for Omy-105 in Wings Pond to 0·90 for Sfo-23 in Skippers Pond (Table II).

Pair-wise exact tests revealed significant ($P < 0{\cdot}001$ for all pair-wise tests) differences in global allele frequencies among all geographical samples. The overall

TABLE II. Range in allele size (bp), allele number, expected heterozygosity and number of samples successfully amplified for each site and locus

Locus		Sfo-18	Sfo-23	Omy-38a	Omy-38b	Omy-105
All populations	Range in allele size	169–205	141–203	94–104	132–214	105–109
	Number of alleles	17	25	6	17	2
	Heterozygosity	0·73	0·89	0·53	0·61	0·42
	Samples amplified	465	439	479	487	464
Moccasin Pond	Range in allele size	173–197	145–197	94–100	138–214	105–109
	Number of alleles	6	13	4	8	2
	Heterozygosity	0·74	0·88	0·53	0·73	0·44
	Samples amplified	28	30	30	30	28
Southern Pond	Range in allele size	173–197	145–203	98–104	138–204	105–109
	Number of alleles	8	16	3	6	2
	Heterozygosity	0·51	0·84	0·48	0·48	0·29
	Samples amplified	133	129	133	139	132
Fourth Pond	Range in allele size	173–205	145–193	96–104	138–214	105–109
	Number of alleles	12	15	4	7	2
	Heterozygosity	0·87	0·89	0·58	0·48	0·38
	Samples amplified	50	48	50	50	50
Alleys Pond	Range in allele size	173–201	145–193	94–104	134–210	105–109
	Number of alleles	10	18	6	9	2
	Heterozygosity	0·79	0·89	0·45	0·62	0·42
	Samples amplified	107	99	112	113	102
Wings Pond	Range in allele size	169–197	141–201	98–100	132–210	105–109
	Number of alleles	9	14	2	8	2
	Heterozygosity	0·73	0·89	0·41	0·63	0·27
	Samples amplified	48	41	48	48	47
Fourmile Pond	Range in allele size	173–199	145–191	96–100	138–206	105–109
	Number of alleles	7	13	3	7	2
	Heterozygosity	0·74	0·84	0·52	0·67	0·51
	Samples amplified	44	40	44	46	45
Second Pond	Range in allele size	173–201	145–197	98–100	138–206	105–109
	Number of alleles	8	13	2	4	2
	Heterozygosity	0·60	0·87	0·40	0·59	0·45
	Samples amplified	28	25	31	31	31
Skippers Pond	Range in allele size	173–203	145–193	98–104	136–206	105–109
	Number of alleles	9	16	3	9	2
	Heterozygosity	0·82	0·90	0·55	0·74	0·50
	Samples amplified	25	27	29	28	28

θ-value was 0·063, indicating a significant ($P < 0·001$) level of population differentiation. As a relative measure, the θ-values indicate that Fourth Pond and Second Pond were the most divergent brook charr populations ($\theta = 0·132$), while Alleys Pond and Moccasin Pond brook charr were most similar ($\theta = 0·015$) (Table III). All pair-wise θ-values were significantly different from zero after Bonferroni corrections, congruent with the results of the pair-wise exact tests.

*R*st was calculated to be 0·052 for all populations over all loci and was significantly different from zero ($P < 0·001$). Pair-wise *R*st-values were similar to the θ-values (simple linear regression, $r^2 = 0·53$, $P < 0·001$) (Table III). Fourth Pond

TABLE III. Pair-wise θ- and *R*st-values for all geographical samples from the Indian Bay watershed. *R*st-values are above the diagonal. NS, values not significantly different from zero after Bonferroni correction ($P = >0{\cdot}001$)

Lake	Moccasin	Southern	Fourth	Alleys	Wings	Fourmile	Second	Skippers
Moccasin Pond	—	0·032	0·028	0·011 (NS)	0·034	0·023	0·104	0·001 (NS)
Southern Pond	0·088	—	0·030	0·043	0·009 (NS)	0·086	0·092	0·071
Fourth Pond	0·038	0·075	—	0·001 (NS)	0·042	0·094	0·127	0·040
Alleys Pond	0·015	0·076	0·016	—	0·057	0·071	0·141	0·021
Wings Pond	0·067	0·050	0·075	0·072	—	0·10	0·116	0·74
Fourmile Pond	0·046	0·093	0·068	0·051	0·081	—	0·079	0·048
Second Pond	0·097	0·109	0·132	0·118	0·099	0·054	—	0·103
Skippers Pond	0·023	0·071	0·033	0·030	0·056	0·033	0·058	—

and Second Pond populations were again among the most divergent (Rst = 0·1269), second only to the divergence between Alleys Pond and Second Pond (Rst = 0·1405). Pair-wise Rst-values for four population comparisons, including that between Moccasin and Alleys Ponds, did not differ significantly from zero (Table III).

D_A (Nei *et al.*, 1983) (Fig. 2) provides estimates of genetic distance that correspond well with the results presented thus far; they are similar to the pair-wise θ-values (simple linear regression, $r^2 = 0{\cdot}87$, $P < 0{\cdot}001$) with Fourth Pond and Second Pond brook charr being the most divergent ($D_A = 0{\cdot}277$), and Moccasin and Alleys Pond populations being the least divergent ($D_A = 0{\cdot}030$). All distances were significantly different from zero after Bonferroni corrections ($P < 0{\cdot}001$).

Individual assignment tests were able to assign 68·3% of the individuals back to their source population (Table IV). The proportion of individuals assigned directly to their source population ranged from 0·58 in Second Pond to 0·80 in Skippers Pond. The proportion of individuals assigned to the incorrect population did not appear to be correlated with the D_A among populations or the number of tagged recaptures among lakes. The proportion of individuals rejected from non-source populations ranged from 5 to 86% (Table IV). The mean proportion rejected from non-source populations was 46%. Rejection of individuals from their source population was <1% in all cases (Table IV). Only five individuals were rejected from all populations. The five unassigned individuals were from Alleys Pond, Fourmile Pond, Moccasin Pond, Second Pond and two from Wings Pond.

The multi-dimensional scaling plots based on D_A, θ-values, and Rst relate well to the geography of the watershed (Figs 1 and 2). Southern Pond and Wings Pond are found on one tributary of the watershed and consistently cluster together. Fourth Pond and Alleys Pond are also found on one tributary and cluster together consistently. An exception to this generally clear pattern is the Moccasin Pond population, which consistently clusters with Alleys Pond and Fourth Pond even though it is found on a separate tributary of the watershed. Mantel tests comparing geographical distance and genetic distances among lakes indicated no significant correlation.

Pair-wise comparisons of allele frequencies, using exact tests, showed no significant difference in allele frequencies between years within lakes ($P = 0{\cdot}66$ for Southern Pond and 0·14 for Alleys Pond). Pair-wise θ-values and Rst-values were also calculated for the year-to-year samples. θ-values were 0·005 and 0·003 for Southern Pond and Alleys Pond, respectively, neither of which was significantly different from zero ($P = 0{\cdot}88$ and 0·18). Rst-values were 0·001 and 0·091 for Southern Pond and Alleys Pond, respectively, neither of which was significantly different from zero ($P = 0{\cdot}54$ and 0·11).

Tagging data collected during the trap net population surveys indicated very low levels of movement among lakes. In the lakes sampled for the tagging study, 9813 brook charr were tagged over three sampling seasons. Five hundred and fifty-six of these brook charr were recaptured (5·6%). Only 12 of these 9813 individuals were recaptured in lakes in which they were not tagged (Tables V and VI). These 12 individuals represent only 0·02% of the 9813 brook charr tagged, or 2·2% of the fish recaptured. Only seven of a possible 55 lake-to-lake migration destinations had recaptures of migrants (Table VI), reflecting very limited movement by brook charr among lakes.

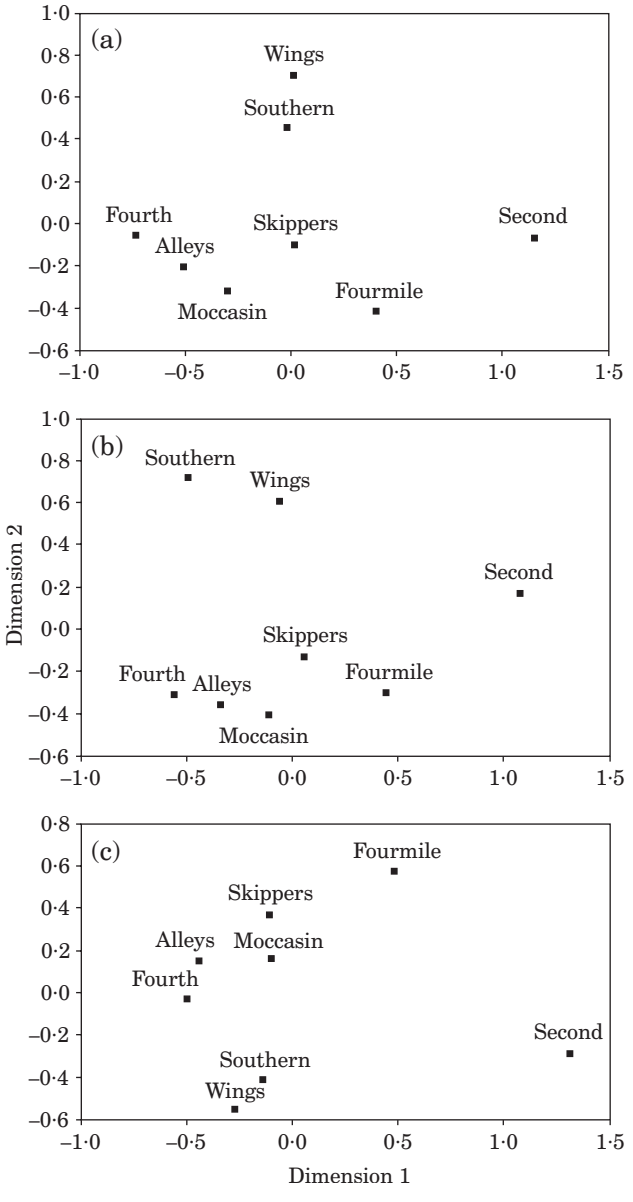

FIG. 2. Multi-dimensional scaling plots based on Nei's genetic distance (a), θ (b), and *R*st (c) proximity matrices for all populations. The proportions of variance in the data accounted for by scaled distances were 0·90, 0·863 and 0·874, respectively.

DISCUSSION

Microsatellite and tagging data suggest that brook charr in the Indian Bay watershed are subdivided by lake into reproductively isolated populations. The exact tests comparing allele frequencies, θ-values, *R*st, D_A and the individual assignment tests, all provide similar and strong evidence of differentiation. Tagging

TABLE IV. The proportion of individuals assigned correctly to their source population and those assigned to incorrect populations for each sample. The proportion rejected from each population based on the Geneclass simulation procedure with a 0·05 rejection probability is given in parentheses

Source population	Moccasin	Southern	Fourth	Alleys	Wings	Fourmile	Second	Skippers
Moccasin Pond	**0·73 (0·03)**	0·07 (0·74)	0·00 (0·50)	0·07 (0·50)	0·03 (0·57)	0·03 (0·60)	0·03 (0·37)	0·03 (0·70)
Southern Pond	0·02 (0·57)	**0·70 (0·07)**	0·00 (0·23)	0·09 (0·23)	0·04 (0·16)	0·00 (0·34)	0·07 (0·05)	0·09 (0·41)
Fourth Pond	0·08 (0·78)	0·00 (0·86)	**0·74 (0·02)**	0·09 (0·46)	0·02 (0·72)	0·04 (0·76)	0·02 (0·42)	0·00 (0·80)
Alleys Pond	0·05 (0·58)	0·10 (0·59)	0·11 (0·18)	**0·73 (0·02)**	0·04 (0·46)	0·03 (0·49)	0·03 (0·13)	0·03 (0·53)
Wings Pond	0·02 (0·25)	0·17 (0·42)	0·02 (0·25)	0·04 (0·25)	**0·67 (0·08)**	0·02 (0·32)	0·02 (0·27)	0·02 (0·46)
Fourmile Pond	0·02 (0·33)	0·00 (0·67)	0·06 (0·33)	0·06 (0·50)	0·02 (0·46)	**0·69 (0·04)**	0·02 (0·23)	0·14 (0·46)
Second Pond	0·06 (0·66)	0·06 (0·77)	0·03 (0·58)	0·10 (0·48)	0·03 (0·71)	0·03 (0·61)	**0·58 (0·03)**	0·10 (0·65)
Skippers Pond	0·00 (0·40)	0·06 (0·63)	0·00 (0·80)	0·10 (0·23)	0·03 (0·33)	0·00 (0·33)	0·03 (0·13)	**0·80 (0·00)**

TABLE V. Numbers of brook charr tagged and released, and sampling effort, in lakes of the Indian Bay watershed between 1996 and 1998. Third Pond, Back-up Pond and Little Bear Cave Pond were not sampled for the genetic study

Lake	Effort (net sets) (1996)	Number tagged (1996)	Number captured (1996)	Effort (net sets) (1997)	Number tagged (1997)	Number captured (1997)	Effort (net sets) (1998)	Number tagged (1998)	Number captured (1998)
Moccasin Pond	20	90	0	—	—	—	12	36	0
Southern Pond	20	—	—	20	328	18	8	110	2
Fourth Pond	20	159	0	20	419	23	20	442	22
Alleys Pond	100	205	8	140	1681	89	20	331	13
Wings Pond	20	28	0	20	—	—	20	146	2
Fourmile Pond	100	81	0	140	1002	152	56	78	0
Second Pond	100	248	3	140	472	22	84	155	4
Back-up Pond	20	310	1	20	376	3	8	252	7
Little Bear Cave Pond	100	343	34	140	407	58	70	101	4
Third Pond	100	816	59	20	128	0	42	301	6
Skippers Pond	20	—	—	20	412	22	20	356	3

TABLE VI. A matrix of lake-to-lake recaptures for the 11 lakes sampled during tagging in 1996–1998. The lakes in the vertical column are the lakes where fish were initially tagged. Third Pond, Back-up Pond and Little Bear Cave Pond were not sampled for the genetic study

Lake	1	2	3	4	5	6	7	8	9	10	11
Moccasin Pond (1)	–										
Southern Pond (2)	0	–									
Fourth Pond (3)	0	0	–								
Alleys Pond (4)	0	0	0	–							
Wings Pond (5)	0	0	0	0	–						
Fourmile Pond (6)	0	0	0	0	0	–					
Second Pond (7)	0	0	0	0	1	0	–				
Back-up Pond (8)	0	0	0	2	0	0	0	–			
Little Bear Cave Pond (9)	0	0	0	0	0	0	3	0	–		
Third Pond (10)	0	0	3	1	0	0	1	0	0	–	
Skippers Pond (11)	0	0	0	0	0	0	0	1	0	0	–

data also indicate little or no movement among lakes. As far as is known, this is the first study to combine tagging and microsatellite genetic data in an analysis of fine-scale brook charr population structure.

The brook charr populations of Indian Bay exhibited high levels of heterozygosity, comparable to those reported elsewhere for this species. Overall heterozygosity at the loci Sfo-18 and Sfo-23 ranged from 0·27 to 0·90, compared to a range of 0·62–0·89 at the same loci for brook charr of La Mauricie National Park, Québec (Angers & Bernatchez, 1998), the only other brook charr populations for which data are available at these loci. As with the La Mauricie National Park populations, the brook charr populations were differentiated at the scale of individual lakes.

The results of the pair-wise comparisons of the sampled populations with θ, *R*st and D_A appear to make biological and geographical sense. Comparing the multi-dimensional scaling plots (Fig. 2) to the map of the watershed (Fig. 1), it is clear that the lakes in the Wings Pond and Alleys Pond tributaries of the watershed are consistently found together in the same quadrant of the multi-dimensional scaling plots (*e.g.* Wings Pond–Southern Pond, Alley's Pond–Fourth Pond–Skippers Pond). This suggests a relationship between genetic distance and geographical distance (measured along the contours of the connecting streams). Although a Mantel test revealed no significant relationship between these two variables, the measure of geographical distance used here may not accurately reflect the relative difficulty of migrating among lakes.

Streams can vary in width, depth, water velocity, pH and conspecific density, all of which can affect the ease and success of migration, independent of stream length. Angers *et al.* (1999) found a relationship between change in altitude and level of gene flow, using canonical correspondence analysis. Changes in relief are minimal throughout the Indian Bay watershed (<30 m over the length of most second and third order streams); thus changes in altitude are unlikely to be a good predictor of gene flow. The notable exception to the grouping by tributary was

Moccasin Pond, which grouped with Alley's Pond and Fourth Pond consistently (Fig. 2) even though it is in a different tributary (Fig. 1). A close examination of the geographical relationship between these lakes reveals that headwater streams entering Moccasin Pond and Alley's Pond originate in a peat bog <50 m from each other. Historically these streams may have been linked. It may also be possible that these streams still make occasional contact during peak spring water levels. Moccasin Pond, however, also had the smallest sample size and the grouping with Alley's Pond and Fourth Pond may be an artifact. Although Fourmile Pond was closest of the lakes to Second Pond in the multi-dimensional scaling plots, the co-ordinates of these two lakes were much farther apart than expected based on their close geographical proximity (Figs 1 and 2). This is difficult to explain. Second Pond has experienced the highest level of anthropogenic disturbance, including damming that lead to a significant flooding of the riparian habitat. This disturbance may have affected the genetic variation in this population or the gene-flow with surrounding brook charr populations. It is worth noting that Fourmile Pond was still the closest lake to Second Pond in the multi-dimensional scaling plots based on all three metrics.

Similarities in genetic variation as suggested by pair-wise estimates of θ, D_A and *R*st are surprising, given that each metric is based on different assumptions about mutation mechanisms. Even though *R*st is based on the SMM, it often performs well (low variance and bias) even when coalescence times are short and drift is the primary cause of divergence (Slatkin, 1995). Slatkin (1995) suggested that increases in *R*st above θ may be due to extra mutations over and above the variation generated by genetic drift. Thus, *R*st may out-perform (*i.e.* provide a more accurate measure of divergence) θ when coalescence time is long and mutation is a significant determinant of genetic variation. In such a situation, θ will underestimate population differentiation (Wenburg *et al.*, 1998). The similarity in these estimates for Indian Bay brook charr suggests that genetic drift may be the primary source of population divergence (Wenburg *et al.*, 1998; Tessier & Bernatchez, 2000). D_A is most appropriate for long-term evolution with drift and mutation effecting divergence (Weir, 1996). Qualitative relationships among populations that have diverged due to drift alone, however, can still be elucidated with some confidence (Paetkau *et al.*, 1997).

The individual assignment tests did not assign individuals to their source populations with the high levels of accuracy often reported in the literature ($>90\%$; King *et al.*, 2001). Many of the studies that reported high fidelity of individuals to source populations, however, also reported substantially higher θ and *R*st and D_A (King *et al.*, 2001; Potvin & Bernatchez, 2001) or compared only two samples (Pettersson *et al.*, 2001). The individuals who were not assigned to their source populations were not distributed among the other populations in any discernable pattern. The proportion of individuals rejected from non-source populations using the Geneclass simulation option was relatively high given the number of loci used and the relatively small sample size for some of the populations. The proportion of individuals rejected from non-source populations also did not appear to correlate with the geography of the watershed. It is worth noting that the Second Pond brook charr were relatively likely to be rejected from the other brook charr populations (0·65, mean proportion rejected), while brook charr from other lakes were relatively unlikely to be rejected from the Second Pond brook charr population,

(0·23, mean proportion rejected) suggesting an admixture of genotypes in this population (Ruzzante *et al*., 2001). Second Pond is essentially the 'hub' of this system, with all the major tributaries running into this lake before continuing to the sea along the main stem of the watershed. These assignment data suggest that Second Pond may be subject to immigration from other areas of the watershed.

The extremely low number of lake-to-lake recaptures revealed by the tagging data supports the conclusion of reproductive isolation among Indian Bay brook charr populations. A quantitative comparison of individual assignment tests and tagging recaptures was not possible because of the very low number of recaptures. A regression analysis comparing genetic distances and the number of lake-to-lake recaptures indicated no significant relationship, which is not surprising given the large number of lakes with no interlake recaptures. Global values of N_m based on θ, *R*st and Slatkin's private allele method were 2·49, 3·24 and 9·38 effective migrants per generation, respectively, slightly lower than expected based on tagging data. For example, the median size of Indian Bay brook charr populations is *c*. 4500 individuals (van Zyll de Jong *et al*., 1999). For two populations with 9000 individuals in total and a 2·0% migration rate per generation, based on 12 of 556 recaptures being lake-to-lake, a minimum of 180 migrants per generation (3 years) would be expected between the two lakes. These data suggest that the minimum N_m could be significantly higher than those indicated by the genetic data. It is important, however, to distinguish between N_m and the number of effective migrants. A brook charr must not only travel to a new lake (N_m) but also survive to reproduce successfully (number of effective migrants). For example, even if migration carries no cost, brook charr in the Indian Bay watershed experience 65–85% annual mortality rates (Adams, 1999). These mortality rates suggest that a significant portion of migrants do not survive to reproduce. Thus, the N_m estimated by tagging data should tend to overestimate the number of effective migrants. Despite the difficulty in making direct quantitative comparisons between tagging and genetic data, the qualitative results of both metrics suggest low levels of migration. The study suggests that both microsatellite markers and tagging data can be effective in defining microgeographical population structure in salmonid fishes. In future studies, however, it might be more informative to compare tagging and genetic data from a system where substantial migration is expected.

Life-history data collected from three of the Indian Bay lakes suggest significant differences in growth rate, survival, fecundity and age at maturity (Adams, 1999). Although it is difficult to determine if those life-history differences have a genetic or phenotypic basis, the genetic and tagging data suggest the potential for local adaptation exists.

The brook charr of Indian Bay are subdivided into reproductively isolated lacustrine populations. Tagging data are congruent with genetic data suggesting that both techniques can be an effective tool for defining salmonid population structure. The data from the present study, and that of Angers & Bernatchez (1998), suggest that brook charr genetic conservation and fisheries management should be based at the scale of individual lakes whenever possible.

We would like to thank three anonymous reviewers whose comments lead to a greatly improved manuscript. We are also grateful to W. Norris, R. Pritchett and B. Wicks of the Indian Bay Ecosystem Corporation (IBEC) and M. van Zyll de Jong of Inland Fisheries,

Government of Newfoundland and Labrador for their field and technical support of this research. D. Cook, T. McParland and M. Jones provided assistance in the laboratory. M. Ferguson, J. Wright, R. Myers and M. Jones provided helpful comments on earlier versions of the manuscript. The research was supported financially by IBEC and by an NSERC (Canada) Research Grant to JAH.

References

Adams, B. (1999). Variation in genetic structure and life histories among populations of brook char *Salvelinus fontinalis*, from insular Newfoundland, Canada. MSc Thesis. Dalhousie University, Halifax.

Angers, B. & Bernatchez, L. (1998). Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. *Molecular Biology and Evolution* **15,** 143–159.

Angers, B., Bernatchez, L., Angers, A. & Desgroseillers, L. (1995). Specific microsatellite loci for brook charr reveal strong population subdivision on a microgeographic scale. *Journal of Fish Biology* **47** (Suppl. A), 177–185.

Angers, B., Magnan, P., Plantes, M. & Bernatchez, L. (1999). Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology* **8,** 1043–1053.

Beacham, T. D., Miller, K. M. & Withler, R. E. (1996). Minisatellite DNA variation and stock identification of coho salmon. *Journal of Fish Biology* **49,** 411–429.

Bentzen, P., Taggart, C. T. & Cook, D. (1996). Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences* **53,** 2706–2721.

Brykov, V. A., Polyakova, L. A., Skurikhina, L. A. & Kukhlevsky, A. D. (1996). Geographical and temporal mitochondrial DNA variability in populations of pink salmon. *Journal of Fish Biology* **48,** 899–909.

Cornuet, J. M., Pirys, S., Luikart, G., Estoup, A. & Solimac, M. (1999). New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153,** 1989–2000.

Danzman, R. G., Ihssen, P. E. & Hebert, P. D. N. (1991). Genetic discrimination of wild and hatchery populations of brook charr, *Salvelinus fontinalis* (Mitchill), in Ontario using mitochondrial DNA analysis. *Journal of Fish Biology* **39,** 69–77.

Ferguson, M. M. & Danzmann, R. G. (1998). Role of genetic markers in fisheries and aquaculture: useful tools or stamp collecting. *Canadian Journal of Fisheries and Aquatic Sciences* **55,** 1553–1563.

Gagiotti, O. E., Lange, O., Rassman, K. & Gliddon, C. (1999). A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* **8,** 1513–1520.

Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L. & Feldman, M. W. (1995). An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139,** 463–471.

Goodman, S. J. (1997). RST CALC: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* **6,** 881–885.

Groot, C., Bailey, R. E., Margolis, L. & Cooke, K. (1989). Migratory patterns of sockeye salmon (*Oncorhyncus nerka*) smolts in the strait of georgia, British Columbia, as determined by analysis of parasite assemblages. *Canadian Journal of Zoology* **67,** 1670–1678.

Guo, S. W. & Thompson, E. A. (1992). Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* **48,** 361–372.

Hansen, M. M. & Loeschcke, V. (1996). Genetic differentiation among Danish brown trout populations, as detected by RFLP analysis of PCR amplified mitochondrial DNA segments. *Journal of Fish Biology* **48,** 422–436.

Healey, M. C. (1978). The dynamics of exploited lake trout populations and implications for management. *Journal of Wildlife Management* **42,** 307–328.
Hilborn, R. (1990). Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. *Canadian Journal of Fisheries and Aquatic Sciences* **47,** 635–643.
Hutchings, J. A. (1993). Adaptive life histories effected by age-specific survival and growth rate. *Ecology* **74,** 673–684.
Hutchings, J. A. (1994). Age- and size-specific costs of reproduction within populations of brook trout, *Salvelinus fontinalis*. *Oikos* **70,** 12–20.
Hutchings, J. A. (1996). Adaptive phenotypic plasticity in brook trout, *Salvelinus fontinalis*, life histories. *Ecoscience* **3,** 25–32.
King, T. L., Kalinowski, S. T., Schill, W. B., Spidle, A. P. & Lubinski, B. A. (2001). Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular Ecology* **10,** 807–821.
Krebs, C. J. (1998). *Ecological Methodology.* Menlo Park, CA: Addison-Welsey Educational Publishers, Inc.
Lester, R. J. G., Sewell, K. B., Barnes, A. & Evans, K. (1988). Stock discrimination of orange roughy, *Hoplostethus atlanticus*, by parasite analysis. *Marine Biology* **99,** 137–143.
McConnell, S., Hamilton, L., Morris, D., Cook, D., Paquet, D., Bentzen, P. & Wright, J. M. (1995). Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. *Aquaculture* **137,** 19–30.
Myers, R. A., Hammill, M. O. & Stenson, G. B. (1997). Using mark-recapture to estimate the numbers of a migrating stage structures population. *Canadian Journal of Fisheries and Aquatic Sciences* **54,** 2097–2104.
Nei, M., Tajima, F. & Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19,** 153–170.
O'Connell, M., Skibinski, D. O. F. & Beardmore, J. A. (1996). Allozyme and mtDNA divergence between Atlantic salmon populations in North Wales. *Journal of Fish Biology* **48,** 1023–1026.
Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L. & Strobeck, C. (1997). An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* **147,** 1943–1957.
Pettersson, J. C. E., Hansen, M. M. & Bohlin, T. (2001). Does dispersal from landlocked trout explain the coexistence of resident and migratory trout females in a small stream? *Journal of Fish Biology* **58,** 487–495. doi: 10.1006/jfbi.2000.1468.
Potvin, C. & Bernatchez, L. (2001). Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multilocus individual assignment and mixed stock analysis. *Molecular Ecology* **10,** 2375–2388.
Raymond, M. & Rousset, F. (1995). Genepop (Version 1.2): population genetics software for exact test and ecumenism. *Journal of Heredity* **86,** 248–249.
Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43,** 223–225.
Ruzzante, D. E. (1998). A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Sciences* **55,** 1–14.
Ruzzante, D. E., Hansen, M. M. & Meldrup, D. (2001). Distribution of individual inbreeding coefficients, relatedness and influence of stocking on native anadromous brown trout (*Salmo trutta*) population structure. *Molecular Ecology* **10,** 2107–2128.
Schneider, S., Kuffer, J. M., Roessli, D. & Excoffier, L. (1998). *Arlequin, Version 1.1: A Software for Population Genetic Data Analysis.* Geneva: Genetic and Biometry Laboratory, University of Geneva.
Shriver, M. D., Jin, L. & Boerwinkle, E. (1995). A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Molecular Biology and Evolution* **12,** 914–920.
Slatkin, M. (1985). Rare alleles as indicators of gene flow. *Evolution* **39,** 53–65.
Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139,** 457–462.

Small, M. P., Beacham, T. D., Withler, R. E. & Nelson, R. J. (1998). Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser river, British Columbia, using microsatellite DNA markers. *Molecular Ecology* **7,** 141–155.

Takezaki, N. & Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144,** 389–399.

Tessier, N. & Bernatchez, L. (2000). A genetic assessment of single versus double origin of landlocked Atlantic salmon (*Salmo salar*) from Lac Saint-Jean, Quebec, Canada. *Canadian Journal of Fisheries and Aquatic Sciences* **57,** 797–804.

Weir, B. S. (1996). *Genetic Data Analysis II*. Sunderland, MA: Sinauer Associates Inc.

Weir, B. S. & Cockerhan, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38,** 1358–1370.

Wenburg, J. K., Bentzen, P. & Foote, C. J. (1998). Microsatellite analysis of genetic population structure in an endangered salmonid: the coastal cutthroat trout (*Oncorhynchus clarki clarki*). *Molecular Ecology* **7,** 733–749.

Whitlock, M. C. & McCauley, D. E. (1999). Indirect measures of gene flow and migration Fst $\neq 1/(4Nm+1)$. *Heredity* **82,** 117–125.

Wilder, D. G. (1952). A comparative study of anadromous and freshwater populations of brook trout (*Salvelinus fontinalis* (Mitchill)). *Journal of the Fisheries Research Board of Canada* **9,** 169–203.

van Zyll de Jong, M. C., Lester, N. P., Korver, R. M., Norris, W. & Wicks, B. L. (1999). Brook trout (*Salvelinus fontinalis* Mitchill.) population dynamics and recreational fishery in Indian Bay Brook, Newfoundland (1995–98). *Canadian Stock Assessment Secretariat Research Document 99/97*. Ottawa: Department of Fisheries and Oceans.

Electronic Reference

Miller, M. P. (1997). *Tools for Population Genetic Analysis (TFPGA) 1.3: A Windows Program for Analysis of Allozyme and Molecular Population Genetic Data*. http://bioweb.usu.edu/mpmbio/index.htm.