Sea Lamprey (*Petromyzon marinus*) Populations in Northeastern North America: Genetic Differentiation and Affinities[1]

Janet Wright

Section of Ecology and Systematics, Cornell University, Ithaca, NY 14853, USA

Charles C. Krueger

Department of Fisheries and Wildlife, University of Minnesota, St. Paul, MN 55108, USA
and
Department of Natural Resources, New York State College of Agriculture and Life Sciences, Cornell University, Ithaca, NY 14853, USA

and Peter F. Brussard and Marjorie Collings Hall[2]

Section of Ecology and Systematics, Cornell University, Ithaca, NY 14853, USA

Wright, J., C. C. Krueger, P. F. Brussard, and M. C. Hall. 1985. Sea lamprey (*Petromyzon marinus*) populations in northeastern North America: genetic differentiation and affinities. Can. J. Fish. Aquat. Sci. 42: 776–784.

We examined genetic population structure in 3253 sea lampreys (*Petromyzon marinus*) sampled at 53 sites in northeastern North America and the British Isles. Hierarchical *F*-statistics of allozyme frequencies indicated that genetic divergence from lake to lake and river to river was greater than divergence among sites within lakes or rivers and also greater than that attributable to the separation of landlocked freshwater from coastal anadromous systems. Lampreys from different lakes varied considerably in the amount of spatial differentiation evident. New York freshwater populations of uncertain invasion history were no more genetically differentiated than were recently introduced upper Great Lakes populations. The apparent reproductive isolation of lampreys in Lake Erie from Lake Ontario populations, despite canal connections, suggests that movement is limited between lakes. Future studies should determine the cause so that management techniques to restrict movement can be developed to control populations.

Les auteurs ont étudié la structure démographique génétique chez 3 253 lamproies marines (*Petromyzon marinus*) échantillonnées à 53 emplacements du nord-est de l'Amérique du Nord et des îles Britanniques. Les statistiques hiérarchiques *F* sur les fréquences des allozymes révèlent que la divergence génétique d'un lac à l'autre et d'un cours d'eau à l'autre était plus élevée que la divergence entre les emplacements lotiques ou lénitiques et plus importante que celle attribuable à la séparation des populations dulçaquicoles confinées aux eaux intérieures et des populations anadromes en milieu côtier. La valeur de la différenciation spatiale évidente variait considérablement chez les lamproies prélevées dans différents lacs. Les populations dulçaquicoles de l'État de New York, dont on ne connaît pas bien les antécédents en matière d'invasion, n'étaient pas plus différentes génétiquement que des populations récemment apparues dans les Grands Lacs supérieurs. L'isolement apparent pour ce qui est de la reproduction des populations de lamproies des lacs Érié et Ontario, malgré la présence d'un canal, porte à croire que les déplacements entre les lacs sont limités. On devrait étudier la cause de cet isolement afin d'élaborer des mesures gestionnelles pour restreindre les déplacements des lamproies et ainsi lutter contre ce prédateur.

Received May 7, 1984
Accepted January 14, 1985
(J7778)

Reçu le 7 mai 1984
Accepté le 14 janvier 1985

From the 1930s through 1950, sea lampreys (*Petromyzon marinus*) invaded the upper Great Lakes from Lake Ontario and underwent spectacular population expansion. Since then, the impact of the adult predatory phase of this species on commercially important fish stocks has caused fishery biologists to develop and implement methods to reduce sea lamprey populations. Chemical lampricides applied to spawning streams have had considerable success in control of populations in the upper Great Lakes. However, treated streams typically become recolonized, and concern over long-term chemical application has called for closer study of sea lamprey biology (Smith et al. 1974; Smith and Tibbles 1980; Pearce et al. 1980). In particular, a knowledge of the genetics of population structure would be desirable for geographically delineating populations, identifying present migration routes, and tracing the probable origin of populations whose history is poorly known. Not only would this knowledge be useful in management planning, but it would also be important for understanding the sea lamprey as an outstandingly successful colonizing species (Parsons 1983).

Sea lamprey populations contain considerable genetic variation detectable by electrophoretic separation of enzyme variants (Krueger 1980; Brussard et al. 1981; Krueger and Spangler

[1]Paper No. 13995, Scientific Journal Series, Minnesota Agricultural Experiment Station, St. Paul, MN 55108, USA. Research Contribution No. 160. This work is the result of research sponsored by Minnesota supported by the NOAA Office Sea Grant, Department of Commerce, under grant No. NA79AA-D-00025.

[2]Present address: 32 Ann Street, Bracebridge, Ont. P0B 1C0.

1981), but previous studies have also brought up some problems of data interpretation. On the one hand, electrophoretic analysis of sea lampreys in Lake Superior was reported to indicate regional populations (Krueger and Spangler 1981), but similar data from Lakes Michigan and Huron have not revealed such populations (L. Jacobson, University of Minnesota, unpubl. data). On the other hand, electrophoretic data have indicated genetic affinities among sea lampreys of New York waters but have not shown unequivocally whether lampreys in inland New York lakes are endemic or recently introduced, because of the difficulty of predicting how differentiated these populations should have become over a given time span (Brussard et al. 1981). In both cases, a broadening of the geographic scale on which the patterns are considered could potentially resolve the problems of interpretation. We examined electrophoretic variation among collections of sea lampreys at 53 sites over a geographic area encompassing northeastern North America and the British Isles with the aim of clarifying these patterns.

Methods and Materials

Our analysis was based on sea lamprey collections made as part of a cooperative project between the University of Minnesota and Cornell University. Previous papers described genetic structure of lamprey populations in Lake Superior (Krueger and Spangler 1981) and New York and nearby waters (Brussard et al. 1981). This report incorporates these results in addition to second-year samples from the original New York sites and 22 sites not previously described.

Most lampreys were collected as ammocoetes by electroshocking stream habitat; some specimens were collected as spawning adults. Among collections from Lakes Superior and Huron, some were made during chemical treatments of streams. Details of collection methods have been reported (Brussard et al. 1981; Krueger and Spangler 1981).

Enzyme variants of individuals were resolved by horizontal starch gel electrophoresis according to previously published methods (Krueger 1980; Brussard et al. 1981). Although Krueger (1980) reported variation at the *PGI-1* (phosphoglucose isomerase) locus in some samples and Brussard et al. (1981) found variation in *PGM-2* (phosphoglucomutase) and occasionally in *IDH-2* (isocitrate dehydrogenase), for purposes of standardization our analysis was based on only those four polymorphic loci scored in common in all samples: *AGP* (alpha-glycerophosphate dehydrogenase), *PGI-2*, *PGM-1*, and *MDH-1* (malate dehydrogenase).[3] To assure correspondence of results between Cornell and Minnesota laboratories, one sample (site 16) containing all the major electromorphs was split and a portion of it run in each laboratory. Bands of differing mobility were assumed to represent allelic variants; designations were based on mobility relative to a value of 100 for the common electromorph.

Analysis of the genetic data employed the *F*-statistics (gene diversity) measures of inbreeding and population structure developed by Wright (1965, 1978) and Nei (1975, 1977) and previously applied to data from Lake Superior sea lampreys (Krueger and Spangler 1981). The basic approach of an *F*-statistics analysis is to compare amounts of genetic differentiation that have accrued between various subdivisions of the total population in time or space. Our primary interest was in spatial differentiation, for which we needed to be able to assume that each individual sample contained members of no more than one deme, and that "noise" from year-to-year variation was low enough that it did not obscure geographic variation on the scale under consideration. Deme integrity within samples was assessed by conformance to Hardy–Weinberg expectation. For each sample at each locus, deviations from expectation were calculated as the fixation index F_{IS} (Nei 1977, p. 228), and their significance was tested by log-likelihood *G*-test (Sokal and Rohlf 1981, p. 704). The significance of temporal variation was tested by heterogeneity *G*-test (Sokal and Rohlf 1981, p. 724) between paired samples at eight sites where collections were made in two successive years and also in three cases where collections included two life stages (ammocoete and adult). For these doubly sampled sites, the magnitude of temporal variation was further compared with that of geographic variation in a variance components analysis of arcsine-transformed frequencies of the common ("*100*") allele (Snedecor and Cochran 1967, p. 279).

With regard to spatial organization, the first question to be asked was one of scale. Was there greater genetic differentiation among samples in an average lake, for instance, than between the total populations of two lakes? To examine such differences, populations were arrayed in a nested scheme of three levels: (1) samples within each of 13 lakes or rivers (we chose to consider the coastal rivers Connecticut, Delaware, and Hudson as analogous to inland lakes, since spawning streams were identifiable in each); (2) lakes or rivers within each of two water systems (inland landlocked, coastal sea-run); and (3) inland or coastal systems within the total. A standardized genetic variance ("F_{ST}" of Wright 1965, as approximated by "G_{ST}" of Nei 1977) was calculated at each level of the hierarchy and designated F_{SL} for samples within lakes (or rivers), F_{LW} for lakes (or rivers) within the inland (or coastal) water systems, and F_{WT} for inland and coastal systems within the total. This hierarchical approach estimated the proportion of total observed genetic variation that could be attributed to differentiation among the subunits of that level. Thus, for the inland/coastal water systems as subunits and all lampreys as total, $F_{WT} = (H_T - H_W)/H_T$, where H_T is expected heterozygosity in the total "population" for that level ($H_T = 1 - \Sigma p_i^2$; p_i = frequency of allele i) and H_W is the mean of expected heterozygosities of each of the two subunits weighted by sample size (Nei 1977). Significance of the degree of differentiation at each level was evaluated by *G*-test for heterogeneity of allele counts among the subunits of that level (Sokal and Rohlf 1981, p. 724).

Next, the individual *F*'s within each level were examined comparatively to address two main questions: (1) Do lampreys in all lakes show equal levels of differentiation over space? (2) Can levels of differentiation among population units of known history indicate the history of populations of unknown origin?

To supplement the *F*-statistics analysis and to allow genetic affinities to be diagrammed, we subjected geographically contiguous sets of samples to a simultaneous test procedure (STP; Sokal and Rohlf 1981, p. 253) to pool together the largest groups of samples that showed no significant ($p < 0.05$) genetic heterogeneity at any locus. Nei's (1975, p. 175) genetic distances, based on allele frequencies at the four polymorphic loci, were

[3]Designations for loci and alleles in this paper follow Krueger (1980). Different designations (in parentheses) were used by Brussard et al. (1981) for the following: *AGP 100, 146 (GPDH^b, GPDH^a); MDH-1 −100, −165 (MDH-2^b, MDH-2^a); PGI-1 (PHI-2); PGI-2 100, 106, 122, 92 (PHI-1^b, PHI-1$^{a'}$, PHI-1^a, PHI-1^c); PGM-1 100, 148, 69 (PGM-2^b, PGM-2^a, PGM-2^c); PGM-2 (PGM-1); IDH-2 (IDH-1).*

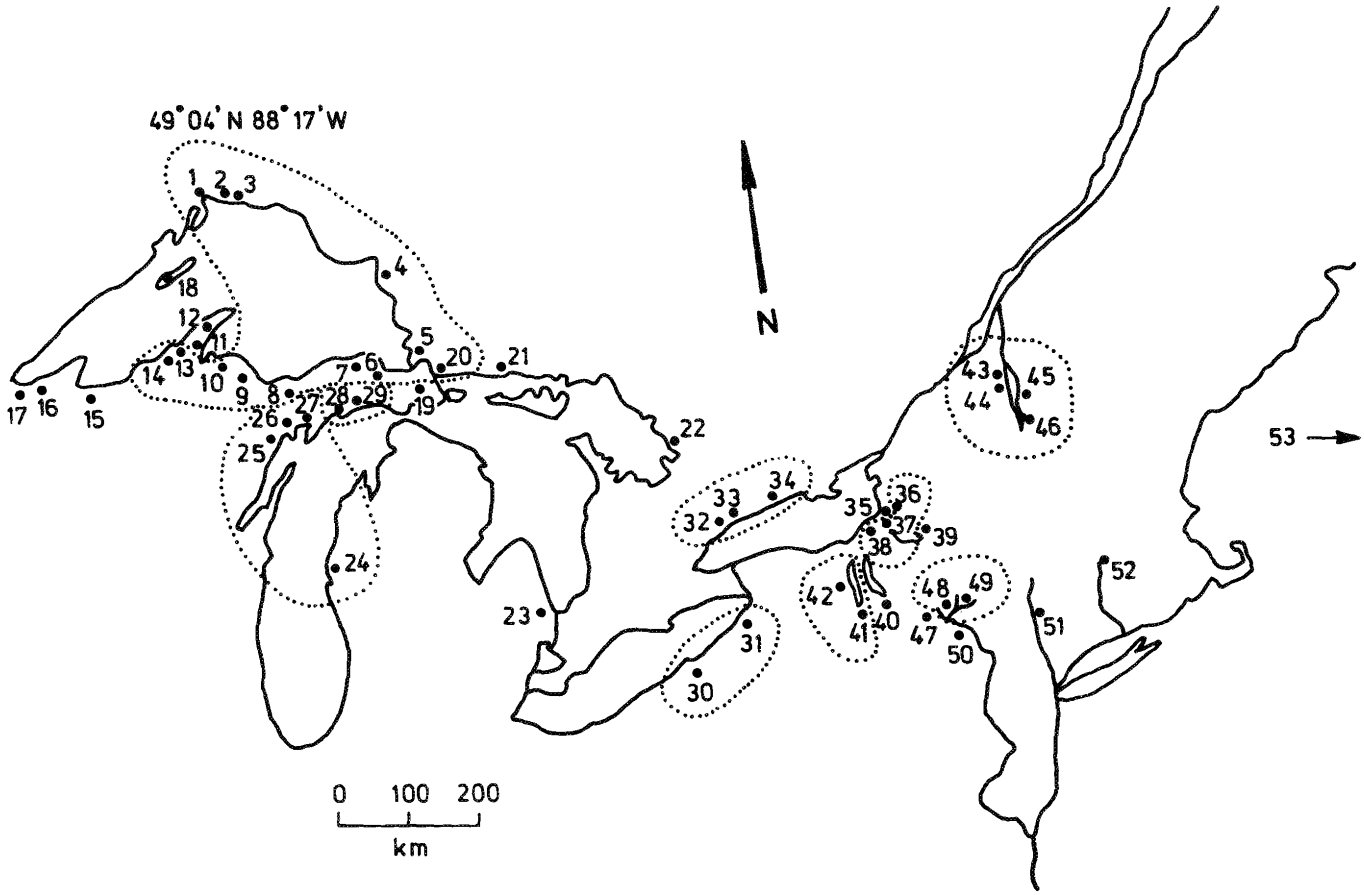

FIG. 1. Locations where sea lamprey samples were collected for genetic analysis (see Appendix). Dotted lines enclose genetically homogeneous clusters as determined by *G*-test for heterogeneity.

calculated among the sample groups thus identified. Cluster analysis (unweighted pair group method analysis; Sneath and Sokal 1973, p. 230) was applied to the matrix of genetic distances to produce a dendrogram of distances among the populations.

Results

Fifty-three collecting localities were represented in this analysis (Fig. 1; Appendix). Genotypic frequencies within samples at the four polymorphic loci (Table 1) generally approximated Hardy–Weinberg expectation; of the total of 212 *G*-tests for the 53 samples, only 11 deviated significantly ($p < 0.05$; Table 1; Brussard et al. 1981; Krueger and Spangler 1981). The fixation indices (F_{IS}) for nine of these deviant comparisons were positive, indicating heterozygote deficiencies which could suggest population substructuring at these localities (Wahlund 1928). The number of deviations observed, however, was no greater than that expected by chance alone.

Temporal genetic variation within lampreys from single collection localities was analyzed in two ways, with the general result that variation over time was less than that over space at the scale we examined (Tables 2 and 3). First, in *G*-tests of heterogeneity, although successive-year samples of ammocoete allele frequencies at four sites were each statistically heterogeneous ($p < 0.05$) at one of their four loci, at four other sites there was no significant heterogeneity from year to year, and there were no significant differences in allele frequencies between ammocoetes and adults collected together at a site (Table 2). Second, in variance components analysis, year-to-year changes at a site were negligible compared with site-to-site differences, and at two of three loci the differences among spatially separated sites contributed substantially more to total variance than did the differences between ammocoetes and adults collected simultaneously at a site (Table 3). For purposes of the geographic analysis below, successive-year and life-stage collections at a site were pooled if they showed no significant differences at any enzyme locus (heterogeneity *G*-test, $p < 0.05$); for sites at which samples were not completely homogeneous, only the collection with the larger sample size was considered.

Hierarchical *F*-statistics analysis over geographic localities showed stronger differentiation at some levels of the hierarchy than at others (Table 4). For each polymorphic locus, and for all loci averaged, the *F* values partition observed genetic variability into within-lake (or river), among-lake (or river), and between-system components. Average differentiation among lakes and rivers (F_{LW}) was greater (0.061) than site-to-site differentiation within lakes and rivers ($F_{SL} = 0.024$), which in turn was greater than differentiation between coastal and inland systems ($F_{WT} = 0.013$; Table 4). Variability unaccounted for at these three levels (the "$1 - F_{ST}$" column) is the proportion attributable to within-sample variation (0.902; Table 4).

Although site-to-site differentiation of lampreys within lakes and rivers was on the average less than lake-to-lake or river-to-river differentiation, lamprey population groups in some lakes were more internally subdivided than in others (Table 5). Site-to-site differentiation (F_{SL}) in the New York inland lakes

TABLE 1. Allele frequencies, sample sizes, and deviation from Hardy–Weinberg expectation (F_{IS}) at four polymorphic loci for previously unreported populations of sea lamprey. Site numbers as in Fig. 1. All are 1980 collections, except the 1979 adults at site 40. The 1979 data from sites 35, 38, 39, 41, 46, 47, 50, 51, and 52 and 1980 data from sites 30 and 31 (Lake Erie) and 52 (Connecticut River) are given in Brussard et al. (1981). For data from sites 1–18 (Lake Superior), see Krueger and Spangler (1981). *$p < 0.05$ (log-likelihood test); **$p < 0.01$; ND = no data.

		AGP						*PGI-2*						*PGM-1*					*MDH-1*			
Drainage	Site No.	*100*	*146*	*158*	*190*	*N*	F_{IS}	*100*	*106*	*122*	*92*	*N*	F_{IS}	*100*	*148*	*69*	*N*	F_{IS}	*−100*	*−165*	*N*	F_{IS}
L. Huron	19	0.47	0.53	0.00	0.00	32	0.16	0.82	0.00	0.13	0.06	36	0.01	0.63	0.38	0.00	36	−0.01	0.82	0.18	36	−0.22
L. Huron	20	0.53	0.47	0.00	0.00	31	−0.13	1.00	0.00	0.00	0.00	32	0.00	0.64	0.36	0.00	32	0.39*	0.81	0.19	32	−0.03
L. Huron	21	0.45	0.55	0.00	0.00	40	0.19	0.99	0.00	0.13	0.00	40	−0.01	0.73	0.28	0.00	40	−0.13	0.91	0.09	40	−0.10
L. Huron	22	0.37	0.63	0.00	0.00	23	0.35	1.00	0.00	0.00	0.00	39	0.00	0.42	0.58	0.00	37	0.17	0.73	0.27	39	−0.11
L. Huron	23	0.66	0.34	0.00	0.00	38	−0.17	0.88	0.00	0.12	0.00	38	−0.13	0.58	0.42	0.00	37	−0.28	0.66	0.34	40	−0.29
L. Michigan	24	0.59	0.41	0.00	0.00	34	−0.09	0.99	0.00	0.01	0.00	36	−0.01	0.55	0.45	0.00	38	−0.17	0.80	0.20	40	−0.25
L. Michigan	25	0.52	0.48	0.00	0.00	30	−0.13	0.95	0.00	0.03	0.03	38	−0.04	0.66	0.34	0.00	38	0.06	0.79	0.21	40	−0.12
L. Michigan	26	0.57	0.44	0.00	0.00	23	−0.06	0.96	0.00	0.03	0.01	36	−0.03	0.68	0.32	0.00	36	0.04	0.85	0.15	40	−0.18
L. Michigan	27	0.46	0.54	0.00	0.00	25	−0.21	0.95	0.00	0.01	0.04	39	−0.04	0.50	0.50	0.00	39	−0.08	0.80	0.20	40	−0.25
L. Michigan	28	0.68	0.33	0.00	0.00	20	−0.25	0.90	0.00	0.00	0.10	26	−0.11	0.67	0.33	0.00	24	0.06	0.90	0.10	26	−0.11
L. Michigan	29	0.53	0.47	0.00	0.00	32	0.00	0.86	0.00	0.01	0.13	35	−0.15	0.66	0.34	0.00	40	−0.17	0.89	0.11	40	−0.13
L. Ontario	32	0.50	0.50	0.00	0.00	4	0.00	0.88	0.00	0.13	0.00	48	−0.14	0.90	0.10	0.00	48	−0.12	0.89	0.12	48	−0.13
L. Ontario	33	0.67	0.33	0.00	0.00	3	−0.50	0.91	0.00	0.09	0.00	66	0.30	0.84	0.16	0.00	63	−0.19	0.86	0.14	63	0.09
L. Ontario	34	0.54	0.46	0.00	0.00	39	−0.03	0.89	0.00	0.12	0.00	39	−0.13	0.91	0.09	0.00	39	−0.10	0.90	0.10	39	0.16
L. Ontario	35	0.46	0.54	0.00	0.00	49	0.22	0.85	0.00	0.14	0.01	111	0.05	0.93	0.07	0.00	113	0.07	0.86	0.14	100	−0.08
L. Ontario	36	0.65	0.35	0.00	0.00	52	0.24	0.78	0.00	0.22	0.01	86	−0.17	0.92	0.08	0.00	86	−0.08	0.94	0.06	86	0.15
L. Ontario	37	0.51	0.49	0.00	0.00	77	−0.01	0.88	0.00	0.10	0.02	77	0.01	0.91	0.09	0.00	80	0.04	0.95	0.05	80	0.47**
L. Ontario	38	0.47	0.53	0.00	0.00	47	−0.11	0.80	0.00	0.16	0.04	51	−0.08	0.88	0.12	0.00	51	−0.13	0.96	0.04	51	−0.04
L. Ontario (adults)	38	ND						0.90	0.00	0.10	0.00	30	−0.11	0.82	0.18	0.00	30	0.44*	0.92	0.08	30	−0.09
Oneida L.	39	0.44	0.56	0.00	0.00	151	0.00	0.77	0.00	0.22	0.01	195	0.13	0.96	0.04	0.00	196	−0.04	0.97	0.03	190	0.14
Cayuga L.	40	0.47	0.53	0.00	0.00	125	0.00	0.84	0.00	0.15	0.01	264	0.01	0.96	0.04	0.00	273	−0.04	0.97	0.03	275	−0.03
Cayuga L. (adults 1979)	40	0.55	0.45	0.00	0.00	21	−0.06	0.91	0.00	0.09	0.00	35	0.21	0.90	0.10	0.00	35	−0.11	0.94	0.06	35	−0.06
Cayuga L. (adults)	40	ND						0.84	0.00	0.16	0.00	66	0.04	0.97	0.03	0.00	66	−0.03	0.98	0.02	64	−0.02
Seneca L.	41	0.33	0.67	0.00	0.00	29	−0.02	0.82	0.00	0.19	0.00	62	0.09	1.00	0.00	0.00	61	0.00	1.00	0.00	61	0.00
Seneca L.	42	0.33	0.67	0.00	0.00	6	−0.50	0.83	0.00	0.17	0.00	6	−0.20	1.00	0.00	0.00	6	0.00	1.00	0.00	6	0.00
L. Champlain	43	0.79	0.21	0.00	0.00	62	−0.07	0.92	0.00	0.08	0.00	66	−0.09	0.99	0.02	0.00	66	−0.02	1.00	0.00	66	0.00
L. Champlain	44	0.68	0.32	0.00	0.00	44	−0.15	0.91	0.00	0.09	0.00	44	−0.10	0.99	0.01	0.00	44	−0.01	1.00	0.00	44	0.00
L. Champlain	45	0.84	0.16	0.00	0.00	35	−0.19	0.93	0.00	0.07	0.00	30	−0.07	0.97	0.03	0.00	30	−0.03	1.00	0.00	30	0.00
L. Champlain	46	0.73	0.27	0.00	0.00	153	−0.03	0.90	0.00	0.10	0.00	161	0.02	0.98	0.02	0.00	161	−0.02	0.98	0.02	161	−0.02
Delaware R.	47	0.62	0.37	0.01	0.00	71	−0.06	0.96	0.01	0.01	0.02	93	−0.03	0.97	0.22	0.01	93	−0.03	0.81	0.19	91	0.20
Delaware R.	48	0.65	0.34	0.00	0.01	76	0.00	0.94	0.00	0.05	0.01	76	−0.06	0.99	0.01	0.00	76	−0.01	0.80	0.20	76	0.09
Delaware R.	49	0.72	0.27	0.01	0.01	85	0.09	0.92	0.00	0.04	0.04	86	0.10	0.97	0.03	0.00	86	−0.03	0.80	0.20	86	−0.18
Delaware R.	50	0.61	0.39	0.00	0.00	99	−0.08	0.92	0.00	0.02	0.06	134	0.02	0.96	0.03	0.02	135	0.14	0.81	0.19	134	−0.09
Hudson R.	51	0.69	0.31	0.00	0.00	32	0.13	0.96	0.00	0.00	0.04	56	−0.04	0.97	0.01	0.02	56	−0.02	0.77	0.23	56	0.30*
British Is.	53	1.00	0.00	0.00	0.00	7	0.00	1.00	0.00	0.00	0.00	7	0.00	1.00	0.00	0.00	7	0.00	0.29	0.71	7	0.30
Total for 53 samples						2495						3253					3238				3318	

TABLE 2. Temporal variation in sea lamprey samples as indicated by G-test for heterogeneity of allele frequencies between paired samples at individual sites (*$p < 0.05$; **$p < 0.01$; ND = no data).

Location	Site No.	*AGP* G	df	*PGI-2* G	df	*PGM-1* G	df	*MDH-1* G	df
Ammocoete samples collected in successive years 1979–80									
Lake Ontario	35	0.05	1	0.81	2	1.84	1	16.01**	1
Oneida Lake	39	0.90	1	9.88**	2	0.17	1	0.30	1
Cayuga Lake	40	1.05	1	1.77	2	0.86	1	1.15	1
Seneca Lake	41	1.13	1	0.05	1	0.00	0	0.00	0
Lake Champlain	46	0.00	1	4.38*	1	0.00	1	0.08	1
Delaware River	47	1.36	2	11.41**	3	2.68	2	0.01	1
	50	0.24	1	3.47	2	1.75	2	3.09	1
Hudson River	51	0.07	1	2.25	2	2.64	2	0.38	1
Adult and ammocoete samples collected simultaneously									
Cayuga Lake 1979	40	1.94	1	4.35	2	0.06	1	3.56	1
Cayuga Lake 1980	40	ND		2.32	2	0.37	1	0.02	1
Lake Ontario	38	ND		5.05	2	1.31	1	1.35	1

TABLE 3. Temporal and spatial variance components in the frequency of the common (*"100"*) allele at each locus for the paired sea lamprey samples in Table 2 (—, insufficient data).

Factor	% of total variance explained by factor *AGP*	*PGI-2*	*PGM-1*	*MDH-1*
Ammocoete samples collected in successive years 1979–80				
Site	89.18	56.20	76.77	87.73
Year	0.00	5.99	0.00	0.00
Error	10.82	37.81	23.23	12.27
Adult and ammocoete samples collected simultaneously				
Site	—	18.04	76.92	68.15
Stage	—	72.45	18.02	2.82
Error	—	9.51	5.06	29.02

TABLE 4. Components of genetic variability in sea lamprey. H_T, the total expected heterozygosity, measures overall variation at a locus. Succeeding columns partition the proportion of H_T that can be attributed to differentiation among samples within a single lake or river (F_{SL}), differentiation among freshwater inland lakes or among anadromous coastal river populations (F_{LW}), and differentiation between inland and coastal systems (F_{WT}). Residual proportion ($1 - F_{ST}$) represents within-sample variation. **$p < 0.01$, heterogeneity G-test of allele frequencies among subunits of the total at each level.

Locus	N	H_T	F_{SL}	F_{LW}	F_{WT}	$(1 - F_{ST})$
AGP	2493	0.483	0.030**	0.031**	0.004**	0.939
PGI-2	3301	0.162	0.019**	0.048**	0.004**	0.935
PGM-1	3287	0.300	0.023**	0.124**	0.032**	0.805
MDH-1	3318	0.219	0.022**	0.041**	0.011**	0.928
Mean		0.291	0.024	0.061	0.013	0.902

Seneca and Champlain, as well as in Lake Erie, was relatively low and approximated site-to-site differentiation in the coastal Delaware River system. The Great Lakes showed more within-lake differentiation, with highest value for Lake Huron (F_{SL} = 0.051) and lesser ones for Lakes Superior, Michigan, and Ontario (F_{SL} = 0.040, 0.020, and 0.016, respectively; Table 5).

TABLE 5. Geographic variability in genetic relationships among population units of sea lamprey. F_{LW} values are the contributions of coastal and inland systems to weighted mean F_{LW} of Table 4, and F_{SL} values are the contributions of individual lakes and rivers to weighted mean among-site variation (F_{SL}) of Table 4. Nei's genetic distance among sites was calculated from data for the four polymorphic loci. (F_{SL} and genetic distance values were not calculated for the Connecticut and Hudson Rivers nor for Oneida and Cayuga lakes, since each was represented by a single sample site.)

Population unit	No. of samples	F-statistic	Avg. genetic distance among samples
Coastal rivers	6	0.009 (F_{LW})	0.002
Delaware River	4	0.006 (F_{SL})	0.002
Inland lakes	46	0.071 (F_{LW})	0.038
L. Huron	5	0.051 (F_{SL})	0.025
L. Superior	18	0.040 (F_{SL})	0.019
L. Michigan	6	0.020 (F_{SL})	0.010
L. Erie	2	0.004 (F_{SL})	0.005
All samples above Niagara Falls	31	0.042	0.019
L. Ontario	7	0.016 (F_{SL})	0.007
L. Champlain	4	0.006 (F_{SL})	0.002
Seneca L.	2	0.000 (F_{SL})	0.000
Seneca–Cayuga–Oneida–Ontario samples	11	0.023	0.012
All N. American samples below Niagara Falls	21	0.052	0.020
All *Petromyzon marinus*	53	0.098 (F_{ST})	0.045

Affinities among sea lamprey population units, based on cluster analysis of Nei's distances (Fig. 2), were congruent with the results from the F-statistics analysis. In the dendrogram, the original 53 samples were reduced to 27, for clarity, by pooling samples that formed the homogeneous collections from the STP tests (Fig. 1). Several patterns were evident from this analysis:

(1) Closer affinity of freshwater Lake Champlain populations to coastal sea-run lampreys than to other freshwater populations.

(2) Closer affinity of populations in inland New York lakes Cayuga and Oneida to those of Lake Ontario than to nearby Seneca Lake.

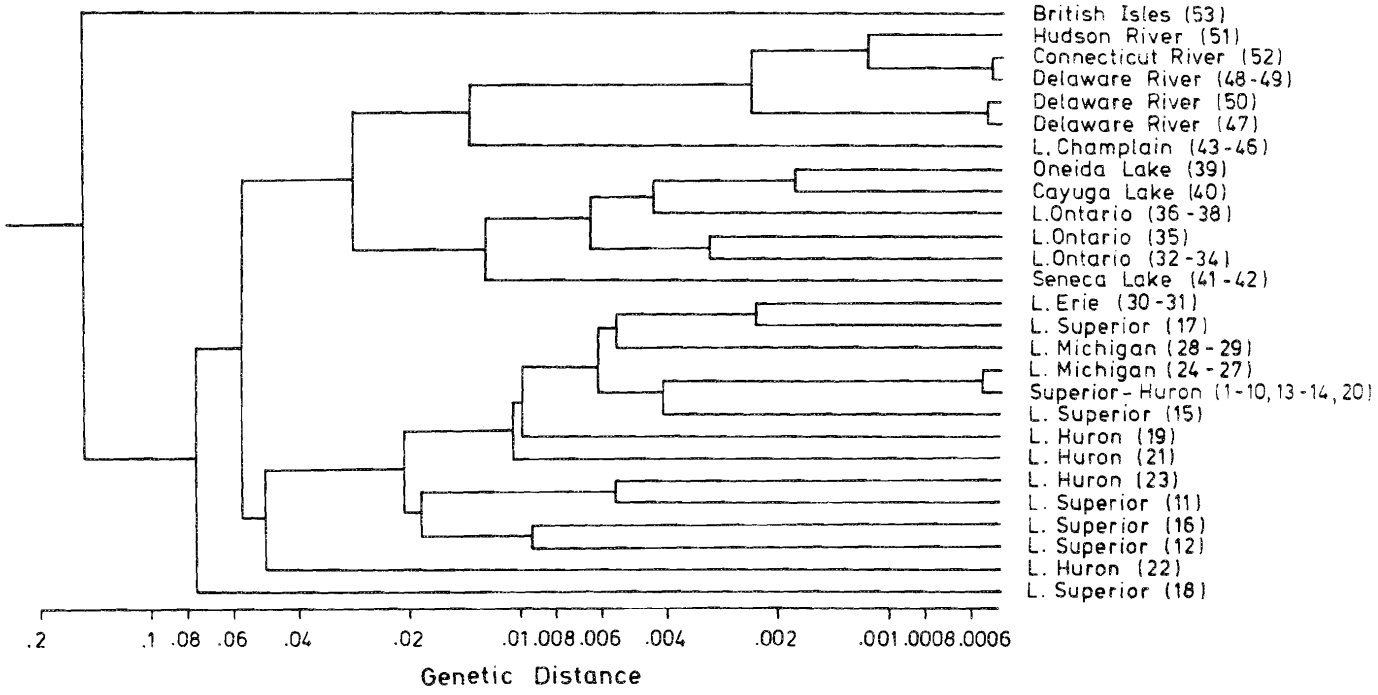

FIG. 2. UPGMA dendrogram of relationships among sea lamprey populations based on Nei's genetic distance at four polymorphic loci. Numbers in parentheses following locality are site designations (see Fig. 1).

(3) Close genetic similarity of upper Great Lakes collections.

(4) Closer affinity of Lake Erie to upper Great Lakes populations than to those of Lake Ontario.

(5) Closer affinity of inland New York lakes and Lake Ontario populations to Lake Champlain and the coastal sea-run system than to lampreys of upper Great Lakes.

(6) Closer affinity of all North American lampreys to each other than to a sea-run sample from the British Isles.

Discussion

Spatial Differentiation of Populations

Variation within samples accounted for 90% of the observed variation in sea lamprey populations (Table 4; summarized in Fig. 3); thus, the average sample of *Petromyzon marinus* from a single site would contain the majority of the genetic variability of the species as a whole. This proportion is typically high in other species as well. Within-sample variation approaches 100% in highly vagile organisms such as migrant monarch butterflies (*Danaus plexippus*) over their summer range (99.6%; Eanes and Koehn 1978), but even in such notably philopatric or sedentary species as red-bellied newts (*Taricha rivularis*) in southern California (Hedgecock 1978) and *Helix* snails among cities (Selander and Kaufman 1975), it can be substantial (97.4 and 83.8%, respectively).

Nevertheless, for any species it is the remainder of the genetic variation that provides the basis for determining how populations are structured and interrelated. It was this residual variation in sea lampreys that we partitioned into its components by means of hierarchical *F*-statistics (Fig. 3). In sea lampreys, the largest source of residual variation (6% of total variation) was the differentiation from lake to lake and among coastal rivers (Table 4; Fig. 3). With an *F*-statistics approach similar to ours, Avise and Felley (1979) examined genetic differentiation among reservoir populations of bluegills (*Lepomis macrochirus*) at three hierarchical levels. They found differentiation in this species to be greatest among reservoirs and less pronounced

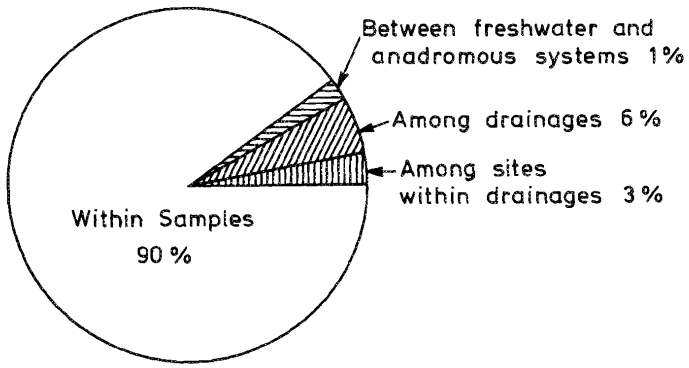

FIG. 3. Summary of partitioned genetic variation in sea lamprey based on *F*-statistic analysis of genetic data from 53 collections.

at the next higher (between river systems) and next lower (among-site but within-reservoir) levels. Their results and ours suggest that genetic mixing of fish populations is limited as expected by physical barriers but is apparently relatively insensitive to geographic distance alone.

The second largest source of residual variation (close to 3% of total) was that attributable to lamprey differentiation within lakes and rivers (Table 4; Fig. 3). Sea lamprey samples do not appear equally differentiated in all lakes, however (Table 5). Statistically significant differences in allelic frequencies were most apparent among collections within Lakes Huron, Michigan, and Ontario and western Lake Superior, whereas collections within other areas showed little or no significant differentiation (Fig. 1). Low levels of differentiation, as we found in the genetically homogeneous clusters of Fig. 1, are most logically and parsimoniously attributable to relative panmixis as a result of high local effective population size, high numbers of migrants, or both (Allendorf and Phelps 1981). Parallel selection pressures acting on kinetic variants of these polymorphic enzymes, however, cannot totally be ruled out (see Powers and Place 1978).

High levels of site-to-site differentiation within Lakes Huron, Michigan, Superior, and Ontario are more difficult to interpret. One possible explanation is that each lake is composed of multiple small populations that exchange few genes (Krueger and Spangler 1981). A second possibility is that each stream collection of ammocoetes is not representative of the spawning adults that use that stream, perhaps because a large percentage of ammocoetes may be the offspring of only a few adult pairs. Jacobson et al. (1984) favored the latter interpretation and showed that a statistically significant amount of genetic variation exists among collection sites and year classes within single drainages to Lakes Michigan and Huron. Since we found only slight changes at a site from year to year in our samples (Tables 2 and 3), we feel that these alternative hypotheses deserve further study. Whichever explanation is correct, the fact remains that ammocoete sites within some bodies of water are locally genetically differentiated, whereas in others they are not. Avise and Felley's (1979) hierarchical data for bluegills similarly showed some reservoirs to be much more internally differentiated than others. What these *F*-statistics indicate is that, over a species' range, degree of spatial differentiation is not a fixed characteristic but varies from place to place. In terms of population management or control, baseline data from the particular geographic area of interest are necessary for understanding of effective population size and identification of populations. We suggest that if genetic data imply presence of multiple populations in one lake and their absence in another, this pattern may well reflect real differences in population structure from lake to lake rather than contradictory results.

A number of studies have calculated site-to-site F_{ST} in fish species. It might appear that these statistics could be used to compare population structure among species (e.g. Winans 1980), but our analysis shows that without regard for the hierarchical way in which populations might be arranged, such comparisons make little biological sense. An F_{ST} value for sea lampreys calculated from all our samples would have been 0.098 (Table 5). We know from the hierarchical analysis, however, that this value would be a composite of at least three sources of differentiation: divergence from lake to lake, differentiation or sampling variance among local populations, and the split between sea-run and freshwater systems. The extent to which each of these sources of divergence contributes to differentiation among a given set of samples depends on the spatial scale on which the sampling was done. This may explain why attempts to associate species-to-species *F* values with such features as chromosomal diversity (Sites and Greenbaum 1983), larval dispersal probability (Winans 1980), and adult vagility (Eanes and Koehn 1978) have not been particularly successful. We suggest that *F*-statistics (F_{ST}, or gene diversity statistics, G_{ST}) are best used not for comparing one species with another but for detecting levels or areas of relatively greater or lesser differentiation within a species (e.g. Chesser 1983).

Genetic distance as shown in the dendrogram (Fig. 2) gives a different perspective on interrelationships among sea lamprey population units. Because genetic distance is calculated directly from allele frequencies, relationships among small samples with high sampling variance can be distorted. Krueger and Spangler (1981) reduced this confusion by pooling samples of $D < 0.06$ (Rogers' distance coefficient) before clustering Lake Superior samples; here instead we have pooled the groups of statistically homogeneous neighboring samples from the earlier part of the statistical analysis. This difference, as well as the addition of more samples, results in an overall dendrogram that is slightly different in detail from previous ones for Lake Superior (Krueger and Spangler 1981) and eastern populations (Brussard et al. 1981) but, in general, the pattern is virtually identical: for Lake Superior, an eastern and a western cluster with one aberrant sample (Isle Royale); for eastern lampreys, an anadromous – Lake Champlain cluster and another encompassing Lake Ontario and the New York Finger Lakes. In addition, analysis of all 53 sites in the present study shows a rather complex pattern of interlake affinities and makes clear a distinct genetic split between lamprey populations above and below Niagara Falls.

Genetic Divergence and Population History

An issue that originally interested us was the degree of endemicity of sea lampreys within New York State (Brussard et al. 1981). Although recent introduction of lampreys into the upper Great Lakes is well documented (Smith and Tibbles 1980) and introduction into New York waters via the canal system is widely assumed, there is a possibility that sea lampreys may have occupied New York waters since the last glacial retreat (D. A. Webster, pers. comm.). If so, they might best be regarded as part of a system coevolved with native salmonids and calling for different management perspectives than those applied in the upper Great Lakes (Lamsa et al. 1980; Pearce et al. 1980).

One method available for projecting population history from genetic data is Nei's (1975, p. 193) formula for time since divergence as a linear function of genetic distance, in isolated equilibrium population units subjected to a fixed gene substitution rate. Ferguson and Mason (1981) and Ryman and Ståhl (1981), with apparent success, have used Nei's formula with allozyme data to estimate divergence time for populations of brown trout (*Salmo trutta*) and of Atlantic salmon (*Salmo salar*). We attempted to apply this technique to sea lamprey populations, but even under the conservative assumption that as many as 85% of loci were monomorphic (Krueger and Spangler 1981), it produced divergence-time estimates far too large to be accurate. In some cases it projected that within-lake populations had differentiated several thousand years before the lakes themselves were exposed in the last glacial retreat. For the sea lamprey system, the assumptions of Nei's model (isolation, no selection, no genetic bottlenecks, no migration) were apparently violated so far as to make the model inapplicable. We mention this result as a caution against relying on Nei's estimate in cases where the validity of its assumptions cannot be assessed.

Although theoretical divergence time was not useful, the expanded data base available allowed us to compare New York lampreys of uncertain history with those of Lakes Michigan, Huron, and Superior, where invasion of sea lampreys from Lake Ontario has been "pinpointed" at 50 yr ago (Smith and Tibbles 1980). The observed genetic distance among these upper Great Lakes populations serves as an empirical standard for comparison with New York populations without the necessity for equilibrium assumptions. Examined in this way, populations of sea lamprey of the Cayuga – Seneca – Oneida – Lake Ontario complex were no more genetically distant from each other than were upper Great Lakes populations which began to diverge half a century ago (Table 5). On the basis of divergence between Lake Superior and the New York lakes, we earlier (Brussard et al. 1981) favored the endemicity hypothesis for New York lampreys; but the present analysis based on genetic structure throughout eastern North America forces the conclusion that either the introduction of sea lampreys into the eastern lakes is

quite recent, or the population-differentiating process among these lakes has been strikingly retarded relative to that in the upper Great Lakes. We have showed that population structure for this species may well vary from region to region, but differences of the magnitude that would be required to make endemic populations less differentiated than newly introduced ones seem unlikely. Possibly the recent connections among these water bodies (e.g. the Erie Barge Canal, etc.) may have allowed gene flow among previously isolated, more genetically distinct, populations (see Morman et al. 1980). Another possible explanation is that the extensive sea lamprey control program in the upper Great Lakes since 1958 has speeded the differentiation process through periodic catastrophic reduction of population numbers in localized areas. Thus, the issue of endemicity of sea lampreys in New York inland waters remains incompletely resolved.

Management Implications

The genetic differentiation of lampreys above and below Niagara Falls suggests future research that may have application to the present sea lamprey control program in the Great Lakes. We recommend that an investigation be conducted into factors that may restrict movement of lampreys between Lakes Ontario and Erie through the Welland Canal. The genetic differentiation observed between the upper Great Lakes and Lake Ontario lampreys probably resulted from "founder effects" by Lake Ontario lampreys colonizing Lake Erie. These effects, which caused significant differences in allele frequencies, were probably then maintained by some type of restriction to movement and subsequent mating of lampreys between the two lakes. The completion of the Welland Canal in 1829 provided access to Lake Erie around Niagara Falls for Lake Ontario lampreys (Morman et al. 1980). Although lampreys were recorded in Lake Ontario at that time (Lark 1973), they did not appear to colonize Lake Erie until much later. The first lamprey recorded from Lake Erie was in 1921 (Dymond 1922) and spawning runs were first noted in 1932 (Creaser 1932). The Welland Canal was made wider and deeper in 1851 and 1881 (Atkinson 1979) and these modifications may have improved access for the lampreys. Substantial restriction of lamprey movement through the canal, however, must still exist, since significant genetic differences occur between Lake Erie and Lake Ontario sea lampreys. Understanding the cause for the lack of interbreeding (and hence, probably movement) between lampreys in these two lakes may suggest new control measures to restrict movement.

Acknowledgments

We are grateful to the field personnel of the U.S. Fish and Wildlife Service, the Canada Department of Fisheries and Oceans, and the New York State Department of Environmental Conservation for helping us obtain specimens and providing equipment. Dr. Peter S. Maitland, Institute of Terrestrial Ecology, Scotland, graciously supplied lampreys from the British Isles for this study. Vicky Meretsky, John Wilson, Mark Angevine, James Stewart, Donald Schreiner, and James Luey assisted in field and laboratory. Richard Baker, Robert Lacy, L. D. Jacobson, and Michael Soulé made helpful comments on the manuscript. We thank Robert Lacy for discussion and advice regarding *F*-statistics, L. D. Jacobson for access to unpublished data, and Dwight Webster for his long-standing support of this project. This research was coordinated on behalf of the Great Lakes Fishery Commission by W. L. Hartman and J. Kutkuhn of the Great Lakes Fishery Laboratory, U.S. Fish and Wildlife Service. This study was generously supported by the Great Lakes Fishery Commission, Minnesota Sea Grant Program, and Minnesota Agriculture Experiment Station through grants to P. F. Brussard and G. R. Spangler.

References

ALLENDORF, F. W., AND S. R. PHELPS. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquat. Sci. 38: 1507–1514.

ATKINSON, C. H. 1979. Aspects of engineering of the Welland Canals, p. 35–64. *In* J. Bartniak and W. B. Turner [ed.] Proc. First Annual Niagara Peninsula History Conference, Brock University, St. Catharines, Ont.

AVISE, J. C., AND J. FELLEY. 1979. Population structure of freshwater fishes. I. Genetic variation of bluegill (*Lepomis macrochirus*) populations in man-made reservoirs. Evolution 33: 15–26.

BRUSSARD, P. F., M. C. HALL, AND J. WRIGHT. 1981. Structure and affinities of freshwater sea lamprey (*Petromyzon marinus*) populations. Can. J. Fish. Aquat. Sci. 38: 1708–1714.

CHESSER, R. K. 1983. Genetic variability within and among populations of the black-tailed prairie dog. Evolution 37: 320–331.

CREASER, C. W. 1932. The lamprey *Petromyzon marinus* in Michigan. Copeia 1932: 157.

DYMOND, J. R. 1922. A provisional list of the fishes of Lake Erie. Univ. Toronto Stud. Biol. Ser. Publ. Ont. Fish Res. Lab. 4: 57–73.

EANES, W. F., AND R. K. KOEHN. 1978. An analysis of genetic structure in the monarch butterfly, *Danaus plexippus* L. Evolution 32: 784–797.

FERGUSON, A., AND F. M. MASON. 1981. Allozyme evidence for reproductively isolated sympatric populations of brown trout *Salmo trutta* L. in Lough Melvin, Ireland. J. Fish Biol. 18: 629–642.

HEDGECOCK, D. 1978. Population subdivision and genetic divergence in the red-bellied newt, *Taricha rivularis*. Evolution 32: 271–286.

JACOBSON, L. D., R. L. TORBLAA, AND R. H. MORMAN. 1984. Problems with samples of sea lamprey (*Petromyzon marinus*) ammocoetes used in stock identification studies. Can. J. Fish. Aquat. Sci. 41: 1421–1427.

KRUEGER, C. C. 1980. Detection of variability at isozyme loci in sea lamprey, *Petromyzon marinus*. Can. J. Fish. Aquat. Sci. 37: 1630–1634.

KRUEGER, C. C., AND G. R. SPANGLER. 1981. Genetic identification of sea lamprey (*Petromyzon marinus*) populations from the Lake Superior Basin. Can. J. Fish. Aquat. Sci. 38: 1832–1837.

LAMSA, A. K., C. M. ROVAINEN, D. P. KOLENOSKY, AND L. H. HANSON. 1980. Sea lamprey (*Petromyzon marinus*) control — Where to from here? Report of the SLIS Control Theory Task Force. Can. J. Fish. Aquat. Sci. 37: 2175–2192.

LARK, J. G. I. 1973. An early record of the sea lamprey (*Petromyzon marinus*) from Lake Ontario. J. Fish. Res. Board Can. 30: 131–133.

MORMAN, R. H., D. W. CUDDY, AND P. C. RUGEN. 1980. Factors influencing the distribution of sea lamprey (*Petromyzon marinus*) in the Great Lakes. Can. J. Fish. Aquat. Sci. 37: 1811–1826.

NEI, M. 1975. Molecular population genetics and evolution. American Elsevier Publishing Co., New York, NY. 288 p.

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

PARSONS, P. A. 1983. The evolutionary biology of colonizing species. Cambridge University Press, Cambridge. 262 p.

PEARCE, W. A., R. A. BRAEM, S. M. DUSTIN, AND J. J. TIBBLES. 1980. Sea lamprey (*Petromyzon marinus*) in the lower Great Lakes. Can. J. Fish. Aquat. Sci. 37: 1802–1810.

POWERS, D. A., AND A. R. PLACE. 1978. Biochemical genetics of *Fundulus heteroclitus* (L). I. Temporal and spatial variation in gene frequencies of *Ldh-B*, *Mdh-A*, *Gpi-B*, and *Pgm-A*. Biochem. Genet. 16: 593–607.

RYMAN, N., AND G. STÅHL. 1981. Genetic perspectives of the identification and conservation of Scandinavian stocks of fish. Can. J. Fish. Aquat. Sci. 38: 1562–1575.

SELANDER, R. K., AND D. W. KAUFMAN. 1975. Genetic structure of populations of the brown snail (*Helix aspersa*). I. Microgeographic variation. Evolution 29: 385–401.

SITES, J. W., AND I. F. GREENBAUM. 1983. Chromosome evolution in the Iguanid lizard *Sceloporus grammicus*. II. Allozyme variation. Evolution 37: 54–65.

SMITH, B. R., AND J. J. TIBBLES. 1980. Sea lamprey (*Petromyzon marinus*) in Lakes Huron, Michigan and Superior: history of invasion and control. Can. J. Fish. Aquat. Sci. 37: 1780–1801.

SMITH, B. R., J. J. TIBBLES, AND B. G. H. JOHNSON. 1974. Control of the sea lamprey (*Petromyzon marinus*) in Lake Superior, 1953–70. Great Lakes Fish. Comm. Tech. Rep. 76: 60 p.

SNEATH, P. H. A., AND R. R. SOKAL. 1973. Numerical taxonomy. W. H. Freeman, San Francisco, CA. 573 p.

SNEDECOR, G. W., AND W. G. COCHRAN. 1967. Statistical methods. 6th ed. Iowa State University Press, Ames, IA. 593 p.

SOKAL, R. R., AND F. J. ROHLF. 1981. Biometry. 2nd ed. W. H. Freeman, San Francisco, CA. 859 p.

WAHLUND, S. 1928. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas 11: 65–106.

WINANS, G. A. 1980. Geographic variation in the milkfish *Chanos chanos*. I. Biochemical evidence. Evolution 34: 558–574.

WRIGHT, S. 1965. The interpretation of population structure by *F*-statistics with special regard to systems of mating. Evolution 19: 395–420.

1978. Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago, IL. 580 p.

APPENDIX. Location of sea lamprey collections analyzed (M = University of Minnesota; C = Cornell University).

Sample No.	Source of data	Locality	State or province	Latitude N	Longitude W
1	M	Nipigon River	Ont.	49°04′	88°17′
2	M	Cypress River	Ont.	48°56′	87°52′
3	M	Mountain Bay	Ont.	48°55′	87°48′
4	M	Michipicoten River	Ont.	47°55′	84°50′
5	M	Goulais River	Ont.	46°44′	84°27′
6	M	Tahquamenon River	MI	46°36′	85°12′
7	M	Sucker River	MI	46°36′	85°57′
8	M	Au Train River	MI	46°25′	86°50′
9	M	Big Garlic River	MI	46°41′	87°35′
10	M	Huron River	MI	46°55′	88°06′
11	M	Sturgeon River	MI	46°47′	88°37′
12	M	Traverse River	MI	47°13′	88°16′
13	M	Misery River	MI	46°59′	88°59′
14	M	Firesteel River	MI	46°55′	89°11′
15	M	Bad River – Potato River	WI	46°28′	90°36′
16	M	Nebagamon River	WI	46°32′	91°36′
17	M	Middle River	WI	46°39′	91°48′
18	M	Washington Creek	MI	47°56′	89°08′
19	M	Little Munuscong River	MI	46°16′	84°21′
20	M	Garden River	Ont.	46°33′	84°09′
21	M	Missisagi River	Ont.	46°10′	83°02′
22	M	Sturgeon River	Ont.	44°44′	79°45′
23	M	Mill Creek	MI	43°13′	82°32′
24	M	Big Manistee River	MI	44°15′	86°20′
25	M	Ford River	MI	45°41′	87°08′
26	M	Whitefish River	MI	46°09′	86°55′
27	M	Fishdam River	MI	45°54′	86°35′
28	M	Marblehead Creek	MI	45°58′	86°07′
29	M	Black River	MI	46°06′	85°21′
30	C	Crooked Creek	PA	41°59′	80°36′
31	C	Cattaraugus Creek	NY	42°26′	78°50′
32	C	Duffin Creek	Ont.	43°46′	79°00′
33	C	Graham Creek	Ont.	43°46′	78°55′
34	M	Shelter Valley Creek	Ont.	43°58′	78°00′
35	C	Little Sandy Creek	NY	43°38′	76°10′
36	C	Lindsey Creek	NY	43°41′	76°00′
37	C	Trout Brook	NY	43°35′	76°03′
38	C	Grindstone Creek	NY	43°33′	76°10′
39	C	Fish Creek	NY	43°22′	75°50′
40	C	Cayuga Inlet Creek	NY	42°25′	76°28′
41	C	Catherine Creek	NY	42°22′	76°52′
42	C	Keuka Lake Outlet	NY	42°41′	76°56′
43	C	Great Chazy River	NY	44°57′	73°26′
44	C	Little Ausable River	NY	44°35′	73°28′
45	M	Lewis Creek	VT	44°15′	73°16′
46	C	Putnam Creek	VT	43°57′	73°27′
47	C	Oquaga Creek	NY	42°07′	75°29′
48	C	E. Branch Delaware River	NY	42°03′	75°07′
49	C	Beaver Kill	NY	41°56′	75°05′
50	C	Bouchouxville Creek	NY	41°52′	75°08′
51	C	Roeliff – Jansen Kill	NY	41°58′	73°55′
52	C	Holyoke Dam	MA	42°12′	72°36′
53	M	British Isles: River Kent (6)		54°14′	02°49′
		River Leven (1)		55°58′	04°36′