GREAT LAKES FISHERY COMMISSION

1983 Project Completion Report[1]

Sea Lamprey (*petromyzon marinus)* in Northeastern North America: Population Structure and Genetic Affinities

by:

Peter F. Brussard[2] and Janet Wright[2]

[2]Ecology & Systematics
Corson Hall
Cornell University
Ithaca, NY 14853-0239

July 1983

[1]Project completion reports of Commission-sponsored research are made available to the Commission's Cooperators in the interest of rapid dissemination of information that may be useful in Great Lakes fishery management, research, or administration. The reader should be aware that project completion reports have not been through a peer review process and that sponsorship of the project by the Commission does not necessarily imply that the findings or conclusions are endorsed by the Commission.

Final rept.

SEA LAMPREY (*Petromyzon marinus*) IN NORTHEASTERN NORTH AMERICA: POPULATION STRUCTURE AND GENETIC AFFINITIES

Peter F. Brussard and Janet Wright

Ecology & Systematics
Corson Hall
Cornell University
Ithaca, NY 14853-0239

ORIGINAL REPORT

(For Duplication Only)

Final program report to the Great Lakes Fishery Commission
July 27, 1983

INTRODUCTION

Freshwater populations of sea lamprey, Petromyzon marinus, show considerable genetic variation detectable by electrophoretic separation of enzyme variants. This variation has been reported to indicate regional stocks in Lake Superior (Krueger and Spangler 1981), but similar data from Lakes Michigan and Huron have not revealed such stocks (L. Jacobson, pers. comm.). In a different context, electrophoretic data have been used to indicate genetic affinities among sea lamprey of New York waters but have not shown unequivocally whether lamprey in inland New York lakes are endemic or recently introduced (Brussard et al. 1981). Here we consider electrophoretic variation among collections of sea lamprey at 53 sites over a broader geographic area encompassing northeastern North America and the British Isles with the aim of resolving these issues and characterizing overall variation in this highly variable species in both landlocked and anadromous populations.

METHODS AND MATERIALS

The present analysis is based on collections of Petromyzon marinus 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 New York sites and 22 sites not previously described.

Most lamprey were collected as ammocoetes by electroshocking in short sections of silty stream habitat (usually not more than 100 m); a few specimens were collected as spawning adults. Details of collection have been reported already (Brussard et al. 1981, Krueger and Spangler 1981).

Enzyme variants of individuals were resolved using horizontal starch gel electrophoresis according to previously published methods (Krueger 1980, Brussard et al. 1981). Krueger (1980) reported slight variation at the PHI-2 (phosphohexose isomerase) locus in some samples, and Brussard et al. (1981) found variation in PGM-1 (phosphoglucomutase) and occasionally in IDH-1 (isocitrate hydrogenase); but for purposes of standardization, the present analysis is based upon the four polymorphic loci scored in common among all samples: alpha-glycerophosphate dehydrogenase (AGP), PHI-1, PGM-2, and malate dehydrogenase (MDH-2). Bands of differing mobility are assumed to represent allelic variants; designations are based on mobility relative to a value of 100 for the common electromorph.

Genotype frequencies in individual samples were tested for fit to Hardy-Weinberg expectation (log-likelihood-test, deviation considered significant at $p < .05$). At some sites samples were taken in successive years or at two life stages; such paired samples were examined for significant differences (log-likelihood test for heterogeneity, $p < .05$). Those samples at a site that showed no heterogeneity at any locus were pooled for the geographic analysis; in the cases with between-year or between-stage heterogeneity, only the larger sample was used.

To examine the question of differences in population structure from place to place, variation was partitioned by an F-statistics (gene diversity) analysis (Wright 1978, Nei 1975). Populations were first arrayed in an hierarchic scheme: samples within each of 13 lakes or rivers; lakes or rivers within each of two water systems (landlocked, sea-run), and landlocked or sea-run systems within the total. An F value was calculated (Wright 1965, as interpreted by Nei 1977) at each level of the hierarchy (F_{SL}, F_{LW}, and F_{WT} respectively). This approach estimates the proportion of total

observed genetic variation at a level that can be attributed to differentiation among the subunits of that level. Thus for the landlocked/sea-run water systems as subunits and all lamprey as total,

$$F_{WT} = (H_T - H_W)/H_T \ ,$$

where H_T is expected heterozygosity in the total "population for the level ($H_T = 1 - p_i^2$, p_i = freq. of allele i) and H_W is weighted mean of expected heterozygosities of each of the two subunits (Nei 1977). In analogous fashion, a weighted mean F for lakes or rivers within systems or samples within lakes or rivers was taken to characterize differentiation at those levels. (Subunits represented by a single sample with resulting F necessarily equal to zero were excluded from the mean calculation.)

In addition, the individual F's within each level were examined comparatively to answer such questions as: Is there greater differentiation among lake populations of landlocked lamprey than among river populations of the anadromous system? Do all lakes show equal levels of differentiation, or can some be subdivided into genetically distinct stocks?

To examine affinities and historical patterns among lamprey populations, Nei's (1977) genetic distance among each of the pairs of populations was calculated over the four polymorphic loci, and Unweighted Pair Group Method Analysis was applied to produce a dendrogram of distances among the samples. To clarify these rather complicated relationships, a second dendrogram was produced, in which the number of population units was first reduced by pooling groups of geographically contiguous samples that showed (by heterogeneity G-test) no significant differences at any locus. Finally, Nei's (1975) model of time divergence and calculations based on "known" introductions were used to compare patterns of differentiation in New York waters with those in the upper Great Lakes in order to estimate probable time of colonization of New York inland lakes.

RESULTS

Fifty-three collecting localities are represented in this analysis (Figure 1). Table 1 shows allele frequencies at the four polymorphic loci and deviation from Hardy-Weinberg expectation (F_{IS}) at each locus . Of the total of 212 G-tests for the 53 samples, 11 departed significantly from expectation; nine of these involved heterozygote deficiency.

A useful measure of total variation in a species is expected heterozygosity over all samples and all loci ($\bar{H}$). For the four polymorphic loci reported here, overall expected heterozygosity is 0.256 If the data for monomorphic loci (17 loci in Cornell-collected samples, 21 in U. Minneosta ones) are included, $\bar{H}$ for Petromyzon marinus is estimated as 0.044 the proportion of loci heterozygous in an average individual.

Several measures of the temporal variation that occurs between collections from a single site appear in Table 2. The analysis of variance result partitions variance into among-site as opposed to between-year or between-life-stage components; only the among-site portions are significant.
An F-statistic that expresses the proportions of total variation in this set of samples attributable to between-year (F_{YS}) or between-stage (F_{AS}) variation has been calculated for comparison to the F-statistics expressing levels of geographic variation. In addition, the average genetic distance value between collections in successive years or between adult-ammocoete stages can be compared to the distance measure calculated among sites. In the remaining results each site is represented by a single sample (the pooled year samples if genetically homogeneous, the larger year sample if not).

Results of the hierarchical F-statistics analysis over geographic localities are shown in Table 3. For each polymorphic locus, and for all loci averaged, the F_{ST} values partition observed genetic variability into between-system,

among-lake or river, and within-lake or river components. Variability unaccounted for at these three levels (i.e. the $1-F_{ST}$ column) is the proportion of variation attributable to within-sample variation (0.9018) (Fig. 2). Differentiation among lakes and rivers F_{LW} is of greater magnitude than site-to-site variation within lakes and rivers (F_{SL}) which in turn is greater than differentiation between anadromous and landlocked systems (F_{WT}; Table 3). F-values for all the components of geographic variation are from 2 to 10 times as great as the average values for temporal differentiation at a site (Table 2).

While Table 3 expresses average amounts of differentiation at each level of the hierarchy, these averages mask considerable variation in F value within each level (Table 4). For example, river-to-river differentiation is much less for anadromous populations (F_{LW} = 0.0092) than is lake-to-lake differentiation among freshwater ones (F_{LW} = 0.0713) within the anadromous system; site-to-site variation (as seen in the Delaware River, F_{LW} = 0.0056) is also very low. Among freshwater lakes, almost no site-to-site variation is evident in some (Seneca Lake, Lake Champlain and Lake Erie), but others (the remaining Great Lakes) show relatively great within-lake differentiation, with the highest value for Lake Huron (F_{SL} = 0.0511) and lesser ones for Lake Superior (F_{SL} = 0.0403) and Lake Michigan (F_{SL} = 0.0203).

The possibility arises for these latter cases that site-to-site differentiation is not evenly spread over the lake but that there may be some regions where sites are genetically more alike. To detect such regional effects we tested clusters composed of neighboring sites for statistical heterogeneity; resulting "within-region" F's for the most comprehensive clusters thus identified are shown in the last column of Table 4.

Lakes were individualistic in their regional patterns. For example, Lake Huron samples were genetically heterogeneous. Those in Lake Superior fell into an undifferentiated eastern group and a heterogeneous western group. Lake Michigan was divisible into a northern cluster of two populations and one of four populations farther south; and in Lake Ontario, northwestern samples formed one homogeneous group and three eastern ones another, while a remaining sample clumped with neither group. Extending this approach over all lakes, we identified a group of samples, in the area where Lakes Superior, Michigan and Huron conjoin, in which there was no heterogeneity at any locus ("Three-lake region"' Table 4). Farther away from this conjunction in each of the lakes, samples again became heterogeneous.

Overall genetic relationships among sea lamprey population units are expressed in the genetic distance dendrogram in Figure 3. The original 53 samples have been reduced to 27 by pooling statistically homogeneous samples within each drainage (the homogeneous "regions" outlined above, except that the "Three-lake " group is diagrammed as a Superior-Huron and separate Lake Michigan cluster). Several patterns are evident here, including:

1. Closer affinity of Connecticut River to Hudson River populations than to those of the Delaware River.

2. Closer affinity of Lake Champlain populations to sea-run lamprey than to other freshwater populations.

3. Closer affinity of Lake Erie to upper Great Lakes populations than to those of Lake Ontario.

4. Closer affinity of Cayuga and Oneida Lakes populations to those of Lake Ontario than to Seneca Lake.

5. Closer affinity of Finger Lakes and Lake Ontario populations to Lake Champlain and the sea-run system than to lamprey of upper Great Lakes.

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

7. Lack of correspondence between natural drainage units and genetic clusters in some areas even after statistically non-heterogeneous adjacent localities were combined.

DISCUSSION

The estimates of overall genetic variability for Petromyzon marinus are approximately the same as those made for other fishes (0.051, n = 57, Nevo 1978); for the only other agnathan studied, the brook lamprey, $\bar{H}$ = 0.076 (Ward et al. 1981). This level of genetic variability in Petromyzon marinus makes it possible to detect spatial and temporal population trends and to delineate genetic affinities.

At individual sites, there are few departures from equilibrium expectations for genotype frequencies; since an equivalent number of significant deviations would occur by chance alone, there is little reason to suspect that the population units represented by these samples violate Hardy-Weinberg assumptions.

Within-sample variation accounts for 90 percent of the observed variation in Petromyzon marinus (Figure 1). One implication of this result is that a sample of sea lamprey from a single site is a reasonably valid representative of genetic variability in the species as a whole. This proportion is high in all species that have been studied. Within-sample variation approaches 100 percent in highly vagile organisms such as migrant monarch butterflies over their summer range (99.6 percent; Eanes and Koehn 1978), but even in such notably philopatric or sedentary species as red-bellied newts in Southern

California (Hedgecock 1978) and Helix snails among cities (Selander and Kaufman 1975) it is substantial (97.4 percent and 83.8 percent respectively).

For any species it is the remainder of the variation that is interesting from the standpoint of population structuring and tracing stock affinities; and it is this residual variation that we sought to partition by means of hierarchical F-statistics. In the case of sea lamprey, the largest source of residual variation is differentiation from lake to lake and river to river (Table 3, Figure 2).

With a design similar to this one, Avise and Felley (1979) examined differentiation among reservoir populations of bluegills (Lepomis macrochirus) at three hierarchical levels (Table 5). They found differentiation in this species to be greatest among reservoirs and less pronounced at the next higher (between river systems) or next lower (among-site but within-reservoir) levels. The large number of physical and biological variables between these studies, (e.g. size and age of water bodies, relative vagility, breeding biology and genetic constitution of each species; loci chosen for study, etc.) make generalization from two examples somewhat tenuous; but several patterns emerge, among them that genetic mixing of fish populations is limited by physical barriers but apparently relatively insensitive to distance alone.

Table 5 lists F-statistics reported for other species of fish, along with the geographic area represented by the samples. Winans (1980) has also calculated F statistics for several other marine and coastal fish which range from 0.004 for plaice in the North Sea area to 0.3542 for Menidia in eastern U.S. All these F's estimate differentiation among samples relative to the total collection made, but without an hierarchical structure it is impossible to discern at what level divergence has occurred. The high F value for Salmo

clarki, for instance (Table 5) may reflect thirty genetically isolated population pockets, but it could equally well result from extreme divergence between two internally-panmictic drainages. Single F statistics ("F_{ST}," or gene diversity statistics "G_{ST}") without reference to the area represented or to the population sampling scheme employed say little about population organization as a species characteristic. Thus, it is not too surprising that attempts to relate species-to-species F_{ST} 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 very enlightening, because the scaling effects for the species involved are unknown. On the other hand, hierarchical analysis within a species is a promising tool for discerning the level at which populations differentiate (e.g. Chesser 1983).

The comparative lake-to-lake analysis for Petromyzon (Table 4) points out another fallacy of using a single F-statistic to characterize a species. Sea lamprey samples are not equally differentiated in all lakes; and within-lake homogeneous and heterogeneous groupings show they are not even equally differentiated within a single lake. All the F values in Columns 2 and 3 of Table 4 represent differentiation among sets of sea lamprey samples, and there is a tenfold difference from least to greatest. However, any one might have been published as "F_{ST} for Petromyzon marinus" if the lamprey had been sampled over a more limited geographic area. These F values are most productively used as indicators of relative genetic mixing over various parts of the sea lamprey's range. It is interesting to note that Avise and Felley's (1979) hierarchical study of bluegills also showed some reservoirs to be much more internally differentiated than others. Thus the major generalization from these F-statistics seems to be that population structure over a species' range 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 stocks.

In the absence of selective forces, low F values, as seen in sea-run populations, Lake Champlain, and the "Three-Lake" area, can result from high local effective population size, high numbers of migrants, or both (Allendorf and Phelps 1981). In addition, selection pressures acting on genetic loci in parallel at various sites could retard genetic divergence, whereas differing selection pressures could speed divergence. Although kinetic variants in fish are known for some of the polymorphic enzymes we analyzed (see review in Powers and Place 1978), such selective effects would be difficult to detect within lakes. Selection cannot be ruled out entirely as being partially responsible for some of the patterns of differentiation seen here, but we believe that the most logical and parsimonious explanation for the variation in F values within lakes represents variation in sea lamprey population structure from region to region, superimposed on an overall pattern of divergence from lake to lake among fresh waters of the eastern U.S. These differences in deviation from panmixis probably result from both founder effects and the physical attributes of each water body.

Genetic distance as shown in the dendrogram (Figure 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 effect by pooling samples of $D < .06$ (Rogers' coefficient) before clustering Lake Superior samples; here we instead have pooled groups of statistically homogeneous neighboring samples to reflect more

closely the regional differentiation revealed in the F-statistics analysis. This difference, as well as the inclusion of additional 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 lamprey, an anadromous-Lake Champlain cluster and another encompassing Lake Ontario and the New York Finger Lakes. Inclusion of samples from the other Great Lakes makes clear a distinct split between lamprey populations above and below Niagara Falls and shows a rather complex pattern of inter-lake affinities.

The picture of lamprey genetic structure that emerges from the combined F-statistics and genetic distance analysis can be summarized as follows: The genetic pattern of Petromyzon marinus in the upper Great Lakes is what one would expect to see for a species with considerable genetic interchange in the region of the St. Mary's River-Mackinac Straits while occupying isolated sites farther away from this center. The pattern in New York waters is consistent with relatively frequent or recent internal exchange of genetic material; and the genetic pattern of sea-run lamprey in the western Atlantic must indicate virtual panmixis.

One issue that originally interested us was the question of endemicity in sea lamprey in inland New York waters. Although recent introduction of lamprey 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, Webster (1979) has suggested that lamprey may have occupied New York waters since the last glacial retreat. If so, they might best be regarded as part of a system coevolved with native salmonids which calls for different management perspectives than those applied in the upper Great Lakes.

We have examined the endemicity issue by using Nei's (1975) estimate of divergence times based on measures of genetic distance. This analysis assumes isolated equilibrium population units subjected to a fixed gene substitution rate. Ferguson and Mason (1981) and Ryman and Stahl (1981) have used this formula to estimate divergence time among popluations of brown trout and Atlantic salmon. The resulting estimates for some population units of interest appear as "Nei time" in Table 6. Since the Great Lakes were occupied by glaciers until approximately 10-13,000 years ago (Bailey and Smith 1981), in situ divergence estimates of ca. 20,000 yr are clearly erroneous. A relatively large number of loci are involved here, and there is no reason to doubt approximate mutation rate; but violation of assumptions of isolation and equilibrium (no selection, no genetic bottlenecks, no migration) apparently renders Nei's divergence time model invalid for this system.

However, given the knowledge that historical introduction time of sea lamprey into Lakes Michigan, Huron and Superior can be established rather accurately at 50 years ago (average sightings, data from Smith and Tibbles 1981), the observed genetic distance among these populations can serve as a standard for comparison to New York populations. Column 3 of Table 6 shows that on average, populations of sea lamprey of the Cayuga-Seneca-Oneida-Lake Ontario complex are no more genetically distant from each other than are upper Great Lakes populations which began to diverge half a century ago. On the basis of differences between Lake Superior and the New York lakes we earlier (Brussard et al.1981) favored the endemicity hypothesis; but the present analysis based on genetic structure throughout eastern North America forces the conclusion that either introduction of sea lamprey into the eastern lakes is quite recent, or that 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 does apparently 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 very unlikely. On the other hand, the recent connections among these water bodies (e.g. the Erie Barge Canal, etc.) may have allowed gene flow among previously isolated, and more genetically distinct populations. Thus, the endemicity issue still remains unresolved.

ACKNOWLEDGMENTS

We are grateful to field personnel of the Great Lakes Fishery Commission and the New York State Department of Environmental Conservation for providing equipment and helping us obtain specimens. Vicky Meretsky, John Wilson and Mark Angevine assisted in field and laboratory. We thank Robert Lacy for discussion and advice regarding F-statistics, Larry Jacobson for access to unpublished data, and Dwight Webster for his longstanding support of this project.

This study was generously supported by a grant from the Great Lakes Fishery Commission.

REFERENCES

Allendorf, F. W. and S.R. Phelps. 1981. Use of allelic frequencies to describe population structure. Can. J. Fish. Aquatic Sci. 38: 1507-1514.

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.

Bailey, R. M. and G. R. Smith. 1981. Origin and geography of the fish fauna of the Laurentian Great Lakes Basin. Can. J. Fish.Aquatic Sci. 38: 1539-1561.

Brussard, P. F., M. C. Hall, and J. Wright. 1981. Structure and affinities of freshwater sea lamprey (Petromyzon marinus)populations. Can. J. Fish. Aquatic Sci. 38: 1708-1714.

Chesser, K. 1983. Genetic variability within and among populations of the black-tailed prairie dog. Evolution 37 (2): 320-331.

Eanes, W. F. and R. K. Koehn. 1978. An analysis of genetic structure in the monarch butterfly, Danaus plexippus. 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.

Krueger, C. C. 1980. Detection of variability at isozyme loci in sea lamprey, Petromyzon marinus. Can. J. Fish. Aquatic 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. Aquatic Sci. 38: 1832-1837.

Loudenslager, E. J. and G. A. E. Gall. 1980. Geographic patterns of protein variation and subspeciation in cutthroat trout, Salmo clarki. Syst. Zool. 29: 27-42.

Nei, M. 1975. Molecular population genetics and evolution. New York, Elsevier Pub. Co. 288 p.

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

Nevo, E. 1978. Genetic variation in natural populations: patterns and theory. Theoretical Population Bio. 13: 121-177.

Powers, D.A. and A.R. Place. 1981. 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. Stahl. 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. Aquatic Sci. 37: 1780-1801.

Stahl, G. 1981. Genetic differentiation among natural populations of Atlantic salmon (Salmo salar) in Northern Sweden. Ecol. Bull. (Stockholm) 34: 95-105.

Ward R. D., B. J. McAndrew and G. P. Wallis. 1981. Enzyme variation in the brook lamprey, Lampetra planeri (Bloch), a member of the vertebrate group Agnatha. Genetica 55: 67-73.

Winans, G. A. 1980. Geographic variation in the milkfish Chanos chanos. I. Biochemical evidence. Evolution 34 (3): 558-574.

Wright, S. 1978. Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. Chicago, U. Chicago Press.

		AGP				N	F_{IS}	PHI			
		100	146	158	190			100	106	122	92
Lake Huron	19	0.47	0.53	0	0	32	0.16	0.82	0	0.13	.06
	20	0.53	0.47	0	0	31	-0.13	1.00	0	0	0
	21	0.45	0.55	0	0	40	0.19	0.99	0	0.13	0
	22	0.37	0.63	0	0	23	0.35	1.00	0	0	0
	23	0.66	0.34	0	0	38	-0.17	0.88	0	0.12	0
Lake Michigan	24	0.59	0.41	0	0	34	-0.09	0.99	0	0.01	0
	25	0.52	0.48	0	0	30	-0.06	0.95	0	0.03	0.03
	26	0.57	0.44	0	0	23	-0.21	0.96	0	0.03	0.01
	27	0.46	0.54	0	0	25	-0.25	0.95	0	0.03	0.03
	28	0.68	0.33	0	0	20	0	0.90	0	0	0.10
	29	0.53	0.47	0	0	32	-0.17	0.86	0	0.01	0.13
Lake Ontario	32	0.50	0.50	0	0	4	-0.18	0.88	0	0.13	0
	33	0.67	0.33	0	0	3	-0.11	0.92	0	0.08	0
	34	0.54	0.46	0	0	39	-0.09	0.89	0	0.12	0
	35	0.46	0.54	0	0	49	0	0.85	0	0.14	0
	36	0.65	0.65	0	0	52	-0.07	0.78	0	0.22	.01
	37	0.51	0.49	0	0	77	0.24	0.88	0	0.10	.02
	38	0.47	0.53	0	0	47	-0.01	0.80	0	0.16	.04
Oneida Lake	39	0.44	0.56	0	0	151	0	0.77	0	0.22	.01
Cayuga Lake	40	0.47	0.53	0	0	125	0.13	0.84	0	0.15	.01
Seneca Lake	41	0.33	0.67	0	0	29	-0.03	0.82	0	0.19	0
	42	0.33	0.33	0	0	6	-0.26	0.83	0	0.17	0
Lake Champlain	43	0.79	0.21	0	0	62	-0.15	0.92	0	0.08	0
	44	0.68	0.32	0	0	44	-0.06	0.91	0	0.09	0
	45	0.84	0.16	0	0	35	-0.03	0.93	0	0.07	0
	46	0.73	0.27	0	0	153	-.08	0.90	0	0.10	0
Delaware River	47	0.62	0.37	.01	0	71	0.09	0.96	0.01	0.01	0.02
	48	0.65	0.34	0	.01	76	-0.50	0.94	0	0.05	0.01
	49	0.72	0.27	.01	.01	85	0	0.92	0	0.04	0.04
	50	0.61	0.39	0	0	99	0	0.92	0	0.02	0.06
Hudson River	51	0.69	0.31	0	0	32	-0.50	0.96	0	0	0.04
British Isles	53	1.00	0	0	0	7	-0.19	1.00	0	0	0
TOTAL for 53 Samples						2495					

TABLE I. Allele frequencies, sample sizes and deviation from Hardy-Weinberg expectation (F_{IS}) at the polymorphic loci for populations of *P. marinus*. Sample numbers refer to localities in Figure 1. For data from samples 1 - 18 (Lake Superior) see Krueger and Spangler (1981); for samples 30 and 31 (Lake Erie: Crooked Creek, Catarraugus Creek), 35 (Lake Ontario: Little Sandy Creek), and 52 (Connecticut River) see Brussard *et al.* (1981).

Single asterisks indicate $p < .05$ (log-likelihood test), double asterisks $p < .01$.

TABLE I - continued

		N	F_{IS}	PGM-2			N	F_{IS}	MDH		N	F_{IS}
				100	148	69			-100	-165		
L. Huron	19	36	0.01	0.63	0.38	0	36	-0.01	0.82	0.18	36	-0.22
	20	32	0	0.64	0.36	0	32	0.39*	0.81	0.19	32	-0.03
	21	40	-0.01	0.73	0.28	0	40	-0.13	0.91	0.09	40	-0.10
	22	39	0	0.42	0.58	0	37	0.17	0.73	0.27	39	-0.11
	23	38	-0.13	0.58	0.42	0	37	-0.28	0.66	0.34	40	-0.29
L. Michigan	24	36	-0.01	0.55	0.45	0	38	-0.17	0.80	0.20	40	-0.25
	25	38	-0.04	0.66	0.34	0	38	0.06	0.79	0.21	40	-0.12
	26	36	-0.03	0.68	0.32	0	36	0.04	0.85	0.15	40	-0.17
	27	38	-0.04	0.50	0.50	0	39	-0.08	0.80	0.20	40	-0.25
	28	26	-0.11	0.67	0.33	0	24	0.63	0.90	0.10	26	-0.11
	29	35	-0.15	0.66	0.34	0	40	-0.17	0.89	0.11	40	-0.13
L. Ontario	32	48	-.14	0.90	0.10	0	48	-0.12	0.89	0.12	48	-0.13
	33	66	-0.09	0.84	0.16	0	63	-0.19	0.86	0.14	63	0.09
	34	39	-0.13	0.91	0.09	0	39	-0.10	0.90	0.10	39	0.16
	35	111	0.05	0.93	0.07	0	113	0.07	0.86	0.14	100	-0.08
	36	86	-0.17	0.92	0.08	0	86	-0.08	0.94	0.06	86	0.15
	37	77	0.01	0.91	0.09	0	80	0.04	0.95	0.05	80	0.47**
	38	51	-0.08	0.88	0.12	0	51	-0.13	0.96	0.04	51	-0.04
Oneida L.	39	195	0.13	0.96	0.04	0	196	-0.04	0.97	0.03	190	0.14
Cayuga L.	40	264	0.01	0.96	0.04	0	273	-0.04	0.97	0.03	275	-0.03
Seneca L.	41	62	0.09	1.00	0	0	61	0	1.00	0	61	0
	42	6	-0.20	1.00	0	0	6	0	1.00	0	6	0
L. Champlain	43	66	-0.09	0.99	0.02	0	66	-0.02	1.00	0	66	0
	44	44	-0.10	0.99	0.01	0	44	-0.01	1.00	0	44	0
	45	30	-0.07	0.97	0.03	0	30	-0.03	1.00	0	30	0
	46	161	0.03	0.98	0.02	0	161	-0.02	0.98	0.02	161	-0.16
Delaware R.	47	93	-0.03	0.97	0.22	0.01	93	-0.03	0.81	0.19	91	0.20
	48	76	-0.06	0.99	0.01	0	76	-0.01	0.80	0.20	76	0.09
	49	86	0.10	0.97	0.03	0	86	-0.03	0.80	0.20	86	-0.18
	50	134	0.02	0.96	0.03	0.02	135	0.14	0.81	0.19	134	-0.09
Hudson River	51	56	-0.04	0.97	0.01	0.02	56	-0.02	0.77	0.23	56	0.30*
British I.	53	7	0	1.00	0	0	7	0	0.29	0.71	7	0.30
TOTAL:		3302					3287				3318	

A. ANALYSIS OF VARIANCE

Source of Variation	df	AGP SS	AGP MS	AGP F	PHI SS	PHI MS	PHI F	PGM-2 SS	PGM-2 MS	PGM-2 F	MDH SS	MDH MS	MDH F
Among sites	7	0.658	0.094	11.275**	1.200	0.171	3.814*	15.748	2.250	47.085***	23.089	3.299	51.338***
Between years	1	0.002	0.002	0.248	0.091	0.091	2.032	0.009	0.009	0.181	0.045	0.045	0.705
Error	7	0.058	0.008		0.341	0.045		0.334	0.048		0.450	0.064	
Total:	15												

For successive-year samples at a site

Average genetic distance = 0.004
Average F_{YS} = 0.007

B. ANALYSIS OF VARIANCE

Source of Variation	df	AGP SS	AGP MS	AGP F	PHI SS	PHI MS	PHI F	PGM-2 SS	PGM-2 MS	PGM-2 F	MDH SS	MDH MS	MDH F
	2				0.053	0.026	3.639	0.871	0.436	48.426*	0.486	0.243	8.60
	1	Insufficient data			0.125	0.125	17.356	0.175	0.175	19.397*	0.025	0.024	0.869
	2				0.011	0.007		0.018	0.009		0.057	0.028	
	5												

For adult and ammocoete samples at a site

Average genetic distance = 0.003
Average F_{AS} = 0.007

TABLE 2. Temporal variation at single collecting sites of sea lamprey.

A. Variation between ammocoete samples collected in successive years (1979-1980) at Sites 35, 39, 40, 41, 46, 47, 50 and 51 (Figure 1).

B. Variation between adults and ammocoetes collected simultaneously at Sites 40 (1979, 1980) and 38 (Figure 1).

F STATISTICS ANALYSIS

PETROMYZON MARINUS

LOCUS	N	H_T	F_{WT}	F_{DW}	F_{SD}	$(1-F_{ST})$
AGP	2493	0.4831	0.0044**	0.0314**	0.0301**	0.9391
PHI	3301	0.1622	0.0036**	0.0482**	0.0189**	0.9351
PGM-2	3287	0.3000	0.0318**	0.1244**	0.0228**	0.8050
MDH	3318	0.2185	0.0108**	0.0411**	0.0222**	0.9282
MEAN		0.2910	0.0127	0.0613	0.0235	0.9018

TABLE 3. Components of genetic variability in Petromyzon marinus. H_T, total heterozygosity, measures overall variation at a locus. Succeeding columns partition the proportion of H_T that can be attributed to differentiation between landlocked and anadromous waters (F_{WT}), among drainages within the two systems (F_{DW}), and among samples within a drainage (F_{SD}). Residual proportion ($1 - F_{ST}$) represents within-sample variation.

** at each level indicates significant ($p < .01$) differentiation at that level.

POPULATION UNIT	NO. OF SAMPLES	F_{DW}	F_{SD}	F_{SR}
Sea-run (4 rivers)		0.0092		
Delaware R.	4		0.0056	
Landlocked (9 lakes)		0.0713		
L. Huron	5		0.0511	
L. Superior	18		0.0403	
Eastern	12			0.0150
Western	6			0.0594
L. Michigan	6		0.0203	
Northern	2			0.0083
Southern	4			0.0106
L. Ontario	7		0.0162	
Northwestern	3			0.0048
Eastern	4			0.0158
L. Erie	2		0.0038	
L. Champlain	4		0.0057	
Seneca L	2		Ca. 0	
"Three-lake" Region	15			0.0144

TABLE 4. Geographic variability in differentiation among population units of Petromyzon marinus.

Column 3: weighted mean differentiation among drainages (F_{DW} of Table F) subdivided into sea-run and landlocked components.

Column 4: contributions to among-site variation (F_{SD}) for individual lakes and rivers.

Column 5: "F_{SR}" values for regional variation in among-site differentiation within some lakes; see text for details.

TABLE 5. F-STATISTICS FOR SOME FISH SPECIES

SPECIES	# SAMPLES	"F_{ST}"	RANGE	REFERENCE
Salmo clarkii	30 (5 subsp)	0.703	Western U.S.	Loudenslager & Gall 1980
Salmo salar	6	0.092	Northern Sweden	Stahl 1981
Chanos chanos	14	0.041	Western Pacific	Winans 1980
Lepomis macrochiris	64	0.012	Sites within reservoirs	Avise & Felley 1979
	8	0.305	Reservoirs within rivers	
	2	.041	Drainages in S. Carolina	
Petromyzon marinus	53	0.029	Sites within drainages	This study
	13	0.061	Drainages within fresh & anadromous systems	
	2	0.013	System in northeastern N. America	

POPULATION UNIT	NEI TIME (Yr.)	HISTORICAL TIME (Yr.)
Lake Superior	20,552	47
Lake Michigan	12,670	29
Lake Huron	31,282	72
Sup-Hur-Mich (all samples)	21,419	49
Lake Erie	5,099	11
Lake Ontario	8,884	20
Lake Champlain	3,007	7
Seneca Lake	119	0
Cayuga-Oneida-Seneca	4,449	10
Anadromous W. Atlantic	2,107	4
All samples east of L. Huron (N. America)	24,605	57
All Petromyzon	49,369	114

TABLE 6. Estimated divergence time among Petromyzon marinus population units.

"Nei Time" is divergence time estimated by Nei's (1975) model; "Historical Time" is standardized to known introduction times in upper Great Lakes (see text).

Figure 1. Collecting sites for 53 Petromyzon marinus samples.
Details of locality data are listed in Appendix.

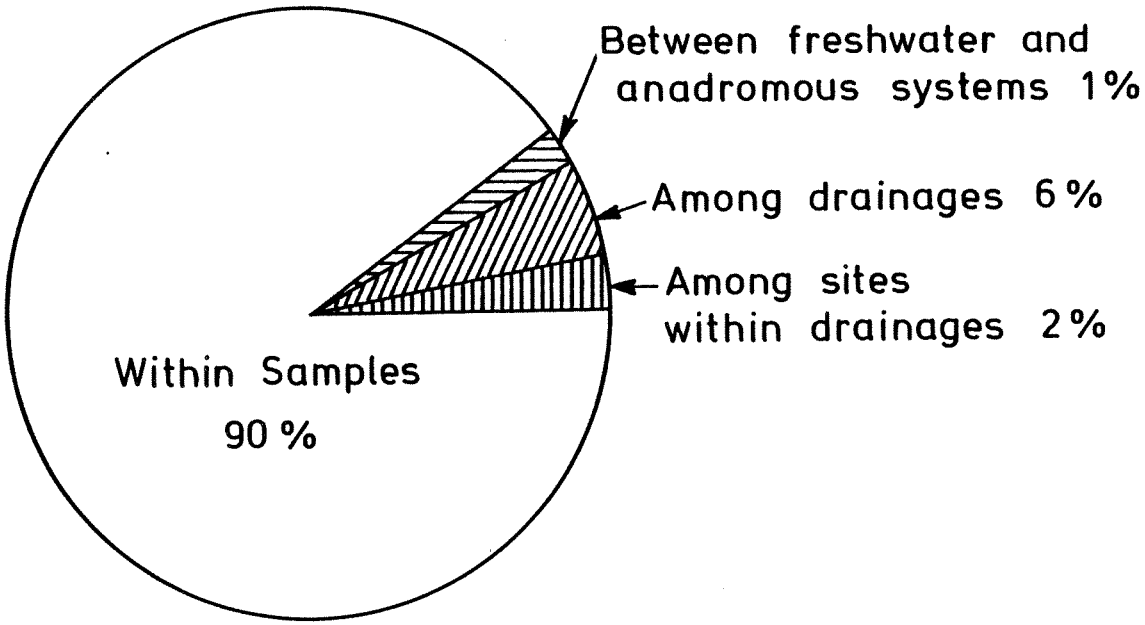

Figure 2. Partitioning of genetic variation in Petromyzon marinus based on F-statistics across 53 samples.

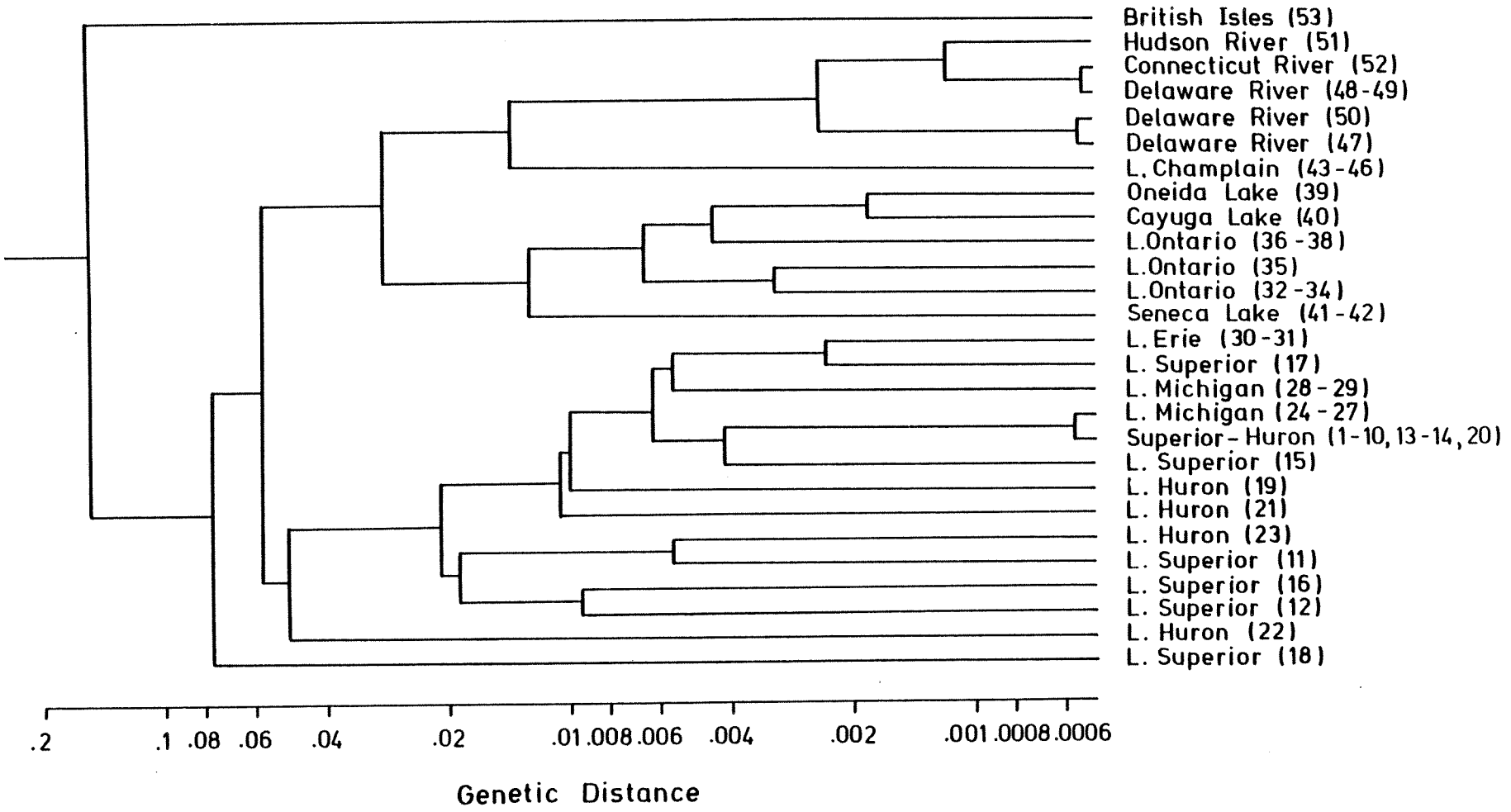

Figure 3. UPGMA dendrogram of relationships among sea lamprey populations based on Nei's (1975) genetic distance at four polymorphic loci. Numbers in parentheses following locality are site designations (see Figure 1).