Characterization of winter flounder (*Pseudopleuronectes americanus*) larval genetic stock structure within eastern Long Island Sound: estimation of larval entrainment and recruitment

A report made to the Millstone Environmental Laboratory, Millstone Power Station, Waterford, CT

By: Joseph F. Crivello

Introduction

This is a report of the activities of a three-year project designed to determine the genetic stock structure of winter flounder larvae in eastern Long Island Sound and to use genetic differences to determine the likely source populations entrained by the Millstone Power Station. The goals of this project were to: provide proof-of-principle for the scientific approach, i.e., that genetic differences among larvae groups were large enough to allow them to be used to identify entrained larvae source populations (1st year). The goal of the 2nd year was to use the information gained in the 1st year to attempt to estimate proportion of larvae from identifiable spawning stocks entrained by the Millstone Power Station. The goal of the 3rd year was to further assess entrainment sources and to provide an indication of the temporal fluctuations. In addition, in the 2nd and 3rd years of the project, the goal was to use developed information to examine recruitment of larvae to young-of-the-year juveniles in the Niantic River.

This project was funded by Northeast Utilities (2000) and Dominion Nuclear Connecticut (2001, 2002) in a contract to Dr. J. Crivello, University of Connecticut, Department of Physiology & Neurobiology.

Materials and Methods

Larval collection

Winter flounder larvae were collected from three sites within eastern Long Island Sound (Figure 1), the Thames River, Niantic River, and just to the west of the mouth of the Connecticut River (at Plum Bank in 2000, or off of Westbrook in 2001 and 2002). These areas were known to be, or adjacent to winter flounder nursery locations in the Connecticut waters of eastern LIS relatively near the Millstone Power Station. Mostly Stage 1 (yolk-sac) and 2 (pre-flexion) and some Stage 3 (flexion) winter flounder larvae were collected with a bongo net sampler with a 202-μm mesh net at variable depths from late February through April of each year. Larvae were placed in 70% ethanol and sorted in the laboratory.

Larvae were also collected using a 333-μm mesh net from seawater entrained at Millstone Power Station from late March through mid-June of each year. Larval samples collected during this time period using a stratified random design (in 2001 and 2002) so as to provide an accurate cross-section of larvae entrained at the Station (i.e., most samples were collected during those weeks when the greatest number of larvae were expected to be entrained at the Station, based on a long-term average distribution).

The number of larvae, by developmental stage (described in DNC 2003), collected from field locations, and by entrainment at the Millstone Power Station, are given in Table 1. There were 134, 164, and 171 staged larvae collected from the Niantic River in 2000, 2001, and 2002, respectively. There were 144, 174, and 151 larvae collected from the Thames River in 2000, 2001, and 2002, respectively. There were 144, 197, and 239 larvae collected from the Connecticut River in 2000, 2001, and 2002, respectively.

There were 359, 1050, and 969 entrained larvae collected in 2000, 2001, and 2002, respectively.

Young-of-the-year collection

Following metamorphosis and settlement, juvenile winter flounder (10-71 mm) were collected with a 1-m beam trawl at two locations in the Niantic River: LR, approximately 500 m west of the river mouth and WA, about 1 km north of the mouth near the eastern shoreline (Figure 2). Juveniles were placed in 70% ethanol and then transferred to the lab for analysis. A small (5 x 5 mm) piece of muscle tissue was

used to isolate genomic DNA from each juvenile. There were 157, 211, and 180 juveniles collected in 2000, 2001, and 2002, respectively (Table 2).

Genetic analysis

Genomic DNA was extracted by the method of Kaplan et al. (2001) from each individual larvae or juvenile muscle sample. Genomic DNA was quantified with Pico Green™ (Molecular Probes, Inc., Eugene, OR) and comparison to a DNA standard curve (0 to 100 ng). Stage 1 larvae provided 75-100 ng of genomic DNA, sufficient for analysis of 6 microsatellite loci. Stage 2 through 4 larvae gave relatively large amounts of genomic DNA.

To analyze each microsatellite locus, 10 ng of genomic DNA was added to a solution containing 10 mM Tris, 50 mM KCl, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 μM forward and reverse primers to a final 10 μl volume. The forward primer was covalently modified with a D2, D3, or D4 fluorescent tag (Research Genetics Inc., Huntsville, AL). The sequence of primers is included in Table 3, as well as the PCR conditions for each primer set. Primers A441, I29, D34 and J42 were from McGowan and Reith (1999).

After PCR, the samples were precipitated by addition of 2 μL 3 M Sodium Acetate pH 5, 2 μL of a 1 mg/ml glycogen solution and 50 μL of absolute ethanol. The samples were frozen at –70°C for 10 minutes and then spun at 30,000*xg* for 15 minutes. The supernatant was discarded and each sample was washed with 75 μl of 70% ethanol. The samples were dried and re-suspended in 30 μL of formamide that contained a 60-400 base pair DNA standard labeled with a D1-fluorescent tag (Beckman Instruments, Palo Alto, CA). The samples were then analyzed on the Beckman Seq-2000™ Capillary Electrophoresis System. Microsatellite products were identified by size with an accuracy of 0.25 base pair by comparison to standards. Table 3 contains the size ranges and number of alleles for each locus.

Statistical analysis

Several different approaches were used for the statistical analyses of pair-wise comparisons between populations; i.e., F_{ST}, Rho (an intra-class relatedness coefficient that permits comparison among different ploidy levels), R_{ST} (an F_{ST} analogue based on allele size), D_s (Nei's 1978 standard genetic distance) and $\delta\mu^2$ (a D_s analogue based on allele size)(Slatkin 1995; Goldstein and Pollock 1997; Ronfort

et al. 1998). All of these statistics were computed for each locus and a multi-locus weighted average was determined. A jackknife re-sampling method for each locus (i.e., deleting information from one locus at a time) was used to determine approximate standard errors for the multi-locus estimates of genetic differences. The allele frequencies for each population were also determined.

Analysis of entrainment

An estimation of the proportion of entrained larvae from identifiable spawning stocks was examined initially by discriminant analysis (Ernesto Lorda, Millstone Power Station). In the 2nd and 3rd years of this project, entrainment was examined by a multi-layer feed-forward neural network trained on supervised procedures, i.e., the construction of a model based on examples of data with known outputs (Lek and Guegan 1999; Wu-Catherine and McLarty 2000; Brosse et al. 2001). These networks construct models solely from provided examples, which are assumed to implicitly contain the information necessary to establish the correct relationship. The structure of these models is a layered feed-forward network in which non-linear elements (neurons) are arranged in successive layers and information flows unidirectionally, from input to output layer, through hidden layers. Networks with an arbitrary number of hidden neurons have been shown to be universal approximators (Cybenko 1989) for continuous maps and can be used to implement any defined function. In this type of neural network, backpropagation algorithms correct errors during the learning process. After training, the network is tested by a 'holdout procedure' (Kohav, 1995). This requires the network to be trained on a randomly selected half of a set of known samples and then applied to the remaining half to assess the accuracy of assignment to the correct population. The holdout procedure is repeated 100 times on a sample set in which the order and the assignment to the trained and assessed halves is randomized (i.e., essentially a boot-strapping procedure). An error rate is determined from the holdout procedure, as well as 95% and 99% confidence limits. The distribution of the classification values generated by the holdout procedure is used to calculate a classification value cutoff for the assignment of unknown samples to known spawning groups.

Analysis of recruitment

The above approach can also be used to estimate recruitment, that is, the estimation of the proportion of young-of-the-year juveniles derived from an identifiable spawning group. In the 2nd and 3rd years of this project, the multi-layer, feed-forward trained neural network was used to examine the recruitment of larvae to young-of-the-year juveniles.

Results

Microsatellite loci

The characteristics of the six assayed microsatellite markers are given in Tables 3 and 4. All of the microsatellite loci were polymorphic and heterozygotic with no evidence of inbreeding (Table 4). There are 135 possible alleles among the six microsatellite loci. An example of the microsatellite alleles is given in Figure 3.

The determination of population genetic structure is made by pair-wise statistical examination of the differences in allele frequencies (i.e., microsatellite products) in each population. The statistic F_{ST} reflects the proportion of the observed genetic variation that can be explained by partitioning between populations (between 0.0 and 1.0)(Wright 1969). A value of 1 indicates that the populations share no alleles in common, and a value of 0 indicates that the populations are identical. A F_{ST} value greater than 0.05 suggests that significant genetic differences exist between two populations, and some factor is acting to limit gene flow.

The statistical analyses of pair-wise comparisons between populations for all three years are given in Table 5. The genetic differences between populations were calculated using different approaches. The greatest variance was seen with D_S (Nei's genetic difference) but R_{ST} and $\delta\mu^2$, that have been developed for the use of microsatellite information, gave very similar results to the standard F_{ST}. The genetic variance is more significant than what might be predicted on the basis of geographical distance.

There was an apparent temporal stability to the genetic difference between collection sites. There was also a geographic influence with the most distant populations (Thames and Connecticut Rivers, with the average = 0.0507) having the greatest genetic difference, and the closest populations (Thames and Niantic Rivers, with the average = 0.0311) having the least difference. The intermediate populations (Niantic and Connecticut Rivers, with the average = 0.0431) had an intermediate level of genetic difference.

Larval Classification

There was sufficient genetic difference between the possible source populations (i.e., genetic differences greater than 0.05) to allow for the estimation of entrained larvae from identifiable spawning stocks by a discriminant program and a multi-layer feed-forward trained neural network. The accuracy of the discriminant analysis is given in Table 6, and the accuracy of the multi-layer feed-forward trained neural network is given in Table 7.

Discriminant analysis is designed expressly to determine what best distinguishes or tells apart groups. The discriminant analysis had 85%, 89% and 92% accuracy in assigning larvae to the known population. The discriminant analysis was used to assign entrained larvae to the most likely source population (Table 7). The neural network approach was used in the 2nd and 3rd years of this project because of its greater accuracy in assigning known samples to known populations (Table 8). The accuracy of the neural network is determined by a holdout procedure repeated 100 times on a sample set in which the order and the assignment to the trained and assessed halves are randomized (i.e., essentially a boot-strapping procedure). An error rate is determined from the holdout procedure as well as 95% and 99% confidence limits. The holdout procedure is also used to develop a classification confidence value. The distribution of the classification values generated by the holdout procedure is used to calculate a mean and standard deviation of the values. Greater than 80% of the classification values are between 0.8 and 1.0 and >90% of the values were greater than or equal to a classification value of 0.75. A classification value equal to or greater than 0.75 was used to classify an unknown sample to a specific spawning group. A value less than 0.75 resulted in the sample being assigned to an unknown group. This gave four possible groups: Niantic River, Thames River, Connecticut River, or an unknown

location. The network was trained on the microsatellite allele frequency differences between the test larval spawning groups and typically generated 75-85 hidden neurons. A single qualitative output neuron was used to classify the juvenile samples. The classification of the 2001 and 2002 samples by the neural network is given in Table 8.

From the data presented in Table 8, the peak fractional entrainment of larvae from the Niantic River occurs in late April through early May. The peak fractional entrainment from the Thames River occurs in April, as does the peak fractional entrainment from the Connecticut River area. The peak entrainment of larvae occurred in late April through early May in both 2001 and 2002. The peak entrainment from the Thames and Connecticut Rivers occurred in late May and early April, and early May and mid April in 2001 and 2002, respectively. We have not attempted to make the same estimation of the 2000 data, since it was determined in a different manner and there were far fewer entrained larvae.

Juvenile recruitment

Proportional stock contribution of larvae to juvenile stage (age-0) winter flounder was estimated by the use of the neural network analysis of allele frequency differences in the same manner as estimation of entrainment. The analysis of recruitment is given in Table 9. LR station, located near the mouth of the river, had initially fewer juveniles classified to the Niantic spawning stock in June 2000 than the WA site that is more upriver. However, smaller differences were seen in September of 2000, 2001, and 2002. The great majority of juveniles were classified to the Niantic River and Connecticut River areas and not to the Thames River. Also of interest was the size of juveniles assigned to source populations, i.e., larger juveniles were produced within the Niantic River as opposed to smaller fish entering the River from other areas. A one-way analysis of variance indicated no significant difference in the size of juveniles by year, month, and station, except in September 2000 at the lower river site, where juveniles originating from the Connecticut River were significantly larger (mean of 43 mm) than those from the Niantic (36 mm) or Thames Rivers (35 mm).

Discussion and Conclusions

Previous work suggests that there are discrete breeding stocks of winter flounder, based on morphometric, meristic, tagging studies and other factors (Perlmutter 1947). In general, marine fish show less genetic differentiation than freshwater or anadromous fishes, since marine environments are less fragmented than freshwater environments (Carvalho and Hauser 1994; Ward et al. 1994). Marine organisms with a planktonic phase have a high potential for physically and biologically mediated dispersal. Nonetheless, evidence does suggest that larval retention (Jordan et al. 2000), cohort fidelity (Sinclair 1988), geographical structures and impediments (Ruzzante et al. 1998) and natal homing instincts (Nielsen et al. 1999) may limit gene flow. Marine species such as cod, hake, herring and squid have previously shown little genetic population differentiation by allozyme markers and were thought to be homogeneous over large geographical ranges. However, recent examination with microsatellite loci has revealed fine levels of population structure in a wide range of marine species (Bentzen et al. 1996; O'Connell et al. 1998; Lundy et al. 1999; Shaw et al. 1999).

Microsatellite loci form a class of highly polymorphic and informative regions of chromosomal DNA that have found great usage for studies of intra-specific population structures, as well as hybridization events, linkage mapping, paternity testing and pedigree analysis (Hughes and Queller 1993; Roy et al. 1994; Dowling et al. 1997). Although statistically significant genetic differences do not always have biological significance (Waples 1998), they can play an important role in fishery management issues or in instances of efforts to recover commercial fishing industries.

Winter flounder has experienced a significant decline in both commercial and recreational catches. Though efforts have been made to limit fishing to increase breeding stocks, no information about any genetically based population structures has been obtained. LIS represents the central portion of the natural geographical range of this fish. Work over the past few decades by the Connecticut Department of Environmental Protection has suggested that distinct spawning and nursery areas for winter flounder exist within LIS (Howell et al. 1999). One such nursery area is the Niantic River that is nearby to Millstone Power Station.

During normal Station operations, cooling water withdrawn by the station causes the entrainment and loss of millions of winter flounder larvae. The issue of plant impact due to winter flounder larval entrainment has been addressed by population dynamics modeling (Lorda et al. 2000). This modeling has focused on the Niantic River winter flounder stock and among the information required is an estimate of the annual reproductive output removed by entrainment. This fraction was determined using an indirect method (i.e., the mass-balance model). However, the present genetic analysis provides a more direct quantitative estimate of entrainment loss by source population and can serve to corroborate modeling approaches.

Larvae in early developmental stages were collected from areas known to be near spawning grounds, as the young Stage 1 and 2 larvae appear to have limited geographical dispersal and, hence, likely retain genetic differences. Comparison of genetic differences among these groups demonstrates a relatively high degree of difference. This genetic differentiation is geographically based, with the greatest difference seen between the Connecticut and the Thames River source areas that are separated by 33 km. Though only separated by 8 km, the Thames and Niantic River source areas also had relatively substantial genetic separation between them, suggesting that one or more factors might be limiting gene flow. Very similar genetic differences were seen in larvae collected in 2000, 2001, and 2002, suggesting a temporal stability to these genetic differences.

There is no apparent geographical or physical barrier to gene flow between these areas, so natal homing instincts or selective pressures may be responsible. These differences were sufficient to provide the resolution to assign entrained larvae or older, settled juveniles to the tested populations. Many different approaches have been used to assign individuals of unknown origin to populations based on the genetic distance between individuals and populations (e.g., neighbor-joining trees, likelihood of the multi-locus genotype and Bayesian (Cornuet et al. 1999). A key component of any of these approaches is that they have the ability to correctly assign individuals, but a common drawback is that if the origin of the individual is not represented in the reference populations, most methods will still designate a wrong population of origin (Cornuet et al. 1999). Many approaches are based on two explicit assumptions that all loci are at Hardy-Weinberg equilibrium, and at linkage equilibrium. Other constraints are the levels of differentiation between tested populations, the ability to sample all, or virtually all of the potentially

contributing stocks, the temporal stability of the microsatellite markers and a large sample size that contains representation for all donor populations (Smouse et al. 1990; Letcher and King 1999). Newer approaches have attempted to overcome these limitations (e.g., Bayesian), and maximizing the accuracy of assignment depends, in part, on the F_{ST} values, population sizes and loci number. The relatively large differences in F_{ST} values among tested reference populations in this work, coupled with large population sizes (on average 150 individuals) and loci with large numbers of alleles (135) allows for the maximal accuracy in assignment to reference populations. It is clear that we have not sampled from all of the possible donor stocks in the Connecticut waters of eastern LIS or surrounding areas, but focusing on a smaller geographical range across which all stocks can be sampled adequately minimizes this problem.

Recently, several investigators have begun to use non-supervised or artificial neural networks to assign individuals to populations based on genetic differences (Brosse et al. 1999, 2001; Wu-Catherine and McClarty 2000). These multi-layer feed-forward neural networks are trained on populations with known genetic differences and then applied to unknown individuals. These neural networks have been used in a wide range of areas including assessments of fish abundance and spatial occupancies (Brosse et al. 1999). The neural network trained on the genetic differences between winter flounder larval populations estimated that the fractional entrainment in 2001 from the Niantic River stocks was 0.22; that is in very close agreement with a mass-balance model of 0.215 (DNC 2002). In 2003, the genetic analysis indicated that the Niantic River fraction was 0.123 and the mass-balance model gave 0.142 (DNC 2003). Peak fractional entrainment from the Niantic River had been estimated by the mass-balance model to occur in spring, which is supported by the results obtained by this neural network classification.

In early spring, winter flounder larvae displaced from the Niantic River make up a greater proportion of all larvae found in Niantic Bay. As the season progresses, more larvae from other sources are transported by tidal currents into this area and locally produced larvae are less dominant. Large assignments were made to the tested population that lies 25 km to the west near the Connecticut River mouth. The Thames River, which is 8 km to the east, contributed far fewer larvae. This suggests that currents and tidal flow transport winter flounder larvae in April and May predominantly from a west-to-east direction. The assignment of larvae to the Connecticut River area reaches a peak in mid-May, when the

contribution of larvae from the Niantic River has reached a nadir. The peak entrainment at the Millstone Power Station occurs in early April, when the contributions from the Niantic and Connecticut River areas are about equal.

Approximately 77% of the larvae could be definitively assigned to one of the tested populations with at least a 0.75 confidence value. Of the remaining 23%, they may have arisen from larval source populations that were not tested. These populations are perhaps more likely west of Millstone Power Station than east, because of predominant current flows.

This approach is also able to classify recruitment between larvae and young-of-the-year juvenile flounder. We analyzed the recruitment of juvenile flounder of the 2000, 2001, and 2002 year classes that were collected in July and September of each year from two sites in the Niantic River. The lower river site had initially a smaller proportion of juveniles assigned to the Niantic River than the mid-river site. Also, there were substantial numbers of juvenile flounder from the Connecticut River area that were captured in the Niantic River. Given the apparently stable genetic differences between larval populations in these areas, the juvenile flounder that are found in the Niantic River that arise from other areas likely do not contribute genetically to the larval stocks.

For fisheries management, this approach allows for the estimation of the impact of a commercial operation on specific flounder spawning groups and also allows for the direct linkage between winter flounder spawning stocks and recruitment to juveniles. Both of these factors are critical components of any fishery management plan that attempts to increase winter flounder biomass in LIS.

References

Bentzen P, CT Taggart, DE Ruzzante, and D Cook. 1996. Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Science, 53:2706-2721.

Brosse S, JF Guegan, JN Tourenq, and S Lek. 1999. The use of neural networks to assess fish abundance and spatial occupancy in the littoral zone of a mesotrophic lake. Ecological Modeling, 120:299-311.

Brosse S, JL Giraudel, and S Lek. 2001. Utilization of non-supervised neural networks and principal component analysis to study fish assemblages. Ecological Modeling, 146:159-166.

Carvalho GR and L Hauser. 1994. Molecular genetics and the stock concept in fisheries. Reviews of Fish Biology, 4:326-350,

Cornuet JM, S Piry, G Luikart, A Estoup, and M Solignac. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics, 153:1989-2000.

Cybenko G. 1989. Approximation by superpositions of a sigmoidal function. Mathematics of control signals and systems, 2: 303-314.

DNC (Dominion Nuclear Connecticut Inc.). 2002. Winter Flounder Studies. Monitoring the marine environment of Long Island Sound at Millstone Nuclear Power Station, Waterford, CT. Pages 167-287 in the 2001 Annual Report.

DNC (Dominion Nuclear Connecticut Inc.). 2003. Winter Flounder Studies. Monitoring the marine environment of Long Island Sound at Millstone Nuclear Power Station, Waterford, CT. Pages 9-127 in the 2002 Annual Report.

Dowling TE, RE Broughton, and BD DeMarais. 1997. Significant role for historical effects in the evolution of reproductive isolation: Evidence from patterns of introgression between the cyprinid fishes, *Luxilus comutus* and *Luxilus chrysocephalus*. Evolution, 51:1574-1583.

Goldstein DB and DD Pollock. 1997. Launching microsatellites: a review of mutation processes and methods for phylogenetic inference. Journal of Heredity, 88:335-342.

Howell PT, DR Molnar, and RB Harris. 1999. Juvenile winter flounder distribution by habitat type. Estuaries, 22:1090-1095.

Hughes CR and DC Queller. 1993. Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. Molecular Ecology, 2:131-137.

Jordan RC, AM Gospodarek, ET Schultz, RK Cowen and K Lwiza. 2000. Spatial and temporal growth rate variation of bay anchovy (*Anchoa mitchilli*) larvae in the mid Hudson River Estuary. Estuaries, 23:683-689.

Kaplan LAE, J Leamon, and JF Crivello. 2001. The development of a rapid and sensitive, high-through-put protocol for RNA-DNA ratio analysis. Journal of Aquatic Animal Health, 13:246-256.

Kohavi R. 1995. A study of cross-validation and bootstrap for estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence. Morgan Kaufmann Publishers Inc., 1137-1143.

Lek S and JF Guegan. 1999. Artificial neural networks as a tool in ecological modeling, an introduction. Ecological Modeling, 120: 65-73.

Letcher BH and TL King. 1999. Targeted stock identification using multilocus genotype 'familyprinting'. Fisheries Research, 43:99-111.

Lorda E, DJ Danila, and JD Miller. 2000. Application of a population dynamics model to the probabilistic assessment of cooling water intake effects of Millstone Nuclear Power Station (Waterford, CT) on a nearby winter flounder spawning stock. Environmental Science and Policy, 3:S471-S482.

Lundy CJ, P Moran, C Rico, RS Milner, and GM Hewitt. 1999 Macrogeographical population differentiation in oceanic environments: a case study of European hake (*Merluccius merluccius*), a commercially important fish. Molecular Ecology, 8:1889-1898

McGowan C and ME Reith. 1999. Polymorphic microsatellite markers for Atlantic Halibut, *Hippoglossus hippoglossus*. Molecular Ecology, 8:1761-63.

Nei M. 1978. Estimation of the average Heterozygosity and genetic distance for small numbers of individuals. Genetics. 89:583-590.

Nielsen EE, MM Hansen, and V Loeschcke. 1999. Genetic variation in time and space: Microsatellite analysis of extinct and extant populations of Atlantic salmon. Evolution, 53:261-268.

O'Connell M, MC Dillon, JM Wright, P Bentzen, S Merkouris, and J Seeb. 1998. Genetic structuring among Alaskan Pacific herring populations identified using microsatellite variation. Journal of Fish Biology, 53:150-163.

Perlmutter A. 1947. The blackback flounder and its fishery in New England and New York. Bulletin of the Bingham Oceanographic Collection, 11:1-92.

Ronfort J, E Jenczewski, T Bataillion, and F Rousset. 1998. Analysis of population structure in autotetraploid species. Genetics, 150: 921-930.

Roy K, D Jablonski, and JW Valentine. 1994. Eastern Pacific molluscan provinces and latitudinal diversity gradient: No evidence for "Rapoport's rule". Proceedings of the National Academy of Sciences, USA., 91:8871-8874.

Ruzzante DE, CT Taggart, and D Cook. 1998. A nuclear DNA basis for shelf- and bank-scale population structure in Northwest Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. Molecular Ecology, 7:1663-1680.

Shaw PW, C Turan, JM Wright, M O'Connell, and GR Carvalho. 1999. Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. Heredity, 83:490-499.

Sinclair M. 1988. Marine populations. An essay on population regulation and speciation. In: Recruitment of Fish in Oceanography., Volume 1, Washington Sea Grant, Seattle, WA, 252 pp.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139:1463-1476.

Smouse PE, RS Waples, and JA Tworek. 1990. A genetic mixture analysis for use with incomplete source population data. Canadian Journal of Fisheries and Aquatic Sciences, 47:620-634.

Waples RS. 1998. Separating the Wheat From the Chaff: Patterns of Genetic Differentiation in High Gene Flow Species. Journal of Heredity, 89:438-450.

Ward RD, M Woodwark, and DOF Skibinski. 1994. A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. Journal of Fisheries Biology, 44:213-232.

Wright S. 1969. Evolution and the Genetics of Populations, Vol. 2, In: The Theory of Gene Frequencies, University of Chicago Press, Chicago, Ill., 511 pp.,.

Wu-Catherine H and JW McLarty. 2000. Neural networks and genome informatics. In: Methods in computational biology and biochemistry, (eds. Wu-Catherine, H. and J.W. McLarty), Volume 1., Elsevier, New York, 205 pp.

Table 1. The number of winter flounder larvae collected by stage at the three sites in eastern Long Island Sound (source populations) and entrained at the Millstone Station.

Year	Niantic River		Thames River		Connecticut River		Entrained	
2000	134	stage 1	10	stage 1	144	stage 1	51	stage 2
			134	stage 2			249	stage 3
							59	stage 4
	134	total	144	total	144	total	359	total
2001	85	stage 1	43	stage 1	60	stage 1	199	stage 2
	79	stage 2	131	stage 2	86	stage 2	741	stage 3
					51	stage 3	111	stage 4
	164	total	174	total	197	total	1050	total
2002	137	stage 1	125	stage 1	179	stage 1	1	stage 1
	22	stage 2	26	stage 2	50	stage 2	85	stage 2
	12	stage 3			10	stage 3	752	stage 3
							129	stage 4
	171	total	151	total	239	total	969	total

Winter flounder larvae were collected from source populations in February, March, and April of each year. Entrained samples were collected in April, May, and June of each year.

Table 2. The number of juveniles collected from the Niantic River in 2000, 2001, and 2002.

Year	Station LR		Station WA	
	Date	Number	Date	Number
2000	June 22	57	June 22	35
	September 18	30	September 18	35
	Total	87	Total	70
2001	July 3	70	July 3	78
	September 18	45	September 18	18
	Total	115	Total	96
2002	June 13	45	June 13	11
	June 24	49	June 24	38
	September 20	31	September 18	6
	Total	125	Total	55

Table 3. Primers and PCR conditions, microsatellite sizes, and number of alleles for the six tested microsatellite loci.

Loci	Primer Sequence[a]	Products (bp)	Annealing T°C	Allele Number
P157	D_2-AGTGCAACAACAGATTCCAG(+) GCAGAATGAGTGAAATGTGG(-)	93-195	50°C	40
P159	D_3-GTGTGGAGGTCAATGC(+) GGAGCATCATTCATACAC(-)	85-209	53°C	11
A441	D_2-CAACTGTGGGTATGTGCCTG(+) GTGTCAGCACTGTGCTTAAACC(-)	89-213	55°C	25
D34	D_4-GCCTGGTCTCATTGTGTTCC(+) AGGTTAAATGATTTCCTGAAGCTG(-)	89-315	55°C	27
I29	D_3-GCTTCGGTTACACCTTTGC(+) AGGACAGTGAGGATGTCCG(-)	91-223	55°C	4
J42	D_4-CACAAACTCAAGATGTTGCG(+) AAGCTCACTGGAAAATAATACCC(-)	95-185	55°C	28

[a] D_2, D_3 & D_4 refer to fluorescent tags on the forward primer.

Table 4. Summary statistics, by source area, for the six microsatellite loci for all three years combined.

Population	Microsatellite Loci					
	P157	P159	A441	D34	I29	J42
Niantic River						
P value[a]	0.0000	0.2798	0.0000	0.0000	0.2759	0.0000
Het_{obs}[b]	0.7481	0.3926	0.6000	0.7185	0.1667	0.5333
N[c]	459	459	459	459	459	459
Thames River						
P value	0.0000	0.9817	0.0000	0.0000	0.0000	0.0000
Het_{obs}	0.8129	0.1719	0.5971	0.5972	0.1020	0.3669
N	469	469	469	469	469	469
Connecticut River						
P value	0.0000	0.8865	0.0000	0.0000	0.1594	0.0000
Het_{obs}	0.7744	0.2359	0.6821	0.4974	0.1103	0.3538
N	587	587	587	587	587	587

[a] P values indicate the probability of conformity to Hardy-Weinberg expectations by the Chi-squared method. [b] Het_{obs} refers to observed Heterozygosity. [c] N is the number of assayed larvae.

Table 5. Genetic differences between larval sampling populations for the past three years (2000 through 2002).

Year	Location[a]	F_{ST}	*Rho*	R_{ST}	D_S	$\delta\mu^2$
2000	Niantic vs. Thames	0.0301	0.0300	0.231	0.0273	0.0224
2001	"	0.0268	0.0268	0.0254	0.0307	0.0189
2002	"	0.0365	0.0365	0.0371	0.0551	0.0251
2000	Niantic vs. Connecticut	0.0553	0.0553	0.0551	0.0503	0.0345
2001		0.0430	0.0430	0.0534	0.0668	0.0312
2002		0.0311	0.0310	0.0362	0.0389	0.0411
2000	Thames vs. Connecticut	0.0441	0.0441	0.0413	0.0379	0.0434
2001		0.0503	0.0503	0.0372	0.0785	0.0214
2002		0.0578	0.0578	0.0462	0.0781	0.0516

Table 6. Discriminant analysis: The accuracy of the discriminant method in assigning known samples to the correct source population (2000 samples only).

From site	Niantic River	Thames River	Connecticut River	Unknown	Total
Assigned to					
Niantic River	94	7	8	1	110
Thames River	11	85	1	0	97
Connecticut River	6	3	103	1	113
Total	111	95	112	2	320
Accuracy	84.7%	89.5%	92.0%		

Table 7. The error rates and confidence values for the accuracy of the neural network assignment of larvae to known source populations (2001 and 2002 data sets).

2001 data	Error rate ± SD	Confidence Value (95%)	Confidence Value (99%)
Niantic River	15.57±3.31	0.0313	0.0063
Thames River	18.13±4.45	0.0429	0.0086
Connecticut River	21.67±4.28	0.0405	0.0081
2002 data			
Niantic River	14.31 ± 3.04	0.0288	0.0058
Thames River	16.65 ± 4.09	0.0394	0.0079
Connecticut River	17.37 ± 3.43	0.0325	0.0065

Table 8. Fractional assignment of entrained larvae to source populations by the multi-layer feed-forward trained neural network approach.

Date	Niantic River	Thames River	Connecticut River	Unknown Site
2000*				
4.05.00	0.11	0.00	0.89	0.00
4.11.00	0.07	0.00	0.89	0.04
4.19.00	0.38	0.03	0.59	0.00
4.26.00	0.06	0.00	0.85	0.08
5.05.00	0.20	0.00	0.80	0.00
5.11.00	0.89	0.00	0.11	0.00
5.19.00	0.79	0.03	0.18	0.00
5.24.00	0.78	0.00	0.22	0.00
6.02.00	0.46	0.00	0.54	0.00
6.09.00	0.80	0.00	0.20	0.00
2001				
3.28.01	0.46	0.08	0.15	0.31
4.09.01	0.15	0.29	0.35	0.21
4.16.01	0.30	0.23	0.25	0.22
4.23.01	0.36	0.23	0.22	0.19
4.30.01	0.27	0.15	0.34	0.24
5.07.01	0.16	0.14	0.52	0.18
5.14.01	0.17	0.14	0.46	0.23
5.21.01	0.22	0.30	0.32	0.16
5.28.01	0.19	0.29	0.29	0.23
6.04.01	0.07	0.24	0.48	0.21
6.11.01	0.19	0.19	0.37	0.25
6.17.01	0.23	0.39	0.08	0.30
2002				
4.01.02	0.07	0.25	0.57	0.11
4.08.02	0.09	0.03	0.77	0.11
4.15.02	0.09	0.06	0.69	0.16
4.22.02	0.17	0.04	0.59	0.20
4.29.02	0.14	0.00	0.57	0.29
5.06.02	0.31	0.02	0.41	0.26
5.13.02	0.12	0.02	0.72	0.14
5.20.02	0.12	0.02	0.63	0.23
5.28.02	0.27	0.00	0.43	0.30
6.03.02	0.04	0.08	0.72	0.16
6.10.02	0.57	0.08	0.08	0.27

*The discriminant analysis was used to assign entrained larvae in 2000; the multi-layer feed-forward neural network was used to assign entrained larvae in 2001 and 2002.

Table 9. Classification of juvenile winter flounder to source areas using the trained neural networks

	Classification (values are individuals assigned to a source area)			
Year				
2000	June LR	September LR	June WA	September WA
Niantic River	10	7	18	9
Thames River	7	3	4	10
Connecticut River	23	11	6	13
Unknown	17	9	7	3
Total	57	30	35	35
2001				
Niantic River	14	13	12	3
Thames River	17	4	10	2
Connecticut River	19	16	35	6
Unknown	20	12	21	7
Total	70	45	78	18
2002				
Niantic River	24	9	11	1
Thames River	7	2	5	0
Connecticut River	53	18	24	4
Unknown	10	2	8	1
Total	94	31	49	6

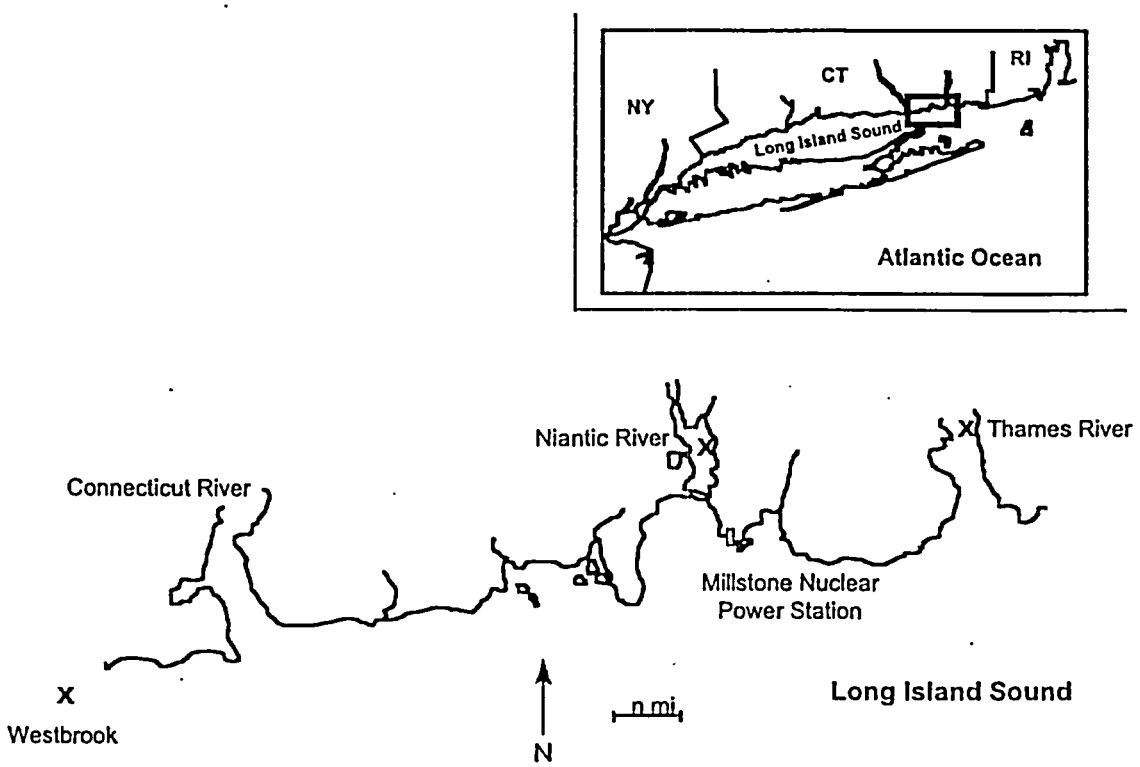

Figure 1. Approximate locations in eastern Long Island Sound (X) where larval winter flounder were collected for genetic stock identification studies during 2000. 2001, and 2002.

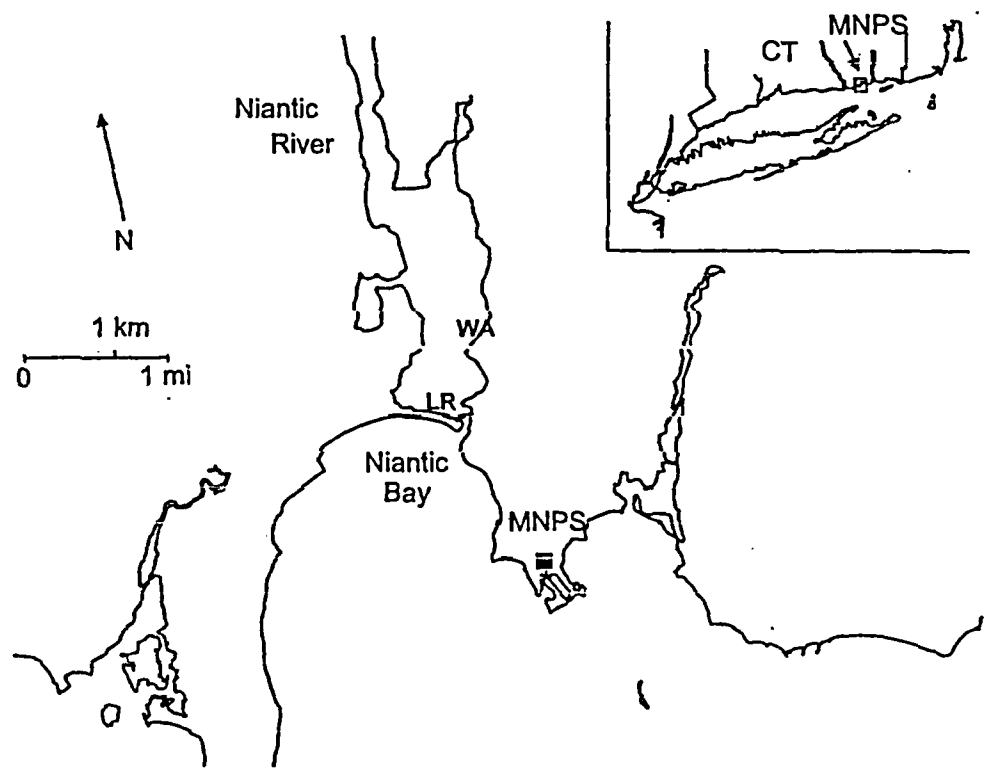

Figure 2. Approximate locations in the Niantic River (LR and WA) where age-0 juvenile winter flounder were collected for genetic stock identification studies during 2000, 2001, and 2002.

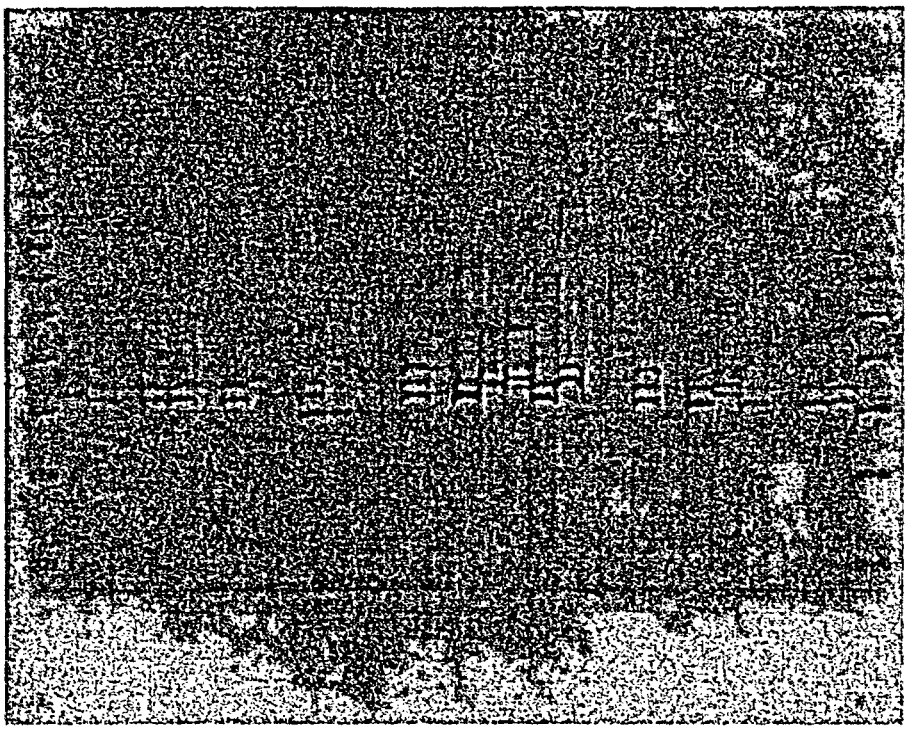

Fig 3. An example of a PAGE (polyacrylamide gel electrophoresis) analysis of microsatellite locus PCR products