Genetic assessment of ancestry of wild-caught Muskellunge (*Esox masquinongy*) in Vermont waters of Lake Champlain

Chris Wilson
Fisheries Genetics Lab, Natural Resources DNA Profiling and Forensics Centre,
Trent University, Peterborough, ON Canada K9J 7P8
chris.wilson@nrdpfc.ca;
phone (705) 755-2260

Shawn Good
Fisheries Biologist, Vermont Department of Fish & Wildlife
271 North Main St., Suite 215, Rutland, VT USA 05701-2423
shawn.good@state.vt.us
phone (802) 786-3862

February 22, 2010

Background

Decline and loss of highly valued sport fish species and populations present major challenges in fisheries management. Muskellunge (*Esox masquinongy*) is a highly prized sport fish species (Crossman 1986), but as territorial top predators with specific habitat requirements at multiple life stages, muskellunge are also particularly sensitive/vulnerable to both exploitation and habitat alteration (Scott and Crossman 1973, Crossman 1986).

Muskellunge are known to be native to Lake Champlain and it's various tributaries, including the Missisquoi River. The first records of the species in Lake Champlain, on both the Vermont and New York side, were in the mid-1800's. In 1849, muskellunge, or "muskalonge" were reported to be uncommon, but apparently widely distributed in Lake Champlain, including the Big (Great) Chazy River (Greeley, 1930). Between 1847 and 1849, Zadock Thompson captured specimens he later identified as muskellenge from the mouth of the Lamoille River (Thompson, 1853). Thompson reported the abundance of this fish in Lake Champlain as "rare". Another report from the late 1800's reported muskellunge to be "seldom found in Lake Champlain but a few are caught in the Missisquoi River" (Vermont Commissioners of Fisheries and Game, 1898). Sometime before 1970 the Missisquoi River between the Swanton dam and the Highgate Falls dam became what is now believed to be the last remaining native, naturally reproducing muskellunge population in the State.

Muskellunge surveys in this area of Lake Champlain in the 1970's resulted in only a small number of specimens being collected. In 1979, a spill of untreated paper mill waste into the upper Missisquoi River caused a massive fish kill, including the Swanton dam to Highgate Falls dam reach, and this is thought to have eradicated the river's muskellunge population. The Vermont Department of Fish and Wildlife has sampled the reach for the species several times since the mid 1980's without observing any muskellunge.

Although Lake Champlain's native muskellunge population was thought to have been eradicated in the late 1970's, small numbers of muskellunge have been caught in recent years by anglers in the lower Missisquoi River and the Missisquoi Bay area of Lake Champlain. The origin of these muskie is of great importance to Vermont's muskellunge restoration program. It is possible these fish originated from previous stocking programs in the Lake Champlain basin. From 1980 through 1986, the Fish & Wildlife Department annually stocked 10,000 to 100,000 muskie fry from

Pennsylvania into the Lemon Fair River and upper Otter Creek, tributaries to Lake Champlain's main basin. Additionally, New York State Department of Environmental Conservation has stocked Chautauqua Lake muskie fry into the Great Chazy River (New York tributary to Lake Champlain) above the Waterworks dam for nearly 30 years. As a result, the occasional muskie caught on the Vermont side of the lake and in the lower Missisquoi River below Swanton dam may be stray fish from New York or the vestiges of the Otter Creek stockings 20 years ago. However, there is an outside chance these fish are descendants of the original Lake Champlain strain fish from the upper Missisquoi River.

There is considerable interest in determining whether these recently-captured fish represent native remnants, stocked fish, or naturalized fish from previous stocking events. In addition, two wild esocids were captured which were tentatively identified as pike x pickerel hybrids. Both were captured in Missisquoi Bay, Lake Champlain in 2004 (VTDFW #MIS-093004-001 and VTDFW #MIS-093004-002).

In this study, genetic markers (species-specific microsatellite DNA loci) were employed to determine the native versus stocked origins of these fish, using individual-based analyses to resolve the ancestry of individual fish. The effectiveness of this approach has been demonstrated in lake trout (Piller et al. 2005), and walleye (Wilson et al. 2007), where microsatellite genotyping and individual assignment tests were able to both quantify contributions of stocking sources to wild populations and detect individuals with pure and partial native ancestry. Ancestry of the putative hybrids was assessed using microsatellite genotyping and amplification and sequencing of a 560-nucleotide segment of the mitochondrial cytochrome b gene, which has been shown to be diagnostic among *Esox* species (Grande et al. 2004, Kyle and Wilson 2007).

Methods

Muskellunge samples were provided from five locations (Table 1). Two muskellunge were caught by anglers below the Swanton dam in 2003 and 2005. For comparison a muskellunge angled above the Swanton dam was also genotyped, as this individual might represent a descendant of the translocated Lake Champlain population. Muskellunge from Otter Creek, captured by angling and electroshocking above and below Vergennes dam, were also genotyped (Table 1). For comparison,

muskellunge from two stocking sources were also genotyped, in order to genetically assign or exclude the wild-caught muskellunge from these potential sources.

Genomic DNA was extracted using a lysis and precipitation method as described by Wilson et al. (2007). Briefly, approximately 5mg of tissue from each sample were individually lysed using 400µl of lysis buffer (50 mM Tris pH 8, 100 mM NaCl, 1 mM EDTA, 1% SDS w/v) with 1.5 uL Proteinase K (20mg/ml). Samples were incubated overnight at 37°C, then DNA was precipitated by adding two volumes (800µl) of 80% isopropanol and centrifuging at 6,000 RCF for 30 minutes. The resulting pellets were washed with 1000µl of 70% ethanol to remove excess salts, re-centrifuged, air-dried, and re-suspended in 200µl of 1X TE (10mM Tris, 1mM EDTA). DNA yield and quality were assessed by horizontal electrophoresis on a 1% agarose gel alongside molecular mass ladders.

All muskellunge provided by VTDFW were genotyped using 15 species-specific microsatellite loci developed by Sloss et al. (2008) [*Ema*A5, *Ema*A10, *Ema*A11, *Ema*A102, *Ema*A104, *Ema*B110, *Ema*B120, *Ema*B126, *Ema*C1, *Ema*D4, *Ema*D5, *Ema*D6, *Ema*D12a, *Ema*D114, and *Ema*D116]. Microsatellites were PCR-amplified in 10µl reactions, containing 1µl of PCR buffer, 2mM $MgCl_2$, 0.4 µl of 10 mM dNTP's, 0.3 µl of each primer [10 mM], 0.25U of *Taq* DNA polymerase (Qiagen, Inc), and 10 ng of genomic DNA. Target DNA was amplified in a PTC-100 thermocycler (MJ Research) using an initial denaturation step of 95°C for 4 minutes, followed by 35 cycles of 94°C for 45 seconds, 58°C for 1 minute, 72°C for 45 seconds, with a final extension step of 60°C for 30 minutes. The resulting amplicons were diluted 1:30 in distilled water and mixed with formamide and ROX350 size standard before loading onto an AB3730 automated sequencer for electrophoresis. Scoring of the resultant genotypes was done using GeneMapper software (Applied Biosystems) and manual proofreading.

For sequence determination of the two unknown esocids, a 500 bp fragment of the mitochondrial cytochrome b gene was amplified using conditions described by Kyle and Wilson (2007). Each 20 ml PCR reaction contained 10 ng of genomic DNA, 2 ml of PCR buffer (Qiagen, Inc) with 2 mM MgCl2, 0.4 µl of 200 mM dNTP's, 0.6 µl of each primer [10 mM] and 0.5U of *Taq* DNA polymerase (Qiagen, Inc). Target DNA was amplified in a PTC-100 thermocycler (MJ Research) under the following cycling conditions: an initial denaturation step of 95°C for 2 minutes, followed by 30 cycles of 92°C for 2 minutes, 50°C for 1 minute, 72°C for 45 seconds, followed by a

final extension of 72°C for 5 minutes. Amplified products were cleaned using ExoSAP (InVitroGen) and sequenced in both directions using the amplification primers using dye-terminator cycle sequencing (Perkin-Elmer) with the manufacturer's protocol. Sequences were electrophoresed on an AB3730 automated sequencer (Applied Biosystems), and proofread and aligned using Sequencher v3.8 software (GeneCodes, Inc.).

Data analysis

Ancestry of the wild-caught muskellunge, as well as those from the two source populations, was determined using individual assignment tests (Cornuet *et al.* 1999, Pritchard *et al.* 2000) and ordination of multilocus genotypes from individual fish using principal coordinate analysis (PCoA) (Peakall and Smouse 2006). Genetic ancestry of individuals was assessed using a partial Bayesian approach (Rannala and Mountain 1997) as implemented in GeneClass 2.0 (Piry *et al.* 2004). Individuals were considered as reliably assigned to a single source population if the probability of their genotype occurring by chance in that population, based on baseline allele frequency data, was significantly greater than that for the next most probable source (*i.e.,* if the log-likelihood ratio ≥ 2). These assignments were tested by complementary exclusion analyses, using simulated population sizes of 10,000 fish for each source population (Cornuet *et al.* 1999) to confirm exclusion of individual genotypes from non-source populations at a rejection level of $p = 0.01$. These combined analyses allowed for several possible outcomes, allowing individuals to be assigned to either a single source population (sole ancestry), to two or more populations due to mixed ancestry, or to be excluded from all known source populations. Assignment interpretations were independently assessed using Structure 2.2 (Pritchard *et al.* 2000), which also allowed estimation of mixed-ancestry contributions for the wild-captured fish. Independent analyses were run with and without using population identifiers for fish sampled from the source populations for comparison; all fish sampled from Missisquoi River and Otter Creek were considered as unknowns. The program Structure was also used to estimate the number of genetic groups (K) present within the dataset, using the complementary hypotheses of panmixia (one genetic group among all samples), two genetic groups (two recognizably distinct stocking sources, with wild-caught muskellunge being solely derived from stocking events), three genetic groups (two stocking sources, plus a remnant Lake Champlain gene pool), to up to ten genetic groups (allowing for substructure within each of the above groups). Analytical runs in Structure 2.2 relied on 50,000 sampling cycles prior to data collection, followed by 100,000 resampling steps, using 10 replicate trials for each estimation of K.

To assess the potential maternal ancestry of the putative esocid hybrids, the mitochondrial sequences from these individuals were uploaded to the GenBank database (http://www.ncbi.nlm.nih.gov/blast) and subjected to a *blastn* search to identify homologous sequences to identify their maternal parents. This search engine calculates the homology and statistical significance of nucleotide sequence matches between query and references sequence pairs (Altschul et al. 1997). In order to assess whether a given sequence alignment using the *blastn* program constituted evidence for homology, bit scores and expectation values (E) were also used to evaluate how strong an alignment can be expected from chance alone. The bit score gives an indication of how good the alignment is; the higher the score, the better the alignment. This score is calculated from a formula that takes into account the alignment of similar or identical residues, as well as any gaps introduced to align the sequences (Altschul et al. 1997). The E-value gives an indication of the statistical significance of a given pairwise alignment and reflects the size of the database and the scoring system used. The lower the E-value, the more significant the hit. The E-value is a parameter that describes the number of hits one can ‘‘expect’’ to see by chance when searching a database of a particular size; the lower the E-value, or the closer it is to ‘‘0’’, the higher is the ‘‘significance’’ of the match.

Results

The small number of wild-caught muskellunge precluded population-based assessment of diversity or relationships between these fish and stocking source populations. Analysis of the two stocking sources confirmed that all microsatellite loci conformed to Hardy-Weinberg equilibrium expectations after sequential Bonferroni correction for multiple comparison tests (data not shown), and had sufficient diversity to be useful for individual assignment tests (Appendix I).

With one exception, all individual-based analyses of the wild-caught muskellunge were consistent with patterns predicted under the hypothesis of their having resulted from previous stocking from one of the two stocking sources (but see below for the 1980 Missisquoi River sample (fish 3)). No evidence of unique genetic characters that would separate them from stocking sources was detected in these wild-caught fish. Multivariate ordination of genotypes from individual fish using Principal Coordinate analysis (PCoA) implemented using GenAlEx (Peakall and Smouse

2006) showed substructure within both stocking sources, with all wild-caught muskellunge falling within the gene pool variance of the stocked fish (Figure 1).

Assignment of muskellunge from the two stocking sources back to their true source population using GeneClass2 (Piry et al. 2004) gave a quality index of 96%, with 97% of all fish correctly assigned to their true source with 95% certainty (data not shown). This measure gives some indication of the power of the assignment tests, and therefore the extent to which the assignment of the wild-caught muskellunge can be relied on. When the wild-caught muskellunge were assessed for assignment back to these two potential sources, all but one of the wild-caught fish were assigned to the potential sources with values that fell within the probability distributions for the source fish themselves (Table 2; Figure 2). One muskellunge from Otter Creek showed comparable assignment probability to both stocking sources, reflecting its having common alleles that were well-represented in both potential sources (Appendix I). Specific exclusion tests failed to exclude any of these wild-caught fish from the stocking sources, thus failing to refute stocking as a possible origin of these fish. This further indicated that no unique alleles or remnant gene pools were detected among these wild-caught fish, suggesting that these recent captures represent stocked rather than remnant muskellunge.

The one exception to the general pattern of wild muskellunge originating from stocking sources was the the 1980 angler-caught sample from the Missisquoi River (fish 3, Table 1), which showed poor assignment probabilities to both sources (log likelihood values of 10^{-22} and lower for both sources) (Table 2, Figure 2). The axes in Figure 2 represent log-transformed probability values, and show that the probability of this individual from either stocking source is more than 10^{6} times less likely than any fish from the hatchery populations (Figure 2). Probability and exclusion tests refuted the probability of this individual having originated from either stocking source ($p<<0.001$) using a *t*-test of a sample against a distribution (Sokal and Rohlf 1995). This individual was similarly excluded from all other characterized muskellunge populations and stocking sources with genetic data available, ranging from Wisconsin, the Great Lakes, and the Ottawa River and Kawartha Lakes in Ontario (Miller et al. 2009, C. Wilson unpubl. data) These results strongly suggest this individual was not a stocked fish, and was likely a member or survivor of the original Lake Champlain population.

Running Structure (Pritchard *et al*. 2000) and allowing for admixture (interbreeding) between genetic groups, using hypotheses of one to ten genetic groups (panmixia versus all sample sets as distinct genetic groups, and allowing for unrecognized substructure within groups) resulted in maximum resolution for $K = 5$ clusters or genetic groups [ln(P|D) = -6707]. Parallel analyses that did not allow for admixture resulted in optimal solutions for K = 2 [ln(P|D) = -6707] and K = 6 [ln(P|D) = -6707] (Figure 3). Using source population information to update allele frequencies within the two stocking sources, leaving the wild-caught fish as unknowns, resulted in an optimal solution for K = 2 (Figure 3). In all instances, wild-caught fish assigned to the two stocking sources with no evidence of distinct ancestry for any of these fish (Figure 3). The failure of the 1980 sample from the Missisquoi River to be recognized as distinct in this analysis may reflect the program algorithm to minimize overall variance in the dataset, rather than maximize resolution of individual genotypes at the expense of the larger dataset (Pritchard et al. 2000).

Ancestry of potential hybrids

The two suspected esocid hybrids showed limited amplification success for the microsatellite data, with only 6 of the 15 loci resulting in successful genotyping (Appendix I). For the six loci where alleles were successfully amplified, the unknown esocids showed unique alleles at three loci, clearly showing their distinction from the muskellunge samples.

Sequence analysis of the two suspected esocid hybrids confirmed both as pike x pickerel hybrids. Both had pickerel mitochondrial DNA, so resulted between successful mating between female pickerel and male pike. A complication for fully resolving their ancestry is that both unknowns had mitochondrial sequences that were identical (100% concordance) with voucher specimens of both *Esox americanus americanus* and *Esox niger* (Grande et al. 2004; GenBank accession numbers AY497431- AY497432 and AY497438- AY497441 respectively) (Appendix II). As these pickerel species themselves have interbred themselves, such that mitochondrial introgression has occurred within these species, it cannot be determined which species of pickerel interbred with male pike in Lake Champlain.

Discussion and Conclusions

The genetic analyses indicated that all but one of the wild-caught muskellunge were derived from previous stocking events from the two stocking sources, and provided no evidence of survival of the historical population in Lake Champlain. This was clearly indicated by the principal coordinate (PCoA) ordination, where all wild-caught fish were grouped within the data clouds from the stocking-source fish. By contrast, native wild fish should have been noticeably distinct; the two pickerel x pike hybrids were widely separate when included in a parallel ordination (data not shown).

Individual assignment tests using both GeneClass2 and Structure showed that the majority of the wild-caught muskellunge were not genetically distinct from the two stocking sources. Accordingly, these fish do not appear to represent remnants of the original Lake Champlain population, as the historical population would be expected to be recognizably different from allopatric populations. Increasing the number of potential clusters in Structure did not increase resolution within or separation between the stocking source populations (data not shown), and resulted in detection of family substructure within both sources. As the samples from the Great Chazy River were fry, this result was anticipated, but substructure representing family groups was evident for samples from Pymatuning Reservoir as well.

The individual assignment tests shown in Figure 2 showed that the one fish caught in 1980 most likely was a surviving member of the original Lake Champlain population. Unfortunately, no other evidence of native Lake Champlain muskellunge was detected among the other wild-caught fish.

Differences in resolution and interpretation among the three analytical approaches reflect their underlying structure. As methods of multivariate ordination, the primary function of both the PCoA and Structure analyses is to maximize resolution or pattern within datasets by minimizing overall variance (Pritchard et al. 2000, Peakall and Smouse 2006). As such, resolving ancestry of a small number of individuals may be overwhelmed by larger total variance, unless those individuals are extremely divergent as was the case for the interspecific hybrids. By contrast, the approach used by Piry et al. (2004) estimates assignment (membership probabilities) of specific individuals to predefined sources, and so was able to resolve the separate ancestry of the muskellunge caught in 1980 (Figure 2).

The genetic results for the two unknown esocids confirmed their suspected hybrid status and showed their clear distinction from the muskellunge that were the primary focus of this study. The microsatellite results confirmed that the larger esocid parent was not a muskellunge. The inability to amplify alleles at nine of fifteen species-specific loci indicated the divergence ancestry of both esocids from muskellunge; this was further supported by the presence of unique or extremely rare alleles at three of the six loci where microsatellite genotypes were obtained. Although northern pike have yet to be screened for the loci developed by Sloss et al (2008), the large size of the unknown esocids and ubiquitous distribution of northern pike makes it probable that this species was involved in their hybrid formation.

Sequence analysis of the two unknown esocids confirmed both as pickerel x pike hybrids. The presence of pickerel mitochondrial haplotypes in both fish showed that both hybrids resulted between successful mating between female pickerel and male pike. As the mitochondrial sequences for both hybrids were identical with voucher specimens of both *Esox americanus americanus* and *Esox niger*, however, it cannot be resolved which of the two pickerel species was the maternal parent for these hybrids based on these data.

Acknowledgements:

Samples and funding for this project were provided by Vermont Department of Fish and Wildlife. Shawn Good (VTDFW) provided valuable background information and discussions during and after the project. David Gillett, Kristyne Wozney, and Kate Pammett (Natural Resources DNA Profiling and Forensics Centre) provided valuable assistance with genetic data collection and analysis.

References

Altschul, S.F., T.L. Madden, A.A. Schaffer, J.Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25: 3389–3402.

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

Grande, T., H. Laten, and J.A. Lopez. 2004. Phylogenetic relationships of extant esocid species (Teleostei: Salmoniformes) based on morphological and molecular characters. Copeia 2004: 743-757.

Greeley, J.R. 1930. Fishes of the Lake Champlain Watershed. Pages 44-87 *in* State of New York. A Biological Survey of the Lake Champlain Watershed. J.B. Lyons Company, Albany. 321 pp.

Kyle, C.J. and C.C. Wilson. 2007. Mitochondrial DNA identification of game and harvested fish species. Forensic Science International 166: 68-76.

Miller, L.M., S.W. Mero, and J.A. Younk. 2009. The genetic legacy fo stocking muskellunge in a northern Minnesota lake. Transactions of the American Fisheries Society 138: 602-615.

Peakall, R., and P. E. Smouse. 2006. GenAlEx 6: Genetic Analysis in Excel, population genetic software for teaching and research. Molecular Ecology Notes 6: 288–295.

Piller, K. R., C. C. Wilson, C. E. Lee, and J. Lyons. 2005. Conservation genetics of inland lake trout in the upper Mississippi River basin: stocked or native ancestry? Transactions of the American Fisheries Society 134: 789–802.

Piry, S., A. Alapetite, J.-M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95: 536–539.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Rannala, B., and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences of the United States of America 94: 9197–9201.

Sloss, B.L., R.P. Franckowiak, and E.L. Murphy. 2008. Development of new microsatellite loci and multiplex reactions for muskellunge (*Esox masquinongy*). Molecular Ecology Notes (in press).

Sokal, R.R. and F.J. Rohlf. 1995. Biometry (3rd edition). W.H. Freeman and Sons, New York.

Thompson, Z. 1853. Natural History of Vermont. Charles E. Tuttle Company, Rutland, Vermont.

Vermont Commissioners of Fisheries and Game. 1898. Fourteenth biennial report of the Commissioners of Fisheries and Game, ending with the fiscal year, June 30, 1898. Argus and Patriot Press, Montpelier.

Wilson, C.C., T.M. Lavender, and J. Black. 2007. Genetic assessment of walleye (*Sander vitreus*) restoration efforts and options in Nipigon Bay and Black Bay, Lake Superior. Journal of Great Lakes Research 33 (Supplement 1): 133-144.

Table 1: Muskellunge samples examined in this study, showing source location, field sampling method, and tissues used for DNA extraction.

Geographic location	N	Site details	Capture method	Date	Source tissue
Missisquoi River	1	below Swanton dam	angling	August 2003	dried fin
Missisquoi River	1	below Swanton dam	angling	September 2005	fin tissue (95% ethanol)
Missisquoi River	1	above Swanton dam	angling	June 1980	dried fin
Otter Creek	1	above Vergennes dam	angling	June 2005	muscle (95% ethanol)
Otter Creek	1	below Vergennes dam	angling	August 2005	fin tissue (95% ethanol)
Otter Creek	1	below Vergennes dam	angling	September 2006	fin tissue (95% ethanol)
Otter Creek	1	below Vergennes dam	VTDFW electrofishing	June 2006	fin tissue (95% ethanol)
Otter Creek	1	below Vergennes dam	VTDFW electrofishing	June 2006	fin tissue (95% ethanol)
Great Chazy R (NY)	166		hatchery fry	May 2006	preserved fry (95% ethanol)
Pymatuning Reservoir (PA)	75		trapnets	1997, 2004	dried scales

Table 2: Assignment of individual wild-caught muskellunge from the Missisquoi River and Otter Creek to two potential stocking sources based on individual-assignment tests implemented in GeneClass2 (Piry et al. 2004), showing the relative likelihoods of assignment to most probable source (A) or second most probable source (B), as well as their absolute log likelihood values of assignment to each source.

unknown	**Assigned source A**	**% likelihood**	**Assigned source B**	**% likelihood**	**-log(L_A)***	**-log(L_B)***
Missisquoi R 1	Pymatuning R	92.202	GreatChazy R	7.798	14.478	13.405
Missisquoi R 2	Great Chazy R	99.581	Pymatuning R	0.419	4.491	6.867
Missisquoi R 3	Great Chazy R	**excluded***	Pymatuning R	**excluded***	**22.025**	**23.241**
Otter Creek 1	Great Chazy R	90.477	Pymatuning R	9.523	15.886	16.863
Otter Creek 2	Pymatuning R	80.561	Great Chazy R	19.439	12.951	12.334
Otter Creek 3	Pymatuning R	99.966	Great Chazy R	0.034	15.855	12.389
Otter Creek 4	Great Chazy R	57.386	Pymatuning R	42.614	14.317	14.446
Otter Creek 5	Pymatuning R	87.683	Great Chazy R	12.317	14.236	13.384

Note: percent likelihood values reflect the ratio between the two log-likelihood assignment probabilities; probability values below 10^{-15} (ie -log(L)$\geq$15) should be considered improbable as occurring from the source(s) listed. The 1980 sample from the Missisquoi River (fish 3) was excluded from both stocking sources ($p<<0.0001$).

Appendix I (attached spreadsheet)**:** Multilocus microsatellite genotypes of muskellunge screened in this study, listed by sampling source/location and individual, based on microsatellite loci developed by Sloss et al. (2008).

Appendix II: Homology of mitochondrial sequences obtained from two suspected *Esox* interspecific hybrids with *Esox* reference sequences in GenBank (http://www.ncbi.nlm.nih.gov/BLAST) from voucher specimens from Grande et al. (2004), based on amplification and sequencing of a 500-base region of the mitochondrial cytochrome *b* gene, using protocols described by Kyle and Wilson (2007). Both putative hybrids had the same mitochondrial sequence.

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
AY497431.1	*Esox americanus americanus* voucher LUF09815 cytochrome b (cytb) gene, complete cds; mitochondrial	898	898	100%	0.0	99%
AY497432.1	*Esox americanus americanus* voucher LUF09817 cytochrome b (cytb) gene, complete cds; mitochondrial	898	898	100%	0.0	99%
AY497441.1	*Esox niger* voucher LUF09832 cytochrome b (cytb) gene, complete cds; mitochondrial	898	898	100%	0.0	99%
AY497440.1	*Esox niger* voucher LUF09831 cytochrome b (cytb) gene, complete cds; mitochondrial	898	898	100%	0.0	99%
AY497439.1	*Esox niger* voucher LUF09824 cytochrome b (cytb) gene, complete cds; mitochondrial	898	898	100%	0.0	99%
AY497438.1	*Esox niger* voucher LUF09823 cytochrome b (cytb) gene, complete cds; mitochondrial	898	898	100%	0.0	99%

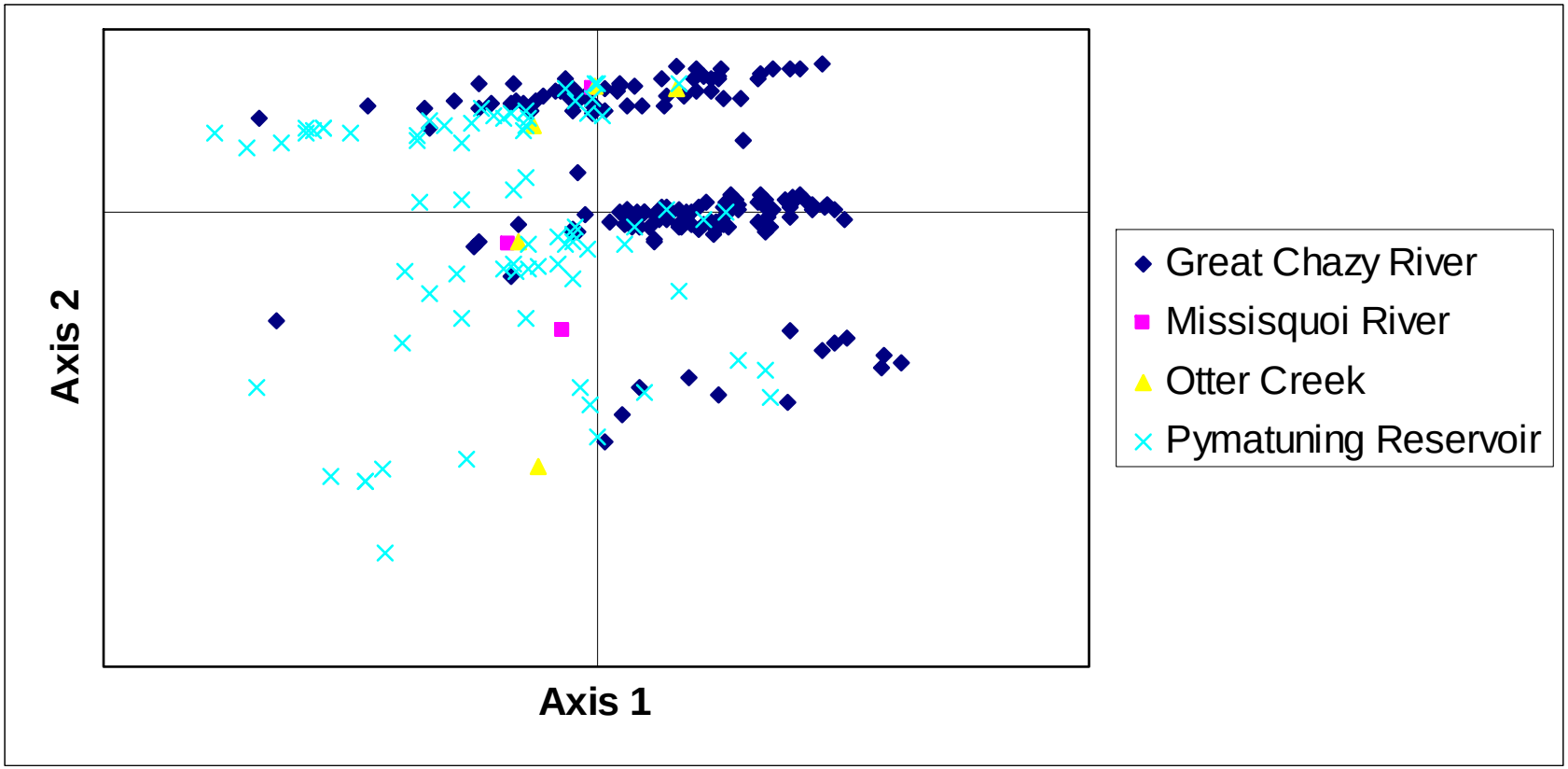


Figure 1: Principal coordinate analysis (PCoA) ordination of multilocus microsatellite genotypes of muskellunge samples from Missisquoi River and Otter Creek, VT, in comparison with potential stocking sources (Great Chazy River and Pymatuning Reservoir). The first two axes (shown) accounted for 54.8% of the overall variation. None of the wild-caught muskellunge showed significant divergence from the two stocking sources with this approach.

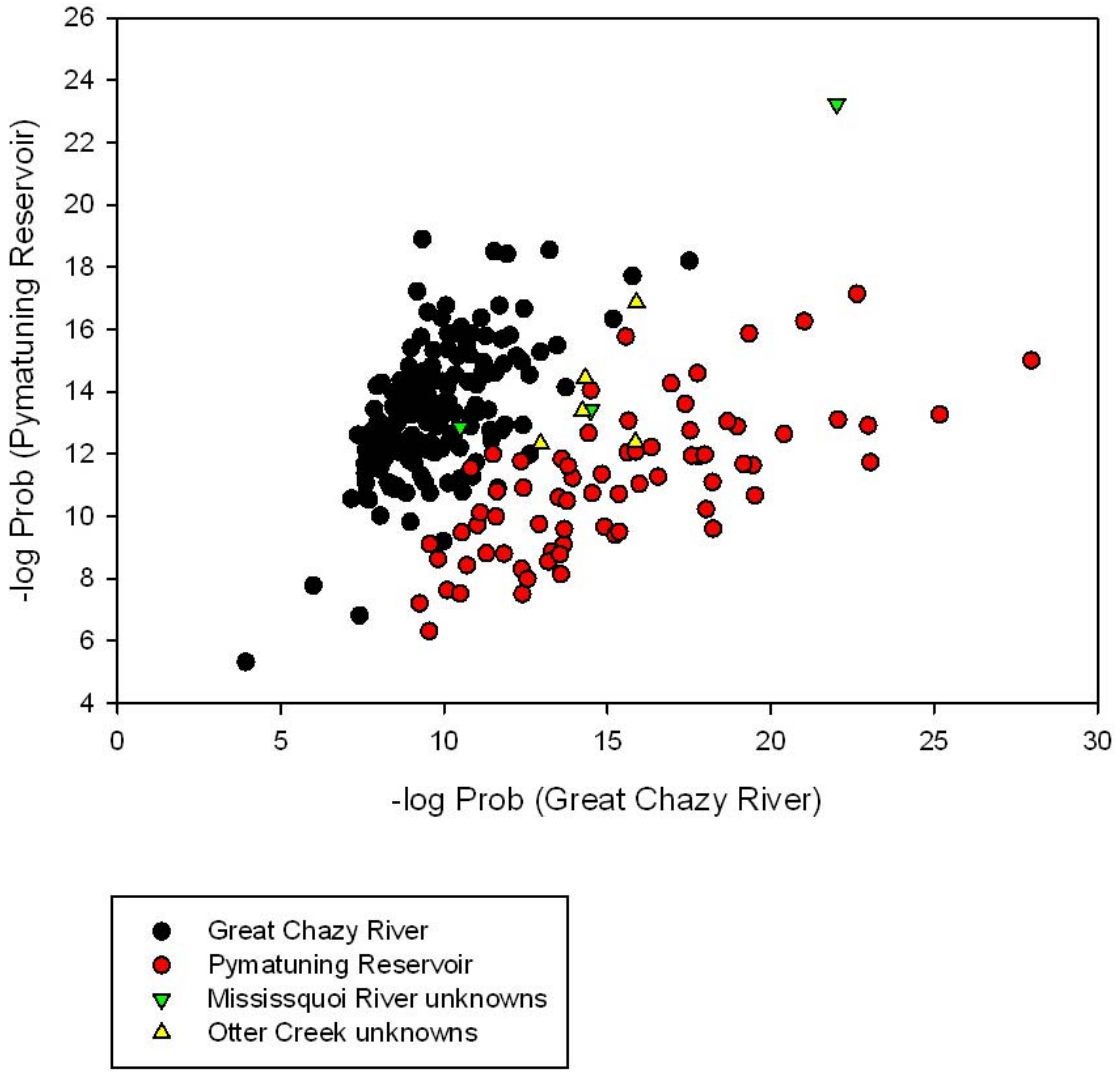
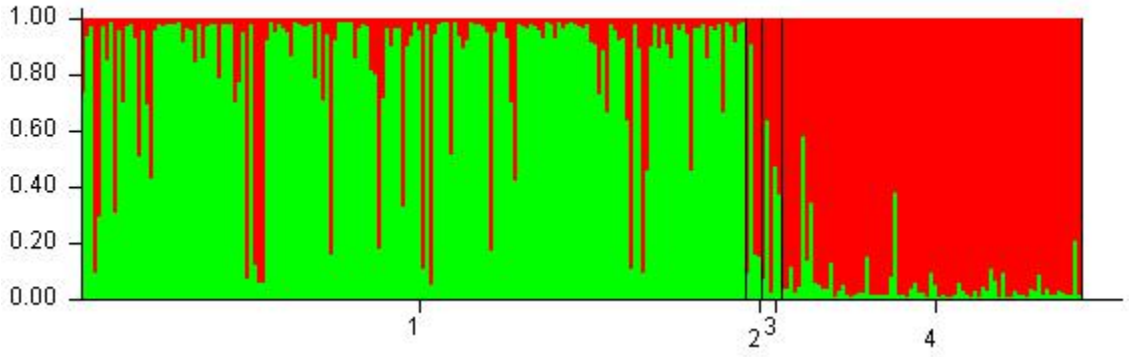


Figure 2: Output results from GeneClass (Piry et al. 2001), showing membership probabilities of individual muskellunge to stocking sources used for Lake Champlain (Great Chazy River and Pymatuning Reservoir), based on multilocus genotype log-likelihood probabilities. Wild-caught muskellunge are shown as green (Missisquoi River) and yellow (Otter Creek) triangles. The outlier individual in the upper right represents the historical (1980) muskellunge captured in the Missisquoi River, and is significantly different from both stocking sources ($p << 0.0001$).

Figure 3: Output results from STRUCTURE (Pritchard et al. 2000), showing fractional memberships of individual muskellunge to genetic clusters, based on differing numbers of genetic groups. Each vertical bar represents the genetic membership for one individual; different colours represent distinct genetic groups based on generalized centroid clustering. Solutions shown are (a and b) admixture (K=2 and 5, respectively); (c-d) no admixture (K=2 and 6, respectively); and (e) using population information to update stocking source genetic assignment. For (a) to (d), represented populations are in alphabetical order: (1) Great Chazy River, NY; (2) Missisquoi River; (3) Otter Creek; (4) Pymatuning Reservoir, PA. For solution (e) (using source population information), population order is (1) Great Chazy River, NY; (2) Pymatuning Reservoir, PA; (3) Missisquoi River; (4) Otter Creek.

(a) admixture (K=2)



(b) admixture (K=5)

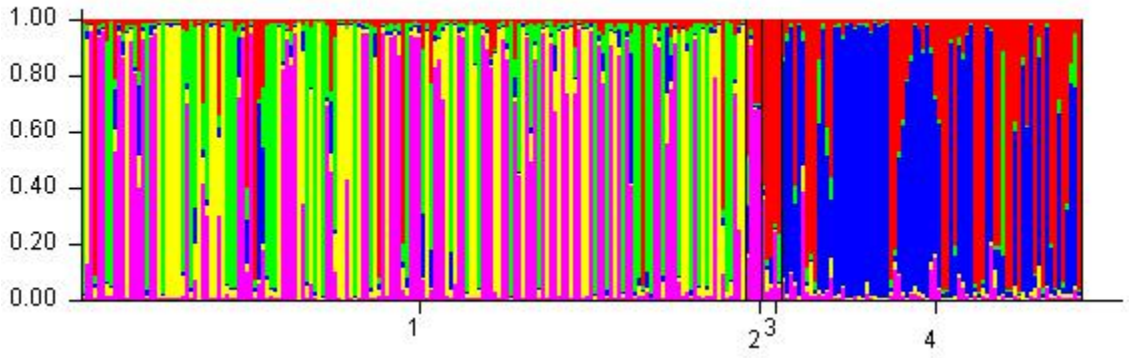
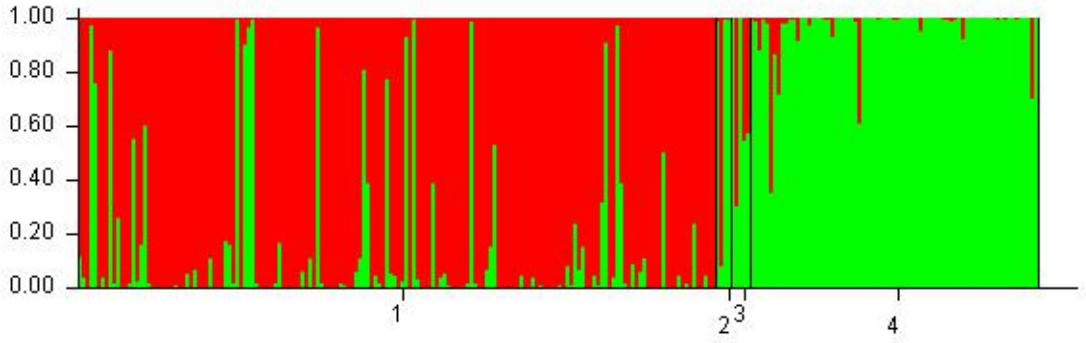
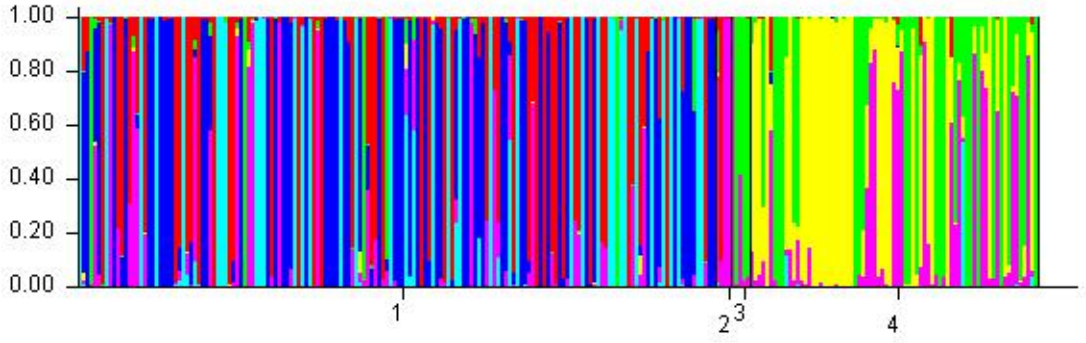


Figure 3 (continued)

(c) no admixture (K=2)



(d) no admixture (K=6)



(e) using source population information (K=2)