Molecular Ecology (2003) **12**, 1689–1702 doi: 10.1046/j.1365-294X.2003.01849.x

Microsatellite variation within and among North American lineages of *Phragmites australis*

K. SALTONSTALL*
Department of Ecology and Evolutionary Biology, Yale University, New Haven, USA

Abstract

Over the past century, the spread of the common reed (*Phragmites australis*) has had a dramatic impact on wetland communities across North America. Although native populations of *Phragmites* persist, introduced invasive populations have dominated many sites and it is not clear if the two types can interbreed. This study compares patterns of differentiation in 10 microsatellite loci among North American and European *Phragmites* individuals with results obtained from sequencing of noncoding chloroplast DNA. Three population lineages (native, introduced and Gulf Coast) were previously identified in North America from chloroplast DNA and similar structuring was found in the nuclear genome. Each lineage was distinguished by unique alleles and allele combinations and the introduced lineage was closely related to its hypothesized source population in Europe. Size homoplasy and diagnostic base substitutions distinguishing lineages were evident at several loci, further emphasizing that native, introduced and Gulf Coast North American *Phragmites* lineages are genetically distinct. Gene flow between lineages was low and invasive introduced populations do not represent a hybrid population type.

Keywords: hybridization, invasive species, microsatellite DNA markers, *Phragmites australis*, population structure

Received 18 October 2002; revision received 18 February 2003; accepted 18 February 2003

Introduction

Exotic invasive species pose one of the greatest threats to species and habitat diversity today, second only to habitat loss (Wilcove *et al*. 1998). In addition to their impacts at the community and ecosystems levels, such introductions pose a number of threats at the genetic level to closely related species and locally adapted populations of the same species. Perhaps most significant are the effects of hybridization with closely related species (Anderson & Stebbins 1954; Lewontin & Birch 1966), which can quickly change the genetic composition of native populations and cause loss of local adaptations (Anttila *et al*. 1998; Huxel 1999). It may also stimulate invasiveness in introduced populations by increasing genetic variability, creating favourable gene combinations, masking deleterious alleles (Ellstrand & Scheirenbeck 2000; Lee 2002), and causing changes in ploidy levels which often increases fitness (Soltis & Soltis 2000).

In natural plant populations, microsatellites have great potential for helping to understand what determines patterns of genetic variation, particularly when used in concert with chloroplast DNA (cpDNA) markers. Their utility has been demonstrated in studies of genetic diversity (Morand *et al*. 2002), mating systems (Durand *et al*. 2000), pollination biology (White *et al*. 2002) and seedling establishment (Dow & Ashley 1996). Few studies have been carried out using microsatellites in analysis of population structure of polyploid species. This is likely due to the problems in analysing polyploid data as well as difficulties in amplifying loci, possibly because of differences in the parental genomes of polyploids (Röder *et al*. 1995).

An area of concern when using microsatellite frequencies to estimate population parameters is size homoplasy which can cause underestimation of the number of alleles, levels of heterozygosity and genetic diversity when only allele sizes are considered (reviewed in Estoup *et al*. 2002). A number of studies have demonstrated that microsatellite alleles of the same size can arise from mutation events

Correspondence: K. Saltonstall. *Present address: Horn Point Laboratory, University of Maryland Center for Environmental Science, PO Box 775, Cambridge, MD 21613, USA. Fax: +1 410 221 8490; E-mail: kristin.saltonstall@aya.yale.edu

which either interrupt repeat units or occur in the regions flanking the repeat region. This has been shown to occur both within (Angers & Bernatchez 1997; Viard *et al*. 1998) and among populations (Estoup *et al*. 1995; Viard *et al*. 1998) and closely related species (Peakall *et al*. 1998; van Oppen *et al*. 2000). One approach to minimizing the risk of misinterpretation of genetic information is to characterize different electromorphs by sequencing, particularly in cases in which other genetic data (e.g. chloroplast or mitochondrial sequences) suggest strong levels of genetic structuring that is not being detected by microsatellite analysis.

The common reed (*Phragmites australis*, hereafter *Phragmites*) has a worldwide distribution and is considered native to North America. However, its distribution and abundance have increased dramatically over the past 150 years in North America and it is considered a nuisance species in many regions. *Phragmites* is a perennial wetland grass that is tolerant of both freshwater and brackish conditions and typically forms a tall dense monoculture in the systems that it invades. Its ability to readily colonize open habitat both by spread of seed and fragments of rhizome material has undoubtedly contributed greatly to the invasive spread of the species. *Phragmites* shows high variation in ploidy levels throughout its distribution. Throughout the genus, the base chromosome number is 12 (Tucker 1990). Although the diploid number ($2n = 24$) has never been reported in a natural population, a wide range of polyploid and aneuploid population types have been reported (reviewed in Clevering & Lissner 1999). Karyotypic studies of North American populations indicate the presence of $3x$, $4x$, $6x$ and $8x$ plants, with $4x$ being the dominant chromosome number in modern day populations (Chambers *et al*. 1999).

In this study, 10 microsatellite loci were used to assess the genetic structure of *Phragmites* populations in North America and Europe at the level of nuclear DNA. Recent evidence based on sequencing of cpDNA indicates that a single nonnative lineage of *Phragmites* has been introduced to North America and is likely to be responsible for the dramatic spread of the species across North America (Saltonstall 2002). Native individuals persist in many locations, such as in the Midwest and western parts of the country, but introduced populations have overrun the Atlantic coast region and eliminated the majority of native populations (Saltonstall 2003). Five diagnostic chloroplast mutations were found to distinguish native and introduced *Phragmites* lineages in North America (Saltonstall 2002). However, because cpDNA has only maternal inheritance patterns, further analysis of whether the two lineages are hybridizing, whether hybrid individuals may be responsible for the widespread invasions, and the extent of nuclear genetic variability in the introduced lineage is needed.

In this context, the specific objectives of this study were to analyse nuclear microsatellite DNA polymorphisms in order to: (i) explore whether patterns of genetic structure seen in cpDNA are also found in the nuclear genome; (ii) examine genetic diversity in native and introduced *Phragmites* haplotype lineages; and (iii) assess whether hybridization between native and introduced *Phragmites* lineages has occurred and what role it might play in the invasions of this species in North America.

Materials and methods

Sampling

Leaf tissue samples were collected throughout the range of *Phragmites* with an emphasis on Europe and North America. In total, 654 modern clones were sampled, with 533 from North America, 70 from Europe and 51 from elsewhere in the world. In addition, 65 herbarium samples collected between 1852 and 1910 in North America were obtained for comparisons between modern and historical populations. Detailed sampling locations are available by request from the author. Tissues were stored dry until they arrived in the laboratory where they were frozen. DNA was extracted using a 2% CTAB extraction protocol (Doyle & Dickson 1987). A subset of samples was used from each region for cpDNA analysis (Saltonstall 2002, 2003) and all samples were used for microsatellite analysis. Sample ploidy levels were unknown.

Microsatellite development

Primers were developed using the microsatellite enrichment technique of Fischer & Bachmann (1998), modified as follows: genomic DNA from a sample collected in Connecticut was digested with *Rsa*I and *Hae*III using universal KGB digestion buffer (McClelland *et al*. 1988) and two oligo adaptors (Edwards *et al*. 1996) were ligated to the digested fragments. Pooled $(dGA)_{10}$ and $(dGT)_{10}$ biotinylated oligo probes were hybridized to the cut DNA and isolated using streptavidin-coated magnetic beads. Polymerase chain reactions (PCRs) on the resulting microsatellite-enriched eluate were performed and the enrichment process repeated. Following this second round of enrichment and PCR amplification, the enriched library was ligated into pcR 2.1 vector (Invitrogen) and transformed into TOP10 cells. Plasmid DNA from clones was column purified (Qiagen) and sequenced on either an ABI 373 or ABI 377 automatic sequencer (PE Applied Biosystems) using dye terminator chemistry.

Detection of length polymorphism and size homoplasy

Primers for PCR amplification were designed in microsatellite flanking regions by eye, leaving at least 30 bp of flanking sequence on either side of the microsatellite

region. PCRs contained 50–100 ng DNA, 1× PCR buffer (10 mM Tris–HCl, 50 mM KCl, 0.001% w/v gelatin), 1.5 mM $MgCl_2$, 0.5 mM each dNTP, 2 pmol of each primer (FAM, TET or HEX end-labelled), and 1 unit of AmpliTaq Gold polymerase (Applied Biosystems) in a total volume of 10 μL. Thermal cycling was carried out on a Hybaid Multiblock thermocycler using the following conditions: an initial denaturation at 94 °C for 12 min, followed by 35 iterations of 94 °C for 30 s, 50–56 °C for 30 s and 72 °C for 4 s, and finally a 2-min extension at 72 °C. Products were run on 6% acrylamide gels on an ABI 373 automatic sequencer using Genescan 500 size standard (Applied Biosystems).

To check for size homoplasy of alleles between lineages, the majority of alleles were sequenced from individuals who displayed homozygosity for a particular allele. Alleles were amplified in 50 μL PCRs, gel purified (Bio 101), and sequenced using BigDye chemistry on an ABI 377 machine. Sequences were aligned using SEQUENCHER 4.1 (GeneCodes Corp.).

Data analysis

Population lineages were defined based on genetic structuring found in cpDNA analysis of a subset of the samples analysed here (Saltonstall 2002; 2003). In this previous study, three major haplotype lineages were identified within North America: native (haplotypes A–H, S, Z, AA), introduced (haplotype M) and Gulf Coast (haplotype I). Europe was added as a fourth group in this analysis, to allow for a comparison of the introduced lineage with its putative source and to facilitate detection of hybrids between the native and introduced lineages. Although samples from other parts of the world (e.g. South America, Asia, Australia, Africa) were also collected (Other in Fig. 1), they were included only in the analysis of allelic diversity and genetic distance due to the small sample sizes obtained from other locations. Complete genotypic profiles of all samples are available from the author by request.

Owing to the polyploid nature of *Phragmites*, the number of alleles detected per individual varied by locus. Although microsatellites generate codominant markers, problems can arise during the identification of polyploid genotypes because it is difficult to determine the number of copies of an allele in heterozygotes. Thus, banding patterns observed at particular loci are referred to as 'allele phenotypes' (Becher *et al*. 2000). Allelic diversity was measured as the total number of alleles per locus. Total numbers of allele phenotypes were calculated as was observed heterozygosity (H_O), which is the proportion of individuals displaying a heterozygotic allele phenotype at a locus. Genetic distances between populations were calculated using the program POPDIST (Guldbrandtsen *et al*. 2000)

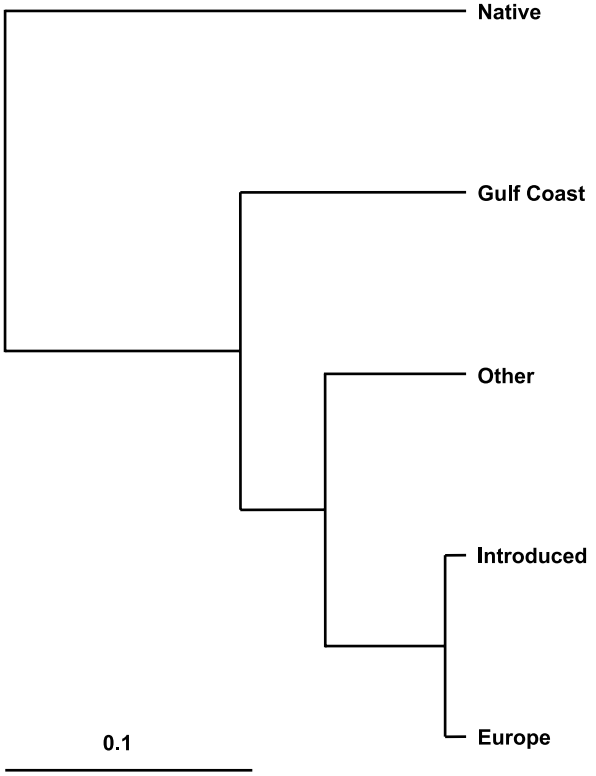

Fig. 1 UPGMA tree constructed from D_{TL} distances showing the relationships of five population groups of *Phragmites*. Genetic distances are based on allele phenotypes at 10 microsatellite loci. Native: North American samples possessing cpDNA haplotypes A–H, S, Z, AA ($n = 137$); Gulf Coast: North American samples possessing cpDNA haplotype I ($n = 27$); Introduced: North American samples possessing cpDNA haplotype M ($n = 171$); Europe: European samples (haplotypes L–O, T; $n = 70$); Other: samples collected in Asia, Australia, Africa and South America (haplotypes I–M, O–R, U–Y; $n = 51$).

based on the measure of Tomiuk & Loeschcke (1991, 1996) which calculates genetic distance (D_{TL}) based on genotype frequencies and allows for comparison of polyploid populations. Based on the number of alleles amplified per sample, all haplotype lineages appeared to be of mixed ploidy levels, thus all samples were coded as hexaploids for this calculation as the maximum number of alleles found in any single sample was six. A UPGMA tree based on the D_{TL} distance was constructed using the NEIGHBOR program in PHYLIP 3.6a3 (Felsenstein 1995) and viewed using TREEVIEW (Page 1996).

A more detailed analysis of population structure in North America and Europe was performed using six loci (*PaGT4*, *PaGT8*, *PaGT9*, *PaGT13*, *PaGT14*, *PaGT16*) from which the majority of samples amplified only one or two alleles and thus displayed an inheritance pattern similar to diploids. The remaining samples amplified three alleles, many of which were identical and belonged to the Gulf Coast lineage. These samples were then coded to display

Table 1 Primer sequences and allelic diversity measures for ten microsatellite loci of *Phragmites australis* pooling 719 individuals collected worldwide. Allele phenotypes are the number of different multiallelic profiles found at each locus. H_O: proportion of individuals displaying a multiallelic banding pattern. The clone sequences from which the primers were designed have GenBank Accession Nos. AY230868–AY230876

Locus	Primer sequences (5′–3′)	No. repeat units	Maximum no. alleles per sample	Allele size range (bp)	Total no. Alleles	No. allele phenotypes	H_O
PaGT4	F: TGCTCCCTGCCAGTTTCTTG R: TATCCACCCTTCGAAGGCAC	$(CA)_9$	3	266–284	10	30	0.35
PaGT8	F: TCTGAACATAATCCTGGTGG R: TCTGTGTGAAGCAGTTCTGC	$(CA)_8$	2	170–193	13	19	0.20
PaGT9	F: CCATGTGTTAATGTTGTCC R: ATTGAATCCACACGTTTCCG	$(CA)_{10}$	3	188–224	17	52	0.49
PaGT11	F: CAACTCCGTGAATGACATGC R: CAGTTTGTGCACTAATGGAC	$(CA)_8$	3	142–151	5	15	0.69
PAGT12	F: CTTCCTAGGTCAGTATCATCC R: GTGGCAGCTGATTGATTTGG	$(CA)_9$	4	151–196	10	39	0.51
PaGT13	F: CTCATGCATCACTTCACAGG R: ACACGGACCTAACATCAACC	$(CA)_9$	3	206–224	8	19	0.30
PaGT14	F: GTTGCAGCAAGTATTTGG R: CAAGCATTCTAGTAGTAGC	$(CA)_7$	4	169–198	18	74	0.43
PaGT16	F: ACCAATCAGTCAGACTAGCC R: GTTCTCATGTTGGAGAAGCC	$(CA)_{10}$	3	231–298	17	44	0.41
PaGT21	F: GCTACTCAACAGGTATACGG R: ATTGAGGATTGAGGTGGTGG	$(CA)_5(AT)_6$ $(CA)_6$	3	138–199	21	47	0.35
PaGT22	F: TTGAGTGCCTGGTGTATTCG R: AAGCTTCTGTCATGGAACCG	$(AC)_8CTT$ $(GA)_5$	3	159–209	22	63	0.52

only two alleles at the locus (*PaGT4*, *PaGT14*) and the data set analysed like diploid samples. Any other samples amplifying three alleles at a locus were coded as missing data. Although this method of analysis assumes that allelic dosage follows fixed parental inheritance patterns and that the presence of two alleles implies two copies of each allele rather than one copy of one and three of the other (in the case of a tetraploid), it allowed for a more detailed analysis of population structure and assignment of samples with an unknown chloroplast type to a lineage. Given that *Phragmites* is an allopolyploid (Raicu *et al*. 1972), which have disomic inheritance patterns from parental genomes (Soltis & Soltis 2000), this assumption is plausible.

In this six-locus analysis, linkage disequilibrium between loci and deviations from Hardy–Weinberg equilibrium (HWE) were tested for each population type across all six loci using Fisher exact tests with GENEPOP (Raymond & Rousset 1995). Levels of differentiation among populations were tested using individual assignment tests as described by Cornuet *et al*. (1999). The 'leave one out' procedure was used when calculating allele frequencies and a cut-off value of < 5% was used to assign individuals to a population. Both Bayesian and genetic distance estimators were used as assignment methods. Assignment values were calculated for each sample belonging to population lineages represented by chloroplast haplotypes. These lineages were then used as reference populations to assign 208 samples with unknown cpDNA haplotypes to the most likely population lineage. Samples that were assigned to more than one lineage were considered to be potential hybrids and their allelic profiles examined more closely. This included a comparison of allele phenotypes between historical and modern native and Gulf Coast lineages with introduced populations to look for changes in population structure over time. Regional groupings of Atlantic Coast (New Brunswick and Nova Scotia south to South Carolina), Midwest (Great Lakes region and southern Canada), West (Pacific northwest and southwest) and Gulf Coast (Florida west to the Gulf of California) samples were compared to look for genetic structuring at smaller geographical scales as well.

Results

Primer development and screening of allelic diversity

Ten primer pairs were selected for further study (Table 1). Nearly all of the microsatellite sequences were perfect with no interruption to the core-motif repeat sequence (Table 1). Average length of the repeat unit was 9 repeats (range 7–10). One to four alleles amplified within individual samples and between five (*PaGT11*) and twenty-two alleles

Table 2 Summary of allele sequences from selected *Phragmites* microsatellite loci displaying size homoplasy and the effect of indel variation in flanking regions on allele size. Microsatellite repeat regions are shown in bold text. Variations in flanking regions are shown in bold, underlined text

Locus	Size (bp)	
PaGT4	266	AACAATGAAAAATGGGAAGAA**ACACACACAC**::::::::::TTATCCCCCAAGAAGGAAGAACG**C**GAGTTCCTCACCACGA GAAAACAGACCAGAAATCGAGAAGCCAAA**A**TCAA
	274	AACAATGAAAAATGGGAAGAA**ACACACACACACACACAC**::TTATCCCCCAAGAAGGAAGAACG**C**GAGTTCCTCACCACGA GAAAACAGACCAGAAATCGAGAAGCCAAA**A**TCAA
	274	AACAATGAAAAATGGGAAGAA**ACACACACACACACACAC**::TTATCCCCCAAGAAGGAAGAACG**A**GAGTTCCTCACCACGA GAAAACAGACTAGAAATCGAGAAGCCAAA**T**TCAA
	276	AACAATGAAAAATGGGAAGAA**ACACACACACACACACACAC**TTATCCCCCAAGAAGGAAGAACG**A**GAGTTCCTCACCACGA GAAAACAGACTAGAAATCGAGAAGCCAAA**T**TCAA
PaGT8	176	TGGCGCAGCTAGTTAGCAGTACATGCATTT::**GTGTGTGTGTGTGT**::::GAACTGGTGATCAAATGCTCAGAGCAACCTT TGTGTGTGTGGACTGTAAATGCTCTCTGTAGA
	178	TGGCGCAGCTAGTTAGCAGTACATGCATTT::**GTGTGTGTGTGTGTGT**::GAACTGGTGATCAAATGCTCAGAGCAACCTT TGTGTGTGTGGACTGTAAATGCTCTCTGTAGA
	180	TGGCGCAGCTAGTTAGCAGTACATGCATTT::**GTGTGTGTGTGTGTGTGT**GAACTGGTGATCAAATGCTCAGAGCAACCTT TGTGTGTGTGGACTGTAAATGCTCTCTGTAGA
	182	TGGCGCAGCTAGTTAGCAGTACATGCATTT**TT**GTGTGTGTGTGTGTGTGTGTGAACTGGTGATCAAATGCTCAGAGCAACCTT TGTGTGTGTGGACTGTAAATGCTCTCTGTAGA
PaGT16	239	CCCT**GTGTGT**::::::::::::::::::ATGTATGT::AGATCACTC:::::::::::::::::::::::::::::::: ::::::::::::::::::ATTTCTCTCATATTG
	261	CCCT**GTGTGTGTGTGTGTGTGT**GAAATGTATGTATGT**GT**AGATCACTC:::::::::::::::::::::::::::::::: ::::::::::::::::::ATTTCTCTCATATTG
	265	CCCT**GTGTGT**::::::::::::::::::ATGTATGT::AGATCACTC**CCTGTGTGTATGTATGTAGATCACTC**::::::: ::::::::::::::::::ATTTCTCTCATATTG
	290	CCCT**GTGTGT**::::::::::::::::::ATGTATGT::AGATCACTC**CCTGTGTGTATGTATGTAGATCACTCCCCTGTGT GTATGTATGTAGATCACTC**ATTTCTCTCATATTG

(*PaGT22*) per locus. The maximum number of allele phenotypes from a single locus was 74 (*PaGT14*). H_O ranged between 20% (*PaGT8*) and 69% (*PaGT11*) of samples depending on the locus (Table 1).

Sequence analysis

Sequence data from amplified alleles revealed multiple mutations in the regions flanking the microsatellites (GenBank Accession nos AY233803–AY233855). Insertion/deletion (indel) mutations accounted for many of the large size changes in alleles (Table 2). In some cases, indels had the effect of changing allele sizes in a direction which did not match the pattern of change seen in the microsatellite repeat region itself. For example, at locus *PaGT16* the 261 bp allele had nine repeat units in the microsatellite, whereas the 290 bp allele had only three repeats (Table 2). Indels have had the effect of increasing the size of the 290 bp allele to 41 bp longer than is reflected by the microsatellite repeat alone. Indel mutations also created different sized amplification products which had identical numbers of microsatellite repeats. This was seen at locus *PaGT8* where four of the alleles (176, 178, 180, 182) vary in multiples of two nucleotides. However, when these alleles were sequenced, the size variation creating the 182 bp allele was due to a 2-bp insertion adjacent to the microsatellite (Table 2). Further, base substitutions caused allelic variation that was not detectable based on allele sizing (size homoplasy), such as the 274 bp allele at locus *PaGT4* where some North American populations have three base substitutions (shared with the 266 bp allele) that are not found in populations originating from other parts of the world with a 274-bp allele (Table 2). Size homoplasy was also found at loci *PaGT12* (173 bp allele) and *PaGT13* (208 bp allele).

Twelve locus analyses

Levels of polymorphism by haplotype lineage. Within North American haplotype groups, the introduced lineage showed the highest number of alleles per locus and Gulf Coast the fewest (Table 3). Although introduced samples also shared alleles with samples from other parts of the world (data not shown), their allelic distribution was most similar to European samples. H_O was higher in the Gulf Coast lineage across most loci than other North American lineages. This is due in part to the fact that several loci were nearly fixed for certain heterozygous allele phenotypes in this lineage (*PaGT4*, *PaGT8*, *PaGT9*, *PaGT22*). The native lineage had the lowest heterozygosity levels across all loci, despite having a moderately high number of alleles per locus (Table 3). This was due to rare alleles occurring in a

Table 3 Genetic characteristics of 10 microsatellite loci in four *Phragmites* haplotype lineages. *n*: Number of samples genotyped; A_o: observed number of alleles; Dominant phenotypes: dominant allele phenotypes, values in parentheses are the frequency of each phenotype; H_O: observed heterozygosity

	Haplotype lineage			
Locus	Introduced	Native	Gulf Coast	Europe
All loci				
Mean A_o	6.5 ± 3.5	5.9 ± 2.2	3.6 ± 1.2	8.0 ± 3.6
Mean H_O	0.46 ± 0.21	0.27 ± 0.27	0.76 ± 0.34	0.54 ± 0.24
PaGT4				
n	150	125	21	57
A_o	9	3	3	6
Dominant	274 (0.12)	266 (0.73)	274/276/280 (0.86)	274 (0.40)
phenotypes	276 (0.43)	274 (0.16)		276 (0.11)
	274/276 (0.30)			274/276 (0.19)
H_O	0.40	0.10	0.90	0.33
PaGT8				
n	131	92	22	57
A_o	2	3	2	4
Dominant	176 (0.76)	178 (0.71)	176/189 (0.95)	176 (0.49)
phenotypes	176/178 (0.19)	180 (0.14)		178 (0.18)
				176/178 (0.25)
H_O	0.19	0.04	0.95	0.33
PaGT9				
n	121	109	21	45
A_o	10	9	3	13
Dominant	198 (0.34)	210 (0.62)	192/196 (0.95)	198/204 (0.09)
phenotypes	198/202 (0.24)	210/212 (0.20)		198/206 (0.11)
H_O	0.55	0.31	1.00	0.83
PaGT11				
n	137	110	17	69
A_o	4	3	3	5
Dominant	142/147 (0.59)	145 (0.84)	142/147 (0.71)	142 (0.13)
phenotypes	142/147/149 (0.24)		142/145/147 (0.29)	142/147 (0.70)
H_O	0.91	0.10	1.00	0.86
PAGT12				
n	127	97	22	61
A_o	6	7	6	6
Dominant	176 (0.29)	176 (0.25)	176 (0.64)	176 (0.18)
phenotypes	176/180 (0.18)	182 (0.18)		176/180 (0.16)
	176/184 (0.16)	176/180 (0.21)		176/178/180 (0.10)
H_O	0.53	0.53	0.18	0.70
PaGT13				
n	137	101	22	55
A_o	2	7	4	3
Dominant	208 (0.55)	208 (0.40)	210 (0.27)	208 (0.71)
phenotypes	208/210 (0.34)	218 (0.37)	208/210 (0.68)	210 (0.12)
H_O	0.34	0.18	0.73	0.11
PaGT14				
n	136	127	23	64
A_o	12	8	5	11
Dominant	189 (0.27)	181 (0.83)	177/189 (0.52)	181 (0.20)
phenotypes	183/189 (0.15)		177/185/189 (0.3)	189 (0.09)
	187/189 (0.12)			187/189 (0.08)
H_O	0.56	0.06	0.96	0.59
PaGT16				
n	128	101	23	58
A_o	4	6	3	11

Table 3 *Continued*

	Haplotype lineage			
Locus	Introduced	Native	Gulf Coast	Europe
Dominant	261 (0.37)	261 (0.20)	261 (0.13)	261 (0.10)
phenotypes	290 (0.26)	265 (0.44)	265/290 (0.74)	263 (0.12)
	261/290 (0.30)	261/265 (0.21)	265/290 (0.74)	261/263 (0.10)
				261/290 (0.14)
H_O	0.33	0.31	0.83	0.67
PaGT21				
n	104	50	11	27
A_o	9	6	4	10
Dominant	173 (0.22)	138/158 (0.74)	173 (0.45)	183 (0.11)
phenotypes	191 (0.18)	158/173 (0.1)	175 (0.36)	195 (0.19)
	193 (0.10)			199 (0.11)
	195 (0.31)			175/193 (0.11)
				175/195 (0.15)
H_O	0.23	0.90	0.09	0.44
PaGT22				
n	95	25	7	36
A_o	7	7	3	11
Dominant	181 (0.20)	185 (0.24)	175/183/197 (1.00)	183 (0.08)
phenotypes	197 (0.13)	191 (0.40)		195 (0.14)
	181/193 (0.21)	193 (0.12)		199 (0.08)
	181/197 (0.18)			195/199 (0.08)
H_O	0.58	0.12	1.00	0.58

few samples, with the majority of samples dominated by common alleles. For example, at locus *PaGT9*, 89% of the samples had the 210 bp allele with all other alleles occurring in fewer than 10% of the samples in the native lineage (Table 3).

The dominant allele phenotypes varied by haplotype lineage as well (Table 3). Common allele phenotypes in the native group were typically different from other lineages, as seen in loci *PaGT4*, *PaGT8*, *PaGT9*, *PaGT11*, *PaGT13*, *PaGT14*, *PaGT16*, *PaGT17*, *PaGT21* and *PaGT22* (Table 3). Furthermore, many of these alleles are essentially private alleles as they were not found in high frequencies in other populations. Sequence data also confirm base substitutions in several alleles that were found in native samples but not in others (loci *PaGT4*, *PaGT9*, *PaGT11*, *PaGT12*, *PaGT13*, *PaGT14*, Table 2). Gulf Coast samples were largely heterozygous and often amplified three alleles per sample indicating that they may be hexaploid.

Relatedness of haplotype lineages. D_{TL} distance estimates varied between 0.017 (introduced and European lineages) and 0.529 (native and Gulf Coast lineages). In comparison with other North American lineages, the native group seems to be most closely related to the introduced lineage (D_{TL} = 0.173). However, the UPGMA tree (Fig. 1) clearly groups the introduced and European lineages and links them most closely with samples collected outside of North America and Europe (Other category), whereas the native lineage is clearly distinguished from all others. The Gulf Coast lineage is most closely related to the introduced lineage (D_{TL} = 0.109).

Six locus analyses

Genetic diversity. All haplotype lineages showed significant departure ($P < 0.001$) from HWE values under the diploid model. The exact test for nonrandom association of genotypes across loci gave significant values ($P < 0.001$) between all lineages except for locus *PaGT13* where the European and introduced groups were not significantly different.

Structuring among regional population groups. Similar allele phenotypes were found between historic and modern samples in the four regional groups (Atlantic Coast, Midwest, West, and Gulf Coast) defined by cpDNA haplotypes (Saltonstall 2003, Table 4). The most common allele phenotypes in the native lineage were found in all three geographical regions at most loci, although their frequencies changed between the two time periods studied here due to the prevalence of introduced haplotype M individuals in modern samples. Unique allele phenotypes were also found in each region. Historic Atlantic Coast samples showed an allele of 274 bp at locus *PaGT4* found both as a homozygote and in combination with the 266 bp allele. The 274 bp allele was not found in the Midwestern group and in only one heterozygous western sample.

Table 4 Patterns of regional structuring in genetic characteristics between historical and modern North American populations of *Phragmites* at six microsatellite loci. Historic Native = populations collected before 1910 showing a Native cpDNA haplotype; Modern Native = populations collected after 1990 showing a Native cpDNA haplotype; Introduced = populations collected after 1990 showing the Introduced haplotype M cpDNA; Historic and Modern samples in the Gulf Coast region have cpDNA haplotype I. *n*: number of samples genotyped; A_o: number of alleles present at the locus; DP: dominant allele phenotypes in the population group, numbers in parentheses are the frequencies of each phenotype within the time period in which the samples were collected

	Regional Population											
	Atlantic Coast			Midwest			West			Gulf Coast		
	Historic Native	Modern Native	Introduced	Historic Native	Modern Native	Introduced	Historic Native	Modern Native	Introduced	Historic GC	Modern GC	Introduced
PaGT4												
n	28	117	106	12	50	25	25	24	8	5	16	3
A_o	4	8	7	2	4	3	1	3	3	3	2	2
DP	266 (0.18) 274 (0.43) 266/274 (0.25)	266 (0.03) 274 (0.04) 266/274 (0.02)	276 (0.36) 274/276 (0.33)	266 (0.92)	266 (0.48) 276 (0.02)	274/276 (0.12)	266 (1.00)	266 (0.66) 266/274 (0.03) 266/280 (0.03) 274/280 (0.03)	276/278 (0.13)	274/276/ 280 (1.00)	274 (0.05) 274/276 (0.05) 274/276/ 280 (0.68)	274 (0.11)
PaGT8												
n	18	101	90	12	48	26	15	25	7	6	20	4
A_o	3	2	2	1	2	2	3	1	1	2	2	1
DP	176 (0.11) 178 (0.22) 180 (0.56)	176 (0.02) 178 (0.04) 180 (0.03)	176 (0.68) 176/178 (0.19)	178 (1.00)	176 (0.02)	176 (0.40) 176/178 (0.10)	170 (0.27) 178 (0.60)	170 (0.08) 178 (0.60)	176 (0.28)	176/189 (1.00)	176 (0.05) 176/189 (0.75)	176 (0.20)
PaGT9												
n	19	96	85	10	46	23	21	26	7	6	18	3
A_o	7	9	8	1	9	7	4	5	2	2	5	4
DP	210 (0.21) 210/212 (0.47)	210 (0.05) 210/212 (0.01)	198 (0.29) 198/202 (0.25)	210 (1.00)	210 (0.43)	198 (0.15) 198/202 (0.11)	210 (0.48) 210/212 (0.38)	210 (0.54) 210/212 (0.12)	198 (0.23)	192/196 (1.00)	192/196 (0.78)	192/204 (0.11)

Table 4 *Continued*

	Regional Population											
	Atlantic Coast			Midwest			West			Gulf Coast		
	Historic Native	Modern Native	Introduced	Historic Native	Modern Native	Introduced	Historic Native	Modern Native	Introduced	Historic GC	Modern GC	Introduced
PaGT13												
n	19	114	103	9	47	22	14	27	7	5	19	3
A_o	4	2	5	2	4	2	5	4	2	2	2	1
DP	208 (0.74) 212 (0.11) 218 (0.11)	208 (0.04) 212 (0.00) 218 (0.04)	208 (0.50) 208/210 (0.30)	218 (0.78) 220 (0.11) 218/220 (0.11)	218 (0.38) 220 (0.02) 218/220 (0.02)	208 (0.21) 208/210 (0.23)	208 (0.79)	208 (0.30)	208 (0.19)	210 (0.40) 208/210 (0.60)	210 (0.21) 208/210 (0.63)	208 (0.16)
PaGT14												
n	31	107	96	12	50	24	21	31	9	7	18	3
A_o	5	9	7	2	8	8	1	7	4	5	6	4
DP	181 (0.81)	181 (0.04)	189 (0.25) 183/189 (0.17) 187/189 (0.10)	181 (0.92)	181 (0.38)	189 (0.16)	181 (1.00)	181 (0.58)	187 (0.10)	177/189 (0.43) 183/187 (0.14) 177/185/ 189 (0.43)	177/189 (0.44) 183/187 (0.00) 177/185/ 189 (0.22)	175 (0.06) 173/175 (0.06) 185/189 (0.06)
PaGT16												
n	16	99	88	12	53	28	19	31	9	6	19	2
A_o	4	3	4	3	4	3	5	4	3	3	3	3
DP	261 (0.31) 263 (0.13) 265 (0.19) 261/265 (0.13) 261/290 (0.13)	261 (0.02) 263 (0.00) 265 (0.08) 290 (0.00) 261/290 (0.01)	261 (0.33) 290 (0.24) 261/290 (0.26)	265 (0.25) 261/265 (0.50)	265 (0.23) 261/265 (0.09)	261 (0.19) 290 (0.13) 261/290 (0.17)	261 (0.16) 265 (0.21) 261/265 (0.53)	261 (0.13) 265 (0.48) 261/265 (0.00)	261 (0.10) 261/290 (0.10)	261/290 (0.17) 265/290 (0.83)	261/290 (0.05) 265/290 (0.63)	261 (0.05) 263/290 (0.05)

Table 5 Assignment tests in five chloroplast haplotype groups of *Phragmites* based on six microsatellite loci displaying one or two alleles per sample. Values are the proportion of individuals 'assigned' or 'not assigned' to any of the five haplotype lineage groups. 'Not assigned' samples had multilocus genotypes with a probability of belonging to any cpDNA haplotype group lower than 5%. Only samples collected in North America and Europe were considered. Introduced/Gulf Coast and Native/Introduced refer to samples that were assigned to more than one population. Bayesian method of classification

cpDNA haplotype lineage	Introduced/ European	Native	Gulf Coast	Introduced/ Gulf Coast	Native/ Introduced
Introduced (n = 139)					
Assigned	0.91*	0.00	0.00	0.04	0.00
Not Assigned	0.05				
Native (n = 118)					
Assigned	0.01	0.77	0.00	0.00	0.10
Not Assigned		0.12†			
Gulf Coast (n = 22)					
Assigned	0.00	0.00	0.91	0.09	0.00
Not Assigned			0.00		
Europe (n = 57)					
Assigned	0.86	0.00	0.00	0.00	0.02‡
Not Assigned	0.12				
Unknowns (n = 208)					
Assigned (n = 204)	0.78§	0.12	0.01	0.06	0.01¶

*Three of these samples had a < 10% probability of also being assigned to the Gulf Coast category.
†All 14 of these samples had a < 5% probability of being assigned to the native category and a 0% probability of being assigned to the introduced or Gulf Coast categories.
‡This sample had a 100% probability of belonging to the European category and a 18% probability of belonging to the native category. It was also missing data at two of the six loci analysed.
§Nine of these samples were also assigned to the Gulf Coast category at a probability < 10%.
¶One of these samples had a 100% probability of being assigned to the native category and 79% probability of being introduced, one had 100% probability of being native and 8% chance of being introduced.

Western samples had a unique allele of 170 bp at locus *PaGT8*. The Atlantic Coast and Western groups shared common allele phenotypes at loci *PaGT9* and *PaGT13* that were not common in Midwestern samples. The Gulf Coast lineage showed unique allele phenotypes at most loci.

Within the introduced lineage, the highest levels of diversity both in terms of the number of alleles and allele phenotypes expressed were found in Atlantic Coast samples (Table 4). Although sample sizes were much smaller in these other groups, Midwestern and Western introduced samples displayed a subset of these alleles and allele phenotypes across most loci. Gulf Coast introduced samples showed unique alleles at loci *PaGT9* and *PaGT14* (Table 4).

Assignment tests. In all haplotype lineages the majority of individuals were correctly assigned to their respective lineage in the assignment test using both Bayesian (76–91%; Table 5) and allele-sharing distance based ($D_{AS;}$ 81–91%; results not shown) methods of classification.

The 208 samples with unknown cpDNA haplotypes also had high levels of assignment to their putative lineage groups (2% not assigned for the Bayesian simulation and 6% not assigned in the distance calculation; Table 5). One hundred and fifty-nine samples were assigned to the introduced, 25 to the native and two to the Gulf Coast haplotype lineages. These assignments corresponded well with geographical patterns of genetic diversity seen in cpDNA based on sample collection locations. No unknown samples from the Atlantic Coast region were assigned to the native lineage. The majority of samples not assigned or assigned to more than one lineage were most likely to be introduced or native.

Assessment of hybridization between native and introduced lineages. No sample with a cpDNA haplotype belonging to the introduced lineage was misassigned to the native category. However, 12 native samples were assigned to both the native and introduced lineages and one sample was misassigned to the introduced lineage. All 13 of these ambiguous samples had missing data at two of the six loci used in this analysis. Further, 11 of them were historical samples from Connecticut and Massachusetts, many of which had a 274-bp allele at locus *PaGT4*. Although a 274-bp allele is also common in the introduced lineage (Table 3), sequencing of this allele revealed that in fact size homoplasy exists between the introduced and native lineages at this particular allele size with native samples

having a 274-bp allele that differs from the introduced 274 bp allele by three base substitutions (Table 2). Also, the 218 bp allele at locus *PaGT13*, which is diagnostic for native samples elsewhere across North America, is rare in historical Atlantic Coast native samples which commonly show a 208-bp allele at this locus, similar to introduced samples (Table 4). Although this could be evidence for the introduced lineage having bred with the native lineage, the 208 bp allele at *PaGT13* is also very common in both historical and modern samples from the West and therefore could represent an ancestral allele in the New England population as well. Thus, the ambiguity involving the assignment of these 13 native samples is likely due to a combination of missing data and shared allele sizes between haplotype lineages.

Discussion

Genetic diversity and population structure

Varying levels of intrapopulation diversity and high genetic differentiation between *Phragmites* lineages defined by cpDNA haplotypes was found. This structure reflects the genetic isolation between populations on different continents and effects of recent introduction of European *Phragmites* stock to North America.

Because *Phragmites* is a polyploid, the allelic dosage could not be determined. Therefore, allele frequencies and H_E values could not be calculated as they can in diploid species. H_O varied greatly across loci and between haplotype lineages, with the Gulf Coast lineage showing the highest levels and native the lowest. The introduced North American lineage showed similar levels to European samples (46 vs. 54%; Table 3). However, the variation observed across many of the loci appeared to be fixed within lineages which could be the result of selection at linked loci, epistatic interactions, or fixed heterozygosity as a consequence of allopolyploidy (Kreitman & Akashi 1995; Soltis & Soltis 2000). This is particularly apparent in the Gulf Coast lineage, where at some loci all or nearly all samples had identical heterozygous genotypes (*PaGT4*, *PaGT8*, *PaGT9*, *PaGT22*; Table 3).

In contrast, the native lineage had very high levels of homozygosity at most loci. There are several possible explanations for this heterozygote deficiency. First, many loci appear to be linked and there could be a biased transmission of alleles. Second, the presence of null alleles (alleles that do not amplify due to mutations in the primer region) must always be suspected for microsatellite loci (Lehmann *et al*. 1996). Problems associated with null alleles are typically greatest when using primers designed for other species (Pemberton *et al*. 1995) and the primers used in this study were all designed from *Phragmites*. However, these primers were designed from an introduced sample of *Phragmites* found along the Connecticut shoreline. It is possible that genetic differentiation between the native North American and introduced/European lineages is great enough that null alleles may be present in the native lineage, which showed the highest levels of homozygosity across loci.

Diversity within native and Gulf Coast lineages

Allelic variation was very similar among regional groupings of native samples collected across North America. In general, all samples had similar allele sizes and very high levels of homozygosity. Interestingly, native Atlantic Coast and Western samples were more similar to each other than to the native Midwestern samples (Table 4), However, all regional groups shared alleles at most loci. In addition, Western samples had more unique allele phenotypes than other groups. With an increased sample size, more differences might be seen.

Gulf Coast samples displayed unique alleles and allele phenotypes at most loci, some of which were nearly fixed in all samples (Table 3). This lineage is defined by a single cpDNA haplotype (haplotype I) which is also found in South America and Asia and is genetically distinct from the 11 native North American haplotypes and the introduced haplotype M (Saltonstall 2002). Although there were not enough haplotype I samples from other parts of the world to compare their allelic distribution with that of Gulf Coast populations, it is clear from these data that Gulf Coast populations are distinct from native North American populations at the level of nuclear DNA as well (Fig. 1).

Diversity within introduced populations

Highest allelic diversity was found in introduced samples collected along the Atlantic Coast with intermediate levels in the Midwest and lowest levels in the West. This is consistent with what would be expected of an invasion moving from the East to West coasts, with highest diversity found at the point(s) of introduction and loss of alleles with greater distance from these points due to genetic drift. An interesting observation, however, is in the introduced samples from the Mississippi delta in the Gulf Coast region where three alleles were found (192 and 212 bp at *PaGT9*, 173 bp at *PaGT14*) that were not found in other introduced samples. This may suggest an independent introduction of these populations. Pellegrin & Hauber (1999) found similar structure in their isozyme analysis of the same individuals in that they grouped independently from other cpDNA haplotype M samples from the Midwest and Atlantic Coast regions. However the sample size of introduced populations along the Gulf Coast was very small ($n = 4$) and further analysis of more

introduced samples from this region is needed to assess the possibility of an independent introduction event.

Levels of hybridization between populations

It appears that hybridization between the introduced and native lineages is occurring very rarely, if at all. Several alleles considered diagnostic for the native lineage were found in introduced samples but at very low frequencies (e.g. 210 bp at *PaGT9*, 145 bp at *PaGT11*, 218 bp at *PaGT13*, 181 bp at *PaGT14* and 191 bp at *PaGT22*, data not shown). However, all of these alleles were also found in European samples at low frequencies and may have arrived in North America via introduction rather than entering the introduced lineage via sexual reproduction with native individuals. Further, the majority of introduced samples have allele phenotypes not seen in native or Gulf Coast type samples at most loci (Tables 3 and 4). Thus it is extremely unlikely that hybrid individuals, possessing cpDNA haplotype M, are responsible for the rapid spread of the species seen over the past century in North America. Similarly, many alleles that are very common among introduced samples are also found in native samples (Table 3), although at much lower frequencies in most cases. Some of these common allele sizes have been shown to be a result of size homoplasy between the two lineages (see above). Others could have entered the native lineage via hybridization.

In contrast, it appears that Gulf Coast and introduced populations may have the potential to interbreed. Two samples showing haplotype I in the cpDNA analysis clustered with the introduced lineage in the assignment test (Table 5). Although both of these samples were missing data at two loci, they also lack the diagnostic allele phenotypes that appear to be fixed in the majority of Gulf Coast samples at several loci. Although sexual reproduction is considered to be extremely rare in Gulf Coast populations (Hauber *et al*. 1991; Pellegrin & Hauber 1999), it appears that these two samples may differ from others in the region due to some form of genetic recombination — either through loss of the diagnostic allele combinations from sexual reproduction with other Gulf Coast lineage individuals or from interbreeding with the introduced lineage. Introduced populations are, however, quite rare in this part of North America and those sampled in this study appear to contain novel alleles not found in other introduced individuals across the rest of the country.

The lack of hybridization between introduced and native North American lineages is curious. Although sexual reproduction is considered to occur very rarely in *Phragmites* in general, given the numbers of clones of each type present today in North America it seems plausible that interbreeding at some level should have occurred given that they are the same species and that the introduced lineage has been present for at least 150 years (Saltonstall 2002). It is possible that assortative mating, such as different flowering times of native and introduced populations, is preventing cross-pollination. Preliminary observations of several native populations along the Atlantic Coast indicate that these populations may mature and senesce earlier in the growing season than introduced populations (personal observation). This could create a temporal barrier to cross-pollination. However, additional work is needed to confirm that this pattern is seen annually and across a broader geographical area. Further, because establishment of new populations from seed is rare (Haslam 1972), it is possible that hybrid seeds occur but were not detected in this study due to the rarity of them forming new populations. Further studies involving controlled crosses of the native and introduced lineages are needed to fully assess whether there is an incompatibility between the two population lineages at the level of fertilization or whether viable seeds can indeed be produced between them.

Conclusions

Despite the limitations posed by working with samples of unknown ploidy levels, these microsatellite markers provided an effective way of comparing the genetic structure of North American *Phragmites* lineages at the nuclear level with previous results obtained from sequencing of cpDNA (Saltonstall 2002). Nearly all individuals were assigned to their correct haplotype lineage with as few as five microsatellite loci. Sequencing of alleles revealed size homoplasy and indel substitutions occurring at several loci. This points to the importance of understanding the underlying DNA sequence when using microsatellite markers in population genetic studies, as the assignment of alleles based solely on PCR product size could significantly bias estimates of genetic relatedness between populations and classification of individuals of unknown origin in the absence of a more detailed analysis of allele sequence data. In this particular case, the observed homoplasy was used to point out allelic differences in the ancestral state of native and introduced North American *Phragmites* lineages, further confirming the patterns seen in the genetic structure of cpDNA in the species.

What is clear from this analysis is that the genetic differentiation found at the nuclear level among introduced, native and Gulf Coast *Phragmites* lineages in North America correlates with the patterns seen in cpDNA (Saltonstall 2002, 2003). The native lineage is strongly differentiated from the other two and the introduced lineage is most closely related to European populations, which are the hypothesized source of this lineage in North America. Samples showing chloroplast haplotype M do not represent a hybrid between the native and introduced lineages. In parts of the country where the introduced lineage

predominates, such as along the Atlantic coast, the majority of native genetic variability has been lost since only remnant native populations remain. These native individuals also show little evidence for hybridization with the introduced lineage, thus it appears that native North American *Phragmites* stock is genetically pure.

Acknowledgements

I thank J.R. Powell, A. Caccone, L. Beheregaray, J.S. Hall and three anonymous reviewers for their helpful comments on the manuscript. Dr Jeffrey R. Powell kindly provided the laboratory space and facilities for this work. I thank the US National Herbarium, Harvard University Herbaria, New England Botanical Society, George Safford Torrey Herbarium at the University of Connecticut, CT Botanical Society and Yale University Herbarium for providing herbarium samples. This research was funded by an EPA STAR graduate fellowship, The Nature Conservancy CT chapter, The Long Island Sound Fund administered by the Connecticut Department of Environmental Protection through the sale of Long Island Sound license plates and contributions, the NJ Public Service, Electric and Gas Company and the US Fish and Wildlife Service through the Biological Control of Non-Indigenous Plant Species Program at Cornell University, and the NOAA Office of Sea Grant and Extramural Programs, U.S. Department of Commerce, under grant no. MERP/SG 99–21. The U.S. Government is authorized to produce and distribute reprints for governmental purpose notwithstanding any copyright notation that may appear hereon NJSG-02-498.

References

Anderson E, Stebbins GL (1954) Hybridization as an evolutionary stimulus. *Evolution*, **8**, 378–388.

Angers B, Bernatchez L (1997) Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution*, **14**, 230–238.

Anttila CK, Daehler CC, Rank NE, Strong DR (1998) Greater male fitness of a rare invader (*Spartina alterniflora*, Poaceae) threatens a common native (*Spartina foliosa*) with hybridization. *American Journal of Botany*, **85**, 1597–1601.

Becher SA, Steinmetz K, Weising K *et al.* (2000) Microsatellites for cultivar identification in *Pelargonium*. *Theoretical and Applied Genetics*, **101**, 643–651.

Chambers RM, Meyerson LA, Saltonstall K (1999) Expansion of *Phragmites australis* into tidal wetlands of North America. *Aquatic Botany*, **64**, 261–273.

Clevering OA, Lissner J (1999) Taxonomy, chromosome numbers, clonal diversity and population dynamics of *Phragmites australis*. *Aquatic Botany*, **64**, 185–208.

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

Dow BD, Ashley MV (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Molecular Ecology*, **5**, 615–627.

Doyle JJ, Dickson EE (1987) Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, **36**, 715–722.

Durand J, Garnier L, Dajoz I, Mousset S, Veuille M (2000) Gene flow in a facultative apomictic Poacea, the savanna grass *Hyparrhenia diplandra*. *Genetics*, **156**, 823–831.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques*, **20** (758), 760.

Ellstrand NC, Scheirenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the USA*, **97**, 7043–7050.

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology*, **11**, 1591–1604.

Estoup A, Tailliez C, Cornuet J-M, Solignac M (1995) Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution*, **12**, 1074–1084.

Felsenstein J (1995) PHYLIP *(Phylogeny Inference Package)*. Department of Genetics, University of Washington, Seattle. http://evolution.genetics.washington.edu/phylip.html.

Fischer D, Bachmann K (1998) Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *Biotechniques*, **24**, 796–802.

Guldbrandtsen B, Tomiuk J, Loeschcke V (2000) POPDIST, Version 1.1.1: a program to calculate population genetic distance and identity measures. *Journal of Heredity*, **91**, 178–179.

Haslam SM (1972) Biological flora of the British Isles: *Phragmites communis* Trin. *Journal of Ecology*, **60**, 585–610.

Hauber DP, White DA, Powers SP, DeFrancesch FR (1991) Isozyme variation and correspondence with unusual infrared reflectance patterns in *Phragmites australis* (Poaceae). *Plant Systematics and Evolution*, **178**, 1–8.

Huxel GR (1999) Rapid displacement of native species by invasive species: effects of hybridization. *Biological Conservation*, **89**, 143–152.

Kreitman M, Akashi H (1995) Molecular evidence for natural selection. *Annual Review of Ecology and Systematics*, **26**, 403–422.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386–391.

Lehmann T, Hawley WA, Collins FH (1996) An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics*, **144**, 1155–1163.

Lewontin RC, Birch LC (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution*, **20**, 315–336.

McClelland M, Hanish J, Nelson M, Patel Y (1988) KGB — A single buffer for all restriction endonucleases. *Nucleic Acids Research*, **16**, 364.

Morand ME, Brachet S, Rossignol P, Dufour J, Frascaria-Lacoste N (2002) A generalized heterozygote deficiency assessed with microsatellites in French common ash populations. *Molecular Ecology*, **11**, 377–385.

van Oppen MJH, Rico C, Turner GF, Hewitt GM (2000) Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in lake Malawi cichlids. *Molecular Biology and Evolution*, **17**, 489–498.

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, **12**, 357–358.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology and Evolution*, **15**, 1275–1287.

Pellegrin D, Hauber DP (1999) Isozyme variation among populations of the clonal species, *Phragmites australis* (Cav.) Trin. ex Steudel. *Aquatic Botany*, **63**, 241–259.

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249–252.

Raicu P, Staicu S, Stoian V, Roman T (1972) The *Phragmites communis* Trin. chromosome complement in the Danube delta. *Hydrobiologia*, **39**, 83–89.

Raymond M, Rousset F (1995) GENEPOP (v. 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Röder MS, Plashke J, König SU *et al.* (1995) Abundance, variability, and chromosomal location of microsatellites in wheat. *Molecular and General Genetics*, **246**, 327–333.

Saltonstall K (2002) Cryptic invasion by a non-native genotype of *Phragmites australis* into North America. *Proceedings of the National Academy of Sciences of the USA*, **99**, 2445–2449.

Saltonstall K (2003) Genetic variation among North American populations of *Phragmites australis*: implications for management. *Estuaries*, in press.

Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the USA*, **97**, 7051–7057.

Tomiuk J, Loeschcke V (1991) A new measure of genetic identity between populations of sexual and asexual species. *Evolution*, **45**, 1685–1694.

Tomiuk J, Loeschcke V (1996) A maximum-likelihood estimator of the genetic identity between polyploid species. *Journal of Theoretical Biology*, **179**, 51–54.

Tucker GC (1990) The genera of Arundinoideae (Graminae) in the southeastern United States. *Journal of the Arnold Arboretum*, **71**, 145–177.

Viard F, Franck P, Dubois M-P, Estoup A, Jarne P (1998) Variation of microsatellite size homoplasy across electromorphs, loci, and populations in three invertebrate species. *Journal of Molecular Evolution*, **47**, 42–51.

White GM, Boshier DH, Powell W (2002) Increased pollen flow counteracts fragmentation in a tropical dry forest: an example from *Swietenia humilis* Zuccarini. *Proceedings of the National Academy of Sciences of the USA*, **99**, 2038–2042.

Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Quantifying threats to imperiled species in the United States. *Bioscience*, **48**, 607–615.

This study is part of the author's doctoral thesis. She is interested in the ecology of invasive species and the application of molecular markers in conservation biology.
