American Fern Journal 101(4):213–230 (2011)

Broad-Scale Integrity and Local Divergence in the Fiddlehead Fern *Matteuccia struthiopteris* (L.) Todaro (Onocleaceae)

DANIEL M. KOENEMANN
Rancho Santa Ana Botanic Garden, 1500 North College Avenue, Claremont, CA 91711-3157, e-mail: daniel.koenemann@gmail.com

JACQUELINE A. MAISONPIERRE
University of Vermont, Department of Plant Biology, Jeffords Hall, 63 Carrigan Drive, Burlington, VT 05405, e-mail: jmaisonp@gmail.com

DAVID S. BARRINGTON
University of Vermont, Department of Plant Biology, Jeffords Hall, 63 Carrigan Drive, Burlington, VT 05405, e-mail: dbarring@uvm.edu

ABSTRACT.—*Matteuccia struthiopteris* (Onocleaceae) has a present-day distribution across much of the north-temperate and boreal regions of the world. Much of its current North American and European distribution was covered in ice or uninhabitable tundra during the Pleistocene. Here we use DNA sequences and AFLP data to investigate the genetic variation of the fiddlehead fern at two geographic scales to infer the historical biogeography of the species. *Matteuccia struthiopteris* segregates globally into minimally divergent (0.3%) Eurasian and American lineages. These two clades have little to no variation even at large geographic scales. Within hemisphere, patterned genetic variation was evident only in the AFLP data and only locally. Genetic variation within Vermont was greater within the westward-trending Winooski River watershed than in the Passumpsic River watershed, which drains east into the Connecticut River. We suggest that historical factors have created this pattern; a Mississippi Valley Pleistocene refugium for the American lineage of the species seems plausible.

KEY WORDS.—*Matteuccia*, Onocleaceae, Pleistocene, Biogeography, AFLP, *PgiC*

Matteuccia struthiopteris (L.) Todaro (Onocleaceae) has a present-day distribution across much of the north-temperate and boreal regions of the world. It is commonly known as the fiddlehead fern in the northeast of North America because of the resemblance of its elegant croziers to their namesake (Kato, 1993). *Matteuccia struthiopteris* is found most commonly on river banks and flood plains. However, the species can also be found in uplands anywhere there are rich alluvial soils (Smith, 1993). Though robust in full sun, we have found that the ferns are more common in either full or partial shade, sheltered by broadleaf deciduous trees. They often form large colonies on wooded floodplains via prolific spread by stolons. The fronds are dimorphic; large (0.3 to greater than 1.5 m) sterile fronds produced in the spring are followed by smaller (0.3 to 0.7 m) fertile fronds produced in late summer, which have their sporangia rolled up inside of the pinnules (Prange and von Aderkas, 1985). The sterile fronds are oblanceolate, tapering gradually at the base and abruptly at the tip; the widest part of the frond is subapical (Prange and von Aderkas, 1985). Sterile fronds present two morphological variants that we call *downy* and *smooth*. In the downy variant the petioles of young fronds are covered with a minute (less than 0.5 mm),

colorless, non-laminar indument. In the smooth variant the petiole lacks this indument. This downy indument dissipates in early summer, so its presence cannot be scored later in the growing season.

Matteuccia struthiopteris is distributed throughout much of the northern hemisphere in boreal deciduous forests (Lloyd, 1971); it is collected as a seasonal edible green in New England and the eastern provinces of Canada (von Aderkas, 1984), as well as in Japan (Miyazawa *et al.*, 2007). Fiddlehead fern collection is documented for several Native American groups, including primarily the Abenaki Indians of New England and Malecite of New Brunswick. European colonists learned of this harvest practice from these native groups on their arrival in North America and collected fiddleheads as part of a subsistence diet (von Aderkas, 1984). In Vermont, we have seen a recent explosion of interest in buying and preparing fiddleheads with the rapid expansion of the local-foods movement. In the last three years, landowners have begun to post their lands with no-harvesting signs for the first time.

Little is known about the impact of fiddlehead harvesters on the demography or genetic structure of *Matteuccia struthiopteris*, as their activity is not regulated. However, it is known that the species' health can be significantly affected by the removal of young fronds. Bergerson and Lapointe (2001) conducted a five-year study in Québec, Canada, investigating the impacts of harvesting on *Matteuccia struthiopteris*. They found that harvesters should collect no more than half the croziers produced by a shoot, as removal of all shoots in a single year impaired carbohydrate accumulation and crozier production in subsequent years.

With the advent of molecular techniques, there has been substantial improvement in our understanding of the geographic distribution of genetic variation in plant populations. In a historical context, these distributional patterns are now commonly used to interpret the recent distributional history of species (Donoghue *et al.*, 2001; Hewitt, 2000). Analyses of these patterns make the basic assumption, an assumption that we will maintain, that populations near Pleistocene refugia, having remained established for the longest period of time, will display greater genetic diversity than more peripheral populations. This concept was first proposed in the stepping-stone model of population structure developed by Kimura and Weiss (1964). Building on this idea, dispersal models constructed by Ibrahim *et al.* (1996) and Hewitt (1996) predicted that as a species expands from its refugium the newly established populations contain less diversity than the parent population. The ideas of Kimura and Weiss (1964), Ibrahim *et al.* (1996), and Hewitt (1996) were confirmed in subsequent studies by Dufresne and Hebert (1997) and most of those reviewed by Barrington and Paris (2007).

On the other hand, genetic variation in populations relates to local evolutionary history (both selection and drift) and factors relating to population size (large central and small peripheral populations can evidence different patterns of genetic diversity). In northeastern North America for instance, the distribution of genetic diversity in *Cypripedium parviflorum* Salisb. appears to be the result of genetic drift related to differences in population size and not historical factors (Wallace and Case, 2000). Given the great differences in population size of

Matteuccia with elevation (very large in alluvial forests low in watersheds and small and isolated in wetlands at high elevations), history and recent evolution must both be considered.

Our interest in the historical biogeography of *Matteuccia struthiopteris* led us to assess genetic diversity of the fiddlehead fern at two scales. First, we used AFLP analysis to document the distribution of genetic diversity within an array of fiddlehead fern to assess variation within and among populations in Vermont. We expected one of two patterns to emerge: 1) concentration of variation in either large lowland or small highland populations implying recent drift or selection or 2) a regional pattern of decreasing diversity implying a historical cause. We also explored the pattern of variation in the downy versus smooth variants in these populations with similar expectations about pattern. Second, we set this local variation in the context of a less detailed view of the global genetic structure for the fiddlehead fern based on chloroplast and nuclear DNA sequences. In this study we sought to assess potential human impact on genetic diversity, as harvesting of these ferns is now intensive in the large populations in lower-elevation river valleys of the region.

Materials and Methods

Taxon sampling.—A total of 83 accessions representing four onocleaceous taxa was included in this study. All but three of these accessions were *Matteuccia struthiopteris*. Unless noted otherwise, at each site samples were collected as follows. Up to ten complete fronds were collected, each from a different shoot. In an effort to avoid collecting multiple ramets of single genets, each frond was taken at a minimum distance of 2 m from any other sample. The maximum distance was 5 m from any other sample. Downy individuals were deliberately included when encountered in a population, and each population was scored as downy abundant, rare, or absent. Material collected from each individual included a sample for DNA analysis stored in silica gel at −80°C and a dried voucher specimen, both from the same frond.

A total of 60 individuals of *Matteuccia struthiopteris* was collected across three watersheds in Vermont: the Passumpsic, Mettawee, and Winooski rivers (Fig. 1). The Passumpsic River is a tributary of the Connecticut River, located entirely within the Northeast Kingdom of Vermont. Three collections were made along the Passumpsic, the highest in elevation in Newark at 358 m, the middle in East Burke at 253 m, and the lowest in Barnet at 156 m (See Appendix 1 for accession data and voucher information).

The Mettawee River is the shortest of the three river systems from which plants were collected. One collection was made along the Mettawee River in Granville, NY at 173 m (Appendix 1).

The Winooski River is the largest of the three watersheds surveyed. The river drains an area of the northern section of the Green Mountains; it is a large tributary to Lake Champlain. Three collections were made along the river. The highest in elevation is in Cabot at 442 m, the next in Plainfield at 238 m, and the lowest in Richmond at 87 m (Appendix 1).

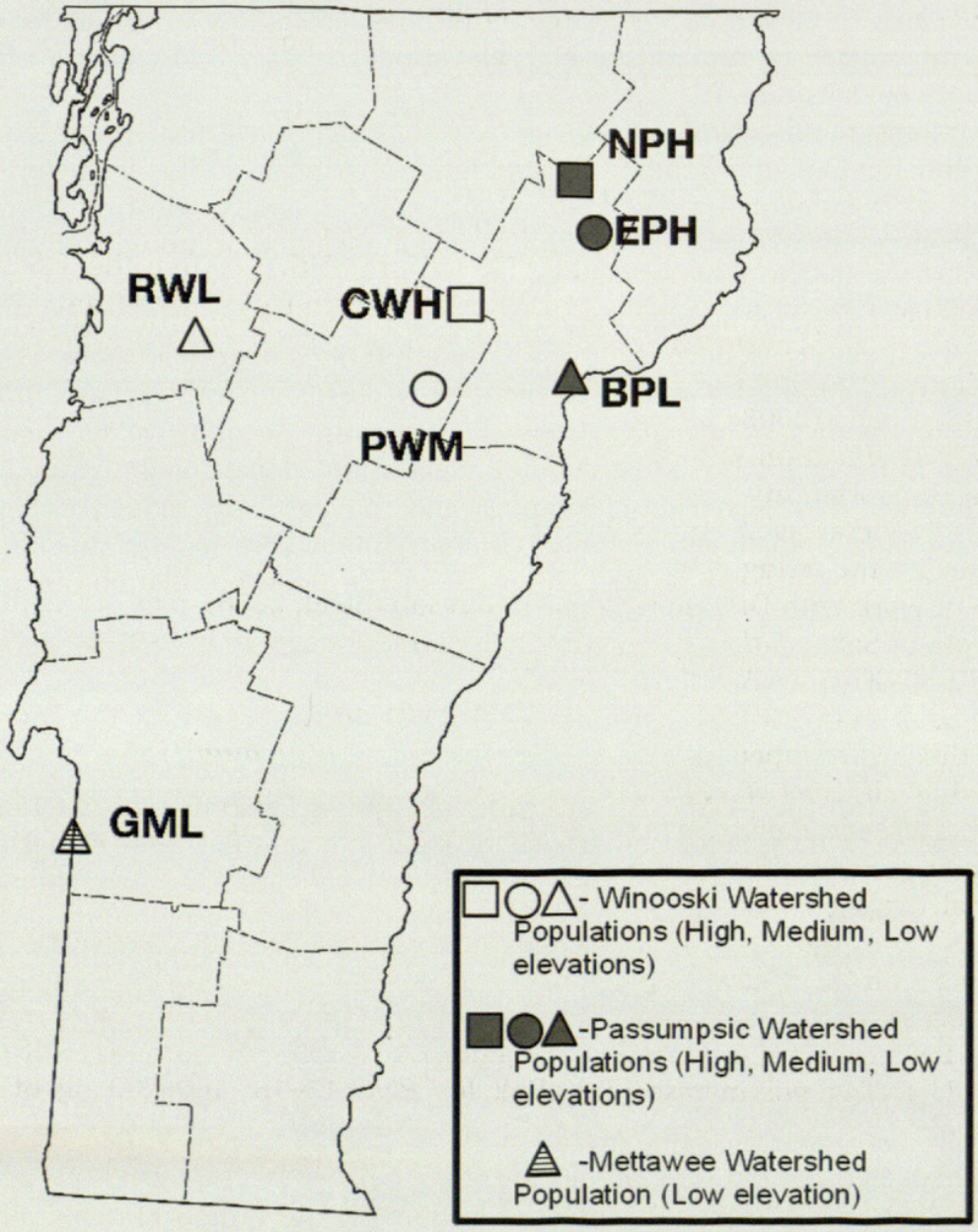

FIG. 1. Source of Vermont *Matteuccia struthiopteris* accessions for the AFLP study.

The remaining 23 accessions comprised two species of the *Matteuccia* segregate genus *Pentarhizidium* (*P. intermedium* (C.Chr.) Hayata and *P. orientale* (Hook.) Hayata), the out-group taxon *Onoclea sensibilis* L., and 20 additional *Matteuccia struthiopteris* accessions including 15 from the New World (Vermont, Maine, Quebec, New Brunswick, Michigan, British Columbia, and Alaska) and four from the Old World (China, Japan, and Sweden). Accessions from China and Sweden were provided by correspondents without vouchers or information on collection method (Appendix 1).

DNA extraction.—Total DNA was extracted from approximately 0.05 g of frond tissue following Dempster *et al.* (1999), which is based on the CTAB protocol of

Doyle and Doyle (1987). The following slight modifications were made in order to acquire samples of the high quality and concentration required for the AFLP technique. Homogenized frond material was allowed to incubate at 55°C for 24 hours in CTAB buffer. After the addition of isopropanol the nucleic acids were allowed to precipitate for one hour at −80°C. DNA was pelleted using centrifugation, then cleaned once in a 70% ethanol wash and once in a 95% ethanol wash. The pellets were dried via tube inversion or Speedvac (Genevac Inc, Gardiner, New York, USA). After the DNA was resuspended in 50 μL of ddH_2O the concentration was quantified using a NanoDrop1000 (Thermo Scientific, Waltham, Massachusetts, USA). Samples with concentrations lower than 100 ng/μL were discarded and re-extracted following Sigel (2008).

PCR amplification.—Screening of seven chloroplast markers and one nuclear marker for variability yielded two variable markers for sequencing, the highly labile cp marker psbA-trnH spacer (*psbA-trnH*, see Kress *et al.*, 2005) and the *PgiC* region including most of intron 15 and a portion of exon 16 (*PgiC* 15–16), labile in work with Dryopteridaceae in our lab (Sigel, 2008). *psbA-trnH* primers are those of Sang *et al.* (1997) with modification for use on ferns (psbA-trnHFor: 5′ GTTATGCATGAACGTAATGCTC 3′, psbA-trnHRev: 5′ CGCGCATGGTG-GATTCACAATCC 3′). The primers for *PgiC* 15–16 were modified for this study from those developed in our lab for use on *Polystichum* (Dryopteridaceae), following Ishikawa *et al.* (2002). They are PgiC15FM (5′ TTTGCTCCTCACATT-CAACA 3′), and PgiC16RM (5′ GTTGTCCATTAGTTCCAGGTTCCCC 3′). These regions were amplified by the polymerase chain reaction (PCR) method with all thermal cycling done using either a model TC-312 or T-3000 thermal cycler (Techne, Burlington, New Jersey, USA). Reactions were completed in 24 μL aliquots with the following reaction components: for psbA-TrnH: 100–200 ng of DNA, 0.1 μmol/L of each primer, 1× ExTaq Buffer (TaKaRa, Madison, Wisconsin, USA), 200 μmol/L of each dNTP, 200 μmol/L of $MgCl_2$, 200 μmol/L of BSA, and 0.625 U ExTaq polymerase (TaKaRa); for *PgiC* 15–16: 100–200 ng of DNA, 0.1 μmol/L of each primer, 1× ExTaq Buffer, 400 μmol/L of each dNTP, 400 μmol/L of BSA, and 0.625 U Ex Taq polymerase. The reaction conditions were as follows for *psbA-trnH*: initial denaturation was for 1 min at 95°C; followed by 30 cycles of 1 min at 95°C, 1 min at 50°C, and 4 min at 65°C; with a final extension of 72°C for 5 minutes. For *PgiC* 15–16: initial denaturation was for 3 min at 95°C; followed by three cycles of 1 min at 94°C, 1 min at 56°C, and 2 min at 72°C; followed by three cycles of 1 min at 94°C, 1 min at 53°C, and 2 min at 72°C; followed by 34 cycles of 45 sec at 94°C, 45 sec at 50°C, and 90 sec at 72°C; with a final extension of 72°C for 8 min. Before sequencing the PCR products were purified using ExoSAP-IT (USB Corp., Cleveland, Ohio, USA).

Agarose gel electrophoresis and extraction.—Prior to sequencing, all amplified samples were visualized on 2% agarose gels with ethidium bromide (1.0 g agarose, 50 mL Tris Borate EDTA buffer, 2.5 μL ethidium bromide). If two or more bands were separated during electrophoresis, the band of the appropriate marker length was excised and the amplified DNA was extracted using the QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, California, USA). Before

sequencing the extraction products were purified using ExoSAP-IT (USB Corp., Cleveland, Ohio, USA).

Sequencing.—Direct sequencing of both markers was undertaken using the ABI BigDye Terminator Cycle Sequence Ready Reaction Kit v. 31 (Perkin-Elmer/Applied Biosystems, Foster City, California, USA). The kit employs the following thermal-cycling protocol: initial denaturation was for 5 min at 80°C followed by 30 cycles of 10 sec at 96°C and five sec at 50°C with a final extension at 50°C for 4 min. An ABI Prism 3130x1 automated sequencer was used to resolve the sequencing products (Vermont Cancer Center DNA Analysis Facility, Burlington, Vermont USA).

Sequence alignment.—Raw forward and reverse sequences were assembled for each sample and ambiguous bases were corrected from inspection of the chromatograms; Sequencher v. 3.1.1 (Nishimura, 2000) and BioEdit v. 7.0.9 (Hall, 1999) were used for sequence editing and consensus-sequence construction. Consensus sequences were first aligned using ClustalX (Larkin *et al.*, 2007), then manually improved using MacClade v.4.08 (Maddison and Maddison, 2005).

Phylogenetic analysis.—Separate analyses were conducted for each data set that was generated (*psbA-trnH, PgiC* 15–16) and the two combined using Maximum Parsimony. For the analysis we used PAUP* version 4.0b10 (Swofford, 2001) with all characters treated as unordered with equal weights. A heuristic search was performed using 10,000 replicates and random taxon addition with ten trees held at the tree-bisection-reconnection (TBR) branch swapping step with each sequence addition. A maximum of ten trees was saved at each step, MulTrees option on, with ACCTRAN character-state optimization. Bootstrapping was performed for 1000 replicates using simple taxon addition, TBR branch swapping, and the MulTrees option on. For this analysis, we retained a single representative accession for sets of identical accessions (see Fig. 2 for details). We considered bootstrap percentages greater than 68 and 95 to be moderate and strong levels of clade support respectively following Driscoll and Barrington (2007). All trees were rooted using *Onoclea sensibilis* as the outgroup. We report only the combined analysis: separate analyses were congruent with but less resolved than the combined analysis.

AFLP protocol.—The AFLP protocol was adapted from Vos *et al.* (1995) with additional modifications as suggested by the Wolf Lab at Utah University (Wolf, 2000), the Gastony Lab at Indiana University (Nakazato *et al.*, 2006), previous studies in the Barrington Lab (Sigel, 2008), as well as modifications specific to this study. Several procedures were performed to minimize genotyping error in the final data. A previously typed sample was included in each round of amplification for comparison between different rounds. Ten percent of the samples were replicated between two and four times to assess the error within samples. Those samples with manifest errors in amplification, such as greatly reduced allele numbers, were discarded and re-amplified following Sigel (2008).

Restriction of the genomic DNA was performed using the restriction endonucleases *Eco*RI and *Mse*I (New England Biolabs, Ipswich, Massachusetts, USA). Ligation of the sticky end-adaptors (Invitrogen, Carlsbad, California, USA) to the restriction sites was also achieved in this initial reaction following the Wolf

(2000) protocol. Modifications from the Wolf Lab protocol (Wolf, 2000) are as follows: 20 μL of the enzyme master mix were used to avoid pipetting volumes under 0.4 μL. The master mix had the following components: 0.02 μL of *Mse*I, 0.05 μL of *Eco*RI, 0.025 μL of T4 Ligase, and 0.655 μL ddH_2O were used per 1 μL of total master-mix product. Each reaction volume totaled 11 μL, each containing 100 ng of DNA suspended in 5.5 μL ddH_2O. Samples were incubated for two hours at 37°C in a thermal cycler. In preparation for pre-selective amplification 3 μL of the restriction/ligation product were diluted with 24.5 μL TE 0.1 buffer.

In the first round of PCR, preselective amplification was achieved with primers that complement the *Eco*RI and *Mse*I adaptors plus one additional nucleotide i.e., *Mse*I+C (5′ GAT GAG TCC TGA GTA AC 3′) and *Eco*RI+A (5′ GAC TGC GTA CCA ATT CA 3′). Each preselective reaction totaled 25 μL, comprising the following: 3 μL diluted restriction/ligation product, 2.5 μL ExTaq Buffer (TaKaRa), 0.1 μL Ex Taq DNA polymerase (TaKaRa), 0.75 μL of 50mM $MgCl_2$, 1.2 μL of 2.5mM dNTPs, 16.45 μL ddH_2O, and 0.5 μL of a 10 μM solution of each of the primers. Reaction samples were initially denatured for 2 min at 72°C, followed by 30 cycles of: 94°C for 30 sec, 56°C for 30 sec and 72°C for 2 min, ending with 30 min at 6°C. Fifteen and a half μL of the preselective amplification product were diluted with 100 μL of TE 0.1 buffer. The quality of the undiluted preselective and restriction/ligation product was visualized on a 1.5% TBE-agarose gel.

In the second round of PCR, selective amplification was achieved using the *Mse*I+CAG and *Eco*RI+AAC (fluorophore NED [yellow] primers); each selective primer has an identical sequence to its corresponding preselective primer but with an additional two bases. The *Eco*RI primer is fluorescently tagged for visualization on the ABI 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, California, USA). Selective reactions were set up in 18 μL aliquots comprising the following: 3.6 μL of the diluted preselective DNA product, 1.8 μL Ex Taq Buffer, 1.8 μL Ex Taq DNA polymerase, 0.5 μL $MgCl_2$, 0.864 μL dNTPs, 1.44 μL of the 0.4mM *Mse*I selective primer (Invitrogen), 3.66 μL of the 0.08mM *Eco*RI selective primer (Invitrogen), 0.144 μL BSA, and 5.872 μL ddH_2O. Reactions were initially denatured at 94°C for 2 min, followed by 13 cycles of 94°C for 30 sec, 65°C for 30 sec (reduced by 0.7°C per cycle), 72°C for 2 min, and 24 cycles of 94°C for 30 min, 56°C for 30 sec, 72°C for 2 min, with a final hold at 72°C for 30 min. Selective amplification products were run on the 4-capillary ABI 3100-Avant Genetic Analyzer at the Vermont Cancer Center DNA Analysis Facility at the University of Vermont.

ALFP data scoring.—The most challenging aspects of AFLP analysis are data scoring and analysis (Bonin *et al.*, 2004; Bonin *et al.*, 2007; Meudt and Clarke, 2007; Pompanon *et al.*, 2005; Vekemans *et al.*, 2002; Whitlock *et al.*, 2008). In principle, the challenge to AFLP data scoring is producing a set of binary phenotypes (termed *loci*) for each individual from the presence/absence of the DNA fragments retrieved, while at the same time excluding experimental artifacts. We used the loci exclusion thresholds and phenotype exclusion thresholds outlined in Whitlock *et al.* (2008) to reduce the likelihood of scoring artifacts while maintaining a maximal number of informative markers. Appendix

B of Sigel (2008) provides a detailed account of AFLP analysis problems and solutions in our lab.

Initial AFLP fingerprints were determined with GeneMapper v. 3.7 (Applied Biosystems). Peak-height data for each individual were exported in tab-delineated form to an Excel worksheet. The top ten percent of the peak heights was trimmed to remove false flares in the recording instruments. The mean of the trimmed data was calculated for each locus across all individuals as well as for each individual across all its loci. These means were used as a basis for defining an array of candidate loci-exclusion and phenotype-calling thresholds. Using the trimmed mean rather than an arbitrary value of 100 relative fluorescence units (rfu) as suggested by GeneMapper (Applied Biosystems) is more accurate as it better represents the specific data set under study (Sigel, 2008). We generated 20 binary datasets with an array of loci-exclusion and phenotype-calling threshold combinations. The mismatch error rate (which represents the number of individuals with inconsistent band data in replicate amplifications) was calculated for each generated dataset. The optimum threshold was determined by selecting the generated dataset that maintained the maximum number of informative loci while retaining a mismatch error rate below five percent (Bonin *et al.*, 2004). Binary phenotype datasets generated using the identified optimum thresholds were used in the data analysis.

Data analysis and interpretation.—The binary dataset with the maximum number of informative loci with an error rate below five percent was used as a basis for data analyses. GENALEX version 6.2 (Peakall and Smouse, 2006) was used for all statistical analyses of the data. A pairwise, individual-by-individual genetic-distance matrix was generated. Each value in the matrix is a tally of differences between two genetic profiles (Peakall and Smouse, 2006, Appendix 1). As an initial probe of the data the genetic-distance matrix was used to do a principal-components analysis (PCA). We explored the PCA by visualizing individual plants on plots of pairs of principal components accounting for substantial variance in the data to search for clusters of individuals according to population, watershed, and elevation.

To assess patterns of genetic diversity across the seven Vermont populations and elevations the AFLP data were subjected to three calculations: the average genetic distance between individuals, the average heterozygosity, and the total number of loci present.

Results

Field observations.—In Vermont the downy trait was absent in the high-elevation population in the Winooski watershed and in both the high and middle populations in the Passumpsic watersheds. The trait was rare in the middle-elevation population in the Winooski watershed. Both watersheds had an abundance of downy individuals in the low-elevation population. The downy character trait was not scorable for individuals in the Mettawee population, as it was collected too late in the season. In the Maritimes, three low-elevation New Brunswick populations (the one near Campbellton New Brunswick, the one at the

intersection of Route 2 Canada and the Canaan River, and the one just west of Fredericton) included downy plants.

Sequence characteristics.—The aligned and concatenated sequence for *psbA-trnH* and *PgiC* 15–16 yielded a matrix of 798 characters, including five indels. *psbA-trnH* comprised 439 of these characters including two of the indels. Within *Matteuccia,* 59 characters (7.4%) were variable, of which 36 including four indels (4.5%) were parsimony informative. Twenty characters separated *Onoclea* (the outgroup) from the study group. In *PgiC,* direct sequencing yielded largely unambiguous sequences, presumably because most plants were homozygous for all nucleotides. An informative exception was our sequence for *Onoclea,* which contained an extensive region of double peaks at the 3′ end. The pattern of double peaks in this region was consistent with the interpretation that they were the result of the insertion of one nucleotide in one of the two homologs; for analysis we retained the allele that was identical to the rest of the study set.

Phylogenetic analysis.—The phylogenetic analysis conducted using Maximum Parsimony (MP) yielded one shortest tree consisting of 53 steps. The analysis (Fig. 2) retrieved a strongly supported clade (BS = 100) comprising the Asian endemics *Pentarhizidium intermedium* and *P. orientale* to be sister to a monophyletic (BS = 100) *M. struthiopteris.* Within our study species, Eurasian and American *M. struthiopteris* clades were moderately well supported (BS value for each is 86). Within each of the two *M. struthiopteris* clades all accessions were unresolved, with the exception of two of the accessions from eastern Canada, which were sister to each other (BS=100). An additional analysis, including the Alaska accession that was received while this contribution was in review, yielded similar results, but with the Old World clade bootstrap support reduced to 66% and a new clade comprising the Old World and Alaska retrieved with 65% bootstrap support.

Looking at individual characters for *Matteuccia struthiopteris,* New World and Eurasian accessions differed by three substitutions (0.3%), two of them synapomorphies for the Old World plants and one for the New World plants. With the addition of the Alaska accession, there was one synapomorphy unique to the Old World plants, one shared by the Old World and Alaska accessions, and one by the New World plants excluding the Alaska accession. The nearly complete lack of resolution within *M. struthiopteris* is due to absence of synapomorphies rather than character incongruence.

AFLP analysis.—AFLP analysis of the sixty *Matteuccia struthiopteris* samples using the *Mse*I+CAG *Eco*RI+AAG primer pair resulted in 254 distinguishable loci. A phenotype-calling threshold of 75% (92.51 rfu) and loci-exclusion threshold of 85% (104.85 rfu) were used to construct the binary data matrix from the raw peak-height data. This combination of thresholds applied to the raw data resulted in a 4.98% mismatch error rate between replicate samples.

The percent polymorphic loci across all populations was 86.22 while the percent polymorphic loci within the individual populations ranged from 33.36 in the Barnet population (middle elevation population in the Passumpsic watershed) to 61.81 in the Plainfield population (middle elevation population in the Winooski watershed, Table 1).

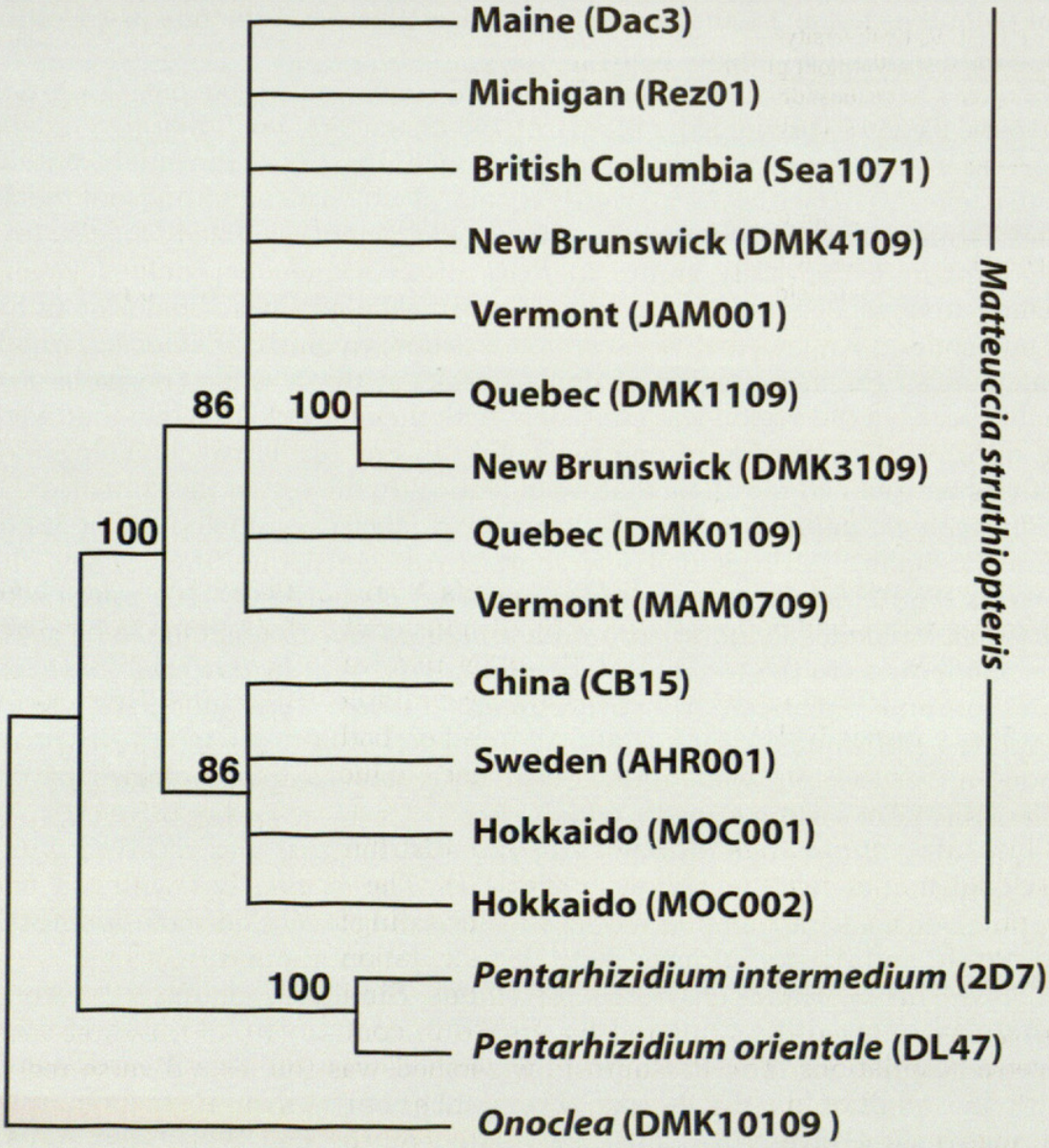

FIG. 2. Bootstrap analysis of *Matteuccia struthiopteris* and allies based on *psbA-trnH* and *PgiC* 15–16 combined, with *Onoclea sensibilis* as the outgroup. Numbers above common ancestors are bootstrap percentages. See methods for bootstrap analysis conditions. Identical sequences not included in this analysis were 1) DSB 2323, (same as Dac3, Rez01, Sea1071), 2) DMK 5109, 7109, 8109, and DMK9109 (same as DMK 4109 and JAM001), and 3) DMK2109 and 6109 (same as DMK 0109).

The principal-components analysis revealed substantial geographic clustering of genetically related individuals. The first three components retrieved from the analysis accounted for 75.52% of the variance; they represented 32.72%, 29.95%, and 12.85% of the total variance in the data respectively. Principal components 1 and 3 were most powerful in portraying the relationships among individuals and populations (Fig. 3). Most of the Passumpsic-watershed plants lay in a tight cluster; they were largely separated from the Winooski-watershed and Mettawee-watershed plants on principal component 1. In contrast the Winooski-watershed

TABLE 1. Genetic diversity calculations (number of loci, mean genetic distance, mean heterozygosity) for each of seven Vermont populations of *Matteuccia struthiopteris*. Rank is categorical order of each population; 1 is the most diverse and 7 the least diverse. Within watershed, the populations are listed from highest to lowest in elevation.

Watershed	Population	Number of Loci/ (Rank)	Mean Genetic Distance/ (Rank)	Mean Heterozygosity/ (Rank)
Winooski	Cabot (CWH)	139 (2)	50.944 (2)	0.152 (1)
	Plainfield (PWM)	160 (1)	55.222 (1)	0.145 (2)
	Richmond (RWL)	132 (3)	46.464 (3)	0.107 (5)
Passumpsic	Newark (NPH)	95 (6)	36.393 (6)	0.096 (6)
	East Burke (EPM)	108 (5)	41.238 (5)	0.112 (4)
	Barnet (BPL)	85 (7)	25.311 (7)	0.058 (7)
Mettawee	Granville, NY (GML)	126 (4)	44.444 (4)	0.127 (3)

plants were widespread on both components 1 and 3. Plants from the single Mettawee-watershed population were tightly clustered on both components 1 and 3; they were largely separated from the other two watersheds on component 3. Within the Passumpsic watershed, the high and middle populations (Newark and East Burke respectively) were tightly clustered on both components 1 and 3; the low-elevation population from this watershed included outliers that clustered with each of the other two watersheds.

The three approaches to assessing the distribution of genetic diversity revealed similar regional patterns (Table 1). The Winooski watershed was overall the most genetically diverse region according to all three calculations. In the number of loci and genetic distance calculation the Winooski watershed contained the three most diverse populations. Similarly, the mean heterozygosity data showed the Winooski watershed to contain two of the three most diverse populations. The Passumpsic watershed was the least diverse region according to all of the calculations. The number of loci as well as the genetic-distance metric revealed this watershed to contain the three least diverse populations, while the heterozygosity data indicates that two of the three least diverse populations lie in this watershed. The mean heterozygosity values for all populations were low to about average relative to other North Temperate ferns, ranging from 0.058 to 0.152 (compare, e.g., Kirkpatrick *et al.*, 1990; Suter *et al.*, 2000).

DISCUSSION

Phylogeny.—Our analysis of two highly labile DNA markers, one chloroplast and one nuclear, reveals *Matteuccia struthiopteris* to comprise a globally cohesive and minimally divergent system of populations. The relationships revealed by the phylogeny in this study are congruent with the relationships proposed by Gastony and Ungerer (1997). In both cases the clade containing *Pentarhizidium orientale* and *P. intermedium* is situated almost equally distant from *Onoclea sensibilis*

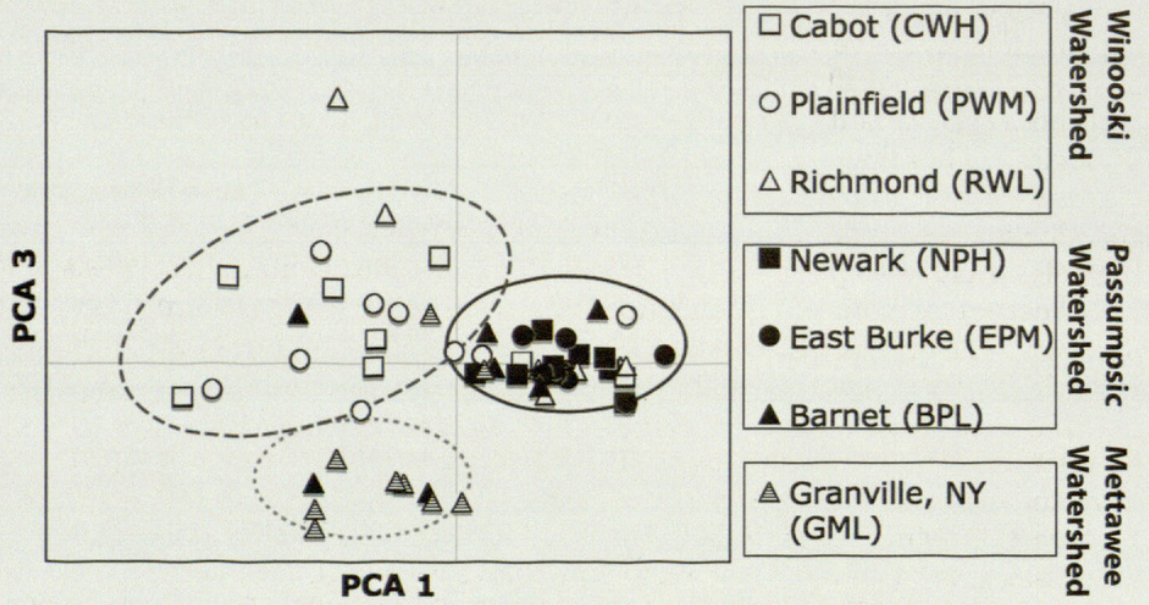

FIG. 3. PCA analysis of AFLP variation in *Matteuccia struthiopteris* by Vermont population and watershed. Shapes represent elevational category: square is high, circle is medium and triangle is low.

(Gastony and Ungerer: 49 steps, this study: 37 steps) and the clade containing *M. struthiopteris* (Gastony and Ungerer: 45 steps, this study: 40 steps).

Our recovery of a strongly supported and divergent clade comprising *Pentarhizidium orientale* and *P. intermedium*—corroborating the results of Gastony and Ungerer (1997)—reinforces the support for recognizing this pair in a genus separate from *Matteuccia*. Though Gastony and Ungerer's data suggest that *Onocleopsis hintonii* is better treated as a species of *Matteuccia sensu stricto*, we were unable to explore this possibility as we were unable to retrieve nuclear-DNA sequence data from this species.

AFLP analysis.—Analysis of the structure of genetic diversity in 60 *Matteuccia struthiopteris* samples revealed that the plants were differentiated between watersheds, rather than between elevations; we take this to suggest that historical biogeography rather than local recent events in Vermont populations accounts for the pattern we retrieved.

The large low-elevation population in the Winooski watershed was low in heterozygotes and lowest (within its watershed) in genetic diversity assessed from both number of loci and mean genetic distance. This population is located in Richmond, an area that receives high fiddlehead harvesting pressures (Maisonpierre, 2009). The low diversity in this population may be anthropogenic in nature. However, a relatively recent origin or expansion from a bottlenecked population is also possible.

Downy and smooth variants.—The geographic distribution of downy individuals presented a pattern at odds with the molecular-genetic signal in the sampled populations. Given the absence of molecular signal suggesting overall greater genetic diversity in the larger rivers, it appears that the downy variant may

be selected against at higher elevations. (Genetic drift seems a less likely explanation, since we would expect at least some high-elevation populations to be downy under a drift scenario). We are left with the working hypothesis that the morphological variant, unlike the molecular variation taken as a whole, has an ecological rather than a historical explanation.

Quaternary biogeography of Matteuccia struthiopteris.—The concentration of genetic diversity in the Winooski watershed (a geographic region), rather than at high elevations (i.e., peripherally) or low elevations (i.e., centrally), suggests that historical rather than environmental factors are driving the pattern of genetic diversity. The greater genetic diversity in the watershed draining westward into Lake Champlain suggests that it may be nearer to the location of populations that lay in a Pleistocene refugium. In this context, the relative decrease in genetic diversity and heterozygosity seen outside of the Champlain Valley may be a result of historical (post-glacial) expansion of the species from west to east across the northeast. A Mississippi Valley refugium for *Matteuccia struthiopteris* seems plausible, as refugia in the American South have been suggested for a set of North American species (Barrington and Paris, 2007; Davis 1983; Hewitt 2000, 2004; Willis and Hewitt, 2004).

In an attempt to gain insight into Holocene expansion in *Matteuccia struthiopteris*, we conducted a larger AFLP analysis (Koenemann, 2009). This larger analysis included *M. struthiopteris* populations from across northeastern North America and adjacent Québec and New Brunswick. However, while the results of this study did corroborate our suggestion that genetic diversity in the fiddlehead fern decreases as one moves from the south and west to the north and east, geographic patterning of the genetic diversity in populations was lost at this larger scale.

Conclusions.—*Matteuccia struthiopteris* is a globally cohesive species, characterized by minimal genetic variation, within which Eastern North American and Eurasion populations—the latter including the Alaskan population—form separable evolutionary lineages. The lower genetic diversity to the east in Vermont suggests expansion eastward in northeast North America. In the context of other studies, migration north and east from a Pleistocene refugium in the Mississippi Valley seems plausible as a working hypothesis.

ACKNOWLEDGMENTS

This work is the result of two University of Vermont undergraduate honors theses: the first (chronologically) by Maisonpierre yielded the population-level insights using AFLP data, and the second by Koenemann yielded the world-level insights using DNA sequence data. The work was supported by funding for work on the fiddlehead fern in Vermont, CSREES Grant VT-H01405 to Barrington and University of Vermont HELiX undergraduate research awards to both Koenemann and Maisonpierre. A remarkable community of correspondents provided plant material for the analysis, including Donald Farrar, Kate Mohatt, Li Chun-xiang, Audrey Reid, Anthony Reznicek, A. K. M. Golam Sarwar, Christopher Sears, and Mary Stensvold. Joshua Fontaine helped Koememann with fieldwork and collecting. In the lab, we had help from graduate students Stacy Jorgensen, Monique McHenry, and Erin Sigel. The Alaska sequences were the first that undergraduate Brendan Lyons ever produced. We are especially grateful to Erin Sigel for her lucid and deeply informed development of the AFLP analysis.

Literature Cited

Barrington, D. S. and C. A. Paris. 2007. Refugia and migration in the Quaternary history of the New England flora. Rhodora 109:369–386.

Bergerson, M. E. and L. Lapointe. 2001. Impact of one year crozier removal on long-term frond production in *Matteuccia struthiopteris*. Canadian Journal of Plant Science 81:155–163.

Bonin, A., E. Bellemain, P. B. Eidensen, F. Pompanon, C. Brochmann and P. Taberlet. 2004. How to track and assess genotyping errors in population genetic studies. Molecular Ecology 13:3261–3273.

Bonin, A. D. Ehrich and S. Manel. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. Molecular Ecology 16:3737–3758.

Davis, M. B. 1983. Quaternary History of Deciduous Forest of Eastern North America and Europe. Annals of the Missouri Botanical Garden 70:550–563.

Dempster, E. L., K. L. Pryor, D. Francis, J. E. Young and H. J. Rogers. 1999. Rapid DNA extraction from ferns for PCR-based analyses. Biotechniques 27:66–68.

Donoghue, M. J., C. D. Bell and J. Li. 2001. Phylogenetic patterns in Northern Hemisphere plant geography. International Journal of Plant Science 162:S41–S52.

Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh frond tissue. Phytochemical Bulletin 19:11–15.

Driscoll, H. E. and D. S. Barrington. 2007. Origin of Hawaiian *Polystichum* (Dryopteridaceae) in the context of a world phylogeny. Amer. J. Bot. 94:1413–1424.

Dufresne, F. and P. D. N. Hebert. 1997. Pleistocene glaciations and polyphyletic origins of polyploidy in an arctic cladoceran. Proceedings of the Royal Society of London 264:201–206.

Gastony, G. J. and M. C. Ungerer. 1997. Molecular systematics and a revised taxonomy of the onocleoid ferns (Dryopteridaceae: Onocleeae). American Journal of Botany 84:840–849.

Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid Symposium Series 41:95–98.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role, in divergence and speciation. Biological Journal of the Linnean Society 58:247–276.

Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society B 359:183–195.

Ibrahim, K. M., R. A. Nichols and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity 77:282–291.

Ishikawa, H., Y. Watano, K. Kano, M. Ito and S. Kurita. 2002. Development of primer sets for PCR amplification of the PgiC gene in ferns. Journal of Plant Research 115:65–70.

Kato, M. 1993. Biogeography of ferns: dispersal and vicariance. Journal of Biogeography 20:265–274.

Kimura, M. and H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576.

Kirkpatrick, R. E. B., P. S. Soltis and D. E. Soltis. 1990. Mating system and distribution of genetic variation in *Gymnocarpium dryopteris* subsp. *disjunctum*. American Journal of Botany 77:1101–1110.

Koenemann, D. M. 2009. The Fiddlehead Fern (*Matteuccia struthiopteris*) in the Holocene of North America. Undergraduate Honors Thesis, University of Vermont, Burlington, Vermont, USA.

Kress, W. J., K. J. Wurdack, E. A. Zimmer, L. A. Weigt and D. H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences. 102:8369–8374.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948.

Lloyd, R. M. 1971. Systematics of the onocleoid ferns. University of California Publications in Botany 61:1–86.

Maddison, D. and W. Maddison. 2005. *MacClade: analysis of phylogeny and character evolution: version 4.03*. Sinauer Associates, Sunderland, Massachusetts, USA.

MAISONPIERRE, J. A. 2009. Impacts of harvesting practices on genetic diversity in Vermont's fiddlehead fern *Matteuccia struthiopteris* (L.) Todaro. Undergraduate Honors Thesis, University of Vermont, Burlington, Vermont, USA.

MEUDT, H. M. and A. C. CLARKE. 2006. Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends in Plant Science 12:106–117.

MIYAZAWA, M., E. HORIUCHI and J. KAWATA. 2007. Components of the essential oil from *Matteuccia struthiopteris*. Journal of Oleo Science 56:457–461.

NAKAZATO, T., M. K. JUNG, E. A. HOUSWORTH, L. H. RIESEBERG and G. J. GASTONY. 2006. Genetic map-based analysis of genome structure in the homosporous fern *Ceratopteris richardii*. Genetics 173:1585–1597.

NISHIMURA, D. 2000. Sequencher 3.1.1. Biotech Software and Internet Report 1:24–30.

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.

POMPANON, F., A. BONIN, E. BELLEMAIN and P. TABERLET. 2005. Genotyping errors: causes, consequences, and solutions. Nature Reviews Genetics 6:847–859.

PRANGE, R. K. and P. VON ADERKAS. 1985. The biological flora of Canada. 6. *Matteuccia struthiopteris* (L.) Todaro, ostrich fern. Canadian Field Naturalist 99:517–532.

SANG, T., D. J. CRAWFORD and T. F. STUESSY. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). American Journal of Botany 84:1120–1136.

SIGEL, E. M. 2008. Polyphyly of polyploid species: an inquiry into the origin of *Dryopteris campyloptera* and *Dryopteris dilatata* (Dryopteridaceae). Master's Thesis, University of Vermont, Burlington, Vermont, USA.

SMITH, A. R. 1993. *Matteuccia*. In: Flora of North America Editorial Committee, eds 1993. *Flora of North America North of Mexico*. Vol. 2 Oxford University Press, New York and Oxford.

SUTER, M., J. J. SCHNELLER and J. C. VOGEL. 2000. Investigations into the genetic variation, population structure, and breeding systems of the fern *Asplenium trichomanes* subsp. *quadrivalens*. International Journal of Plant Sciences 161:233–244.

SWOFFORD, D. L. 2001. *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*. Sinauer Associates, Sunderland, Massachusetts, USA.

VEKEMANS, X., T. BEAUWENS, M. LEMAIRE and I. ROLDÁN-RUIZ. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Molecular Ecology 11:139–151.

VON ADERKAS, P. 1984. Economic history of ostrich fern, *Matteuccia struthiopteris*, the edible fiddlehead. Economic Botany 38:14–23.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. Y. D. LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KULPER and M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407–4414.

WALLACE, L. and M. CASE. 2000. Contrasting allozyme diversity between northern and southern populations of *Cypripedium parviflorum* (Orchidaceae): Implications for Pleistocene refugia and taxonomic boundaries. Syst. Bot. 25:281–296.

WHITLOCK, R., H. HIPPERSON, M. MANNARELLI, K. BUTLIN and T. BURKE. 2008. An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. Molecular Ecology Resources 8:725–735.

WILLIS, K. J. and G. M. HEWITT. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions Royal Society London B 359:183–195.

WOLF, P. G. 2000. AFLP Protocol [online]. Available from http://bioweb.usu.edu/wolf/aflp_protocol.htm

APPENDIX 1. Accessions used in the study. Species numbers correspond to the following, 1: *Matteuccia struthiopteris*, 2: *Onoclea sensibilis*, 3: *Pentarhizidium intermedium*, 4: *Pentarhizidium orientale*. Vouchers and silica-dried material, where available, are at the University of Vermont, Burlington, VT 05405. The Genbank numbers marked with an asterisk are for plant JAM001.

Accessions Used at Location	Accession Number	Genbank Accession Numbers (*PgiC*/ *psbA-trnH*)	Population Location	Taxon	Watershed & Elevational category (AFLP)	Latitude	Longitude	Elevation (m)
1	DMK10109	HQ243688/ HQ243704	Burlington, VT, USA	2	NA	44.284°N	73.111°W	69 m
9	JAM001-JAM009	HQ243679/ HQ243695*	Cabot, VT, USA	1	Winooski – High	44.457° N	72.326° W	442 m
9	JAM011-JAM019	-	Plainfield, VT, USA	1	Winooski –Medium	44.285° N	72.415°W	238 m
8	JAM021-JAM028	-	Richmond, VT, USA	1	Winooski – Low	44.398°N	72.990°W	87 m
8	JAM031-JAM038	-	Newark, VT, USA	1	Passumpsic – High	44.702°N	72.005°W	358 m
10	JAM041-JAM050	-	East Burke, VT, USA	1	Passumpsic-Medium	44.599°N	71.949°W	253 m
7	JAM051-JAM057	-	Barnet, VT, USA	1	Passumpsic – Low	44.326°N	72.037°W	156 m
1	DSB2323	-	Winooski, VT, USA	1	NA	44.531°N	73.269°W	30 m
1	MAM0709	HQ243683/ HQ243699	Stowe, VT, USA	1	NA	44.474°N	72.221°W	235 m
1	Dac3	HQ243675/ HQ243691	Dacey Pond, ME, USA	1	NA	45.611°N	68.820°W	152 m
1	DMK9109	-	Lake George, ME, USA	1	NA	44.764°N	69.592°W	70 m
9	JAM061-JAM069	-	Granville, NY, USA	1	Mettawee – Low	43.444°N	73.282°W	178 m
1	Rez01	HQ243676/ HQ243692	Ann Arbor, MI, USA	1	NA	42.125°N	83.445°W	194 m

APPENDIX 1. Continued.

Accessions Used at Location	Accession Number	Genbank Accession Numbers (*PgiC*/*psbA-trnH*)	Population Location	Taxon	Watershed & Elevational category (AFLP)	Latitude	Longitude	Elevation (m)
1	DMK0109	HQ243682/ HQ243698	Rivière Verte, QC, Canada	1	NA	48.042°N	69.302°W	30 m
1	DMK1109	HQ243680/ HQ243696	Rivière Saumon, QC, Canada	1	NA	48.697°N	67.931°W	65 m
1	DMK2109	–	Rivière Matane, QC, Canada	1	NA	48.735°N	67.495°W	33 m
1	DMK3109	HQ243681/ HQ243697	Campbelton, NB, Canada	1	NA	47.982°N	66.282°W	6 m
1	DMK4109	–	Redmondville, NB, Canada	1	NA	46.907°N	65.199°W	53 m
1	DMK5109	–	Moncton, NB, Canada	1	NA	46.315°N	64.624°W	15 m
1	DMK6109	–	Canon River Bank, NB, Canada	1	NA	45.969°N	65.196°W	29 m
1	DMK7109	–	Waterborough, NB, Canada	1	NA	45.871°N	66.007°W	51 m
1	DMK8109	–	Southampton, NB, Canada	1	NA	45.952°N	67.289°W	64 m
1	Sea1071	HQ243677/ HQ243693	Chilliwack, BC, Canada	1	NA	49.114°N	121.555°W	1012 m
1	MCS001	JF912402/ JF912403	Russian River Campground, Alaska	1	NA	60.484°N	149.973°W	300 m
1	MOC001	HQ243689/ HQ243705	Sapporo, Hokkaido, Japan	1	NA	43.067°N	141.350°E	25 m

APPENDIX 1. Continued.

Accessions Used at Location	Accession Number	Genbank Accession Numbers (*PgiC*/ *psbA-trnH*)	Population Location	Taxon	Watershed & Elevational category (AFLP)	Latitude	Longitude	Elevation (m)
1	MOC002	HQ243690/ HQ243706	Sapporo, Hokkaido, Japan	1	NA	43.067°N	141.350°E	25 m
1	AHR001	HQ243685/ HQ243701	Västerås, Sweden	1	NA	59.611°N	16.551°E	17 m
1	CB15	HQ243684/ HQ243700	Mt. Changbaishan, Heilongjian Province, China	1	NA	NA	NA	NA
1	2D7	HQ243686/ HQ243702	Zhongdian, Yunnan Province, China	3	NA	NA	NA	NA
1	DL47	HQ243687/ HQ243703	Dali, Yunnan Province, China	4	NA	NA	NA	NA