

# Chloroplast haplotype variation among monoecious and dioecious populations of *Sagittaria latifolia* (Alismataceae) in eastern North America

M. E. DORKEN and S. C. H. BARRETT

Department of Botany, University of Toronto, 25 Willcocks Street, Toronto, ON, M5S 3B2, Canada

## Abstract

Aquatic plants commonly have extensive geographical distributions, implying few restrictions to dispersal. Here we investigate the postglacial history of an aquatic plant with contrasting sexual systems (monoecy and dioecy), which are predicted to affect dispersal ability. We examined the distribution of cpDNA haplotypes using polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) among 76 populations (32 monoecious, 38 dioecious, two mixed and four undetermined populations) of *Sagittaria latifolia* sampled throughout eastern North America. We also use these data to investigate the polarity of the evolutionary transition between monoecy and dioecy. Using PCR–RFLP, we identified eight cpDNA haplotypes. All haplotypes were found in unglaciated areas of the species' range, clustered primarily in the southeastern United States, providing evidence that glacial refugia probably occurred in this area. Genetic diversity ( $h_T$ ) was more than six times greater among monoecious compared to dioecious populations. All seven of the haplotypes for which the sexual system could be determined were represented among monoecious populations. In contrast, only four haplotypes were detected in dioecious populations and 94% of individuals from dioecious populations possessed a single haplotype. Monoecious populations possessing this widespread haplotype were restricted to the southern portion of the range, indicating that dioecy probably originated in this region and then spread northwards. The distribution of cpDNA haplotypes in dioecious populations represents a subset of the variation found in monoecious populations, a pattern expected if dioecy has evolved from monoecy in *S. latifolia*.

*Keywords:* aquatic plants, dioecy, monoecy, PCR–RFLP, phylogeography, postglacial history

Received 17 February 2004; revision received 30 April 2004; accepted 30 April 2004

## Introduction

Aquatic plants typically have much broader geographical distributions compared to their terrestrial counterparts. This has led to the suggestion that there are fewer restrictions to dispersal in aquatic plants (Darwin 1859; Arber 1920; Ridley 1930; Sculthorpe 1967; Barrett *et al.* 1993). In temperate regions, the dispersal ability of plants affected their response to the cyclical advance and retreat of glaciers in the Pleistocene (Delcourt & Delcourt 1987). Examining

the historical patterns of dispersal provides insights into the response of species to changes in climate (Hewitt 2000) and can aid in the interpretation of phenotypic variation among geographical regions (e.g. Abbott & Comes 2004). However, for herbaceous plants in general and for aquatic plants in particular, patterns of dispersal following the retreat of the glaciers in North America are poorly understood. Recent investigations of these topics have been facilitated by the use of molecular markers (e.g. Soltis *et al.* 1992; Abbott *et al.* 2000; Griffin & Barrett 2004).

Because reproductive systems affect the ability of plants to establish new populations, differences in dispersal ability will be influenced by variation in their modes of reproduction (Holsinger 1986; Barrett 1996; Heilbut *et al.* 2001). For example, plants from dioecious populations (i.e. populations

Correspondence and current address: M. E. Dorken, Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, UK. Fax: +44 (0) 1865 275074; E-mail: marcel.dorken@plants.ox.ac.uk

comprised of females and males) may be less likely to colonize unoccupied habitat than plants from hermaphroditic populations, which often require only one individual to establish new populations (i.e. Baker's Law, Baker 1955; Pannell & Barrett 1998). Variation in allocation to sexual vs. asexual propagation will also affect dispersal ability in plants. Because clonal propagules are typically larger and more prone to desiccation than seeds, they are usually less important for dispersal among habitats (Abrahamson 1980), particularly for many aquatic species with seed dispersal by waterfowl (Santamaría 2002). As a result, plants from sexual hermaphroditic populations might be expected to have a dispersal advantage over plants from clonal dioecious populations. In temperate areas, this difference should be most pronounced at the geographical margin of a species' range, where variation in dispersal ability will affect establishment during migration into previously glaciated areas.

To examine the historical patterns of dispersal for an aquatic plant we investigated *Sagittaria latifolia* Willd., a widespread species from eastern North America occurring in both previously glaciated and unglaciated regions of the continent (Bogin 1955; Haynes & Hellquist 2000). This species is unusual among flowering plants because there are both monoecious (hermaphrodites with unisexual flowers) and dioecious populations (Wooten 1971; Sarkissian *et al.* 2001; Dorken *et al.* 2002) determined by the simple Mendelian segregation of alleles at nuclear loci (Dorken & Barrett 2004). Plants of the two sexual systems have contrasting patterns of investment in life-history traits affecting dispersal (Dorken & Barrett 2003). In monoecious populations, plants have greater investment in sexual propagation than those from dioecious populations, which are typically clonal with limited sexual recruitment. Populations of the two sexual systems are geographically widespread and often occur in close proximity. However, in areas of geographical overlap, restricted gene flow maintains the integrity of the two sexual systems and mixed populations are uncommon (Dorken *et al.* 2002).

The goal of this study was to examine intraspecific patterns of chloroplast DNA (cpDNA) variation in populations sampled throughout the geographical range of *S. latifolia* in eastern North America. Because the chloroplast genome is haploid, nonrecombining and maternally inherited through seeds, such variation should enable investigations of the historical patterns of dispersal (e.g. Stehlik 2002), as well as the genealogical relationships among haplotypes (e.g. Maskas & Cruzan 2000). Specifically, cpDNA variation can be organized into ordered networks of descent, providing a basis for understanding historical associations between genes and geography (reviewed in Ennos *et al.* 1999). In addition, differences in the partitioning of cpDNA variation between the two sexual systems in *S. latifolia* may provide information on the evolutionary relationships

between monoecious and dioecious populations. In particular, because there is limited gene flow between the two sexual systems, if dioecy is more recently derived in *S. latifolia* we might predict that only a restricted subset of the cpDNA haplotypes found in monoecious populations would be represented among dioecious populations. Although it is generally assumed that dioecy is the derived condition (Bull & Charnov 1985), few studies have explicitly examined the polarity of this transition in the flowering plants (but see Weiblen 2000; Renner & Won 2001; Joussetin *et al.* 2003) and none have attempted this at the intraspecific level. Determining the sequence of evolutionary changes in reproductive systems is important for understanding the ecological and genetic factors involved in the evolution and maintenance of sexual systems.

Here we use a broad geographical sample of monoecious and dioecious populations of *S. latifolia* to address the following specific questions: (1) what is the genetic structure of populations as revealed by cpDNA variation, and what does this structure indicate about the postglacial history of *S. latifolia* in eastern North America? (2) Are there differences in the geographical distribution of monoecious populations compared to dioecious populations? Specifically, in previously glaciated areas, do monoecious populations of *S. latifolia* occur further north than dioecious populations, as expected if individuals from monoecious populations have greater dispersal ability? (3) Is there a difference in the distribution of cpDNA haplotypes among monoecious vs. dioecious populations? If so, does this difference reveal the polarity of the transition between the sexual systems?

## Materials and methods

### Study system

*Sagittaria latifolia* is an emergent, clonal aquatic common to a variety of wetland habitats throughout eastern North America (Smith 1894; Bogin 1955). Monoecious and dioecious populations grow in habitats that differ in the importance of disturbance vs. competition (Dorken & Barrett 2003). Specifically, monoecious populations occur typically in more ephemeral and disturbed habitats, while dioecious populations are found more commonly in extensive wetlands associated with large river systems (Dorken & Barrett 2003; see Wooten 1971). Corms are the sole means of perennation in this species and can float, potentially facilitating dispersal within wetlands. However, seeds are more likely to be involved in dispersal between watersheds because of their smaller size and tolerance to desiccation. Flowering shoots produce racemes with three unisexual flowers at each node. These flowers are visited by a variety of generalist pollinators, including flies, bees and wasps (Muenchow & Delesalle 1994).

### *Distribution of monoecious vs. dioecious populations*

To evaluate the hypothesis that monoecious populations should occur further north than dioecious populations, we conducted analysis of variance (ANOVA) in latitude among the sampled populations (see below). We chose populations by visiting localities identified from herbarium records. Populations that were found by driving between these sites were also sampled, yielding a total of 76 populations of *S. latifolia* from throughout the eastern United States and southeastern Canada. Because we were particularly interested in examining the northern half of the species' range, and thus the portion of the range most affected by the advance and retreat of the glaciers in eastern North America, we included the region in which populations were sampled (i.e. northern vs. southern populations) as a fixed effect in a two-way ANOVA. We defined northern and southern populations as those occurring north and south of the degree of latitude midway between the southern- and northern-most populations in the sample, respectively. We also investigated the correlation between haplotype diversity and latitude by dividing the difference between the minimum and maximum latitudes by eight and assigning populations to one of these eight categories.

### *Tissue sampling*

In July–August 2001 and August–September 2002, we collected leaf samples from the populations described above. We sampled tissue from eight individuals per population and recorded the sexual system of each population. To reduce repeated sampling of the same clonal genotype, we collected samples at 2–10 m intervals, depending on the spatial extent of clones. We found two populations that appeared to contain mixtures of 'monoecious' and 'dioecious' individuals, and four nonflowering populations for which we were unable to determine the sexual system. These populations were removed from the analyses of genetic differentiation between the sexual systems, yielding 32 monoecious and 38 dioecious populations in the sample. Leaf material was dried and stored in silica gel until DNA isolation was performed.

### *cpDNA isolation and polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP)*

We extracted total DNA using a simplified version of Doyle & Doyle (1987; see Griffin & Barrett 2004 for details) and conducted PCR–RFLP analysis to detect intraspecific cpDNA polymorphisms. We began by screening one individual from a subset of the populations using 20 universal cpDNA primer pairs corresponding to noncoding intergenic regions of cpDNA (Taberlet *et al.* 1991; Demesure *et al.* 1995; Dumolin-Lapegue *et al.* 1997). Of these, seven

primer pairs provided reliable amplification (TFc-d, TFe-f, CS, ML, K1K2, ST, TC). These amplified regions were screened for polymorphisms using 16 restriction enzymes [4-, 5- and 6-base pair (bp) cutters]. We performed DNA amplification using a Hybaid PCRexpress® using the following conditions: 1.5 mM MgCl<sub>2</sub>, 200 µM of each dNTP, 0.00625 U/µL *Taq*, and 1 × PCR buffer; denaturing for 1 min at 94 °C, annealing at: 56 °C for 30 s for TFc-d; 60 °C for 30 s for TFe-f; 60 °C for 1 min for CS, K1K2 and ST; 57 °C for 1 min for ML; and 58 °C for 1 min for TC; and extension for 2 min at 72 °C for 40 cycles. We used 1.6% agarose gels stained with ethidium bromide to assay the cpDNA regions.

Our initial screen yielded five polymorphic regions of cpDNA in *S. latifolia*. Two of these were sequence polymorphisms (gain/loss of a restriction site): TFe-f cut with the restriction enzymes *SspI* and *HinFI*. This inference was verified by sequencing the TFe-f fragment (Dorken 2004). The TFc-d, K1K2 and ST amplified regions each had insertion/deletion (indel) polymorphisms that were also resolvable using a number of restriction enzymes; however, *AluI*, *Bsh1236I* and *MboI*, respectively, provided the best resolution of differences among the fragments. To determine the distribution of these polymorphisms within and among populations, we screened eight individuals per population ( $n = 608$  individuals from 76 populations) for each polymorphic fragment × restriction enzyme combination.

### *Analysis of cpDNA variation*

We used analysis of molecular variance (AMOVA) as well as measures of genetic diversity ( $h_S$ ,  $h_T$ ) and differentiation ( $G_{ST}$ ,  $N_{ST}$ ) to evaluate the partitioning of haplotype variation among populations. Our primary goal was to use these techniques to compare patterns of haplotype variation between monoecious and dioecious populations. AMOVA was conducted using the squared Euclidean distances among 32 monoecious and 38 dioecious populations using ARLEQUIN (version 2.000; Schneider *et al.* 2000). We calculated measures of diversity ( $h_S$  and  $h_T$ , the gene diversity averaged within populations and among all populations, respectively) and differentiation ( $G_{ST}$  and  $N_{ST}$ ) in the chloroplast genome following the methods described by Pons & Petit (1995, 1996) using the program PERMUT, where  $G_{ST}$  is based on haplotype frequencies and  $N_{ST}$  takes similarities between haplotypes into account. We compared values of  $h_T$  between populations occurring in the northern vs. the southern halves of the sampled range, and values of  $N_{ST}$  between monoecious and dioecious populations using a permutation test with 1000 permutations (Burban *et al.* 1999). We also used these measures of diversity and differentiation to examine geographical patterns in the distribution of haplotype variation. Specifically, we compared the values obtained for these parameters between populations from

the northern vs. southern portion of the species' range, as defined above. For this analysis we also included populations for which the sexual system was unknown, yielding 31 southern and 45 northern populations. Finally, we examined the relatedness of cpDNA haplotypes of *S. latifolia*, by constructing a minimum-spanning tree using ARLEQUIN. The reliability of this network was determined by calculating the coefficient of cophenetic correlation between the tree distance and raw distance matrices (Smouse 1998).

## Results

### *Distribution of monoecious vs. dioecious populations*

As predicted, monoecious populations of *S. latifolia* occurred further north than dioecious populations. Analysis of variance of population latitude revealed a significant interaction between region (northern vs. southern) and sexual system (monoecious vs. dioecious;  $F_{1,63} = 4.8, P < 0.05$ ). This interaction is the result of the higher average latitude of monoecious (average =  $44.3^\circ \text{N} \pm 0.6 \text{SE}$ ) compared to dioecious (average =  $42.2^\circ \text{N} \pm 0.5 \text{SE}$ ) populations in the northern portion of the species range in eastern North America (least squares means contrast:  $F_{1,63} = 7.4, P < 0.01$ ). In the southern portion of the species' range, there was no difference in average latitude between monoecious (average =  $33.3^\circ \text{N} \pm 0.7$ ) and dioecious (average =  $34.0^\circ \text{N} \pm 0.8$ ) populations.

### *Geographical distribution of PCR-RFLP haplotypes*

Using PCR-RFLP, we detected eight cpDNA haplotypes for *S. latifolia* (Table 1). All haplotypes were represented in the unglaciated regions of eastern North America, and all but one haplotype could be found in the region bounded by Alabama, Mississippi, Tennessee and Florida's panhandle (Fig. 1). Only two haplotypes (D and G) were found in areas north of the last glacial maximum. Of these two haplotypes, one (haplotype G) was widespread, occurring in all but the northernmost part of the species' range and representing 64% of all individuals sampled. The other haplotype (haplotype D) occurred almost exclusively in glaciated eastern United States and Ontario, with one population in unglaciated southern Indiana. Reduced haplotype diversity in northern populations resulted in lower

**Table 1** Description of the RFLP haplotypes detected from four PCR products cut with five restriction enzymes (indicated in brackets) in *Sagittaria latifolia*. Columns describe the pattern of RFLPs for each PCR product and restriction enzyme combination. Rows describe the combined PCR-RFLP phenotype for each haplotype

Haplotype	K1K2 ( <i>Bsh1236I</i> )	ST ( <i>MboI</i> )	TFc-d ( <i>AluI</i> )	TFe-f ( <i>HinfI</i> )	TFe-f ( <i>SspI</i> )
A	1	3	2	1	0
B	1	3	1	1	0
C	1	3	2	1	1
D	1	3	3	1	1
E	1	1	2	0	1
F	2	1	3	0	1
G	2	1	1	0	1
H	2	2	1	0	1

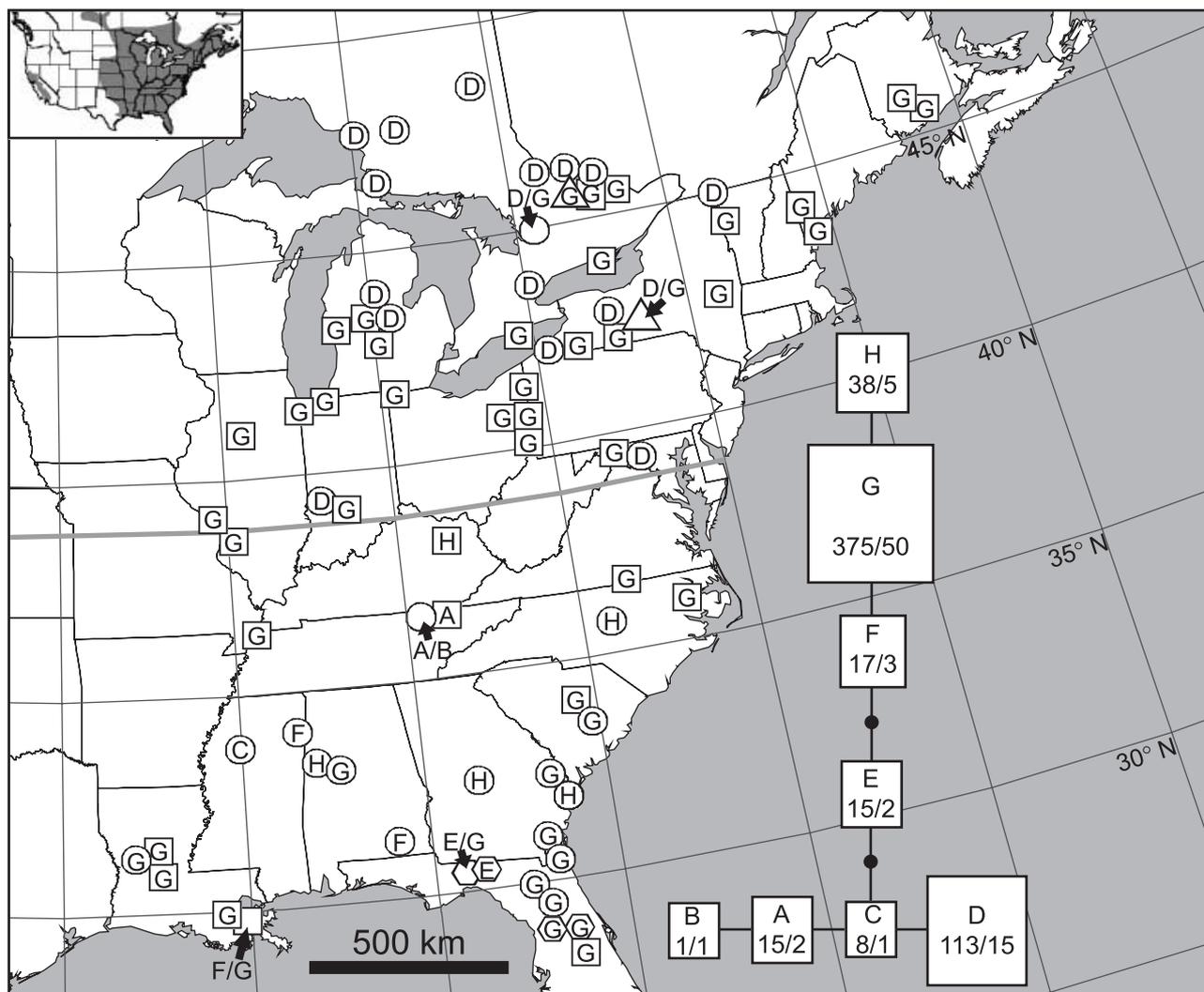
estimates of genetic diversity among populations ( $h_T = 0.46$ ) compared to southern populations ( $h_T = 0.59$ ; Table 2). There was less than 5% overlap between permuted  $h_T$  values between the two regions, indicating that the structuring of genetic diversity differed significantly between the two regions. Reduced diversity in northern areas also resulted in a negative correlation between latitude and haplotype diversity ( $r = -0.74$ ;  $n = 8$  categories of latitude;  $P < 0.04$ ). Throughout the range, there were few populations that were polymorphic for cpDNA markers, resulting in substantial values of  $G_{ST}$  and  $N_{ST}$  (Table 2).

### *Distribution of haplotypes in monoecious vs. dioecious populations*

PCR-RFLP haplotypes were unequally distributed between monoecious and dioecious populations *S. latifolia* (Fig. 2). We were able to determine the sexual system for all populations for seven of the eight haplotypes detected. All seven of these haplotypes were found in monoecious populations. In contrast, only four haplotypes were found in dioecious populations, and 94% of the individuals sampled from these populations had the same haplotype (haplotype G). The monoecious populations representing these four haplotypes all occurred in the southern portion of the species' range. The contrasting distribution of haplotypes between the sexual systems was evident from our

Region	$G_{ST}$	$N_{ST}$	$h_T$	$h_S$
Northern populations	0.95 (0.03)	0.95 (0.03)	0.46 (0.05)	0.02 (0.02)
Southern populations	0.93 (0.04)	0.91 (0.06)	0.59 (0.09)	0.04 (0.02)
Monoecious populations	0.97 (0.02)	0.98 (0.02)	0.70 (0.06)	0.02 (0.02)
Dioecious populations	0.94 (0.07)	0.96 (0.05)	0.11 (0.07)	0.01 (0.01)
All populations	0.95 (0.02)	0.94 (0.03)	0.54 (0.06)	0.03 (0.01)

**Table 2** Differentiation and diversity ( $\pm \text{SE}$ ) of the chloroplast genome for 76 populations of *Sagittaria latifolia* in eastern North America



**Fig. 1** Distribution of PCR-RFLP haplotypes among monoecious (circles), dioecious (squares) and mixed (triangles) populations of *Sagittaria latifolia* in eastern North America. Hexagons indicate populations for which the sexual system was unknown. The grey line below 40° latitude indicates the divide between northern and southern populations in this study (see Methods). The minimum-spanning tree indicates genealogical relationships among the PCR-RFLP haplotypes detected. The tree was a 'good fit' (Rohlf 1997) with the raw haplotype distances (coefficient of cophenetic correlation:  $r = 0.89$ ). The numbers below each haplotype identify the number of individuals and populations representing a given haplotype, respectively. Black dots indicate linking haplotypes that were not detected in the PCR-RFLP analysis. The inset map shows the distribution of *S. latifolia* in North America (from Haynes & Hellquist 2000). Chloroplast DNA sequence and PCR-RFLP data from the closely related *Sagittaria australis* and *S. cuneata* indicate that haplotypes A, B and C are likely to be most ancestral in *S. latifolia* (Dorken 2004).

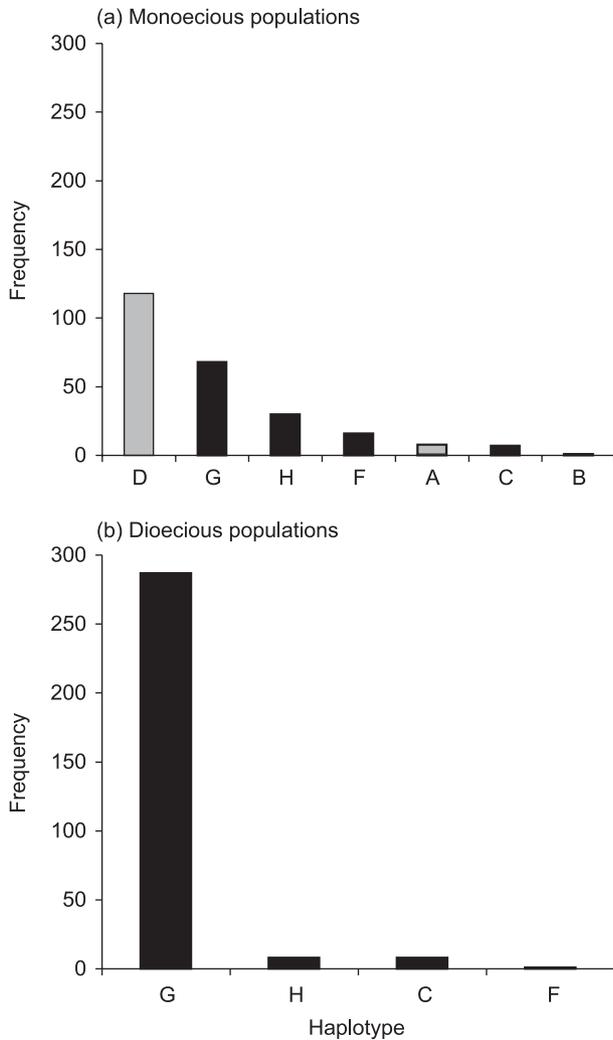
estimate of total genetic diversity ( $h_T$ ). Monoecious populations had a value of  $h_T \approx$  sixfold greater than for dioecious populations (Table 2). Genetic diversity in monoecious populations was greater in both the northern ( $h_T = 0.15 \pm 0.11$  SE) and southern ( $h_T = 0.72 \pm 0.09$  SE) portions of the sampled range compared to dioecious populations for which no haplotype variation was found in the north, and in the south  $h_T = 0.34 \pm 0.16$  SE.

Analysis of molecular variance revealed that the disparity in the distribution of haplotypes between monoecious and dioecious populations of *S. latifolia* resulted in substantial

differentiation between monoecious and dioecious populations (Table 3). Indeed, 46% of the variation in pairwise distances among populations was explained by differences in the distribution of PCR-RFLP haplotypes between monoecious and dioecious populations.

## Discussion

All eight PCR-RFLP haplotypes detected in *S. latifolia* occurred in unglaciated regions of the species' range. Because of the slow rate of cpDNA evolution (Zurawski *et al.* 1984),



**Fig. 2** Distribution of PCR-RFLP haplotypes in: (a) monoecious and (b) dioecious populations of *Sagittaria latifolia*. Black bars indicate haplotypes that were found in both sexual systems, while shaded bars indicate haplotypes that were found only in monoecious populations.

these haplotypes probably predate the most recent glaciation. Indeed, with an average synonymous substitution rate of 0.1–0.3% per Myr (Wolfe *et al.* 1987), and only four nucleotide substitutions in a 238 bp region of cpDNA among haplotypes of *S. latifolia* (Dorken 2004), the divergence

among the majority of haplotypes is likely to have occurred several million years ago, and almost certainly well before deglaciation (and see Stehlik *et al.* 2002). If so, the geographical distribution of haplotypes likely reflects their sorting in eastern North America, not evolutionary diversification following migration. Thus, for the levels of cpDNA polymorphism found, explicit phylogeographical approaches, such as nested clade analysis (Templeton 1998) are less appropriate than for more rapidly evolving genomes (and see Knowles & Maddison 2002), for which the evolutionary diversification of haplotypes in a geographical context can be explicitly studied (Avice 2000).

#### Geographic distribution of cpDNA haplotypes

The presence of all but one *S. latifolia* haplotype in the region bounded by Alabama, Mississippi, Tennessee and Florida's panhandle indicates that the most important areas of glacial survival for *S. latifolia* were likely to have been in this zone. This area includes the southern Appalachians and the Mississippi Embayment, landforms that have remained uncovered by ancient seas and glacial ice since the close of the Palaeozoic and Tertiary, respectively (Stuckey 1993). It is therefore interesting that the three PCR-RFLP haplotypes that appear to be most ancestral (i.e. haplotypes A, B and C, Fig. 1) are restricted to this area. This region was also important for the survival of terrestrial plant species (Davis 1983; Delcourt & Delcourt 1987; Stuckey 1993; Griffin & Barrett 2004) and a variety of animal groups (e.g. Zamudio & Savage 2003), including freshwater fish (Bernatchez & Wilson 1998) during the most recent glaciation.

Only two haplotypes (D and G) were found in the glaciated portion of the species' range, consistent with the common observation that genetic variation is reduced in these areas (Hewitt 2000). These haplotypes were also found in unglaciated regions in areas close to where the glaciers reached their southern limits, indicating possible northern areas of survival during glaciation. For haplotype D, its southernmost occurrence in the sample was in southern Indiana. During the Wisconsinan glaciation, this region was likely to have been dominated by boreal vegetation (Delcourt & Delcourt 1987). Because populations of *S. latifolia* with this haplotype occur in the boreal region of

Source of variation	d.f.	Sum of squares	Variance components	% of variation
Between sexual systems	1	148.7	0.52	46.4
Among populations within sexual systems	68	317.8	0.58	52.1
Within populations	490	8.6	0.02	1.6
Total	559	475.2	1.12	

**Table 3** Analysis of molecular variance for 32 monoecious and 38 dioecious populations of *Sagittaria latifolia* from eastern North America. All fixation indices and variance components were highly significant ( $P < 0.001$ )

$$F_{SC} = 0.97, F_{ST} = 0.98, F_{CT} = 0.46.$$

central Ontario (Fig. 1), it seems likely that they could have survived in similar vegetation zones during the glacial maximum. Moreover, the existence of northern refugia for *S. latifolia* is supported by the close tracking of fossilized *Sagittaria* pollen in lake sediments to the retreating edge of the glacier (Dieffenbacher-Krall & Jacobson 2001). Northern areas of survival have also been proposed for the ericaceous shrub *Calluna vulgaris*, which occurs in the formerly glaciated and unglaciated areas of Europe (Rendell & Ennos 2002).

Haplotype G is widespread in the glaciated portion of the range of *S. latifolia*. However, this haplotype was also common throughout the species' range, occurring in several portions of the range considered to be important refugia for aquatic plants, including the Atlantic and Gulf of Mexico coastal plains, the Mississippi Embayment and the Appalachian mountains (Stuckey 1993). The presence of this haplotype in several glacial refugia appears likely to have facilitated its spread into previously glaciated areas.

The occurrence of widespread haplotypes combined with the general absence of geographical clustering (Fig. 1) suggests that dispersal is an important feature of the colonization dynamics of *S. latifolia*. However, high values of  $G_{ST}$  and  $N_{ST}$  are more consistent with restricted gene flow among populations (Table 2). This discrepancy in patterns of genetic differentiation with the dynamics of population turnover may be resolved by the occurrence of long-distance dispersal events in combination with founder events. If these events were important during the colonization of eastern North America, this would explain the absence of populations that were polymorphic for cpDNA markers, and thus high values of  $G_{ST}$  and  $N_{ST}$ . Moreover, patterns of allozyme variation from a subset of populations from the species' range in southern Ontario indicate that there is little evidence for isolation-by-distance among populations (Dorken *et al.* 2002), a finding also consistent with the importance of long-distance dispersal and founder events in *S. latifolia*.

#### *The evolution of dioecy in S. latifolia*

Although dioecy is rare among angiosperms, occurring in only 6% of species, it is reported from 45% of flowering plant families (Renner & Ricklefs 1995). This distribution indicates that dioecy tends to occur in the derived portions of angiosperm clades, an inference that is supported by detailed investigations of the phylogenetic relations among hermaphroditic and dioecious taxa (Weiblen 2000; Renner & Won 2001). Indeed, Bull & Charnov (1985) considered the evolution of dioecy to be an irreversible evolutionary change, although more recent studies indicate that reversions are possible (Huff & Wu 1992). Because virtually all other species of *Sagittaria* are monoecious (Bogin 1955), it also seems likely that dioecy has evolved from monoecy in *S. latifolia*.

As predicted, if dioecy is the derived condition in *S. latifolia*, the haplotypes found in dioecious populations represented only a restricted subset of those found in monoecious populations. Indeed, the majority of individuals from dioecious populations had the same haplotype (haplotype G). In contrast, there was a more even distribution of haplotype diversity in monoecious populations (Fig. 2). This difference in the distribution of cpDNA haplotypes between the sexual systems was also evident from the low values of genetic diversity in dioecious compared to monoecious populations (see values of  $h_T$  in Table 2). Taken together, these findings are consistent with monoecy being the ancestral condition with a more recent origin of dioecy, possibly in populations possessing haplotype G.

Non-historical mechanisms to explain the difference in haplotype distribution between monoecious and dioecious populations seem less likely. For example, reduced haplotype variation in dioecious populations could reflect differences in the total amount of genetic variation that can be maintained across geographical regions at equilibrium between the sexual systems. Specifically, if patterns of colonization and extinction differ between the sexual systems, then differences in haplotype diversity between monoecious and dioecious populations would be expected (Pannell & Charlesworth 2000). Indeed, our observations from natural populations indicate that population turnover is more common in monoecious compared to dioecious populations (Dorken & Barrett 2003). However, this pattern should yield the opposite result to that found in this study, with lower  $h_T$  in monoecious, not dioecious populations. Therefore, it seems more likely that the observed differences in haplotype distribution between populations of the two sexual systems are the result of historical factors associated with the origin of dioecy in *S. latifolia*.

Assuming that dioecy is the derived condition, when and where did it evolve? Because dioecious populations share the same haplotypes with monoecious populations, it is not possible to provide an estimate of the timescales involved using the data in this study. However, the restricted distribution of chloroplast haplotypes in dioecious compared with monoecious populations, in spite of complete interfertility between the sexual systems (Dorken & Barrett 2003), suggests that dioecy has evolved recently. Because monoecious and dioecious populations occur, for the most part, in the same geographical regions, given enough time, dioecious populations would be expected to possess most, if not all, of the haplotypes found in monoecious populations if periodic gene flow occurred between the sexual systems. Indeed, during the northward migration of *S. latifolia* following the retreat of the glaciers, there should have been many opportunities for gene flow between the sexual systems, potentially resulting in the homogenization of haplotypes among monoecious and dioecious populations. Even low rates of gene flow are

sufficient for the introgression of chloroplast haplotypes (e.g. Dorado *et al.* 1992). Although in *S. latifolia* such gene flow appears to have been greatly restricted, it may account for the occurrence of a dioecious population with what appears to be an ancestral haplotype (haplotype A). The apparent rarity of gene flow between monoecious and dioecious populations, particularly where both sexual systems co-occur in the same watersheds (e.g. Ontario), indicates that there is likely to be strong selection against the products of gene flow between monoecious and dioecious populations (and see Dorken & Barrett 2003).

Although alternative explanations are possible, the geographical region in which dioecy originated in *S. latifolia* seems likely to have been in the southern part of the species' range. This is because, if dioecy had evolved in the northern portion of the range, we would expect dioecious populations to be primarily composed of haplotype D, which is the most common haplotype in this area. The fact that dioecious populations are instead predominantly composed of haplotype G, and that monoecious populations with this haplotype occur only in the south of the range supports a southern origin for dioecy followed by subsequent migration northwards.

### Acknowledgements

We thank Bill Cole, Kay Hodgins and Rachel Sturge for help in the field, Pauline Wang, Steve Griffin and David Guttman for their advice and assistance with PCR-RFLP and sequencing techniques, Ivana Stehlik, Monica Geber, John Pannell and especially Richard Ennos for comments on an earlier version of the manuscript. This study was funded with the assistance of NSERC through a post-graduate scholarship to MED and a research grant to SCHB.

### References

- Abbott RJ, Comes HP (2004) Evolution in the Arctic: a phylogeographic analysis of the circumarctic plant, *Saxifraga oppositifolia* (purple saxifrage). *New Phytologist*, **161**, 211–224.
- Abbott RJ, Smith LC, Milne RI, Crawford M, Wolff K, Balfour J (2000) Molecular analysis of plant migration and refugia in the Arctic. *Science*, **289**, 1343–1346.
- Abrahamson WG (1980) Demography and vegetative reproduction. In: *Demography and Evolution in Plant Populations* (ed. Solbrig OT), pp. 89–106. Blackwell, Oxford.
- Arber A (1920) *Water Plants: a Study of Aquatic Angiosperms*. Cambridge University Press, Cambridge, U.K.
- Avise JC (2000) *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge, MA.
- Baker HG (1955) Self-compatibility and establishment after 'long-distance' dispersal. *Evolution*, **9**, 347–348.
- Barrett SCH (1996) The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London B*, **351**, 725–733.
- Barrett SCH, Eckert CG, Husband BC (1993) Evolutionary processes in aquatic plants. *Aquatic Botany*, **44**, 105–145.
- Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431–452.
- Bogin C (1955) Revision of the genus *Sagittaria* (Alismataceae). *Memoirs of the New York Botanical Garden*, **9**, 179–233.
- Bull JJ, Charnov EL (1985) On irreversible evolution. *Evolution*, **39**, 1149–1155.
- Burban C, Petit RJ, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology*, **8**, 1593–1602.
- Darwin C (1859) *On the Origin of Species*. John Murray, London, UK.
- Davis MB (1983) Quaternary history of deciduous forests of eastern North America and Europe. *Annals of the Missouri Botanical Garden*, **70**, 550–563.
- Delcourt PA, Delcourt HR (1987) *Long-Term Forest Dynamics of the Temperate Zone, Ecological Studies 63*. Springer-Verlag, New York.
- Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**, 129–131.
- Dieffenbacher-Krall AC, Jacobson GL Jr (2001) Post-glacial changes in the geographic ranges of certain aquatic vascular plants in North America. *Proceedings of the Royal Irish Academy*, **101B**, 79–84.
- Dorado O, Rieseberg LH, Arias DM (1992) Chloroplast DNA introgression in southern California sunflowers. *Evolution*, **46**, 566–572.
- Dorken ME (2004) *The evolution and maintenance of monoecy and dioecy in Sagittaria latifolia (Alismataceae)*. PhD Thesis, University of Toronto.
- Dorken ME, Barrett SCH (2003) Life-history differentiation and the maintenance of monoecy and dioecy in *Sagittaria latifolia* (Alismataceae). *Evolution*, **57**, 1973–1988.
- Dorken ME, Barrett SCH (2004) Sex determination and the evolution of dioecy from monoecy in *Sagittaria latifolia* (Alismataceae). *Proceedings of the Royal Society of London B*, **271**, 213–219.
- Dorken ME, Friedman J, Barrett SCH (2002) The evolution and maintenance of monoecy and dioecy in *Sagittaria latifolia* (Alismataceae). *Evolution*, **56**, 31–41.
- Doyle JJ, Doyle JL (1987) A rapid isolation procedure for small quantities of fresh tissue. *Phytochemical Bulletin*, **19**, 11–15.
- Dumolin-Lapegue S, Pemonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393–397.
- Ennos RA, Sinclair WT, Hu X-S, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: *Molecular Systematics and Plant Evolution* (eds Hollingsworth PM, Bateman RM, Gornall RJ), pp. 1–9. Taylor & Francis, London.
- Griffin SR, Barrett SCH (2004) Post-glacial history of *Trillium grandiflorum* (Melanthiaceae) in eastern North America: inferences from phylogeography. *American Journal of Botany*, **91**, 465–473.
- Haynes RR, Hellquist CB (2000) Alismataceae. In: *Flora of North America* (ed. Flora of North America Editorial Committee), vol. 22, pp. 7–23. Oxford University Press, New York.
- Heilbut JC, Ilves KL, Otto SP (2001) The consequences of dioecy for seed dispersal: modeling the seed-shadow handicap. *Evolution*, **55**, 880–888.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Holsinger KE (1986) Dispersal and plant mating systems: the evolution of self-fertilization in subdivided populations. *Evolution*, **40**, 405–413.

- Huff DR, Wu L (1992) Distribution and inheritance of inconstant sex forms in natural populations of dioecious buffalograss (*Buchloe dactyloides*). *American Journal of Botany*, **79**, 207–215.
- Jousselin E, Rasplus J, Kjellberg F (2003) Convergence and coevolution in a mutualism: evidence from a molecular phylogeny of *Ficus*. *Evolution*, **57**, 1255–1269.
- Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623–2635.
- Maskas SD, Cruzan MB (2000) Patterns of intraspecific diversification in the *Piriqueta caroliniana* complex in southeastern North America and the Bahamas. *Evolution*, **54**, 815–827.
- Muenchow G, Delesalle V (1994) Pollinator response to male floral display size in two *Sagittaria* (Alismataceae) species. *American Journal of Botany*, **81**, 568–573.
- Pannell JR, Barrett SCH (1998) Baker's Law revisited: reproductive assurance in a metapopulation. *Evolution*, **52**, 657–668.
- Pannell JR, Charlesworth B (2000) Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London B*, **355**, 1851–1864.
- Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity I. Haploid locus. *Theoretical and Applied Genetics*, **90**, 462–470.
- Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.
- Rendell S, Ennos RA (2002) Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology*, **11**, 69–78.
- Renner SS, Ricklefs RE (1995) Dioecy and its correlates in the flowering plants. *American Journal of Botany*, **82**, 596–606.
- Renner SS, Won H (2001) Repeated evolution of dioecy from monoecy in Siparunaceae (Laurales). *Systematic Biology*, **50**, 700–712.
- Ridley HN (1930) *The Dispersal of Plants Throughout the World*. Reeve, Ashford, UK.
- Rohlf FJ (1997) *NTSYS-PC. Numerical Taxonomy and Multivariate Analysis System*, version 2.0. Applied Biostatistics Inc., Setauket, NY.
- Santamaría L (2002) Why are most aquatic plants widely distributed? Dispersal, clonal growth and small-scale heterogeneity in a stressful environment. *Acta Oecologia*, **23**, 137–154.
- Sarkissian TS, Barrett SCH, Harder LD (2001) Gender variation in *Sagittaria latifolia* (Alismataceae): is size all that matters? *Ecology*, **82**, 360–373.
- Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN Version 2.000: a Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Sculthorpe CD (1967) *The Biology of Aquatic Vascular Plants*. Edward Arnold, London, UK.
- Smith JG (1894) A revision of the North American species of *Sagittaria* and *Lophocarpus*. *Annual Report of the Missouri Botanic Garden*, **6**, 27–64.
- Smouse PE (1998) To tree or not to tree. *Molecular Ecology*, **7**, 399–412.
- Soltis DE, Soltis PS, Kuzoff RK, Tucker TL (1992) Geographic structuring of chloroplast DNA genotypes in *Tiarella-trifoliata* (Saxifragaceae). *Plant Systematics and Evolution*, **181**, 203–216.
- Stehlik I (2002) Glacial history of the alpine herb *Rumex nivalis* (Polygonaceae): a comparison of common phylogeographic methods with nested clade analysis. *American Journal of Botany*, **89**, 2007–2016.
- Stehlik I, Blattner FR, Holderegger R, Bachmann K (2002) Nunatak survival of the high alpine plant *Eritrichium nanum* (L.) Gaudin in the central Alps during the ice ages. *Molecular Ecology*, **11**, 2027–2036.
- Stuckey RL (1993) Phytogeographical outline of aquatic and wetland angiosperms in continental eastern North America. *Aquatic Botany*, **44**, 259–301.
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.
- Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.
- Weiblen GD (2000) Phylogenetic relationships of functionally dioecious *Ficus* (Moraceae) based on ribosomal DNA sequences and morphology. *American Journal of Botany*, **87**, 1342–1357.
- Wolfe KH, Li W, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences USA*, **84**, 9054–9058.
- Wooten JW (1971) The monoecious and dioecious conditions in *Sagittaria latifolia* L. (Alismataceae). *Evolution*, **25**, 549–553.
- Zamudio KR, Savage WK (2003) Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, **57**, 1631–1652.
- Zurawski G, Clegg MT, Brown AHD (1984) The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. *Genetics*, **106**, 735–749.

---

This work was part of Marcel Dorken's Ph.D. thesis research at the University of Toronto investigating the evolution and maintenance of monoecy and dioecy in *Sagittaria latifolia*. Dorken's main interests are in the evolution of strategies for sexual reproduction and vegetative growth, particularly in clonal plants. Spencer Barrett works on the evolution and ecology of plant reproductive systems and the genetics of plant colonization and migration.

---

