ISSN 0962-1083

VOLUME 17
NUMBER 22
NOVEMBER
2008

MOLECULAR ECOLOGY

Published by
Wiley-Blackwell

Molecular Ecology (2008) **17**, 4827–4844 doi: 10.1111/j.1365-294X.2008.03963.x

Clonal dynamics in western North American aspen (*Populus tremuloides*)

K. E. MOCK*, C. A. ROWE*, M. B. HOOTEN†, J. DEWOODY‡ and V. D. HIPKINS‡
*Wildland Resources Department, †Mathematics and Statistics Department, Utah State University, Logan, UT 84322, USA,
‡USDA Forest Service, PSW Research Station, National Forest Genetics Laboratory, 2480 Carson Road, Placerville, CA 95667 USA

Abstract

Clonality is a common phenomenon in plants, allowing genets to persist asexually for much longer periods of time than ramets. The relative frequency of sexual vs. asexual reproduction determines long-term dominance and persistence of clonal plants at the landscape scale. One of the most familiar and valued clonal plants in North America is aspen (*Populus tremuloides*). Previous researchers have suggested that aspen in xeric landscapes of the intermountain west represent genets of great chronological age, maintained via clonal expansion in the near absence of sexual reproduction. We synthesized microsatellite data from 1371 ramets in two large sampling grids in Utah. We found a surprisingly large number of distinct genets, some covering large spatial areas, but most represented by only one to a few individual ramets at a sampling scale of 50 m. In general, multi-ramet genets were spatially cohesive, although some genets appear to be fragmented remnants of much larger clones. We conclude that recent sexual reproduction in these landscapes is a stronger contributor to standing genetic variation at the population level than the accumulation of somatic mutations, and that even some of the spatially large clones may not be as ancient as previously supposed. Further, a striking majority of the largest genets in both study areas had three alleles at one or more loci, suggesting triploidy or aneuploidy. These genets tended to be spatially clustered but not closely related. Together, these findings substantially advance our understanding of clonal dynamics in western North American aspen, and set the stage for a broad range of future studies.

Keywords: aspen, clonality, *Populus tremuloides*, triploid, Utah, vegetative

Received 26 April 2008; revision received 31 July 2008; accepted 7 August 2008

Introduction

Clonality has long been an intriguing subject for ecologists and evolutionary biologists, challenging anthropocentric views of individuality and mortality. In plants, a clonal growth habit can produce a mosaic of sexually produced genets, each consisting of varying numbers of ramets. Genet patch size is determined by genet age, rate of ramet production, rate of root system expansion, extent and timing of sexual reproduction, and spatio-temporal disturbance regimes. The pattern and grain of this landscape-scale mosaic of genets, then, can give important clues regarding the relative importance of these factors. From a management perspective, conservation of clonal species can be particularly challenging, since spatial expansion or contraction of species coverage may be decoupled from increases or decreases in numbers of genets. The diversity of genets present in an area, however, is an important metric of evolutionary potential. Further, genet patch size in a dominant species may determine the scale and pattern of ecological responses to environmental factors. In the present study, we describe the landscape-scale pattern of genet distribution in an ecologically and economically important tree species, using the results to make inferences about clonal dynamics and natural history.

Quaking aspen (*Populus tremuloides*) is the most broadly distributed tree species in North America, extending from southern Mexico to northern Alaska and Newfoundland (Little 1971) and spanning an altitudinal range of 0–3700 m above sea level (a.s.l.) (Jones & DeByle 1985). Perhaps not

Correspondence: Karen Mock, Fax: (435)797-3796;
E-mail: karen.mock@usu.edu

surprisingly given its geographic range, seed mobility, and clonal nature, aspen is also one of the most genetically diverse tree species (Cheliak & Dancik 1982; Jelinski & Cheliak 1992; Kanaga *et al*. in press). In landscapes of intermountain western North America, aspen is frequently the dominant deciduous tree species, and aspen forest types are associated with disproportionately high levels of plant, bird, and butterfly biodiversity (Stohlgren *et al*. 1997a, b; Mills *et al*. 2000; Rumble *et al*. 2001; Simonson *et al*. 2001). Aspen are capable of reproducing sexually or vegetatively, although the proportion of these reproductive strategies varies across the species' range, likely due to variation in conditions suitable for seedling establishment (Kemperman & Barnes 1976).

As *Populus trichocarpa* is the first tree to have its genome completely sequenced (Tuskan *et al*. 2006), we anticipate that its close relative *P. tremuloides*, with its large geographic spatial extent and existence in a broad range of habitat types, will become an important system for the study of adaptive molecular evolution in natural populations. In order to effectively address evolutionary and adaptive processes in aspen, however, it is imperative to improve our understanding of reproductive modes and clonal dynamics in this species. The development of effective management strategies for aspen across its vast range also requires an accurate understanding of these dynamics on a regional basis.

Throughout its range, aspen ramets are regenerated primarily via vegetative reproduction (Barnes 1966; Schier 1973). Although viable seeds are produced (Faust 1936; Einspahr & Winton 1976), seedling establishment occurs only in a relatively narrow range of circumstances (Moss 1938; Pauley *et al*. 1963; McDonough 1979). This tendency toward vegetative reproduction is pronounced in the landscapes of intermountain western North America, leading to the establishment of large, multi-ramet clones, particularly in portions of the central and southern Rocky Mountains. One aspen clone in southern Utah has even been nominated as the largest organism in existence, and subsequently named 'Pando' (Grant *et al*. 1992). The size and spatial distribution of clones is thought to be a function of age, inter- and intraspecific competitive interactions, disturbance frequency, and biotic or abiotic habitat components (Pearson 1914; Barnes 1966; Namroud *et al*. 2005; Romme *et al*. 2005).

Seedling establishment, although it has been observed on many occasions in these landscapes, is considered to be an exceedingly rare event (Faust 1936; Kay 1993 and examples therein; Turner *et al*. 2003; Romme *et al*. 2005). From a management perspective, current sexual reproduction is considered to be a negligible contribution to genetic diversity. This combination of circumstances, along with large and well-demarcated clonal boundaries visible in intermountain western landscapes, has led to the general perception that aspen in these landscapes is dominated by a low number of ancient clones. The description of very large clones in the central and southern Rocky Mountain region by Kemperman & Barnes (1976), based on phenotypic traits of ramets, reinforced this perception and led to the hypothesis that aspen in more xeric portions of the western USA may represent ancient genets persisting vegetatively since the last glacial maximum, perhaps 8000–10 000 years ago (Barnes 1975; Kemperman & Barnes 1976; Cheliak & Dancik 1982; Jelinski & Cheliak 1992). Barnes (1975) suggested that some of the large clones in unglaciated portions of the intermountain western USA could have been established as early as the Pliocene. However, there have been very few landscape-scale studies of western aspen using genetic techniques to characterize clonal size, fragmentation, and juxtaposition (but see Turner *et al*. 2003; Hipkins & Kitzmiller 2004; Romme *et al*. 2005).

If the majority of aspen clones in intermountain western landscapes represent ancient genets that have persisted for thousands of years without appreciable sexual reproduction, we might expect the following two patterns. First, genets that have achieved a large spatial extent at some point during their existence will have undergone fragmentation as a result of small-scale disturbance and encroachment of other clones and vegetation types. Second, genets will have accumulated many somatic mutations, particularly at rapidly evolving neutral loci. These mutations should result in a strong spatial clustering of mutational variants arising from the same parent genet, a pronounced deviation from Hardy–Weinberg expectations in local populations, genotypic disequilibrium among loci, and pronounced differences in allele frequencies between regions that would otherwise be homogenized by gene flow via wind-dispersed pollen and seeds.

We used nuclear microsatellites to characterize patterns of landscape-scale genetic variation in two large study areas in Utah, USA. Specifically, our objectives in both study areas were: (i) to characterize the number, size, and juxtaposition of unique aspen genets, and (ii) to assess spatial patterns of genetic variation and similarity among genets and stands. We sought to determine whether the observed patterns were consistent with the hypothesis of a low number of ancient aspen clones dominating landscapes of the intermountain western USA.

Materials and methods

Study areas and sample collection

We chose two large study areas in Utah, USA: Swan Flats (SF) and Fish Lake (FL) (Fig. 1). The SF study area consisted of an approximately 2100-ha area containing 30 spatially distinct stands of varying sizes dominated by aspen, located 41°58′N, 111°28′W and spanning 2286–2560 m a.s.l.

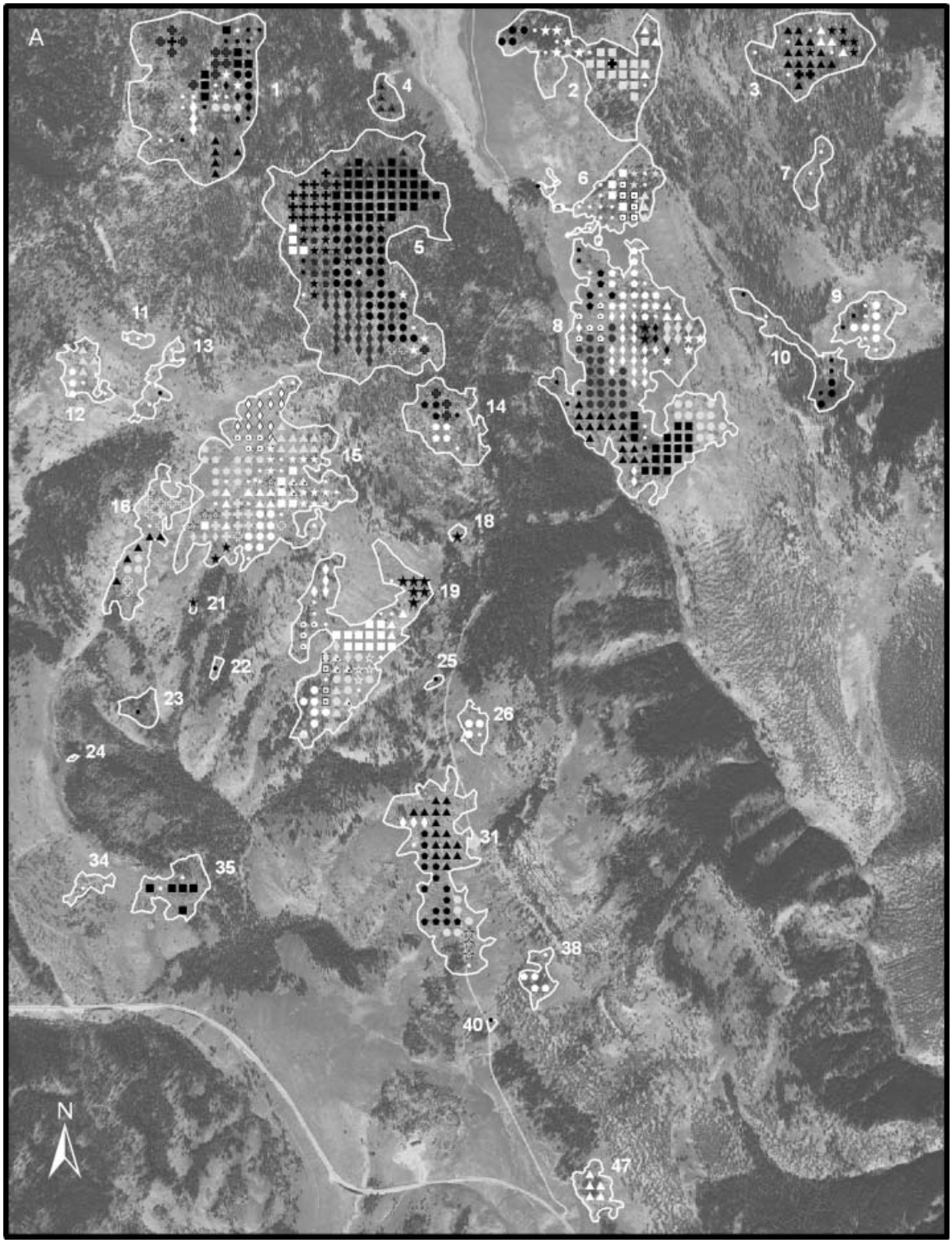

Fig. 1 Genet-specific map for (A) Swan Flats (SF) and (B) Fish Lake (FL) study areas. White lines in 1A represent stand boundaries. Unique symbols within stands represent unique genets at individual sampling points on a 50-m grid. Identical symbols in different stands represent different genets, with the exception of the '+' symbols, which represent genets found in multiple stands (Swan Flats only). Black and dark grey symbols denote genets with one or more loci containing triple alleles. In SF, overlain symbols represent sampling points where two samples were taken from different ramets representing different genets. In FL, an overstorey (upper symbol) and an understorey ramet (lower symbol) was sampled at each sampling point where they were available. In both maps, small dots are used as a general symbol to represent genets which were observed twice or fewer times among different sampling points.

Fig. 1 *Continued*

in altitude. Stands were defined as polygons of contiguous aspen coverage, delineated using aerial photographs available through the USDA Forest Service. The FL study area consists of approximately 360 ha located at 38°31′N, 111°45′W, ranging in altitude from 2700–2790 m a.s.l. and containing two distinct stands dominated by aspen (80 and 20 ha). The larger of these stands contains the Pando clone. The SF and FL study areas were separated by a linear distance of approximately 390 km. In both areas, sampling occurred on a 50-m grid which was overlain onto digitized, orthorectified aerial photographs of all stands in the form of Universal Transverse Mercator (UTM) coordinates. In SF, a sample was collected within a 5-m radius of each sampling point, if aspen was present. A leaf sample was obtained if accessible, but a small percentage of samples were represented by cambium samples. This sampling protocol likely led to an overrepresentation of smaller trees with accessible leaves. In FL, samples were collected in a similar manner, except that an overstorey tree and an understorey tree were sampled at each point.

Table 1 Microsatellite primer information and allelic observations among the Swan Flats (SF) and Fish Lake (FL) study areas. Dye labels were used on the 5′ end of forward primers in all cases

Locus	Linkage group	Dye label	Repeat motif	Allelic size range observed		No. of alleles observed		No. of genets with triple alleles	
				SF	FL	SF	FL	SF	FL
WPMS15*	V§	HEX	$[CCT]_n$	181–202	184–199	8	5	14	0
WPMS14*	V¶	HEX	$[CGT]_n$	191–239	200–233	17	9	32	4
WPMS20*	XIII¶	HEX	$[TTCTGG]_n$	204–235	204–234	10	7	13	2
GCPM970-1†	Unknown	FAM	$[TGC]_n$	119–134	119–134	6	4	10	0
PMGC2571†	X§	FAM	$[GA]_n$	84–120	88–114	18	10	27	2
PMGC433†	VIII§	TAMN	$[GA]_n$	178–208	182–204	15	6	19	3
PMGC576†	V§	HEX	$[GA]_n$	148–200	156–184	20	9	16	1
PTR14‡	III**	HEX	$[TGG]_n$	151–211	n/a	9	n/a	1	n/a
WPMS16*	VII§	HEX	$[GTC]_n$	172–205	n/a	7	n/a	16	n/a

Primer sequence sources are denoted as follows: *(Smulders *et al*. 2001); †(http://www.ornl.gov/sci/ipgc/ssr_resource.htm); ‡(Rahman *et al*. 2000); §(Tuskan *et al*. 2006); ¶(Gaudet *et al*. 2008), **(Hall *et al*. 2007).

In SF, samples were collected from 810 sampling points in September and October of 2004 and 2005. The majority of these samples were whole leaves collected directly from trees, placed into a paper envelope, and submerged in a silica gel desiccant at ambient temperature. At 15 of the sampling points, leaves were not accessible and a cambium sample was obtained by removing the bark from an area approximately 3 cm^2 using a clean, single-edged razor blade to strip the cambium layer. In FL, sample collection was performed in the same manner in September and November of 2006, and in June of 2007, except that both an understorey and an overstorey ramet were sampled at each sampling point when available. A total of 298 understorey and 263 overstorey ramets were sampled, 23 of which were cambium tissue.

Collection of genetic data

Leaf or cambium tissue was extracted using a QIAGEN DNEasy 96 Plant Kit. Microsatellite loci were amplified (Table 1) using 1.5 μL of template and final concentrations of 1.8 mM $MgCl_2$, 0.20 mM each dNTP, 0.25 μM of each primer, 0.3 U *Taq* polymerase, and 1× PCR buffer in a total reaction volume of 10 μL. Thermocycling conditions were 95 °C for 2 min, followed by 30 cycles (94 °C for 30 s; a primer-specific annealing temperature for 40 s; 72 °C for 50 s), and a final extension step of 72 °C for 10 min. Primer-specific annealing temperatures were 55 °C (WPMS14, WPMS15, WPMS16 and PTR14), 56 °C (PMGC2571), 57 °C (GCPM970–1), 59 °C (PMGC433), and 52 °C (PMGC576). Polymerase chain reaction (PCR) products were separated using an ABI 3730 DNA sequencer with a LIZ500 size standard and were analyzed using ABI GeneScan software. Chromatograms were scored using ABI GeneMapper software. All nine loci were amplified for SF samples, but only seven for FL samples (PTR14 and WPMS16 excluded). Evenly numbered samples from FL were amplified at the USDA National Forest Genetics Laboratory in Placerville, California, USA (NFGEL), following similar protocols and using the same equipment, with exceptions noted in DeWoody *et al*. in press. All other samples were extracted and amplified at Utah State University (USU). Amplification and scoring error rates within and between laboratories were assessed by including replicates aliquoted prior to PCR. At USU, 10–20 replicated templates were included for every locus in each of the study areas, and over these replicates, only a single mistyped allele was detected at a single sample for a single locus. At NFGEL, at least 40 templates were replicated per locus, and the overall error rate was 0.047 mistypes per allele. A total of 127 templates out of the 561 samples were replicated between laboratories for the FL samples, and no genotypic discrepancies were detected at any of the loci.

Genet assignment

Unique multilocus genotypes (MLGs) may be distinct because of somatic mutations or sexual reproduction. Because microsatellite loci are highly polymorphic and have a relatively rapid rate of mutation, and because of the known clonal nature of western aspen, we expected that many of the unique MLGs would be due to somatic mutations. The assignment of MLGs to distinct genets (i.e. genetic individuals originating from a single seed) is an important aspect of our study, and thus, we used input from several different approaches to optimize genet assignments.

Initially, we constructed a simple genetic distance matrix among all pairs of unique MLGs in each study area, with genetic distance calculated as the number of allelic

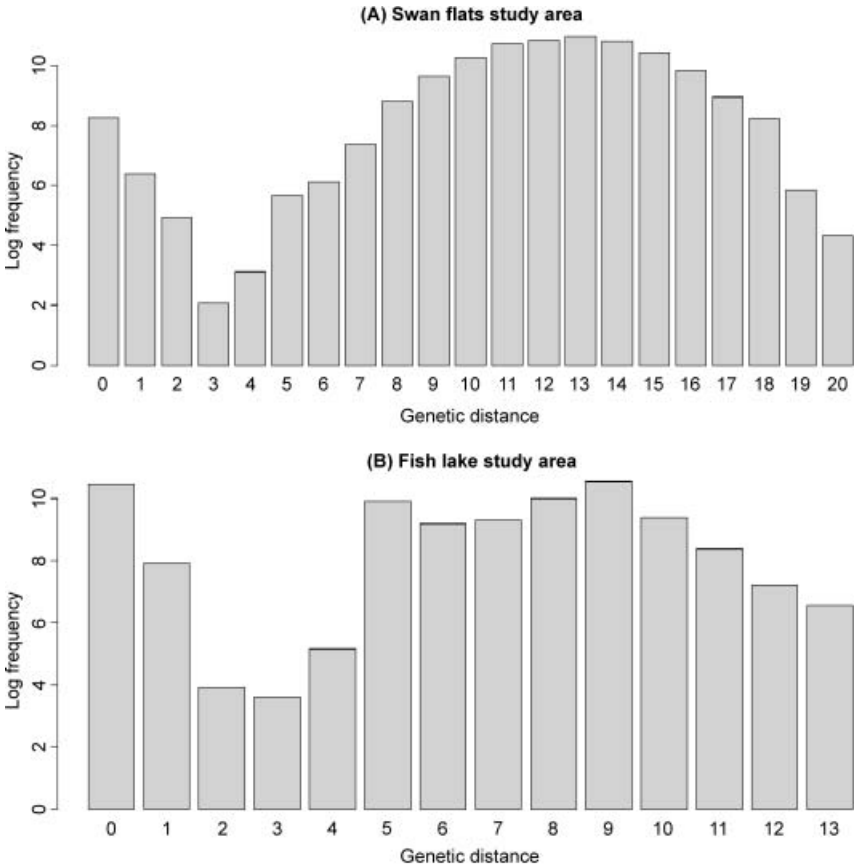

Fig. 2 Frequency distribution of mismatched alleles between genotypes of ramets sampled in the Swan Flats and Fish Lake study areas. Frequency distributions based on differences among unique multilocus genotypes (not shown) had similar patterns in both study areas.

mismatches between MLGs. Next, we used a classification technique known as partitioning around medoids (PAM; Kaufman & Rousseuw 1990) to identify clusters of genetically similar MLGs, based on the pairwise genetic distance matrix. In general, PAM identifies k optimal clusters by minimizing within-cluster dissimilarity based on medoids (in our application, the MLG in each cluster with the minimum distance to all other MLGs in that cluster). The optimal number of clusters is determined by computing a quality index for numerous values of k and choosing the k-value associated with the highest cluster quality. We used the silhouette index (Rousseuw 1987) as a measure of quality, which is computed as follows: for the i^{th} MLG, the average genetic distance from all other MLGs in the cluster (d_i) is calculated. The average genetic distance over all MLGs in other clusters (d_{ic}, where i = MLG and c = cluster) is also calculated. For each i, $d_{ic}min$ is the minimum distance between a particular MLG and medoids of all other clusters. The cluster which attains this minimum is called the 'neighbour of MLG$_i$' and can be thought of as the second best cluster for MLG$_i$. The silhouette values, s_i, are then defined as the difference between $d_{ic}min$ and d_i divided by the maximum of d_i and $d_{ic}min$. The quality index is the silhouette coefficient SC$_k$, which is the average of s_i over all MLGs in the data set. The coefficient SC$_k$ is calculated for numerous values of k, and the clustering arrangement yielding the largest SC$_k$ is chosen. It is important to note that SC$_k$ values may be close together for various k; thus, some post-clustering splitting or combining may be warranted based on other information. PAM clustering of MLGs provided a genetic distance-based set of clusters which could be compared with clustering patterns produced by other methods. Using this approach, we determined the optimal k and then made minor adjustments to the MLG membership in these clusters by considering: (i) the shape of the distribution of allelic mismatches between ramets (Fig. 2), and (ii) the spatial distribution of ramets within each genet.

A frequency distribution of allelic mismatches among all pairs of ramets was constructed for each study area to assess bimodality representing somatic vs. sexual genetic distances (Arnaud-Haond *et al*. 2007). No assumptions were made concerning ploidy levels when tallying mismatches; apparent differences in locus-specific ploidy levels were scored as a single mismatch. The bimodality of this distribution allowed us to assess the probability of pairs of MLGs being derived somatically or sexually.

We assessed the degree of undersplitting in the PAM clusters by first measuring the average Euclidian distance

among ramets within each putative genet, and dividing that figure by the number of ramets represented in the genet. PAM clusters were subdivided if: (i) they either had values falling outside the 99th percentile, or they had within-cluster genetic distances suggestive of sexual origin according to the mismatch distribution, and (ii) if subdivision produced clusters lacking these problems. We assessed the degree of genet oversplitting in the PAM clusters by screening pairs of genets in the lowest two genetic distance categories and determining whether any of these pairs: (i) were significantly nearer spatially than the other pairs of genets in those categories, and (ii) could be joined without creating a hypothesized genet exceeding the number of MLG mismatches considered to be due to sexual origin.

Finally, we made an independent assessment of hypothetical genet membership under the full range of mismatch thresholds using GenoType software (Meirmans & Van Tienderen 2004). We tracked the clustering of each individual MLG through the range of threshold levels possible in each study area. We noted the point at which hypothesized clusters containing each MLG collapsed into groups which (i) spanned genetic distances greater than those likely explained by somatic variation within the MLG mismatch distribution, and (ii) showed a marked increase in spatial disjunction. MLGs were assigned to the cluster just below the point of this collapse. We termed this the 'threshold' approach to genet assignment.

The genet memberships identified using the PAM and threshold-clustering approaches were compared. In the few cases where there were discrepancies, we used a somewhat subjective combination of ramet spatial distributional information and mismatch distributional probabilities to assign genet membership (see Results). To confirm that distinct clusters of genetically similar genets did not exist, final genet assignments were also subject to a principal coordinates analysis (PCoA) based on a matrix of pairwise genetic distances. This method for visualizing genetic space allows for the detection of over-splitting of genet memberships.

Within assigned genets, we estimated the probability of observing individual MLGs under the assumption of random mating, excluding loci with allelic variants presumed to be of somatic origin (Parks & Werth 1993; Sydes & Peakall 1998). We expected these probabilities to be quite low if genets had been correctly identified.

In FL, we had the opportunity to assess the pattern of genetic variance in overstorey vs. understorey trees. If sexual reproduction in this study area is an uncommon but ongoing event, one might expect to find a greater number of genets among younger (presumably smaller) ramets. Under this scenario, new genets would be continually added to the youngest cohort of ramets in the stand. Alternatively, if sexual reproduction is episodic, one might expect that genets are lost over time between episodes of successful seedling establishment, and that more genets would be found among older ramets. Under this scenario, the older ramets would represent this earlier episode of sexual reproduction, and younger ramets would represent a subset of that diversity. This question was assessed by comparing the number of genets that were represented only by understorey ramets to the number of genets that were represented only by overstorey ramets.

Analysis of genetic data

Assessment of Hardy-Weinberg (HW) equilibrium and genotypic disequilibrium (GD) was complicated by the observation of three alleles in many of the genets (see Results). If these results are due to triploidy or a high level of aneuploidy, it is possible that there are many undetected alleles in these samples, preventing the accurate assessment of HW equilibrium and GD. Therefore, in order to assess HW equilibrium and GD, we used a subset of genets (see below) in which no triple alleles were observed. In SF, this subset consisted of 136 genets out of 189 total; in FL, this subset consisted of 51 genets out of 60 total (see Results). While this does not eliminate the possibility of allelic dominance in our data set, this possibility should be reduced with an apparently diploid subset. HW equilibrium was assessed using a Markov chain approach as implemented in GenePop software to estimate per-locus probabilities (Guo & Thompson 1992; Raymond & Rousset 1995). GD was assessed using the contingency table approach in GenePop, with a Markov chain approach to estimate probabilities for co-occurrence of alleles at each pair of loci. Results for both HW equilibrium and GD assessment were interpreted using the Bonferroni correction for multiple tests. In both study areas, the multilocus probability of identity (Waits *et al*. 2001) was estimated, using the subset of genets with no triple allele observations and using GenAlEx software (Peakall & Smouse 2006).

Assessment of genetic differentiation among study areas

Under the ancient clones hypothesis, we would expect allele frequency differences among study areas to be quite distinct. This expectation presumes that most genetic variation within study areas would be due to the accumulation of random somatic mutations in a small number of remaining genets. Unfortunately, because of the possibility of a high proportion of triploid genotypes in both populations, estimation of allele frequencies is potentially problematic. As with assessment of HW equilibrium and GD, we estimated allele frequencies in each area using a subset of genets with no triple alleles observed. This data subset was then used to estimate the

extent of genetic subdivision (F_{ST}) among study areas using an analysis of molecular variance (AMOVA; Excoffier *et al*. 1992) as implemented in GenAlEx software (Peakall & Smouse 2006), using 1000 permutations to assess significance. To make the data between study areas comparable in this analysis, we only included data from the seven loci common to both study areas. Using this data set, genetic differences among study areas were further assessed using frequency-based assignment testing (Paetkau *et al*. 1995, 2004), using the 'leave one out' approach, with GenAlEx software (Peakall & Smouse 2006).

We also estimated allele frequencies in the two study areas using a simple count of alleles encountered, including genotypes with triple alleles at some loci. These data were used to compare the study areas with respect to the number of alleles observed. Because the number of samples can influence the number of alleles observed, we used a simple rarefaction procedure available in PAST (Hammer *et al*. 2001) software to estimate the number of alleles that would have been observed in the larger study area (SF) had our sampling been limited to the level in the smaller study area (FL).

Results

Genet assignment

The frequency distributions of allelic mismatches for both study areas were distinctly bimodal, with the frequency of mismatches putatively due to somatic mutations being substantially smaller than (and somewhat obscured by) the frequency of mismatches putatively due to sexual reproduction (Fig. 2). Predictably, the genetic distance separating these two distributions was greater in SF, where genotypes consisted of two additional microsatellite loci.

Overall, genet assignment ambiguities were a very small proportion of the total number of genets identified. From the 812 ramets sampled in SF, 248 unique multilocus genotypes (MLGs) were detected, with PAM clustering identifying 191 genets. These results corresponded to 10 discrepancies with the threshold methodology, which were resolved by splitting genets assigned by PAM (one case), merging genets separated by PAM (three cases), splitting genets according to the PAM clusters which were originally joined by the threshold approach (three cases), and joining genets according to the PAM clusters which were originally split by the threshold approach (three cases). This refinement produced a total of 189 genets. From the 561 ramets sampled in FL, 83 MLGs were detected and 60 genets were defined. The four discrepancies identified between methods were resolved by joining PAM-proposed clusters of MLGs (two cases) and splitting the PAM-proposed clusters (two cases). PCoA plots summarizing the pairwise genetic distance matrix of individual genets produced a single cloud of points for each of the study areas (not shown), and did not indicate any further splitting or grouping of genets. Genet locations and membership are shown in Fig. 1.

Within putative SF genets, the average genotype probability (P_{gen}) for genotypes differing by one to five loci were (i.e. excluding non-identical loci from the calculation): (i) 9.83×10^{-10}; (ii) 3.64×10^{-7}; (iii) 3.18×10^{-6}; (iv) 8.79×10^{-5}; and (v) 1.67×10^{-4}. In FL, the average P_{gen} of MLGs within genets differing by one allele was 2.73×10^{-5} (range 1.06×10^{-8} to 1.57×10^{-4}). Probabilities of MLGs containing triple alleles at one or more locus within genets were not estimated due to dominance effects at potentially triploid loci. These low genotype probabilities suggest that the MLGs within our assigned genets are unlikely to be of independent sexual origin.

The number of mutational variants within genets was low, even in the largest genets (Table 2). Within SF genets, a maximum of seven MLGs were observed within a genet, with a maximum of six mutational steps between genotypes (Table 2). In FL, a maximum of seven MLGs were observed in a single genet, with only two mutational steps separating genotypes (Table 2) In the genet with the largest number of ramets, the 'Pando' clone in FL, we detected only six mutational variants, all only a single mutational step away from the dominant MLG. However, the correspondence between genet size (number of ramets) and either number of MLGs or maximum genetic distance within genets was not pronounced.

A number of the MLGs in both study areas included three alleles at various loci (Table 3). The phenomenon of triple alleles was observed at all loci, although no MLG had triple alleles at all loci. There was a higher proportion of genets with triple alleles in SF (28%) than FL (15%). This difference was not simply due to the analysis of more loci in SF; the additional two loci in SF genotypes only resulted in the detection of one additional genet with triple alleles. The frequency of triple alleles at particular loci was loosely correlated with the number of observed alleles per locus (SF $r^2 = 0.417$; FL $r^2 = 0.307$) (Tables 1 and 3). In SF, 75.5% of the genets with triple alleles at one locus had triple alleles at multiple loci. This effect was less pronounced in FL (33.3%). There was a tendency for genets with triple alleles to be represented by more ramets than genets without triple alleles (Fig. 3). In SF, genets with triple alleles ($n = 53$) had an average of 10.7 ramets, while genets without triple alleles ($n = 136$) had an average of 1.8 ramets. In FL, genets with triple alleles ($n = 9$) had an average of 46 ramets, while genets without triple alleles ($n = 51$) had an average of two ramets. A heteroscedastic *t*-test indicated that the mean number of ramets in genets with triple alleles was significantly larger than in genets without triple alleles ($P = 0.02$). In SF, there was a notable stand-level clustering of genets with triple alleles (Fig. 1A). Genetic distances among triple-

Table 2 Numbers of multilocus genets (MLGs), maximum genetic distance measured as the number of allelic mismatches, and number of ramets included in each genet containing more than a single MLG. Genet numbers are arbitrary and for laboratory reference only

	Genet ID	No. of MLGs in genet	Maximum genetic distance within genet	No. of ramets in genet
Swan Flats study area	141	7	2	43
	53	2	1	39
	105	2	1	25
	165	4	3	24
	173	4	3	24
	16	2	1	15
	5	2	1	13
	14	4	5	13
	8	5	6	12
	162	4	2	12
	40	2	1	11
	139	2	1	11
	189	2	2	11
	1	3	2	9
	83	2	1	7
	171	2	1	7
	146	2	1	6
	183	3	2	6
	18	2	1	5
	27	3	2	5
	42	2	2	5
	99	2	1	5
	21	2	5	4
	45	2	1	4
	111	2	1	4
	163	3	2	4
	167	2	1	4
	176	2	1	4
	182	2	1	4
	9	2	1	3
	54	2	3	3
	94	2	4	3
	185	2	1	3
	15	2	1	2
	23	2	3	2
	37	2	1	2
	59	2	5	2
	121	2	2	2
	138	2	3	2
Fish Lake study area	1	7	2	256
	39	3	2	71
	13	6	2	62
	12	4	2	7
	53	2	1	6
	48	2	1	4
	58	2	1	4
	18	2	1	3
	29	2	1	2

allele genets within stands were not significantly different from genetic distances between those chosen randomly from among stands ($P = 0.99$ based on 1000 random permutations of triple-allele genets across all stands, with the same frequency of occurrence as that observed in the data).

Spatially, genets tended to be cohesive at our sampling scale (Fig. 1). In SF, 164 genets (87%) consisted of ramets which were continuously adjacent at the 50- or 100-m scale. Of the 25 instances of fragmentation, most were ≤ 300 m apart within stands or adjacent stands, but four gaps of

Table 3 Distinct stands (polygons) sampled in the Swan Flats (SF) and Fish Lake (FL) study areas. In both areas, sampling locations were on a 50-m grid. In SF, a single sample was collected at each location when aspen was present. In FL, an overstorey ramet and an understorey ramet were sampled when present. The number of unique genets detected among all samples is presented (see text for details on genet assignment), along with the number of genets with one or more triple alleles present. The total number of genets detected in SF is less than the numerical total across stands because of the occurrence of some genets in more than one stand. Area was not provided for stands represented by only one or two ramets

Study area	Stand ID	Area (ha)	No. of ramets sampled	No. of genets detected	No. of genets with triple alleles (%)	
SF	1	23.6	57	23	13	56
SF	2	12.4	34	10	3	30
SF	3	7	32	9	3	33
SF	4	0.2	4	1	1	100
SF	5	49.1	164	15	11	74
SF	6	11.5	37	18	1	6
SF	7	—	2	1	0	0
SF	8	53.1	152	24	11	46
SF	9	1.7	12	5	1	2
SF	10	5.1	8	5	4	80
SF	11	—	2	1	0	0
SF	12	2	9	3	0	0
SF	13	1.7	7	7	1	14
SF	14	2.2	12	5	3	60
SF	15	36.6	121	31	1	3
SF	16	5	19	5	1	20
SF	18	—	1	1	1	100
SF	19	24.5	69	18	1	6
SF	21	—	1	1	1	100
SF	22	—	1	1	1	100
SF	23	—	1	1	0	0
SF	24	—	1	1	0	0
SF	25	—	1	1	1	100
SF	26	0.2	4	2	0	0
SF	31	13.2	44	7	2	29
SF	34	—	1	1	0	0
SF	35	1.5	7	3	1	33
SF	38	0.88	5	2	0	0
SF	40	—	1	1	1	100
SF	47	0.38	5	1	0	0
Total SF	**30**	**252.2**	**812**	**189**	**53**	**28**
FL	1	71.25	465	53	8	15
FL	2	13.25	96	7	1	14
Total FL	**2**	**84.5**	**561**	**60**	**9**	**15**

over 500 m were observed. In all four cases, these were single ramets spatially separated from multiple stems of the remainder of the genet.

In FL, there were no instances of genet recurrence across stands. Again, most genets were contiguous at the 50-m scale, although there were two instances where ramets at a single sampling point were disjunct (500–600 m) from a cluster of ramets in the same genet. There was a pronounced tendency for genets with single or low numbers of ramets to be along the uphill edge of the stand (Fig. 1B).

In order to assess whether patterns of spatial and genetic distances within each study area were due to reproductive processes (e.g. seed dispersal patterns) or oversplitting of MLGs arising from the same propagule during genet assignment, we calculated the minimum genetic distance (numbers of allelic mismatches) and the minimum Euclidean distance (m) between ramets in all pairs of genets. The resulting plots of these data indicate that there is a weak but nonsignificant tendency for genetically similar genets to be spatially close on average in first two genetic distance categories (Fig. 4). Mantel tests for correlation between genetic distance and spatial distance between MLGs (SF: $P = 0.187$, FL: $P = 0.507$) indicate no significant spatial/genetic correlation at either site.

In FL, for the 239 sampling points where both an understorey and overstorey ramet was sampled, 203 of these pairs had identical MLGs; 16 of these pairs (6.7%) belonged to the same genet but had different MLGs, and 20 (8.4%) belonged to different genets. There were fewer genets represented among understorey ramets than among overstorey ramets: five genets were represented only by understorey ramets (2.3% of the total understorey ramets sampled), and 11 genets were represented only by overstorey ramets (3.8% of the total overstorey ramets sampled).

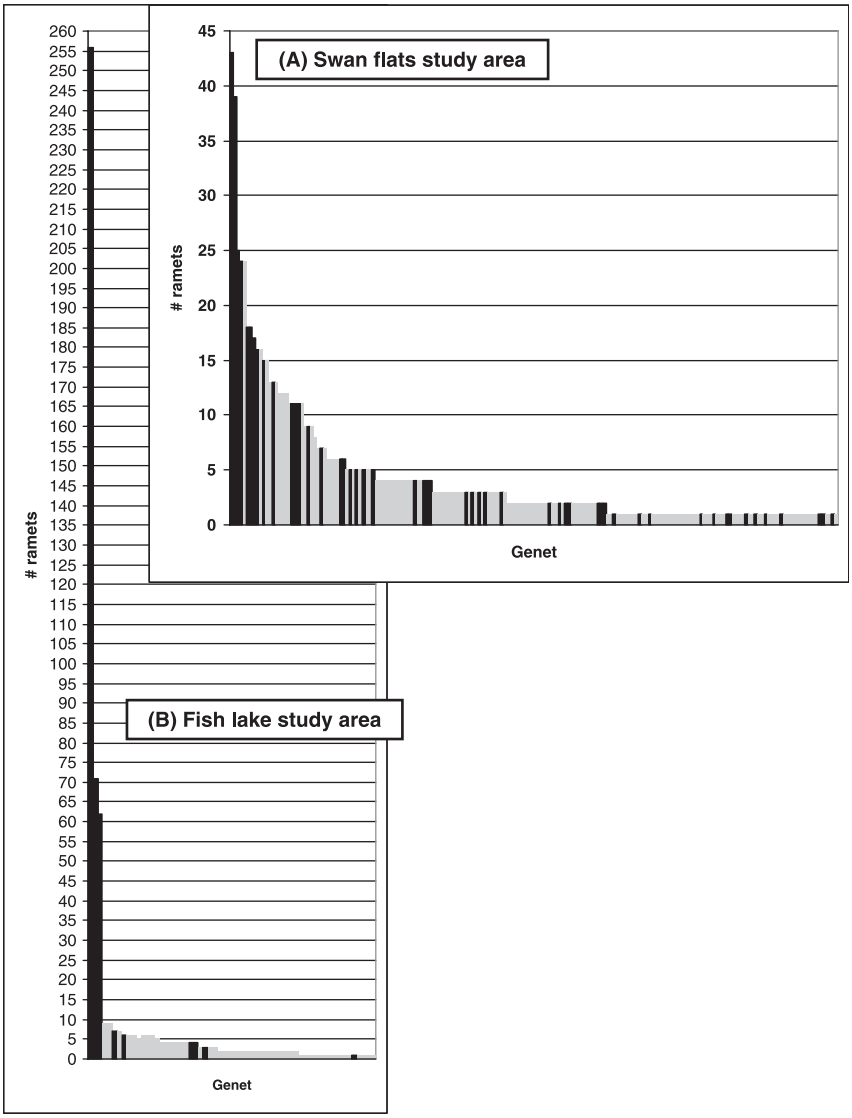

Fig. 3 Distribution of sampled ramets per genet detected in the Swan Flats (SF) and Fish Lake (FL) study areas. Vertical bars represent individual genets. Black bars represent ramets with one or more instances of triple alleles at a locus. In FL, up to two ramets per sample point (overstorey and understorey) were collected, inflating the total number of ramets by comparison to SF.

Analysis of genetic data

The data subset lacking genets with triple alleles was used to assess HW equilibrium and GD. Two of the loci (WPMS20 and PMGC576) showed significant heterozygote deficiencies relative to HW equilibrium expectations in SF, but no HW deviations were found in FL. There were no instances of significant GD in either of the study areas. The probability of identity based on nine loci for SF was 3.8×10^{-10} for random individuals and 3.1×10^{-4} for full siblings. The probability of identity based on seven loci for FL was 1.1×10^{-6} for two random individuals and 0.001 for two full siblings.

Assessment of genetic differentiation among study areas

Allelic size ranges in FL were a subset of those observed in SF (Table 4). The two most common alleles at each locus were identical across study areas with a single exception (second most common allele at PMGC 2571 differed). The number of alleles observed per locus was consistently higher in SF than in FL, an expected result given the differences in spatial area and sample numbers. However, even with rarefaction to compensate for unequal sample sizes, SF had greater allelic richness across all loci (Table 4).

Allele frequencies calculated using only genets lacking triple alleles were quite similar to allele frequencies calculated using all samples (data not presented); the two most common alleles at each locus were identical between these data sets, and rare allele frequencies showed only minor deviations. Thus, it is unlikely that the exclusion of genets with triple alleles had a major effect on our results.

Genetic differentiation among study areas was low: F_{ST} estimated using an AMOVA approach was 0.028 ($P = 0.001$), with variance among study areas explaining only 3% of the

Table 4 Microsatellite allele sizes and allelic richness (A) observed in each study area (SF and FL), including allele frequencies (Freq) adjusted via rarefaction to numbers observed in FL (A_r). Most common two alleles in each study area are in bold, italicized. n/o, not observed

Locus											
WPMS15			WPMS14			WPMS20			GCPM970-1		
Allele	SF Freq	FL Freq	Allele	SF Freq	FL Freq	Allele	SF Freq	FL Freq	Allele	SF Freq	FL Freq
202	0.008	n/o	239	0.002	n/o	235	0.003	n/o	134	0.067	0.125
199	0.023	0.008	236	0.005	n/o	234	0.015	n/o	131	0.010	n/o
196	0.212	0.125	233	0.007	0.040	228	0.010	n/o	128	0.062	0.050
193	***0.245***	***0.500***	230	0.063	0.153	224	0.061	0.107	125	0.026	n/o
190	0.077	0.008	227	0.027	n/o	222	0.020	0.016	122	***0.585***	***0.692***
187	0.120	n/o	224	0.124	0.065	218	0.087	0.025	119	***0.250***	***0.133***
184	***0.258***	***0.358***	221	0.071	n/o	216	***0.425***	***0.500***	**$A(A_r)$**	**6(5.7)**	**4**
181	0.059	n/o	218	0.015	n/o	212	***0.279***	***0.205***			
$A(A_r)$	**8(7.6)**	**5**	215	0.022	0.065	210	0.015	0.008			
			212	***0.244***	***0.242***	204	0.084	0.139			
			209	0.124	0.113	**$A(A_r)$**	**10(8.8)**	**7**			
			206	***0.159***	***0.242***						
			203	0.051	0.040						
			200	0.061	0.040						
			197	0.005	n/o						
			194	0.002	n/o						
			191	0.017	n/o						
			$A(A_r)$	**17(14.0)**	**9**						

Locus												
PMGC2571			PMGC433			PMGC576			PTR14		WPMS16	
Allele	SF Freq	FL Freq	Allele	SF Freq	FL Freq	Allele	SF Freq	FL Freq	Allele	SF Freq	Allele	SF Freq
120	0.002	n/o	208	0.020	n/o	200	0.010	n/o	196	0.741	193	0.411
114	0.069	***0.107***	204	0.003	0.008	198	0.005	n/o	157	0.177	187	0.241
112	0.017	0.057	201	0.028	n/o	190	0.010	n/o	151	0.021	181	0.211
110	0.010	n/o	198	0.008	0.008	188	0.003	n/o	202	0.021	175	0.089
108	0.010	n/o	196	0.126	0.211	186	0.003	n/o	199	0.013	199	0.041
106	0.012	n/o	194	0.018	n/o	184	0.003	0.008	187	0.013	205	0.005
104	0.005	n/o	192	0.038	n/o	182	0.013	n/o	211	0.005	172	0.003
102	0.059	0.107	190	0.020	n/o	176	0.010	n/o	166	0.005	**A**	**7**
100	0.052	0.033	189	0.005	n/o	174	0.008	0.008	163	0.003		
99	0.010	n/o	188	0.144	0.114	173	0.030	n/o	**A**	**9**		
98	***0.274***	***0.467***	186	***0.350***	***0.415***	172	0.005	0.050				
96	0.148	0.074	185	0.005	n/o	170	0.020	0.017				
94	0.059	0.041	184	0.003	n/o	168	***0.297***	***0.314***				
92	0.022	0.008	182	***0.184***	***0.244***	166	0.071	0.066				
90	***0.160***	0.082	178	0.050	n/o	164	0.028	0.017				
88	0.072	0.025	**$A(A_r)$**	**15(12.1)**	**6**	162	0.190	0.124				
86	0.010	n/o				160	0.003	n/o				
84	0.007	n/o				156	***0.256***	***0.397***				
$A(A_r)$	**18(15.2)**	**10**				154	0.015	n/o				
						148	0.020	n/o				
						$A(A_r)$	**20(14.8)**	**9**				

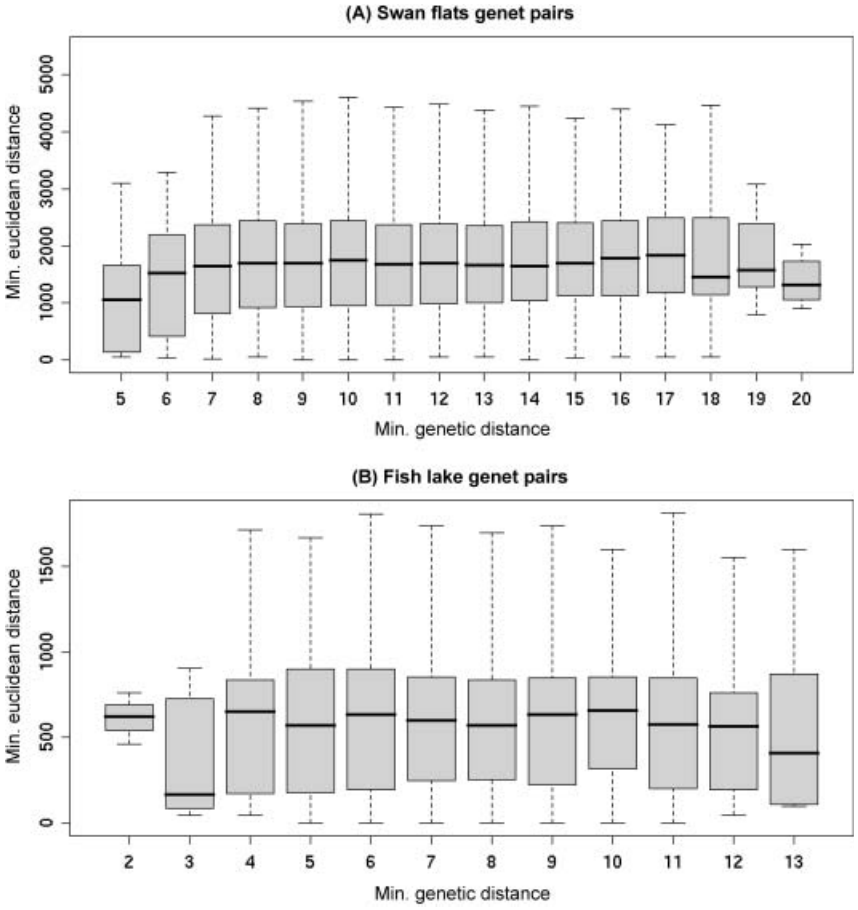

Fig. 4 Minimum genetic distance (number of allelic mismatches) and minimum Euclidean distance (m) between ramets in all pairs of genets in the (A) Swan Flats and (B) Fish Lake study areas. Dashed lines represent the range of observations, shaded areas represent interquartile range (25th and 75th quantiles), and the horizontal bar represents the median.

total variance. Assignment tests detected structure among study areas, although 27% of the genets from SF and 16% of the genets from FL were misassigned.

Discussion

Testing the ancient clones hypothesis

We systematically genotyped aspen ramets in two large study areas in Utah, USA, both representing some of the more xeric conditions found in the central and southern Rocky Mountain region. Under the hypothesis that such landscapes are populated by a low number of ancient clones, we expected to find a low number of genets, large genet size for most genets, a high incidence of somatic mutations within genets at microsatellite loci, spatial fragmentation of genets, and pronounced genotypic disequilibrium and Hardy–Weinberg disequilibrium among genets. Our results were generally contrary to these expectations.

We found a large number of genets in both study areas, despite our sampling density of 50 m. In SF, we found 189 distinct genets out of the 812 ramets sampled, 72 of these being single observations. In FL, we found 60 distinct genets out of the 561 ramets sampled (over 319 sampling points), 28 of these being single-point genet observations. These numbers far exceeded our expectations and those of local foresters (J. Long, personal communication), and suggest that an even greater number of genets would be found if sampling density was increased. These findings also suggest that for at least a portion of the landscape, clone sizes in these study areas may be comparable to those typically described for more mesic, eastern North American and northern Rocky Mountain landscapes (Barnes 1969; Steneker 1972). However, some genets were quite large (Table 2, Fig. 3), consistent with the findings of Kemperman & Barnes (1976). In both study areas, the distribution of clone sizes was skewed (Fig. 3), with a low number of spatially dominant clones and a large number of spatially small clones. This trend was particularly pronounced in FL, which was dominated by three large clones.

In FL, clonal boundaries and the existence of the celebrated 'Pando' clone had previously been described using morphological techniques by Kemperman & Barnes (1976) and confirmed elsewhere using genetic tools (DeWoody *et al*. in press). We were able to confirm the approximate boundaries of the second large clone in FL initially deline-

ated by Kemperman & Barnes (1976), and identified a third large clone within the same stand as Pando. Thus, morphological criteria, when meticulously applied (Barnes 1969), can be useful in identifying clonal boundaries when morphologies are distinct from adjacent clones. However, given the prevalence of spatially small clones in our study, we caution that clonal richness may be greatly underestimated based on morphological criteria.

We did find somatic mutations in many genets, although the number of these was quite low overall (Table 2). In general, the somatic variants within genets represented a single observation or a small number of variants relative to the dominant MLG, in contrast to the expectation for ancient clones with many neutral microsatellite variants that would have accumulated over several thousands of years of cell division. This finding suggests that even the largest clones (e.g. the 'Pando' clone) may be of relatively young age, although a molecular clock for these markers does not exist and evolution of microsatellite markers in general is poorly understood. It is possible that a portion of these variants could be due to laboratory error, resulting in even fewer 'real' somatic variants. It is also possible that some genetically similar clusters of somatic variants were inappropriately assigned to distinct genets, although if this was a frequent problem, we would expect to see more pronounced spatial proximity among genetically similar genets, as well as HW disequilibrium and GD (see below).

Under the hypothesis of ancient clones dominating the landscape, one would expect that large clones persistent over long periods of time would become physically fragmented by other aspen genets or other vegetation types following various disturbances. Again, this is not the general pattern that we found in our study areas: genets tended to be continuously adjacent at the 50-m scale. Our findings are generally consistent with a 'phalanx' growth form (Lovett Doust 1981; Cheplik 1997) whereby clones are spatially cohesive (Barnes 1966), although a finer scale might reveal more interdigitation of genets (Namroud *et al*. 2005).

There were, however, six instances (four in SF, two in FL) of apparent genet fragmentation resulting in ramets separated by more than 500 m. Given the low probability of identical, sexually produced genotypes within putative genets, these ramets probably represent isolated remnants of genets that once covered larger contiguous geographic areas. The general pattern of genet cohesion with a few examples of fragmentation was also reported for aspen in the Sierra Nevada of western North America using allozyme data (Hipkins & Kitzmiller 2004).

Five of the six highly fragmented genets had triple alleles at one or more loci; a number greater than expected, given the overall proportion of genets with triple alleles in this study area (28%). This finding could be the result of triploid clones expanding rapidly and being at greater risk of fragmentation simply because of their size, or it could be that these represent older clones with higher rates of somatic aneuploidy. In the latter case, one would also expect to see greater within-clone genetic distances in fragmented clones. However, in the six cases of highly fragmented genets, four of these genets included only a single MLG, while the other two genets had multiple MLGs but with only a maximum genetic distance of two mutational steps. Thus, although our sample size was quite small and the genetic distances quite low overall, we did not see a correspondence between genet fragmentation and within-genet genetic distance. The issue of triploidy vs. aneuploidy is further discussed below.

Under the ancient clones hypothesis, if the genetic variation among assigned genets in our study areas was the result of many accumulated somatic mutations among a few genets instead of being generated via sexual reproduction, we would expect that these genets would show marked GD as well as HW disequilibrium. Instead, we found that none of the pairs of loci showed evidence of GD in either study area, and that the genet genotypes did not deviate from expected HW proportions at all loci in FL, and in all but two loci in SF. The two loci showing HW disequilibrium showed a significant deficit of heterozygotes, in contrast to the expected direction of deviation in the case of accumulated somatic mutations under the infinite alleles model (see review in Cheliak & Dancik 1982). Given the proportion of genets with triple alleles observed in both study areas, it is possible that frequencies of common alleles may have been underestimated (a dominance effect), resulting in HW genotypic ratio expectations that were skewed towards heterozygotes. This may have been responsible for the heterozygote deficiency observed in SF. The observation of HW equilibrium and a lack of GD in this study would be difficult to explain under a model whereby most genetic variation was due to accumulated somatic mutations in a low number of ancient, sexually established genets.

Additionally, the accumulation of somatic mutations within ancient genets would be expected to result in a strong correlation between genetic distance and spatial distance among MLGs. While we found this to be the case within our assigned genets, we found that this was not generally true between genets (Fig. 4). There was a nonsignificant tendency in SF for genets in the lowest two genetic distance categories to also be spatially closer than other genet pairs. This pattern in SF could be the result of a few genets being oversplit by our clustering techniques or it could be a half-sibling clustering phenomenon.

At a larger spatial scale, we would expect under the ancient clones hypothesis that allele frequencies and allelic size ranges from one geographic region to another would be quite different if most genetic variance was due to the accumulation of somatic mutations under an infinite alleles model. In contrast to these expectations, we found

that allele frequencies and size ranges were quite similar between our two study areas, which are approximately 390 km apart. The same most common two alleles were identical in all but one locus, which had the same most common three alleles in both populations (Table 4). Furthermore, the population subdivision between the two study areas was quite low but significant ($F_{ST} = 0.028$, $P = 0.001$), with a high proportion of individuals assigned to the wrong study area (SF 27%; FL 16%) in assignment tests. These findings suggest a high degree of gene flow between these two study areas on an evolutionary timescale, consistent with wind dispersal of *Populus* seeds and pollen. While it is possible that selective constraints on particular allele sizes could contribute to size homoplasy in microsatellites between study areas, gene flow is a more parsimonious explanation than somatic mutation accumulation given the similarity of allele frequency distributions between these areas (Table 4).

Overall, these data suggest that sexual reproduction among aspen in our study areas is a stronger contributor to standing genetic diversity at the population level than the accumulation of somatic mutations. Sexual reproduction may well be rare and episodic, but is not a negligible process in these landscapes. This finding is particularly striking because our study areas represent the portion of the species' range which is thought to be most dominated by large, ancient clones, and wherein sexual reproduction would have been limited or absent for thousands of years (Barnes 1966; Kemperman & Barnes 1976; Mitton & Grant 1996). Jelinski & Cheliak (1992) suggested that episodes of successful establishment of seedlings, although rare, could present 'windows of opportunity' for recruitment of aspen seedlings based on their findings of high levels of genetic diversity in Alberta, Canada. Our results suggest that even in these xeric study environments, these 'windows of opportunity' may be more important, and perhaps more frequent, than previously recognized.

Clone/stand age

The ability to estimate an absolute or even a relative age for aspen clones has remained elusive, although ages can be bound at the low end by clonal size and ramet age. Rapidly evolving molecular markers such as microsatellites may prove useful for this purpose, although mutation rates, evolutionary patterns, and evolutionary constraints are likely to be locus-specific variables. Landscape histories in our study areas are not well enough known to make inferences about maximum clonal ages, and even relative clonal ages cannot be assessed with confidence since current clonal sizes may be the result of many different histories. We did assess correspondence between genet size and maximum genetic distance among ramets within a genet, but were unable to draw conclusions from this data since there were so few somatic variants found and since genets were generally represented by just a few ramets.

We were surprised to find so few mutational variants in the Pando genet; out of 256 individual ramet samples, there were only six mutational variants, all single-step variants of the dominant MLG in this genet. This finding suggests a relatively young age even for this very large clone, as somatic microsatellite variants are expected to occur and be propagated even during the lifetime of a single tree (O'Connell & Ritland 2004) or during short-term sucker propagation (Rahman & Rajora 2001).

Landscape clonal dynamics

In both sample locations, there was a large range of genet sizes, suggesting differences in the timing of establishment and/or rates of expansion/decline among genets. In FL, there was a particularly interesting pattern of clustering of small genets on the north and northeast (uphill) edges of the stand (Fig. 1). This pattern could be indicative of successful seeding event(s) following the establishment and spread of the Pando clone, or discontinuities in disturbance histories or environmental conditions.

The two study areas differed in several respects. SF had a more continuous distribution of genet sizes (Fig. 3), greater allelic richness, more pronounced spatial fragmentation of genets, and a higher proportion of genets with triple alleles. These differences may simply be due to our sampling of a larger spatial area and more stands in SF, along with the existence of one enormous clone (Pando) in FL. Alternatively, these observations may reflect fundamental differences in landscape disturbance histories and/or seedling establishment patterns and timing.

Ploidy

One of the interesting findings in our data was the frequent occurrence of triple alleles in both study areas, particularly in the larger genets (Fig. 3). Triple alleles could be due to laboratory error, amplification of multiple loci in a diploid genome, aneuploidy resulting in duplication of some loci in some ramets/genets, or polyploidy. We can rule out laboratory error as the primary explanation because of our low error rates with replicated samples, as well as the consistent amplification of triple alleles across all ramets of genets with triple alleles. Large regions of the *P. trichocarpa* genome are known to have undergone genome duplication (Tuskan *et al*. 2006), and thus, it is possible that locus-specific gene duplication may explain our findings. However, triple alleles were observed at every locus in at least one genet, and we never observed four alleles at any locus. Further, the microsatellite loci represented several linkage groups (Table 1), although duplication among these linkage groups does exist (Tuskan *et al*. 2006). Aneuploidy, a

common form of somatic variance in plants, is a possibility, and without chromosome counts, this cause cannot be clearly distinguished from full triploidy. We did note that triple alleles were particularly common among spatially fragmented clones, a result consistent with the expectation that older clones will have accumulated more somatic aneuploidy and also have an increased probability of spatial fragmentation due to small-scale disturbances. We found four instances of individual ramets which were identical to others in the genet except for additional third alleles, and in each case, the difference was due to a single locus. This is a pattern that might be expected with somatic aneuploidy, but could also be explained by somatic mutation in a triploid genet where triploidy had not been detected at other alleles.

Our finding that the largest clones in both study areas tended to be those with triple alleles is also consistent with expectations for full triploids. Notably, one of the genets in which we observed triple alleles was the giant 'Pando' clone in FL. Triploid genets are likely to have reduced fertility (Ramsey & Schemske 1998), be more vegetatively vigorous, and have a more rapid stem growth rate than diploid genets (Van Buijtenen *et al*. 1958). The presence of a high proportion of triploid genets in our study areas, and their large size, would be consistent with the suggestion of Eckert (2002) that sexuality in clonal plants may be particularly reduced in marginal habitats. Triploidy has been previously documented in Utah aspen via chromosome counting (Every & Wiens 1971), but we have, to date, been unable to harvest suitable tissue for chromosome counting in our study areas.

Another striking finding with respect to genets with triple alleles is that they tended to cluster within stands in SF (Fig. 1A). This spatial clustering is not likely to be the result of common ancestry, as triple-allele genets within stands were not more similar genetically than triple-allele genets among stands. Instead, we suggest that this clustering may be due to particular landscape features or histories selecting for triploid genets over time, perhaps between episodes of disturbances favouring sexual reproduction. Alternatively, the accumulation of somatic aneuploidy over time may also explain the tendencies for genets with triple alleles to have large numbers of ramets and be more spatially fragmented than genets without triple alleles. Under this scenario, spatial clustering of genets with triple alleles could be due to stand age. Chromosome counts in these genets would resolve these issues, although collection of appropriate tissue under field conditions can be technically challenging.

Implications for ecological studies

Regardless of the potential effect of ploidy issues, the enormous variance in clonal sizes within stands, in conjunction with the very high heritability of most phenotypic traits within aspen (Kanaga *et al*. in press), could seriously confound ecological studies in aspen forests. If clonal structure is not accounted for in experimental designs, studies on aspen regeneration, disease susceptibility, ungulate herbivory, carbon dynamics, age structure, etc. could be making inferences about only particular clones instead of about the species in general. We recommend that study plots in natural areas take clonal boundaries (and perhaps ploidy) into account. On a more positive note, the existence of genets containing many ramets and potentially covering various ecological gradients or contrasts presents an excellent opportunity to study trait plasticity and ecological trade-offs in energy investment.

Acknowledgements

We would like to acknowledge the field and technical assistance provided by Ryan Shaw, Jer Pin Chong, Eric O'Neill, Megan Kanaga, Sean Kelly, Jon Mock, Richie Gardner, Brian and Anna Cardall, Masako Niwa, Justin DeRose, Olivia Salmon, Ricardo Hernandez, Kristin Motz, Robert C. Saich, and Ron Ryel. Sherm Jensen of Whitehorse Associates assisted with sampling map construction, and Robert Campbell and Dale Bartos offered many helpful suggestions during the study. The authors would also like to thank Adele Cutler for statistical suggestions and advice, and Paul Wolf and James Long for many helpful comments on the manuscript. Funding for the project was provided by the USU ADVANCE programme, the USU Community University Research Initiative, the USU Office of Extension, and the USFS Rocky Mountain Research Station in Logan, Utah.

References

Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA (2007) Standardizing methods to address clonality in population studies. *Molecular Ecology*, **16**, 5115–5139.

Barnes BV (1966) The clonal growth habit of American aspens. *Ecology*, **47**, 439–447.

Barnes BV (1969) Natural variation and delineation of clones of *Populus tremuloides* and *P. grandidentata* in northern Lower Michigan. *Silvae Genetica*, **18**, 130–142.

Barnes BV (1975) Phenotypic variation of trembling aspen in western North America. *Forest Science*, **21**, 319–328.

Cheliak WM, Dancik BP (1982) Genic diversity of natural populations of a clone forming tree *Populus tremuloides*. *Canadian Journal of Genetics and Cytology*, **24**, 611–616.

Cheplik GP (1997) Responses to severe competition stress in a clonal plant: differences between phenotypes. *Oikos*, **79**, 581–591.

DeWoody J, Rowe C, Hipkins VD, Mock KE (in press) Pando lives: molecular genetic evidence of a giant aspen clone in central Utah. *Western North American Naturalist*.

Eckert CG (2002) The loss of sex in clonal plants. *Evolutionary Ecology*, **15**, 501–520.

Einspahr DM, Winton LL (1976) *Genetics of Quaking Aspen*. US Department of Agriculture, US Forest Service Research Paper WO-25, Washington D.C.

Every AD, Wiens D (1971) Triploidy in Utah aspen. *Madroño*, **21**, 138–147.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitocondrial DNA restriction sites. *Genetics*, **131**, 479–491.

Faust ME (1936) Germination of *Populus grandidentata* and *P. tremuloides* with particular reference to oxygen consumption. *Botanical Gazette*, **97**, 808–821.

Gaudet M, Jorge V, Paolucci I, Beritognolo I, Scarascia Mugnozza G, Sabatti M (2008) Genetic linkage maps of *Populus nigra* L. including AFLPs, SSRs, SNPs, and sex trait. *Tree Genetics and Genomes*, **4**, 25–36.

Grant MC, Mitton JB, Linhart YB (1992) Even larger organisms. *Nature*, **360**, 216.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.

Hall D, Luquez V, Garcia VM, St. Onge KR, Jansson S, Ingvarsson PK (2007) Adaptive population differentiation in phenology across a latitudinal gradient in European aspen (*Populus tremula, L.*): a comparison of neutral markers, candidate genes and quantitative traits. *Evolution*, **61**, 2849–2860.

Hammer Ø, Harper DAT, Ryan PD (2001) PAST: palaeontological statistics software package for education and data analysis. *Palaeontologia Electronica*, **4**, 1–9.

Hipkins VD, Kitzmiller JH (2004) Genetic variation and clonal distribution of quaking aspen in the central Sierra Nevada. *Transactions of the Western Section of the Wildlife Society*, **40**, 32–44.

Jelinski DE, Cheliak WM (1992) Genetic diversity and spatial subdivision of *Populus tremuloides* (Salicaceae) in a heterogeneous landscape. *American Journal of Botany*, **79**, 728–736.

Jones JR, DeByle NV (1985) Morphology. In: *Aspen: Ecology and Management in the Western United States* (eds DeByle NV, Winokur RP) pp. 11–18. General Technical Report RM-119, US Department of Agriculture, US Forest Service, Rocky Mountain Forage and Range Experiment Station, Fort Collins, Colorado.

Kanaga MK, Ryel RJ, Mock KE, Pfrender ME (in press) Quantitative genetic variation in morphological and physiological traits within a quaking aspen (*Populus tremuloides*) population. *Canadian Journal of Forest Research*, **38**, 1690–1694.

Kaufman L, Rousseeuw PJ (1990) *Finding Groups in Data*. John Wiley & Sons, New York.

Kay CE (1993) Aspen seedlings in recently burned areas of Grand Teton and Yellowstone National Parks. *Northwest Science*, **67**, 94–104.

Kemperman JA, Barnes BV (1976) Clone size in American aspens. *Canadian Journal of Botany*, **54**, 2603–2607.

Little EL Jr (1971) *Checklist of United States Trees (Native and Naturalized).* US Department of Agriculture, Agriculture Handbook 541. Washington D.C.

Lovett Doust L (1981) Population dynamics and local specialization in a clonal perennial. *(Ranunculus repens).* I. The dynamics of ramets in contrasting habitats. *Journal of Ecology*, **69**, 743–755.

McDonough WT (1979) Quaking aspen-seed germination and early seedling growth. US Department of Agriculture, Forest Service Research Paper INT-234, Ogden Utah.

Meirmans PG, Van Tienderen PH (2004) GenoType and GenoDive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Mills TR, Rumble MA, Flake LD (2000) Habitat of birds in ponderosa pine and aspen/birch forest in the Black Hills, South Dakota. *Journal of Field Ornithology*, **71**, 187–206.

Mitton JB, Grant MC (1996) Genetic variation and the natural history of quaking aspen. *Bioscience*, **46**, 25–31.

Moss EH (1938) Longevity of seed and establishment of seedlings in species of *Populus*. *Botanical Gazette*, **49**, 529–542.

Namroud M-C, Park A, Tremblay F, Bergeron Y (2005) Clonal and spatial genetic structures of aspen (*Populus tremuloides* Michx.). *Molecular Ecology*, **14**, 2969–2980.

O'Connell LM, Ritland K (2004) Somatic mutations at microsatellite loci in western red cedar (*Thuja plicata*: Cupressaceae). *Journal of Heredity*, **95**, 172–176.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347–354.

Paetkau D, Slade R, Burdens M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.

Parks JC, Werth CR (1993) A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, **80**, 537–544.

Pauley SS, Johnson AG, Santamour FS Jr (1963) Results of aspen screening tests. I. Seed sources of quaking aspen (*P. tremuloides* Michaux). *Schoological of Forestry, University of Minnesota, Minnesota Forestry Notes*, **136**, 1–2.

Peakall R, Smouse P (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pearson GA (1914) The role of aspen in the reforestation of mountain burns in Arizona and New Mexico. *Plant World*, **17**, 249–260.

Rahman MH, Rajora OP (2001) Microsatellite DNA somaclonal variation in micropropagated trembling aspen *Populus tremuloides*. *Plant Cell Reports*, **20**, 531–536.

Rahman MH, Dayanandan S, Rajora OP (2000) Microsatellite DNA markers in *Populus tremuloides*. *Genome*, **43**, 293–297.

Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploidy formation in flowering plants. *Annual Reviews in Systematics*, **29**, 467–501.

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

Romme WH, Turner MG, Tuskan GA, Reed RA (2005) Establishment, persistence, and growth of aspen (*Populus tremuloides*) seedlings in Yellowstone National Park. *Ecology*, **86**, 404–418.

Rousseuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, **20**, 53–65.

Rumble MA, Flake LD, Mills TR, Dykstra BL (2001) Do pine trees in aspen stands increase bird diversity? In: *Sustaining Aspen in Western Landscapes: Symposium Proceedings: 13–15 June 2000, Grand Junction, Colorado* (eds Sheppard WD *et al.*), pp. 185–191. US Department of Agriculture, Forest Service Rocky Mountain Research Station Publication RMRS-P-18.

Schier GA (1973) Origin and development of aspen root suckers. *Canadian Journal of Forest Research*, **3**, 45–53.

Simonson SE, Opler PA, Stohlgran TJ, Chong GW (2001) Rapid assessment of butterfly diversity in a montane landscape. *Biodiversity and Conservation*, **10**, 1369–1386.

Smulders MJM, Van der Schoot J, Arens P, Vosman B (2001) Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). *Molecular Ecology Notes*, **1**, 188–190.

Steneker GA (1972) *Size and suckering of trembling aspen* (P. tremuloides Michaux) *clones in Manitoba*. PhD thesis. University of Michigan, Ann Arbor, Michigan.

Stohlgren TJ, Chong GW, Kalkhan MA, Schell LD (1997a) Multiscale sampling of plant diversity: effects of minimum mapping unit size. *Ecological Applications*, **7**, 1064–1074.

Stohlgren TJ, Coughenour MB, Chong GW *et al*. (1997b) Landscape analysis of plant diversity. *Landscape Ecology*, **12**, 155–170.

Sydes MA, Peakall R (1998) Extensive clonality in the endangered shrub *Haloragodendron lucasii* (Haloragacae) revealed by allozymes and RAPDS. *Molecular Ecology*, **7**, 87–93.

Turner MG, Romme WH, Reed RA (2003) Postfire aspen seedling recruitment across the Yellowstone (USA) landscape. *Landscape Ecology*, **18**, 127–140.

Tuskan GA, DiFazio S, Jansson S *et al*. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, **313**, 1596–1604.

Van Buijtenen JP, Joranson PN, Einspahr DW (1958) Diploid versus triploid aspen as pulpwood sources with reference to growth, chemical, physical, and pulping differences. *Technical Association of the Pulp and Paper Industry*, **41**, 170–175.

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249–256.