Molecular Ecology (2008) **17**, 4897–4911 doi: 10.1111/j.1365-294X.2008.03962.x

Can clone size serve as a proxy for clone age? An exploration using microsatellite divergence in *Populus tremuloides*

D. ALLY,*† K. RITLAND† and S. P. OTTO*
Department of Zoology, 6270 University Blvd, †Department of Forest Sciences, Forest Sciences Center, 2424 Main Hall, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

Abstract

In long-lived clonal plants, the overall size of a clone is often used to estimate clone age. The size of a clone, however, might be largely determined by physical or biotic interactions, obscuring the relationship between clone size and age. Here, we use the accumulation of mutations at 14 microsatellite loci to estimate clone age in trembling aspen (*Populus tremuloides*) from southwestern Canada. We show that the observed patterns of genetic divergence are consistent with a model of increasing ramet population size, allowing us to use pairwise genetic divergence as an estimator of clone age. In the populations studied, clone size did not exhibit a significant relationship with microsatellite divergence, indicating that clone size is not a good proxy for clone age. In *P. tremuloides*, the per-locus per-year neutral somatic mutation rate across 14 microsatellite loci was estimated to lie between 6×10^{-7} (lower bound) and 4×10^{-5} (upper bound).

Keywords: clone age, clone size, coalescence, microsatellite divergence, *Populus tremuloides*, somatic mutation

Received 3 April 2008; revision received 10 June 2008; accepted 15 June 2008

Introduction

Although it is relatively straightforward to measure lifespan in unitary organisms, in long-lived clonal plants, it is complicated to determine how old is old. Clonal plants produce new physiological modules (ramets) by asexual processes (Orive 1995; Gardner & Mangel 1997), so that no one ramet carries the age signature of the entire clone (the genet). To circumvent the issue, researchers have used the area covered by a clone as a proxy for clone age (Vasek 1980; Steinger *et al*. 1995; Escaravage *et al*. 1998; Reusch *et al*. 1998; Tyson *et al*. 1998). Using this method, one clonal plant, a Creosote bush, *Larrea tridentata*, was estimated to be 11 700 years old (Vasek 1980). But using area as a proxy for clone age implicitly assumes that the area covered by a clone is a perfect circle, whose radius increases linearly with time (Fisher 1937; Lewis 2000). It is unlikely that most clones grow in a perfect circle because heterogeneity in physical resources necessary for growth can constrain the total size and/or the shape of a clone. Furthermore, events such as disease or fire can cause a clone to contract. Indeed, in species such as trembling aspen (*Populus tremuloides*), the entire aboveground portion of a stand can die back (Hogg *et al*. 2002; Frey *et al*. 2004), and the clone can regenerate itself from root stock (Maini 1965a, b; Peet 1981). Such events result in the uncoupling of clone age and size (Hughes 1984; Tanner 2001).

The only alternative that has been used to assess clone age is radiocarbon dating. This method can be used in species where the oldest portion of the clone can be identified. This is possible in some species where the ramets remain attached to the genet, such that the central core of the clone is usually the oldest, e.g. *Larrea tridentata* (Harris & Lovell 1980; Vasek 1980; Steinger *et al*. 1995). For many clonal plant species, however, not only is it difficult to identify the oldest ramet, but ramet turn-over makes it possible for clones to be older than their parts.

Trembling aspen is a prime example of a species in which it is difficult to determine clone age. Each ramet of *P. tremuloides* develops a large distal and a small proximal root (Fig. 1). The lateral roots give rise to suckers that can develop into independent physiological units, which in turn produce their own lateral roots, which can give rise to additional ramets (Schier 1973). Because individual aspen trees can die and the root system is large and complex,

Correspondence: Dilara Ally, Fax: (604) 822-2416; E-mail: ally@zoology.ubc.ca

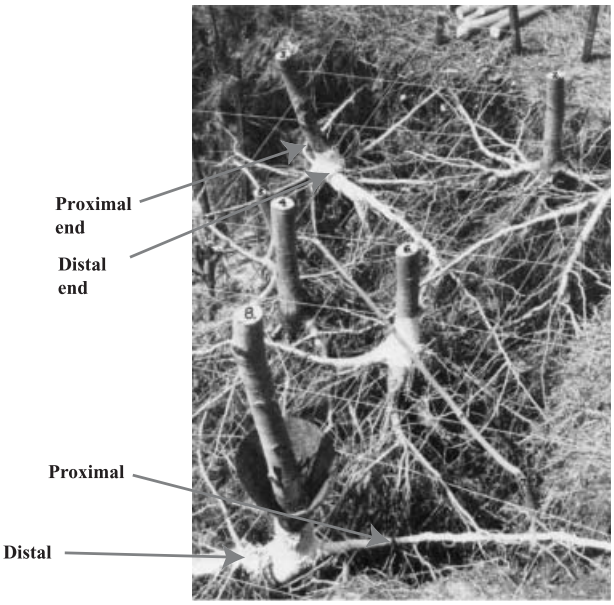

Fig. 1 A hydraulic excavation of part of a *Populus tremuloides* clone shows the pattern of interconnection among ramets. The root system of a ramet has a distal and a proximal end. The distal end, the portion of the root that is enlarged, is characteristic of the parent root from which the ramet arose. Photograph taken from DeByle & Winokur (1985).

there is no straightforward way to identify the oldest component.

During the repeated cycle of ramet birth, development, and death, somatic mutations arise (Klekowski 1997), making the clone a mosaic of mutant and non-mutant cell lineages (Fig. 2). This offers an alternative method to identify clone age. If mutations occur at a constant rate, age can be estimated by the amount of neutral genetic divergence that has accumulated within a clone. Here, we use microsatellite divergence as an 'ontogenetic molecular clock,' to estimate clone age in *P. tremuloides*.

Microsatellites are ideal for estimating genetic divergence within a clone for three reasons. First, microsatellite loci typically have high mutation rates, on the order of 10^{-3} to 10^{-6} per sexual generation (Chakraborty *et al*. 1997; Schug *et al*. 1997; Schlotterer *et al*. 1998; Vazqeuz *et al*. 2000; Udupa & Baum 2001; Thuillet *et al*. 2002; Ellegren 2004; O'Connell & Ritland 2004). Higher mutation rates provide greater resolution for determining the age of a clone. In addition to mutations altering the number of repeats, we can also observe the loss of heterozygosity at polymorphic microsatellite loci, where processes such as mitotic recombination and gene conversion cause homozygotes to appear within otherwise heterozygous clones. Second, length variants are easy to score using standard methods [polymerase chain reaction (PCR) and sequencing gels). Third, dinucleotide repeats, which are the focus of our study, are virtually always in non-coding sequences, because changes in repeat length within coding sequences would cause missense mutations. As a consequence, microsatellites are more likely to evolve neutrally than if we were to track sequence changes within protein-coding genes.

In addition to replication-dependent mechanisms like mitotic recombination, recent evidence suggests that replication-independent mechanisms may lead to the accumulation of mutations over time. First, taxa whose seeds undergo longer periods before germination have more mutations at neutral loci per unit time (Whittle 2006). Second, no differences were found in molecular evolutionary rates between taxa with higher and lower pollen:ovule ratios (Whittle 2006). This suggests that replication-independent events are more important or that cell divisions during reproduction had little impact on the number of mutations. Third, when compared to young tissue, senescent tissue showed a higher frequency of point mutations largely due to a longer exposure to oxidative stress (Pla *et al*. 2000). These data taken together suggest that mutations can accumulate on a per unit time basis. Thus, microsatellite divergence captures both mutations that accrue via cell division and through time, making it a good proxy for clone age.

Estimating clone age using genetic divergence

To relate molecular divergence to clone age, we first specify a model by which the ramets of a clone arise and die over time. This model can then be used to predict how mutations should accumulate within a clone. Here, we contrast two different demographic models, which have often been used to interpret molecular divergence patterns. The first demographic model is one of constant ramet population size (as in the classic Wright–Fisher model), while the second demographic model is one of population growth, where the population has expanded so recently that most lineages trace their ancestry back to a point in time near the common ancestor of the entire clone (exhibiting a 'star phylogeny' (Slatkin & Hudson 1991)). In the following, we describe the predicted patterns of molecular divergence under these two demographic models. We can then use these predictions to determine the most appropriate model and to estimate clone age for the populations of trembling aspen that we have studied.

Constant population size

If a clone has occupied a patch for a long period of time and if physical or biotic constraints prevent patch expansion, the total effective number of diploid ramets, N_e, may be approximately stable over time, with the number of ramet births matching the number of deaths. We have used the total effective number of diploid ramets, N_e, to account for variation in ramet production per parent ramet. Many populations exhibit higher variation in reproductive success

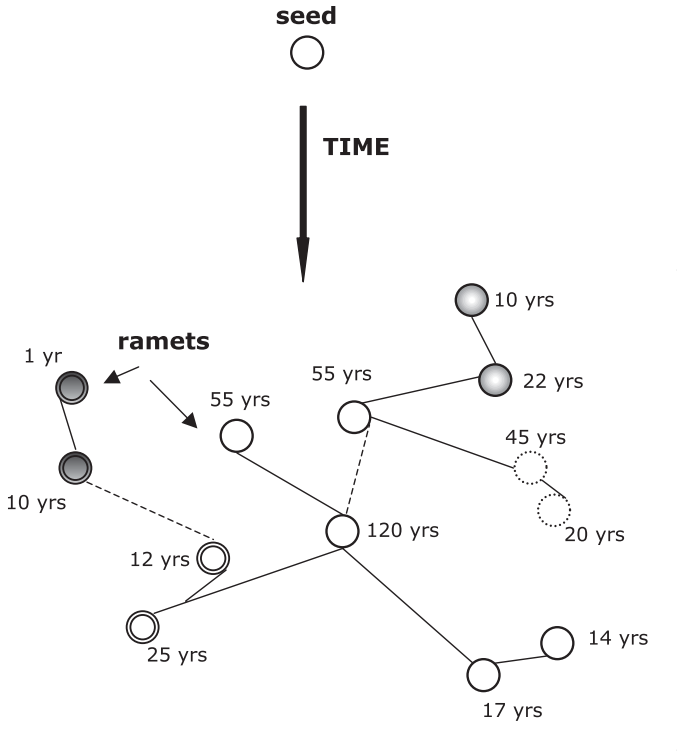

Fig. 2 Clonal growth through time from a common cellular ancestor, the seed. As a genet grows, there are repeated bouts of ramet birth and death, possibly accompanied by mutations, resulting in a clone becoming a genetic mosaic. We represent the different somatic mutations carried by ramets by the shape and shading of circles. Without knowing the original genotype of the seed, it is impossible to determine the true ancestral lineage within a clone. For simplicity's sake, we designate the most frequent genotype as the ancestral lineage (open circle). The genet is a collection of ramets of varying physiological ages (numbers beside the circles). In *Populus tremuloides*, ramets often remain connected to each other via lateral roots (solid lines), but sometimes these root connections degrade (dashed lines), fragmenting the individual.

than expected based on random allocation of offspring to parents; this variation causes the population to evolve more like an idealized population whose size is N_e than one whose size is the census size, N. The expected time to the most recent common ancestor of any two lineages (the coalescence time) is N_e ramet generations (Hudson 1990). Note that we use the term 'lineage' to refer to the ancestral ramets that gave rise to a diploid ramet; because the patch is clonal, this lineage describes the pattern of descent of both alleles carried by the ramet. If mutations occur at rate 2μ per diploid ramet (μ per allele), then the probability that a mutation has accumulated in either of the two lineages since they split from their common ancestor is approximately $\theta = 4N_e\mu$.

Two statistics can be used to measure the level of genetic divergence within a clone and can be related to θ. The first is the average number of pairwise differences per locus for the k^{th} clone:

$$\pi_k = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} s_{ij} \qquad \text{(eqn 1)}$$

(Slatkin 1996; Wakeley 2007), where n is the number of sampled ramets and s_{ij} is the number of differences between ramet i and ramet j, averaged across loci. The second statistic, S_k, is the proportion of loci exhibiting a polymorphism within the clone. For a population of constant size, π_k and $S_k \Big/ \sum_{i=1}^{n-1} \frac{1}{i}$ are both expected to equal $\theta = 4N_e\mu$ (this is the basis of Tajima's 1989 neutrality test; Hudson 1990). Thus, if trembling aspen clones have been stable in size, we would predict a 1 : 1 relationship between π_k and $S_k \Big/ \sum_{i=1}^{n-1} \frac{1}{i}$.

In this case, however, genetic divergence, measured by either π_k or S_k, would not estimate the age of a clone; instead, it would estimate the effective population size of ramets ($\theta = 4N_e\mu$).

Growing population size

If a clone has been expanding since it was first established from a seed, then coalescent events are much more likely to happen early in the clone's history (when few ramets were present), causing the relationships among ramets to form a 'star phylogeny' (Slatkin & Hudson 1991). In this case, the coalescence of any two lineages would trace back to a point in time soon after the clone was established, T_{CCA} (time to the common cellular ancestor). If we compare two lineages that have been accumulating mutations at rate 2μ per diploid ramet (μ per allele), then the probability that a mutation has accumulated in either of the two ramet lineages (in either allele at a locus) will be approximately $4\mu T_{CCA}$. This probability can be estimated by the average number of pairwise differences, π_k, observed per locus.

Similarly, the probability that a mutation has accumulated at a locus in any of the lineages leading to a sample of n ramets (thus causing the locus to appear polymorphic within the clone) should be $2n\mu T_{CCA}$, which can be estimated from the observed proportion of polymorphic loci, S_k. Thus, if trembling aspen clones have been growing over time, we would predict a 1 : 1 relationship between π_k and $2S_k/n$. If such a relationship is observed, then either π_k or S_k could be used to estimate T_{CCA}, and hence, provide an estimate for the age of the clone. As discussed in Supporting Information – Section A, π_k provides an estimate for T_{CCA} that is more robust to departures from a star-like phylogeny than S_k (e.g. if some sampled ramets happen to be closely related). We thus focus on the use of π_k as an estimator of clone age.

In this study, we test the assumption that clones grow as a perfect circle and also explore the relationship between clone size and genetic divergence in two different populations of trembling aspen to determine whether size can be used as a proxy for clone age. The fact that clone size is quite variable in the western half of *Populus tremuloides*' range, varying from 1.5 to 43.2 ha (Kemperman & Barnes 1976), provides us with an ideal opportunity to examine the relationship between size and genetic divergence at microsatellite loci.

Materials and methods

Data collection

We collected foliage for microsatellite analysis from 871 trees of *Populus tremuloides* sampled from two different populations: Riske Creek, British Columbia (RC) and Red Rock, Waterton Lakes National Park, Alberta (RR). Trees in both populations were sampled in two ways: on the perimeter of a patch and systematically along two or three 30–50 m transects within the patch. On average 30–50 individuals were sampled per patch. No tree less than 1.5 m in height was sampled, and only patches separated by at least 1 km of terrain lacking aspen trees were used. Trees on the perimeter and along transects were physically mapped using both a measuring tape and a handheld Global Positioning System (GPS) unit. In addition to collecting foliage and GPS measurements on all transect trees, we measured ramet height and diameter at breast height (d.b.h.). The heights of trees were measured using a clinometer and tape measure, and diameters at breast height were measured with a steel d.b.h. tape. Finally, an increment core was taken from a sample of ramets belonging to each patch.

Within the grasslands at Riske Creek, *P. tremuloides* grows in discontinuous, clearly defined, patches across the landscape with the occasional contiguous forested area; 17 patches were assayed, involving 719 ramets. Within the montane region at Red Rock, trees grew sparingly and were interspersed among coniferous trees along a ridge. Every attempt was made to sample in the same manner as in the grasslands, but in most cases all the trees in the designated patch were sampled. A total of 152 ramets in 10 patches were sampled at Red Rock.

Criteria for somatic mutations

A preliminary survey of 100 potential microsatellite loci were assayed in samples of *P. tremuloides* to identify a subset of primers for which fragments could be reliably re-amplified, that yielded fragments with the least stutter and mispriming, and that were polymorphic across clones. Fourteen microsatellite loci were subsequently used to identify potential somatic mutations in all sampled ramets (see Tables 1 and 2). Eight of the loci are known to map to seven of the 19 linkage groups in the *Populus* genome (Cervera *et al*. 2001; Yin *et al*. 2004). For the other six loci, we have no mapping information. In all PCRs, we used a hi-fidelity Ampli*Taq* DNA polymerase (Roche). For each 10 μL reaction, we used 1.0 μL of 10× PCR buffer (Roche) ($MgCl_2$: 1.5 mM) with 0.2 mM dNTP (to ensure specificity, no additional $MgCl_2$ and dNTP were added). The maximum number of PCR cycles was 30. In cases where PCR failed, we performed several additional reactions with altered PCR conditions. Failed PCRs may or may not be caused by mutations in the flanking primer regions, resulting in null alleles (no PCR fragment). A failed PCR reaction was not counted as a somatic mutation, and the locus was excluded from the analysis for that clone.

PCR samples and a negative control were loaded on a LI-COR 4200 automated sequencer, and microsatellite products were detected by M13-tailed or end-labelled primers. DNA fragments were sized using the software analysis program SAGA (LI-COR) and LI-COR size standards (catalogue no. 4200–44), labelled with IRD dye. We reanalysed our data on two separate occasions, both with the software and by visual inspection.

In order to discriminate between a true somatic mutation and PCR error, we performed two additional PCRs on ramets carrying potential mutations. To maintain a consistent sizing across different gels, we also performed two additional PCRs on a clonemate in which no mutations were observed. Throughout, we used high quality DNA template obtained from freshly opened buds and young leaves, reducing the potential for allele dropouts, allelic competition, and stuttering (Dewoody *et al*. 2006).

In general, *Taq* DNA polymerase is expected to have a base pair (bp) substitution error rate that is a function of the starting amount of DNA, the number of cycles, and the amount of $MgCl_2$ and dNTPs. For the concentrations we used, the error rate of *Taq* has been estimated at 1.33×10^{-6} point mutations per bp per PCR cycle (Cline *et al*. 1996). Given 30 PCR cycles, the probability of falsely observing a

Table 1 The 14 microsatellite loci used in the current study. Twelve primers were originally designed for *Populus trichocarpa* and *P. deltoides* (Cervera 2001; Tuskan *et al.* 2004; van der Schoot *et al.* 2000; Yin *et al.* 2004), while the remaining two were designed for *P. tremuloides* (Dayanandan 1998)

Locus	Reference	Linkage group in *P. trichocarpa*	Repeat motif observed in *P. trichocarpa*	No. of alleles observed	Size range observed (bp)
PMGC 486	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	3	$(GA)_{n.a.}$	20	132–298
PMGC 510	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	10	$(GA)_{n.a.}$	21	145–213
PMGC 575	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	1	$(GA)_{n.a.}$	16	172–230
PMGC 2274	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	3	$(GA)_{n.a.}$	10	129–147
PMGC 2658	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	13	$(GA)_{n.a.}$	15	237–277
WPMS17	(van der Schoot *et al.* 2000)	7	$(CAC)_{15}$	11	139–204
ORPM 276	(Tuskan *et al.* 2004)	19	$(TA)_{6}$	18	128–228
GCPM 2768	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	2	$(GA)_{11}$	13	193–217
WPMS19 (lower)	(van der Schoot *et al.* 2000)	unknown	Unknown	19	178–250
WPMS19 (upper)	(van der Schoot *et al.* 2000)	unknown	$(CAG)_{28}$	22	109–286
GCPM 1065	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	unknown	$(AC)_{9}$	12	145–209
PMGC 2731	http://www.ornl.gov/sci/ipgc/ssr_resource.htm	unknown	$(GA)_{n.a.}$	29	150–254
PTR-2	(Dayanandan *et al.* 1998)	unknown	$(TGG)_{8}$	14	213–246
PTR-3	(Dayanandan *et al.* 1998)	unknown	$(TC)_{11}$	19	198–258

Table 2 Somatic mutations detected within two populations of *Populus tremuloides* (Riske Creek and Red Rock). Repeat type refers to dinucleotide (di) or trinucleotide (tri) repeat. Loss of heterozygosity (LOH) was inferred when a ramet was homozygous at a locus that was otherwise heterozygous within the clone. LOH can result from mitotic recombination or gene conversion

			Mutational change					Mutant repeat size (bp)		Most frequent genotype (bp)	
Clone	Locus	Repeat type	+1 step	–1 step	+*k* step	–*k* step	LOH	Allele 1	Allele 2	Allele 1	Allele 2
RC24-1	PMGC 575	di				2		178	186	182	186
RC9-1	PMGC 575	di					x	194	194	190	194
RC9-3*	PMGC 575	di				2		139	169	143	169
RC17-3	PMGC 575	di		x				176	194	176	196
RC23-1	PMGC 575	di	x					184	194	184	192
RC25-2	PMGC 575	di				3		184	186	186	190
RC15-1	PMGC 510	di	x					163	169	165	169
RC22-2*	PMGC 510	di				2		165	169	169	169
RR4-1	PMGC 510	di			7			153	171	153	157
RC13-1	W19U	tri	x					229	274	229	271
RC16-1	W19U	tri					x	235	235	235	241
RR1-1	W19U	tri			3			238	241	241	247
RC17-3	O276	di	x					212	214	212	212
RC17-3	O276	di				9		194	212	212	212
RC20-1	O276	di	x					214	224	212	224
RC1-1	PMGC 486	di			2			136	142	132	142
RC3-1	PMGC 486	di					x	224	224	142	224
RC4-1	PMGC 2274	di			2			137	143	133	143
RR9-1	PMGC 2274	di					x	142	142	142	144
RC20-1	PMGC 2731	di		x				218	222	220	222
RC20-4	PMGC 2731	di	x					214	216	214	214
RR3-1	PTR-2	tri					x	218	218	218	222
RC13-1	PTR-2	tri		x				228	232	232	232
RC15-1	G2768	di	x					209	213	195	209
RC24-1	G2768	di				5		201	203	201	213
RC25-1b	W17	tri			7			154	158	140	154
RC13-1	PMGC 2658	di				2		251	259	251	263
RC15-18	PTR-3	di	x					230	248	228	248
RR9-1	G1065	di				3		165	171	171	171

*not included in further analysis because the number of ramets sampled for the clone was below five.

microsatellite mutation three times in a row due to PCR error is approximately 4.8×10^{-7} ($=(1.33 \times 10^{-6} \times 30 \times 196)^3$), assuming an average fragment length of 196 bp as observed in this study. This calculation is conservative in that it assumes that all three PCR errors appear identical on a sequencing gel, suggesting that we are very unlikely to mistake PCR errors for a somatic mutation.

Alleles were scored as somatic mutations when an individual ramet in a clone differed by one allele at one locus but shared the same alleles as the most frequent genotype at all other loci. We accepted alleles as mutants regardless of the change in repeat number. This approach to scoring microsatellites does not allow for multiple mutations in a lineage leading to a particular ramet. While a ramet is limited to a maximum of one mutational event, clones can have more than one somatic change. We performed a sensitivity analysis to determine if accepting two changes within a single ramet lineage affected the outcome (see Results).

Ramet population growth models

From the identified somatic mutations, we estimated π_k and S_k, as described above, and plotted estimates of π_k vs. $S_k \Big/ \sum_{i=1}^{n-1} \frac{1}{i}$ (slope of one predicted for the constant population-size model) and π_k vs. $2S_k/n$ (slope of one predicted for the clonal growth model).

Relationship between clone size and time

Estimates of clone age based on area rely on the assumption that clones are perfectly circular, with a radius that increases linearly with time. In reality, aspen patches often differ substantially from a circle in shape. To quantify both patch and clone shape, we measured the degree of invagination (*DI*) and elongation (*E*) for each patch. Degree of invagination is measured as

$$DI = \frac{PM_{\mathrm{OBS}}}{2\sqrt{\pi\, Area}}, \qquad \text{(eqn 2)}$$

where PM_{OBS} is the observed perimeter. For a perfectly circular patch, *DI* equals one, while $DI > 1$, indicates that the patch has a perimeter more extensive than expected from a circle. Elongation measures the extent to which a patch is stretched such that the maximum distance between any two ramets, $D_{\max}$, is greater than that expected for a circle of the same area, which is quantified by

$$E = \frac{D_{\max}\sqrt{\pi}}{2\sqrt{Area}}. \qquad \text{(eqn 3)}$$

When $E > 1$, the clone exhibits a stretched form.

We then explored the relationship between π_k as an estimate of neutral genetic variation and clone size. Three measures of clone size were explored: area (*A*), perimeter (*PM*), and maximum distance between any two ramets in a clone ($D_{\max}$). For area, we analysed the relationship as a linear function of area and as a square-root function of area, because the latter is theoretically predicted to grow linearly with time (Fisher 1937). To calculate *A* and *PM*, we used ArcView GIS 3.3 (2002 Environmental Systems Research Institute). ArcView GIS 3.3 creates polygons from the GPS data and then sums the lengths of all the edges of the polygon to obtain a measure of perimeter. To obtain $D_{\max}$, we estimated all possible pairwise physical distances (*m*) separating any two ramets and selected the maximum distance.

Although it is recognized that clonal populations consisting of a single genotype are rare, it is typically assumed, especially in aspen, that a single patch within a population consists of one genotype (Cheliak & Dancik 1982; Jelinski & Cheliak 1992). We found, however, that many patches within our populations consisted of multiple clones (Ally 2008). In these cases, we analysed the relationship between size and microsatellite divergence in two ways. First, we estimated the area of each clone from the physical location of the sampled ramets for that clone (for clones consisting of at least five ramets). Second, we considered the entire patch as if it consisted entirely of the largest clone within that patch.

To select the best predictor of microsatellite divergence, the predictors [area, sqrt(area), perimeter, $D_{\max}$] were subject to separate linear regressions. Linear model evaluation included examining scaled residual error and goodness of fit tests. All statistical tests and analyses were performed using R version 2.3.1 (R Development Core Team 2006).

Estimates of clone age

As described in the Results, the relationship between π_k and S_k is consistent with a growth model, allowing us to estimate the time to common cellular ancestor, T_{CCA}, from the average number of pairwise differences, π_k, observed per locus for clone *k*: $T_{\mathrm{CCA}} = \pi_k/(4\mu)$. Unfortunately, there are no estimates of the microsatellite mutation rate in *Populus*. We thus used two approaches to estimate the mitotic mutation rate and clone age. First, we estimated the number of cell divisions separating all pairs of ramets across all clones to obtain an overall estimate of the mutation rate per cell division from the observed number of pairwise genetic differences; this overall mutation rate was then used to estimate the ages of each clone in number of cell divisions. Second, we placed bounds on the maximum and minimum ages of each clone based on glaciation records and tree ring data, respectively, to estimate a mutation rate per year, which was then used to age all

clones. This is because a clone cannot be younger than the age of its oldest ramet and the oldest a clone could possibly be is the time since the glaciers receded from that region.

First, we estimated the average number of cell divisions separating any two ramets based on their physical distance. The average pairwise distance between two ramets measured in metres, $\bar{D}_{pairs}$, was obtained by averaging the physical distances of all pairs of ramets within a clone and then averaging this value across all clones. We then translated this distance measure from metres to cell divisions as follows. The rate of root growth in *Populus* was measured at ~9.5 mm per day, during the ~20 days of the growing season when most root production occurs (Coleman *et al*. 1996). From this, we can estimate growth per cell division (this includes both the formation of daughter cells and root elongation; Beemster & Baskin 1998), given an estimate of the rate of cell division during periods of rapid growth. We were not able to find a species-specific estimate of the rate of cell division for *P. tremuloides*; thus, we used an estimate of 18 h obtained from cortical cells in *Arabidopsis thaliania* (Beemster & Baskin 1998). This gives us a conversion factor from metres to cell divisions of:

$$C = \frac{0.0095\ m}{24\ hr} \times \frac{18\ hr}{cell\ division} = \frac{0.0071\ m}{cell\ division} \qquad \text{(eqn 4)}$$

Although the rate of growth is slower at other times of the year, C can still be used as a conversion factor throughout the year, assuming that the rate of root elongation (in metres) and the rate of cell division per day slow down proportionately. An estimate of the mutation rate per locus per cell division was then obtained by dividing the number of pairwise genetic differences, $\bar{\pi}$ (per locus and averaged across clones, k), by the physical distance in metres, $\bar{D}_{pairs}$, and multiplying by the conversion factor, C, $\mu_{cell} = \bar{\pi}C/\bar{D}_{pairs}$.

This method of inferring the mutation rate is problematic because it assumes that root connections between ramets are linear. But previous work has shown that *P. tremuloides* grows in a guerrilla formation and that recruitment and death do not simply occur at the edge of a patch but also in the centre of a clone, resulting in a complicated root system under the clone (Ally 2008). To assess whether genetic distance increases with physical distance between pairs of ramets, we regressed π_k against $D_{pairs,k}$ (Fig. S3, Supporting Information). Although we found a significant positive relationship in Riske Creek ($F_{1,18} = 12.36$, p value = 0.00247, $R^2 = 0.407$), this regression is no longer significant when we removed a single outlier ($F_{1,17} = 0.3007$, p value = 0.591, $R^2 = 0.017$). Furthermore, we did not detect a significant relationship in Red Rock, although as indicated, the power for this data set was quite low ($F_{1,5} = 0.8013$, p value = 0.412, $R^2 = 0.14$, $1 - \beta_{err\ prob} = 0.12$). These results suggest that the aboveground physical distance between two ramets provides a poor estimate of the total distance of the roots connecting them, which is the more relevant measure when predicting genetic distance. Assuming a linear root connection would underestimate the true distance of the root connection, and hence, would overestimate the mutation rate per cell division, μ_{cell}. An additional problem is that we used the cell cycle duration from a short-lived annual, which need not apply for a longer-lived species such as *p. tremuloides*. Indeed, there is a known positive relationship between heterochromatin and cell cycle duration (Bennett 1972; Nagl 1974), suggesting that cells would divide slower in tree species, such as trembling aspen, with a great deal of heterochromatin. For example, a cell cycle length of 25.7 h was estimated for the gymnosperm *Pinus banksiana* (Miksche 1967). Using a cell cycle that is too short would overestimate the distance in cell divisions between ramet pairs, and hence, would underestimate the mutation rate per cell division, μ_{cell}. While these two sources of error act in opposite directions, it is difficult to know by how much, and we emphasize that our estimates of mutation rate per cell division based on physical distance between ramets should be treated with a great deal of caution.

The second method used the maximum and minimum possible clone ages to determine lower and upper bounds on the mitotic mutation rate per year, respectively. According to the glacial history of British Columbia, the ice sheets fully retreated by about 10 000 years before present (BP) (Clague 1989; Huntley & Broster 1997; Clague & James 2002), and pollen records suggest that *P. tremuloides* first appeared in southern British Columbia approximately 8000 years BP (Alley 1976; Cawker 1983). Because aspen could not have established in British Columbia before 10 000 years BP, we assumed conservatively that the clone with the most neutral divergence, π_{max}, was a maximum of T_{CCA} = 10 000 years old, giving a lower bound estimate of the mutation rate per year: $\mu_{lower} = \pi_{max}/(4 \times 10\,000$ years) (i.e. using an upper limit on the age of a clone requires fewer mutations per year to account for the observed genetic divergence). Conversely, our tree ring data could be used to estimate the minimum age of each clone. Because a clone must be at least as old as the oldest ramet within it, the minimum age of a clone, T_{min}, was set to the age of the oldest sampled ramet, as inferred from increment cores. We then averaged T_{min} and π_k across clones to obtain an upper bound estimate of the mutation rate per year:

$$\mu_{upper} = \bar{\pi} / \left(4\,\bar{T}_{min}\right)$$

(i.e. using a lower limit on the age of a clone requires more mutations per year to account for the observed genetic divergence). As one clone appeared to be substantially older than the rest (with a higher value of π_k), this clone was excluded from these averages as ramet turn-over in very old clones is likely to erase the relationship between tree

ring data and clone age. A similar procedure was used in Red Rock, except that here the last glaciation ended approximately 12 000 years BP, although *P. tremuloides* pollen has only been found in this region approximately 6600 years BP (Bujak 1974; Clague 1989). Thus, for the Red Rock population, we assumed that the clone with the most neutral divergence was a maximum of T_{CCA} = 12 000 years old.

Results

Somatic mutations

We genotyped 715 ramets in Riske Creek and found a total of 24 somatic changes across 14 microsatellite loci. In Red Rock, 147 ramets were genotyped, with five somatic changes inferred across the same 14 loci. Four samples from Riske Creek and five from Red Rock did not amplify at any loci and were dropped from all analyses. In both Riske Creek and Red Rock, we found somatic mutations that were shared by neighbouring ramets. While it is unknown which ramet had the first mutational event, it is likely that the neighbouring ramets inherited the mutation from a shared ancestor. Thus, these cases were treated as single mutational events in all of our analyses. Before describing the relationship between microsatellite divergence and clone size, we briefly discuss the types of mutations observed.

Eleven of the 29 somatic mutations involved one-step increases (eight) or decreases (three) in the number of repeats (Table 2). Thirteen mutations involved multi-step increases or decreases. The remaining five somatic mutations involved loss of heterozygosity, potentially due to a homologous recombination event during mitosis. Mitotic recombination is more likely to lead to the loss of heterozygosity for loci far from their centromeres, but map information indicating the position of these markers relative to the centromere was not available. Of the changes in repeat number, we found a relatively high percentage of large-step increases or decreases in repeat number. A previous study of microsatellite mutations in maize found that only 5% of mutations involved more than two-step changes in dinucleotide repeat number (Vigouroux *et al*. 2002). By comparison, restricting our attention to dinucleotide repeats, we found that 20% of changes in repeat number involved more than two-step changes.

Studies examining PCR errors involving mononucleotides and dinucleotide repeat sequences have found that PCR amplification introduces systematic decreases, most of which are losses of a single repeat (Walsh *et al*. 1996; Clarke *et al*. 2001). By contrast, we found approximately equal numbers of increases (13) and decreases (11) in repeat number, supporting our claim that we are unlikely to have mistaken PCR errors for somatic mutations.

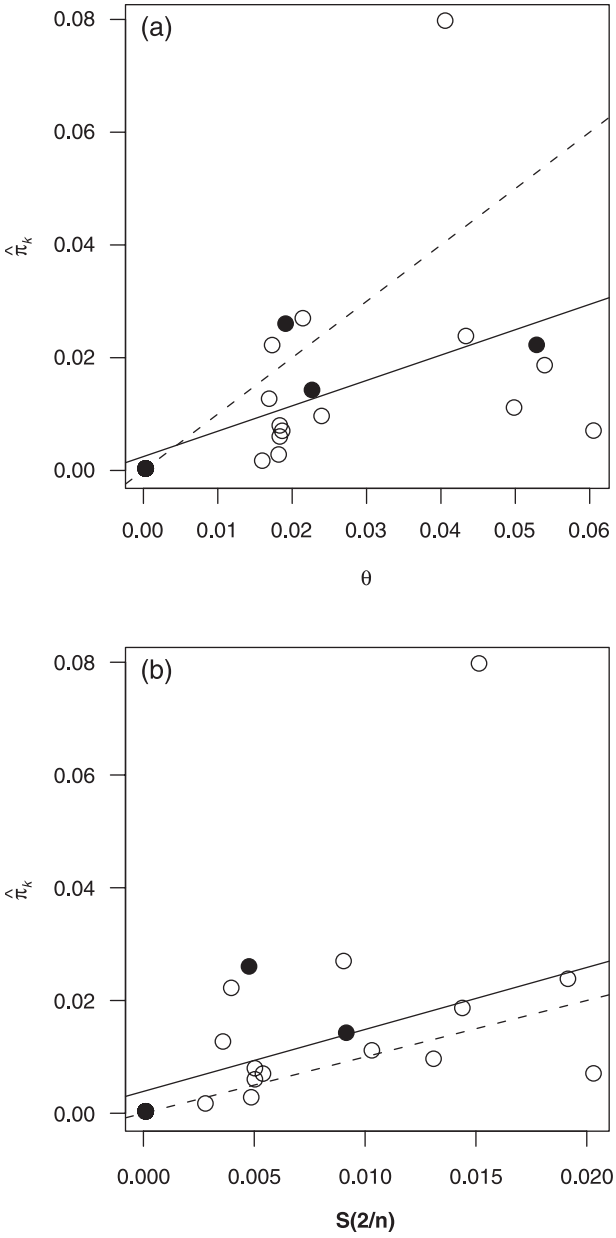

Fig. 3 A comparison of two different growth models for Red Rock, solid circles, and for Riske Creek, open circles. (a) Assuming a constant population size with continual ramet turn-over, the slope of π_k vs. $S_k / \sum_{i=1}^{n-1}(1/i)$ should exhibit a 1 : 1 relationship (dashed). The observed line (solid) had a significantly lower slope of 0.478 ± 0.145 ($F_{1,25}$ = 10.83, p value: 0.002971, R^2 = 0.3023). (b) Assuming a growing population size, the slope of π_k vs. $2S_k/n$ should exhibit a 1 : 1 relationship (dashed). The observed line (solid) had a slope of 1.099 ± 0.329 ($F_{1,25}$ = 11.17, p value: 0.002613, R^2 = 0.3089), which did not differ significantly from the expectation.

Ramet population growth models

As illustrated in Fig. 3a, the slope of a regression of π_k on $S_k \Big/ \sum_{i=1}^{n-1} \frac{1}{i}$ was significantly different from the value of one

expected if clone size were approximately constant over time with continual ramet turn-over ($\beta_{constant} = 0.48 \pm 0.15$, p value = 0.0014). In contrast, the slope of a regression of π_k on $2S_k/n$ was not significantly different from the value of one expected for clones that have been growing over time since their establishment from seed ($\beta_{growth} = 1.09 \pm 0.33$, p value = 0.765). Thus, a model of population growth was more consistent with the genetic relationships among clones than a constant population size model. These results did not change qualitatively if we accepted two changes in any one lineage as a somatic mutation rather than assigning the lineage to a new clone ($\beta_{constant} = 0.446 \pm 0.139$, p value = 0.00054; $\beta_{growth} = 0.944 \pm 0.313$, p value = 0.86).

The fact that the growth model is consistent with the relationship between pairwise genetic diversity and the number of polymorphic loci justifies our use of $\pi_k/4\mu$ as an estimator for clone age, T_{CCA}. We thus proceed to ask whether this estimator is related to the size of a clone.

Relationship between clonal divergence and size

Most clones do not grow as perfect circles, as illustrated by the degree of invagination (*DI*) and elongation (*E*) (Fig. 4). In fact, most clones were either highly convoluted in shape (average value of *DI* = 1.46, SD = 0.40) and/or elongated (average value of *E* = 2.40, SD = 1.1). The fact that clones are not perfect circles indicates that clone size cannot be directly translated into clone age, because radial growth models do not apply.

In Riske Creek, we found little to no relationship between microsatellite divergence and size of a genet (Table 3). Neither area (itself or its square-root) nor perimeter was significantly related to π_k. Although D_{max} was significantly positively related to microsatellite divergence, this was driven by the inclusion of a single data point. This data point has a residual that is beyond two standard deviations of the mean of the jackknifed residuals and two times the average leverage value. When the data were reanalysed using transformed data that better fit the assumptions of a linear regression model (square root-transformed π_k and log-transformed D_{max}), the regression was not significant (Table S1 and Fig. S4, Supporting Information). Similarly, a non-parametric Spearman's rank correlation did not find a significant relationship between D_{max} and π_k (rho = 0.312, S = 915.173, p value = 0.1807). We conclude that size is a poor proxy for the age of a clone, as measured by the degree of genetic divergence within it.

In the clonal literature, when patch size is used as a proxy for age, it is generally assumed that a single genotype makes up the entire patch. We have found, however, that several patches of *Populus tremuloides* consist of multiple genotypes (Ally 2008). Typically, one clone dominates a patch, and the other genotypes are found at the edge of the patch. Thus, we investigated whether the size of the entire patch might serve as a reasonable proxy for the age of the largest clone. We defined the largest clone as the patch inhabiting the largest area within the patch. Again, we found no significant relationship between patch size and π_k for the predominant clone (Table S2, Supporting Information).

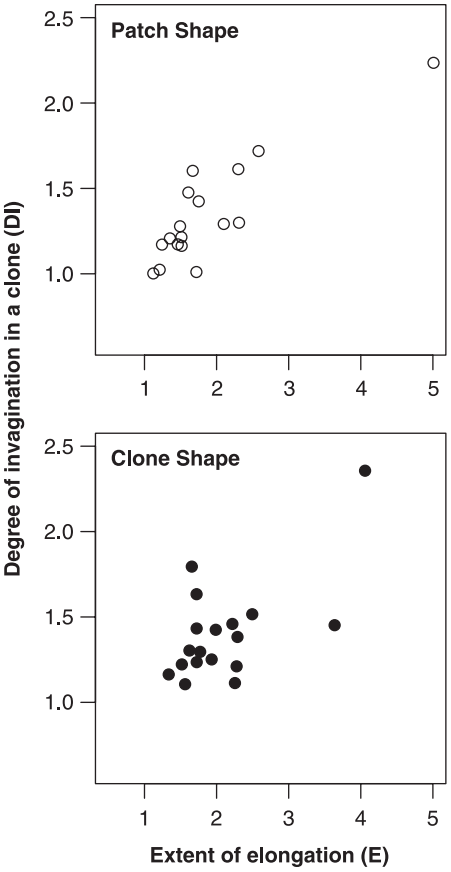

Fig. 4 Neither patch nor clone shape follows a perfect circle (*DI* = 1 and *E* = 1) in Riske Creek.

We next analysed the relationship between clone size and genetic diversity in Red Rock. Again, we found no significant relationship between any of the size metrics examined (area, square-root of area, perimeter, or D_{max}) and π_k (Table 3). This was true whether we examined the size of each clone (Table 3) or the size of the entire patch vs. π_k for the most common clone (Table S2).

Estimates of clone age in cell divisions

In Riske Creek, we obtained a microsatellite mutation rate of $\mu_{cell} = 1.8 \times 10^{-6}$ per locus per cell division by dividing the average number of pairwise genetic differences ($\bar{\pi} = 0.01032$) by the average pairwise physical distance ($\bar{D}_{pairs} = 41\ m$), and multiplying by the conversion factor ($C = 0.0071\ m$/cell division). Among clones where we obtained non-zero estimates of π_k, estimates of clone age

Table 3 Linear regression analyses between microsatellite divergence, π_k, and clone size: area, square-root of area, perimeter (*PM*), and maximum distance between two ramets (D_{max}) for two different aspen populations

Model	F_{stat} p value	d.f.	Estimate ± SE				R^2
			Intercept	Area	Perimeter	D_{max}	
Riske Creek							
Intercept only		19	0.0103 ± 0.004				
			$p = 0.0185$				
Model A1	0.0450		9.48×10^{-3}	3.2×10^{-7}			
Area	$p = 0.834$	1,18	$\pm 5.69 \times 10^{-3}$	$\pm 1.50 \times 10^{-6}$			0.0025
			$p = 0.113$	$p = 0.834$			
Model A2	0.164	1,18	7.52×10^{-3}	6.35×10^{-5}			0.0090
Sqrt (Area)	$p = 0.691$		$\pm 8.05 \times 10^{-3}$	$\pm 1.6 \times 10^{-4}$			
			$p = 0.363$	$p = 0.691$			
Model B	0.120		7.8×10^{-3}		1.14×10^{-5}		
PM	$p = 0.733$	1,18	$\pm 8.34 \times 10^{-3}$		$\pm 3.31 \times 10^{-5}$		0. 0066
			$p = 0.362$		$p = 0.733$		
Model C	11.3		-6.06×10^{-3}			1.46×10^{-4}	
D_{max}	$p = 0.00340$	1,18	$\pm 5.84 \times 10^{-3}$			$\pm 4.33 \times 10^{-5}$	0.380
			$p = 0.313$			$p = 0.00343$	
Model C	0.0497		5.96×10^{-3}			7.10×10^{-6}	
(no outlier)	$p = 0.826$	1,17	$\pm 3.69 \times 10^{-3}$			$\pm 3.18 \times 10^{-5}$	0.0006
			$p = 0.125$			$p = 0.826$	
Red Rock							
Intercept only		6	0.0878 ± 0.0317				
			$p = 0.0325$				
Model A1	1.61		8.4×10^{-3}	6.14×10^{-6}			
Area	$p = 0.260$	1,5	$\pm 6.5 \times 10^{-3}$	$\pm 4.84 \times 10^{-6}$			0.24
			$p = 0.257$	$p = 0.260$			
Model A2	2.72	1,5	0.0028 ± 0.008	0.00045			0.35
Sqrt (Area)	$p = 0.160$		$p = 0.747$	± 0.00027			
				$p = 0.160$			
Model B	2.06		4.82×10^{-3}		3.59×10^{-5}		0.29
PM	$p = 0.211$	1,5	$\pm 7.86 \times 10^{-3}$		$\pm 2.5 \times 10^{-5}$		
			$p = 0.567$		$p = 0.211$		
Model C	1.97		5.88×10^{-3}			7.30×10^{-5}	0.28
D_{max}	$p = 0.219$	1,5	$\pm 7.39 \times 10^{-3}$			5.20×10^{-5}	
			$p = 0.462$			$p = 0.219$	

using $T_{CCA} = \pi_k / 4\mu_{cell}$ ranged from 193 to 10 980 cell generations in Riske Creek. Across all clones, including those with no observed somatic mutations, the average clone age was 2194 cell generations (SE = 628).

In Red Rock, we estimated a mutation rate of $\mu_{cell} = 2.9 \times 10^{-6}$ per locus per cell division. This was based on an average number of pairwise genetic differences of $\bar{\pi} = 0.01374$, an average pairwise physical distance of $\bar{D}_{pairs} = 34\ m$, and $C = 0.0071\ m$/cell division. Among clones in Red Rock with non-zero values of π_k, clones ranged in age from 1227 to 2632 cell divisions. The average clone age across all clones in Red Rock was 2117 cell generations (SE = 232).

Estimates of clone age in years

Using our second approach, the maximum possible clone age in Riske Creek based on glacial records (10 000 years) was assigned to the clone exhibiting the largest amount of genetic diversity ($\pi_{max} = 0.079$). This allowed us to obtain a lower-bound estimate of the mutation rate: $\mu_{lower} = \pi_{max} / (4 \times 10\,000 \text{ years}) = 2.0 \times 10^{-6}$ per locus per year. Among clones with non-zero estimates of π_k, estimates of clone age using $T_{CCA} = \pi_k / (4\mu_{lower})$ ranged from 175 to 10 000 years. Based on μ_{lower}, the average age of all clones, including those with no observed somatic mutations, is estimated to be 1998 years (SE = 571). Conversely, we can use the minimal clone age based on tree ring data to obtain an upper-bound estimate of the mutation rate,

$$\mu_{upper} = \bar{\pi} / \left(4\,\bar{T}_{min}\right) = 2.3 \times 10^{-5}$$

per locus per year, from the average pairwise genetic diversity ($\bar{\pi} = 0.00667$) and the average number of tree rings of the oldest ramet ($\bar{T}_{min} = 71$ years), where both averages excluded

the clone with greatest divergence. Among clones with non-zero estimates of π_k in Riske Creek, estimates of clone age using $T_{CCA} = \pi_k/(4\mu_{upper})$ ranged from 14 to 846 years. Based on μ_{upper}, the average age of all clones, including those with no observed somatic mutations, is 169 years (SE = 48). The upper- and lower-bound estimates of the time to common cellular ancestor differ by only about an order of magnitude, and thus provide a reasonable range of estimates for clone age.

In Red Rock, the lower and upper bounds around the mitotic mutation rate per year were $\mu_{lower} = 6.0 \times 10^{-7}$ and $\mu_{upper} = 3.8 \times 10^{-5}$. The lower bound was based on glacial records dating back 12 000 years BP and $\pi_{max} = 0.029$. The upper bound was based on the average $\bar{\pi} = 0.0111$, and $\bar{T}_{min} = 73$ years, where both averages excluded the clone with the greatest divergence. Thus, in Red Rock, among clones with non-zero estimates of π_k, clone age ranges from 5595 to 12 000 years (based on μ_{lower}) or from 92 to 198 years (based on μ_{upper}). The average age of all clones, including those with no observed somatic mutations, is 9652 years (based on μ_{lower}; SE = 1058) and 160 years (based on μ_{upper}; SE = 17). These bounds are wider in Red Rock than in Riske Creek, reflecting the fact that we had less data (in particular, fewer patches) in Red Rock.

Discussion

Long-lived clonal plants are notoriously difficult to age. Age estimates based on patch size are problematic because the growth of a clone is not perfectly circular and is often constrained by the clone's environment. Here we applied an ontogenetic 'molecular clock' to patches of *Populus tremuloides* in order to estimate clone age from the degree of within-clone genetic diversity at 14 microsatellite loci. To do so, we first had to confirm that genetic divergence among ramets reflects the time since establishment of the clone from a seed. As confirmation, we showed that the relationship between average pairwise genetic diversity (π_k) and the number of segregating polymorphisms (S_k) per locus is consistent with a model of clonal growth from a common cellular ancestor that existed at time T_{CCA} in the past (Fig. 3). This justifies the use of π_k as an estimator of this coalescence time for a clone ($T_{CCA} = \pi_k/4\mu$). We used π_k rather than S_k as a surrogate of age because the former is less sensitive to the existence of clusters of related ramets (see Supporting Information – Section A).

Next, we explored the relationship between clone size and π_k as a surrogate measure of age. The assumption of a positive linear relationship between clone size and age has not been explored to any significant extent in the literature on clonal plants. Despite variation in both size and microsatellite divergence within a clone, our data exhibited little to no relationship between clone size and π_k. A lack of relationship between size and a genetic estimate of clone age is not surprising given that the physical space occupied by a clone depends not only on time since seed establishment but also on the clone's competitive ability and its biotic and abiotic environment. In particular, events such as fires and disease can cause a stand to contract and die back, causing the clone that regenerates from the rootstock to be reduced in size.

The location of the somatic mutants that we observed may provide some insight into the clonal growth of aspen. Growth habit of a clone, i.e. whether it expands vertically or horizontally, can affect where mitotically derived somatic mutations are found and whether mutations are shared by other ramets. If horizontal growth predominates, then more shared mitotically derived mutations may be found among ramets. If, however, the development of a genet is primarily vertical, then fewer mutations will be found in the rootstock and fewer will be shared by neighbouring ramets. In both populations, we found seven different clones (five in Riske Creek and two in Red Rock) where somatic mutations were shared by at least one neighbouring ramet. The remaining 22 somatic mutations could have occurred either in the shoots or in the rootstock. In addition, in two clones, where ramets shared somatic mutations, the physical distance between two of the neighbouring ramets was substantial. In one case, two ramets with a shared mutation were 21 m and in the other case they were 58.2 m apart. This fact coupled with our finding that multiple clones occupy a patch suggests that the root connections may be extensive and that aspen may not follow strictly a phalanx growth.

To use somatic mutations to estimate clone age, we had to first infer a mutation rate appropriate for these microsatellite loci in *P. tremuloides*. Using the distance between pairs of ramets in metres and relating this to the number of cell divisions separating them, we obtained an estimate of the rate at which somatic mutations arise per locus per cell division of $\mu_{cell} = 1.8 \times 10^{-6}$ (Riske Creek) and $\mu_{cell} = 2.9 \times 10^{-6}$ (Red Rock). Both these estimates include repeat length mutations and mitotic recombination events within somatic cells. We also present estimates of the somatic mutation rate per year, based on maximum possible clone ages given by the last glacial retreat from these regions ($\mu_{lower} = 2.0 \times 10^{-6}$ in Riske Creek, $\mu_{lower} = 6.0 \times 10^{-7}$ in Red Rock) and based on minimum possible clone ages from tree ring counts ($\mu_{upper} = 2.3 \times 10^{-5}$ in Riske Creek, $\mu_{upper} = 3.8 \times 10^{-5}$ in Red Rock). Comparing these mutation rates per cell division and per year suggests that very few cell divisions occur per year, but it must be remembered that 1 year of root growth and ramet production (involving several cell divisions) could be followed by several years of stasis (involving few additional cell divisions but ongoing mutation). Furthermore, the mutation rate per cell division was estimated assuming linear root connections between ramets; more circuitous connections (as must surely be the case) would lower estimates of μ_{cell} and predict more cell divisions per year.

Finally, these mutation rates were used to estimate average clone age from the relationship

$$T_{CCA} = \bar{\pi} / 4\mu .$$

For Riske Creek, we estimated average clone age to be 2194 cell generations (from μ_{cell}), or between 169 and 1998 years (from μ_{upper} and μ_{lower}, respectively). For Red Rock, we estimated average clone age to be 2117 cell generations (from μ_{cell}), or between 160 and 9652 years (from μ_{upper} and μ_{lower}, respectively). Our estimates of clone age indicate that clones at Riske Creek and Red Rock are similar in age, on average, regardless of the method used to estimate mutation rate.

Caveats

There are several caveats that must be kept in mind when interpreting our results. Despite the very large sampling effort (862 ramets genotyped at 14 microsatellite loci), we observed only a modest number of somatic mutations (29 across all clones). Thus, our estimate of genetic diversity for any one clone contains substantial measurement error, as does any estimate of the age of a particular clone. For this reason, we have focused our discussion on statements about the average age of clones within these regions. With the same number of loci, however, our method of estimating clone age from microsatellite divergence may prove more successful in some herbaceous perennials (e.g. ferns), where microsatellite mutation rates are higher.

A related point is that the lack of a relationship between genetic diversity and clone size (Table 3) might simply reflect the imprecision of our genetically based estimates of age, rather than problems with size-based metrics. Personal observations of the patches, however, suggest that patch size was often physically constrained, especially in the rocky and montane habitat of Red Rock. Patches were generally not circular (Fig. 4), as is assumed when using a size-based metric of age. Even in cases where clones do develop with a phalanx arrangement (aggregated or circular), habitat constraints (space, nutrients, competition, etc.) are likely to restrict the upper limit of clone size. Even as clone size reaches its maximum size, mutations can continue to accrue, making size a less useful estimate of clone age than genetic divergence. In our populations, we found ramets were variably spaced with some longer connections between individual modules within some genets, typical of the guerrilla form of growth. We also found out that patches were often composed of multiple clones (Ally 2008), which introduces error when estimating the size of a clone from the size of a patch. For all of these reasons, as well as the lack of a relationship between genetic diversity and size, we believe that size-based metrics cannot be used to estimate clone age with any accuracy in *P. tremuloides*. If the age of a particular clone is desired, a more precise estimate of age could be obtained by increasing the number of microsatellite loci scored.

Our ability to detect somatic mutations may have been influenced by the interaction between sampling strategy used in the field and patterns of clonal growth. If the birth of new ramets occurs primarily at the perimeter of a clone, while death occurs in the centre, then the ability to detect somatic mutations will be greater near the edge of a patch. For example, higher soil temperatures, which may occur at the edge of a patch, stimulate early initiation of suckers, while very dry or water-saturated soils reduce sucker initiation (Maini & Horton 1966; Schier 1982). A one-tailed Fisher exact test, however, showed that the detection of somatic mutations was not affected by the location of the ramet (i.e, whether ramets occurred on the edge of a patch or in the centre, p value = 0.1404) (Table S3 and Fig. S5, Supporting Information). Thus, our estimates of age do not appear to be strongly influenced by where sampling occurred.

A final caveat is that we have assumed an infinite sites model when measuring genetic divergence, implying that we could detect each mutational event. With high rates of mutation at microsatellite loci, it is possible to have parallel mutations, thus reducing the observed number of segregating sites under the infinite sites model (Bertorelle & Slatkin 1995). In addition, there is a strong downward mutation bias in long microsatellites, which increases the possibility of parallel back mutations (Wierdl *et al*. 1997). However, the relatively young ages of our clones makes homoplasy less likely.

Comparisons to other mutation rate estimates

Before concluding, it is worth comparing our estimated mutation rate per cell division to other mutation rates estimated for microsatellites in plants. The majority of estimates in the literature provide mutation rates per sexual generation. For example, mutation rates per locus per sexual generation have been estimated in *Triticum turgidum* (2.4×10^{-4} averaged across 10 loci, Thuillet 2002) and in *Thuja plicata* (3.1×10^{-4} averaged across 8 loci, O'Connell & Ritland 2004). In *Zea mays* (ssp. *mays*), the mutation rate estimated from 142 loci was 7.7×10^{-4} per sexual generation (Vigouroux *et al*. 2002). Using an estimated number of cell divisions of ~50 from zygote to zygote in *Z. mays* (Otto & Walbot 1990), this suggests a mutation rate of 1.5×10^{-5} per cell division. While this estimate is 10-fold higher than the mutation rate per cell division estimated herein, this factor is within the realm of variation expected among species and among different types of repeats (dinucleotide vs. longer repeats; (Vigouroux *et al*. 2002)).

One study has estimated the somatic mutation rate per cell generation in *Pinus strobus* (Cloutier *et al*. 2003). In that study, the number of cell divisions per metre was estimated (14 205–47 170) from the observed range of cell sizes in

the shoot. Although no somatic mutations were detected, the authors were able to use the zero term of a Poisson distribution to estimate a per-locus mutation rate of $\mu_{cell} = 1.4 \times 10^{-7}$ (based on the larger number of cell divisions) and 4.6×10^{-7} (based on the smaller number of cell divisions). These estimates are lower than the above estimates of the mutation rate per cell division. This might reflect differences in mutation rates in the rootstock (the focus of our study) vs. aboveground plant mass (the focus of Cloutier *et al*. 2003). However, one reason to suspect that Cloutier *et al*. (2003) might have overestimated the number of cell divisions per metre (and underestimated the mutation rate) is that the number of cell divisions separating the closest pair of sampled tissues (15.6 m apart) would be at least 221 598 cell divisions using their method (assuming linear growth at the meristem and the smaller estimate of 14 205 cell divisions per metre). Given cell cycle lengths in pine of ~24 h (Miksche 1967), this number of cell divisions would require nearly 500 years of growth, yet the breeding orchard from which the data were obtained was only established in the 1970s. One potential explanation for this discrepancy, raised by the authors, is that growth can be accomplished by cell divisions outside of the apical region, whereas only mutations in the apical region contribute to genetic differentiation of distant plant tissues.

In our study, we avoided the difficulty of translating distance into cell divisions using cell size by focusing instead on the maximum growth rate of roots in metres and the maximum rate of cell division. This procedure allowed us to infer the rate of growth including both cell divisions inside and outside of the apical meristem as well as cell elongation over the time period expected of a cell division. That said, our estimates of the somatic mutation rate per cell are themselves rough, because we approximated total root distance separating two ramets by the linear distance between them, and because we used estimates from other species of the maximum rate of cell division and from other locales for the maximum rate of root growth.

Conclusions

Clone age is not only important to a fundamental understanding of the ecology and evolution of clonal plants, but it also has additional applications. Knowing clone age can provide a metric for examining associated community characteristics, like changes in understorey species richness or abundance over time. The maximum age of an aspen clone in a given patch would be the upper bound on the community age and thus offer a scale of temporal progression against which changes in diversity of herbaceous understorey species could be assessed. This study is the first to examine the relationship between microsatellite divergence (a measure of clone age) and clone size, motivated by the desire to have a quick and easy proxy to assess clone age. Unfortunately, our results indicate that no simple size metric can be used to estimate clone age and that more labour-intensive estimates of genetic divergence are needed. Furthermore, our analysis was hampered by the need to estimate mutation rates using indirect methods, before we were able to age clones of *P. tremuloides*. Additional empirical work that examines ramets and/or clones of known age (e.g. from provenance trials) would provide more direct estimates of mutation rate that could then be applied to data from natural populations. As more information becomes available on clonal growth, development, and responses to habitat, and as genomic methods to score somatic mutations become more accessible and less costly, improved estimates of the lifespan of clonal plants will become possible.

Acknowledgements

The authors would like to acknowledge the help of Carol Ritland (Genetic Data Centre), Jodie Krakowski, Laura Glaubach, and Cyndi Smith (WLNP). Funding and support for the project to came from grants from Waterton Lakes National Park (to D.A.), The Nature Trust British Columbia (Brink-McLean Grassland Fund to D.A.), Kathy Martin, NSERC (PGSB to D.A. and Discovery grants to K.R. and S.P.O.), and the Zoology Computing Unit. We would also like to thank the following people for their helpful and insightful comments on the manuscript: Sally Aitken, Rosie Redfield, Michael Doebeli, Quentin Cronk, Carol Ritland, Ruth Shaw and Rémy Petit, and the two anonymous reviewers.

References

Alley NF (1976) The palynology and paleoclimatic significance of a dated core of Holocene peat, Okanagan Valley, southern British Columbia. *Canadian Journal of Earth Sciences*, **13**, 1131–1144.

Ally D (2008) *The Cost of Longevity: Loss of Sexual Function in Natural Clones of Populus Tremuloides*. PhD Dissertation, Department of Graduate Program of Genetics University of British Columbia, Vancouver, Canada.

Beemster GTS, Baskin TI (1998) Analysis of cell division and elongation underlying the developmental acceleration of root growth in *Arabidopsis thaliana*. *Plant Physiology*, **116**, 1515–1526.

Bennett MD (1972) Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society B: Biological Sciences*, **181**, 109–135.

Bertorelle G, Slatkin M (1995) The number of segregating sites in expanding human populations, with implications for estimates of demographic parameters. *Molecular Biology and Evolution*, **12**, 887–892.

Bujak CA (1974) *Recent Palynology of Goat Lake and Lost Lake, Waterton Lakes National Park*. MSc. Geography, University of Calgary, Calgary, Alberta, Canada.

Cawker KB (1983) Fire history and grassland vegetation change: three pollen diagrams from southern British Columbia. *Canadian Journal of Botany*, **61**, 1126–1139.

Cervera M, Storme V, Ivens B *et al*. (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics*, **158**, 787–809.

Chakraborty R, Kimmel M, Stivers David N, Davison Leslea J, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proceedings of the National Academy of Sciences, USA*, **94**, 1041–1046.

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

Clague JJ (1989) Quaternary geology of the Canadian Cordillera. In: *Quaternary Geology of Canada and Greenland* (ed. Fulton RJ), pp. 17–96. Geological Survey of Canada, Geology of Canada.

Clague JJ, James TS (2002) History and isostatic effects of the last ice sheet in southern British Columbia. *Quaternary Science Reviews*, **21**, 71–87.

Clarke LA, Rebelo CS, Goncalves J, Boavida MG, Jordan P (2001) PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences. *Molecular Pathology*, **54**, 351–353.

Cline J, Braman JC, Hogrefe HH (1996) PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Research*, **24**, 3546–3551.

Cloutier D, Rioux D, Beaulieu J, Schoen DJ (2003) Somatic stability of microsatellite loci in Eastern white pine, *Pinus strobus* L. *Heredity*, **90**, 247–252.

Coleman MD, Dickson RE, Isebrands JG, Karnosky DF (1996) Root growth and physiology of potted and field-grown trembling aspen exposed to tropospheric ozone. *Tree Physiology*, **16**, 145–152.

Dayanandan S, Rajora OP, Bawa KS (1998) Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). *Theoretical Applied Genetics*, **96**, 950–956.

DeByle NV, Winokur RP (1985) *Aspen: Ecology and Management in the Western United States*. United States Department of Agriculture, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Dewoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes*, **6**, 951–957.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, **5**, 435–445.

Escaravage N, Questiau S, Pornon A, Doche B, Taberlet P (1998) Clonal diversity in a *Rhododendron ferrugineum* L. (Ericaceae) population inferred from AFLP markers. *Molecular Ecology*, **7**, 975–982.

Fisher RA (1937) The wave of advance of advantageous genes. *Annals of Eugenics*, **7**, 355–369.

Frey BR, Lieffers VJ, Hogg EH, Landhausser SM (2004) Predicting landscape patterns of aspen dieback: mechanisms and knowledge gaps. *Canadian Journal of Forest Research*, **34**, 1379–1390.

Gardner SN, Mangel M (1997) When can a clonal organism escape senescence? *The American Naturalist*, **150**, 462–490.

Harris GR, Lovell PH (1980) Localized spread of *Veronica filiformis*, *V. agrestis* and *V. persica*. *Journal of Applied Ecology*, **17**, 815–826.

Hogg EH, Brandt JP, Kochtubajda B (2002) Growth and dieback of aspen forests in northwestern Alberta, Canada, in relation to climate and insects. *Canadian Journal of Forest Research*, **32**, 823–832.

Hudson RR (1990) Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology*, **7**, 1–44.

Hughes TP (1984) Population dynamics based on individual size rather than age: a general model with a reef coral example. *The American Naturalist*, **123**, 778–795.

Huntley DH, Broster BE (1997) The Late Wisconsinan deglacial history of the east-central Taseko Lakes area, British Columbia. *Canadian Journal of Earth Sciences*, **34**, 1510–1520.

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.

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

Klekowski EJ (1997) Somatic mutation theory of clonality. In: *The Ecology and Evolution of Clonal Plants* (ed. De Kroon H, Van Groenendael J), pp. 227–241. Backhuys Publishers, Leiden, The Netherlands.

Lewis MA (2000) Spread rate for a non-linear stochastic invasion. *Journal of Mathematical Biology*, **41**, 430–454.

Maini JS (1965a) On the organization and growth of aspen roots. II. Rapidly tapering and cord-like roots. In: *Canadian Journal of Botany*, pp. 0–24. Canadian Department Forestry, Forest Research Branch, Ottawa, Ontario, Canada.

Maini JS (1965b) On the organization and growth of aspen roots. III. Relationship between length of root cutting and production of suckers. Canadian Dept. Forestry, Forest Research Branch, Ottawa, Ontario, Canada.

Maini JS, Horton KW (1966) Vegetative propagation of *Populus* spp. I. Influence of temperature on formation and initial growth of aspen suckers. *Canadian Journal of Botany*, **44**, 1183–1189.

Miksche JP (1967) Radiobotanical parameters of *Pinus banksiana*. *Naturwissenschaften*, **54**, 322–322.

Nagl W (1974) Role of heterochromatin in the control of cell cycle duration. *Nature*, **249**, 53–54.

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

Orive ME (1995) Senescence in organisms with clonal reproduction and complex life-histories. *The American Naturalist*, **145**, 90–108.

Otto SP, Walbot V (1990) DNA methylation in eukaryotes: kinetics of demethylation and *de novo* methylation during the life cycle. *Genetics*, **124**, 429–437.

Peet RK (1981) Forest vegetation of the Colorado Front Range: composition and dynamics. *Vegetatio*, **45**, 3–75.

Pla M, Jofre A, Martell M, Molinas M, Gomez J (2000) Large accumulation of mRNA and DNA point modifications in a plant senescent tissue. *FEBS Letters*, **472**, 14–16.

R Development Core Team (2006) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Reusch TBH, Stam WT, Olsen JL (1998) Size and estimated age of genets in eelgrass, *Zostera marina*, assessed with microsatellite markers. *Marine Biology*, **133**, 519–525.

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

Schier GA (1982) Sucker regeneration in some deteriorating Utah aspen stands: development of independent root systems. *Canadian Journal of Forest Research*, **12**, 1032–1035.

Schlotterer C, Ritter R, Harr B, Brem G (1998) High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. *Molecular Biology and Evolution*, **15**, 1269–1274.

van der Schoot J, Pospiskova M, Vosman B, Smulders MJM (2000) Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). *Theoretical Applied Genetics*, **101**, 317–322.

Schug MD, Mackay TFC, Aquadro CF (1997) Low mutation rates of microsatellite loci in *Drosophila melanogaster*. *Nature Genetics*, **15**, 99–102.

Slatkin M (1996) Gene genealogies within mutant allelic classes. *Genetics*, **143**, 579–587.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Steinger T, Körner C, Schmid B (1995) Long-term persistence in a changing climate: DNA analysis suggests very old ages of clones of alpine *Carex curvula*. *Oecologia*, **105**, 94–99.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tanner JE (2001) The influence of clonality on demography: patterns in expected longevity and survivorship. *Ecology*, **82**, 1971–1981.

Thuillet A, Bru D, David J, Roumet P, Santoni S, Sourdille P, Bataillon T (2002) Direct estimation of mutation rate for 10 microsatellite loci in durum wheat, *Triticum turgidum* (L.) Thell. ssp *durum* desf. *Molecular Biology and Evolution*, **19**, 122–125.

Tuskan GA, Gunter LE, Yang ZK *et al*. (2004) Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forest Research*, **34**, 85–93.

Tyson M, Vaillancourt RE, Reid JB (1998) Determination of clone size and age in a mallee eucalypt using RAPDs. *Australian Journal of Botany*, **46**, 161–172.

Udupa S, Baum M (2001) High mutation rate and mutational bias at $(TAA)_n$ microsatellite loci in chickpea (*Cicer arietinum* L.). *Molecular Genetics and Genomics*, **265**, 1097–1103.

Vasek FC (1980) Creosote bush: long-lived clones in the Mojave Desert. *American Journal of Botany*, **67**, 246–255.

Vazqeuz JF, Perez T, Albornoz J, Dominguez A (2000) Estimation of microsatellite mutation rates in *Drosophila melanogaster*. *Genetical Research Cambridge*, **76**, 323–326.

Vigouroux Y, Jaqueth JS, Matsuoka Y *et al*. (2002) Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology and Evolution*, **19**, 1251–1260.

Wakeley J (2007) *Coalescent Theory: An Introduction*. Roberts & Company Publishers, Greenwood Village, Colorado, USA.

Walsh PS, Fildes NJ, Reynolds R (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus. *Nucleic Acids Research*, **24**, 2807–2812.

Whittle CA (2006) The influence of environmental factors, the pollen: ovule ratio and seed bank persistence on molecular evolutionary rates in plants. *Journal of Evolutionary Biology*, **19**, 302–308.

Wierdl M, Dominska M, Petes TD (1997) Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics*, **146**, 769–779.

Yin T, DiFazio S, Gunter L, Riemenschneider D, Tuskan G (2004) Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. *Theoretical and Applied Genetics*, **109**, 451–463.

Dr. Dilara Ally investigates the evolutionary processes that lead to diversity within clonal organisms. Her PhD dissertation coupled neutral genetic divergence within natural populations of *Populus tremuloides* clones with pollen fitness data to infer the rate and effect of mildly deleterious mutations. Currently, she is using bacteriophage to examine the effects of spatial structure on genetic events that occur during adaptive evolution under environmental shifts. Dr. Sally Otto's research focuses on how organismal diversity arises by exploring the influence of environmental circumstances (biotic and abiotic) on the outcome of evolution, primarily using mathematical models. Current research investigates the evolutionary forces acting on genome structure and mating systems to account for the remarkable diversity that exists in these fundamental features of an organism. Dr. Kermit Ritland conducts research in both basic and applied aspects of population and quantitative genetics, mainly in plants, and is best known for his work on estimation of plant mating systems. He is currently involved with a large scale genome project with spruce.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Departures from a pure star-like coalescent history.

Fig. S2 Error introduced by assuming a star-like coalescent history.

Fig. S3 The relationship between microsatellite divergence (π_k) and the average pairwise distance between two ramets within the k^{th} clone ($D_{pairs,\,k}$) in Riske Creek and Red Rock.

Fig. S4 The relationship between microsatellite divergence and maximum distance between two ramets across clones in Riske Creek.

Table S1 The relationship between microsatellite divergence and clone size for Riske Creek

Table S2 The relationship between microsatellite divergence and the most frequent clone in a patch for Riske Creek

Table S3 The number of somatic mutations detected was not affected by whether a ramet was found on the perimeter or within the centre of the patch

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.