Molecular Ecology Resources (2008) 8, 74–82 doi: 10.1111/j.1471-8286.2007.01910.x

TECHNICAL ADVANCES

The importance of setting the right genetic distance threshold for identification of clones using amplified fragment length polymorphism: a case study with five species in the tropical plant genus *Piper*

E. LASSO

Smithsonian Tropical Research Institute, Unit 0948 APO AA 34002–0948, Panama, University of Illinois at Urbana-Champaign, 265 Morrill Hall, 505 Goodwin Ave, IL 61801, USA

Abstract

Amplified fragment length polymorphism (AFLP) has been widely used for clone identification, but numerous studies have shown that clonemates do not always present identical AFLP fingerprints. Pairwise AFLP distances that distinguish known clones from nonclones have been used to identify a threshold genetic dissimilarity distance below which samples are considered to represent a single clone. Most studies to date have reported threshold values between 2% and 4%. Here, I determine the consistency of the clonal threshold across five species in the tropical plant genus *Piper*, and evaluate the sensitivity of genetic diversity indices and estimates of frequency of clonal reproduction to the threshold value selected. I sampled multiple ramets per individual from widely distributed plants for each of the five *Piper* species to set a threshold at the point where the error rate of clonal assignments was lowest. I then sampled all individuals of each shade-tolerant species in a 1-ha plot, and of each light-demanding species in 25 × 35-m plot, to estimate the frequency of asexual recruitment in natural populations using a series of different thresholds including the threshold set with the preliminary sampling. Clonal threshold values for the different species ranged from 0% to 5% AFLP genetic dissimilarity distance. To determine the sensitivity of estimates of clonal reproduction, I calculated several clonal diversity indexes for the natural populations of each of the five species guided by the range in clonal threshold values observed across the five *Piper* species. I show that small changes in the value of the clonal threshold can lead to very different conclusions regarding the level of clonal reproduction in natural populations.

Keywords: AFLP, asexual recruitment, clones, evenness index, fingerprint dissimilarities, genetic diversity

Received 1 April 2007; revision accepted 19 June 2007

Asexual reproduction is a common phenomenon in plants representing many habits and environments (Stuefer *et al.* 2001). The recognition that the balance between sexual and asexual reproduction in a population can have major ecological and evolutionary consequences has motivated numerous studies designed to evaluate the incidence of clonal reproduction in natural populations and to understand the effect of clonal spread on: genetic diversity (Ellstrand & Roose 1987), the evolution of breeding systems (Handel 1985; Eckert 2000; Charpentier 2001), fine- and large-scale genetic structure (Chung & Epperson 2000; Chung *et al.* 2005; Clark-Tapia *et al.* 2005), and population dynamics in general (Thomas & Dale 1975; Eriksson 1994; Mandujano *et al.* 2001).

One critical step in the evaluation of clonal reproduction in natural populations is the ability to accurately distinguish clones from genetically distinct individuals. Sometimes

Correspondence: Eloisa Lasso, Fax: (507)-2128790/8791; E-mail: elasso@uiuc.edu

clones are identified by following root connections (Reinartz & Popp 1987; Barsoum 2001). However, studying root connectivity can give estimates of clonal frequency that are biased upwards if roots graft naturally, or biased downwards if connections among roots are lost over time (Miwa *et al.* 2001). The frequency of asexual reproduction has also been assessed by counting the proportion of plants in a given area that possess callus tissue growth between the roots and shoot, which is supposed to be present only in plantlets that originated by fragmentation (Sagers 1993). Estimates obtained using this method overestimated clonal frequency compared to estimates from molecular data in the same area (Bush & Mulcahy 1999). Moreover, large-scale excavation to determine callus growth is often not desirable, for example in protected areas.

With the development of molecular markers, it has become easier to study clonal structure over large spatial scales and with minimal impact on populations (only one to a few leaves are collected per plant). However, molecular methods also have pitfalls that need to be taken into consideration when designing a sampling strategy. For example, amplified fragment length polymorphism (AFLP) analysis (Vos *et al.* 1995) has been used for clone identification (Escaravage *et al.* 1998; Kollmann *et al.* 2000; van der Hulst *et al.* 2000) because the technique can reveal a large number of markers with a high degree of reproducibility (Jones *et al.* 1997); and markers can be developed easily and at relatively low cost when no prior genetic information is available for the study species (Mueller & Wolfenbarger 1999). However, samples coming from known clonemates (i.e. leaves from same shoot or individuals connected below ground) do not always have identical fingerprints (Arens *et al.* 1998; Winfield *et al.* 1998; van der Hulst *et al.* 2000; Douhovnikoff & Dodd 2003). Some of the approaches to determine clonality assume that clones are always identical and they identify clones by calculating the likelihood of observing at least n identical multilocus genotypes in a specific sample from a population in Hardy–Weinberg equilibrium (Parks & Werth 1993; Ivey & Richards 2001; Stenberg *et al.* 2003). However, these approaches are inadequate if somatic mutation takes place and individuals produced asexually are not genetically identical. Disregarding this fingerprint variability within clones can lead to biased estimates of clonal diversity and asexual reproduction frequencies in the studied populations if individuals are incorrectly assigned to different clonal lineages whenever they are not identical.

To circumvent these drawbacks, pairwise genetic distance comparisons among replicate runs or samples have been used to identify a threshold to distinguish putative clones from nonclonal individuals (Douhovnikoff & Dodd 2003). Most studies to date have found threshold values between 2% and 4% (Arens *et al.* 1998; Winfield *et al.* 1998; van der Hulst *et al.* 2000; Douhovnikoff & Dodd 2003; Douhovnikoff *et al.* 2004). That is, any pair of individuals having less than 2–4% differences in their fingerprint profiles were classified as part of the same clone. Because these threshold values are so consistent across studies, it is temping to use them universally. However, before assuming that a single threshold value can be assumed in all situations, some fundamental questions need to be addressed. How does the threshold value vary among species? Can closely related species be analysed using the same threshold value? How sensitive are estimates of clonal frequency and genetic diversity to the threshold used? In this work, I aimed to answer these questions by investigating clonal threshold value across five species in the genus *Piper*.

Materials and methods

Study species and study site

The genus *Piper* constitutes an important component of tropical forests worldwide and is represented by > 1000 species. The study site is located in a tropical moist semideciduous forest on Barro Colorado Island, Panama (BCI; 9°09′N, 79°51′W), where the genus *Piper* is represented by 22 species (Croat 1978). Species of *Piper* have a remarkable ability to regenerate by fragmentation and from pinned down branches (Greig 1993; E. Lasso & J. Dalling, unpublished data), but it remains to establish how important this regeneration pathway is in sustaining natural populations. However, because I found that AFLP profiles are not identical among samples of the same individual (DNA extracted from different leaves), a threshold had to be determined before frequency of clonality can be evaluated for these species.

Threshold setting procedure

To identify the amount of variation on AFLP fingerprints and to find out the threshold to use for each species, I collected two to three separate leaves from different ramets from the same genet for a total of 11–24 genetically distinct plants per species (Table 1). To ensure that I was collecting genetically distinct plants, I collected only samples from plants that were more than 100 m apart. In order to include as much variation as possible, I selected plants from across Barro Colorado Island (BCI) and two plants per species from mainland (16.5 km away).

The threshold indicates the maximum dissimilarity (genetic distance) that is allowed between two individuals to still be considered clonemates, with the 'same' genotype. I classified each sample either as a clone or unique genotype under different threshold scenarios using the software GENOTYPE (Meirmans & Van Tienderen 2004). GENOTYPE uses pairwise genetic distances to classify samples as members of a clonal group or as unique genotypes based

Species	No. of leaves (threshold setting procedure)	No. of plants (threshold setting procedure)	No. of plants (estimates in natural populations)
*Piper darienensis**	59	20	182
*Piper cordulatum**	70	24	72
*Piper aequale**	67	23	166
Piper dilatatum	33	11	43
Piper marginatum	38	13	43

Table 1 Number of samples collected. The first two columns are the number of samples collected per species for the threshold setting procedure. The third column is the number of samples collected to estimate frequency of clonal reproduction in natural populations

*Shade-tolerant species.

on the threshold value selected by the user. Therefore, in each separate analysis with each threshold, a different set of samples were assigned to clones or to unique genotypes. As the origin of each sample was known, I was able to calculate an error rate of assignment for each threshold as $(x + y)/n$; where x is the number of leaves from the same plant that were wrongly assigned as being from a different plant and y is the number of leaves from different plants that were wrongly assigned as being from the same plant, and n is the total number of samples. The threshold with the lowest error rate was then considered the appropriate threshold for the species to be used in future analysis to calculate the frequency of clonal reproduction in natural populations.

A histogram of the pairwise genetic distances for ramets and genets was created to examine whether the threshold with the lowest error rate coincided with the point where the distribution of genetic distances of genets and ramets overlapped. The genetic distance in GENOTYPE was calculated using the Dice similarity option for dominant data which is then transformed by the software to a distance measure. The distance equation is $\{1 - [2a/(2a + b + c)]\} \times 100$; where a is the number of bands shared by both individuals, b is the number of bands present in the first individual but not in the second, and c is the number of band present in the second individual but not in the first individual.

Sensitivity of clonal genetic diversity indexes to the threshold selected

To evaluate how sensitive estimates of genetic diversity are to changes in the threshold used, I sampled one natural population for each species and then calculated several indices of clonal diversity for each population using threshold values ranging from 0 to 14. This range of values was selected to observe the behaviour of the indices to changes across an ample range of threshold values. Natural populations of three shade-tolerant species were sampled in a 1-ha plot in the forest. Natural populations of two light-demanding species were sampled across a 35 × 25-m plot in a clearing. All plants present in the plots that were located at least 5 cm apart were sampled (Table 1). One leaf from each plant was collected for AFLP analysis.

Using the software GENODIVE (Meirmans & Van Tienderen 2004), I calculated the indices of clonal diversity most commonly used in the literature. Following Ellstrand & Roose (1987), I calculated the genetic diversity index 'Proportion distinguishable' PD, as G (number of genotypes)/N (number of samples), which is the proportion of distinguishable genotypes or the proportion of individuals in the population that were recruited sexually. I also calculated the Simpson genetic diversity index corrected for sample size which is $D = n/n - 1 * (\Sigma\ p_i^2)$; where n is the sample size; and p_i is the frequency of genotype i. The evenness which is an indicator of how evenly the genotypes are divided over the population is calculated as $1/s * 1/\Sigma\ p_i^2$; where p_i is the frequency of genotype i; and s is the number of genotypes.

Application of the threshold: frequency of clonality in natural populations of Piper

The frequency of asexual reproduction for *Piper* species was assessed for the same populations described above. I determined the proportion of individuals in the population that were recruited asexually as: $1 - G$ (number of genotypes)/N (number of samples). The number of genotypes was obtained using GENOTYPE and using the threshold set already with the preliminary sampling.

DNA isolation and AFLP procedure

Leaves where collected and kept in ice until they were processed in the laboratory. On the day of collection, leaves were surface cleaned with 95% alcohol and left to dry in silica gel for 1 week. Twenty milligrams of dry tissue was ground using the FastPrep FP120 (Qbiogene). DNA was extracted using DNeasy 96 plant extraction kit (QIAGEN) and following the manufacturer's protocol. DNA concentrations were established by running DNA samples with DNA mass ladder of known concentration on agarose gels.

AFLP analysis followed the method of Vos *et al.* (1995) with some modifications. Restriction digestion and ligation

were performed separately. The restriction digestion was performed in a 20-μL solution containing 50–200 ng DNA; 0.9 U of *Mse*I, 6 U of *Eco*RI, BSA 1×, NE buffer 0.5× and *Eco*RI buffer 0.5× (New England Biolabs). The solution was incubated at 37 °C for 3 h, and then at 65 °C for 20 min. To this 20 μL of digested DNA, I added 20 μL of the ligation solution that contained 0.25 μM of *Eco*RI adapter, 2.5 μM of *Mse*I adapter, 0.6 U of T4 DNA ligase (Fisher), T4 ligase buffer 1×. The ligation reaction was incubated at 16 °C for 3 h and then at 70 °C for 10 min. Then a final restriction–digestion–ligation reaction was carried out to re-restrict any fragments accidentally re-ligated during the ligation. This method was found to give more repeatable bands across samples. To the 40 μL of the solution with restricted and ligated DNA, I added 15.5 μL of the final restriction–ligation solution. This solution contained 0.5 U *Mse*I, 5 U *Eco*RI, BSA 1.3′×, NaCl 0.03 M, 0.3 U of T4 DNA ligase, and T4 ligase buffer 1.3×. This solution was incubated at 25 °C for 24 h and then 65 °C for 20 min. The product of the restriction–ligation was then diluted with 50 μL $TE_{0.1}$. Preselective amplification was performed in 20 μL of reaction mixture containing 4 μL of the restriction–ligation product and 16 μL of preselective amplification solution that contained 0.3 μM Eco+1 A primer and Mse+1C primer, 0.8 mM dNTPs, 4 mM $MgCl_2$, polymerase chain reaction (PCR) buffer 1×, 1 U of *Taq* polymerase (QIAGEN). After an initial denaturation at 72 °C for 2 min, 20 PCR cycles of 30 s at 94 °C, 30 s at 56 °C, and 2 min at 72 °C were performed, followed by a final 30-min elongation at 60 °C. Ten microlitres of pre-amplified products was diluted in 250 μL of $TE_{0.1}$ buffer. Selective amplification was performed in 25 μL of reaction mixture containing 6.5 μL of preselective amplification product and 18.5 μL of selective amplification solution that contained 0.24 μM *Eco*RI selective primer labelled with a fluorescent marker (6-FAM), 0.3 μM *Mse*I primer, 0.2 mM dNTPs, 2.0 mM $MgCl_2$, 1× PCR buffer, 0.5 U ampli*Taq* (Applied Biosystems). The selective PCR had two cycles set after an initial denaturation at 94 °C for 2 min: 14 cycles of 30 s at 94 °C, 30 s at 65 °C (reduced by –0.7 deg/cycle), and 60 s at 72 °C, followed by 19 cycles of 30 s at 94 °C, 30 s at 56 °C and 60 s at 72 °C. Four primer combinations per species were used to obtain AFLP fingerprints; those were chosen after screening 20 selective primer pairs. Three of them were used across all species; *Eco*RI-ACG/*Mse*I-CTG, *Eco*RI-*AGG*/*Mse*I-CCA, *Eco*RI-ATG/*Mse*I-CAG. The fourth primer pair used per species is as follow: *Eco*RI-ACG/*Mse*I-CTA for *Piper darienensis* C.DC. and *Piper dilatatum* L.C.Rich, *Eco*RI-AGT/*Mse*I-CTG for *Piper aequale* Vahl, *Eco*RI-AGG/*Mse*I-CAG for *Piper cordulatum* C.DC., and *Eco*RI-AGT/*Mse*I-CAT for *Piper marginatum* Jacq. For *P. dilatatum*, I used two additional primer pairs, *Eco*RI-ACG/*Mse*I-CTT and *Eco*RI-ACG/*Mse*I-CCA, because most of the loci were monomorphic. Fingerprint data were obtained by running the amplified samples in an ABI PRISM 3130 capillary electrophoresis machine, and presence or absence of fragments were scored using GENESCAN and GENOTYPER software (version 3.7, Applied Biosystems).

To reduce the possibility of collecting leaves with endophytes, whose DNA could introduce upward biases to estimates of intraclonal variation in AFLP fingerprints, I always collected only young leaves. Additionally, to exclude putative fragments coming from endophyte DNA, I used a series of three leaves of the same individuals (10 individuals × species) to evaluate each band from each primer combination. The following criteria were used to filter loci: if a band was present only in one plant, I only retained that band if it was consistently scored as present in all three leaves from that same plant. I also retained bands that were present in all three leaves from at least 60% of the plants. This is an arbitrary cut off but it reduces the probability of including endophytes because of the reduced likelihood that 60% of the plants sampled will have the same endophyte in three of their leaves; and at the same time, it leaves room to include natural polymorphism due to somatic mutations.

Results

The four primer combinations yielded 559–770 clearly identifiable bands per species. However, after the filtering of loci, only 15–27% of fragments were selected for fingerprinting (Table 2).

Threshold setting

Piper species varied in the threshold dissimilarity distances that minimized errors in clone assignment. For *Piper darienensis*, a threshold of 3% dissimilarity gave us the lowest error rate for this species (error rate = 5.08%; Fig. 1) meaning that individuals that are 97% or more similar in their fingerprints should be considered as part of the same clone. For *Piper cordulatum*, the threshold was 5% (error

Table 2 Number of bands detected using four primer pairs (for *Piper dilatatum*, six primer pairs were used), number of bands selected for fingerprinting and number of polymorphic bands per species

Species	No. of fragments detected	No. of loci used	No. of polymorphic loci
P. darienensis	770	120	87
P. cordulatum	736	102	68
P. aequale	694	124	95
P. dilatatum	739	140	49
P. marginatum	559	153	110

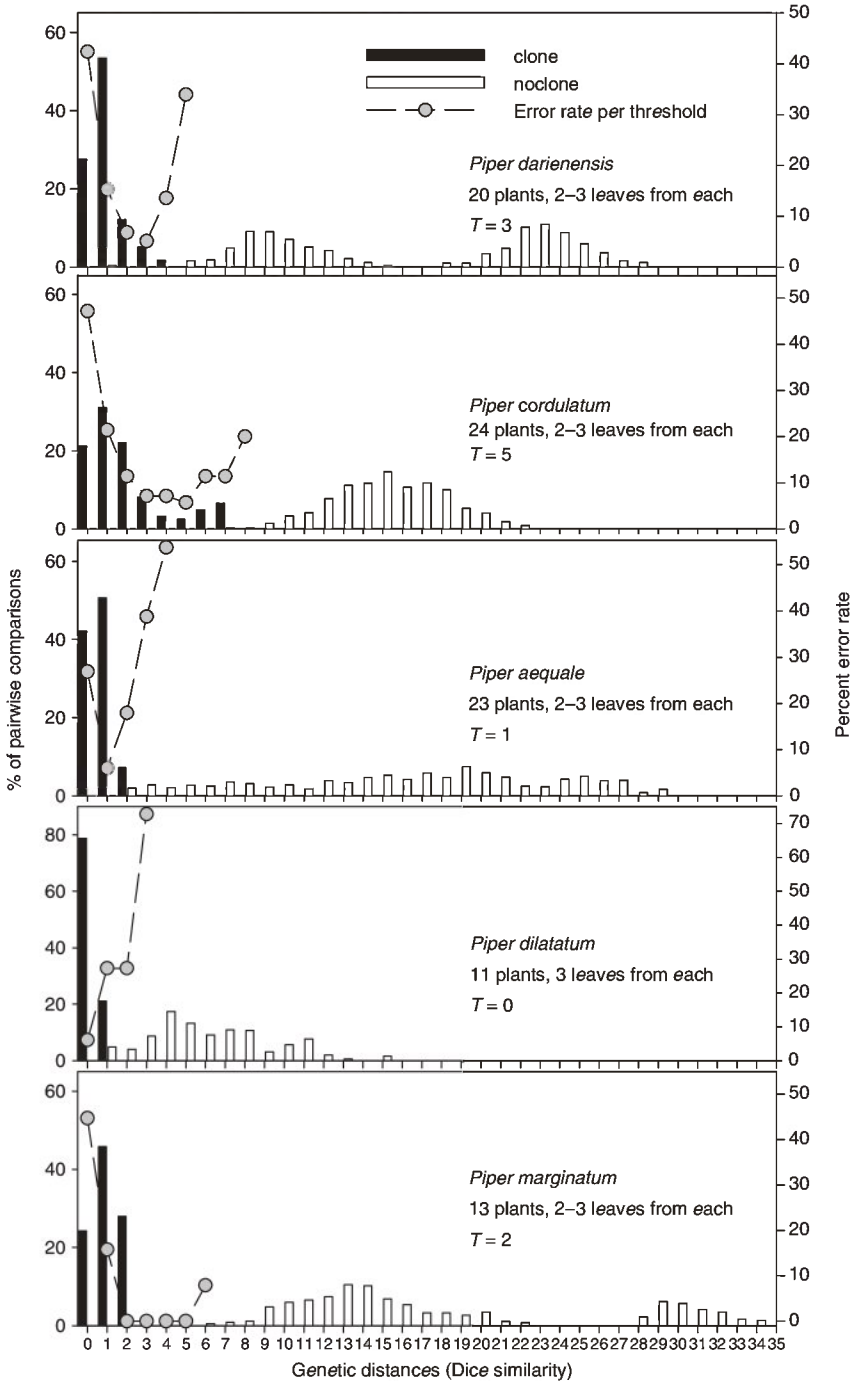

Fig. 1 Pairwise genetic distances for samples collected from ramets (black bars) and samples collected from genets (white bars) located in various geographical points. Round symbols (right axes) represent the error rate of clonal assignment for each threshold tested. The *T* value in the figure represents the threshold with the lowest error rate.

rate = 5.7%), for *Piper aequale*, was 1% (error rate = 5.97%). For *Piper dilatatum*, the threshold was 0% (error rate = 6.06%) meaning that for this species only individuals that are identical in their fingerprints should be considered as part of the same clone. However, for *Piper marginatum*, all thresholds from 2% to 5% dissimilarity gave a 0% error rate (Fig. 1).

Sensitivity of clonal genetic diversity indexes to the threshold selected

The most sensitive of the three diversity indexes calculated to changes to the threshold value used, was PD = *G*/*N*, which is the proportion of distinguishable genotypes or the proportion of individuals in the population that were

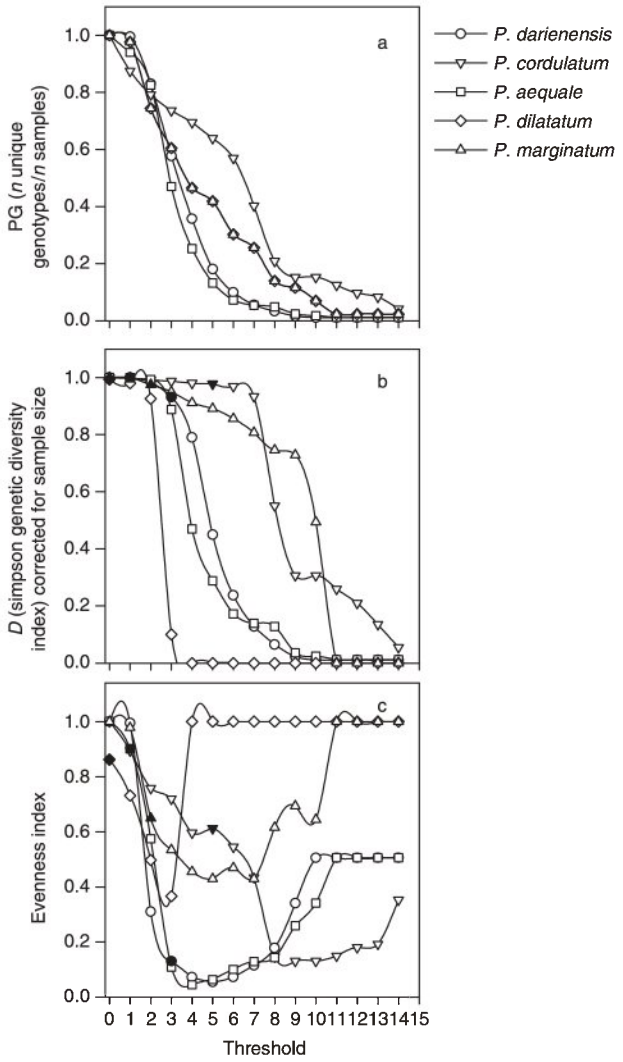

Fig. 2 Changes in three estimates of genotypic diversity when different threshold values are used to assign genotypes to clonal lineages. The estimates are: (a) 'Proportion distinguishable' PD; (b) Simpson genetic diversity index corrected for sample size; and (c) evenness index E. In all figures, dark symbols represent the estimate for the right threshold of the species found with the preliminary sampling. PD = 1 when all individuals are unique genotypes; PD = 0 when all individuals are clones.

sexually recruited (Fig. 2a). The Simpson genetic diversity index was less sensitive; with similar values across 0–3 threshold values (Fig. 2b). The evenness index is also very sensitive to changes in the threshold value used, changing in an inconsistent way across different thresholds (Fig. 2c).

Application of the threshold: frequency of clonality in natural populations

Piper species differ in their reproductive strategies. Two of the understorey species, *P. darienensis* and *P. cordulatum*, were the species with the highest frequency of clonality; 42% and 36%, respectively, of the individuals in their population were recruited by asexual means. This was followed by *P. marginatum* with 27%; and then by *P. aequale* and *P. dilatatum* with less than 10% of the individuals in their population being asexually recruited.

Discussion

The low frequency of seedling establishment success in understory *Piper* species combined with observations that *Piper* can successfully regenerate from leaf and stem fragments (Greig 1993; Lasso & Dalling, unpublished data) suggest that clonal reproduction may be important in allowing these species to maintain populations in forest understory. This study confirms that clonal reproduction does occur in natural populations of *Piper*; but before any definite conclusion can be drawn, a thorough study is needed.

What looks like a discrepancy between percent of clones identified in the population and the percent of pairwise comparisons having a distance measure lower than the threshold selected (Fig. 3) is an unavoidable side-effect of

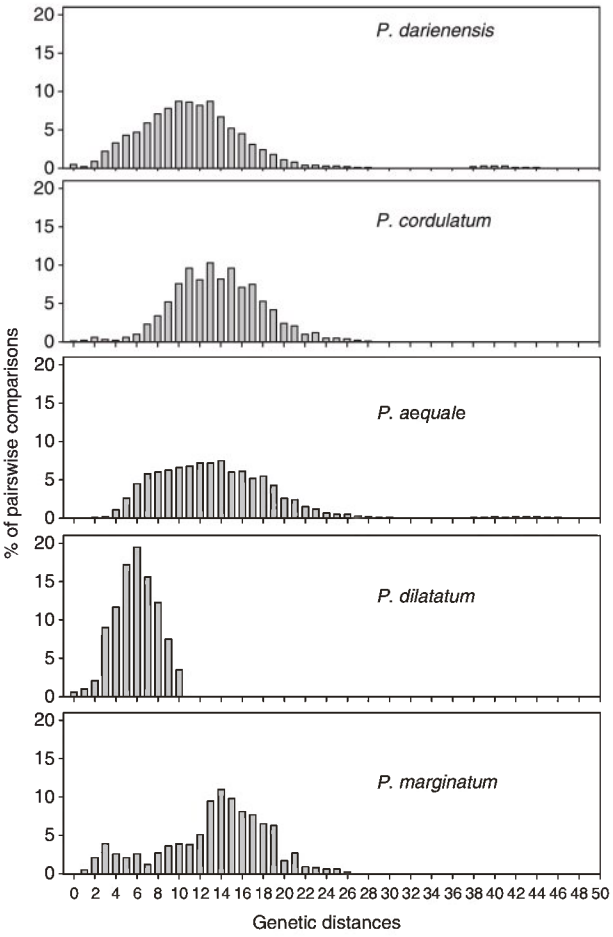

Fig. 3 Distribution of pairwise genetic distances in natural populations of *Piper* species on Barro Colorado Island, Panama.

using a single variable (clone) to summarize multivariate data (the allelic data). Sometimes a clone can include individuals with some of their pairwise comparisons distance higher than the threshold selected.

Threshold setting procedure

AFLP is considered a reliable marker to estimate frequency of asexual recruitment and genetic diversity in clonal plants. They are similar to microsatellites in consistency and resolution for the detection of unique genotypes in clonal plants (van der Hulst *et al.* 2003). It is a good marker for clonal studies as long as it is acknowledged that clones are not always identical and that degree of intraclonal genetic differences can vary across species.

Other approaches, different to the one I present here, have been used to set a threshold to distinguish putative clones from nonclonal individuals. Douhovnikoff & Dodd (2003) set a threshold base on the mean and standard deviation of the two peaks in a bimodal frequency distribution of genetic distances among siblings and clones. They assessed the level of variation of AFLP fingerprints among and within clones and siblings raised from seeds. However, although desirable, it can be difficult to include siblings in the analysis of the threshold, as clonally reproducing plants often have low seed production or seed germination success (Eckert 2000) as is the case with *Piper* species. Meirmans & Van Tienderen (2004) developed the software GENOTYPE and GENODIVE to help the user to recognize the threshold for the identification of clones and to assign individuals to clonal lineage once the threshold has been selected. However, they claimed that no prior sampling was needed and that thresholds could be obtained directly from the frequency distribution of all pairwise comparisons because it is expected that for clonal populations, the distribution should be multimodal. Multimodality should arise when the first peak represents the frequency of pairwise genetic distances of all individuals that are clones, including individuals that differ slightly due to scoring error or somatic mutations, while the second peak reflects the presence of putative siblings and closely related individuals. Any further peaks would then represent substructuring in the population. They argued that the threshold can be assigned to the valley between the first and second peak. However, as I found in this study, histograms of genetic distances from samples collected in natural populations are not always multimodal (Fig. 3); therefore it becomes challenging to decide what threshold to use in such situations. Here, I propose an alternative method to set the threshold that takes advantage of the capability of the software GENODIVE, but that requires an additional preliminary sampling of known clones and nonclones to calculate an error rate of assignment of different putative thresholds and select that threshold with the lowest error rate.

I found a large degree of variation in the threshold values, but this variation was comparable to what has been found for other species. I found that for some species, clones have identical AFLP fingerprints (threshold = 0), as in *Piper dilatatum*, or can differ as much as 5% as in *Piper cordulatum*. AFLP fingerprints for sequoia tree clones can have up to 3% differences among them (Douhovnikoff *et al.* 2004); and up to 2% on willow tree clones (Douhovnikoff & Dodd 2003). For black poplar trees, Arens *et al.* (1998) found that clones can have up to 2% differences in their AFLP fingerprint, but Winfield *et al.* (1998) found that clones from this same species can have up to 4% differences in their AFLP fingerprint. This discrepancy between studies indicates that thresholds are not only species specific but also study-site specific and therefore need to be calibrated for each study. Most of these studies, except for Douhovnikoff studies, used only replicate runs of the same leaves; therefore, they were only accounting for laboratory and scoring errors but not for variation within clones. However, a study by Douhovnikoff & Dodd (2003), which also included in the analysis replicates of several leaves per individual, found that genetic variation among stems of the same clone is more important than variation due to AFLP procedures. Somatic mutation may then be the main culprit of intraclonal variation in AFLP profiles, and perhaps some of the previous studies that were accounting for laboratory and scoring error only were setting the threshold too low and underestimating levels of clonality. As is shown next, the threshold value used could indeed drastically affect estimates of clonal reproduction.

Sensitivity of clonal genetic diversity indexes to the threshold selected

Through analysis of all plants in 1-ha plots, I was able to establish how sensitive estimates of the frequency of clonal reproduction and genetic diversity are to threshold values of genetic distance. Here, I demonstrate that slight changes in the threshold can drastically change the estimates of frequency of sexual vs. asexual reproduction in populations (Fig. 2a). For example, using a threshold of 0% dissimilarity for *P. dilatatum* would estimate that 90% of samples were recruited sexually, whereas by using a threshold of 2% dissimilarity one will estimate that 48% of samples were recruited sexually. However, the other species are less sensitive; estimates vary from 100% to 78% sexual recruitment when changing the threshold from 0% to 2%. These results clearly show that a single threshold value cannot be assumed in all studies and species, and that getting the right threshold is crucial for making the right conclusions about species regeneration strategies. Estimates

of genetic diversity (*D*, Simpson diversity index) seem to be less sensitive to the threshold used (Fig. 2b), but the evenness index (Fig. 2c) greatly varies. These results suggest that the evenness index can be useful only when the threshold is assigned accurately. Even so, evenness needs to be interpreted with caution because it is not useful in distinguishing populations in the extreme ranges of high and low diversity where E will approach 1. For example, *Piper marginatum* has an evenness value of 1 when using a threshold of 0% and also when using any threshold value above 11%. In the first situation, using a threshold of 0%, accepting as clones only individuals with identical fingerprints, all plants were classified as unique genotypes (no clones were present) and an evenness index of 1 was achieved. In the second situation, accepting that clones could have up to 11% differences in their fingerprint, all plants in the plot were classified as part of one single clone and again the index was 1. A value E = 1 indicates that all genotypes have equal frequencies. This will occur when (i) all samples belong to one clone, or (ii) when all individuals are unique genotypes, or (iii) when all clones are represented by the same number of individuals. The index will approach zero when genotypes are in uneven frequencies in the population and this can happen with different arrangements of genotypes. It is probably more informative to present a map with the distribution of clones than to present the evenness index.

Even though sampling strategies were different for shade-tolerant and light-demanding species, using the jackknifing method with increasing sample size in GENODIVE, I determined that sample sizes were big enough for all species to be able to estimate PG (*n* unique genotypes/*n* samples) and *D* (Simpson genetic diversity index) without bias. In both cases, the trend in the value leveled off when it reached the actual sample size indicating that the sample size was adequate for all species. However, the evenness index only leveled off for *Piper darienensis* and *Piper aequale*, indicating that good estimates of evenness needs large sample sizes for the other species.

The old dogma that clones are 'genetically identical' has been largely disproved for many taxa, including viruses, plants, fungi and animals because clones may rapidly change genetically due to mutations (review by Lushai & Loxdale 2002). Despite the fact that clonal members can differ genetically, it is still possible to get a good approximation of the level of clonality in a population by first detecting the amount of intraclonal variation for a given species in a given situation and then incorporating that knowledge to set the threshold to detect clones in the wild. Ideally, it would be better to use ramets from different clones, but when they are not available (i.e. when hard to detect morphologically or unable to detect underground connections), multiple samples from single ramets can be used. However, it should be acknowledged that leaves from the same ramet may have more somatic similarities than tissue from distinct ramets.

Even though most studies to date have found threshold values between 2% and 4% (Arens *et al.* 1998; Winfield *et al.* 1998; van der Hulst *et al.* 2000; Douhovnikoff & Dodd 2003; Douhovnikoff *et al.* 2004), they cannot be used universally for two reasons. First, the threshold value varies among species, perhaps due to species differences in somatic mutation rate; and second, estimates of clonal frequency and genetic diversity are very sensitive to the threshold used. Given that any conclusion about the frequency of asexual reproduction in clonal species will be greatly affected by the threshold used, great care in this matter is recommended. No matter what type of marker is used, it is essential to always carry out a preliminary sampling specifically designed to detect the amount of intraclonal genetic variability in the studied species and to set the threshold to identify clones.

Acknowledgements

I want to thank E. Bermingham, J. W. Dalling, J. Wright and A. Jones for early comments in this manuscript. Thanks to O. Sanjur, M. Gonzales, G. Grajales and C. Vergara for all their support and help in the laboratory. Thanks to E. Sanchez, R. Gallery and D. Kikuchi for helping with the collection of samples. Thanks to J. Koontz and A. Hinkle for the orientation with AFLP techniques at the beginning of this journey. Financial support for this work came from (i) Deep Gene Research Training Award (NSF grant DEB-0090227)-UC-Berkeley; (ii) STRI predoc fellowship; (iii) Dissertation Improvement Grant-NSF (grant DEB 05–08471); (iv) Francis M and Harlie M. Clark Research Support Grant from University of Illinois; and (v) SENACYT-IFARHU becas doctorales 2005.

References

Arens P, Coops H, Jansen J, Vosman B (1998) Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. *Molecular Ecology*, **7**, 11–18.

Barsoum N (2001) Relative contributions of sexual and asexual regeneration strategies in *Populus nigra* and *Salix alba* during the first years of establishment on a braided gravel bed river. *Evolutionary Ecology*, **15**, 255–279.

Bush SP, Mulcahy DL (1999) The effects of regeneration by fragmentation upon clonal diversity in the tropical forest shrub *Poikilacanthus macranthus*: random amplified polymorphic DNA (RAPD) results. *Molecular Ecology*, **8**, 865–870.

Charpentier A (2001) Consequences of clonal growth for plant mating. *Evolutionary Ecology*, **15**, 521–530.

Chung MG, Epperson BK (2000) Clonal and spatial genetic structure in *Eurya emarginata* (Theaceae). *Heredity*, **84**, 170–177.

Chung MY, Suh Y, Lopez-Pujol J, Nason JD, Chung MG (2005) Clonal and fine-scale genetic structure in populations of a restricted Korean endemic, *Hosta jonesii* (Liliaceae) and the implications for conservation. *Annals of Botany*, **96**, 279–288.

Clark-Tapia R, Alfonso-Corrado C, Eguiarte LE, Molina-Freaner F (2005) Clonal diversity and distribution in *Stenocereus eruca*

(Cactaceae), a narrow endemic cactus of the Sonoran Desert. *American Journal of Botany*, **92**, 272–278.

Croat TB (1978) *Flora of Barro Colorado Island*. Stanford University Press, Stanford, California.

Douhovnikoff V, Dodd RS (2003) Intra-clonal variation and a similarity threshold for identification of clones: application to *Salix exigua* using AFLP molecular markers. *Theoretical and Applied Genetics*, **106**, 1307–1315.

Douhovnikoff V, Cheng AM, Dodd RS (2004) Incidence, size and spatial structure of clones in second-growth stands of coast redwood, *Sequoia sempervirens* (Cupressaceae). *American Journal of Botany*, **91**, 1140–1146.

Eckert CG (2000) Contribution of autogamy and geitonogamy to self-fertilization in a mass-flowering, clonal plant. *Ecology*, **81**, 532–542.

Ellstrand NC, Roose ML (1987) Patterns of genotypic diversity in clonal plant species. *American Journal of Botany*, **74**, 123–131.

Eriksson O (1994) Stochastic population dynamics of clonal plants: numerical experiments with ramet and genet models. *Ecological Research*, **9**, 257–268.

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.

Greig N (1993) Regeneration mode in Neotropical *Piper*: habitat and species comparisons. *Ecology*, **74**, 2125–2135.

Handel SN (1985) The intrusion of clonal growth patterns on plant breeding systems. *American Naturalist*, **125**, 367–384.

van der Hulst RGM, Mes THM, den Nijs JCM, Bachmann K (2000) Amplified fragment length polymorphism (AFLP) markers reveal that population structure of triploid dandelions (*Taraxacum officinale*) exhibits both clonality and recombination. *Molecular Ecology*, **9**, 1–8.

van der Hulst RG, Mes THM, Falque M, Stam P, den Nijs JCM, Bachmann K (2003) Genetic structure of a population sample of apomictic dandelions. *Heredity*, **90**, 326–335.

Ivey CT, Richards JH (2001) Genotypic diversity and clonal structure of Everglades sawgrass, *Cladium jamaicense* (Cyperaceae). *International Journal of Plant Sciences*, **162**, 1327–1335.

Jones CJ, Edwards KJ, Castaglione S *et al.* (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding*, **3**, 381–390.

Kollmann J, Steinger T, Roy BA (2000) Evidence of sexuality in European *Rubus* (Rosaceae) species based on AFLP and allozyme analysis. *American Journal of Botany*, **87**, 1592–1598.

Lushai G, Loxdale HD (2002) The biological improbability of a clone. *Genetical Research, Cambridge*, **79**, 1–9.

Mandujano MC, Montana C, Franco M, Golubov J, Flores-Martinez A (2001) Integration of demographic annual variability in a clonal desert cactus. *Ecology*, **82**, 344–359.

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.

Miwa M, Tanaka R, Yamanoshita T *et al.* (2001) Analysis of clonal structure of *Melaleuca cajuputi* (Myrtaceae) at a barren sandy site in Thailand using microsatellite polymorphism. *Trees – Structure and Function*, **15**, 242–248.

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution*, **14**, 389–394.

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.

Reinartz JA, Popp JW (1987) Structure of clones of northern prickly ash (*Xanthoxylum americanum*). *American Journal of Botany*, **74**, 415–428.

Sagers CL (1993) Reproduction in Neotropical shrubs: the occurrence and some mechanisms of asexuality. *Ecology*, **74**, 615–618.

Stenberg P, Lundmark M, Saura A (2003) MLGSIM: a program for detecting clones using a simulation approach. *Molecular Ecology Notes*, **3**, 329–331.

Stuefer JF, Erschbamer B, Huber H, Suzuki JI (2001) The ecology and evolutionary biology of clonal plants: an introduction to the proceedings of Clone-2000. *Evolutionary Ecology*, **15**, 223–230.

Thomas AG, Dale HM (1975) The role of seed production in the dynamics of established populations of *Hieracium floribundum* and a comparison with that of vegetative reproduction. *Canadian Journal of Botany*, **53**, 3022–3031.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP — a new technique for DNA-fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Winfield MO, Arnold GM, Cooper F *et al.* (1998) A study of genetic diversity in *Populus nigra* subsp. *betulifolia* in the Upper Severn area of the UK using AFLP markers. *Molecular Ecology*, **7**, 3–10.