Fingerprinting and Pedigree Analysis in *Eucalyptus globulus* Using RAPDs

By K. A. Nesbitt, B. M. Potts, R. E. Vaillancourt and J. B. Reid

Cooperative Research Centre for Temperate Hardwood Forestry and Department of Plant Science, University of Tasmania, GPO Box 252-55, Hobart, TAS 7001, Australia

(Received 23rd April 1996)

Summary

RAPDs were used to confirm clonal fidelity and distinguish individuals of varying degrees of relationship in *Eucalyptus globulus*. RAPD variation was examined amongst ramets derived from 10 ortets which were either full-sibs, half-sibs, or unrelated, as well as their parents. Genetic similarity amongst samples was calculated from 117 RAPD bands using the simple matching coefficient. RAPD variation within clones was trivial compared to the variation found even between full-sib clones. Genetic similarity decreased with pedigree distance. The pedigree of the material was successfully displayed by UPGMA clustering of the genetic similarity matrix, with parents generally clustering with their offspring. In a second experiment, 120 RAPD bands allowed the successful discrimination of selfs from their intra- and interprovenance half-sib outcrosses and again genetic similarity decreased with pedigree distance. These results clearly indicate that RAPDs have great potential for estimation of outcrossing rates and pedigree analysis.

Key words: *Eucalyptus globulus*, RAPD, fingerprinting, breeding system, pedigree analysis, genetic distance.

FDC: 165.3; 165.5; 176.1 *Eucalyptus globulus*.

Introduction

*Eucalyptus globulus* Labill. is a commercially important forestry species. It is widely planted in temperate regions throughout the world and is currently the subject of many genetic improvement programs (Elbridge et al., 1993). The accurate identification of individuals and clones is important in breeding and deployment operations. Several studies have already shown that clone mis-identification may be relatively common (Keil and Griffin, 1994; Widen et al., 1994). It is therefore essential to have a cheap and practical means of proving clonal identity and fidelity. The PCR based RAPD (Randomly Amplified Polymorphic DNA; Welsh and McClelland, 1990; Williams et al., 1990) technique seems an obvious choice for this type of fingerprinting. The technique produces abundant loci, is relatively cheap and the technology is readily accessible to non-specialists (Rafalski and Tingey, 1993). Although the markers are dominant and there have been doubts raised about their reliability and reproducibility (Weeden et al., 1992; Penner et al., 1993), RAPDs have already been successfully used to identify clones (Duriham and Korban, 1994; Keil and Griffin, 1994; Lin et al., 1994) and varieties (Novy et al., 1994). They have also shown high levels of variability in all species so far examined, including *E. globulus* (Nesbitt et al., 1995).

In *Eucalyptus* most genetic improvement programs are currently exploiting families derived from open-pollinated seed collected from native stands where only the maternal parent is known for certain (Borralho and Potts, 1996). Eucalypts have a mixed mating system and such open-pollinated seed may result from selfing, or outcrossing between relatives and unrelated trees (Elbridge et al., 1993). Failure to account for variation in outcrossing rate at the individual level may bias estimates of genetic merit when these are based on open-pollinated progeny (Borralho and Potts, 1996; Hodges et al., 1996). Estimates of outcrossing levels in *Eucalyptus* have previously been made using allozymes (Moran and Bell, 1983; Sampson et al., 1989; House and Bell, 1994). In *Eucalyptus* only a relatively small number of polymorphic allozyme loci are available, particularly from adult leaf material (Moran and Bell, 1983). Therefore alternative markers such as RAPDs deserve investigation for this purpose. RAPDs detect more loci than allozymes, but are dominant and thus less informative per marker (Rafalski and Tingey, 1993). There is also a need to have a practical means of identifying relatives in breeding programs. Inbreeding and reduction of genetic diversity through incorporation of related individuals in breeding populations may prevent the realisation of expected genetic gains. RAPDs may offer a relatively cheap and easy way of identifying, and perhaps quantifying, the pedigree distance between individuals. However, to define the genetic distance between relatives, accurate knowledge of the genetic distance between unrelated individuals in the species is also required (Nesbitt et al., 1995).

The present study aims to confirm the stability of RAPD banding patterns at the clonal level and examine the relationship between RAPD and pedigree distance in *Eucalyptus globulus* ssp. *globulus*. The study is split into 2 main parts: firstly confirmation of clonal identity and the discrimination of unrelated, full-sibling, and half-sibling clones; and secondly the discrimination of selfed from outcrossed progeny. Previous work in *Eucalyptus* (Keil and Griffin, 1994) has shown differences between several full-sib clones, but more extensive studies analysing full-sibs from a wider range of crosses are required to establish confidence in this finding. As yet there are no studies examining the relationship between RAPD distance and pedigree distance in *Eucalyptus*.

Materials and Methods

Plant material

Discriminating clonal progeny

All clonal material are F₁ progeny resulting from interprovenance crosses of *Eucalyptus globulus* ssp. *globulus*. Ten ortets were collected from a trial established by North Forest Products and the CRC for Temperate Hardwood Forestry in NW Tasmania. Three ramets from each ortet were also collected from nearby trials. The parents of the ortets were collected from the Woolnorth seed orchard of North Forest Products. The 10 ortets came from 2 unrelated family groups, each comprising 2 crosses with a common maternal parent (Table 1). Six individual clones were collected from the first family group and 4 from the second. The common parent of family group one (MH4C) was crossed to 2 individuals, each from a different provenance (KISI and T4U), and 3 progeny from each cross were sampled. In the second family group the common parent (SF8N) was crossed to 2 individuals from the same provenance (S80 and S21), and 2 individuals from each cross were
Discriminating self-fertilised and outcrossed progeny

Samples of 10 selfed progeny and 24 outcrossed progeny were collected from a trial in north-west Tasmania (West Ridgley, established by North Forest Products and CSIRO Division of Forestry). All crosses shared the same maternal parent (KI2) and the 24 outcrossed progeny were from 12 controlled cross families, 2 progeny from each family being collected. A tissue sample from the maternal parent was collected from the Woolnorth seed orchard of North Forest Products and samples from the paternal parents were collected from wild stands. The paternal parents came from 2 provenances, 6 from the Taranna provenance and 6 from the King Island provenance of *E. globulus* ssp. *globulus*. The maternal parent was also from King Island, thus samples from 6 inter-provenance outcrosses and 6 intra-provenance outcrosses were collected (Table 2). A total of 47 samples was collected (10 selfed progeny, 12 pollen parents, 1 maternal parent and 24 outcrossed progeny). The progeny were each assigned a random number, in order to avoid bias in the scoring of bands. It was not until all results had been collated that the selfed progeny was distinguished from the outcrossed progeny. This design mimicked the usual case in open-pollinated progeny collection where the investigator does not know whether an individual is the result of an outcross or a self-pollination.

**DNA extraction and RAPD conditions**

Approximately 3 g of fresh, recently expanded leaf tissue was collected from each tree and frozen in liquid nitrogen. Tissue samples were stored at −80°C until use. A modified version of the DNA extraction procedure of *Doyle and Doyle (1990)* was used (*Nesbitt et al., 1995*). Recalcitrant DNA samples were further treated by extraction with phenolchloroform (1:1) before final dilution to approximately 10 ng/µL.

PCR amplification was carried out in a 20 µL volume in microtitre plates using an MJ Research Inc. PTC-100 (MA 02172 USA) programmable thermal controller. The reaction mixture contained 200 µM of each dNTP, 1 X Taq polymerase buffer, 150 mg/mL BSA, 3.0 mM MgCl₂, 10 pmoles primer, 20 ng DNA and 1 unit Taq polymerase. The reactions were overlaid with mineral oil and amplification was performed as follows: initial 1 min 92°C denaturation; 35 cycles of 1 min 92°C, 1 min 35°C, 2 min 72°C; and 5 min 72°C extension, using the fastest possible transitions between each temperature. Amplified fragments were separated in 1.4% agarose gel using 1 X TBE buffer and photographed after staining with ethidium bromide. Ten primers (Operon Technologies Inc., Alameda, CA) previously identified as detecting reproducible polymorphisms in *E. globulus* (*Nesbitt et al., 1995*) were used in discriminating selfed and outcrossed progeny: OPC-01, OPC-19, OPF-01, OPF-04, OPE-03, OPE-07, OPE-18, OPE-14, OPE-20 and OPD-05. The same set of primers, minus OPE-18, was used in the clonal study. Reproducible bands of consistent intensity were scored as present or absent for each individual. Reproducibility was judged from duplicated samples on the gels. In general bands were not scored if they were faint or diffuse, or occurred in regions of the gel demonstrating variable intensity (usually the extremes of the amplified size range). In the clonal study, if an individual could not be scored that whole clonal set (ortet plus three ramets) was repeated. This allowed for the best possible comparison of the clones. Bands that could not always be distinguished in these repeats or showed variable intensity were discarded for all individuals. Only the most reliable, reproducible bands were kept for analysis. Internal controls were maintained in the selfed versus outcrossed study to monitor band reproducibility.

**Statistical analyses**

The presence/absence data from each study was used to calculate a similarity matrix based on the simple matching
coefficient (SOKAL and SNETH, 1963), using the SIMQUAL procedure of NTSYS (NTSYS 1993). The simple matching coefficient is defined as the total number of positive (1) and negative (0) matches between 2 individuals, divided by the total number of bands scored. In the clonal study average genetic similarity for each full-sib and half-sib group of ortets was calculated based on the similarity matrix. UPGMA clustering of the similarity matrix was undertaken, for the clonal study, using the SAHN clustering procedure of NTSYS. For the selfing study, the matrix of simple matching coefficients amongst all individuals was similarly calculated. This matrix was used to compare the similarity of the selfs and outcrosses to the female (K12). The genetic material used in this study had five different levels of relatedness. The mean similarity among individuals at each level was calculated to examine the relationship between pedigree and genetic similarity. In addition, the similarity of each outcrossed progeny with each male parent was calculated and the number of times a progeny was most similar to its pedigree male parent tabulated in order to determine the efficiency of RAPDs for paternity analysis.

Results

Discriminating clonal progeny

The 9 primers amplified a total of 117 scorable bands with 12 to 15 bands amplified per primer. Of the 40 clonal samples, only 4 were found to possess bands not shared with other samples of that clone. In one case the ortet was found to contain a band not present in any of its ramets, and in the other cases the anomalous individual was a ramet. RAPDs were able to clearly discriminate between all clones in this study and UPGMA clustering (Figure 1) revealed the family structure incorporated into this study. The 2 unrelated family groups cluster separately and, with 1 exception, the parents fall within these groups. Within each family cluster, the fullsibs form separate clusters. The supposed common parent of family 2 (SF8N) is an outlier to the whole cluster. However, it is still more closely related to its own progeny than to unrelated progenies (Table 3), and is unlikely to be a pedigree error. The parent common to the half-sibs in family group 1 falls outside the family cluster as it has affinities to both full-sib groups. It is however, still separate from family group 2.

There is a clear association between pedigree relationship and RAPD similarity. Similarity based on RAPDs increases with closer levels of pedigree relationship. This trend is consistent both within each family group (Table 3) and for the combined data (Figure 2a). The genetic similarity amongst unrelated individuals ranges from 0.60 to 0.82 compared with 0.77 to 0.83 for half-sibs and 0.82 to 0.91 for full-sibs. There was little overlap between ranges of genetic similarity observed for full-sibs and half-sibs in both family groups (Table 3). Parent-offspring genetic similarities varied from an average value of 0.76 to 0.87, with individual values ranging from 0.74 to 0.89 (Table 3).

Discriminating self-fertilised and outcrossed progeny

Ten primers detected a total of 120 easily scorable bands. As in the first study, the more 2 individuals were related, the higher was the RAPD similarity between them (Figure 2b). The RAPD similarity between the female parent, K12, and its selfed progeny ranged from 0.85 to 0.93, and averaged 0.90 (Table 4). By contrast, the similarity of the outcrossed progeny to their female parent (K12) ranged from 0.73 to 0.88, and averaged 0.82. There was virtually complete separation of selfed and outcrossed progeny at a cut off similarity of 0.89. The only exception was selfed progeny number 18, but even then this progeny showed greater similarity to K12 (similarity = 0.85) than to any other parent in the study. This differentiation of self from outcrosses is also reflected in the number of non-maternal bands shown in Table 2 and the UPGMA dendrogram based on the similarity matrix shown in figure 3. Of the 24 outcrosses, only 10 showed greater similarity to their pedigree male parent than to any other male parent (data not shown). Therefore, in 10 of 24 outcrosses (42%) a correct paternity identification was made using this approach. The probability of
Figure 2. – Average RAPD similarity (simple matching coefficient) observed for different degree of relationship in (a) the clonal study, and (b) the study of self versus outcross progeny.

correct identification was not different across parents, since there was no tendency for both progeny of the same male to be identified correctly.

Discussion

This study demonstrated that with careful scoring, RAPDs are useful for confirmation of the clonal identity and pedigree structure in *E. globulus*, although their use in paternity analysis may be limited. After an initial screening to select suitable primers, it does not appear necessary to use a large number of primers in order to discriminate even closely related individuals. While some anomalous bands were identified in clonal sets, these differences were insignificant when compared to the differences found even between full-sibs. Similar rare discrepancies have been reported in a study of poplar clones (LIN et al., 1994). Such discrepancies may be due to a variety of factors, including competition amongst sites for primers (HEUN and HELENTJARIS, 1993), and may be reduced further in future studies by more stringent band selection. Such anomalous bands did not arise from ramet mis-labelling as many, if not all, primers would be expected to reveal differences in this case (KEIL and GRIFFIN, 1994; WIDEN et al., 1994). KEIL and GRIFFIN (1994) suggest that the mis-identification of plant material may be common in applied breeding and may have serious economic consequences. The present study clearly indicates that with stringent marker selection RAPDs are an effective and relatively cheap means of monitoring clonal identity.

Cluster analysis based on RAPD similarity was effective in uncovering the pedigree structure in the clonal study. A positive relationship between genetic distance and pedigree distance has been shown using RFLP (BRUNKLAUSJUNG et al., 1993; GERDES and TRACY, 1994; HALLIDEN et al., 1994; LORENZEN et al., 1995) and RAPD (BEEBE et al., 1995) markers in several crop species. In most cases the relationships revealed by markers were consistent with known pedigrees. These findings

Table 4. – Genetic similarities (simple matching coefficient) amongst *Eucalyptus globulus* samples from the study of self-fertilised and outcrossed progeny. The pedigree of the progeny is given in table 2.

<table>
<thead>
<tr>
<th>Pedigree relationship</th>
<th>Average RAPD similarity</th>
<th>Standard deviation</th>
<th>Minimum RAPD similarity</th>
<th>Maximum RAPD similarity</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between unrelated individuals</td>
<td>0.758</td>
<td>0.036</td>
<td>0.658</td>
<td>0.857</td>
<td>462</td>
</tr>
<tr>
<td>Between half-sibs</td>
<td>0.777</td>
<td>0.036</td>
<td>0.641</td>
<td>0.858</td>
<td>264</td>
</tr>
<tr>
<td>Parent/outcross offspring</td>
<td>0.812</td>
<td>0.044</td>
<td>0.700</td>
<td>0.883</td>
<td>48</td>
</tr>
<tr>
<td>Female parent (K2)/outcross</td>
<td>0.820</td>
<td>0.044</td>
<td>0.739</td>
<td>0.880</td>
<td>24</td>
</tr>
<tr>
<td>Self/half sibs</td>
<td>0.806</td>
<td>0.040</td>
<td>0.700</td>
<td>0.894</td>
<td>240</td>
</tr>
<tr>
<td>Between full sibs</td>
<td>0.818</td>
<td>0.040</td>
<td>0.758</td>
<td>0.843</td>
<td>12</td>
</tr>
<tr>
<td>Between selves</td>
<td>0.870</td>
<td>0.028</td>
<td>0.817</td>
<td>0.925</td>
<td>45</td>
</tr>
<tr>
<td>Female parent (K2)/selves</td>
<td>0.902</td>
<td>0.024</td>
<td>0.850</td>
<td>0.923</td>
<td>10</td>
</tr>
</tbody>
</table>

Figure 3. – Dendrogram showing the relationships among *Eucalyptus globulus* progeny and their parents. The dendrogram is based on a similarity matrix, calculated using the simple matching coefficient, and is generated using UPGMA. The numbers correspond to those in table 2. Selfed progeny are shown in bold. The maternal parent is K12.
are supported in both our RAPD studies. The differences among full-sibs, half-sibs and unrelated individuals of *E. globulus* are not large. However, they are consistent across the 2 studies, that use different genetic material. Furthermore, these levels of similarity amongst related trees are larger than the average similarity we calculated between all trees (0.66) in Nesbitt et al’s (1995) range-wide study of *E. globulus* sp. *globulus*, and between trees within their localities (mean for 31 localities = 0.73). Defining genetic distance or similarity ranges for the identification of a wide range of relationships (e.g. half-sib, full-sib and unrelated) may be possible in the future with more extensive sampling. For each species or population, these ranges would be dependent on the total variation in that species or population and the degree of heterozygosity of the parents. Using controls of known pedigree relationship as points of reference, it should be relatively easy to estimate the relationships of individuals of unknown origin. Thus a RAPD screening program may allow the avoidance of inbreeding in breeding populations by identifying related individuals. Screening of breeding populations would be particularly useful in species that are in the early stages of domestication or following mass selection where pedigrees are poorly known. For example, many breeding programs are exploiting open-pollinated seed collections derived from native stands where the male parent is unknown and in some cases the pedigree relationships among the females themselves are also poorly known. Genetic similarity may be of use when: (i) choosing among individuals with similar breeding values in order to maintain genetic diversity and reduce the potential for inbreeding depression; (ii) deciding whether to impose restrictions on representation of individuals from exceptional families or provenances; and (iii) when checking for errors in pedigree.

RAPD's have been successfully used to study the breeding system in several species (Huwy et al., 1993; Marshall et al., 1994). The utility of RAPD's in breeding system analysis is confirmed in the present study where selfs could be distinguished from both intra and inter-provenance half-sib outcrosses, with a high degree of accuracy. However, in the present case RAPD similarity was not as successful in identifying the correct outcross parental parent. This was due to many parents segregating for the same bands, coupled with some progenies possessing bands not found in the known parents. The average percentage of non-parental bands across all progeny was 2.5%. Non-homology of some bands and/or pollen contamination could explain this phenomenon, but high frequencies of such events are unlikely. A low number of non-parental bands were also found in the selfs (average of 1.6% of all bands), which could be due to the factors discussed previously for the clonal study, or have a genetic explanation such as recombination or competition amongst sites for primers (Heun and Helentarris, 1993). Davis et al. (1995) suggest that the formation of heteroduplexes from the amplification of co-dominant markers in heterozygous individuals may be an alternative explanation for rare non-parental RAPD bands.

Outcrossing rates in *Eucalyptus* have traditionally been estimated using allozymes using either single or multi-locus models (Moran and Bell, 1983; Sampson et al., 1989; House and Bell, 1994). Isozyme loci are generally assayed destructively at the germinant or young seedling stage. While extraction of isozymes from adult leaf material is technically feasible, not many loci can be resolved at this ontogenetic stage (Moran and Bell, 1983). As RAPD's do not have these limitations, they offer a means of determining outcrossing rates in later stages of the life cycle, if care is taken to score stable bands. In addition, as they appear to be ontogenetically stable (Keil and Griffin, 1994), they allow direct comparison across developmental stages. This is particularly important as adjustments required in estimating heritabilities from open-pollinated progenies are often required retrospectively, and for later stages in the life cycle (Borralho and Potts, 1996). The large number of loci revealed by RAPD's offers the possibility of differentiating selfing from bi-parental inbreeding (Brown, 1978) and in the absence of cheap co-dominant markers (e.g. microsatellites) they provide an alternative tool for breeding system analysis.

References

Genetic Variation in Disease Resistance of Juniperus virginiana and J. scopulorum Grown in Eastern Nebraska

By J. W. Zhang1), N. B. Kloppenstein2) and G. W. Peterson2)

(Received 6th May 1996)

Summary

Juniperus trees were examined in a plantation located in Horning State Farm, near Plattsmouth in eastern Nebraska, USA. Trees were grown from seeds collected from 131 open-pollinated families within 39 seed-zones in eastern redcedar (J. virginiana L.) and 25 open-pollinated families within 14 seed-zones in Rocky Mountain juniper (J. scopulorum SARG.) from their native ranges in the Great Plains of USA. The plantation was established in spring 1980 to examine genetic variation among genotypes for resistance to Cercospora blight caused by Pseudocercospora juniperi (ELLIS and EVERDL.) SUTTON and HODGES, comb. nov. (formerly Cercospora sequoiae var. juniperi), a major disease that threatens juniper survival east of the Rocky Mountains. All plantation plots were inoculated with P. juniperi in 1982, 1984, 1985, and 1986. Infection was scored in 1987. Kabatina tip blight caused by Kabatina juniperi SCHNEIDER and V. ARX, which occurred naturally, was scored also. Between the two juniper species, we found significant differences in survival and resistance to both diseases (P<0.05). All traits differed among seed-zones and among families within seed-zones of eastern redcedar (P<0.01). With Rocky Mountain juniper, variation was significant among seed-zones for survival (P<0.01), and among families within seed-zones for Cercospora blight and survival in 1994 (P<0.05). Heritabilities and genetic correlations were high for both disease resistance traits in eastern redcedar. Geographic patterns of genetic variation were identified; seed sources from southeastern collection sites of lower elevations tended to exhibit higher resistance to both diseases than seed sources from northwestern collection sites of higher elevations. Disease resistance traits were not correlated with height growth (r<0.20, P>0.05) for either species. The relationship between Cercospora blight resistance and survival in 1994 was significant (r = 0.59, P<0.05). Results indicate that Cercospora blight resistance in eastern redcedar can be improved by selecting resistant seed sources or families for direct reforestation programs or future breeding programs in eastern Nebraska. Additionally, Kabatina tip blight levels were lower on genotypes selected for resistance to Cercospora blight. Moreover, because resistance to Cercospora and Kabatina blights can be selected independently of height growth and survival, there is apparently no need to sacrifice growth and survival characteristics.

Key words: progeny test, disease resistance, Pseudocercospora juniperi, Kabatina juniperi, Cercospora blight, Kabatina tip blight, genetic parameters, selection, geographic variation, Great Plains.

FDC: 123211; 181.4; 165.3; 165.4; 165.5; 443; 172.8 Pseudocercospora juniperi; 172.8 Kabatina juniperi; 174.7 Juniperus; (782).

Introduction

Genetic improvement of any forest tree species requires a long-term commitment to a well designed selection and breeding program. With eastern redcedar (Juniperus virginiana L.) and Rocky Mountain juniper (J. scopulorum SARG.), this commitment is especially critical because, unlike many other coniferous species in which the selection programs are primarily focused on productivity, these species are used in urban and rural environmental plantings, such as shelterbelts, windbreaks, wildlife habitats, and landscapes (CUNNINGHAM, 1993). Therefore, sustainability is paramount when selecting adaptive provenances or families for environmental purposes. Assuring sustainability in stressful regions requires long-term studies on tree performance.

Although eastern redcedar and Rocky Mountain juniper can survive and grow in widely ranging climatic, edaphic, and topographic situations, 2 blights caused by Pseudocercospora juniperi (ELLIS and EVERDL.) SUTTON and HODGES, comb. nov. (formerly Cercospora sequoiae var. juniperi) and Kabatina juniperi SCHNEIDER and V. ARX are considered the greatest threats to long-term survival of these species in the Great Plains (PETERSON, 1981). In eastern Nebraska, for example, successive seasons of severe Cercospora blight can cause mortality within 3 years for 15- to 20-year-old trees (PETERSON, 1977).

Selection of genetic materials that resist these diseases is the primary approach to improve the sustainability of juniper plantings. In a shelterbelt containing a row of both juniper species, damage by Cercospora blight was much heavier in Rocky Mountain juniper than in eastern redcedar (PETERSON, 1981; PETERSON and WYSONG, 1968). However, specific genetic information is lacking for the disease resistance in these species, and general information on other characteristics is limited to a few studies. Working on multiple populations grown at various locations within the Great Plains of the USA, van Haverbeke and King (1990) reported that differentiation among populations was subtle, and genotype by environment interaction was significant in survival and height growth at age 5. Variation in volatile oil composition differed among 106 populations (COMER et al., 1982). Populations also differed in

1) Department of Forestry, Fisheries and Wildlife, University of Nebraska at Lincoln, East Campus-UNL, Lincoln, NE 68583-0822, USA
2) USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, National Agroforestry Center, East Campus-UNL, Lincoln, NE 68583-0822, USA