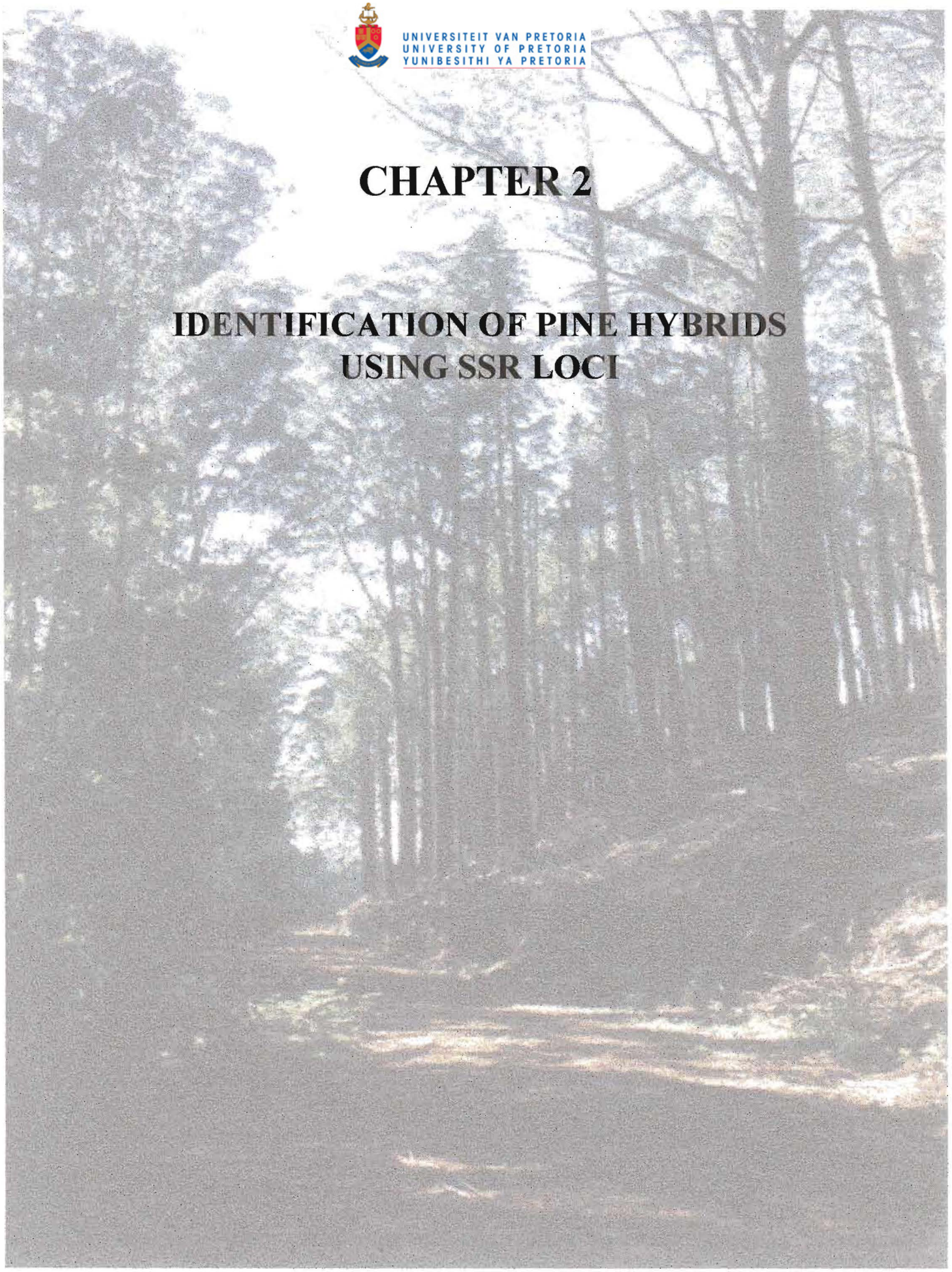

CHAPTER 2

IDENTIFICATION OF PINE HYBRIDS USING SSR LOCI

2.1 CROSS-SPECIES AMPLIFICATION OF SSR LOCI IN THE *PINACEAE* FAMILY

2.1.1 Introduction

Microsatellites, also called "simple sequence repeats", consist of tandemly repeated, short DNA motifs that provide an attractive source of genetic polymorphisms for both mammals and plants (Wang *et al.,* 1994; Weber and May, 1989). They have been found and used for genetic analysis in many mammalian species (Mellersh *et al.,* 1997; Moore *et al.,* 1991; Ostrander *et al.,* 1992) and in other eukaryotes, such as insects (Colson and Goldstein, 1999; Strassmann, 1998), birds (Groenen *et al.*, 1998; Moran, 1993), fish (Rico *et al.,* 1993) and plants (Akkaya *et al.,* 1992; Hokanson *et al.,* 1998; Lavi *et al.,* 1994; Thomas and Scott, 1993; Wang *et al.,* 1994).

In 1989 the first reports of microsatellite isolation and characterization of allelic variability at these loci using PCR were reported. In these early studies microsatellites were either cloned and sequenced (Litt and Luty, 1989; Tautz, 1989), or were identified from sequence databases (Weber and May, 1989). Based on this information primers were designed to recognize sequences flanking the microsatellite region. The polymorphic amplification products obtained after PCR with these specifically designed primers were then separated on polyacrylamide gels (Wu *et al.,* 1994). Using this approach the resolution of alleles differing by as little as one base pair was possible, and several loci could be analyzed together on the same gel (Rafalski and Tingey, 1993).

Given the long and expensive procedures involved in the isolation of SSR loci, it would be advantageous to be able to utilize primer sequences developed in one species in other closely related species. The problem is that while cross-species amplification is not uncommon, it does not always occur and amplification across genus boundaries tends to be unusual. Examples of cross-species amplification have been displayed in both the animal and plant kingdoms. Moore *et al.* (1991) found that high levels of sequence conservation exist within the primer regions of mammalian genomes in closely related species, while White and Powell (1997) found similar sequence conservation among *Meliaceae* species.

The aim of this study was to determine the conservation of microsatellite sequences between species and to use these heterologous PCR primers for the analysis of microsatellites in closely related species. The usefulness of these cross-species microsatellite markers was also investigated with regard to DNA profiling, gene flow and parental contribution determinations in the *P. elliottii x P. caribaea* hybrid.

2.1.2 Materials and Methods

2.1.2.1 Plant material

For the production of the *P. elliottii x P. caribaea* hybrids the *P. elliottii* tree E 503 was used as the female parent. Various *P. caribaea* tree types (i.e. Ach 33, Ach 49, Ach 93, Ach 271 and Pch 23) were used as the pollen donors. All plant material was obtained

from SAFCOL, South Africa. Diploid needle and zygotic embryo tissues, as well as haploid megagametophyte and pollen tissues were used to obtain genomic DNA.

2.1.2.2 DNA preparation

The extraction methods used to obtain genomic DNA's required for this study are summarized in Table 2.1.

Table 2.1. Plant Material and extraction methods used in the study.

Plant Material	Extraction Method	Reference
P. elliottii needle tissue	CTAB	Murray and Thompson, 1980
P. elliottii megagametophyte tissue	Lee – Binelli modification	Lee and Taylor, 1990 Binelli and Bucci, 1994
P. elliottii x P. caribaea hybrid zygotic embryo tissue	Lee – Binelli modification	Lee and Taylor, 1990 Binelli and Bucci, 1994
P. caribaea needle tissue	CTAB	Murray and Thompson, 1980
P. caribaea pollen	PCR directly	Petersen *et al.*, 1996

The extraction of the *P. elliottii* and *P. caribaea* needle tissue DNA was done using a modified CTAB method (Murray and Thompson, 1980). Bulked needle samples of approximately 3 g were ground to a fine powder in liquid nitrogen using a sterile mortar and pestle. Thereafter, 3 ml of heated CTAB extraction buffer (5% (w/v) CTAB, 1.4 M NaCl, 0.2% (v/v) 2-mercaptoethanol, 20 mM EDTA, pH 8.0, 100 mM Tris-HCl, pH 8.0,

1% (w/v) PVP) was added to the frozen tissue, which was then ground further to form a paste. This tissue paste was transferred to a large test-tube, and the mortar was rinsed with a further 3 ml of CTAB extraction buffer. The mixture was incubated at 65°C for one hour, after which an equal volume of phenol / chloroform / isoamyl alcohol (25:24:1) was added and the contents mixed by inverting. The mixture was centrifuged (12 000 x *g* for 15 min) at room temperature, and the aqueous phase was transferred to a new tube. Two and a half (v/v) times ice cold, 100% ethanol was added to the aqueous phase and gently mixed by inverting. The mixture was left on ice for up to 5 min to allow for complete DNA precipitation. After 5 min the precipitated DNA was spooled using a sterile Pasteur pipette. The DNA was washed twice in 70% ethanol; vacuum dried and resuspended in 50 to 100 µl of sterile water. The DNA concentration was determined spectrophotometrically and all samples were subsequently diluted to 50 ng/µl.

The extraction of DNA from *P. elliottii* megagametophyte tissue and *P. elliottii x P. caribaea* hybrid zygotic embryo tissue was performed using a modification of the methods described by Lee and Taylor (1990) and Binelli and Bucci (1994). The procedure is similar to the CTAB extraction method previously described, with the following exceptions. Firstly, the individual zygotic embryos and megagametophytes were ground to a fine powder, in eppendorf tubes, in liquid nitrogen using sterile glass pestles, as opposed to the bulked needle samples ground in a mortar and pestle. Subsequently 200 µl of heated extraction buffer (50 mM Tris-HCl, pH 7.5, 50 mM EDTA, pH 8.0, 3% (w/v) SDS, 1% (v/v) 2-mercaptoethanol), as opposed to 3 ml CTAB extraction buffer, was added and the tubes were vortexed briefly prior to incubation at

65°C for one hour. Secondly 5 µl of 3 M NaOAc, pH 8.0, was added to the aqueous phase prior to the addition of 2.5 (v/v) 100% ice-cold ethanol. Finally, the vacuum dried DNA was resuspended in a maximum of 50 µl of sterile water and was quantified spectrophotometrically.

2.1.2.3 Primers

Eleven different primer pairs specific to microsatellite loci identified in other *Pinus* species were selected for the DNA profiling of the *P. elliottii / P. caribaea* population (Table 2.2). Of the eleven microsatellite markers screened, two were developed in *Pinus radiata* (PR 4.6 and PR 9.3), while six of the eleven were developed in *Pinus taeda* (RPTest 2, RPTest 5, RPTest 6, RPTest 9, RPTest 15 and RPTest 20). The remaining three of the eleven microsatellite markers were developed in *Pinus strobus* (RPS 105, RPS 150 and RPS 160), which is a soft pine of the subgenus *Strobus,* while both *Pinus radiata* and *Pinus taeda* are hard pines of the subgenus *Pinus* (Echt *et al.,* 1999).

2.1.2.4 PCR conditions

PCR conditions were based on a modification of conditions suggested by Smith and Devey (1994). All amplification reactions were carried out in 25 µl reaction volumes. The reaction mixture contained 50 ng/µl template DNA, 1% (v/v) Promega magnesium free, thermophilic DNA polymerase 10X buffer, (10 mM Tris-HCl, pH 9.0, 50 mM KCl, 0.1% (v/v) Triton®X-100), 300 µM dNTPs, 3 mM $MgCl_2$, 0.25 µM of each primer, and

Table 2.2. List of microsatellite primers used in the study.

Primer	Sequence (5' – 3')	Reference
PR 4.6 (forward)	GAAAAAAAGGCAAAAAAAAGGAG	Smith and Devey, 1994
PR 4.6 (reverse)	ACCCAAGGTACATAACTCG	
PR 9.3 (forward)	GAAATTTAACACCACACCGTTG	Smith and Devey, 1994
PR 9.3 (reverse)	TGGGGCTTAAAGTGAAATGG	
RPTest 2 (forward)	AATCCTTGTGCCAGGAACAC	Echt and Burns, 1999
RPTest 2 (reverse)	ATTTTATCAGCGCCACTGCT	
RPTest 5 (forward)	ACAACAATAATAACGGGGGC	Echt and Burns, 1999
RPTest 5 (reverse)	ACGCTTTAGATCCTCCTGCA	
RPTest 6 (forward	AGGATTCCAACAGCATCACC	Echt and Burns, 1999
RPTest 6 (reverse)	CTGAACATGAAGCGCAGTGT	
RPTest 9 (forward)	CCAGACAACCCAAATGAAGG	Echt and Burns, 1999
RPTest 9 (reverse)	GCCTGCTATCGAATCCAGAA	
RPTest 15 (forward)	GAACGTGGTTATGGCGGTAG	Echt and Burns, 1999
RPTest 15 (reverse)	CCAGGGACAGTTACCAGCAT	
RPTest 20 (forward)	GTTCCCACTCAAGGGTTGAA	Echt and Burns, 1999
RPTest 20 (reverse)	ACATCATTTGTTGCCGCATA	
RPS 105 (forward)	TGGACATCCTAGTCGGAACC	Karhu *et al.*, 2000
RPS 105 (reverse)	AAAATCATTTCTGTATCAGAACAA	
RPS 150 (forward)	TCCATCAGTGAGCAGTGG	Karhu *et al.*, 2000
RPS 150 (reverse)	CACTTGGGCTTCCTCTTC	
RPS 160 (forward)	ACTAAGAACTCTCCCTCTCACC	Karhu *et al.*, 2000
RPS 160 (reverse)	TCATTGTTCCCCAAATCAT	

1 U Promega *Taq* DNA polymerase in storage buffer A (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 50% (v/v) glycerol, 0.5% (v/v) Tween®20, 1% (v/v) Triton®X-100). Temperature cycling was done using a Perkin Elmer GeneAmp® PCR system 9700 (Applied Biosystems) thermal cycler programmed for an initial step of 4 min at 94°C, followed by 30 cycles of 45 s at 94°C, 45 s at 55°C and 45 s at 72°C. The amplification cycles were followed by a final elongation step of 5 min at 72°C.

A modification of the protocol described by Petersen *et al.* (1996), in combination with PCR conditions described by Smith and Devey (1994), were used to obtain amplification products directly from *P. caribaea* pollen grains, without prior DNA isolation. The procedure involved the suspension of 50 to 100 pollen grains in distilled water, in an eppendorf tube, and grinding briefly with a glass pestle to facilitate the release of the DNA from the pollen grains. This suspension was then mixed with the Promega thermophilic DNA polymerase 10X buffer, magnesium free, and heated to 95°C for 10 min to destroy all enzymatic activity and to burst any intact pollen grains which might still be present. The PCR amplifications were subsequently performed in 25 µl reactions as described previously.

2.1.2.5 PAGE analysis

The amplification products were analyzed by separation on 7.5% non-denaturing polyacrylamide gels (3.75 % (v/v) FMC® Long Ranger Gel Solution, 1% (v/v) TBE buffer, 0.1% (w/v) APS, 0.004% (v/v) TEMED) in 1 X TBE buffer (45 mM Tris-borate, 1 mM EDTA, pH 8.0) at 26.6 V/cm for one and a half hours. After separation the amplification products were detected by staining with (1:10 000) Sybr®GreenI Nucleic Acid Gel stain (Boehringer Mannheim). The Amplification products were visualized by UV illumination. Boehringer Mannheim 100 bp DNA ladder was included as the molecular size marker in all electrophoresis profiles.

2.1.3 Results

Of the eleven microsatellite markers included in this study two primer pairs were found to produce informative polymorphisms (primer PR 4.6 and primer RPTest 9), five were found to be monomorphic (primer RPTest 5, primer RPTest 15, primer RPTest 20, primer RPS 105 and primer RPS 150) for the entire *P. elliottii* and *P. caribaea* population. Four primer pairs did not result in any amplification products (primer PR 9.3, primer RPTest 2, primer RPTest 6 and primer RPS 160). The results of the eleven SSR markers that were screened are shown in Table 2.3.

The polymorphic microsatellite markers, PR 4.6 and RPTest 9, produced amplification products that could be used to clearly distinguish between the *P. elliottii* and *P. caribaea*

(Figure 2.1, Figure 2.2 and Figure 2.3). These polymorphic amplification products could also be used to determine which parent contributed which allele to the *P. elliottii x P. caribaea* hybrid of a specific cross (Figure 2.1). In Figure 2.1 the 250 bp fragment in lane 4 of the microsatellite fingerprints, of the E 503 x Pch 23 hybrid population, clearly suggest that Pch 23 was definitely used as the pollen donor during the controlled pollination.

Table 2.3. Results of SSR cross-species amplification.

SSR marker	***Pinus* species for which marker was developed**	**Number of loci**	**Polymorphism**
PR 4.6	*Pinus radiata*	2	Polymorphic
PR 9.3	*Pinus radiata*	0	None
RPTest 2	*Pinus taeda*	0	None
RPTest 5	*Pinus taeda*	1	Monomorphic
RPTest 6	*Pinus taeda*	0	None
RPTest 9	*Pinus taeda*	2	Polymorphic
RPTest 15	*Pinus taeda*	1	Monomorphic
RPTest 20	*Pinus taeda*	1	Monomorphic
RPS 105	*Pinus strobus*	1	Monomorphic
RPS 150	*Pinus strobus*	1	Monomorphic
RPS 160	*Pinus strobus*	0	None

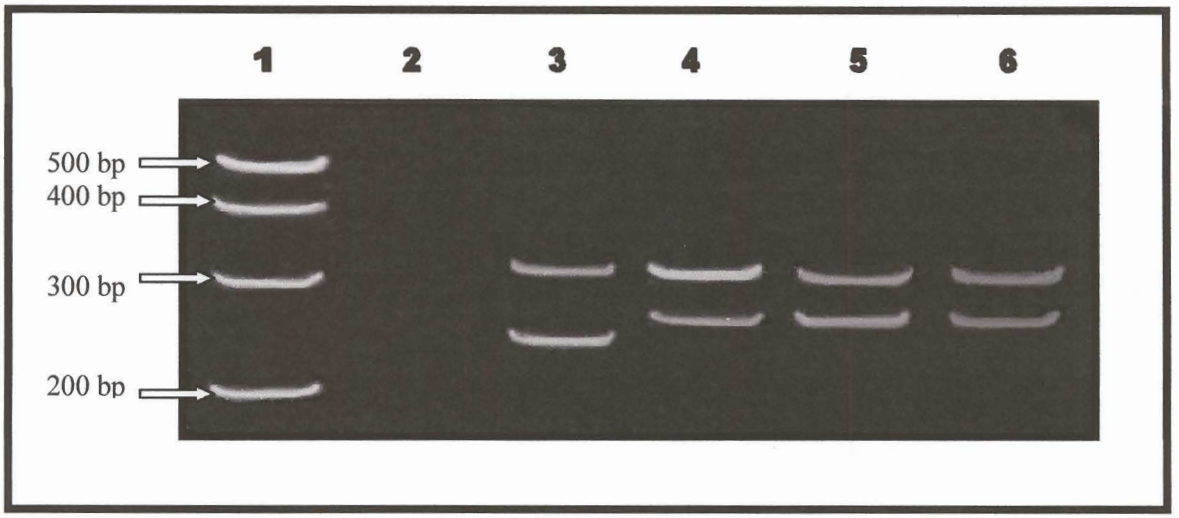

Figure 2.1. Microsatellite fingerprints of the *P. elliottii* (E 503) x *P. caribaea* (Pch 23) hybrid population. The fingerprints were obtained by PCR amplification with primer RPTest 9. Lane 1 contains 100 bp DNA ladder, while lanes 2 to 6 contain the fingerprints obtained using E 503 megagametophyte DNA, E 503 needle DNA, E 503 x Pch 23 embryo DNA, Pch 23 needle DNA and Pch 23 pollen DNA, respectively.

In Figure 2.2, in which the microsatellite fingerprints of the E 503 x Ach 271 hybrid population are recorded, there is no definitive evidence as to the gene flow that occurred during fertilization. If the controlled pollination was successful, then the hybrid must have inherited the polymorphic allele (i.e. the 275 bp fragment in lane 4, Figure 2.2) from *P. elliottii* and not from *P. caribaea.* Another possibility is that the controlled pollination was not successful and that the embryo produced was actually the result of *P. elliottii* self fertilization.

For specific hybrid populations, such as the E 503 x Ach 49 hybrid population presented in Figure 2.3, null alleles were detected using polymorphic marker RPTest 9. The absence of an amplification product of between 250 – 280 bp was verified by altering the stringency of the PCR conditions to ensure that the result was not merely due to poor amplification at that locus.

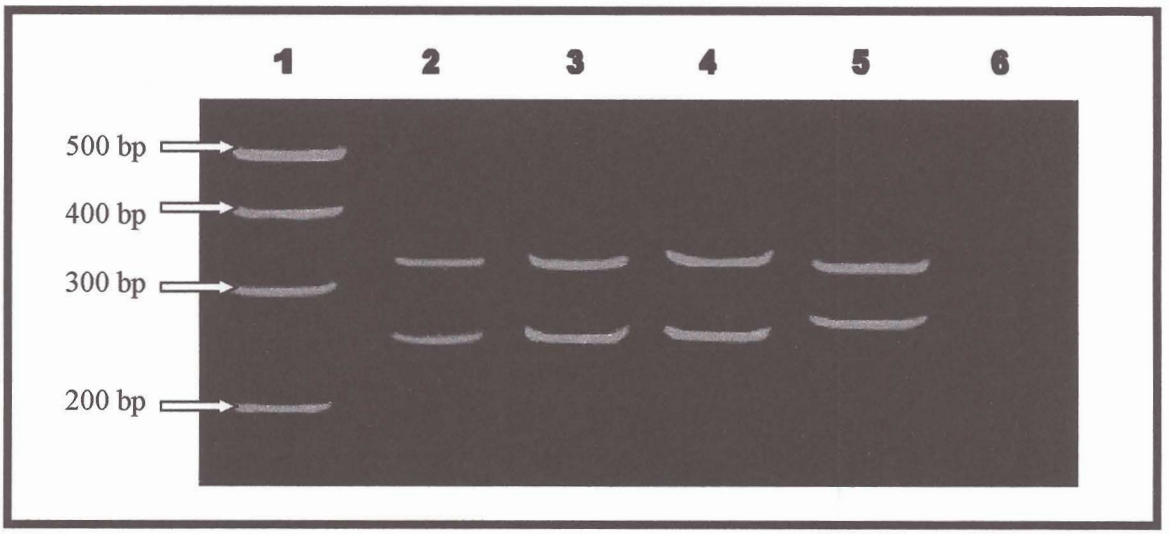

Figure 2.2. Microsatellite fingerprints of the *P. elliottii* (E 503) x *P. caribaea* (Ach 271) hybrid population. The fingerprints were obtained by PCR amplification with primer RPTest 9. Lane 1 contains 100 bp DNA ladder, while lanes 2 to 6 contain the fingerprints obtained using E 503 megagametophyte DNA, E 503 needle DNA, E 503 x Pch 23 embryo DNA, Ach 271 needle DNA and Ach 271 pollen DNA, respectively.

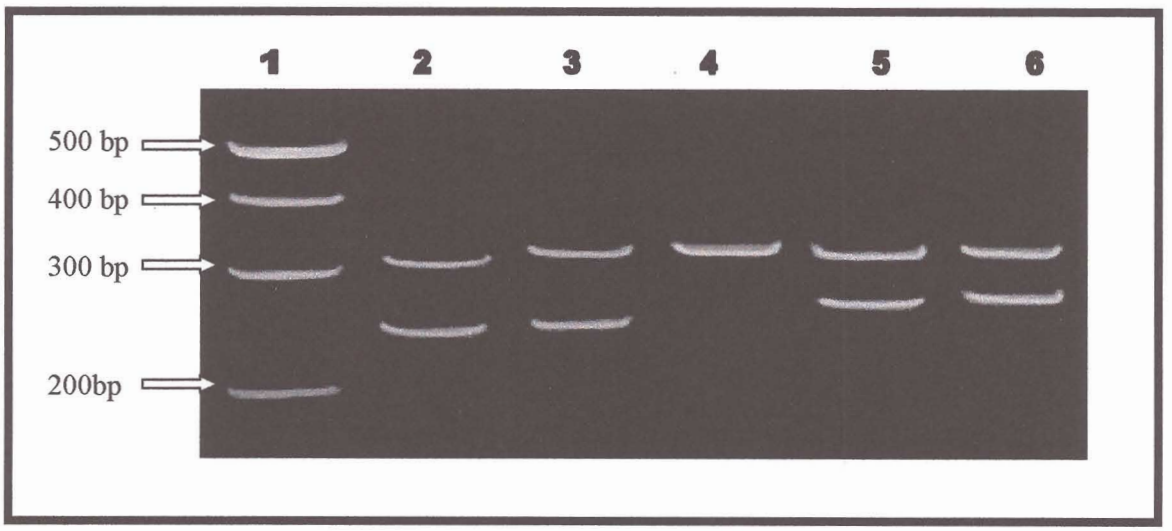

Figure 2.3. Microsatellite fingerprints of the *P. elliottii* (E 503) x *P. caribaea* (Ach 49) hybrid population. The fingerprints were obtained by PCR amplification with RPTest 9. Lane 1 contains 100 bp DNA ladder, while lanes 2 to 6 contain the fingerprints obtained using E 503 megagametophyte DNA, E 503 needle DNA, E 503 x Pch 23 embryo DNA, Ach 49 needle DNA and Ach 49 pollen DNA, respectively.

2.1.4 Discussion

The evolutionary conservation of DNA sequences flanking microsatellite regions allows the cross-species amplification of microsatellite markers in various related species (Echt *et al.*, 1999). According to Primmer *et al.* (1996) the probability of successful cross-species amplification of microsatellites seems to correlate with the taxonomic distance between the species. This would imply that cross-species amplification, while not uncommon, does not always occur and the greater the taxonomic distance between the species the smaller the chance that it will occur.

Pinus elliottii and *Pinus caribaea* both belong to the subgenus *Pinus*, the hard pines (Wright, 1976). The eleven microsatellite markers screened were developed for both the taxonomic subgenera *Pinus* (*P. radiata* and *P. taeda* - the hard pines) and *Strobus* (*Pinus strobus* – the soft pines). Of the eleven microsatellite markers screened, seven (63.6%) produced DNA fingerprints (Table 2.3). Of these seven markers that produced DNA fingerprints, only two (28.6%) produced polymorphic fingerprints, while the remaining five (71.4%) produced monomorphic DNA fingerprints. Both of the polymorphic microsatellite markers were developed using species in the subgenus *Pinus*. The monomorphic microsatellite markers were derived from both the *Pinus* and *Strobus* subgenera. These results are consistent with other cross-species amplification results, where suitable informative markers were obtained from only 20 – 24% of primer pairs screened (Echt *et al.*, 1996; Pfeiffer *et al.*, 1997). Relatedness, therefore, seems to play an

important role in the type of polymorphism (i.e. polymorphic vs. monomorphic) observed, as well as the success rate encountered during cross-species amplification.

The important role which microsatellite markers can play in population studies is demonstrated by their two-fold ability to detect polymorphisms between the amplification products of both *P. elliottii* and *P. caribaea* parents, and to determine gene flow and parental contribution in *P. elliottii x P. caribaea* hybrid progeny. According to Queller *et al.* (1993) if the genotypes of the mother and its offspring are known, then it is straightforward to exclude incompatible paternal genotypes, but all other males must be excluded before the father can be assigned. The results tend to indicate that microsatellite fingerprints, of the hybrid populations, obtained using single polymorphic primer pairs are not always conclusive (Figure 2.2). However, in studies using more than one polymorphic microsatellite marker it is possible to clearly distinguish between species and hybrid material and to eliminate some of the uncertainty surrounding gene flow.

Not all polymorphic markers produce DNA fingerprints that are equally informative. The usefulness of a marker is determined by its degree of polymorphism in the species that is being tested and the extent to which gene flow and parental contribution can be determined from the fingerprint produced. For example, the polymorphic fingerprint produced by primer RPTest 6 on the E 503 x Pch 23 hybrid population (Figure 2.1) could be considered useful, as it is possible to distinguish the E 503 maternal parent from the Pch 23 pollen donor using this marker. In the same way this marker can also be used to determine gene flow and parental contribution, as the E 503 x Pch 23 hybrid fingerprint

mirrors the fingerprint obtained for the Pch 23 pollen donor. This means that based upon the DNA fingerprints produced by primer RPTest 6 it is possible to accept the null hypothesis that the embryos were the result of the controlled pollination between the *P. elliottii* and *P. caribaea* parents.

The polymorphic fingerprint produced by primer RPTest 6 on the E 503 x Ach 271 hybrid population (Figure 2.2) is less informative as the E 503 x Ach 271 hybrid fingerprint mirrors that of the E 503 maternal parent and it is therefore impossible to determine gene flow or parental contribution. In this scenario two possible options exist. Firstly, that the controlled pollination was unsuccessful and that the E 503 x Ach 271 "hybrid" was actually the progeny of E 503 self-fertilization. The second option is that it is assumed that the controlled fertilization between E 503 maternal parent and the Ach 271 pollen parent was successful, in which case the E 503 x Ach 271 hybrid must have inherited the polymorphic allele (i.e. the 275 bp fragment in lane 4, Figure 2.2) from the E 503 maternal parent. This means that based upon the DNA fingerprints produced by primer RPTest 6 it is not possible to reject or accept the null hypothesis that the embryos were the result of the controlled pollination between the *P. elliottii* and *P. caribaea* parents.

Monomorphic microsatellite markers are not useful as it is impossible to distinguish between species or to determine gene flow or parental contribution. However, the fact that cross-species amplification does occur indicates that sequence conservation does exist within the genus *Pinus* and that those specific loci were more conserved than others.

A second type of uninformative microsatellite marker is a marker that produces fingerprints that contain null alleles. Null alleles are the manifestations of mutations in one or both of the microsatellite primer binding regions (Callen *et al.*, 1993). These mutations inhibit the primers from binding and may result in the loss of the PCR product. An example of the uninformative nature of null alleles, in the current study, is evident in the microsatellite fingerprints produced by primer RPTest 6 in the E 503 x Ach 49 hybrid population (Figure 2.3). From these fingerprints it is impossible to distinguish the E 503 maternal parent from the Ach 49 pollen donor as the polymorphic allele (i.e. a fragment of 250 – 280 bp) is absent. This means that based upon the DNA fingerprints produced by primer RPTest 6 it is impossible to accept the null hypothesis that the embryos were the result of the controlled pollination between the *P. elliottii* and *P. caribaea* parents as the polymorphic allele is absent.

The ability to determine gene flow and parental contribution via cross-species amplification of microsatellite markers, holds important implications for the application of this technique on a large scale in the forestry industry. This technique not only reduces the time and costs involved in marker development, but also the time and costs involved during the nursery phase and the selection of superior clones for future plantations. For example, once superior *P. elliottii x P. caribaea* hybrids are identified, based on field trial performance, it will be possible to screen the seedlings produced in order to verify that they really are the progeny of the controlled pollination between the superior parents. This procedure will allow seedlings that do not conform to the microsatellite fingerprints

expected for the hybrid to be discarded early on in the screening phase and ensure that only the correct genotypes enter into the bulking-up phase.

2.1.5 Conclusion

The fact that microsatellite markers developed in other species can be used successfully in the DNA profiling of related species, makes this an extremely powerful and accessible technique. This cross-species amplification also greatly reduces the cost and time requirements involved in DNA profiling studies, by eliminating the need for sequence information and marker development (Echt *et al.,* 1999; Weising and Gardner, 1999).

Not all microsatellite markers produce DNA fingerprints that are equally informative and useful. The usefulness of a marker is determined by its degree of polymorphism in the species that are being tested and the extent to which gene flow and parental contribution can be determined from the fingerprints produced. Polymorphic microsatellite markers tend to be highly informative and are therefore far more useful than monomorphic markers, which tend to be relatively uninformative.

The polymorphism displayed by microsatellites makes them useful as markers in linkage studies (Moore *et al.,* 1991). However, parental determination should generally be approached via a process of elimination in order to accurately accept or reject the null hypothesis that the embryo was the result of successful controlled pollination. It is therefore important that information from several loci be pooled and analyzed together,

as it is unlikely that a single locus will provide enough information upon which the null hypothesis may be rejected or accepted (Strassmann, 1998).

2.2 DEVELOPMENT OF SSR MARKERS IN *PINUS ELLIOTTII X PINUS CARIBAEA* HYBRIDS

2.2.1 Introduction

Microsatellites are repeated sequences that consist of 1 – 5 bp which are scattered throughout the eukaryotic genome (Litt and Luty, 1989). Variations in the number of these tandem repeats result in the generation of extensive polymorphisms (Hearne *et al.*, 1992). These polymorphisms make microsatellite loci extremely useful markers for genetic mapping, DNA typing and population genetic studies (Dayanandan *et al.*, 1998; Dib *et al.*, 1996; Hokanson *et al.*, 1998; Schmidt *et al.*, 2000).

In early studies microsatellites were either cloned and sequenced (Litt and Luty, 1989; Tautz, 1989), or were identified from sequence databases (Weber and May, 1989). Based on this information primers were designed to recognize sequences flanking the microsatellite region. The polymorphic amplification products obtained after PCR with these specifically designed primers were then separated on polyacrylamide gels (Wu *et al.*, 1994).

In 1994 Zietkiewicz *et al.* demonstrated a novel approach to measuring genetic diversity in plants and animals by using primers based on microsatellites. The Zietkiewicz technique combined the benefits of microsatellite and RAPD analysis (Hantula *et al.*, 1996; Hantula and Müller, 1997). Two types of polymorphism can be identified using this random amplified microsatellite (RAMs) technique. Firstly the on / off polymorphisms, similar to those found using RAPD markers, as variation is due to existence or absence of a marker. Secondly, polymorphism due to length differences in the sizes of the amplification products (Hantula *et al.*, 1996). The RAMs technique involved the amplification of DNA between the distal ends of two closely located microsatellite regions, by using random microsatellite primers (Zietkiewicz *et al.*, 1994; Hantula *et al.*, 1996, 1998).

The aim of this study was to develop microsatellite markers in the *P. elliottii x P. caribaea* hybrid population, using the RAMs approach, and to determine how informative these microsatellite markers could be with regard to DNA profiling.

2.2.2 Materials and Methods

2.2.2.1 Plant material and DNA isolation

Needle tissue of the female parent tree, *P. elliottii* E 503, supplied by SAFCOL, South Africa, was used as the main source of plant material in this study. DNA was isolated from this needle tissue using a modified CTAB method (Murray and Thompson, 1980) as described in section 2.1.2.2.

2.2.2.2 Primers and PCR conditions

Three different microsatellite primers were screened in this study, namely CA(8) (Appendix A), CT(8) and GTG(5). These primers were selected for their ability to produce amplification profiles linked to random microsatellites throughout the genome of *P. elliottii.*

All amplification reactions were carried out in 25 µl reaction volumes and DNA template concentrations ranging from 12.5 ng/µl to 75 ng/µl were tested. The reaction mixture contained template DNA, 1% (v/v) Promega thermophilic DNA polymerase 10X buffer, magnesium free (10 mM Tris-HCl, pH 9.0, 50 mM KCl, 0.1% (v/v) Triton®X-100), 300 µM dNTPs, 3 mM $MgCl_2$, 0.5 µM primer, and 1 U Promega *Taq* DNA polymerase in storage buffer A (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.1 mM EDTA 1 mM DTT, 50% (v/v) glycerol, 0.5% (v/v) Tween®20, 1% (v/v) Triton®X-100). Temperature cycling was done using a, Perkin Elmer GeneAmp® PCR system 9700 (Applied Biosystems), thermal cycler programmed for an initial step of 10 min at 94°C, followed by 30 cycles of 30 s at 94°C, 45 s at 47°C and 2 min at 72°C. The amplification cycles were followed by a final elongation step of 7 min at 72°C. The amplification products were analyzed by electrophoresis on 3% agarose gels (Whitehead Scientific) in 1 X TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) at 5.3 V/cm for 30 min. The amplification products were detected by staining with ethidium bromide (0.5 µg/ml) and visualized by exposure to UV illumination. Molecular size marker III, prepared by

digesting lambda DNA with restriction enzymes *Eco*RI and *Hin*dIII, was included in all electrophoresis profiles.

2.2.2.3 Cloning and transformation

In order to ensure the cloning of a variety of amplification products, the profile displaying the greatest number and variety of bands was selected. Impurities, which may have influenced the effectiveness of the cloning and transformation reactions, were removed from the amplification products using the QIAquick™ PCR purification kit (QIAGEN). The amplification products were then randomly cloned into the T7 / SP6 cloning site of the pGEM®-T Easy plasmid vector, before being transformed into *Escherichia coli* JM 109 competent cells using the Promega pGEM®-T EASY vector system. The transformed *E. coli* cells were then plated out, at various concentrations, on LB plates (1% (w/v) Bacto-tryptone, 0.5% (w/v) Bacto-yeast extract, 1% (w/v) NaCl, 1.5% (w/v) Agar, pH 7.4) containing Ampicillin (50 mg/ml), IPTG (100 mM) and XGAL (2% w/v) and incubated at 37°C for 16 to 24 h. Transformed colonies, identified using the blue / white selection criteria of the IPTG / XGAL system, were used to inoculate liquid LB, containing Ampicillin (50 mg/ml), and were incubated for 24 h at 37°C.

2.2.2.4 Colony screening

In order to identify transformed colonies containing inserts of interest screening was done directly on each colony using 25 µl PCR reactions with primers homologous to the flanking regions of the recombinant plasmids. DNA was added to each reaction by lightly pressing a sterile pipette tip to the selected colony, inserting the tip into the eppendorf tube containing the reaction mixture and pipetting up and down. The reaction mixtures contained 1% (v/v) Promega thermophilic DNA polymerase 10X buffer, magnesium free, 100 µM dNTPs, 1 mM $MgCl_2$, 0.5 µM of each microsatellite primer, 0.5 µM primer SP6, 0.5 µM primer T7 and 1 U Promega *Taq* DNA polymerase in storage buffer A. Temperature cycling was done using a Perkin Elmer GeneAmp® PCR system 9700 (Applied Biosystems), thermal cycler programmed for 25 cycles of 30 s at 94°C, 45 s at 50°C and 2 min at 72°C, followed by an elongation step of 5 min at 72°C. The amplification products were analyzed by electrophoresis on 2% agarose gels in 1 X TAE buffer at 5.3 V/cm for 20 min. The amplification products were detected by staining with ethidium bromide (0.5 µg/ml) and visualized by exposure to UV illumination. Molecular size marker III was included in all electrophoresis profiles.

2.2.2.5 Plasmid isolation and PEG precipitation

Based on the amplification profiles obtained during the screening procedure, transformed colonies were selected and their plasmids isolated using the Nucleospin® Plasmid kit (Macherey-Nagel) The isolated plasmid DNA was then purified further by polyethylene

glycol (PEG) precipitation. The precipitation involved the addition of 625 mM NaCl and 9.75% (w/v) PEG_{6000} to the isolated plasmid DNA. The mixture was mixed gently by pipetting and left on ice for 30 min. The mixture was centrifuged at 12 000 x *g*, at 4°C, for 30 min. The aqueous phase was discarded and the pellet was washed twice in 70% ethanol. The pellet was air-dried, at 40°C on a heating block, until all traces of ethanol evaporated. The pellet was resuspended in 20 µl sterile water and the DNA concentration was determined spectrophotometrically.

2.2.2.6 Sequencing

Sequencing of the insert was done on a Perkin Elmer GeneAmp® 96000 thermal cycler, using the ABI Prism™ Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) according to manufacturers recommendations. The cycle sequencing consisted of 25 cycles of 10 s at 96°C, 5 s at 50°C and 4 min at 60°C. The sequencing reaction was followed by 0.7 X (v/v) 95% ethanol and 42 mM sodium acetate, pH 4.6, precipitation. The mixture was then left on ice for 10 min before being centrifuged at 12 000 x *g* for 30 min. The supernatant was discarded and the pellet was rinsed twice in 70% ethanol prior to air-drying on a heating block, at 40°C, until all traces of ethanol had evaporated. The DNA sequence was subsequently obtained using a Perkin Elmer ABI 377 automated sequencer and analyzed using the Sequence Navigator™ programme, version 1.01.

2.2.2.7 Primer development and testing

Once analyzed the sequence was searched for regions containing microsatellites. Primers, of between 17 and 21 nucleotides and of G + C percentages of ca. 50% - 70%, were designed in the regions flanking the microsatellites. The primer design utility on the Cybrgene AB website (http://www.cybergene.se) was used to ensure that the primer sequences selected did not contain hairpin loops or palindromic sequences, which would reduce the efficiency of the primer-template interaction.

All of the primers were subsequently tested in 25 µl PCR reactions. The reaction mixture contained 50 ng/µl *P. elliottii* E 503 needle DNA, 1% (v/v) Promega thermophilic DNA polymerase 10X buffer - magnesium free, 300 µM dNTPs, 3 mM $MgCl_2$, 0.25 µM of each primer, and 1 U Promega *Taq* DNA polymerase in storage buffer A. Temperature cycling was done using a Perkin Elmer GeneAmp® PCR system 9700 (Applied Biosystems). The reaction conditions were optimized for each primer by starting at an annealing temperature 5°C lower than the Tm of the primers and gradually increasing this temperature by 5°C with each subsequent PCR. The thermal cycler was programmed for an initial step of 4 min at 94°C, followed by 30 cycles of 45 s at 94°C, 45 s at various annealing temperatures and 45 s at 72°C. The amplification cycles were followed by a final elongation step of 5 min at 72°C. The PCR reactions were screened for amplification products on 2% agarose gels in 1 X TAE buffer, at 5.3 V/cm for 20 min, and were detected by staining with ethidium bromide (0.5 µg/ml). Molecular size marker III was included in all electrophoresis profiles.

PCR reactions which were found to contain amplification products were further analyzed by separation on 7.5% non-denaturing polyacrylamide gels (3.75 % (v/v) FMC® Long Ranger Gel Solution, 1% (v/v) TBE buffer, 0.1% (w/v) APS, 0.004% (v/v) TEMED) in 1 X TBE buffer (45 mM Tris-borate, 1 mM EDTA, pH 8.0) at 26.6 V/cm for one and a half hours. After separation the amplification products, which were detected by staining with (1:10 000) Sybr®GreenI Nucleic Acid Gel stain (Boehringer Mannheim), were visualized by exposure to UV illumination. Boehringer Mannheim 100 bp DNA ladder was included as the molecular size marker in all electrophoresis profiles.

The amplification products were then cut out of the PAGE gels and extracted from the plugs by using the QIAquick™ PCR purification kit (QIAGEN) according to manufacturers recommendations. The purified amplification products were then cloned into the T7 / SP6 cloning site of the pGEM®-T Easy plasmid vector, before being transformed into *Escherichia coli* JM 109 competent cells using the Promega pGEM®-T EASY vector system. The transformed *E. coli* cells were then plated out, at various concentrations, on LB plates containing Ampicillin (50 mg/ml), IPTG (100 mM) and XGAL (2% w/v) and incubated at 37°C for 16 to 24 hours. Transformed colonies, identified using the blue / white selection criteria of the IPTG / XGAL system, were then selected and their plasmids isolated using the Nucleospin® Plasmid kit (Macherey-Nagel). The isolated plasmid DNA was then purified further by PEG precipitation, before sequencing of the insert was done on a Perkin Elmer GeneAmp® 96000 thermal cycler, using the ABI Prism™ Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) according to manufacturers recommendations.

2.2.3 Results

The careful screening of transformed colonies resulted in the selection of ten colonies that contained inserts potentially suitable for sequencing. The ten sequences obtained are given in Appendix A. The ten sequences were analyzed and subsequently screened for regions containing microsatellites. Primers, of between 17 and 21 nucleotides, were then designed in the regions flanking these microsatellite loci. A summary of the primers designed is presented in Table 2.4, while the position of the primers relative to the sequences from which they were designed can be seen in Appendix A.

Two of the ten microsatellite primer pairs that were designed were found to produce amplification products (primer JD7-AMO and JD9-AMO), while the remaining eight would not produce any amplification products (primers JD1-AMO, JD2-AMO, JD3-AMO, JD4-AMO, JD5-AMO, JD6-AMO, JD8-AMO and JD10-AMO).

The microsatellite markers, JD7-AMO and JD9-AMO, produced monomorphic amplification products. Figure 2.4 shows the microsatellite fingerprints, of the E 503 x Pch 23 hybrid population, which were obtained using microsatellite marker JD9-AMO, while figure 2.5 shows the microsatellite fingerprints of the same population which were obtained using microsatellite marker JD7-AMO.

Microsatellite marker JD9-AMO was designed to produce an amplification product of approximately 300 base pairs. This corresponds with the amplification product shown in

figure 2.4 as well as with the sequencing results obtained, which overlap substantially with the original RAMs sequence 3 (Appendix A, Figure A-3).

Table 2.4. Summary of RAM primers developed.

Primer	Forward sequence (5' – 3')	Reverse Sequence (5' – 3')
JD1-AMO	F1: ATAATCAAGCATACAAG	R1: GGTTGTAGTGGTGTCTC
JD2-AMO	F2: GCCCACATCTACCCAAG	R2: CTGATGTGGAAGGTCTG
JD3-AMO	F3: GCAGGACAACGCCAA	R3: GCAAAACAAGATAAT
JD4-AMO	F4: GCATCTCATTTTCTATT	R4: GCGTGTTTAGATACTTTA
JD5-AMO	F5: CCCAACACTCCTGGGCT	R5: GAGCAGGACAACGCCAA
JD6-AMO	F6: TGGACAAACGCCTAAC	R6: AGGTCAAGGGAGACAA
JD7-AMO	F7.1: TGGATCCCCCAAGTCT F7.2: TAGTCCAGCCTTCAGTG	R7: CAAAGAGGAAACGAACAG
JD8-AMO	F8: CACACACACATGGACGCA TA	R8.1: CGAACTGAATGGGTTGTGTG R8.2: GTTTACCGCAGTTGGTTG
JD9-AMO	F9.1:GCACACACACACACACAAAAAAAACT F9.2: CACCAGGGGAAACCGAG	R9: TTGCTCAAGCCAGCCAG
JD10-AMO	F10: CCCGGACGATTTCCTTG	R10: GTGATTGACCCGCCTCC

Microsatellite marker JD7-AMO was designed to produce an amplification product of approximately 250 base pairs. However, as is shown in figure 2.4, an amplification product of approximately 350 base pairs was obtained. When the annealing temperature

was raised, thereby increasing the stringency of the PCR reaction, the same amplification product was obtained.

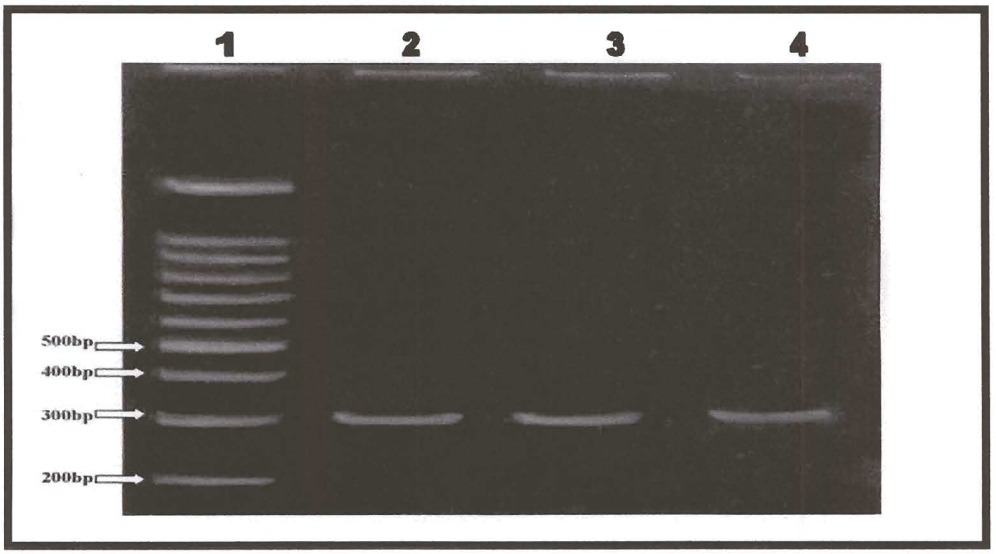

Figure 2.4. Microsatellite fingerprints of the *P. elliottii* (E 503) *x P. caribaea* (Pch 23) hybrid population. The fingerprints were obtained by PCR amplification with primer JD9-AMO. Lane 1 contains 100 bp DNA ladder, while lanes 2 to 4 contain the fingerprints obtained using E 503 needle DNA, E 503 x Pch 23 embryo DNA and PCH 23 needle DNA, respectively.

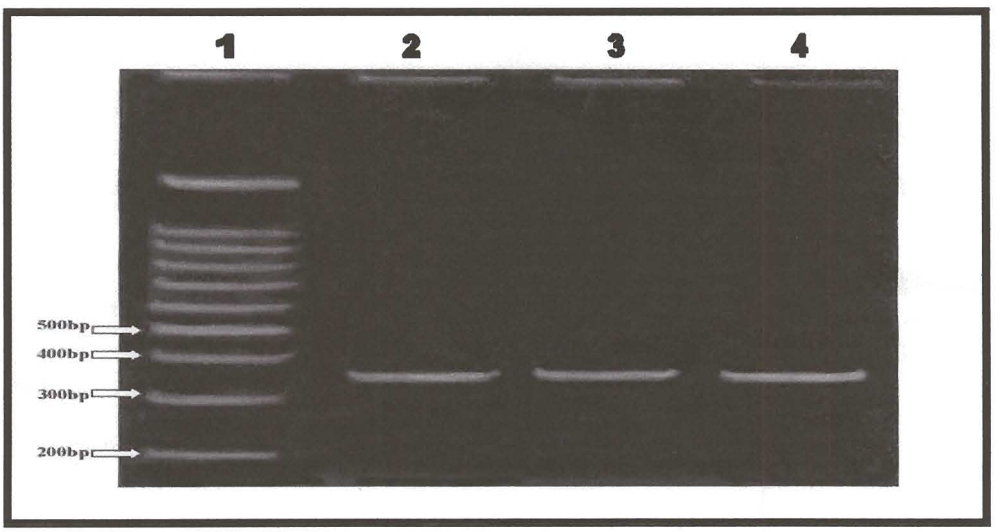

Figure 2.5. Microsatellite fingerprints of the *P. elliottii* (E 503) *x P. caribaea* (Pch 23) hybrid population. The fingerprints were obtained by PCR amplification with primer JD7-AMO. Lane 1 contains 100 bp DNA ladder, while lanes 2 to 4 contain the fingerprints obtained using E 503 needle DNA, E 503 x Pch 23 embryo DNA and PCH 23 needle DNA, respectively.

These results seem to indicate that the JD7-AMO reverse primer bound further down-stream than it was designed to, thereby amplifying an additional, untargeted section of the *Pinus* genome. The sequencing results verify these results as substantial sequence homology was found to exist with that of the original RAMs sequence 6 (Appendix A, Figure A-6).

2.2.4 Discussion

In 1994 Zietkiewicz *et al.* found that RAMs could be used to generate species specific DNA fingerprints. These species-specific DNA fingerprints can provide useful information pertaining to linkage, gene flow and parental contribution within and between populations (Hantula *et al.,* 1996).

Due to the costs involved in sequencing and primer design only ten inserts were sequenced. The sequences were analyzed and ten primer pairs were designed. Upon testing it was found that only two of the ten primer pairs (20%) resulted in the production of amplification products. In both cases the amplification products produced were monomorphic. As discussed in section 2.1.4 monomorphic microsatellite markers are not as useful as polymorphic markers, as it is impossible to distinguish between species or to determine gene flow or parental contribution. Although the amplification products produced were monomorphic the results indicate that this method has potential for microsatellite marker development in pines.

The fact that marker JD-7AMO produced an amplification product which was approximately 100 bp larger than it should have been indicates the need for marker screening to ensure that the amplification product obtained actually corresponds with the expected microsatellite region being targeted. In this scenario the marker does not need to be discarded as it amplified a conserved region in *P. elliottii*, *P. caribaea* and in the *P. elliottii x P. caribaea* hybrid.

In section 2.1.4 the important role which polymorphic microsatellite markers can play in population studies was discussed. Similarly, polymorphic microsatellite markers developed using RAMs would be just as useful for linkage, gene flow and genotype determination.

2.2.5 Conclusion

The RAMs approach to microsatellite marker development, although not as cost effective as simply using markers developed in related species, is a relatively fast way of developing species specific microsatellites. Although the source of variability in the RAMs fragments is still unknown, Hantula *et al.* (1996) proposed that the most likely polymorphisms would be length polymorphisms as variability in these elements is mostly due to differences in the number of repeats. The species specific DNA fingerprints obtained using RAMs are highly reproducible and can be used to identify both inter- and intra-species specific polymorphisms.

The results of this study indicated that the RAMs technique is potentially an effective way of developing microsatellite markers specific to the *P. elliottii x P. caribaea* hybrids. The results are promising enough to justify further investigation into this approach for marker development and perhaps even investigation into the potential of these markers for cross-species amplification.