academicJournals

Vol. 12(29), pp. 4578-4587, 17 July, 2013
DOI: 10.5897/AJB2013.12120
ISSN 1684-5315 ©2013 Academic Journals
http://www.academicjournals.org/AJB

African Journal of Biotechnology

*Full Length Research Paper*

# Analysis of genetic diversity in female, male and half sibs willow genotypes through RAPD and SSR markers

**N. B. Singh[1], M. K. Singh[1], P. K. Naik[2], S. Thakur[1] and J. P. Sharma[1]**

[1]Department of Tree Improvement and Genetic Resources, College of Forestry, Dr. Y. S. Parmar University of Horticulture and Forestry, Nauni, Solan, (H.P.)-173230, India.
[2]Department of Bioinformetics and Biotechnology, J. P. University, Waknaghat, Solan, (H.P.) - 173 234, India.

Accepted 12 July, 2013

**Willows belong to the genus *Salix* (*Salicaceae*) and consist of large number of species with large phenotypic variations. As a result, it has a low diagnostic value for identifying pure species and interspecific hybrids. Genetic characterization of 34 reference genotypes (4 female, 10 male, and 20 half sibs) of *Salix* collected from Naganji Nursery of University of Horticulture and Forestry, Solan, Himachal Pradesh, India were analyzed using 10 SSRs and 15 RAPDs PCR-based molecular markers. RAPD analysis yielded 87 polymorphic fragments (98.9%), with an average of 5.8 polymorphic fragments per primer. Similarly, SSR analysis produced 33 bands, out of which 26 were polymorphic (78.8%) with an average of 2.6 polymorphic fragments per primer. The genetic diversity was high among the genotypes (Nei's genetic diversity = 0.468 and Shannon's information index = 0.659) as measured by combination of both RAPD and SSR markers. The mean coefficient of gene differentiation (Gst) was 0.034, indicating 96.6% of the genetic diversity resided within the genotypes. The genetic diversity among genotypes of *Salix* sp. was found to be high, suggesting the importance and feasibility of introducing elite genotypes from different origins for *Salix* germplasm conservation and breeding programs.**

**Key words:** *Salix* sp., half sibs, molecular markers, genomic DNA fingerprinting.

## INTRODUCTION

The genus *Salix* (family *Salicaceae*) is one of the most important taxonomic entities with over 300 taxa widespread both in the boreal and austral hemisphere. Hybridization between *Salix* species seems to be common (Brunsfeld et al., 1992) with rates higher than those known for many other genera. In addition to the naturally occurring hybrids in various parts of the world, a considerable number of hybrids have been artificially produced by controlled matings. As a result, identification of true species is very difficult at the morphological level (Newsholme, 2003). Natural hybridization is supported by dioecism and is affected by diverse flowering phenologies in different *Salix* species. Allozyme variation studies in the genus *Salix* revealed that differentiation between populations is generally low and the gene flow rather moderate (Purdy and Bayer, 1995). Recent molecular studies carried out in progenies from controlled crosses and field clones suggest that hybrids are rare in the *S. alba* - *S. fragilis* complex (Triest et al., 2000). In particular, molecular markers revealed that both species have kept their gene pools well separated and that interspecific hybridization actually does not seem to be a dominating process. Furthermore, molecular characterization of willow germplasm provides the tools to estimate the genetic variability at species and within-species level. The study of genetic variability at different levels, in terms of allele composition, allele frequency, genetic diversity and differentiation, is a crucial step for the correct taxo-

*Corresponding author. E-mail: mintusingh78@gmail.com. Tel: 01792-252324. Fax: 01792-252318.

**Table 1a.** List of the 34 *Salix* genotypes.

| SN | Male | SN | Half sib | SN | Female |
|---|---|---|---|---|---|
| 1 | *Salix alba* | 11 | Half sib 1 | 31 | *Salix tetrasperma* |
| 2 | *Salix babylonica* | 12 | Half sib 2 | 32 | *Salix matsudana x Salix alba* (J799) |
| 3 | *Salix tetrasperma* | 13 | Half sib 3 | 33 | *Salix matsudana* (PN227) |
| 4 | *Salix matsudana* (PN722) | 14 | Half sib 4 | 34 | *Salix matsudana* (SE-69-002) |
| 5 | *Salix rubence* | 15 | Half sib 5 | | |
| 6 | *Salix udensis* (SX59) | 16 | Half sib 6 | | |
| 7 | *Salix alba* × *Salix babylonica* (131/25) | 17 | Half sib 7 | | |
| 8 | *Salix matsudana* × *Salix alba* (NZ1140) | 18 | Half sib 8 | | |
| 9 | *Salix matsudana* × *Salix alba* (NZ1179) | 19 | Half sib 9 | | |
| 10 | *S.matsudana* × *S.alba* (NZ1002) | 20 | Half sib 10 | | |
| | | 21 | Half sib 11 | | |
| | | 22 | Half sib 12 | | |
| | | 23 | Half sib 13 | | |
| | | 24 | Half sib 14 | | |
| | | 25 | Half sib 15 | | |
| | | 26 | Half sib 16 | | |
| | | 27 | Half sib 17 | | |
| | | 28 | Half sib 18 | | |
| | | 29 | Half sib 19 | | |
| | | 30 | Half sib 20 | | |

nomic classification of critical materials and for identifying the most appropriate sources of individuals to use in wood industries and waste land reclamation.

This research dealt with the utilization of different systems of multilocus PCR-based molecular markers, such as SSRs and RAPDs for molecular characterization and genetic differentiation of female, male and half sib willow genotypes.

## MATERIALS AND METHODS

### Plant material

Thirty four (34) genotypes (four female, ten male and twenty half sib) of *Salix* sp. were collected from Naganji nursery farm of the Dr. Y. S. Parmar University of Horticulture and Forestry, Nauni, Solan (H.P.) India, (Table 1a). Although, these plants showed distinctive taxonomic traits of the different willow species, they were chosen for their great variability in terms of morphological traits such as young and mature leaves, bark colour, etc.

### DNA extraction

The young leaf samples were collected during the period of March to October in sampling bags under aseptic conditions. The leaves were stored at -20°C for DNA extraction. Total genomic DNA was extracted from the frozen leaves (2 g) by the CTAB method (Saghai-Maroof et al., 1984) with minor modifications, which included the use of 200 mg of polyvinyl pyrollidone per sample. The extracted DNA was then treated with 20 µl of 10 mg/ml of RNase and incubated at 37°C for 60 min. After incubation with RNase, equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added and mixed gently by inverting the microcentrifuge tube followed by centrifugation at 10,000 rpm for 5 min at room temperature. The supernatant was pipetted out into a fresh tube. The sample was then extracted twice with equal volume of chloroform:isoamyl alcohol (24:1). The DNA was precipitated by adding 0.6 vol of isopropanol and 2.0 M NaCl. To the aforementioned, 20 µl of sodium acetate and 1 vol of 80% ethanol were added, incubated for 30 min and centrifuged at 5,000 rpm for 3 min to pelleted the DNA. The pellet was then washed with 70% ethanol twice, air-dried and finally suspended in 40 to 50 µl of TE buffer. The yield of the extracted DNA and purity was checked by running the sample on 0.8% agarose gel along with standard (non restriction enzyme digested) lamda DNA marker (Biogene, USA).

The extracted genomic DNA was tested for purity index ($A_{260}/A_{280}$ absorbance ratio) on Nano drop spectrophotometer. A value of 1.8 of extracted DNA samples indicate high purity, whereas, the value <1.8 or >1.8 denotes the contamination of proteins and RNA, respectively (Sambrook et al., 1989).

### RAPD markers

PCR amplification was carried out in a 25 µl total reaction volume containing 30 ng genomic DNA, 1.5 mM $MgCl_2$, 1 µM of primers and 1 unit of Taq DNA polymerase (Pharmacia) (Barcaccia et al., 1997). Amplification was performed in a 9700 Thermal Cycler (PerkinElmer) under the following temperature profile: initial denaturation for 5 min at 95°C was followed by 3 cycles of 2 min at 95°C, annealing temperature of 35°C for 1 min, 72°C for 2 min to extend, 37 cycles at 94°C for 15 s, 36°C for 30 s, 72°C for 1 min and 72°C for 10 min. The rates of temperature change adopted for heating and cooling were +1°C/2.9 s and -1°C/2.4 s, respectively. Amplification products were electrophoresed on 1.5% agarose gels run at constant voltage and 1X TBE buffer for approximately 2 h, visualized by staining with ethidium bromide (Sambrook et al., 1989) and photographed under UV light (using DC120 camera, Kodak).

**Table 1b.** List of primers used for RAPD amplification, GC content, total number of loci, the level of polymorphism, resolving power and PIC value.

| Primer | Primer sequence | % GC content | Annealing temperature | Total number of loci | NPL | PPL | Number of fragments amplified | Rp | PIC |
|---|---|---|---|---|---|---|---|---|---|
| OPA-15 | GGGCCACTCA | 70 | 42 | 6 | 6 | 100 | 123 | 7.235 | 0.785 |
| OPJ-2 | GGAGAGACTC | 60 | 41.7 | 6 | 5 | 83.3 | 144 | 8.471 | 0.795 |
| OPJ-4 | CACAGAGGGA | 60 | 41.7 | 7 | 7 | 100 | 94 | 5.529 | 0.791 |
| OPG-8 | GGGTTTGGCA | 60 | 38 | 5 | 5 | 100 | 116 | 6.824 | 0.783 |
| OPJ-10 | CAAGGGCAGA | 60 | 45.5 | 4 | 4 | 100 | 93 | 5.471 | 0.713 |
| OPJ-19 | GGCAGGCTGT | 70 | 42 | 6 | 6 | 100 | 127 | 7.471 | 0.793 |
| OPJ-8 | AACGGCGACA | 60 | 41 | 6 | 6 | 100 | 122 | 7.176 | 0.782 |
| OPJ-7 | TTCCCCGCGA | 70 | 41 | 6 | 6 | 100 | 118 | 6.941 | 0.779 |
| OPG-6 | AGGACTGCCA | 60 | 45.5 | 6 | 6 | 100 | 119 | 7.00 | 0.781 |
| OPG-11 | GGTGAACGCT | 60 | 43.8 | 6 | 6 | 100 | 123 | 7.235 | 0.785 |
| OPH-10 | CCAACGTCGT | 60 | 42 | 6 | 6 | 100 | 123 | 7.235 | 0.785 |
| OPH-15 | GGTCGGAGAA | 60 | 36 | 6 | 6 | 100 | 122 | 7.176 | 0.782 |
| OPH-18 | TCGGACGTGA | 60 | 36 | 6 | 6 | 100 | 123 | 7.235 | 0.784 |
| OPH-1 | AGACGTCCAC | 60 | 36 | 6 | 6 | 100 | 123 | 7.235 | 0.785 |
| OPH-2 | ACGCATCGCA | 60 | 36 | 6 | 6 | 100 | 125 | 7.353 | 0.783 |

NPL, Number of polymorphic loci; PPL, percentage of polymorphic loci, Rp, resolving power and PIC, polymorphism information content.

## SSR markers

A set of 10 pairs of SSR primers (Table 2) (synthesized by Life Technologies, Inc.) were used in this study. PCR reactions were performed with a protocol reported earlier (Barcaccia et al., 2003b) with minor changes. The volume of PCR solution was 25 µl, containing 10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 300 µM each of dCTP, dGTP, dATP and dTTP, 800 µM of primer, 1.5 U of Taq DNA polymerase (Pharmacia Biotech) and 30 ng of genomic DNA. Amplification reactions were performed in a 9700 Thermal Cycler (PerkinElmer) using a touchdown cycling profile. The optimized PCR amplifying conditions used were: initial denaturation at 95°C for 3 min, followed by 2 cycles of 1 min at 95°C, an annealing temperature of 1 min at 63°C and 2 min at 72°C followed by a reduction in annealing temperature by 1°C every two cycles until a final annealing temperature of 56°C was reached. The last cycle was repeated 26 times and was ended by a final step at 72°C for 10 min. The amplified fragments were separated on 2% agarose gels with 1X TBE buffer (Sambrook et al., 1989) at 150 V for 3 h. Photographs (DC120 camera, Kodak) of the polymerized genomic fragments were taken after staining of the agarose gels with ethidium bromide.

## Data collection and analysis

The genetic relationship among the entire genomic DNA under study was assessed by comparing the RAPD and SSR fragments separated according to their size. The banding pattern of each of the primer was scored as present (1) or absent (0), each of which was treated as an independent character. Only the reproducible bands were observed for scoring and the light bands were omitted as they were not reproducible. The Jaccard's dissimilarity coefficient (J) was calculated, subjected to cluster analysis by bootstrapping and neighbor-joining method using the program DARWIN (version 5.0.158). Statistically unbiased clustering of collected genotypes was performed using STRUCTURE (version 2.3.1). POPGENE software was used to calculate Nei's unbiased genetic distance among different genotypes with all markers.

Data for observed number of alleles (Na), effective number of alleles (Ne), Nei's gene diversity (H), Shannon's information index (I), number of polymorphic loci (NPL) and percentage of polymorphic loci (PPL) across all the 34 genotypes were analyzed by Nei and Li, (1979). Within group diversity (Hs) and total genetic diversity (Ht) were calculated within the species and within three major groups (based on the male, female and half sib genotypes) by using POPGENE software by Nei (1978). The RAPD and SSR data were subjected to a hierarchical analysis of molecular variance (AMOVA) (Excoffier et al., 1992), using three hierarchical levels; individual, population and grouping based on their male, female and half sib genotypes. The non-parametric analysis of molecular variance (AMOVA) was done via Gen Alex (Excoffier et al., 1992), where the variation component was partitioned among individuals within populations, among populations within groups and

**Table 2.** List of primers used for SSR amplification, GC content, total number of loci, the level of polymorphism, resolving power and PIC value.

| Primer | Primer sequence | % GC content | Annealing temperature | Total number of loci | NPL | PPL | Number of fragments amplified | Rp | PIC |
|---|---|---|---|---|---|---|---|---|---|
| SB-243 | FP-ACT TCA ATC TCT CTG TAT TCT | 47 | 53 | 2 | 1 | 50 | 36 | 2.118 | 0.105 |
| | RP-CTA TTT ATG GGT TGG TCG ATC | | | 3 | 2 | 66.7 | 83 | 4.882 | 0.656 |
| SB-38 | FP-CCA CTT GAG GAG TGT AAG GAT | 53 | 54.5 | 3 | 2 | 66.7 | 81 | 4.765 | 0.646 |
| | RP-CTT AAA TGT AAA ACT GAA TCT | | | 2 | 2 | 100 | 61 | 3.588 | 0.450 |
| SB-199 | FP-CTA TTT GGT CTC AAT CAC CTT | 53 | 58 | 3 | 2 | 66.7 | 84 | 4.941 | 0.656 |
| | RP-CTT TAC CTC AGA AAA TCC AGA | | | 5 | 4 | 80 | 96 | 5.647 | 0.736 |
| SB-85 | FP-CTC AGC AAC TTA ATC CAA CTA | 53 | 59 | 6 | 6 | 100 | 117 | 6.882 | 0.777 |
| | RP-GTT TGT TAG GGG AGG TAA GAA | | | 2 | 2 | 100 | 52 | 3.059 | 0.488 |
| SB-80 | FP-TAA TGG AGT TCA CAG TCC TCC | 44 | 54.3 | 4 | 3 | 75 | 56 | 3.294 | 0.563 |
| | RP-ATA CAG AGC CCA TTT CAT CAC | | | 3 | 2 | 66.7 | 57 | 3.353 | 0.526 |

among groups. The resolving power of the RAPD and SSR primers was calculated according to Prevost and Wilkinson (1999). The resolving power (*Rp*) of a primer is:

$Rp = \Sigma IB$

Where, *IB* (band informativeness) takes the value of: 1-[2* (0.5-P)], P being the proportion of the 34 genotypes containing the band.

In order to determine the utility of each of the marker systems, diversity index (DI), effective multiple ratio (EMR) and marker index (MI) were calculated according to Powell et al. (1996). DI for the genetic markers was calculated from the sum of squares of allele frequencies:

$DI_n = 1-\sum pi^2$

Where, pi is the allele frequency of the ith allele.

The arithmetic mean heterozygosity, $DI_{av}$, was calculated for each marker class:

$DI_{av} = \sum Di_{n/n}$

Where, n represents the number of the markers (loci) analyzed.

The DI for the polymorphic marker is:

$DI_{av})p = \sum Di_n/n_p$

Where, $n_p$ is the number of polymorphic loci and n is the total number of loci.

EMR (E) is the product of the fraction of polymorphic loci and the number of polymorphic loci for an individual assay EMR (E) = $n_p(n_p/n)$.

MI is defined as the product of the average diversity index for polymorphic bands in any assay and the EMR for that assay.

$MI = DI_{avp} \times E$

***Polymorphism information content (PIC)***

The frequency of the polymorphism obtained in the genotypes was calculated on the basis of presence (1) and absence (0) of the bands amplified. The PIC was calculated according to Anderson et al. (1993) based on the allele pattern of all the willow genotypes by employing the following formula:

$$PIC_i=1-\sum_{j=1}^{n} Pij^2$$

## RESULTS AND DISCUSSION

### Molecular analysis using RAPD markers

The RAPD technique had been successfully used in a variety of taxonomic and genetic diversity studies and it was found suitable for use with *Salix* sp. genotypes because of its ability to generate reproducible polymerphic markers. A total of 34 plant samples were fingerprinted using 15 RAPD makers. These primers produced multiple band profiles with a number of amplified DNA fragments varying from 4 to 7. All the amplified fragments varied in size from 100 to 2000 bp. Out of 88 amplified bands, 87 were found polymorphic (98.8%) (Table 1b).

The observed high proportion of polymorphic loci suggests that there is a high degree of genetic variation in the *Salix* sp. The resolving power of the 15 RAPD primers ranged from 5.471 for primer OPJ-10 to a maximum of 8.471 for primer OPJ-2. Polymorphism information content (PIC) refers to the value of a marker for detecting polymorphism within a population or set of genotypes by taking into account not only the number of alleles that are expressed but also the relative frequencies of alleles per locus. As evident, RAPD marker 'OPJ-2' showed the highest level of polymorphism with PIC value of 0.795, whereas the PIC values for the rest of the RAPD markers were in the range of 0.713 to 0.791. A dendrogram analysis based on bootstrapping and neighbor joining (NJ) method grouped all the 34 genotypes into three main clusters which are further extensively divided into mini clusters (Figure 1a). An unbiased clustering of genotypes based on STRUCTURE program without prior knowledge about the populations clustered all the 34 genotypes into three major groups. Under the admixed model, STRUCTURE calculated that the estimate of likelihood of the data [LnP(D)] was greatest when K = 3. For K > 3, LnP(D) increased slightly but more or less plateued (Figure 1b), that is, ΔK reached its maximum value when K = 3 (Figure 1c), suggesting that all the populations fell into one of the three clusters albeit small interference (Figure 1d).

This result is almost similar to the splitting in the NJ tree. Overall, the cluster analysis strongly suggested that the 34 sampled genotypes can be divided into 3 clusters, however, there is no distinct clustering of genotypes based on their four female, 10 male and 20 half sibs. The genetic diversity of 34 genotypes was calculated in terms of Na, Ne, H, I, Ht and PPL with respect to 3 different groups such as 4 female, 10 male, and 20 half sibs revealed higher values, indicating more variability among the genotypes (Table 3). Polymorphic loci of 100% were calculated using POPGENE among 4 females, 10 males and 20 half sibs genotypes. Three groups containing genotypes with different sexes such as female, half sibs and males showed Nei's genetic diversity (H): 0.517, 0.460 and 0.413, respectively and of Shannon's information index (I): 0.742, 0.652 and 0.598 (Table 3), respectively showed a higher genetic differentiation within each of the three groups. The respective values for overall genetic variability for Na, Ne, H, I, Ht, Hs, Gst, NPL, PPL and Gene flow (Nm) across all the 34 genotypes were also given in Table 4. The rate of gene flow estimated using Gst value was found to be 0.40 which is very high. Analysis of molecular variance among genotypes based on three major groups with respect to 4 female, 10 male and 20 half sibs plant indicated that majority of genetic variation (100%) occurred among genotypes, while the variation between the three clusters was 0% (Table 5).

### SSR analysis

The 10 SSR primers selected in the study generated a total of 33 SSR bands (an average of 3.3 bands per primer), out of which 26 were polymorphic (78.8%) (Table 2). Among the dinucleotide repeat types, (AG)n and (GA)n produced more number of bands followed by (CT)n and (AC)n. Similarly, among the tri-nucleotide repeat types, (CTC)n produced more number of bands. The primers that were based on the (GA)n, (AG)n and (CT)n motif produced more polymorphism than the primers based on any other motifs used in the present investigation. We obtained good amplification products from primers based on (AG)n and (GA)n repeats, despite the fact that (AT)n di-nucleotide repeats are thought to be the most abundant motifs in plant species (Martin et al., 2000). Similar results were obtained in grapevine (Moreno et al., 1998), rice (Blair et al., 1999), *Vigna* (Ajibade et al., 2000) and wheat (Nagaoka and Ogihara, 1979). A possible explanation of these results is that SSR primers based on AT motifs are self-annealing, due to sequence complementarity, and would form dimers during PCR amplification (Blair, 1999) or it may be due to its-non annealing with template DNA due to its low Tm. The resolving power (Rp) of the 10 SSR primers ranged from 2.118 to 6.882 (Table 2). Similarly, the PIC value ranges from 0.105 to 0.777 demonstrating uniform polymorphism rate among all the 10 SSR primers. The complete data set of 723 bands was used for cluster analysis based on bootstrapping and NJ method. The genotypes were clustered into three major clusters, well supported by bootstrap value of > 20 (Figure 2a). The estimated likelihood [LnP(D)] of the clustering of data using STRUCTURE was found to be optimal when K = 3, LnP(D) increased slightly for K > 3, but more or less plateued (Figure 2b).

ΔK reached its maximum value when K = 3 (Figure 2c), suggesting that all the populations were distributed with high probability into one of the 3 clusters (Figure 2d). The clustering pattern of the genotypes were almost similar to the splitting in the NJ tree, however, there is no distinct clustering of genotypes based on their 4 female, 10 male and 20 half sibs plants. A relatively high genetic variation

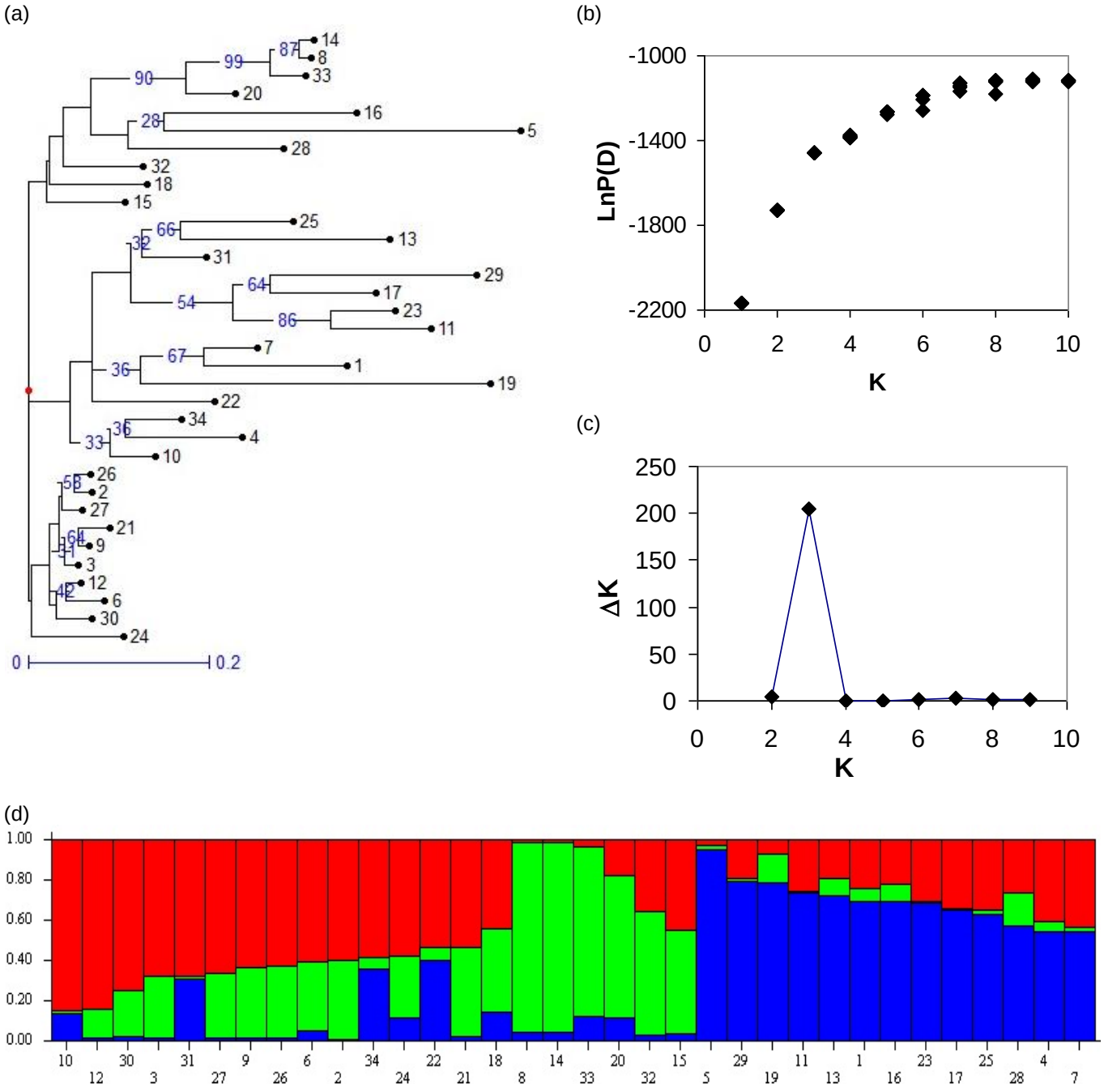


**Figure 1.** (a) Dendrogram generated by Neighbor joining (NJ) clustering technique showing relationships between 34 genotypes of *Salix sp.* based on RAPD profiling. Genotypes from 1 to 10 are male; 11 to 30 are half sibs and 31 to 34 are female. Number indicates bootstrap support values. (b) The relationship between the number of cluster (K) and the estimated likelihood of data [LnP(D)]. A model based clustering of 34 genotypes using STRUCTURE without prior knowledge about the populations and under an admixed model calculated that LnP (D) was greatest when K = 3. (c) The relationship between K and ΔK., that is, ΔK is reaches its maximum when K = 3, suggesting that all genotypes fall into one of the 3 clusters. (d) Grouping of genotypes when K = 3. The genotypes were more likely clustered with respect to one of the 3 clusters. Genotypes from different clusters are represented with different colours: cluster 1 (red), cluster 2 (green) and cluster 3 (blue).

was detected among the genotypes categorized into three different groups. Genetic diversity analysis in terms of Na, Ne, H, I, Ht, Hs and PPL revealed higher value for the group with four female, 10 male and 20 half sib plants. This disparity may be because of more number of genotypes included in the group with 4 female, 10 male

**Table 3.** Summary of genetic variation statistics for the combination of (a) RAPD only (b) SSR only and (c) combination of both RAPD and SSR profiling among the genotypes of *Salix sp.* with respect to their distributions into 3 groups.

| Markers | Sample size | Na | Ne | H | I | Ht | NPL | PPL |
|---|---|---|---|---|---|---|---|---|
| **(a) RAPD markers** | | | | | | | | |
| Male | 10 | 2.00 (0.00) | 1.745 (0.254) | 0.413 (0.100) | 0.598 (0.114) | 0.413 (0.010) | 68 | 100 |
| Half sibs | 20 | 2.00 (0.00) | 1.862 (0.132) | 0.460 (0.041) | 0.652 (0.043) | 0.460 (0.002) | 68 | 100 |
| Female | 4 | 1.738 (0.132) | 1.864 (0.174) | 0.517 (0.063) | 0.742 (0.004) | 0.528 (0.142) | 68 | 100 |
| **(b) SSR markers** | | | | | | | | |
| Male | 10 | 2.00 (0.00) | 1.901 (0.106) | 0.472 (0.031) | 0.665 (0.032) | 0.483 (0.002) | 50 | 100 |
| Half sibs | 20 | 2.00 (0.00) | 1.936 (0.049) | 0.483 (0.013) | 0.676 (0.013) | 0.483 (0.002) | 50 | 100 |
| Female | 4 | 1.980 (0.141) | 1.804 (0.230) | 0.435 (0.088) | 0.622 (0.111) | 0.435 (0.008) | 49 | 98 |
| **(c) Combination of RAPD and SSR markers** | | | | | | | | |
| Male | 10 | 2.00 (0.00) | 1.811 (0.218) | 0.438 (0.084) | 0.626 (0.094) | 0.438 (0.007) | 118 | 100 |
| Half sibs | 20 | 2.00 (0.00) | 1.893 (0.111) | 0.470 (0.034) | 0.662 (0.036) | 0.470 (0.001) | 118 | 100 |
| Female | 4 | 1.924 (0.267) | 1.744 (0.289) | 0.406 (0.132) | 0.581 (0.179) | 0.406 (0.017) | 109 | 92.4 |

Na, Observed number of alleles; Ne, effective number of alleles; H, Nei's gene diversity; I, Shannon's Information index; Ht, heterozygosity; NPL, number of polymorphic loci; PPL, percentage of polymorphic loci.

**Table 4.** Overall genetic variability across all the 34 genotypes of *Salix sp.* based on RAPD only, SSR only and combination of both RAPD and SSR markers.

| Parameter | Sample size | Na | Ne | H | I | Ht | Hs | NPL | PPL | Gst | Nm |
|---|---|---|---|---|---|---|---|---|---|---|---|
| RAPD | 34 | 2.00 (0.00) | 1.836 (0.162) | 0.451 (0.055) | 0.642 (0.060) | 0.451 (0.003) | 0.433 (0.004) | 68 | 100 | 0.040 | 11.980 |
| SSR | 34 | 2.00 (0.00) | 1.964 (0.039) | 0.491 (0.011) | 0.684 (0.011) | 0.491 (0.00) | 0.477 (0.00) | 50 | 100 | 0.028 | 17.653 |
| RAPD+SSR | 34 | 2.00 (0.00) | 1.890 (0.141) | 0.468 (0.047) | 0.659 (0.050) | 0.468 (0.002) | 0.451 (0.003) | 118 | 100 | 0.034 | 13.994 |

**Table 5.** Summary of analysis of molecular variance (AMOVA) based on (a) RAPD only (b) SSR only and (c) combination of both RAPD and SSR markers among the genotypes of *Salix sp.* Levels of significance are based on 1000 iteration steps.

| Source of variation | Degree of freedom | Variance component | Percentage of variation | P-value |
|---|---|---|---|---|
| **(a) Based on RAPD profiling** | | | | |
| Among groups | 2.0 | 7.78 | 0.0 | - |
| Among genotypes | 31.0 | 16.30 | 100 | < 0.992 |
| **(b) Based on SSR profiling** | | | | |
| Among groups | 2.0 | 6.984 | 0.0 | - |
| Among genotypes | 31.0 | 13.005 | 100 | < 0.882 |
| **(c) Based on combination of both RAPD and SSR profiling** | | | | |
| Among groups | 2.0 | 14.760 | 0.0 | - |
| Among genotypes | 31.0 | 29.305 | 100 | < 0.995 |

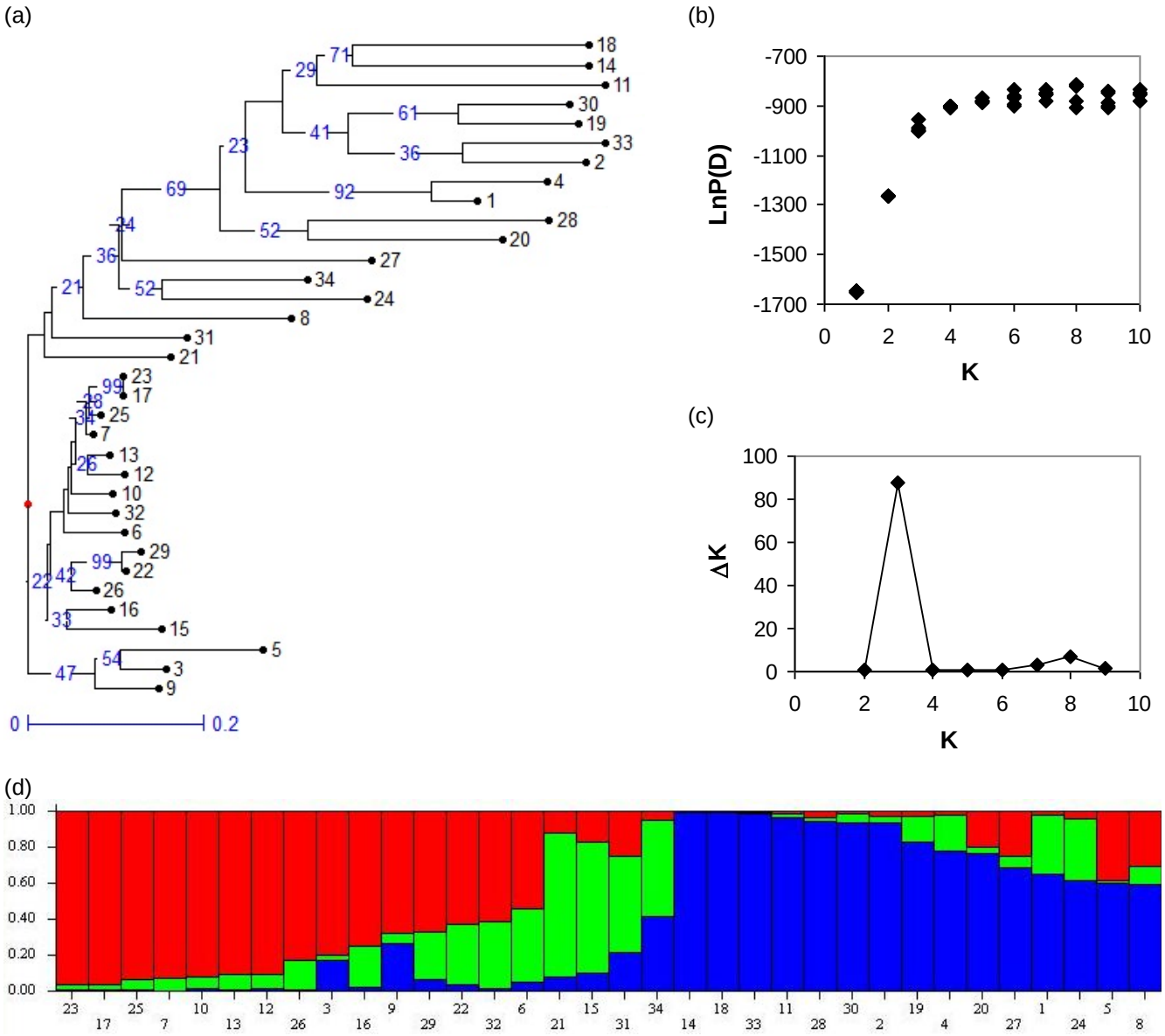


**Figure 2.** (a) Dendrogram generated by Neighbor joining (NJ) clustering technique showing relationships between 34 genotypes of *Salix sp.* based on SSR profiling. Genotypes from 1 to 10 are male; 11 to 30 are half sibs and 31 to 34 are female. Number indicates bootstrap support values. (b) The relationship between the number of cluster (K) and the estimated likelihood of data [LnP(D)]. A model based clustering of 34 genotypes using STRUCTURE without prior knowledge about the populations and under an admixed model calculated that LnP(D) was greatest when K = 3. (c) The relationship between K and ΔK., that is, ΔK reached its maximum when K = 3, suggesting that all genotypes fell into one of the 3 clusters. (d) Grouping of genotypes when K = 3. The genotypes were more likely clustered with respect to one of the 3 clusters. Genotypes from different clusters are represented with different colours: cluster 1 (red), cluster 2 (green) and cluster 3 (blue).

and 20 half sibs (Table 3). Overall genetic variability across all the 34 genotypes in terms of Na, Ne, H, I, Ht, Hs, Gst, NPL, PPL and Gene flow (Nm) were also included in Table 4. The Nei's genetic diversity index was 0.491 and Shannon information index was 0.684 demonstrating high rate of genetic variability. AMOVA for among groups (0%) and among genotypes (100%) indicated that there are more variations across the genotypes and not among the groups (Table 5). The estimated gene flow was 17.653.

## RAPD and SSR combined data for cluster analysis

Based on combined data set of RAPD and SSR markers, the dendrogram obtained gave similar clustering pattern like RAPD and SSR (Figure 3a). This result corroborate that of the STRUCTURE analysis; the estimated likeli-

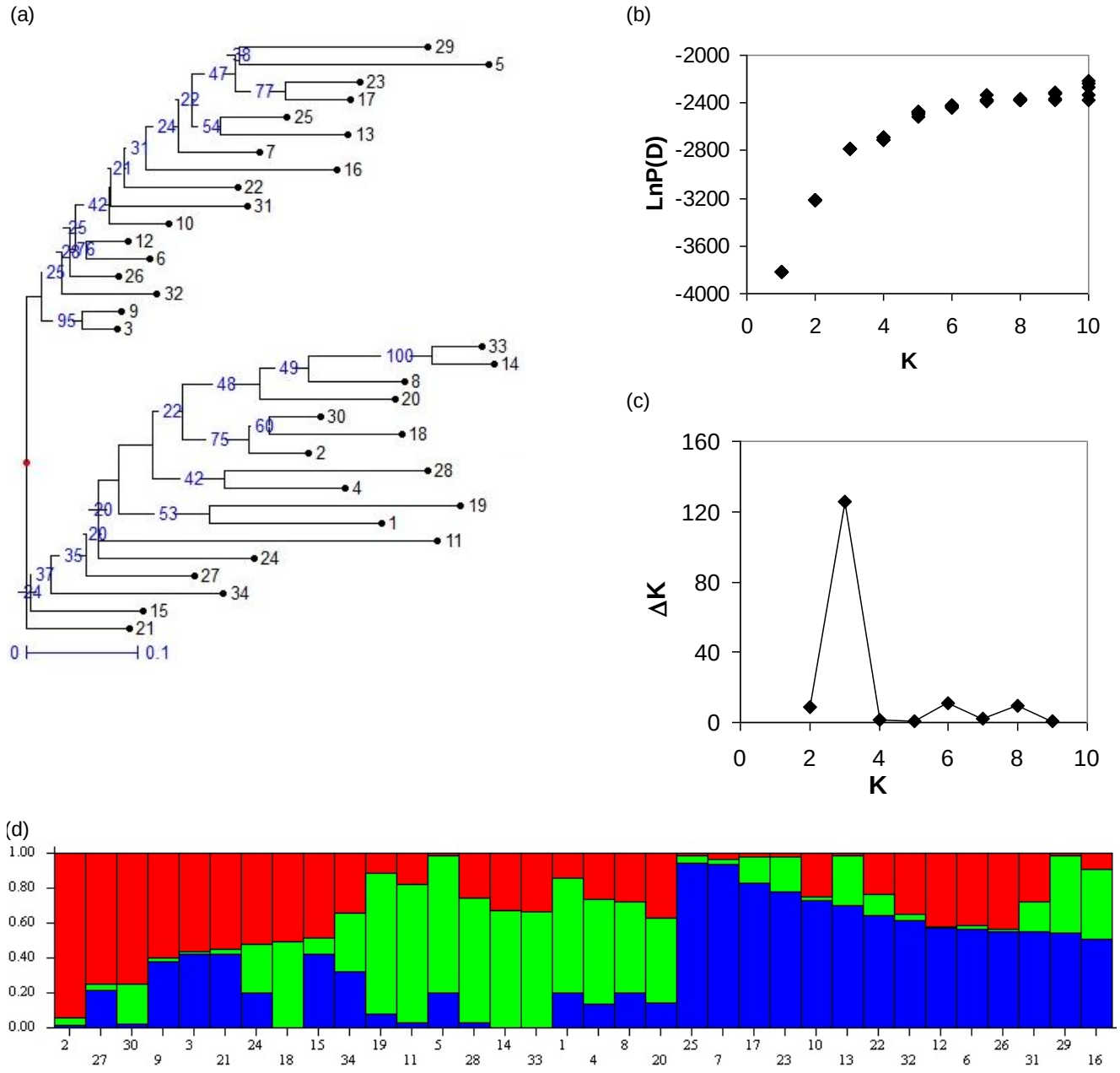


**Figure 3.** (a) Dendrogram generated by Neighbor joining (NJ) clustering technique showing relationships between 34 genotypes of *Salix sp.* based on combination of RAPD and SSR profiling. Genotypes from 1 to 10 are male; 11 to 30 are half sibs and 31 to 34 are female. Number indicates bootstrap support values. (b) The relationship between the number of cluster (K) and the estimated likelihood of data [LnP(D)]. A model based clustering of 34 genotypes using STRUCTURE without prior knowledge about the populations and under an admixed model calculated that LnP(D) was greatest when K = 3. (c) The relationship between K and ΔK., that is, ΔK reached its maximum when K = 3, suggesting that all genotypes fell into one of the 3 clusters. (d) Grouping of genotypes when K = 3. The genotypes were more likely clustered with respect to one of the 3 clusters. Genotypes from different clusters are represented with different colours: cluster 1 (red), cluster 2 (green) and cluster 3 (blue).

hood of distribution [LnP(D)] for all the 34 genotypes was highest when K = 3 (Figure 3b), and ΔK was maximum with K = 3 (Figure 3c); this reveals that all the genotypes were clustered better (with high likelihood probability) with three clusters (Figure 3d). Other genetic variation studies were also performed on RAPD and SSR combined data which are represented in different tables (Tables 3, 4 and 5). The differences found among the dendrograms generated by RAPDs and SSRs could be partially explained by the different number of PCR products analyzed reinforcing again the importance of the number of loci and their coverage of the overall genome, in obtaining reliable estimates of genetic relationships as observed by Loarce et al. (1996) in barley. Another expla-

nation could be the low reproducibility of RAPDs (Karp et al., 1997). The genetic similarity of these genotypes is probably associated with their similarity in the genomic and amplified region.

### Comparative analysis of RAPD with SSR markers

RAPD markers were found more efficient with respect to number of polymorphism detection (based on average NPL value), as they detected 68 polymorphism loci as compared to 50 polymorphic loci for SSR markers. This is in contrast to the results obtained for several other plant species like wheat (Nagaoka and Ogihara, 1997) and *Vigna* (Ajibade et al., 2000). More polymorphism in case of RAPD than SSR markers might be due to the fact that 10 SSR primers used in the study only amplified 723 number of fragments (Table 2) while in case of RAPD, all the 15 primers which were used in the investigation amplified 1795 number of fragments (Table 1). Similar polymorphism pattern was also observed in the case of *Jatropha* (Gupta et al., 2008) and *Podophyllum* (Alam et al., 2009). This shows that RAPD data is more close to RAPD + SSR combined data. A possible explanation for the difference in resolution of RAPDs and SSRs is that the two-marker techniques target different portions of the genome. The mean effective multiplex ratio was more for RAPD (5.744) than that for SSR (2.128) and similarly marker index was more for RAPD (0.772) than that for SSR (0.0.448) markers.

### Conclusion

In this study, we may conclude that molecular analyses of both RAPD and SSR markers were extremely useful for studying the genetic relationships of *Salix* genotypes. The results indicates the presence of high genetic variability, which should be exploited for the future conservation and breeding of willow sp. Since no single, or even a few plants, will represent the whole genetic variability in willow, it is essential to maintain sufficiently large populations in natural habitats to conserve genetic diversity in willow to avoid genetic erosion.

### ACKNOWLEDGMENT

The authors are thankful to Professor R. S. Cauhan, of the Department of Biotechnology and Bioinformatics, J. U. I. T, Waknaghat, Solan, (H.P.) for providing technical help.

#### REFERENCES

Ajibade SR, Weeden NF, Chite SM (2000). Inter simple sequence repeat analysis of genetic relationships in the genus *Vigna.* Euphytica 111:47-55.

Alam A, Gulati P, Gulati A, Mishra GP, Naik PK (2009). Assessment of genetic diversity among *Podophyllum hexandrum* genotypes of the North-western Himalayan region for podophyllotoxin production. Indian J. Biotechnol. 8:391-399.

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993). Optimizing parental selection for genetic linkage maps. Genome 36:181-186.

Barcaccia G, Mazzucato A, Belardinelli M, Pezzotti M, Lucretti S, Falcinelli M (1997). Inheritance of parental genomes in progenies of *Poa pratensis* L. from sexual and apomictic genotypes as assessed by RAPD markers and flow cytometry. Theor. Appl. Genet. 95:516-524.

Barcaccia G, Parrini P, Lucchin M (2003b). Characterization of a flint maize (*Zea mays var. indurata*) Italian landrace. II. Genetic diversity and relatedness assessed by SSR and Inter- SSR molecular markers. Genet. Resour. Crop Evol. 50:253-271.

Blair MW, Panaud O, Mccouch SR (1999). Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa L.*). Theor. Appl. Genet. 98:780-792.

Brunsfeld SJ, Soltis DE, Soltis PS (1992). Evolutionary patterns and processes in *Salix* sect. *Longifoliae*: evidence from chloroplast DNA. Syst. Bot. 17:239-256.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Gupta S, Srivastava M, Mishra Gyan P, Naik PK, Chauhan RS, Tiwari SK, Kumar M, Singh R (2008). Analogy of ISSR and RAPD markers for comparative analysis of genetic diversity among different *Jatropha curcas* genotypes. Afr. J. Biotechnol. 23:4230-4243

Karp A, Edwards KJ, Bruford M, Funk S, Vosman B, Morgante M, Seberg O, Kremer A, Boursot P, Arctander P, Tautz D, Hewitt GM (1997) Newer molecular technologies for biodiversity evaluation: Opportunities and challenges. Nat. Biotechnol. 15:625-628.

Loarce Y, Gallego R, Ferrer EA, (1996). Comparative analysis of genetic relationships between rye cultivars using RFLP and RAPD markers. Euphytica 88:107-115.

Martin JP, Sanchez-yelamo MD (2000). Genetic relationships among species of the genus *Diplotaxis* (*Brassicaceae*) using inter-simple sequence repeat markers. Theor. Appl. Genet. 101: 1234-1241.

Moreno S, Martin JP, Ortiz JM (1998). Inter-simple sequence repeat PCR for characterization of closely related grapevine germplasm. Euphytica 101:117-125.

Nagaoka T, Ogihara Y (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theor. Appl. Genet. 94:597-602.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76: 5269-5273.

Nei M, (1978) Estimation of average heterozygosity and genetic distance from a small number of individual. Genetics, 89: 583-590.

Newsholme C (2003). Willows. The genus *Salix*. Portland, Oregon: Timber Press.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996). The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasms analysis. Mol. Breed. 2:225-238.

Prevost A, Wilkinson MJ (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor. Appl. Genet. 98:107-112.

Purdy BG, Bayer RJ (1995). Allozyme variation in the *Athabasca* sand dune endemic, *Salix silicola*, and the closely related widespread species *S. alaxensis.* Syst. Bot. 20:179-190.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984). Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Nail. Acad. Sci. USA. 81:8014-8019.

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: *A laboratory, manual*. Cold spring harbor press, cold spring harbor, New York.

Triest L, De Greef B, De Bondt R, Van Slycken J (2000). RAPD of controlled crosses and clones from the field suggests that hybrids are rare in the *Salix alba -Salix fragilis* complex. Heredity 84:555-563.