3 Biotech (2018) 8:161
https://doi.org/10.1007/s13205-018-1175-5

ORIGINAL ARTICLE

# Molecular identification of *Saraca asoca* from its substituents and adulterants

**Satisha Hegde[1,2] · Archana Saini[1] · Harsha Vasudev Hegde[1] · Sanjiva D. Kholkute[1] · Subarna Roy[1]**

Received: 24 October 2017 / Accepted: 19 February 2018 / Published online: 5 March 2018


## Abstract

*Saraca asoca* (Roxb.) De Wilde is an important medicinal plant from the Western Ghats of India, traditionally used in treatment of various gynecological disorders. Increasing commercial demand and decreasing numbers has resulted in this plant becoming endangered with crude drug materials being extensively substituted/adulterated with other plant species. The present study was undertaken with the objective of development and evaluation of multivariate cluster analysis of ISSR fingerprints against *rbcL*-based DNA barcodes as tool to understand the relationships and to differentiate common adulterants and substituents from *S. asoca*. ISSR-based Hierarchical Cluster Analysis was carried out on 41 samples of *S. asoca* and 5 each of the 5 common substituent/adulterant plants and the clustering patterns were evaluated against DNA-sequence-based barcoding of *rbcL* region of their plastids. Factorial analysis and Principal Coordinate Analysis revealed distinct groups of genetic pools of respective taxa thereby confirming the utility of ISSR fingerprinting as a useful tool for differentiation between the genuine and the adulterants/substituents. NCBI-BLAST search on DNA barcode *rbcL* region confirmed the results of ISSR assays. Therefore, our study demonstrated the utility of simple, cost-effective method of ISSR fingerprinting coupled with *rbcL* barcoding in differentiating this important medicinal plant from its common adulterants/substituents.

**Keywords** Detection · DNA barcoding · Identification · ISSR · Phylogenetics · *rbcL*

## Abbreviations

| | |
|---|---|
| ISSR | Inter Simple Sequence Repeat |
| *rbcL* | Ribulose-1, 5-bisphosphate carboxylase/oxygenase large subunit |
| PCoA | Principal Coordinate Analysis |
| HCA | Hierarchical Cluster Analysis |
| CP | Cophenetic Correlation Coefficient |
| UPGMA | Unweighted Pair Group Method with Arithmetic Mean |

## Introduction

Allele frequencies in any species may change from generation to generation. This happens due to the influence of multiple forces that may include mutation, selection, drift, and gene flow as well as various other constraints viz. history, demography, development, genomic structure and environment, resulting in 'Biological Evolution' (Jorde and Ryman 1995; Ellstrand 2014). These changes are effected through a plethora of biochemical pathways that comprise an interplay of the actions of many enzymatic reactions that produce primary and secondary metabolites. Evolutionary consequences, coupled with environmental changes may lead to dramatic variations in the genome of a species, vis-à-vis their phytocostituents (Wolf et al. 1998; Briskin 2000; Shukla et al. 2018). Cladistic or phylogenetic analysis of molecular genetic characteristics of a plant species, therefore, has tremendous potential as a tool for use in understanding the relationships, origin or lineage of the plant (Schaal et al. 1998; Carvalho et al. 2012). Cluster analysis through phylogenetics and multivariate analysis has received considerable attention not only for genetic studies, but also for studies involving diverse array of phytoconstituent fluctuations (Palmblad

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s13205-018-1175-5) contains supplementary material, which is available to authorized users.

✉ Subarna Roy
roys@icmr.gov.in; drsubarnaroy@gmail.com

[1] ICMR-National Institute of Traditional Medicine, Indian Council of Medical Research, Department of Health Research, Government of India, Belagavi, Karnataka 590010, India

[2] KLE Academy of Higher Education and Research (Deemed-to-be-University), Dr. Prabhakar Kore Basic Science Research Centre, Belagavi, Karnataka 590010, India

and Deelder 2012). Inferences derived from phylogenetic analysis are finding more and more use in authentication and development of quality control parameters for commercially important plants (Jensen et al. 2012; Feng et al. 2014; Pendkar et al. 2016).

*Saraca asoca* (Roxb.) De Wilde (Caesalpiniaceae), commonly known as "Ashoka" is one of the most highly traded medicinal trees (Nadkarni 1976; Khare 2007; Tandon and Yadav 2017). It is native to south and central Western Ghats in India and grows in Sri Lanka as well. *S. asoca* is greatly valued for treatment of gynecological disorders. The bark extract has been reported to possess a variety of therapeutic effects that include antitumor/anticarcinogenic, and antimicrobial activities and against skin diseases (Nadkarni 1976; Singh et al. 2015). The bark, which is a strong astringent and uterine sedative, has a stimulating effect on endometrium and ovarian tissues (Nadkarni 1976; Tandon and Yadav 2017). It is used as the main ingredient in several commercial Ayurvedic preparations like '*Ashokarishtam*' and '*Ashokaghritham*' and therefore, it is immensely exploited by the phytopharmaceutical industry. *S. asoca* figures in Red List of 'Threatened Species' by International Union for Conservation of Nature (IUCN) and is reported to be endangered. It is suggested that extensive substitution and adulteration of the crude drug might be taking place to match increasing demand (Singh et al. 2015). Urumarudappa et al. 2016 also reported occurrence of extensive adulteration (80%) in crude drug of *S. asoca*. Literature reveals its substitution and adulteration with different plant materials e.g., *Trema orientalis*, *Bauhinia variegata*, *Mesua ferrea*, *Shorea robusta*, and *Polyalthia longifolia* (Sarin 1996; Anonymous 2005; Singh et al. 2015; Hegde et al. 2017a).

*Polyalthia longifolia* (Annonaceae) is such a common adulterant of *S. asoca* that it is also often called locally as "Ashoka". Therefore, *P. longifolia* has earned the common name "False Ashoka" in English. Like *S. asoca*, it is also native to India and Sri Lanka and widely cultivated and distributed also in Bhutan, China and many tropical countries. *Trema orientalis* (Cannabaceae) is another adulterant/substituent of *S. asoca* that is found in Australia, Africa and Asian countries. Leaf, stem and root of this tree have been reported for treatment of diarrhoea and epilepsy (Nadkarni 1976; Khare 2007). *Bauhinia variegata* is native to China, Myanmar, North Thailand, Peoples Democratic Republic of Laos, and North Vietnam. It is used to treat haematuria, and menorrhagia while it also shows antidiabetic, antioxidant and anti-inflammatory properties (Nadkarni 1976; Khare 2007; Farag et al. 2015). *Shorea robusta* is native to the Indian subcontinent and its bark contains a novel benzofuran, shoreaphenol (Patra et al. 1992). It is mainly used as timber in India. However, medicinally, it acts as antidiarrhoeal and antidysentric agent and the oil is used to treat skin diseases (Khare 2007). *Mesua ferrea* is an angiosperm belonging to the family Calophyllaceae which is native to Cambodia, India, Malaysia, Myanmar, Philippines, Singapore, Sri Lanka, Thailand, and Vietnam. It contains compounds like 1, 5-dihydroxyxanthone (II), euxanthone7-methyl ether (IV) and β-sitosterol and is used to treat various diseases like bleeding piles, gastritis and bronchitis (Chow and Quon 1968; Nadkarni 1976; Khare 2007). The present study was undertaken with the aim to develop and evaluate multivariate cluster analysis of ISSR fingerprints against *rbcL*-based DNA barcodes as a tool to understand the relationships and to differentiate common adulterants and substituents from *S. asoca*.

## Materials and methods

### Plant sampling and identification

Leaves from 41 individuals of *S. asoca* were collected, labeled with separate laboratory identification codes, and stored at − 80 °C (Table 1). Collections were made from six different states that consisted of ten localities to ensure that larger/diverse populations are included for assessment of intra-specific variations. For the purpose of differential genetic profiling of the adulterant/substituent plants, leaf samples of five individuals of each species of *Trema orientalis*, *Bauhinia variegata*, *Mesua ferrea*, *Polyalthia longifolia* and *Shorea robusta* were collected, assigned laboratory identification codes and kept separately. Voucher specimens from the collected plant samples were identified by a taxonomist and deposited in the herbarium at ICMR-National Institute of Traditional Medicine, Belagavi, India.

### DNA extraction

Genomic DNA was extracted from leaf samples of all the collected plants individually using modified CTAB method (Richards et al. 1994). The quantity and purity of DNA was measured using Nanodrop Spectrophotometer (JH BIO, USA) and by confirmation of PCR amplification with agarose gel electrophoresis. Final dilution was made up to 40 ng/μL with TE buffer (10 mM Tris HCI, pH 8.0 and 0.1 mM EDTA, pH 8.0) and stored at − 20 °C until further use.

**Table 1** Details of plant sampling including NCBI GenBank accession numbers

| Sl. no. | Plant name | Location/collection place | Accession code (Lab Code) | NCBI GenBank accession number | Geographical attribution: longitude and latitude |
|---|---|---|---|---|---|
| 1 | *S. asoca* | Shravasti, Uttar Pradesh | S 1-5; SA 1-5 | KY678332, KY678333, KY678334, KY678335, KY678336 | 27°34.2952°N 81°35.48°E |
| 2 | *S. asoca* | Thiruvananthapuram, Kerala | KL 1-5; SA 6-10 | KY678327, KY678328, KY678329, KY678330, KY678331 | 08°51.1847°N 76°95.2624°E |
| 3 | *S. asoca* | Alase (Shivamogga), Karnataka | THI 1-5; SA 11-15 | KY678312, KY678313, KY678314, KY678315, KY678316 | 13°83.237′N 075°36.745′E |
| 4 | *S. asoca* | Ghativade, Maharashtra | GHA 1-5; SA 16-20 | KY678322, KY678323, KY678324, KY678325, KY678326 | 15°80.309°N 074°13524°E |
| 5 | *S. asoca* | Bondla, Goa | BON 1-5; SA 21-25 | KY678302,KY678303,KY678304, KY678305,KY678306 | 15°45.760°N 074°09042°E |
| 6 | *S. asoca* | Devimane Ghat, Karnataka | DEV 1-5; SA 25-30 | KY678317, KY678318, KY678319, KY678320, KY678321 | 14°31.255°N 074°33.979°E |
| 7 | *S. asoca* | Navsari, Gujarat | GU 1-5; SA 31-36 | KY678307, KY678308, KY678309, KY678310, KY678311 | 20°55.482N 72°54.565E |
| 8 | *S. asoca* | Kolhapur, Maharashtra | K1, K2; SA 37-38 | KY678337, KY678338 | 16°42.045N 74°13.254E |
| 9 | *S. asoca* | Sangli, Maharashtra | SG1, SG2; SA 39-40 | KY678339, KY678340 | 16°51′520N 74°34′200E |
| 10 | *S. asoca* | Vandane, Karnataka | V-2; 41 | KY678341 | 14°21.413N 074°480E |
| 11 | *P. longifolia* | Belagavi, Karnataka | PL1 | KY654491 | 15°88.332N 074°52.379°E |
| 12 | *P. longifolia* | Belagavi, Karnataka | PL2 | KY654492 | 15°88.334°N 074°52.376°E |
| 13 | *P. longifolia* | Belagavi, Karnataka | PL3 | KY654493 | 15°88.337°N 074°52.381°E |
| 14 | *P. longifolia* | Dharwad, Karnataka | PL4 | KY654494 | 15°43.4652 N 74°98.69.36E |
| 15 | *P. longifolia* | Pashan, Pune | PL5 | KX010596 | 18°54.3294°N 73°78.80.19°E |
| 16 | *T. orientalis* | Dandeli, Karnataka | TO1 | KY654499 | 15°22.20°N; 074°50.07°E |
| 17 | *T. orientalis* | Ambika Nagara, Karnataka | TO2 | KY654500 | 15°19.98°N; 074°64.78°E |
| 18 | *T. orientalis* | Belagavi, Karnataka | TO3 | KY654501 | 15°88.7665N 74°52.2699°E |
| 19 | *T. orientalis* | Belagavi, Karnataka | TO4 | KY654502 | 15°88.7665N 74°52.2698°E |
| 20 | *T. orientalis* | Belagavi, Karnataka | TO5 | KX010598 | 15°88.7665N 74° 52.2696°E |
| 21 | *B. variegata* | Khanapur, Karnataka | BV1 | KY654483 | 15°65.28N 074°50.07°E |
| 22 | *B. variegata* | Belagavi, Karnataka | BV2 | KY654484 | 15°88.7665 N74° 52.261°E |
| 23 | *B. variegata* | Belagavi, Karnataka | BV3 | KY654485 | 15°88.7665 N 74°52.2699°E |
| 24 | *B. variegata* | Belagavi, Karnataka | BV4 | KY654486 | 15°88.7665N 74°52.2681°E |
| 25 | *B. variegata* | Belagavi, Karnataka | BV5 | KX010594 | 15°88.7665N 74°52.2690°E |
| 26 | *M. ferrea* | Belagavi, Karnataka | MF1 | KY654487 | 15°50.112N 074°30.355°E |
| 27 | *M. ferrea* | Belagavi, Karnataka | MF2 | KY654488 | 15°887665N 74°52.2699°E |
| 28 | *M. ferrea* | Belagavi, Karnataka | MF3 | KY654489 | 15°83.2926N 74°50.7335°E |
| 29 | *M. ferrea* | Belagavi, Karnataka | MF4 | KY654490 | 15°83.2930N 74°50.7327°E |
| 30 | *M. ferrea* | Dharwad, Karnataka | MF5 | KX010595 | 15°43.737N 74°98.946°E |
| 31 | *S. robusta* | Dehradun, Uttarakhand | SR1 | KY654495 | 30°16.42N 78°2.51°E |
| 32 | *S. robusta* | Dehradun, Uttarakhand | SR2 | KY654496 | 30°16.42N 78°2.51°E |
| 33 | *S. robusta* | Lucknow, Uttar Pradesh | SR3 | KY654497 | 26°85.817N 80°94.9801°E |
| 34 | *S. robusta* | Laxmanpur, Uttar Pradesh | SR4 | KY654498 | 27°38.51N 82°3.54°E |
| 35 | *S. robusta* | Gonda, Uttar Pradesh | SR5 | KX010597 | 27°0.18N 82°21.12°E |

## ISSR and *rbcL* amplifications

ISSR fingerprinting assays were carried out with well-established UBC primer # set 9 (University of British Columbia, Canada; synthesized from Sigma-Aldrich, India). A primer UBC 815 (3′-CTCTCTCTCTCTCTCTG-5′) that resulted in the most consistent and reproducible bands was selected for all subsequent ISSR assays which were performed in 25 µL reaction volumes containing 40 ng genomic DNA, 10 µM primer, 200 µM of each dNTP, 3U/µL of Taq DNA polymerase (Merck, India) and 10X PCR buffer (Tris HCl, pH 9.0; 15 mM $MgCl_2$; Merck, India). ISSR assays were performed following cycling condition described by Hegde et al. (2017b).

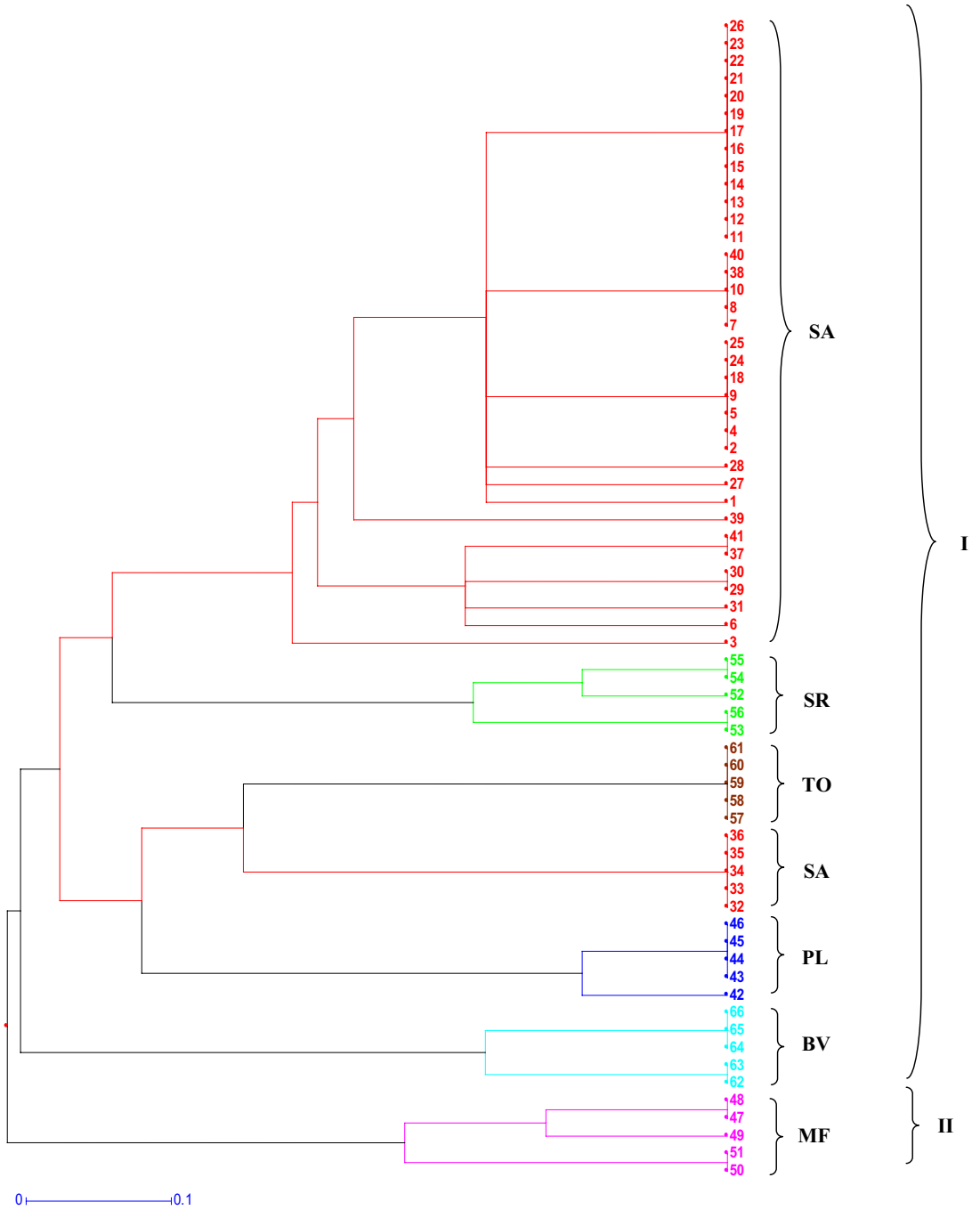


**Fig. 1** ISSR-based Hierarchical Clusters Analysis (HCA) of 41 individuals of *Saraca asoca* (SA), and 5 samples each from *Trema orientalis* (TO), *Bauhinia variegata* (BV), *Mesua ferrea* (MF), *Polyalthia longifolia* (PL), *Shorea robusta* (SR)

ISSR amplifications were performed in Mastercycler® Nexus (Eppendorf AG, Germany) thermocycler. The amplified products were visualized after gel electrophoresis (Bio-Rad, USA) on 1.5% agarose gels in 1X TAE buffer using GelRed (Biotium, USA) as staining dye. The agarose gels were visualized and documented using a gel documentation system (Syngene, UK). Amplification was repeated thrice to confirm the reproducibility of the bands and only the consistent and reproducible bands were considered for analysis.

Amplification of *rbcL* region was performed in 25 µL reaction volumes containing 40 ng genomic DNA, 10 µM primer, 200 µM of each dNTP, 3U/µL of Taq DNA polymerase (Merck, India) and 10X PCR buffer (Tris HCl, pH 9.0; 15 mM $MgCl_2$; Merck, India). To enhance and facilitate the PCR amplification, spermidine (10 mM) and DMSO (100%)

**Table 2** *rbcL* primers and PCR amplification conditions used for the study

| Primer name | Sequence (5′–3′) | Reference |
|---|---|---|
| *rbcL*-F | ATGTCACCACAAACAGAGACTAAAGC | Levin et al. 2003 |
| *rbcL*-R | GTAAAATCAAGTCCACCRCG | Kress et al. 2009 |
| *rbcL* PCR amplification conditions | | |
| 98 °C for 45 s, 98 °C for 10 s, 55 °C for 30 s, 72 °C for 40 s with 35 cycles and 72 °C for 10 min | | |

was used in each assay (Wan et al. 1993; Rasmussen et al. 1994). Details of amplification conditions and primer are provided in Table 2. PCR products were sequenced after confirming amplification in agarose gel, and purification using Mini Elute PCR Purification Kit (Qiagen).

## Data analysis

Amplification of each ISSR PCR fragment from each individual sample was recorded from agarose gel profiles as in binary matrix. Presence (1) or absence (0) of bands were scored for each genotype. Hierarchical Cluster Analysis (HCA) was performed for all samples and a dendrogram was constructed from the scored binary matrix through DARwin 6.0.12 software (Perrier and Jacquemoud-Collet 2006). These scored binary matrixes were also used to compute Jaccard similarity coefficients into distances using program DendroUPGM (Garcia-Vallve et al. 1999). Circular dendrograms were constructed using UPGMA by alignment of leaves and auto sort option through PhyloWidget program using Jaccard matrix (Jordan and Piel 2008). Distance matrix generated by DARwin software was also used to perform Factorial Analysis. Graphical representations of the individual's relationships were performed to visualize grouping through Principal Coordinate Analysis (PCoA) and the percentage of polymorphic loci was calculated using GenAlEx version 6.5. (Peakall and Smouse 2012). Molecular variance was also computed via covariance matrix with data standardization using GenAlEx version 6.5.

The raw sequences were manually trimmed using Bioedit Sequence Alignment Editor Version 7.2.5 (Hall 1999). The sequences were aligned with Clustal W (Thompson et al. 1994) and matched in NCBI—Basic Local Alignment Search Tool (BLAST) search to compare the desired gene. The multiple sequence alignment and UPGMA analyses of sequences was performed using MEGA version 7.0.14 (Kumar et al. 2016). Population genetic parameters like number of segregating sites ($S$); total number of mutations that occurred in that population ($\eta$); haplotype (gene) diversity ($Hd$), nucleotide diversity ($\pi$) were estimated by pair wise comparisons (Nei 1987). Average number of differences genetic diversity (K) per species was calculated using the program DnaSP version 4.0 (Rozas et al. 2003). SIAS server (https://imed.med.ucm.es/Tools/sias.html) was used to compute sequence identity percentage and factorial analysis was performed using DARwin 6.0.12 (Perrier and Jacquemoud-Collet 2006).

# Results and discussions

## Polymorphism

Amplification of genomic DNA by a ISSR primer UBC 815 from *S. asoca* (SA), *M. ferrea* (MF), *P. longifolia* (PL), *T. orientalis* (TO), *S. robusta* (SR), *B. variegata* (BV), yielded total 26 scorable loci. Briefly, 41 individuals of SA produced 7 scorable loci, while the 5 individuals of PL and SR produced 5 each, MF produced 4, TO produced 2 and BV produced 3. Although, 20 primers were screened, primer UBC815 showed the best discrimination power between the species with the samples used in the present investigation (Fig. 1a and b: Supplementary material). The Jaccard similarity coefficient, which measure the similarity between two sets of binary data, viz., the size of the intersection divided by the size of the union of the sample sets was computed through Phylowidget, which ranged from 0.11 to 1.0 for all the 66 individuals taken across species. Cophenetic Correlation Coefficient (CP) is a measure of how faithfully a dendrogram preserves the pairwise distances between the original unmodeled data points and in the present study, it showed a value of 0.930, indicating robust matching of dataset. The mean percentage of polymorphic loci, calculated from the binary matrix through GenAlEx version 6.5, was 13.73% (Standard Error: SE = 6.90%) when taken in total. Individually for SA it was 47.06%, for PL it was 5.88%, MF – 11.76%, SR – 11.76%, and for BV it was 5.88%. No polymorphic bands were observed in case of TO.

The aligned sequences of *S. asoca rbcL* gene were queried against available sequences in the NCBI BLAST and it showed the best match with *S. asoca* (GenBank accession KU499910.1, KU499909.1, KU499908.1) and also with *Saraca palembanica* and *Saraca declinata* (GenBank accession AM234238.1, JX856760.1). The sequences of BV showed similarity with *Bauhinia variegata, Bauhinia galpinii* (viz., GenBank accession JX571784.1, AM234262.1).

Similarly, MF showed similarities with *Mesua ferrea* (viz., GenBank accession GQ436685.1), PL with *Polyalthia longifolia, Monoon coffeoides* (viz., GenBank accession AY319027.1, EU522288.1), SR with *Shorea robusta*, *Shorea obtusa* (viz., GenBank accession JX856763.1, AB925320.1) and TO with *Trema orientalis*, *Trema cannabina* (viz., GenBank accession KM895509.1, KP094231.1). The percentage of sequence identity was calculated and is presented in Table 3. The sequence of all the samples showed 100% identity with their respective species, except PL, which showed minor variation ranging from 99.42 to 99.43% (Table 3). This might be because of geographical constraints and its cultivation as an ornamental tree in different regions of the world. The sequence identity variation of BV1 to BV4 showed 89.84% against PL and rest of the sequences showed above 90% identity. The sequence variation between the species depicts the allelic alteration among the *rbcL* region, which is one of the conserved chloroplast regions (Kress et al. 2005).

Within the entire data set, there was a total of 519 sites excluding sites with gaps/missing data from a total of 564 sites. In that, there were 91 observed number of variable/polymorphic sites (*S*), and 98 total numbers of mutations ($\eta$) (Fu and Li 1993). There was 0.594 (SD: 0.066) average haplotype (gene) diversity (*Hd*), and 0.04038 (SD: 0.00497) average number of nucleotide differences per site between two sequences which is called as nucleotide diversity $\pi$ (per site) (Nei et al. 1987). The nucleotide divergence which in this case is the average proportion of nucleotide difference between species [with Jukes and Cantor = K (JC)], from SA to others are K (JC-Silent): 0.04087; K (JC-Silent): 0.07372; K (JC-Silent): 0.07918; K (JC-Silent): 0.08132; K (JC-Silent): 0.06270 in BV, MF, PL, SR and TO, respectively. Also, as evident from the sequence identity matrix, there was less sequence divergence (Table 3). This might be due to the conserved nature of the *rbcL* region, which also makes it suitable for identification purpose as a DNA barcode (Kress et al. 2005).

## Multivariate cluster analysis

Dendrogram generated from ISSR-based binary matrix using DARwin 6.0.12 software showed two major clusters in which cluster I consisted of SA, SR, TO, PL and MF samples, whereas cluster II contained of only MF samples (Fig. 1). Within cluster I, SA grouped in two different minor clusters with SA 32–36 forming the smaller group and SA 1–41 forming the larger group. SA 32–36 appeared closer to TO in cluster I. SA 1, 3, 6, 26 and 27 appeared very distinct in the larger sub-group of SA in cluster I. BV samples in cluster II had two genetic types as was the case with PL in cluster I. SR and MF exhibited three types. Simplicifolious separation of SA (32–36) and TO (57–61) from others and their occurrence closer to each other is an interesting observation made in the study. The results generated with Phylowidget software also shows very similar clustering patterns as depicted (Fig. 2). Although, in ISSR-derived analysis, all the individual population clusters were distinct, SA, SR and TO tended to be close in both the cluster analyses done. Figure 3 represents UPGMA-based phylogenetic tree generated using *rbcL* sequences from all the 66 samples. The optimal tree with the sum of branch length is 0.218 with 519 positions in the sequence data. It produced two major clusters in which cluster I consisted of five minor clusters. The upper half of the tree contained only SA samples followed by BV, TO, MF, PL (Fig. 3). The sequences from SR samples produced different major cluster (cluster II). Therefore, in the present study, *rbcL*-sequence-based phylogenetic inferences were found position species on clear clusters that could be used easily to find origins and to aid identification (Bodin et al. 2016).

Factorial analysis was carried out for confirming the relationships seen in the genetic dendrogram. It is a different approach in the representation of structure and species distinctness, and considered complementary to each other (Wang et al. 2016). With the ISSR data matrix, although factorial analysis did not separate out all the individuals, it grouped them distinctly according to species affinities. However, *rbcL*-based factorial analysis showed clear representation of each species and clearly differentiated them from amongst each other. Although, in dendrograms, SA was clustered closer to adulterants and substituents (TO/SR/BV), factorial analysis revealed six distinct groups both in *rbcL* (Fig. 4a) and ISSR profiles (Fig. 4b) representing their species and origin. The data similarities or dissimilarities were visualized by ISSR based PCoA analysis (Fig. 5) in which all the groups distinctly separated from SA and corroborated the results of factorial analysis using both ISSR and *rbcL* sequences (Fig. 4). In both factorial analysis and PCoA analyses, the tested individual samples showed distinct groups of genetic pools of respective taxa thereby confirming its utility as a useful tool for differentiation between the genuine and the adulterant/substituent groups.

Although, HCA analyses of ISSR fingerprints and *rbcL* sequences generated for *S. asoca* and its common adulterants/substituents revealed closeness of some genotypes with TO, BV and SR, it was able to group them into distinct clusters. Factorial analysis and PCoA analysis helped in generation of distinct clusters that grouped them according to their species. BLAST search identified the exact matches of the respective sequences and confirmed the results of ISSR assays. Therefore, our study demonstrated the utility of simple, cost-effective method of ISSR fingerprinting in differentiating this important medicinal plant from its common adulterants/substituents. Because of the wider distribution of the substituent/adulterant plants, some of which are

**Table 3** *rbcL*-based sequence identity percentage of *S. asoca* with its adulterants and substituents

| Sl. no. | Samples | KU981156.1/SA (%) | KX010594.1/BV (%) | KX010595.1/MF (%) | KX010596.1/PL (%) | KX010597.1/SR | KX010598.1/TO (%) |
|---|---|---|---|---|---|---|---|
| 1 | KU981156.1/SA | 100 | 96.02 | 92.97 | 90.22 | 91.77 | 93.45 |
| 2 | KX010594.1/BV | 96.02 | 100 | 94.30 | 90.53 | 92.42 | 93.37 |
| 3 | KX010595.1/MF | 92.97 | 94.30 | 100 | 91.46 | 90.70 | 93.73 |
| 4 | KX010596.1/PL | 90.22 | 90.53 | 91.46 | 100 | 90.60 | 90.97 |
| 5 | KX010597.1/SR | 91.77 | 92.42 | 90.70 | 90.60 | 100 | 91.12 |
| 6 | KX010598.1/TO | 93.45 | 93.37 | 93.73 | 90.97 | 91.12 | 100 |
| 7 | BON-1 | 100 | 96.02 | 92.97 | 90.22 | 91.77 | 93.45 |
| 8 | BON-2 | 100 | 96.02 | 92.97 | 90.22 | 90.75 | 91.57 |
| 9 | BON-3 | 100 | 96.02 | 92.97 | 90.22 | 90.75 | 91.57 |
| 10 | BON-4 | 100 | 96.02 | 92.97 | 90.22 | 90.75 | 91.57 |
| 11 | BON-5 | 100 | 96.02 | 92.97 | 90.22 | 90.75 | 91.57 |
| 12 | DEV-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 13 | DEV-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 14 | DEV-3 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 15 | DEV-4 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 16 | DEV-5 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 17 | GHA-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 18 | GHA-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 19 | GHA-3 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 20 | GHA-4 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 21 | GHA-5 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 22 | GU-1 | 99.81 | 96.02 | 92.97 | 90.22 | 92.12 | 93.80 |
| 23 | GU-2 | 99.81 | 96.02 | 92.97 | 90.22 | 92.12 | 93.80 |
| 24 | GU-3 | 99.81 | 96.02 | 92.97 | 90.22 | 92.12 | 93.80 |
| 25 | GU-4 | 99.81 | 96.02 | 92.97 | 90.22 | 92.12 | 93.80 |
| 26 | GU-5 | 99.81 | 96.02 | 92.97 | 90.22 | 92.12 | 93.80 |
| 27 | KL-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 28 | KL-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 29 | KL-3 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 30 | KL-4 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 31 | KL-5 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 32 | S-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 33 | S-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 34 | S-3 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 35 | S-4 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 36 | S-5 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 37 | K-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |

**Table 3** (continued)

| Sl. no. | Samples | KU981156.1/SA (%) | KX010594.1/BV (%) | KX010595.1/MF (%) | KX010596.1/PL (%) | KX010597.1/SR | KX010598.1/TO (%) |
|---|---|---|---|---|---|---|---|
| 38 | K-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 39 | SG-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 40 | SG-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 41 | V-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 42 | THI-1 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 43 | THI-2 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 44 | THI-3 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 45 | THI-4 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 46 | THI-5 | 99.62 | 96.02 | 92.97 | 90.22 | 91.94 | 93.63 |
| 47 | BV-1 | 96.05 | 100 | 94.30 | 89.84 | 92.48 | 93.42 |
| 48 | BV-2 | 96.05 | 100 | 94.30 | 89.84 | 92.48 | 93.42 |
| 49 | BV-3 | 96.05 | 100 | 94.30 | 89.84 | 92.48 | 93.42 |
| 50 | BV-4 | 96.05 | 100 | 94.30 | 89.84 | 92.48 | 93.42 |
| 51 | MF-1 | 93.04 | 94.31 | 100 | 90.60 | 90.78 | 93.79 |
| 52 | MF-2 | 93.04 | 94.31 | 100 | 90.60 | 90.78 | 93.79 |
| 53 | MF-3 | 93.04 | 94.31 | 100 | 90.60 | 90.78 | 93.79 |
| 54 | MF-4 | 93.04 | 94.31 | 100 | 90.60 | 90.78 | 93.79 |
| 55 | PL-1 | 91.13 | 91.09 | 92.03 | 99.43 | 91.50 | 91.32 |
| 56 | PL-2 | 91.13 | 91.09 | 92.03 | 99.43 | 91.50 | 91.32 |
| 57 | PL-3 | 91.82 | 91.44 | 92.20 | 99.42 | 91.44 | 91.25 |
| 58 | PL-4 | 91.82 | 91.44 | 92.20 | 99.42 | 91.44 | 91.25 |
| 59 | SR-1 | 91.77 | 92.42 | 90.70 | 90.60 | 100 | 90.29 |
| 60 | SR-2 | 91.77 | 92.42 | 90.70 | 90.60 | 100 | 90.29 |
| 61 | SR-3 | 91.77 | 92.42 | 90.70 | 90.60 | 100 | 90.29 |
| 62 | SR-4 | 91.77 | 92.42 | 90.70 | 90.60 | 100 | 90.29 |
| 63 | TO-1 | 93.45 | 93.37 | 93.73 | 90.78 | 91.12 | 100 |
| 64 | TO-2 | 93.45 | 93.37 | 93.73 | 90.97 | 91.12 | 100 |
| 65 | TO-3 | 93.45 | 93.37 | 93.73 | 90.97 | 91.12 | 100 |
| 66 | TO-4 | 93.45 | 93.37 | 93.73 | 90.97 | 91.12 | 100 |

*SA*: *S. asoca*, *BV*: *B. variegata*, *MF*: *M. ferrea*, *PL*: *P. longifolia*, *SR*: *S. robusta*, *TO*: *T. orientalis*

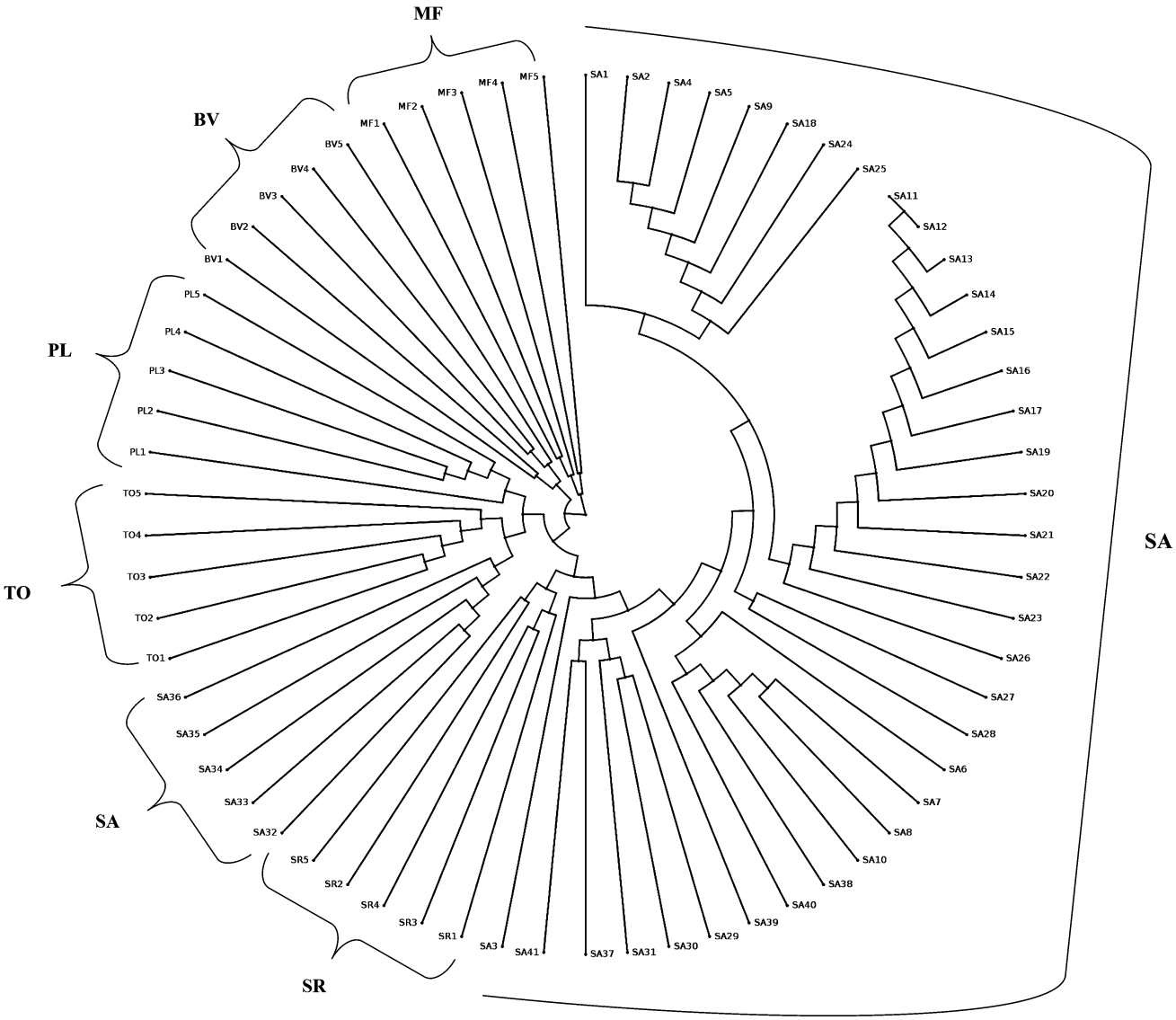


**Fig. 2** UPGMA dendrogram of ISSR fingerprints from 41 individuals of *Saraca asoca* (SA) and 5 samples each from *Trema orientalis* (TO), *Bauhinia variegata* (BV), *Mesua ferrea* (MF), *Polyalthia longifolia* (PL), *Shorea robusta* (SR)

cultivated extensively, inclusion of a larger number of these samples would have been practically beyond the scope of this study and perhaps would not serve the purpose significantly better. Despite this limitation, the results obtained in this study could be well utilized in carrying out further field studies with market samples of the crude drug of *S. asoca*.

Utility of phylogenetic trees in differentiation of *S. asoca* from common substituents/adulterants through powerful visualization using multiple tools including factorial analysis and PCoA (Jordan and Piel 2008; Mooi and Sarstedt 2011) has been well demonstrated in the present study. The findings also warrant further studies on method development for resolving quality issues in medicinal products made from *S. asoca*. To the best of our knowledge, this is one of the first studies that utilizes ISSR markers and *rbcL*-based barcoding in differentiating *S. asoca* from its common adulterants/substituents.

DNA-based techniques, which are independent of environmental dependency have an advantage over other traditional methods and can be exploited for addressing genuineness in crude drugs (Smillie and Khan 2010; Pendkar et al. 2016). Although various DNA-based methods are available, one simple and cost-effective method was chosen for characterization of the samples along with another more reliable and robust confirmatory technique. ISSR is one of the simplest and most cost-effective techniques

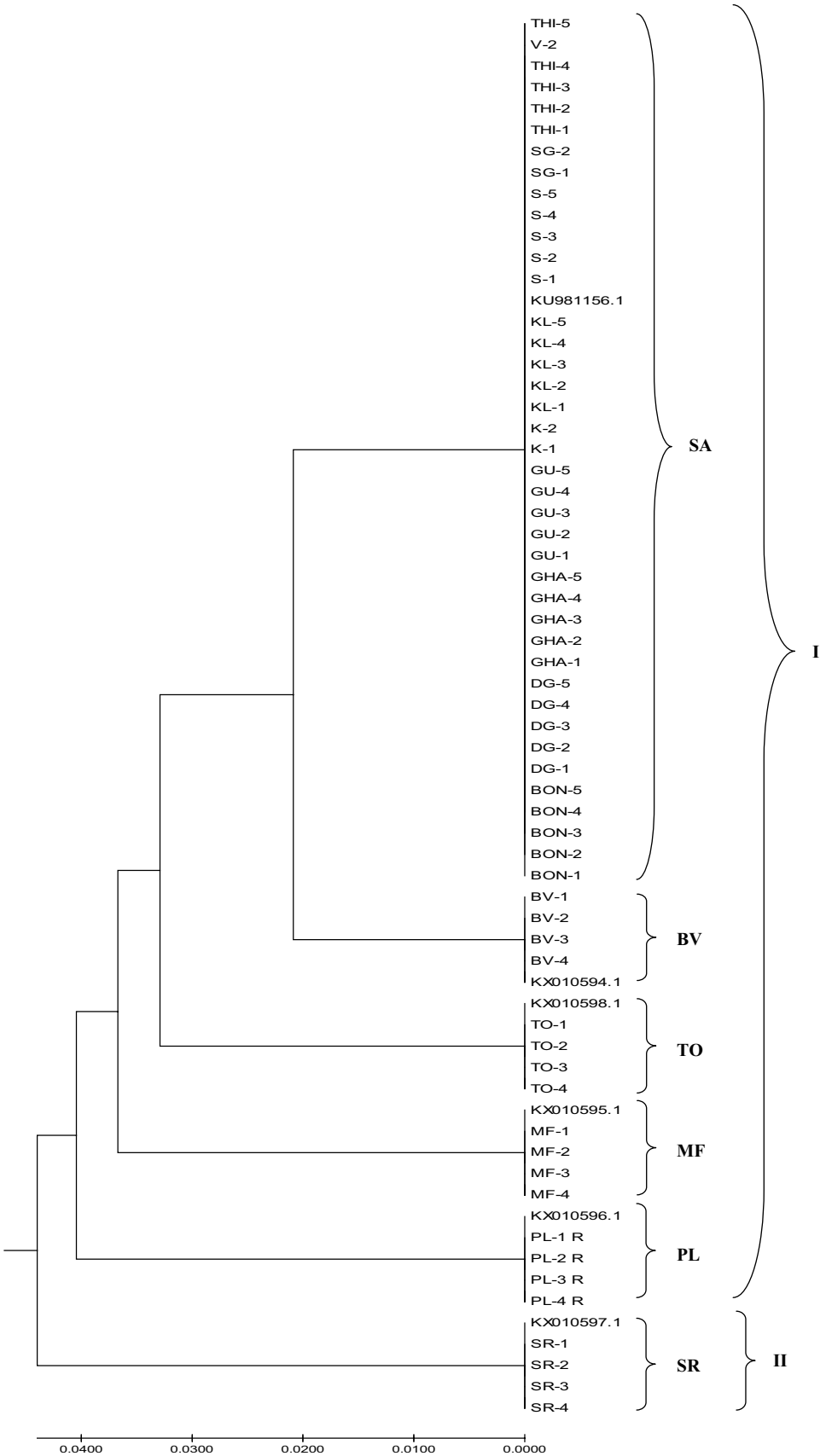


**Fig. 3** UPGMA-based phylogenetic tree of 41 samples of *Saraca asoca* with 5 samples each from *Trema orientalis* (TO), *Bauhinia variegata* (BV), *Mesua ferrea* (MF), *Polyalthia longifolia* (PL), *Shorea robusta* (SR) (see Table 1) using *rbcL* sequences

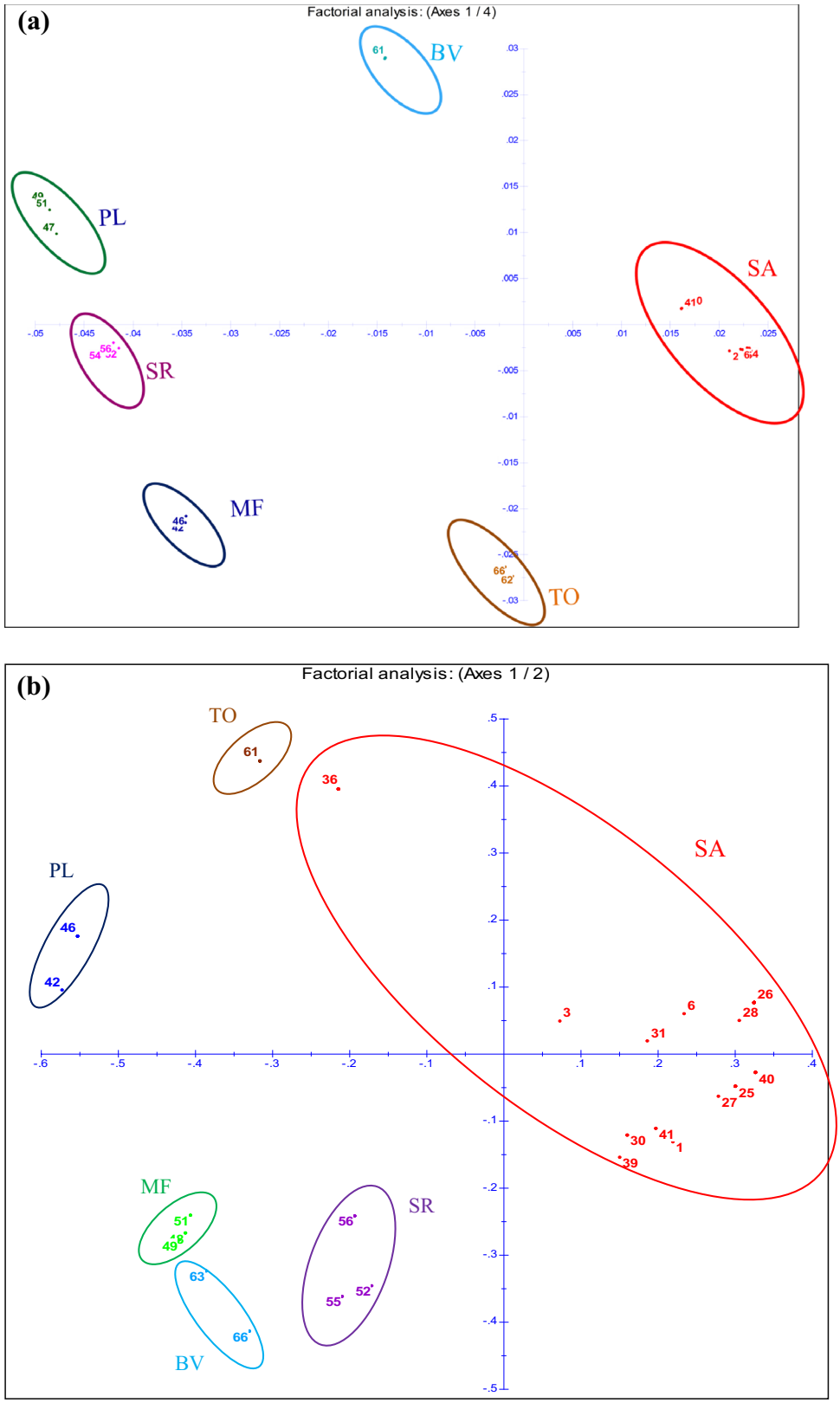


**Fig. 4** **a** Factorial analysis of *rbcL* sequences including 41 samples of *Saraca asoca* (SA) with 5 samples each from *Trema orientalis* (TO), *Bauhinia variegata* (BV), *Mesua ferrea* (MF), *Polyalthia longifolia* (PL), *Shorea robusta* (SR); **b** ISSR-based factorial analysis of 41 samples of *Saraca asoca* (SA) with 5 samples each from *Trema orientalis* (TO), *Bauhinia variegata* (BV), *Mesua ferrea* (MF), *Polyalthia longifolia* (PL), *Shorea robusta* (SR)

available for DNA-based fingerprinting which is considered to be superior to other random fingerprinting methods. This technique can be adopted and carried out in a wider number of laboratories and is therefore practically more feasible than other methods requiring more expertise and high-end technologies. Although a variety of genes

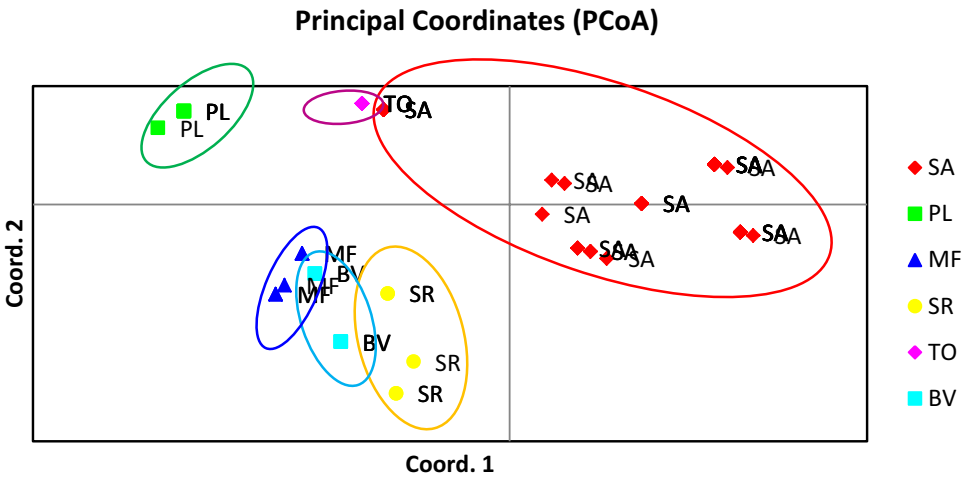


**Fig. 5** ISSR-based Principal Coordinate Analysis (PCoA) of 41 samples of *Saraca asoca* (SA) with 5 samples each from *Trema orientalis* (TO), *Bauhinia variegata* (BV), *Mesua ferrea* (MF), *Polyalthia longifolia* (PL), *Shorea robusta* (SR)

have been tested as candidates for barcoding, in the present study *rbcL* was used as it produced better amplifications than others in preliminary studies (not shown). However, in the present study DNA barcoding with *rbcL* region was used to confirm the findings and utility of ISSR-based fingerprinting analyses and drawing of reliable conclusions.

The present study revealed that multivariate phylogenetic clustering coupled with factorial analysis and PCoA analysis can be effectively used to detect and differentiate *S. asoca* from its common substituents/adulterants. The present study may also help the regulatory authorities in using/recommending use of ISSR markers and/or *rbcL* barcoding as potential tools for differentiating adulterants and substituents of *S. asoca*.

**Acknowledgements** Authors are thankful to the Director-in-Charge, ICMR-NITM, formerly RMRC, Belagavi, India, and Indian Council of Medical Research (Government of India), for supporting this study. Authors are also thankful to Dr. Vidya S. Gupta, Biochemical Science Division, National Chemical Laboratory, Pune, for technical suggestions and Bio-Medical Informatics Centre of ICMR at NITM, Belagavi, for informatics support. Authors extend thanks to Pramod Kumar and Jotiba B. Palekar, ICMR-NITM, Belagavi, Anil Bisht (Dehradun), Jayendrasinh Chavda (Gujarat) and Aparna J (Kerala), India, for help during sample collection. SH is grateful to ICMR for funding.

## Compliance with ethical standards

**Funding** This study was funded by The Indian Council of Medical Research, New Delhi (Grant No. 45/53/2013/BMS/TRM; IRIS ID: 2013-18920).

**Conflict of interest** All authors declare that there is no conflict of interest.

**Ethical approval** Not applicable.

## References

Anonymous (2005) *Saraca asoca*. In: Gupta AK, Tandon N, Sharma M (eds) Quality standards of indian medicinal plants, 2nd edn. Indian Council of Medical Research (ICMR), New Delhi, pp 201–208

Bodin SS, Kim JS, Kim JH (2016) Phylogenetic inferences and the evolution of plastid DNA in Campynemataceae and the Myco-heterotrophic Corsia dispar DL Jones and B Gray (Corsiaceae). Plant Mol Biol Rep 34(1):192–210

Briskin DP (2000) Medicinal plants and phytomedicines: linking plant biochemistry and physiology to human health. Plant Physiol 124(2):507–514

Carvalho FA, Renner SS (2012) A dated phylogeny of the papaya family (Caricaceae) reveals the crop's closest relatives and the family's biogeographic history. Mol Phylogenet Evol 65(1):46–53

Chow YL, Quon HH (1968) Chemical constituents of the heartwood of *Mesua ferrea*. Phytochemistry 7(10):1871–1874

Ellstrand NC (2014) Is gene flow the most important evolutionary force in plants? Am J Bot 101(5):737–753

Farag MA, Sakna ST, El-fiky NM, Shabana MM, Wessjohann LA (2015) Phytochemical, antioxidant and antidiabetic evaluation of eight *Bauhinia* L species from Egypt using UHPLC–PDA–qTOF-MS and chemometrics. Phytochemistry 119:41–50

Feng SG, Lu JJ, Gao L, Liu JJ, Wang HZ (2014) Molecular phylogeny analysis and species identification of *Dendrobium* (Orchidaceae) in China. Biochem Genet 52(3–4):127–136

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133(3):693–709

Garcia-Vallve S, Palau J, Romeu A (1999) Horizontal gene transfer in glycosyl hydrolases inferred from codon usage in *Escherichia coli* and *Bacillus subtilis*. Mol Biol Evol 16(9):1125–1134

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98

Hegde S, Hegde HV, Jalalpure SS, Peram MR, Pai SR, Roy S (2017a) Resolving identification issues of *Saraca asoca* from its adulterant and commercial samples using phytochemical markers. Pharmacognosy Mag 13:S266–S272

Hegde S, Pai SR, Roy S (2017b) Combination of DNA isolation and RP-HPLC analysis method for bark samples of *Saraca asoca* and its adulterant. 3 Biotech 7(3):208

Jensen TL, Stubbendeck RM, Barnes AL, Wichers KM, Keller BJ, Bailey CA (2012) Use of phylogenetics for the purpose of quality control in the green algae genus *Chlorella*. FASEB J 26:9833

Jordan GE, Piel WH (2008) PhyloWidget: web-based visualizations for the tree of life. Bioinformatics 24(14):1641–1642

Jorde PE, Ryman N (1995) Temporal allele frequency change and estimation of effective size in populations with overlapping generations. Genetics 139(2):1077–1090

Khare CP (2007) Indian medicinal plants: an illustrated dictionary. Springer, Heidelberg

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA 102(23):8369–8374

Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R et al (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Natl Acad Sci USA 106(44):18621–18626

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 70 for bigger data sets. Mol Biol Evol 33(7):1870–1874

Levin RA, Wagner WL, Hoch PC et al (2003) Family-level relationships of onagraceae based on chloroplast *rbcL* and *ndhF* data. Am J Bot 90(1):107–115

Mooi E, Sarstedt MA (2011) Concise guide to market research: cluster analysis. Springer, Heidelberg, pp 237–284

Nadkarni AK (1976) Indian meteria medica, vol 1. Popular Prakashan, Mumbai

Nei M (1987) Molecular evolutionary genetics. Columbia Univ Press, New York

Palmblad M, Deelder AM (2012) Molecular phylogenetics by direct comparison of tandem mass spectra. Rapid Commun Mass Spectrom 26(7):728–732

Patra A, Deya AK, Kundu AB, Purushothaman KK, Saraswathy A (1992) Shoreaphenol, a polyphenol from *Shorea robusta*. Phytochemistry 31(7):2561–2562

Peakall ROD, Smouse PE (2012) GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19):2537–2539

Pendkar SK, Hegde S, Nayak SU, Hegde H, Kholkute SD, Roy S (2016) Detection of adulteration by *Wedelia calendulacea* in *Eclipta alba* through ISSR and RAPD Markers. Planta Med Int Open 3(02):e43–e46

Perrier X, Jacquemoud-Collet, JP (2006) DARwin software. http://darwin.cirad.fr/

Rasmussen HN, Rasmussen OF, Andersen JK, Olsen JE (1994) Specific detection of pathogenic *Yersinia enterocolitica* by two-step PCR using hot-start and DMSO. Mol Cell Probes 8(2):99–108

Richards E, Reichardt M, Rogers S (1994) Preparation of genomic DNA from plant tissue. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (Eds), Curr Protoc Mol Biol, John Wiley & Sons, New York, USA

Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19(18):2496–2497

Sarin YK (1996) Illustrated Herbs of Ayurveda New Delhi, India. Council of Scientific and Industrial Research (CSIR) and Indian Council of Medical Research (ICMR), New Delhi

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7(4):465–474

Shukla S, Hegde S, Kumar A, Chaudhary G, Tewari SK, Upreti DK, Pal M (2018) Fatty acid composition and antibacterial potential of *Cassia tora* (leaves and stem) collected from different geographic areas of India. J Food Drug Anal 26(1):107–111

Singh S, Krishna THA, Kamalraj S, Kuriakose GC, Valayil JM, Jayabaskaran C (2015) Phytomedicinal importance of *Saraca asoca* (Ashoka): an exciting past, an emerging present and a promising future. Curr Sci 109(10):1790–1801

Smillie TJ, Khan IA (2010) A comprehensive approach to identifying and authenticating botanical products. Clin Pharmacol Ther 87(2):175–186

Tandon N, Yadav SS (2017) Contributions of Indian Council of Medical Research (ICMR) in the area of Medicinal plants/Traditional medicine. J Ethnopharmacol 197:39–45

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673–4680

Urumarudappa SKJ, Gogna N, Newmaster SG, Venkatarangaiah K, Subramanyam R, Saroja SG, Gudasalamani R, Dorai K, Ramanan US (2016) DNA barcoding and NMR spectroscopy-based assessment of species adulteration in the raw herbal trade of *Saraca asoca* (Roxb) Willd, an important medicinal plant. Int J Legal Med 130(6):1457–1470

Wan CY, Wilkins TA (1993) Spermidine facilitates PCR amplification of target DNA. PCR Methods Appl 3(3):208–210

Wang H, Khera P, Huang B, Yuan M, Katam R, Zhuang W, Harris-Shultz K, Moore KM, Culbreath AK, Zhang X, Varshney RK, Xie L, Guo B (2016) Analysis of genetic diversity and population structure of peanut cultivars and breeding lines from China, India and the US using simple sequence repeat markers. J Integr Plant Biol 58(5):452–465

Wolf JB, Brodie ED III, Cheverud JM, Moore AJ, Wade MJ (1998) Evolutionary consequences of indirect genetic effects. Trends Ecol Evol 13(2):64–69