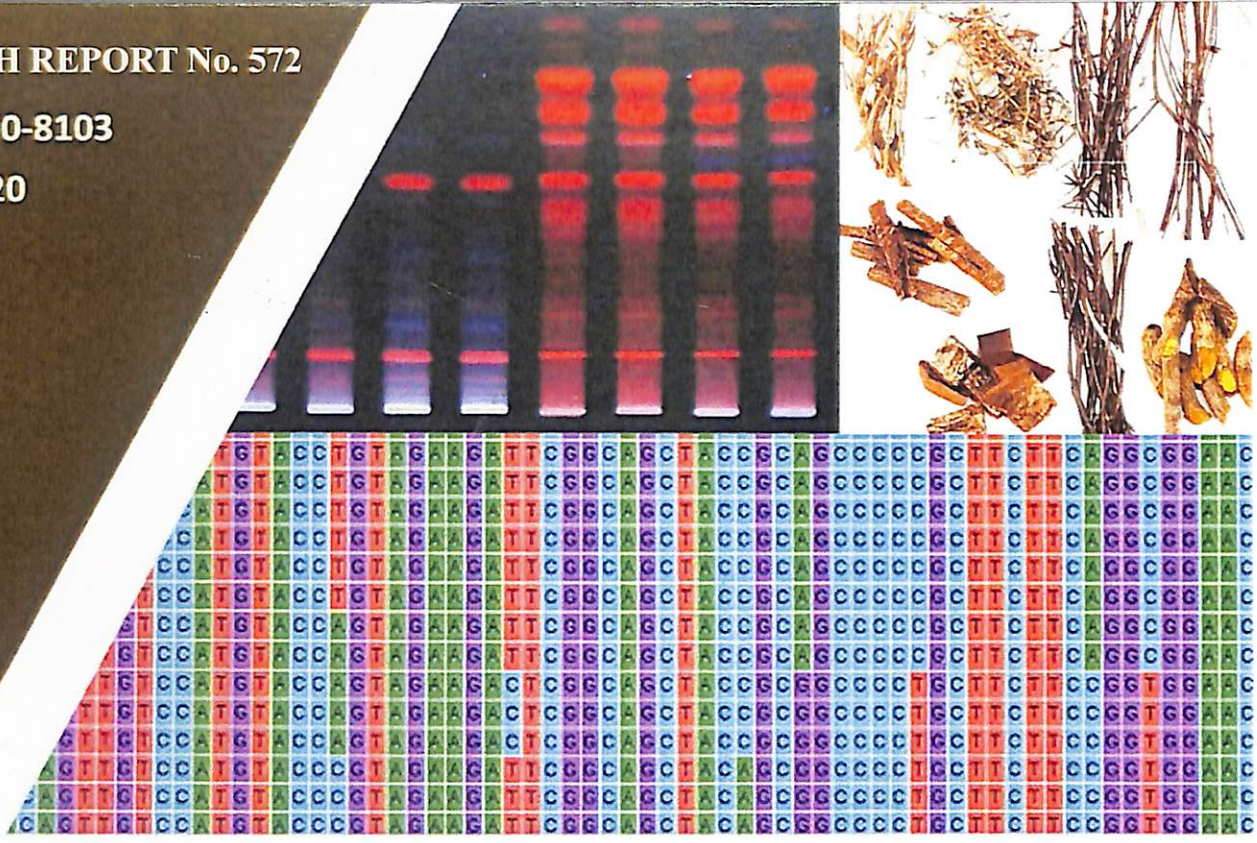
KFRI RESEARCH REPORT No. 572

ISSN 0970-8103

2020



AUTHENTICATION OF MAJOR COMMERCIALLY TRADED RAW DRUGS IN THE AYURVEDIC SYSTEMS OF MEDICINE IN INDIA

SUMA ARUN DEV

R JAYARAJ

P SUJANAPAL

V ANITHA

NATIONAL MEDICINAL PLANTS BOARD
MINISTRY OF AYUSH, NEW DELHI

KSCSTE- KERALA FOREST RESEARCH INSTITUTE
PEECHI, THRISSUR, KERALA - 680653, INDIA
www.kfri.res.in

KFRI RESEARCH REPORT NO. 572
(Final Report of KFRI RP 722/ 2016)

AUTHENTICATION OF MAJOR COMMERCIALLY TRADED RAW DRUGS IN THE AYURVEDIC SYSTEMS OF MEDICINE IN INDIA

INVESTIGATORS

Suma Arun Dev
Genetics & Tree Breeding Department
Forest Genetics and Biotechnology Division

Jayaraj R
Non Wood Forest Products Department
Forest Ecology & Biodiversity Conservation Division

Sujanapal P
Silviculture Department
Sustainable Forest Management Division

Anitha V
Forest Economics Department
Forestry & Human Dimensions Division

KSCSTE-KERALA FOREST RESEARCH INSTITUTE
Peechi 680653, Kerala, India

PROJECT DETAILS

1.	Project No.	KFRI RP 722/2016
2.	Title	Authentication of major commercially traded raw drugs in the ayurvedic systems of medicine in India
3.	Principal investigator	Dr. Suma Arun Dev, Senior Scientist, Forest Genetics & Biotechnology Division
4.	Associate investigators	Dr. Jayaraj R NWFP Department Forest Ecology & Biodiversity Division Dr. Sujanapal P Silviculture Department Sustainable Forest Management Division Dr. Anitha V Economics Department Forestry & Human Dimensions Division
5.	Research Fellow	Ms. Remya Unnikrishnan
6.	Objectives	1. DNA barcode development and HPTLC profiling of the selected ayurvedic raw drugs and its adulterants 2. Validation of the developed integrated approach for raw drug authentication
7.	Duration	4 Years
8.	Funding Agency	National Medicinal Plants Board (NMPB), Ministry of AYUSH, Government of India

ACKNOWLEDGEMENTS

The authors are thankful to Dr. Bransdon S. Corrie and Dr. S. Pradeep Kumar, former Directors as well as Dr. Syam Viswanath, Director, KSCSTE Kerala Forest Research Institute (KFRI) for providing the facilities to perform the project work and for their keen interests. The authors' record their gratitude for the financial support provided by National Medicinal Plants Board (NMPB), Ministry of AYUSH, Government of India, to carry out the research work at Kerala Forest Research Institute (KFRI). We are indebted to Kerala, Karnataka and Tamil Nadu Forest Departments for giving us permission to collect the required raw drug samples from the Protected areas and Reserve forests. The meticulous laboratory and field works carried out by Ms. Remya Unnikrishnan during the tenure of the project is highly appreciated.

CONTENTS

		PAGE NO.
	ABSTRACT	1
1	**INTRODUCTION**	2
2	**MATERIALS AND METHODS**	6
	2.1 Collection of authenticated reference material	6
	2.2 DNA Extraction & Polymerase Chain Reaction of DNA barcode gene regions	12
	2.3 HPTLC analysis	13
	2.4 Sequence data analysis	15
	2.5 Machine Learning Algorithm (MLA) based analysis	16
3	**RESULTS**	17
	3.1 *Coscinium fenestratum* and its adulterants	18
	3.1.1 MLA based analysis of *Coscinium fenestratum* and its adulterants	20
	3.1.2 HPTLC analysis of *Coscinium fenestratum* and its adulterants	21
	3.2 *Desmodium gangeticum* and its adulterants	23
	3.2.1 MLA based analysis of *Desmodium gangeticum* and its adulterants	25
	3.2.2 HPTLC analysis of *Desmodium gangeticum* and its adulterants	26
	3.3 *Saraca asoca* and its adulterants	27
	3.3.1 MLA based analysis of *Saraca asoca* and its adulterants	29
	3.3.2 HPTLC analysis of *Saraca asoca* and its adulterants	30
	3.4 *Terminalia arjuna* and its adulterants	32
	3.4.1 HPTLC analysis of *Terminalia arjuna* and its adulterants	33
	3.5 *Sida alnifolia* and its adulterants	34
	3.5.1 MLA based analysis of *Sida alnifolia* and its adulterants	36
	3.5.2 HPTLC analysis of *Sida alnifolia* and its adulterants	37
4	**DISCUSSION**	39
5	**CONCLUSIONS & FUTURE RECOMMENDATIONS**	43
6	**REFERENCES**	44
	APPENDIX 1	52

ABSTRACT

Huge demand for medicinal plants in India has exerted a heavy strain on the existing natural resources, leading to depletion of highly traded ayurvedic plants. Alongside, adulteration of expensive raw drugs with inferior taxa compromised the quality and safety of herbal products. Therefore, it is imperative to bring forth universally acceptable standard tools to authenticate ayurvedic raw drugs. In this regard, the study addresses the development of an integrated approach involving DNA barcode and High Performance Thin Layer Chromatography (HPTLC) fingerprinting to authenticate selected commercially traded ayurvedic raw drugs (*viz. Saraca asoca* (Roxb.) de Wilde, *Terminalia arjuna* (Roxb. ex DC.) Wight & Arn., *Sida alnifolia* L., *Desmodium gangeticum* (L.) DC. and *Coscinium fenestratum* (Gaertn.) Colebr.) from its adulterants. CBOL recommended DNA barcode gene regions *viz.* nuclear ribosomal-Internal Transcribed Spacer (nrDNA-*ITS*), maturase K (*mat*K), ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbc*L) and *psb*A-*trn*H spacer regions along with HPTLC profiling were experimented for the purpose. Even though, DNA barcode region, *ITS* showed promising results along with other barcode gene regions in *D. gangeticum*, *T. arjuna*, *S. asoca* and *psb*A-*trn*H barcode in *C. fenestratum*, *S. alnifolia*, high number of indels along with huge interspecific variation limited their utility for authentication. Consequently, *rbc*L and *mat*K barcode sequence database which was discriminant enough to identify adulterants were selected to validate the traded raw drugs. HPTLC analysis depicted the quality profile that distinguished original raw drugs from adulterants, though showed profile variations among accessions of species. Further, an integrated analytical approach employing Maximum Likelihood phylogenetic tree and Waikato Environment for Knowledge Analysis (WEKA) were employed to prove efficacy of DNA barcode method. The automated species identification technique, WEKA provided a large platform for rapid and precise authentication analysis of raw drug samples. Along with the recommended organoleptic and analytical methods, an integrated approach involving a DNA barcode tool along with HPTLC fingerprinting can strengthen the existing practice of quality checking and authentication of ayurvedic raw drugs by any of the certification agencies.

ABSTRACT

Huge demand for medicinal plants in India has exerted a heavy strain on the existing natural resources, leading to depletion of highly traded ayurvedic plants. Alongside, adulteration of expensive raw drugs with inferior taxa compromised the quality and safety of herbal products. Therefore, it is imperative to bring forth universally acceptable standard tools to authenticate ayurvedic raw drugs. In this regard, the study addresses the development of an integrated approach involving DNA barcode and High Performance Thin Layer Chromatography (HPTLC) fingerprinting to authenticate selected commercially traded ayurvedic raw drugs (*viz. Saraca asoca* (Roxb.) de Wilde, *Terminalia arjuna* (Roxb. ex DC.) Wight & Arn., *Sida alnifolia* L., *Desmodium gangeticum* (L.) DC. and *Coscinium fenestratum* (Gaertn.) Colebr.) from its adulterants. CBOL recommended DNA barcode gene regions *viz.* nuclear ribosomal-Internal Transcribed Spacer (nrDNA-*ITS*), maturase K (*mat*K), ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbc*L) and *psb*A-*trn*H spacer regions along with HPTLC profiling were experimented for the purpose. Even though, DNA barcode region, *ITS* showed promising results along with other barcode gene regions in *D. gangeticum*, *T. arjuna*, *S. asoca* and *psb*A-*trn*H barcode in *C. fenestratum*, *S. alnifolia*, high number of indels along with huge interspecific variation limited their utility for authentication. Consequently, *rbc*L and *mat*K barcode sequence database which was discriminant enough to identify adulterants were selected to validate the traded raw drugs. HPTLC analysis depicted the quality profile that distinguished original raw drugs from adulterants, though showed profile variations among accessions of species. Further, an integrated analytical approach employing Maximum Likelihood phylogenetic tree and Waikato Environment for Knowledge Analysis (WEKA) were employed to prove efficacy of DNA barcode method. The automated species identification technique, WEKA provided a large platform for rapid and precise authentication analysis of raw drug samples. Along with the recommended organoleptic and analytical methods, an integrated approach involving a DNA barcode tool along with HPTLC fingerprinting can strengthen the existing practice of quality checking and authentication of ayurvedic raw drugs by any of the certification agencies.

1. INTRODUCTION

India is renowned for its traditional medicine systems like Ayurveda, Siddha and Unani and is the largest producer of medicinal herbs in the world (Ganesan et al. 2016). As per the recent report of International Union for Conservation of Nature(IUCN) and World Wildlife Fund for Nature (WWF), globally 50,000 to 80,000 flowering plant species are being utilized for various medicinal purposes (Chen et al. 2016). Ethno-biological Survey of Ministry of Environment and Forests and Climate Change (MOEF & CC), Government of India could identify 8000 plant species utilized in various systems of medicine with approximately 25,000 effective herbal formulations. This huge demand for herbal drugs invariably leads to the upsurge of herbal industries in India. Commercial demand of herbal drugs has been recently valued at INR 300 billion in 2018, which is more than double as compared to preceding years (CISION PR News Wire 2020).

Although the use of herbal medicines has become commercialised, majority of plants are still harvested from the wild. The growing demand exerts a heavy pressure on the available natural resources (Schippmann et al. 2002). Consequently, over exploitation, unscientific extraction and resource limitation has instigated a latest trend to adulterate the potential raw drugs (Chen et al. 2010). Recently, raw drug adulteration has become a burning problem in herbal industries wherein the quality of formulations is compromised with look-alike plants of inferior properties (Mishra et al. 2016; Ouarghidi et al. 2013; Techen et al. 2014; Walker et al. 2012). The percentage of adulterated products varies significantly among countries, the highest percentage herbal product adulteration was reported form Brazil, followed by Taiwan and India (Ichim 2019).

WHO pharmacopeia has implemented certain criteria for proper identification of plant species and quality assessment using potent chemical markers to guarantee the quality of herbal medicines (WHO 2011). Even though there are several recommended methods, right from the traditional taxonomic, chemical/biochemical and organoleptic, it is extremely difficult to ensure authenticity of ayurvedic raw drugs in many instances (Coghlan et al. 2012). The

traditional taxonomic method alone cannot address the adulteration, due to the lack of expertise and limitations of morphological identification of species owing to phenotypic plasticity and cryptic species (Hebert 2003). Since raw drugs are available in extremely dried, shredded or powdered form, species identification using traditional taxonomic as well as organoleptic, macroscopic and microscopic means may not always be possible. In many instances, the quality and quantity estimation of potent herbal medicines are performed using chemical finger printing methods such as TLC (Thin-layer chromatography), HPTLC (High-performance thin layer chromatography), HPLC (High Performance Liquid Chromatography), GC-MS (Gas chromatography–mass spectrometry), NMR (Nuclear magnetic resonance spectroscopy) and FTIR (Fourier-transform infrared spectroscopy) (Mukherjee et al. 2010; Smillie and Khan 2010; Meena Devi et al. 2010). However, chemical fingerprints are influenced by external environmental factors such as age of the plant and storage conditions as well as type of plant parts used (Liu et al. 2011; Kaur et al. 2016).

During early 1990s, researchers shifted their focus to DNA based markers for molecular species identification (Sucher and Carles 2008; Hao et al. 2010). A wide range of molecular markers such as AFLP (Amplified fragment length polymorphism), RFLP (Restriction fragment length polymorphism), RAPD (Randomly Amplified Polymorphic DNA), ISSR (Inter simple sequence repeat), SSR (Simple sequence repeat), SCAR (Sequence characterized amplified region), LAMP (Loop mediated isothermal amplification) and SNPs (Single nucleotide polymorphisms) were extensively used for identification of plants (Sharma et al. 2008; Diao et al. 2009; Tamhankar et al. 2009; Cao et al. 2010). DNA markers gained popularity for its accuracy and proper identification as it remained unaffected by physiological or environmental factors (Ganie et al. 2015). Each DNA marker has its own advantages and limitations and none can be considered as ideal. Selection of markers is purely based on the nature of research, financial stability and technical expertise (Kiran et al. 2010). Thus, there is an ensuing demand for a universally acceptable standard molecular tool to authenticate herbal medicine.

In this regard, DNA barcoding offers a novel prospective tool for taxonomists and has greatly transformed species identification process (Hebert et al. 2003). Short DNA sequences from the conserved gene regions of a genome termed as barcode, is an accurate and reliable alternative to morphological identification of biological material in challenging situations (Hebert et al. 2003; Chen et al. 2010). It can overcome inherent problems associated with traditional taxonomic identification due to phenotypic plasticity, species complexity, difficulty in finding reliable characters due to long maturity period, among others (Kress et al. 2005). *mat*K is one of the most rapidly evolving coding region of the plastid genome whereas *rbc*L is a highly conserved plastid region, both have been widely used in numerous plant taxa. Though barcode regions such as psbA-*trn*H and *ITS* were successful in several instances, single tier approach was not encouraged owing to huge nucleotide variation and inability to resolve the species completely in some instances (Doebley et al. 1990; CBOL 2009). Successively, in 2009, CBOL recommended two locus combination of *mat*K and *rbc*L as a core barcode for plants. Other combinations such as *rbcL, psb*A-*trn*H and *ITS* have also been evaluated for their potential. Combination of *rbc*L + *ITS* region that enables higher retrieval capacity is widely used for identification of herbal products (Newmaster et al. 2006; Hollingsworth et al. 2009; Burgess et al. 2011; Krawczyk et al. 2013; Newmaster et al. 2013; Malik et al. 2018).

Integrated approach

Currently, quality control and safety analysis of herbal drugs and mixtures are progressing in a comprehensive and integrated direction. Integrated method is a combination of two or more diverse techniques which are capable of authenticating a species more precisely. Recently, it is shown that in addition to adulteration, absence of potent principles in the processed products could also be a serious threat to Ayurveda or traditional system of medicine (Palhares et al. 2015). The efficiency of integrated approach involving DNA barcoding and HPTLC was demonstrated in the analysis with 257 samples derived from 8 species recommended by WHO (Palhares et al. 2014). A combination of DNA barcoding and NMR was also performed in species adulteration of *Garcinia* species and *Saraca asoca* (Kumar et al. 2016;

Seethapathy et al. 2018). For conservation and proper use of Brazilian quinas, a comprehensive system of chemical, biological and molecular methods has been used (Palhares et al. 2015). Comprehensive approach of DNA metabarcoding, TLC and HPLC-MS was carried out for the detection of substitution/adulteration of *Hypericum perforatum* (Raclariu et al. 2017). Authentication of *Marsdenia*, was also successfully done by using a multi-tier approach of DNA barcoding coupled with TLC and HPLC (Yu et al. 2018). Integrated approach would be a future promising tool for accurate and reliable qualitative/quantitative authentication of medicinal plants and products.

In this regard the present study was carried out with the following objectives

1. DNA barcode development and HPTLC profiling of the selected ayurvedic raw drugs and its adulterants
2. Validation of the developed integrated approach for raw drug authentication

2. MATERIALS AND METHODS

2.1. Collection of authenticated biological reference material

The highly traded ayurvedic raw drugs and its market adulterants were identified through a preliminary survey as part of an earlier study. Mature stem, leaf and flower samples of *Saraca asoca, Terminalia arjuna (cuneata), Sida alnifolia, Desmodium gangeticum* and *Coscinium fenestratum* and its market adulterants were collected from different geographic locations of its distribution zones in south India and *Berberis aristata* from north India for development of the DNA barcode database (Fig.1). Medicinally important plant parts (mature stem, bark and root) were collected from the respective original raw drug species and its adulterants for HPTLC analysis (Fig. 2). GPS coordinates of the locations and details of the collected plant samples are provided in Table 1. Multiple accessions were collected and specimens were stored in silica gel for further use. The voucher specimens were deposited in the KFRI herbarium.

Table 1. Species collected from multiple locations

Species	Location	GPS coordinates
Original Species		
Coscinium fenestratum	Thrissur	10°31'49.64"N 76°20'48.87"E
	Boys Town	11°50'24.79"N 75°55'07.54" E
	Aralam	11°55'20.38"N 75°47'33.16" E
Adulterants		
Anamirta cocculus	KFRI, Peechi	10°31'49.64"N 76°20'48.87" E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Rose mala	85°54'54.02"N 77°10'12.41" E
Morinda pubescens	Olavakode	10°47'58.29"N 76°38'33.91" E
	Vellanimala	10°32'58.08"N 76°20'08.81" E
	Adoor	9°05'28.38"N 76°51'51.49" E
Diploclisia galucescnes	Nilambur	11°17'07.93"N 76°14'18.89" E

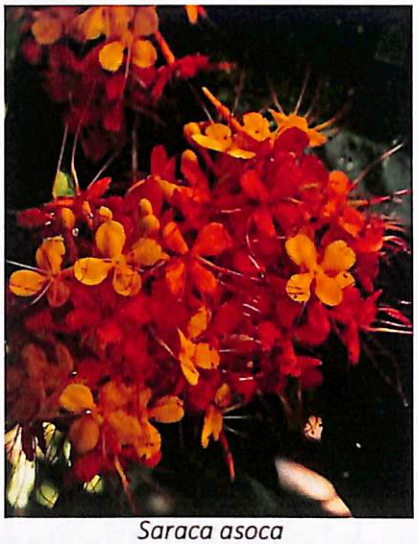
	Periyar Tiger Reserve	9°35'16.45"N 77°10'48.26" E
	Silent valley	11°03'50.86"N 76°32'16.14" E
Berberis aristata	Jammu	33°46'41.43"N 76°34'34.22" E
	Himachal Pradesh	31°07'09.56"N 77°08'22.28" E
	Uttarakahand	30°03'44.29"N 79°00'08.07" E
Original Species		
Sida alnifolia	Nilambur	11°17'07.93"N 76°14'18.89" E
	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
Adulterants		
Sida acuta	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Paramabikulam	10°23'32.55"N 76°46'26.08" E
Sida cordifolia	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Paramabikulam	10°23'32.55"N 76°46'26.08" E
Sida rhombifolia	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Nilambur	11°17'07.93"N 76°14'18.89" E
Sida rhomboidea	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Paramabikulam	10°23'32.55"N 76°46'26.08" E
Urena lobata	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Aralam	11°55'20.38"N 75°47'33.16" E
Orginal Species		
Saraca asoca	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Thrissur	10°31'39.51"N 75°12'51.97" E
	Palakkad	10°47'12.23"N 75°39'17.26" E
Adulterants		

Shorea roxburghii	Parambikulam	10°23'32.55"N 76°46'26.08" E
	Thoonakadavu	10°26'03.00"N 76°46'54.00" E
	Palakkad	10°47'12.23"N 75°39'17.26" E
Polyalthia longifolia	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Palakkad	10°47'12.23"N 75°39'17.26" E
Polyialthia coffeoides	Thamarassery	11°29'54.08"N 76°01'19.13" E
	Thrissur	10°31'39.51"N 75°12'51.97" E
	Wayanad	11°29'54.08"N 76°01'19.13" E
Original Species		
Terminalia arjuna (cuneata)	Thayannamkudi	10°18'24.47"N 75°12'21.69" E
	Chinnar	10°18'24.47"N 75°12'21.69" E
	KFRI Nursery, Peechi	10°31'49.64"N 76°20'48.87"E
Adulterants		
Lagerstroemia microcarpa	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Palakkad	10°47'12.23"N 75°39'17.26" E
	Thrissur	10°31'39.51"N 75°12'51.97" E
Lagerstroemia speciosa	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Periyar Tiger reserve	9°35'16.45"N 77°10'48.26" E
	Thrissur	10°31'39.51"N 75°12'51.97" E
Original Species		
Desmodium gangeticum	KFRI	10°31'49.64"N 76°20'48.87"E
	KFRI nursery	10°31'49.64"N 76°20'48.87"E
	Nilambur	11°17'07.93"N 76°14'18.89" E
Adulterants		
Desmodium pulchellum	Aralam	11°55'20.38"N 75°47'33.16" E
	Aralam farm	11°56'28.46"N 75°44'00.80" E
	Palakkad	10°47'12.23"N 75°39'17.26" E
Desmodium triangulare	Aralam	11°55'20.38"N 75°47'33.16" E
	Aralam farm	11°56'28.46"N 75°44'00.80" E

	Silent valley	11°03'50.86"N 76°32'16.14" E
Desmodium triquetrum	Aralam	11°55'20.38"N 75°47'33.16" E
	Aralam farm	11°56'28.46"N 75°44'00.80" E
	Athirapilly	10°16'27.88"N 76°30'56.10" E
Desmodium vellutinum	KFRI, Peechi	10°31'49.64"N 76°20'48.87"E
	Aralam	11°55'20.38"N 75°47'33.16" E
	Athirapilly	10°16'27.88"N 76°30'56.10" E

Coscinium fenestratum

Desmodium gangeticum



Saraca asoca

Terminalia cuneata

Sida alnifolia

Figure 1. Five selected species of ayurvedic raw drugs

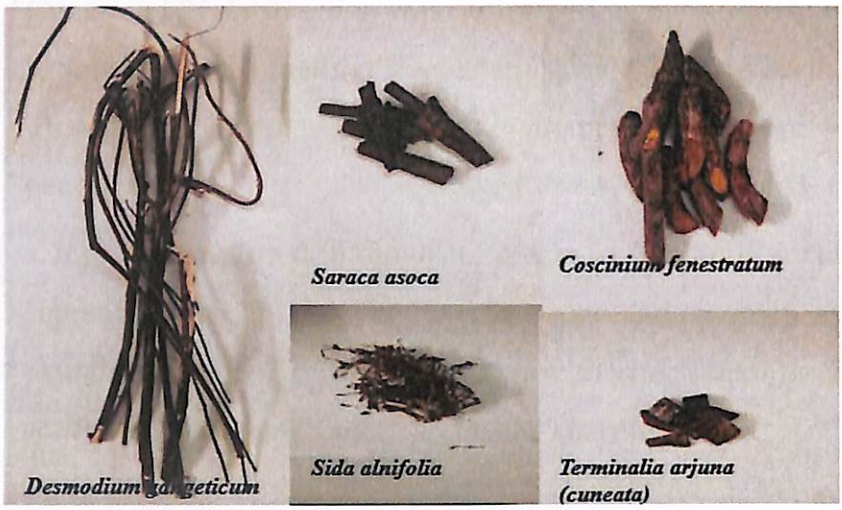


Figure 2. Medicinally important parts of selected Biological Reference Material (BRM)

2.2. Collection of traded samples

Traded samples were collected from the selected authorised dealers of ayurvedic raw drugs and major ayurvedic industries in south India (Table 2). About 100 g of each raw drugs (available in extremely dried and shredded form) was purchased from shops to check adulteration in the traded samples. Each collected raw drug sample was given Herbal Authentication Service Code (HAS) with details of location. To avoid the chances of mixing up, strict attention was followed from collection to final data analysis. Most of the procured raw drugs had not retained any morphological features of the original plant species. Raw drugs of all the selected species were able to purchase from herbal markets except *Terminalia arjuna (cuneata).*

Table 2. **List of collected market samples**

Raw drug samples	Sample Id
Coscinium fenestratum	CF1,CF2,CF3,CF4,CF5
Saraca asoca	SA1,SA2,SA3,SA4,SA5
Desmodium gangeticum	DG1,DG2,DG3,DG4,DG5
Sida alnifolia	S1,S2,S3,S4,S5

2.3. DNA Extraction

Genomic DNA extraction was performed using modified cetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987) as well as DNeasy Plant Mini Kit for difficult samples according to manufacturer's protocol (Qiagen, USA). Total genomic DNA was also extracted from the collected raw drug samples. DNA samples were stored in the deep freezer at 20^0C until further use. The samples were separated on 1.5 per cent agarose gel and stained in ethidium bromide and visualised under UV transilluminator to check quality. It was quantified using a spectrophotometer (Nanodrop Fisher Thermo., USA).

2.4. Polymerase Chain Reaction of DNA barcode gene regions

The taxonomically authenticated Biological Reference Materials (BRM) were used to develop species specific barcodes for four standard barcode gene regions (*rbc*L, *mat*K, *ITS* and *psb*A-*trn*H). The barcode regions, primer sequences and PCR conditions used are provided in Table 3.

Working concentration of genomic DNA was prepared by diluting the stock solution at a concentration 25 ng/ul. 25 µL of PCR reaction mixture comprised of 2.5 µL PCR buffer at 1X (supplied with10X concentration), 1 µL each of forward and reverse primers (5 pmol), 2.5 µL of dNTPs from 10 mM stock, 2 U/25 µL of Taq-polymerase, 1 µL template DNA with the concentration of 25 ng/µL and the final volume of the PCR reaction mixture was made upto 25 µL with sterile distilled water. PCR reaction was performed with the following conditions, initial denaturation of 5 minutes at 94°C, cycle denaturation of 1 minute at 94°C, cycle annealing of 1 min at 60°C and cycle extension of 1 min at 72°C for 35 cycles and a final extension at 72°C for 10 minutes. PCR products were resolved by 2per cent agarose. Electrophoresis was performed on agarose gel by applying constant voltage to resolve the products and documented with Alpha Imager (Alpha Innotech, USA).

Table 3. Details on primers and PCR reaction conditions

Barcode loci	Primer name	Sequence 5'-3'	Primer annealing temperature
ITS	*ITS 1*	TCCGTAGGTGAACCTGCGG	60 ° C
	ITS 2	TCCTCCGCTTATTGATATGC	
psbA-trnH	*psbA*	GTWATGCAYGAACGTAATGCTC	58 ° C
	trnH	CGCGCATGGTGGATTCACAATCC	
matK	*matk* 427 F	CCCRTYCATCTGGAAATCTTGGTT	50 ° C
	matK 1248 R	GCTRTRATAATGAGAAAGATTTCTGC	
rbcL	*rbcL* 1 F	ATGTCACCACAAACAGAAAC	60 ° C
	rbcL 724 R	TCGCATGTACCTGCAGTAGC	

2.5. Elution of PCR products

PCR reaction was scaled up to 50 µL volume for the purpose of elution. Elution of the PCR product was done by Nucleospin gel and PCR clean up kit as per the manufacturer's protocol (Machery-Nagel, U.S.A.). DNA sequencing was performed for the eluted PCR products in both forward and reverse directions employing Sanger's dideoxy method.

2.6. HPTLC analysis

CAMAG Linomat 5 with twin plate chamber and CAMAG TLC scanner instrument programmed through Win CATS software was used for HPTLC finger printing. Medicinally important parts of the selected raw drug species and its adulterants with two individuals each were finely powdered. Ten gram of powder accurately weighted from each sample was extracted using solvents. Methanol extraction was carried out in *Coscinium fenestratum, Sida alnifolia, Sarac asoca, Desmodium gangeticum*

and its adulterants whereas chloroform extraction was performed for *Terminalia arjuna*. Extracts were filtered and concentrated under reduced pressure and made upto 10 ml in standard flasks separately. Stationary phase of aluminium TLC plates pre-coated with Silica gel 60 F_{254} of 0.2 mm thickness and mobile phase for each set was standardised (Table 4). The plate was derivatized with anisaldehyde sulphuric acid reagent for band visualization. Chemical profile of each sample was analysed according to their RF values (Retention factor). Dendrogram was constructed by SPSS v.16.0 (SPSS Inc 2007) using nearest neighbour adopting euclidean distances, which revealed the relation between each species according to their phytochemical constituents.

Table 4. HPTLC sample preparation and mobile phase

Sample set	**Medicinally important part**	**Solvent System**	**Mobile phase**
Coscinium fenestratum *Anamirta cocculus* *Morinda pubescens* *Diploclisia galucescnes* *Berberis aristata*	Stem	Methanol	Toluene:Ethylacetate: Aceticacid
Sida alnifolia *Sida rhombifolia* *Sida rhomboidea* *Sida cordifolia* *Sida acuta* *Urena lobata*	Whole plant	Methanol	Toluene:Ethylacetate: Methanol:Formicacid
Terminalia arjuna *Lagerstroemia microcarpa* *Lagerstroemia speciosa*	Bark	Chloroform	Toluene:Ethylacetate: Aceticacid
Saraca asoca *Polyalthia longifolia* *Polyalthia coffeoides*	Bark	Methanol	Toluene:Ethylacetate: Formicacid

Shorea roxburghii			
Desmodium gangeticum *Desmodium pullchellum* *Desmodium triangulare* *Desmodium triquetrum* *Desmodium velutinum*	Whole plant	Methanol	Toluene:Ethylacetate: Methanol

2.7. Sequence data analysis

Raw chromatograms were edited and trimmed using BioEdit software (Hall 1999). The edited sequences were aligned using Clustal W (Thompson et al. 1994). Homology searches were performed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to confirm the identity of the sequences.

For pair-wise genetic distance (PWG) method, genetic pair-wise distance (interspecific as well as intraspecific distances) was determined by MEGA v.6.0 using Kimura two-parameter distance model (K2P) adopting complete deletion option (Tamura et al. 2013). The interspecific divergence between species was calculated using three parameters; (i) average interspecific distance, (ii) average theta prime (θ') and (iii) minimum interspecific distances. Intraspecific parameters; (iv) average intraspecific distance, (v) theta (θ) and (vi) coalescent depth were also calculated to characterize intraspecific divergences (Meyer and Paulay 2005). DNA barcoding gaps were calculated by comparing intra and interspecific genetic distances (Meyer and Paulay 2005; Meier et al., 2006). Significance of barcoding gap was assessed using Wilcoxon matched pairs signed rank test in SPSS v.16.0 (SPSS Inc 2007). Dendrogram was constructed using developed barcode sequences for authentication of traded samples with BRM, with 1000 bootstrap using MEGAv.0.7 adopting Kimura 2 model (Kumar et al. 2016).

2.8. Machine Learning Algorithm (MLA) based analysis

In MLA, DNA barcoding analysis was performed with a reference data set composed of DNA sequences of known species (BRM) and query data set with the sequence of unknown species (market samples). In the adopted algorithm namely, Waikato Environment for Knowledge Analysis (WEKA), the function-based method Support Vector Machines (SMO) (Suykens and Vandewalle 1999), the rule-based RIPPER (Jrip) (Shahzad et al. 2013), the decision tree C4.5 (J48) (Quinlan 1996) and the Bayesian-based method Naive Bayes (Lewis 1998) were tested on DNA barcodes with 10-fold cross validation. The ".fasta" files of barcode sequences were converted to ".arff" format using "Fasta2Weka" programme for analysis in WEKA (Weitschek et al. 2014). All four classification methods in WEKA were run with four barcode primer sequences of BRM. Best classifier was selected according to their efficiency in species discrimination. Using the best classifier, sequences from traded market samples were further analysed along with BRM sequence database.

3. RESULTS

Good quality genomic DNA was obtained from original plant species as well as most of the raw drug samples (Fig. 3). All four DNA barcode regions (*rbc*L, *mat*K, *psb*A-*trn*H and *ITS*) were successfully amplified from original plants (Fig. 4) whereas impurities hindered primer annealing and subsequent PCR amplification of DNA from raw drugs in some instances. Sequence length and basic sequence statistics like conserved sites, variable sites and singletons were based on the results of CLUSTAL X alignment as well as with alignment explorer in MEGA v.6.0 (Fig. 3). The sequences after alignment were subjected to BLAST sequence similarity search in NCBI GenBank (http://www.ncbi.nlm.nih.gov/genbank/). Interspecific divergence, intraspecific variation and DNA barcoding gap were used to identify the potential barcode region and the DNA barcode database was employed for authentication of traded ayurvedic raw drug samples.

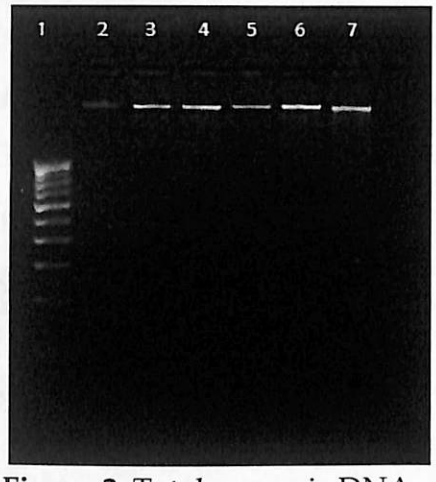


Figure 3. Total genomic DNA

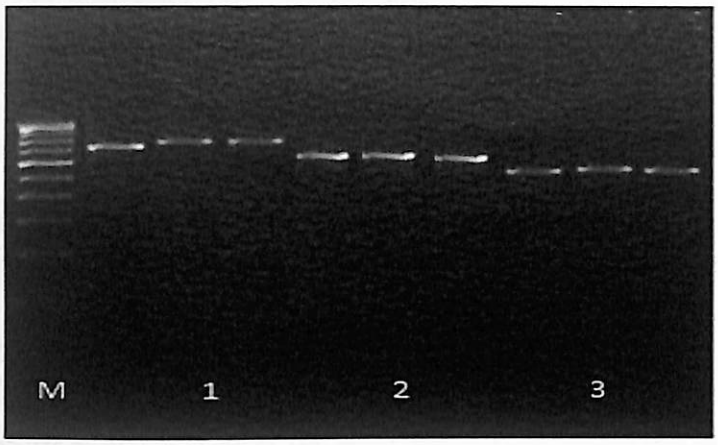


Figure 4. PCR product of *rbc*L, *mat*K, *trn*H-*psb*A barcode regions (1. *Coscinium fenestratum*,2. *Diploclisia glaucescens*, *3. Anamirta cocculus*) (M-100bp ladder)

3.1. *Coscinium fenestratum* and its adulterants

All the analysed barcode regions (*ITS*, *psb*A-*trn*H, *mat*K and *rbc*L) amplified successfully in most samples. Among these barcode regions, *psb*A-*trn*H spacer region showed highest nucleotide variation (460/648), followed by *rbc*L (212/590), *mat*K (200/780) and *ITS* (166/423) regions, respectively (Table 5). Intra and Inter specific genetic divergences analysis of the four barcode regions showed only interspecific divergence and did not show any intra specific divergence. Among these four barcode regions, *psb*A-*trn*H and *rbc*L showed highest inter specific divergences in *Coscinium fenestratum* and its adulterants (Table 6; Fig. 5). The barcode regions *viz. ITS*, *psb*A-*trn*H, *mat*K and *rbc*L showed distinct barcode gap of 0.0276, 0.175, 0.0234 and 0.0632 respectively. Wilcoxon's signed rank test performed to test the significance of interspecific divergence in barcode regions (*psb*A-*trn*H, *mat*K, *rbc*L, and *ITS*) showed significant values. Based on barcode gap analysis, *psb*A-*trn*H and *rbc*L gene regions can be considered as potential barcodes to authenticate *Coscinium fenestratum* from its adulterants.

Further, *rbc*L barcode region alone was adopted for authentication of traded samples successfully owing to the difficulties in PCR amplification of other barcode gene regions from degraded DNA of traded raw drugs. The barcode sequences developed from the traded samples clustered into separate clades corresponding to

the respective sequences of BRM samples (Fig. 6). The phylogenetic tree generated based on these sequences showed a clear clustering of traded samples with those of BRM. Thus, most of the traded *Coscinium fenestratum* samples were clustered along with *B. aristata* and remaining samples showed similarity with *Coscinium fenestratum*

Table 5. Basic sequence statistics of *Coscinium fenestratum* and its adulterants

Comparison	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Sequence length	590	780	648	423
Conserved sites	378	580	188	257
Variable sites	212	200	460	166
Informative site	375	287	455	166
Singleton site	3	0	15	0

Table 6. Genetic divergence parameters of *Coscinium fenestratum* and its adulterants

Parameters	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Average intraspecific distance	0	0	0	0
Average interspecific distance	0.0632±0.0015	0.0234±0.0064	0.175±0.0880	0.0276±0.0064

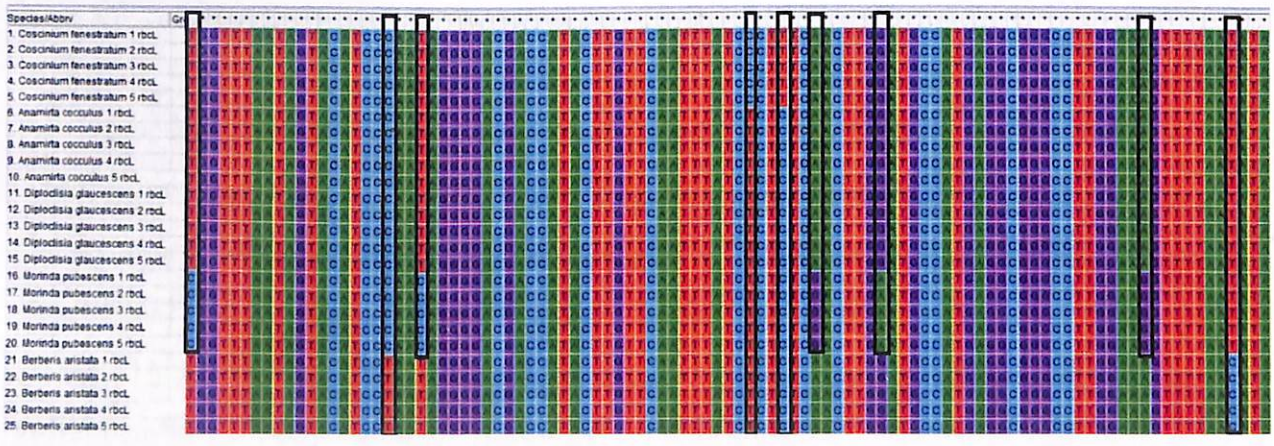


Figure 5. Multiple sequence alignment showing single nucleotide polymorphisms (SNPs) in *rbc*L sequences of Biological Reference Material (BRM) of *Coscinium fenestratum* and its adulterants

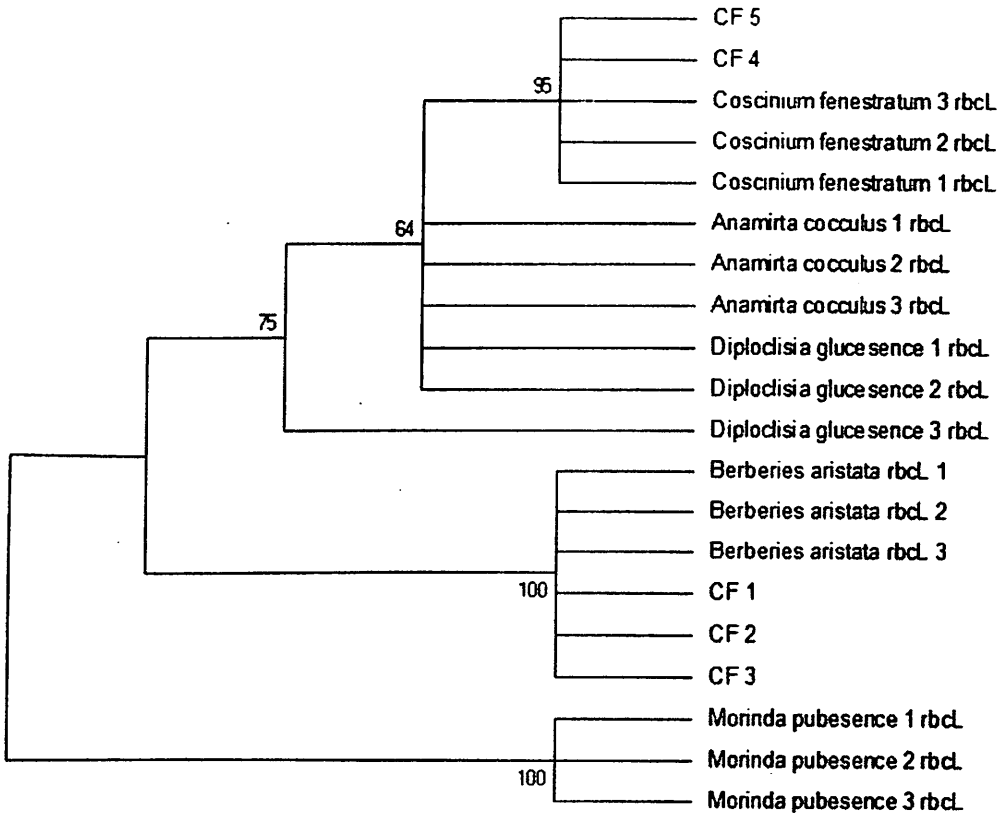


Figure 6. Maximum Likelihood tree (ML) of market samples along with BRM using *rbc*L barcode

3.1.1 MLA based analysis of *Coscinium fenestratum* and its adulterants

All four classification methods were run in WEKA with 10-fold cross validation. Among four machine learning algorithms, Naive bayes and JRip failed to identify the sequences of reference data set as well as the test sequence database. SMO and J48 showed species identification in BRM samples with 100 per cent discrimination power. These two machine learning algorithms were subsequently employed for authentication of sequences of unknown traded samples. When the test data of unknown traded samples with large variations were analysed, the J48 classifier could identify only 35 per cent of the species. Best performance was shown by SMO with 100 % accuracy in authenticating the market samples (Fig. 7). Traded samples

showed similarity with *B. aristata* and *C. fenestratum* provided in the reference data set, which again substantiated the dendrogram based sequence analysis in MEGA.

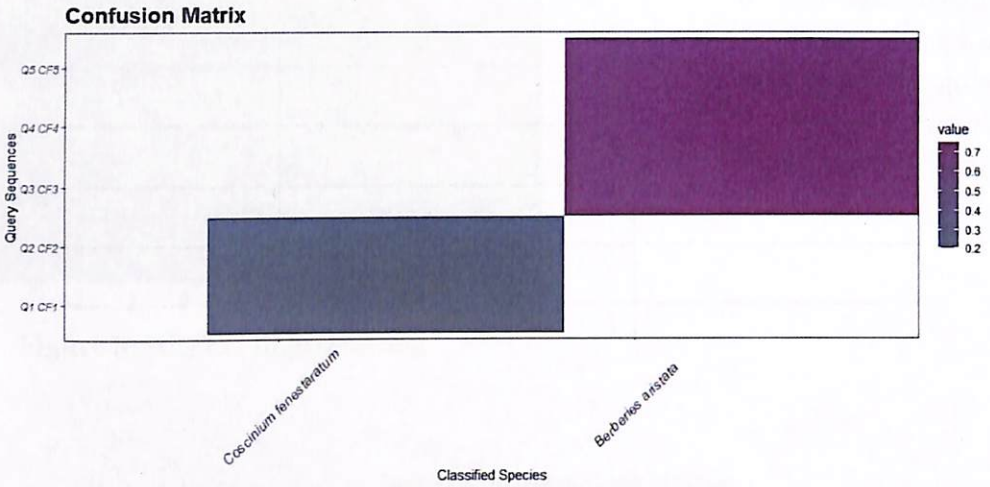
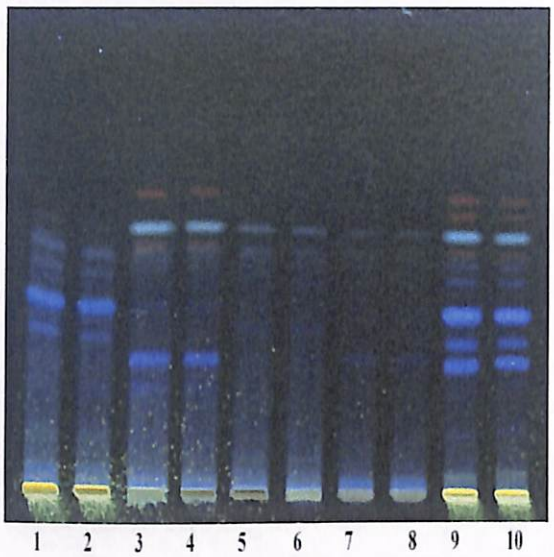


Figure 7. The confusion matrix showing the identification rate of market samples along with BRM based on *rbc*L barcode sequence

3.1.2. HPTLC analysis of *Coscinium fenestratum* and its adulterants

Chemical profile of each sample was analysed according to their Retention factor (RF) values (Fig. 8). Dendrogram generated using RF values was used to analyse the chemical fingerprinting of the selected raw drugs and its adulterants (Fig. 9). Each species showed specific banding pattern with some amount of intra specific variation in the banding pattern of minor bands amongst accessions belonging to different geographical locations and were grouped in different clades.



1 *Coscinium.fenestratum*
2 *Coscinium.fenestratum*
3 *Anamirta.cocculus*
4 *Anamirta.cocculus*
5 *Morinda. pubescens*
6 *Morinda. pubescens*
7 *Diploclisia. galucescnes*
8 *Diploclisia. galucescnes*
9 *Berberis aristata*
10 *Berberis aristata*

Figure 8. HPTLC fingerprinting

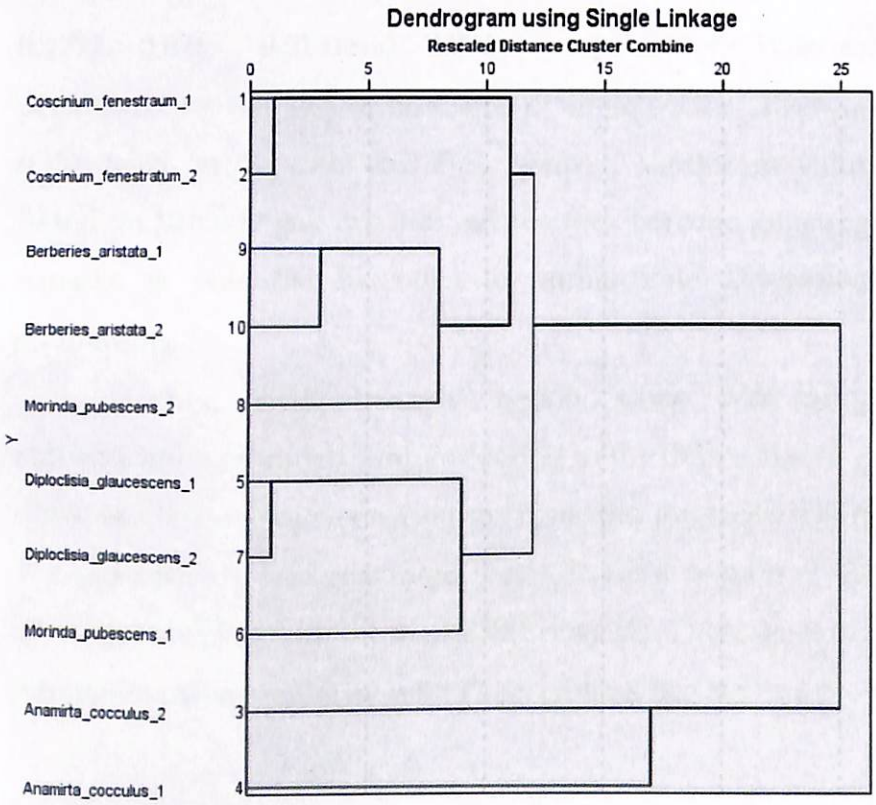


Figure 9. Dendrogram constructed using HPTLC banding pattern

3.2. *Desmodium gangeticum* and its adulterants

All the analysed barcode regions (*ITS*, *psb*A-*trn*H, *mat*K and *rbc*L) amplified successfully with 100 per cent PCR efficiency. Among these barcode regions, *mat*K region showed highest nucleotide variation (190/507), followed by *ITS* (184/289), *psb*A-*trn*H (82/369) and *rbc*L (31/632) regions, respectively (Table 7). Intra and Inter specific genetic divergences analysed from the four barcode regions showed interspecific as well as intra specific divergences. Among these four barcode regions, *ITS* showed highest inter specific divergence and *psb*A-*trn*H showed highest intraspecific variation in *Desmodium gangeticum* and its adulterants (Table 8; Fig. 10). Based on intra and inter specific distances, barcode gap was also estimated. The barcode regions *viz. ITS*, *psb*A-*trn*H, *mat*K and *rbc*L showed distinct barcode gap of 0.1722, 0.0265, 0.0151and 0.00463 respectively. Wilcoxon's signed rank test performed to test the significance of interspecific divergence in barcode regions (*psb*A-*trn*H, *mat*K, *rbc*L and *ITS*), showed significant values for all four regions. Based on barcode gap analysis, all the four barcode gene regions can be considered equally as potential barcodes to authenticate *Desmodium gangeticum* from its adulterants.

Further, *matK* barcode region alone was adopted successfully for authentication of traded samples owing to the difficulties in getting amplification for other barcode regions consistently from the degraded DNA of traded raw drugs. The phylogenetic tree generated based on these sequences showed a clear clustering of traded samples with those of BRM (Fig. 11). Thus, most of the traded *Desmodium* samples showed similarity with *D. pulchellum* instead of the original drug species.

Table 7. Basic sequence statistics of *Desmodium gangeticum* and its adulterants

Comparison	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Sequence length	632	507	369	289
Conserved sites	511	190	287	92
Variable sites	31	317	82	184
Informative site	29	297	82	158
Singleton site	2	19	0	26

Table 8. Genetic divergence parameters of *Desmodium gangeticum* and its adulterants

Parameters	*rbc*L	*mat*K	*psb*A-*trn*H	ITS
Average intraspecific distance	0	0	0.0013±0.0008	0
Average interspecific distance	0.00463±0.001	0.0151±0.0028	0.0278±0.0049	0.1722±0.1061
Average theta	0	0	0.003±0.0006	0
Average theta prime	0	0	0.002±0.003	0
Average coalescent depth	0	0	0.002±0.001	0

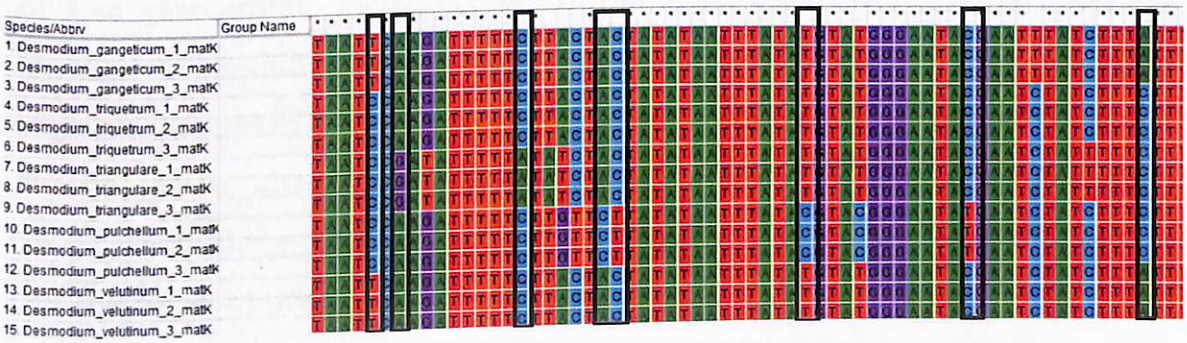


Figure 10. Multiple sequence alignment showing single nucleotide polymorphisms (SNPs) in *mat*K sequences of Biological Reference Material (BRM) of *Desmodium gangeticum* and its adulterants

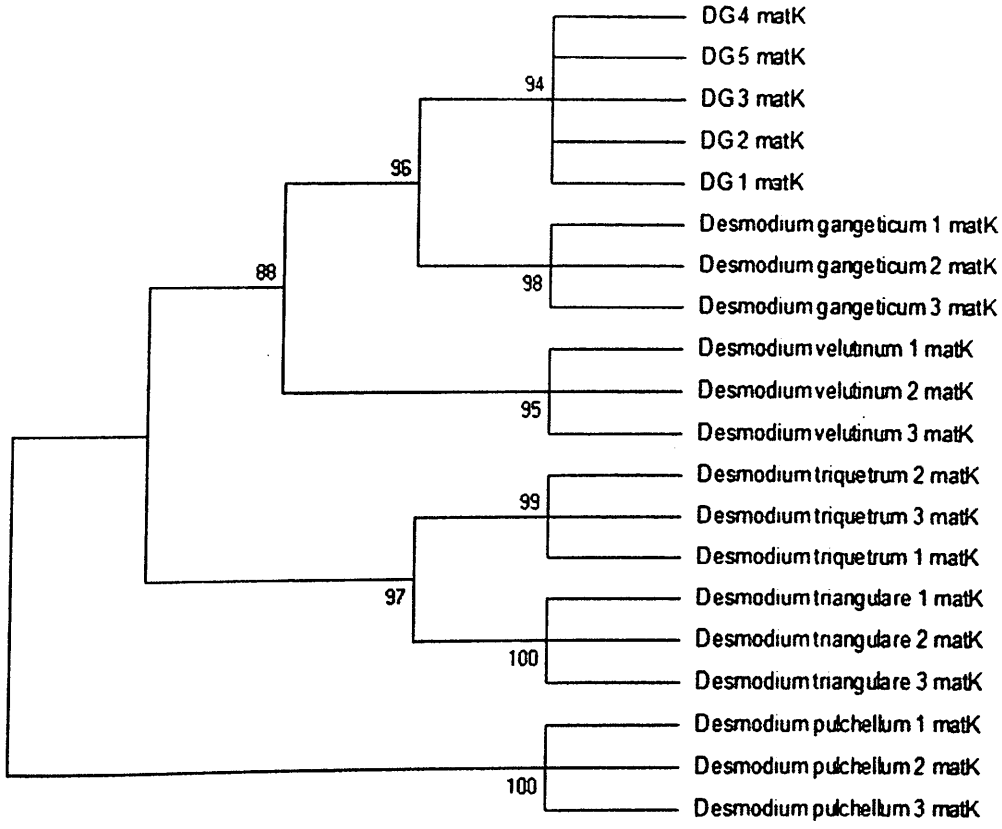


Figure 11. Maximum Likelihood tree (ML) of market samples along with BRM using *mat*K barcode

3.2.1. MLA based analysis of *Desmodium gangeticum* and its adulterants

All four classification methods were run in WEKA with 10-fold cross validation. Among four machine learning algorithms, JRip failed to identify the sequences of reference data set as well as the test sequence database. SMO, Naive bayes and J48 showed species identification in BRM samples with 100 per cent discrimination power. These three machine learning algorithms were subsequently employed for authentication of sequences of unknown traded samples. When the test data of unknown traded samples with large variations were analysed, the Naive bayes classifier could identify only 33 per cent of the species. SMO and J48 showed best performance with 100 % accuracy in authenticating the test data of market samples (Fig. 12). Traded samples showed presence of adulterants which again substantiated the dendrogram based sequence analysis in MEGA.

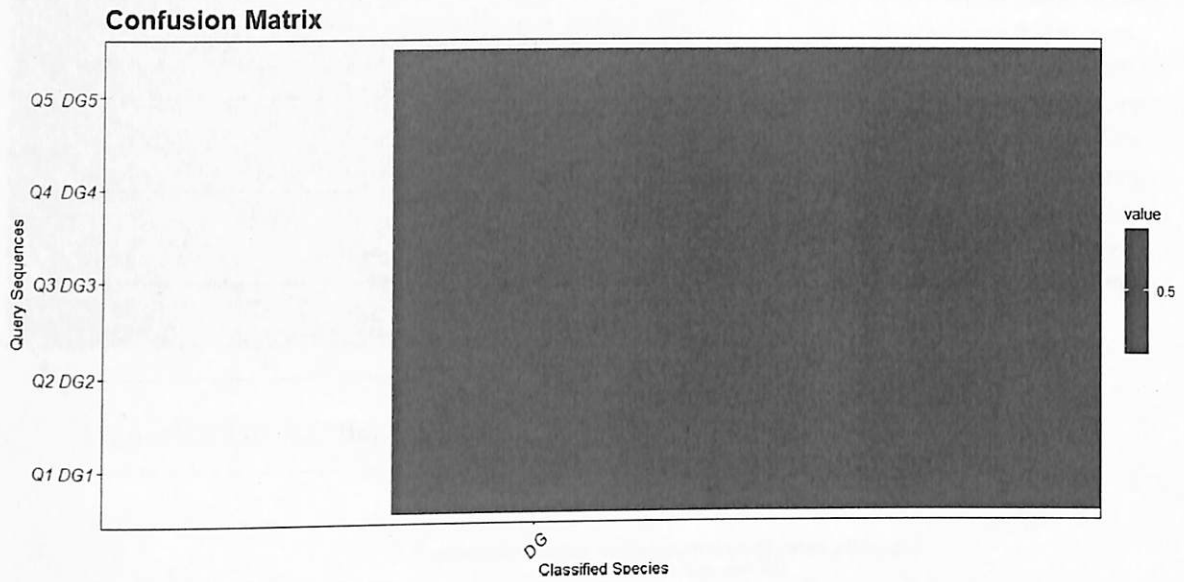
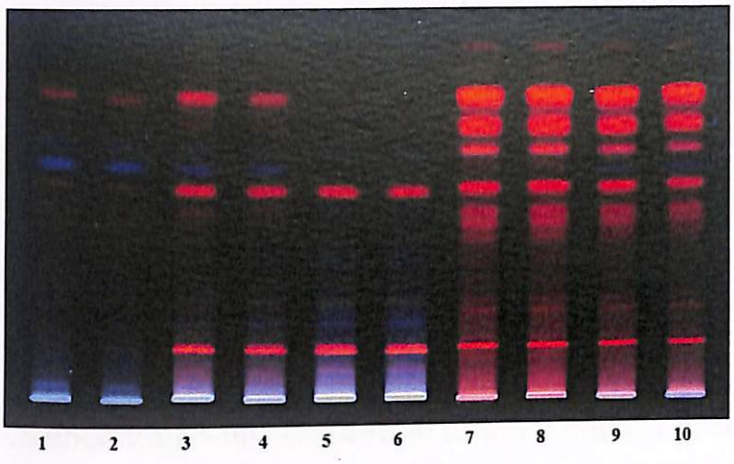


Figure 12. The confusion matrix showing identification rate of market samples along with BRM based on *mat*K

3.2.2. HPTLC analysis of *Desmodium gangeticum* and its adulterants

Chemical profile of each sample was analysed according to their RF values (Retention factor) (Fig. 13). Dendrogram generated using RF values was used to analyse the chemical fingerprinting of the selected raw drugs and its adulterants (Fig. 14). Each species showed specific banding patterns. Compared to other species different accessions of *Desmodium velutinum* and *D. pulchellum* showed similar banding patterns. Hence, HPTLC fingerprinting did not use further for raw drug authentication.



1 *Desmodium gangeticum*
2 *Desmodium gangeticum*
3 *Desmodium triquetrum*
4 *Desmodium triquetrum*
5 *Desmodium triangulare*
6 *Desmodium triangulare*
7 *Desmodium pulchellum*
8 *Desmodium pulchellum*
9 *Desmodium velutinum*
10 *Desmodium velutinum*

Figure 13. HPTLC fingerprinting

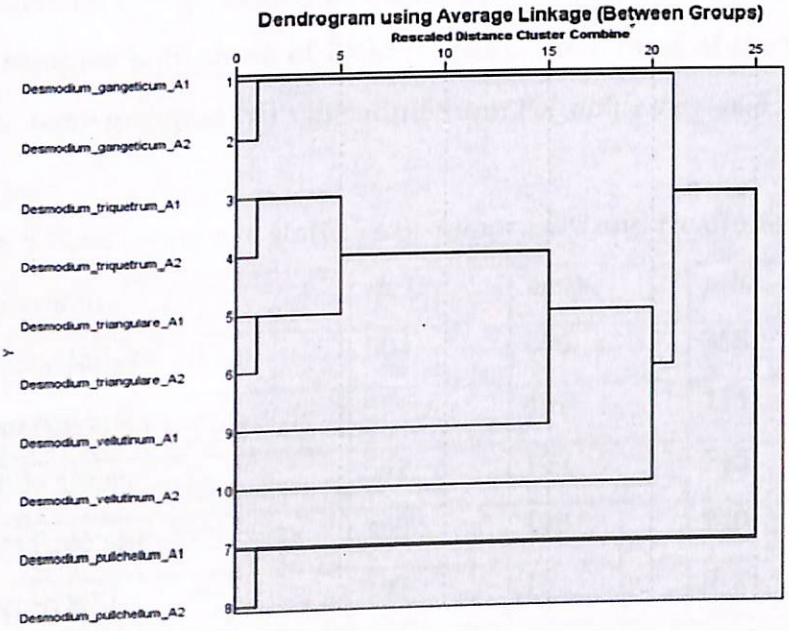


Figure 14. Dendrogram constructed using HPTLC banding pattern

3.3. *Saraca asoca* and its adulterants

All the analysed barcode regions (*ITS*, *psb*A-*trn*H, *mat*K and *rbc*L) amplified successfully with 100 per cent PCR efficiency. Among these barcode regions, *rbc*L region showed highest nucleotide variation (547/604), followed by *psb*A-*trn*H (240/355), *ITS* (220/270) and *mat*K (199/606) regions, respectively (Table 9). Intra and inter specific genetic divergences analysed from the four barcode region showed interspecific but no intra specific divergence. Among these four barcode regions,

*rbc*L showed highest inter specific divergence followed by *psb*A-*trn*H (Table 10; Fig.15). Based on intra and inter specific distances, barcode gap was also estimated. The barcode regions *viz. rbc*L, *psb*A-*trn*H, *ITS* and *mat*K showed distinct barcode gaps of 0.2064, 0.1944, 0.321 and 0:0541 respectively. Wilcoxon's signed rank test performed to test the significance of interspecific divergence in these barcode regions (*psb*A-*trn*H, *mat*K, *rbc*L, and *ITS*), showed significant values. Based on barcode gap analysis, *ITS* and *rbc*L gene regions can be considered as potent barcodes to authenticate *Saraca asoca* from its adulterants. Further, *rbc*L barcode region alone was adopted for authentication of traded samples successfully owing to difficulty in amplification of other barcode regions from degraded DNA of traded raw drugs. The phylogenetic tree generated based on these sequences showed a clear clustering of traded samples with those of BRM (Fig.16). Thus, most of the traded samples of *Saraca asoca* were grouped with the adulterant, *Polyalthia longifolia.*

Table 9. Basic sequence statistics of *Saraca asoca* and its adulterants

Comparison	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Sequence length	604	606	355	270
Conserved sites	57	407	115	50
Variable sites	547	199	240	220
Informative site	355	198	240	220
Singleton site	192	1	0	0

Table 10 : Genetic divergence parameters of *Saraca asoca* and its adulterants

Parameters	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Average intraspecific distance	0	0	0	0
Average interspecific distance	0.2064±0.0761	0.0541±0.104	0.1944±0.208	0.321±0.1297

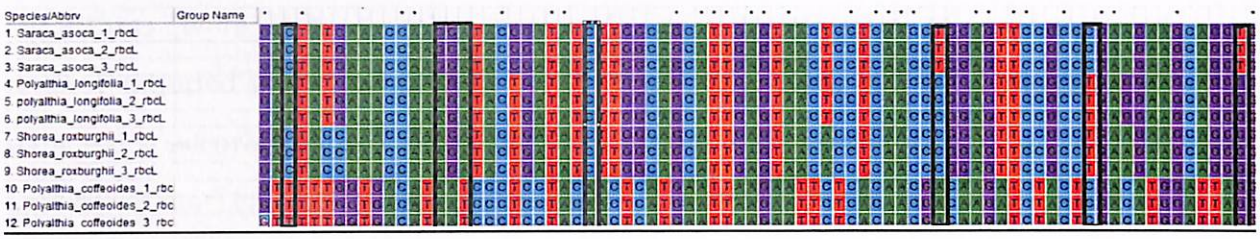


Figure 15. Multiple sequence alignment showing single nucleotide polymorphisms (SNPs) in *rbc*L sequences of Biological Reference Material (BRM) of *Saraca asoca* and its adulterants

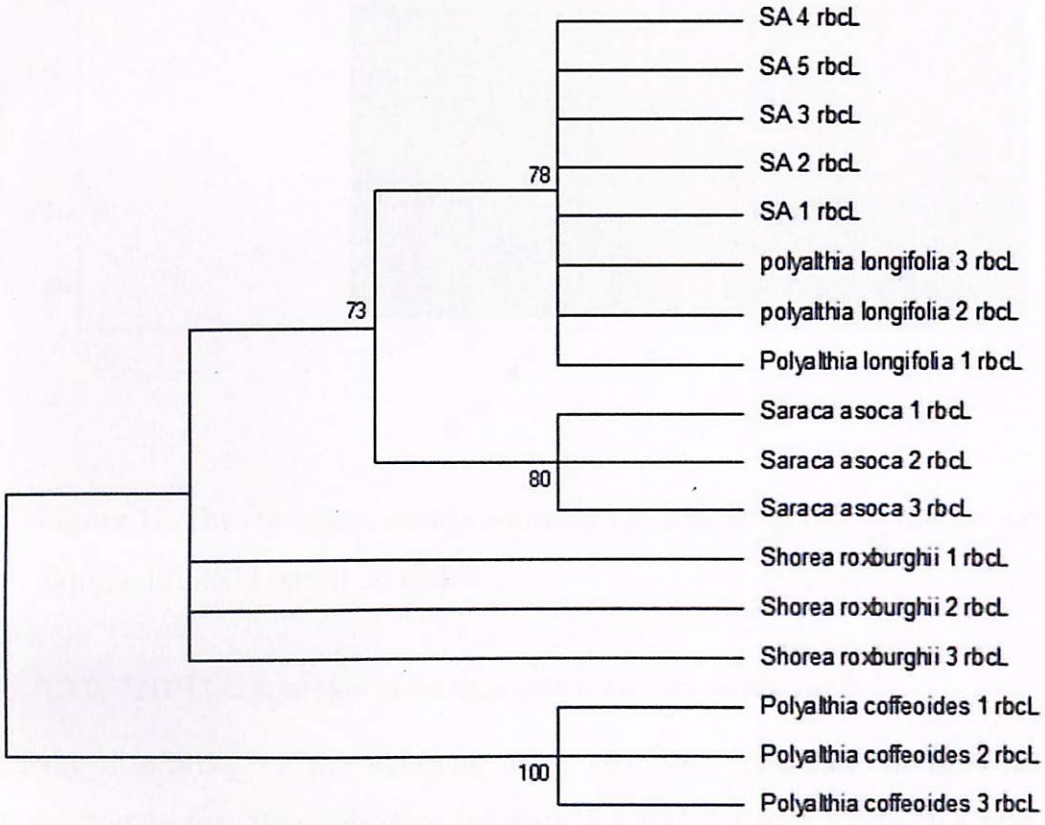


Figure 16. Maximum Likelihood tree (ML) of market samples along with BRM using *rbc*L barcode

3.3.1. MLA analysis of *Saraca asoca* and its adulterants

All four classification methods were run in WEKA with 10-fold cross validation. SMO J48, Jrip and Naive bayes algorithms showed species identification of BRM samples with 100 per cent discrimination power. These four machine learning

algorithms were subsequently employed for authentication of sequences of unknown traded samples. All four classification methods showed good performance with 100 % accuracy in authenticating the test data of market samples (Fig. 17). Test data set of market samples showed similarity with *Polyialthia longifolia* the reference data set (BRM sequence database), which again proved the dendrogram based sequence analysis in MEGA.

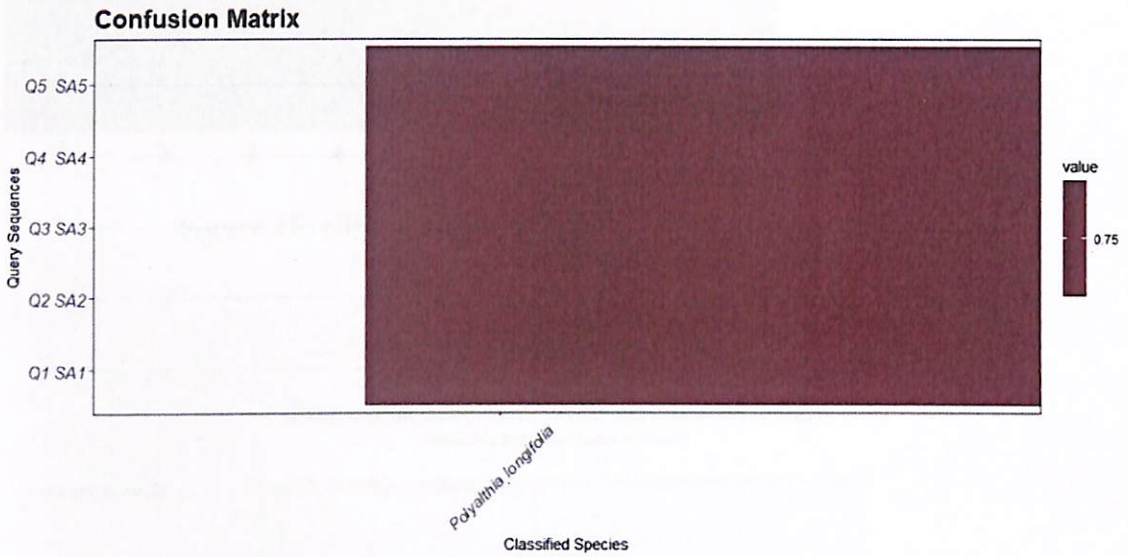
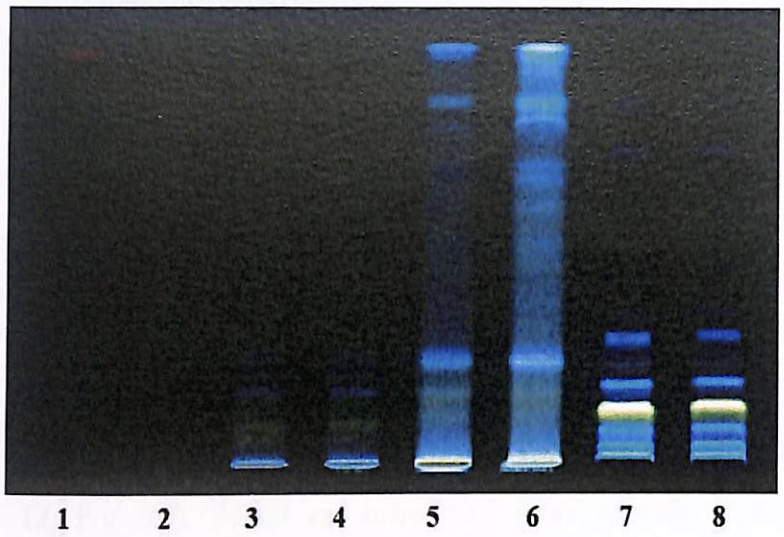


Figure 17. The confusion matrix showing identification rate of market samples along with BRM based on *rbc*L

3.3.2. HPTLC analysis of *Saraca asoca* and its adulterants

Chemical profiles of the samples were analysed according to their RF values (Retention factor) (Fig. 18). Dendrogram generated using RF values was used to analyse the chemical fingerprinting of the selected raw drugs and its adulterants (Fig. 19). Each species showed specific banding pattern with some amount of intra species variation and therefore, accessions belonging to different geographical locations were grouped in different clades. *Saraca asoca* from different geographic locations showed unique banding pattern while adulterant species' showed intraspecific variation. Further, *Shorea roxburghii* from Parambikulam showed similarity with *Polyalthia longifolia* and *Shorea roxburghii* from Palakkad were clustered with *Polyalthia coffeoides*. The discrepancies in banding patterns/RF values could be due to the geographic differences, age and time of collection.



1 *Saraca asoca*
2 *Saraca asoca*
3 *Polyalthia longifolia*
4 *Polyalthia longifolia*
5 *Polyalthia coffeoides*
6 *Polyalthia coffeoides*
7 *Shorea roxburghii*
8 *Shorea roxburghii*

Figure 18. HPTLC finger printing

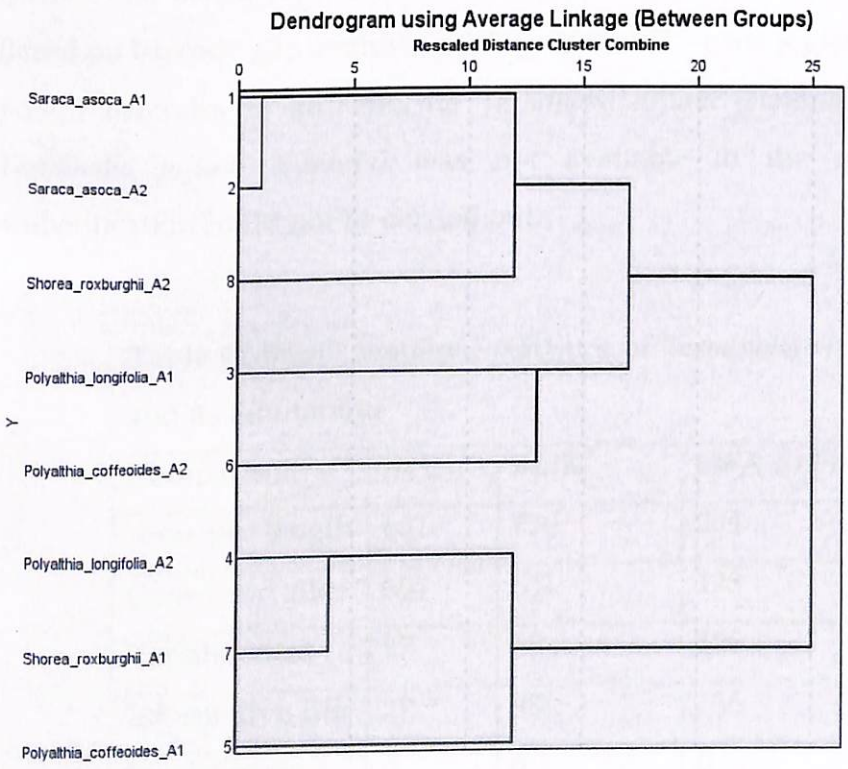


Figure 19. Dendrogram constructed using HPTLC banding pattern

3.4. *Terminalia arjuna (cuneata)* and its adulterants

All the analysed barcode regions (*ITS*, *psb*A-*trn*H, *mat*K and *rbc*L) amplified successfully with 100 per cent PCR efficiency. Among these barcode regions, *Psb*A-*trn*H region showed highest nucleotide variation (547/604), followed by *matK* (240/355), *ITS* (220/270) and *rbc*L (199/606) regions, respectively (Table 11). Intra and Inter specific genetic divergences analysed from the four barcode regions showed interspecific but no intraspecific divergences. Among these four barcode regions, *psb*A-*trn*H showed highest inter specific divergence followed by *ITS* (Table 12; Fig. 20). Based on intra and inter specific distances, barcode gap was also estimated. The barcode regions viz. *rbc*L, *psb*A-*trn*H, *ITS* and *mat*K showed distinct barcode gaps of 2.65, 0.068, 0.066 and 0.040 respectively. Wilcoxon's signed rank test performed to test the significance of interspecific divergence in barcode regions (*psb*A-*trn*H, *mat*K, *rbc*L, and *ITS*), showed significant values for all four regions. Based on barcode gap analysis, *psbA-trnH* and *ITS* gene regions can be considered as potent barcodes to authenticate *Terminalia arjuna (cuneata)* from its adulterants. *Terminalia arjuna (cuneata)* was not available in the market. So raw drug authentication could not be carried out.

Table 11. Basic sequence statistics of *Terminalia arjuna (cuneata)* and its adulterants

Comparison	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Sequence length	651	750	284	264
Conserved sites	654	651	128	206
Variable sites	27	99	156	58
Informative site	27	99	156	56
Singleton site	0	0	0	2

Table 12. Genetic divergence parameters of *Terminalia arjuna (cuneata)* and its adulterants

Parameters	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Average intraspecific distance	0	0	0	0
Average interspecific distance	0.066±0.0024	0.040±0.00487	2.65±1.99	0.068±0.0109

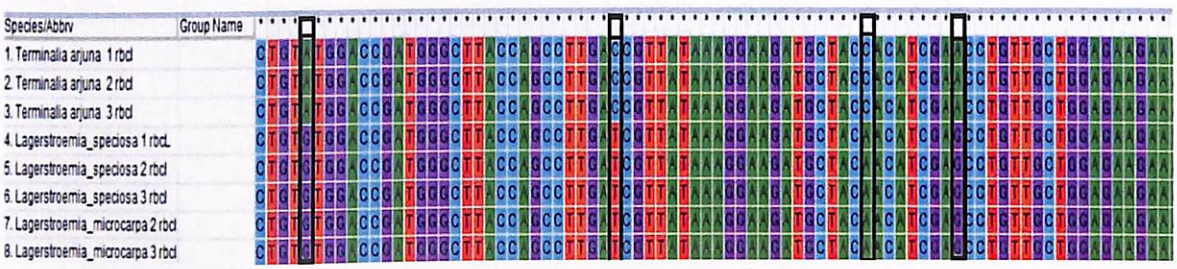
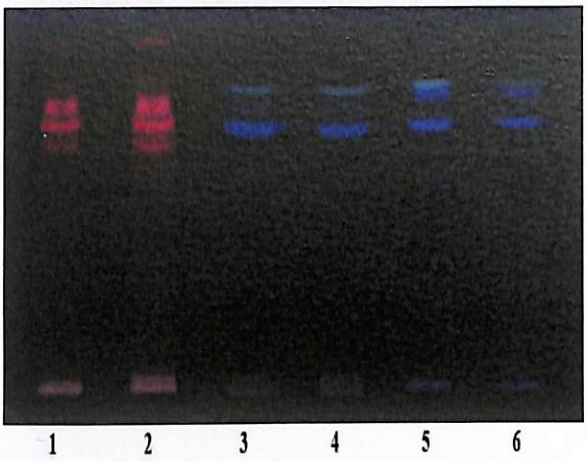


Figure 20. Multiple sequence alignment showing single nucleotide polymorphisms (SNPs) in *rbc*L sequences of Biological Reference Material (BRM) of *Terminalia arjuna (cuneata)* and its adulterants

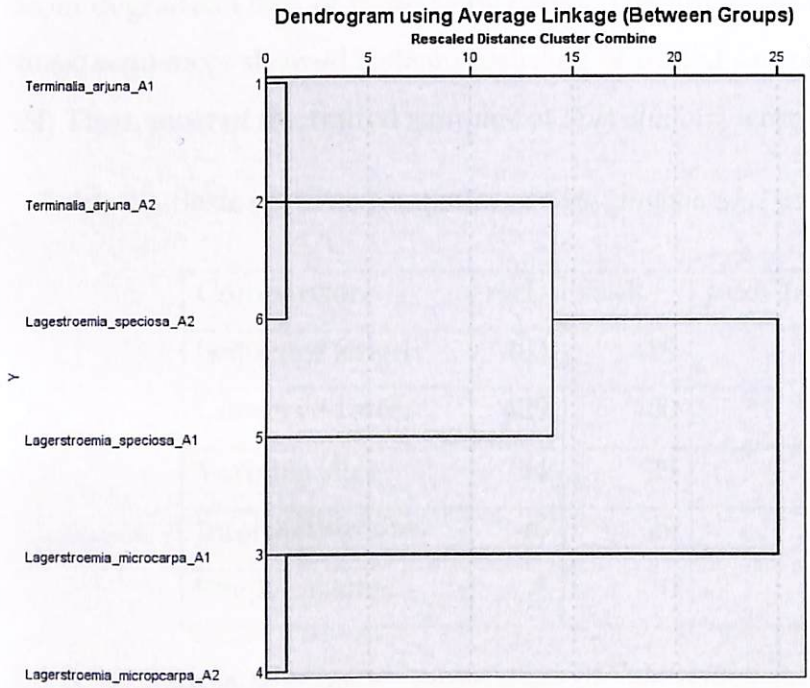
3.4.1. HPTLC analysis of *Terminalia arjuna (cuneata)* and its adulterants

Chemical profile of each sample was analysed according to their RF values (Retention factor) (Fig. 21). Dendrogram generated using RF values was used to analyse the chemical fingerprinting of the selected raw drugs and its adulterants (Fig. 22). Each species showed specific banding patterns. *Lagestroemia speciose* collected from Thrissur showed more similarity to *Terminalia arjuna (cuneata)*. Different accessions of *Lagestroemia microcrapa* showed similar banding pattern and stands as a separate clade.



1 *Terminalia arjuna (cuneata)*
2 *Terminalia arjuna (cuneata)*
3 *Lagestroemia speciosa*
4 *Lagestroemia speciosa*
5 *Lagestroemia microcarpa*
6 *Lagestroemia microcarpa*

Figure 21. HPTLC fingerprinting



lia arjuna and its

-	ITS
	0
1.99	0.068±0.0109

Figure 22. Dendrogram constructed using HPTLC banding pattern

3.5. *Sida alnifolia* and its adulterants

All the analysed barcode regions (*ITS*, *psb*A-*trn*H, *mat*K and *rbc*L) amplified successfully with 100 per cent PCR efficiency. Among these barcode regions, *ITS* region showed highest nucleotide variation (151/547), followed by *psb*A-*trn*H (149/170), *rbcL* (54/483) and *matK* (29/429) regions, respectively (Table 13). Intra and Inter specific genetic divergences analysed from the four barcode region showed

interspecific as well as intra specific divergences. *psb*A-*trn*H and *ITS* regions showed intra specific divergences in *Sida acuta* and *Sida cordifolia*. Among these four barcode regions, *psb*A-*trn*H showed highest inter specific divergence followed by *ITS* (Table 14; Fig. 23). Based on intra and inter specific distances, barcode gap was also estimated. The barcode regions *viz. psb*A-*trn*H, *ITS*, *mat*K and *rbc*L showed distinct barcode gaps of 0.0806, 0.0199, 0.0033 and 0.00083, respectively. Wilcoxon's signed rank test performed to test the significance of interspecific divergence in barcode regions (*psb*A-*trn*H, *mat*K, *rbc*L and *ITS*), showed significant values for all four regions. Further, *rbc*L barcode region alone was adopted for authentication of traded samples successfully owing to difficulty in amplification of other barcode regions from degraded DNA of traded raw drugs. The phylogenetic tree generated based on these sequences showed a clear clustering of traded samples with those of BRM (Fig. 24) Thus, most of the traded samples of *Sida alnifolia* were identified as *Sida cordifolia*.

Table 13. Basic sequence statistics of *Sida alnifolia* and its adulterants

Comparison	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Sequence length	483	429	170	547
Conserved sites	429	400	21	396
Variable sites	54	29	149	151
Informative sites	46	29	149	151
Singleton sites	8	0	170	0

Table 14. Genetic divergence parameters of *Sida alnifolia* and its adulterants

Parameters	*rbc*L	*mat*K	*psb*A-*trn*H	*ITS*
Average intraspecific distance	0	0	0.0006±.0003	0.0008±0.002
Average interspecific distance	0.00083±0.00054	0.0033±0.00096	0.0812±0.0196	0.0207±0.0028
Average theta	0	0	0.0033±0.0012	0.0024±0.0022
Average coalescent depth	0	0	0.0024±0.0019	0.0042±0.001
Average theta prime	0	0	0.0025±0.0020	0.0032±0.0024

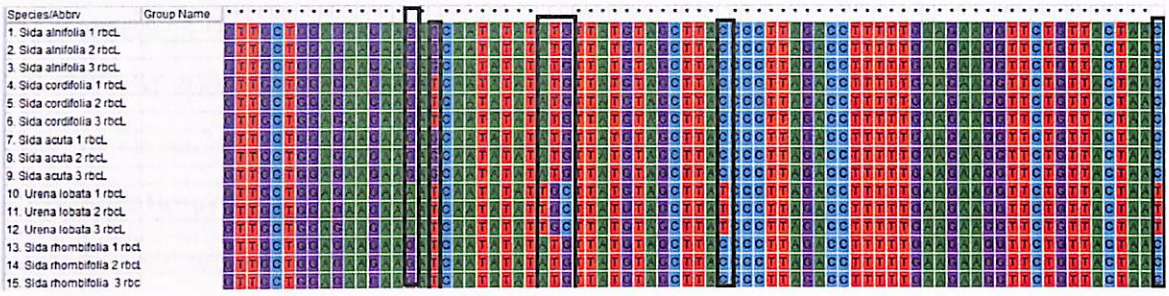


Figure 23. Multiple sequence alignment showing single nucleotide polymorphisms (SNPs) in *rbc*L sequences of Biological Reference Material (BRM) of *Sida alnifolia* and its adulterants

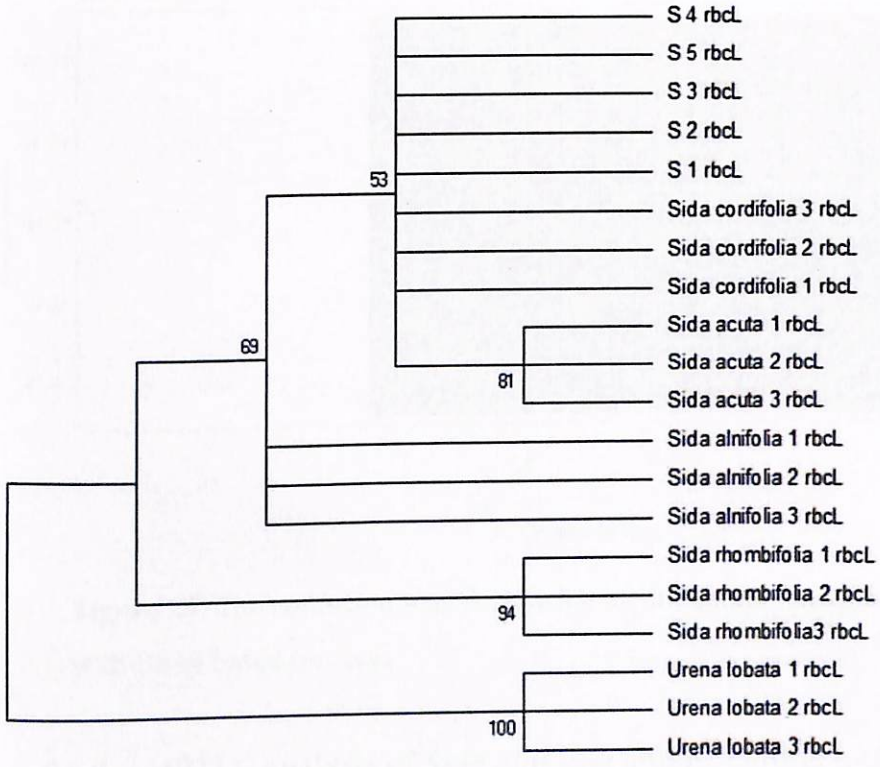


Figure 24. Maximum Likelihood tree (ML) of market samples along with BRM using *rbc*L *barcode*

3.5.1. MLA analysis of *Sida alnifolia* and its adulterants

All four classification methods were run in WEKA with 10-fold cross validation. Among four machine learning algorithms, Naïve bayes and JRip failed to identify

the sequences of reference data set as well as the test sequence database. SMO and J48 showed species identification in BRM samples with 100 per cent discrimination power. These two machine learning algorithms were subsequently employed for authentication of sequences of unknown traded samples. SMO and J48 showed best performance with 100 % accuracy in authenticating the test data of market samples (Fig. 25). Test data set of market samples showed similarity with the reference data set (BRM sequence database), which again corroborated the dendrogram based sequence analysis in MEGA.

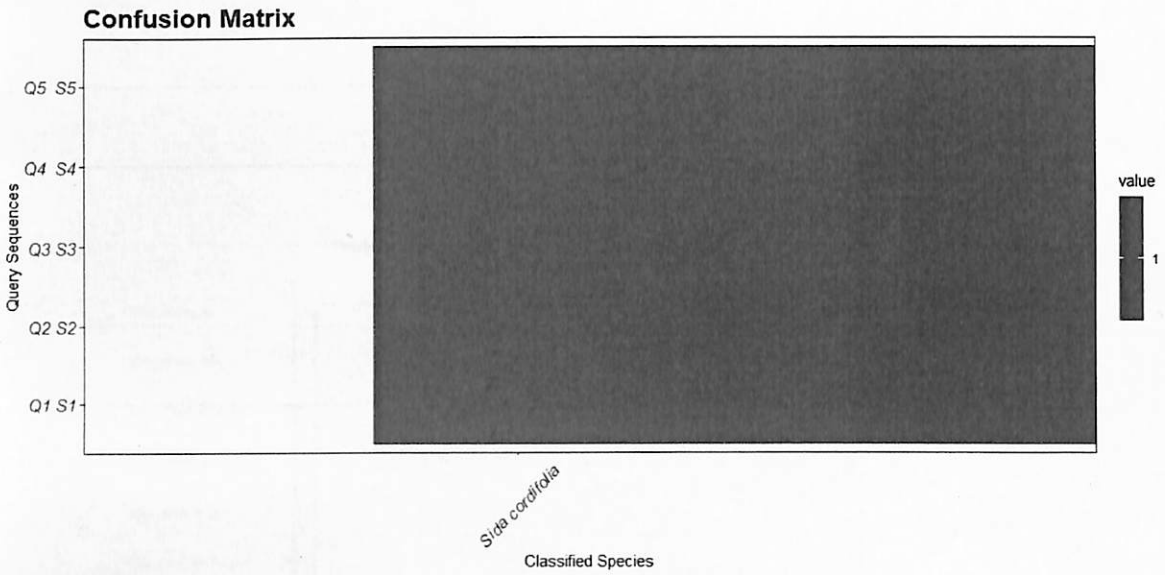
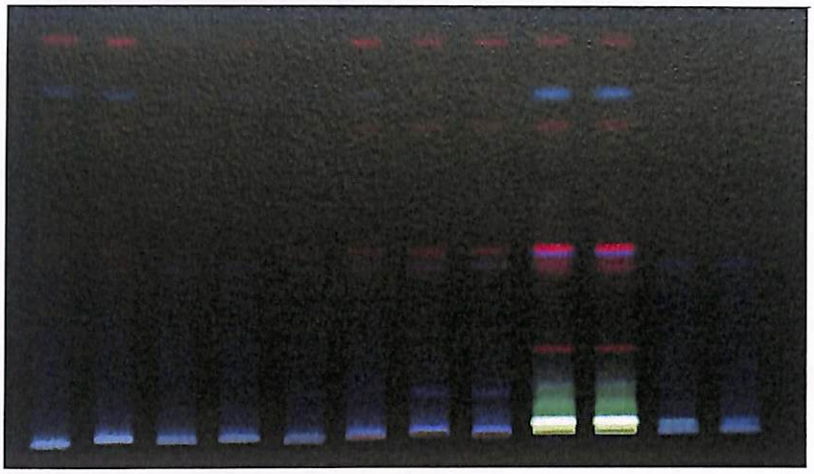


Figure 25. The confusion matrix showing identification rate of market samples along with BRM based on *rbc*L

3.5.2. HPTLC analysis of *Sida alnifolia* and its adulterants

Chemical profile of each sample was analysed according to their RF values (Retention factor) (Fig. 26). Dendrogram generated using RF values was used to analyse the chemical fingerprinting of the selected raw drugs and its adulterants (Fig. 27). Each species showed unique banding patterns. *Sida rhombifoila, S. rhomboidea, S. acuta* and *S. cordifolia* showed similar chemical fingerprinting pattern with the original species *S. alnifolia*, while *Urena lobate* stands as a separate clade.



1 *Sida alnifolia*
2 *Sida alnifolia*
3 *Sida rhombifolia*
4 *Sida rhombifolia*
5 *Sida rhomboidea*
6 *Sida rhomboidea*
7 *Sida cordifolia*
8 *Sida cordifolia*
9 *Sida acuta*
10 *Sida acuta*
11 *Urena lobata*
12 *Urena lobata*

Figure 26. HPTLC Fingerprinting

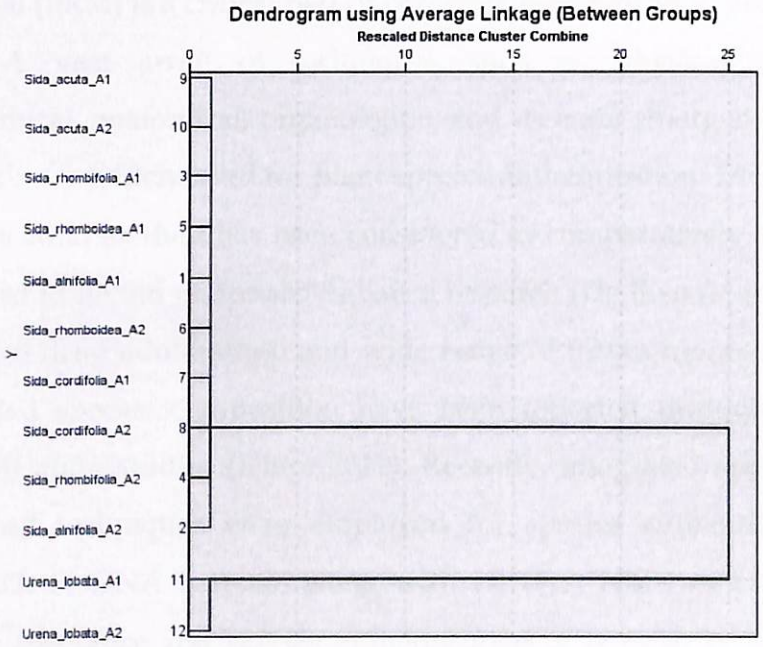


Figure 27. Dendrogram constructed using HPTLC banding pattern

4. DISCUSSION

Authentic herbal drug plays a crucial role in determining the quality, safety and efficacy of herbal formulations. Therefore, to guarantee the quality of herbal medicines, WHO pharmacopeia or International pharmacopeia has implemented certain criteria for proper identification of plant species and quality assessment of potent phytochemical principles (Palhares et al. 2015). Yet adulteration/substitution is a burning problem in ayurvedic industries. Though, ayurvedic medicine has gained much popularity in India and all around the world, proper certification of the products is yet to be in place. Herbal medicine once formulated is easily available to the public through various portals where no mention of any clinical trials or authenticity is cited. This demands a proper industry standard and organized public BRM library for herbal products. The development of voucher Biological Reference Material (BRM) is a critical part of raw drug authentication process.

A vast array of techniques such as physical, chemical (analytical), biochemical, anatomical, organoleptic, and recently emerged DNA based molecular methods are widely used for plant species authentication. DNA barcode-based plant identification method has been considered as comparatively powerful and potential standard in herbal pharmacovigilance research (De Boer et al. 2015). In India, 31 % of herbal drug adulteration and wide range of incongruences between claimed and identified species composition have been reported through DNA barcode-based authentication studies (Ichim 2019). Recently, integrated approach of two or more advanced techniques were employed for species authentication. This multi-tier approach of DNA barcode along with HPTLC, NMR or HPLC were utilized for quality assurance and species authentication of popular species such as *Hamamelis virginiana, Matricaria recutita, Maytenus ilicifolia, Mikania glomerata, Panax ginseng, Passiflora incarnata, Peumus boldus* and *Valeriana officinalis* (Palhares et al. 2014), *Garcinia* species and *Sarca asoca* (Kumar et al. 2016; Seethapathy et al. 2018). However, studies have also reported the inconsistency of chemical markers in delineating medicinal plants, owing to variation with age of plant and environmental heterogeneity (Liu et al. 2011; Kaur et al. 2016; Moustafa et al. 2016;

Cao et al. 2017). Analytical methods also failed to differentiate closely related species containing similar chemical constituents in some instances (Upton et al. 2019)

An integrated approach involving CBOL recommended barcode regions such as *rbc*L, *mat*K, *psb*A, and *ITS* as well as HPTLC profiling was investigated in the present study to create a reference database for majorly traded ayurvedic raw drugs in India. HPTLC fingerprints depicted species specific quality profiles of active principles and were able to distinguish original raw drugs from adulterants in most cases. However, accessions of a species collected from different geographic locations of unknown age, showed variations in the HPTLC fingerprints which restricted their further use in raw drug authentication. Similarly, accessions of species such as *Artemisia japonica, Cinnamomum glaucescens* and *Cymbopogon distans* collected from different geographical locations also showed variation in their chemical pattern (Joshi et al. 2016). In *Tinosporia cordifolia,* sex specific disparity in chemical constituents was earlier reported (Bajpai et al. 2017).

Even though, DNA barcode regions like *ITS* showed promising results in the case of *D. gangeticum, T. arjuna, S. asoca* and *psb*A-*trn*H in *C. fenestratum, S. alnifolia,* high number of indels along with huge interspecific variation failed to provide consistent bidirectional unambiguous sequencing reads which limited their utility for authentication (Chase et al., 2007; Hollingworth et al., 2009; CBOL 2009; Roy et al., 2010). Further, degraded DNA obtained from the market samples with many impurities hindered the primer annealing and subsequent PCR amplification of ITS/*psb*A-*trn*H barcode gene regions as reported earlier by Newmaster et al. 2013. Similar hurdleswere also reported in the traded samples of medicinal plants in Morocco, India and Brazil (Kool et al. 2012; Palhares et al. 2015; Santhosh kumar et al. 2018). Consequently, *rbc*L and *mat*K barcode sequence database which was discriminant enough to identify adulterants in all the cases were used to validate the market samples.

DNA barcode authentication analysis revealed the presence of adulteration in the traded market samples of the studied species. Market samples of *Coscinium fenestratum* showed more similarity with *B. aristata, Saraca asoca* with *Polyalthia*

longifolia, *Sida alnifolia* with *S. cordifolia* and *D. gangeticum* with other similar species of *Desmodium*. Similar vernacular name, presence of potent chemicals, morphological similarity as well as overlapping species distribution are considered as the primary reason for adulteration (Srirama et al. 2017). In the studied species, *C. fenestraium* and *B. aristata* is locally known as 'daruharidra' and contains the potent chemical, berberine. Similarly, *Saraca asoca* and *P. longifolia* known as 'asoka', contains caffeic acid ellagic acid in common. Similar issues of adulteration were reported wherein *Myristica fragrans* adulterated with *M. malabarica*, *Cinnamomum verum* for *C. cassia* and *C. malabatrum* (Swetha et al. 2016, 2017).Consequences of herbal drug adulteration were reported from the countries like Australia, Japan, Taiwan and China, where chronic use of *Artistolochia fangchi* adulterated products led to death of patients due to renal failure (Michl et al. 2013; Jadot et al. 2017). Earlier, US Food and Drug Administration also banned *Piper methysticum* containing products in Germany, Switzerland, France, Canada and UK owing to health issues related to hepatitis, cirrhosis and liver failure (U.S. Food and Drug Administration 2001). Consumer's faith on herbal medicine is in the phase of decline due to extremities in adulteration/substitution (Palhares et al. 2015).

In India, there are no proper guidelines to coordinate and maintain the information related to collection, supply, trade and consumption of botanicals (Kala et al. 2006, Goraya and Ved 2017). Major herbal trade occurs India through conventional collection centres and wholesale markets and most of the herbal drugs available in the market are sourced from wild by informal sectors (Goraya and Ved 2017). There is no codified price for raw drugs, price of collected raw drugs varies from shop to shop. Along with government agencies like Forest Department, tribal cooperative society and Vana Samrakshana Smathi (VSS), there are a number of stakeholders ranging from herb gatherers, local middlemen, urban traders, wholesalers, manufacturers, exporters and herbal healers in the medicinal plants trade sector (Goraya and Ved 2017). Recently, National Medicinal Plant Board (NMPB), Government of India, has launched an online platform e-charak, to create transparent trade linkage among primary collectors to end users of medicinal plant sector.

The resource limitation of ayurvedic raw drugs owing to the escalating demand, leads to adulteration with plants/plant parts of inferior properties. Therefore, to ensure safety and quality of ayurvedic formulations, standard techniques in practice warrant more consistency and precision (Mishra et al. 2016). Raw drugs available in the market need to be analysed critically and strict regulations are needed to monitor the quality of herbal products by authenticating raw dugs from the time of collection, prior to its processing into formulations. Recommendation of a universal tool may not be practical in herbal industry as the analytical methodology solely depends on the type of raw material and the product derived. Along with the recommended standard organoleptic and analytical methods in raw drug authentication, an integrated approach involving a DNA barcoding tool can strengthen the existing practice of quality checking. British pharmacopeia is the first agency to globally implement DNA barcoding method as a tool for authentication, given its ability to identify source of the herbal product accurately compared to traditional methods (Sgamma et al. 2018; Heinrich et al. 2018). The DNA barcode reference library once created and deposited in public domain can be further accessed for authentication of unknown samples, whenever required for certification purposes. Though Ayurvedic medicine has gained much popularity in India and all around the world, proper certification procedures and agencies are yet to be established. It is therefore important to bring forth a statutory body to monitor the proper collection, processing, certification and sale of raw drugs. Substitution of authentic species with species of similar therapeutic effects based on the ancient Ayurveda scripts could reduce the destruction of the existing population of endangered/endemic species. Concurrently, scientific management, restoration and conservation measures should be given utmost priority to augment the depletion of wild resources as well as to meet the rapidly increasing demand of the herbal industries.

5. CONCLUSIONS & FUTURE RECOMMENDATIONS

Authenticity of raw materials (ayurvedic raw drugs) is an essential prerequisite to ensure quality and safety of the consumers. WHO regulatory guidelines and Indian Ayurvedic Pharmacopoeia suggest macroscopic and microscopic evaluation and chemical profiling of the botanical materials for quality control. These methods lack precision when it comes to the identification of extremely dried form of original raw drugs. In this regard, DNA barcode database can offer a foolproof technique to ensure identity of the original raw drug species from its market adulterants. Measures should be taken to include DNA barcoding as a method along with the recommended tools in Indian Ayurvedic Pharmacopeia. Steps may be taken to develop the basic infrastructure and human resource in different parts of the country for the development of DNA barcode database and implementation of the same among the end users/Ayurveda drug manufacturers. National Medicinal Plant Board, Ministry of AYUSH can take an initiative to set up a national certification agency for the certification of ayurvedic raw drugs. Only certified raw drugs would be allowed to use by the ayurvedic industries to ensure the quality and safety of the herbal formulations derived out of that.

REFERENCES

1. Bajpai V, Kumar S, Singh A, Singh J, Negi MPS, Bag SK, Kumar B (2017) Chemometric based identification and validation of specific chemical Markers for geographical, seasonal and gender Variations in *Tinospora cordifolia* stem using HPLC-ESI-QTOF-MS Analysis. Phytochemical Analysis. 28, 277–288. https://doi.org/10.1155/2012/18010310.1002/pca.2673.
2. Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, Barrett SCH (2011) Discriminating plant species in a local temperate flora using the *rbc*L+*mat*K DNA barcode. MEE 2:333–340. https://doi.org/10.1111/j.2041-210X.2011.00092.x.
3. Cao L, Li SX, Wei BY, Huang D, Xu F, Tong ZY (2010) Optimizing RAPD reaction system and authentic genuineness related genetic background of *Fructus Evodia*. Chin Tradit Herb Drugs 41:975-978.
4. Cao Y, Fang S, Yin Z, Fu X, Shang X, Yang W, & Yang H (2017). Chemical fingerprint and multicomponent quantitative analysis for the quality evaluation of *Cyclocarya paliurus* leaves by HPLC–Q–TOF–MS. Molecules. 22, 1927.
5. CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106:12794–12797.
6. Chase MW, Cowan RS, Hollingsworth PM, Van den Berg C, Madrinan S, Petersen G, Seberg O, Jorgsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M (2007). A proposal for a standardised protocol to barcode all land plants. Taxon. 56, 295–299.
7. Chen S, Yao H, Han J, Liu C, Song J, Shi L, Ma X, Zhu Y, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Leon C (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE 5:1. https://doi.org/10.1371/journal.pone.0008613.

8. Chen SL, Yu H, Luo HM, Wu Q, Li CF & Steinmetz A (2016) Conservation and sustainable use of medicinal plants: problems, progress, and prospects. Chinese Medicine, 11(1). doi:10.1186/s13020-016-0108-7

9. CISION, PR News Wire. (2020) India Ayurveda industry outlook 2019–2024 – Market size and growth forecasts, product launches, competitive landscape. Report p. 161. Retrieved from https://www.prnewswire.com/news-releases/india-ayurveda-industry-outlook-2019-2024—market-size-growth-forecasts-product-launches-competitive-landscape-301007651.html.

10. Coghlan ML, Haile J, Houston J, Murray DC, White NE, Mool huijzen P (2012) Deep sequencing of plant and animal DNA contained within Traditional Chinese Medicines reveals legality issues and health safety concerns. PLoS Genet 8: e1002657. https://doi.org/10.1371/journal.pgen.1002657

11. De Boer, HJ, Ichim MC, Newmaster SG (2015) DNA barcoding and pharmacovigilance of herbal medicines. Drug Safety. 38: 611-620. https://doi.org/doi: 10.1007/s40264-015-0306-8.

12. Diao Y, Lin XM, Liao CL, Tang CZ, Chen ZJ, Hu ZL (2009) Authentication of *Panax ginseng* from its adulterants by PCR-RFLP and ARMS. Planta Med 75:557-560. https://doi.org/10.1055/s-0029-1185321.

13. Doebley J, Durbin M, Golenberg EM, Clegg M, Ma DP (1990) Evolutionary analysis of the large subunit of carboxylase *(rbcL)* nucleotide sequence among the grasses (Gramineae). Evolution 44:1097–1108. https://doi.org/10.2307/2409569.

14. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19: 11–15

15. Ganesan A, Marichamy K, Thagamariappan J (2016) Development of medicinal plants sector in India- An Empirical study IJETMAS. 4:82-89.

16. Ganie SH, Upadhyay P, Das S, Prasad Sharma M (2015) Authentication of medicinal plants by DNA markers Plant Gene. 4:83-99. https://doi.org/10.1016/j.plgene.2015.10.002.

17. Goraya GS, Ved DK (2017) Medicinal Plants in India: An Assessment of their Demand and Supply. Dehradun, Ministry of AYUSH.

18. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis: Department of Microbiology, North Carolina State University.

19. Hao DC, Chen SL, Xiao PG, Peng Y (2010) Authentication of medicinal plants by DNA-based markers and genomics. Chin Herb Med 2:250-261.

20. Hebert PDN, Cywinska A, Ball SL, de Waard JR (2003) Biological Identifications through DNA barcodes. Proceedings of the Royal Society B 270:313–322. https://doi.org/10.1098/rspb.2002.2218.

21. Heinrich M, Lardos A, Leonti M, Weckerle C, Willcox M, Apple quist W, Stafford G (2018) Best practice in research: consensus statement on ethnopharmacological field studies–Con SEFS. J Ethnopharmacol 211:329–339.

22. Hollingsworth ML, Andra Clark A, Forrest LL, Richardson J, Pennington RT, Long DG, Hollingsworth PM (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants Mol Ecol Resour. 9:439–457. https://doi.org/10.1111/j.1755-0998.2008.02439.x.

23. Ichim MC (2019) The DNA-based authentication of commercial herbal products reveals their globally widespread adulteration. Frontiers in Pharmacology 10. doi:10.3389/fphar.2019.01227

24. Jadot I, Decleves A, Nortier J and Caron N (2017) An integrated view of aristolochic acid nephropathy: update of the literature. Int J Mol Sci 18:297. https://doi.org/10.3390/ijms18020297.

25. Joshi R, Satyal P, Setzer W (2016) Himalayan aromatic medicinal plants: a review of their ethnopharmacology, volatile phytochemistry and biological activities. Medicines 3(1): 6. doi:10.3390/medicines3010006.

26. Kala C, Dhyani P, Sajwan B (2006). Developing the medicinal plants sector in northern India: challenges and opportunities. Journal of Ethnobiology and Ethnomedicine, 2, 32. https://doi.org/10.1186/1746-4269-2-32

27. Kaur T, Bhat R, Vyas D (2016) Effect of contrasting climates on antioxidant and bioactive constituents in five medicinal herbs in Western Himalayas. J Mt Sci 13:484–492. https://doi.org/10.1007/s11629-014-3380-y.

28. Kiran U, Khan S, Mirza KJ, Ram M, Abdin MZ (2010) SCAR markers: A potential tool for authentication of herbal drugs. Fitoterapia 81:969–976. https://doi.org/10.1016/j.fitote.2010.08.002

29. Kool A, de Boer HJ, Krüger Å, Rydberg A, Abbad A, Björk L, Martin G (2012) Molecular identification of commercialized medicinal plants in southern Morocco. PLoS ONE 7(6): e39459. doi:10.1371/journal.pone.0039459.

30. Krawczyk K, Szczecińska M, Sawicki J (2013) Evaluation of 11 single-locus and seven multilocus DNA barcodes in *Lamium* L. (Lamiaceae). Mol Ecol Resour 14: 272–285. https://doi.org/10.1111/1755-0998.12175.

31. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. Proc. Natl Acad Sci USA 102:8369–8374.https://doi.org/10.1073/pnas.0503123102

32. Kumar SJ, Gogna N, Newmaster SG, Venkatarangaiah K, Subramanyam R, Saroja, SG, 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. International Journal of Legal Medicine 130: 1457–1470. https://doi.org/10.1007/s00414-016-1436-y.

33. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 33: 1870–1874. https://doi.org/10.1093/molbev/msw054.

34. Lewis DD (1998) Naïve (Bayes) at forty: the independence assumption in information retrieval. Mach. Learn. 98: 4–15

35. Liu Y, Reich PB, Li G, Sun S (2011) Shifting phenology and abundance under experimental warming alters trophic relationships and plant reproductive capacity. Ecology 92:1201–1207. https://doi.org/10.1890/10-2060.1.

36. Malik S, Priya A, Babbar S (2018) Employing barcoding markers to authenticate selected endangered medicinal plants traded in Indian markets. PMBP 25:327-337. https://doi.org/10.1007/s12298-018-0610-8.

37. Meena Devi VN, Nagendra Prasad P, Kalirajan K (2010) Infrared spectral studies on Siddha drug- Pavalaparpam. Int. J. Pharma Bio Sci 1:474–483.

38. Meier R, Shiyang K, Vaidya G, Ng PK (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol 55: 715–728. PMID: 17060194.

39. Meyer CP, Paulay G (2005). DNA barcoding: error rates based on comprehensive sampling. Plos One. 3: e422

40. Michl J, Jennings HM, Kite GC, Ingrouille MJ, Simmonds MSJ, Heinrich M (2013) Is aristolochic acid nephropathy a widespread problem in developing countries? A case study of *Aristolochia indica* L. in Bangladesh using an ethnobotanical-phytochemical approach. J Ethnopharmacol 149:235–244. https://doi.org/10.1016/j.jep.2013.06.028.

41. Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, Sundaresan V (2016) DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotechnology J 1:8-21. https://doi.org/10.1111/pbi.12419.

42. Moustafa MF, Hesham AEL, Quraishi MS, Alrumman SA (2016). Variations in genetic and chemical constituents of *Ziziphus spina-christi* L. populations grown at various altitudinal zonation up to 2227m height. Journal of Genetic Engineering and Biotechnology 14: 349–362.

43. Mukherjee PK, Pitchairajan V, Murugan V, Sivasankaran P, Khan Y (2010) Strategies for revitalization of traditional medicine. Chin Herb Med 2:1-15. https://doi.org/10.3969/j.issn.1674-6384.2010.01.001.

44. Newmaster SG, Fazekas AJ, Ragupathy S (2006) DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Can. J. Bot 84:335–341. https://doi.org/10.1139/B06-047

45. Newmaster SG, Grguric M, Shanmughanandhan M, Ramalingam S, Ragupathy S (2013) DNA barcoding detects contamination and substitution in North American herbal products. BMC Med 11:222–35. https://doi.org/10.1186/1741-7015-11-222

46. Ouarghidi A, Powell B, Martin GJ, Hugo Abbad A (2013) Species substitution in medicinal roots and possible implications for toxicity in Morocco. Econ Bot 66: 370-382. https://doi.org/10.1007/s12231-012-9215-2

47. Palhares RM, Drummond MG, Brasil BS, Krettli AU, Oliveira GC, Brandao MG (2014) The use of an integrated molecular, chemical, and biological based approach for promoting the better use and conservation of medicinal species: a case study of Brazilian quinas. J Ethno pharmacol 155:815-22. https://doi: 10.1016/j.jep.2014.06.040

48. Palhares RM, Gonçalves Drummond M, Alves Figueiredo Brasil, BDS, Cosenza GP, Lins Brandao MDG, Oliveira G (2015) Medicinal plants recommended by the World Health Organization: DNA barcode identification associated with chemical analyses guarantees their quality. PLoS ONE 10: e0127866. https://doi: 10.1371/journal.pone.0127866.

49. Quinlan JR (1996) Improved use of continuous attributes in C4.5. Journal of Artificial Intelligence Research 4: 77–90.

50. Raclariu AC, Mocan A, Popa MO, Vlase L, Ichim MC, Crisan G (2017) *Veronica officinalis* product authentication using DNA metabarcoding and HPLC-MS reveals widespread adulteration with *Veronica chamaedrys*. Front. Pharmacol 8:378. https://doi.org/10.3389/fphar.2017.00378.

51. Roy S, Tyagi A, Shukla V, Kumar A, Singh UM, Chaudhary LB, Tuli R (2010). Universal plant DNA barcode loci may not work in complex groups: a case study with Indian *Berberis* Species. PLoS ONE 5(10): e13674. doi:10.1371/journal.pone.0013674.

52. Santhosh Kumar JU, Krishna V, Seethapathy GS, Ganesan R, Ravikanth G, Shaanker RU (2018) Assessment of adulteration in raw herbal trade of important medicinal plants of India using DNA barcoding. 3 Biotech *8(3)*. doi:10.1007/s13205-018-1169-3.

53. Schippmann U, Leaman DJ, Cunningham AB (2002) Impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues, in Biodiversity and the Ecosystem Approach in Agriculture, Forestry and Fisheries, FAO, Rome, Italy.

54. Seethapathy GS, Tadesse M, Urumarudappa SKJV, Gunaga S, Vasudeva R, Malterud KE, Wangensteen H (2018) Authentication of *Garcinia* fruits and food supplements using DNA barcoding and NMR spectroscopy. Scientific Reports 8: 10561. https://doi.org/10.1038/s41598-018-28635.

55. Sgamma T, Masiero E, Mali P, Mahat M and Slater A (2018) Sequence-specific detection of *Aristolochia* DNA - a simple test for contamination of herbal products. Front Plant Sci 9:1828. https://doi.org/10.3389/fpls.2018.01828.

56. Shahzad W, Asad S, Khan MA (2013) Feature subset selection using association rule mining and JRip classifier. International Journal of Physical Science 8: 885–96.

57. Sharma A, Kumar N, Mishra IG (2008) Role of molecular marker in the genetic improvement of the medicinal and aromatic plants. Biotechnological approaches for medicinal and aromatic plants. Kumar, Nitish (Eds.), Biotechnological approaches for medicinal and aromatic plants. Springer Singapore 557–567. https://doi.org/10.1007/978-981-13-0535-1.

58. Smillie TJ, Khan IA (2009) A Comprehensive approach to identifying and authenticating botanical products. Clin Pharmacol Ther 87:175–186. https://doi.org/10.1038/clpt.2009.287

59. SPSS Inc. (2007). Released. SPSS for Windows, Version 16.0. Chicago, SPSS Inc.

60. Srirama R, Santhosh Kumar JU, Seethapathy GS. Newmaster SG, Ragupathy S, Ganeshaiah KN, Ravikanth G (2017) Species Adulteration in the Herbal Trade: Causes, Consequences and Mitigation. Drug Saf 40:651–661 https://doi.org/10.1007/s40264-017-0527-0

61. Sucher NJ, Carles MC (2008) Genome-based approaches to the authentication of medicinal plants. Planta Med 74:603-623. https://doi.org/ 10.1055/s-2008-1074517.

62. Suykens JAK, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Letter. 9: 293–300.

63. Swetha VP, Parvathy VA, Sheeja TE, Sasikumar B (2016) Authentication of *Myristica fragrans* Houtt. using DNA barcoding. Food Control 73:1010–1015. https://doi.org/10.1016/j.foodcont.2016.10.004.

64. Tamhankar S, Ghate V, Raut A, Rajput B (2009) Molecular profiling of Chirayat complex using Inter Simple Sequence Repeat (ISSR) markers. Planta Med 75:1266-1270. https://doi.org/10.1055/s-0029-1185543.

65. Techen N, Parveen I, Pan Z, Khan IA (2014) DNA barcoding of medicinal plant material for identification. Curr Opin Biotechnol. 25:103–110. https://doi.org/10.1016/j.copbio.2013.09.010.

66. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25: 4876–4882.

67. Upton R, David B, Gafner S, Glasl S (2019) Botanical ingredient identification and quality assessment: strengths and limitations of analytical techniques. Phytochemistry Reviews. doi:10.1007/s11101-019-09625-z

68. US Food and Drug Administration (2001) Dietary supplements: aristolochic acid.

69. Walker KM, Apple Quist WL (2012) Adulteration of selected unprocessed botanicals in the US retail herbal trade. Econ Bot 66:321–7. https://doi.org/10.1007/s12231-012-9211-6.

70. Weitschek E, Fiscon G, Felici G (2014) Supervised DNA Barcodes species classification: analysis, comparisons and results. Biodata Minining 7: 4.

71. WHO (2011) The World medicines situation—traditional medicines: global situation, issues and challenges. WHO Press, Geneva, Switzerland.

72. Yu N, Wei Y, Zhu Y, Zhu N, Wang Y, Zhang H, Sun A (2018) Integrated approach for identifying and evaluating the quality of *Marsdenia tenacissima* in the medicine market. PLoS ONE 13: e0195240. https://doi.org/10.1371/journal.pone.0195240.

Appendix 1. Genbank accession numbers generated for 80 samples

Genbank accession numbers for the generated barcode sequences			
Sl No.	**Species**	**Gene region**	**Accession No.**
1	*Coscinium fenestratum*	*rbc*L	MT787043- MT787047
2	*Anamirta cocculus*	*rbc*L	MT787048- MT787052
3	*Morinda pubescens*	*rbc*L	MT787058- MT787062
4	*Diploclisia galucescnes*	*rbc*L	MT787093- MT787097
5	*Berberis aristata*	*rbc*L	MT787053- MT787057
6	*Coscinium fenestratum*	*mat*K	MT787063- MT787067
7	*Anamirta cocculus*	*mat*K	MT787088- MT787092
8	*Morinda pubescens*	*mat*K	MT787083- MT787087
9	*Diploclisia galucescnes*	*mat*K	MT787078- MT787082
10	*Berberis aristata*	*mat*K	MT787068- MT787072
11	*Coscinium fenestratum*	*PsbA-trnH*	MT787108- MT787112
12	*Anamirta cocculus*	*PsbA-trnH*	MT787123- MT787127
13	*Morinda pubescens*	*PsbA-trnH*	MT787118- MT787122
14	*Diploclisia galucescnes*	*PsbA-trnH*	MT787113- MT787117
15	*Berberis aristata*	*PsbA-trnH*	MT787158- MT787162
16	*Coscinium fenestratum*	*ITS*	MT787143- MT787147
17	*Anamirta cocculus*	*ITS*	MT787098- MT787102
18	*Morinda pubescens*	*ITS*	MT787153- MT787157
19	*Diploclisia galucescnes*	*ITS*	MT787148- MT787152
20	*Berberis aristata*	*ITS*	MT787133- MT787137


KFRI