RESEARCH ARTICLE

Employing barcoding markers to authenticate selected endangered medicinal plants traded in Indian markets

Saloni Malik¹ · Akanksha Priya¹ · Shashi B. Babbar¹

Received: 23 May 2018 / Revised: 24 September 2018 / Accepted: 26 September 2018 / Published online: 29 September 2018
© Prof. H.S. Srivastava Foundation for Science and Society 2018

Abstract The high demand of medicinal plants and their unrestricted collection have rendered many of these as rare or endangered. The restrictions imposed on their collection and trade are difficult to implement because of the inability to identify them in fragmented form. The rarity of these plants in nature and lack of their cultivation raise doubt about the authenticity of the herbals sold in markets. Therefore, in the present investigation, ITS/ITS2, *matK*, *rbcL* and *rpoC1* sequences of fourteen species of important medicinal plants, some of which are endangered, were generated and checked for their species-specificity (sequences having maximum similarity only with their own) by BLAST1 and/or BOLD identifications. ITS sequences of 12 species were species-specific. However, ITS2 of only 10 of these 12 species were species-specific. As for the chloroplast loci, *rbcL* and *rpoC1* sequences of all 14 species could be obtained, while *matK* sequences of only 10 of these could be generated. Of the retrieved sequences, *rbcL*, *rpoC1* and *matK* sequences of 7, 11 and 7 species, respectively, were species-specific. The sequences of the targeted loci from the herbal samples of these species were difficult to retrieve because of failure in the amplification or sequencing. Nevertheless, based on ITS2 and/or one or more of the chloroplast loci targeted, the botanical identities of 22 herbal market samples were checked by phylogenetic tree, BLAST1 and BOLD identification methods.

Of these 22 samples, only one of each of *Rauvolfia serpentina* and *Picrorhiza kurroa* were found to be authentic.

Keywords DNA barcoding · Herbals · ITS · *matK* · Medicinal plants

Introduction

India, which is ranked sixth among the 12 mega diversity countries, is home to a great variety of ethno-medicinally important plants (Dawa et al. 2013). According to the National Medicinal Plant Board of India, of the 17,000–18,000 species of flowering plants occurring in India, 6000–7000 are documented as medicinal in traditional systems of medicine, Ayurveda, Siddha, Unani and/or Homoeopathy, or are used as folk medicine. About 960 species of these medicinal plants are traded, with 178 having annual consumption exceeding 100 metric tons (<http://www.medicinalplants.in/trade>). Many medicinal plants in India are at the risk of extinction mainly due to their unrestricted collections and habitat destruction. The raw materials are generally collected as roots, tubers, fruits, seeds, flowers and bark which are the essential regenerative organs of the plants (Hussain and Hore 2008). Moreover, with increase in demand associated with limited availability, problems of adulterations and/or substitutions with related or totally unrelated species have become prevalent (Haider 2015; Parveen et al. 2016). The adulteration with a related species that has similar biochemical composition may result in reduced efficacy of the drug, while in some cases the adulterant could be harmful too. The correct botanical identification of these herbs and their adulterants/substituents is necessary to ensure safety and efficacy of the herbal products (Parveen et al. 2016).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12298-018-0610-8>) contains supplementary material, which is available to authorized users.

✉ Shashi B. Babbar
babbar@rediffmail.com

¹ Department of Botany, University of Delhi, Delhi 110007, India

Traditional methods of authentication viz. morphological, microscopic and chemical based tests, cannot be used for identification of herbal materials available in dried, fragmented or powdered forms. Therefore, additional genome-based methods for the identification and authentication of medicinal plants are much needed in context (Sucher and Carles 2008). DNA-based methods of identification are advantageous over other conventional ones due to their high efficacy, sensitivity, reliability and easy standardization (Chase et al. 2005, 2016; Haider 2011; Mishra et al. 2015). DNA barcoding, which is a robust, rapid and cost effective technique, offers a practical solution for the authentication of highly processed plant materials (Selvaraj et al. 2012; Sarwat and Yamdagni 2014; Mishra et al. 2015; Raclariu et al. 2018). Haider and Wilkinson (2011) designed a set of universal chloroplast DNA-specific primers and discussed their use in identification of plant species.

Hebert et al. (2003a, 2003b, 2004) suggested *Cox1* as the barcode for species level identification based on their studies of insects, birds and fishes. However, because of low substitution rate in mitochondrial genome of plants, this locus was not found to be suitable for plants, except for some macroalgae (Saunders 2005). Therefore, different combinations of loci either from the chloroplast genome alone or along with a locus from the nuclear genome have been suggested as barcodes candidate for plants. These are: *trnH-psbA* spacer + ITS2 (Internal Transcribed Spacer 2), *rpoC1* + *matK* + *rpoB* or *rpoC1* + *matK* + *psbA-trnH*, *rbcL* + *trnH-psbA* and *matK* (Kress et al. 2005; Rubinoff et al. 2006; Chase et al. 2007; Kress and Erickson 2007; Lahaye et al. 2008). Based on the comparison of 12 candidate loci for their amplification, sequencing and species discrimination among 98 land plant taxa, Ford et al. (2009), concluded that for effective diagnosis a combination of loci would be needed and for this, they recommended further investigation of *matK*, *rpoB*, *rpoC1*, *ndhJ*, *ycf5* and *accD*. The Plant Working Group of the Consortium for Barcode of Life (CBOL) recommended *rbcL* + *matK* as a two-locus barcode for plants (Hollingsworth et al. 2009). Chen et al. (2010) and Pang et al. (2010) demonstrated the efficacy of ITS2 and ITS/ITS2 from the nuclear genome, as barcodes for plants. The utility of ITS/ITS2 was amply demonstrated and a strong case for its inclusion in the core barcode comprising *matK* and *rbcL* was advocated by China Plant BOL Group (2011). Rather in one of the recent studies, ITS1 exhibited higher species discrimination power than ITS2 among the members of taxonomically complex sub-tribe Cassiinae of the family fabaceae (Mishra et al. 2016). In our previous study, all the 36 investigated species of *Dendrobium* could be discriminated on the basis of ITS alone (Singh et al. 2012). Though in some studies, dealing with limited number of species in a genus;

single locus has proved to be successful, e.g. *matK* for *Paphiopedilum* spp. (Parveen et al. 2012), ITS for *Dendrobium* spp. (Singh et al. 2012) and *psbA-trnH* spacer for *Myristica* spp. (Swetha et al. 2017), generally, a combination of loci are required for the best species discrimination (Parveen et al. 2017; Mishra et al. 2017). Thus, in a number of studies, ITS, ITS2, *matK* and *rbcL*, along with a few others, such as *rpoC1* and *trnH-psbA* spacer, have been used for the identification of medicinal plants and their herbal samples available in the markets (Song et al. 2009, Kool et al. 2012, Yuan et al. 2015).

In the present investigation, ITS, ITS2, *matK*, *rbcL* and *rpoC1* loci were sequenced and checked for their specificity for 14 medicinal plant species, viz. *Aquilaria malaccensis* (Agarwood), *Coptis teeta* (Mamira), *Dioscorea deltoidea* (Singli-mingli), *Gentiana kurroo* (Trahi-maan), *Kaempferia galanga* (Kapoor Kachri), *Panax pseudoginseng* (Ginseng), *Picrorhiza kurroa* (Kutki), *Pterocarpus santalinus* (Raktachandan), *Podophyllum hexandrum* (Himalayan mayapple, Bankakri), *Rauvolfia serpentina* (Sarpagandha), *Saussurea costus* (Kuth), *Swertia chirayita* (Chiraiyata), *Taxus wallichiana* (Talispatra) and *Vanda coerulea* (Blue Vanda). Eight of these species are listed in Appendix II of CITES and one, *S. costus*, is included in Appendix I. As per International Union for Conservation of Nature (IUCN), *S. costus* and *G. kurroo* are critically endangered, *C. teeta*, *P. santalinus* and *T. wallichiana* are endangered, while *A. malaccensis* is vulnerable. The status of the remaining eight species is “not-assessed” (<http://www.iucnredlist.org/>). The developed barcodes were used to ascertain botanical identity of some herbal samples of the targeted species being sold in markets.

Materials and methods

Plant material

A total of 49 accessions belonging to 14 species were used in the present study (Online Resource 1, Table S1). The plant materials of selected species were collected/procured from different locations in India (Online Resource 1, Table S1). For some of the species, only leaves were provided by Jawaharlal Nehru Tropical Botanic Garden and Research Institute (Kerala), Botanical Survey of India (BSI, Shillong) and BSI (Dehradun), along with accession numbers of the herbariums deposited in their respective repositories (Online Resource 1, Table S1). The botanical identities of rest of the collected plants were verified by comparing with the herbarium specimens available at BSI, Dehradun and with the help of Prof. A. K. Pandey, Professor, Botanical Survey of India. Herbariums of each of

these accessions were prepared and submitted to the Delhi University Herbarium (DUH) and accession numbers were obtained. A set of 28 herbal samples of 13 of these species were procured from three different markets in India, viz. Delhi, Rishikesh and Amritsar (Online Resource 1, Table S2). The herbal sample of *P. hexandrum* was not available in any of the three markets visited.

DNA extraction and Polymerase Chain Reaction (PCR) amplification

CTAB method (Doyle and Doyle 1987) was used for isolation of total genomic DNA from the collected plant materials. As CTAB method did not yield satisfactory results for most of the herbal samples, for the isolation of DNA from such samples, modified CTAB method (Barnwell et al. 1998), the method used by Warude et al. (2003) or DNeasy Plant Minikit (Qiagen) were used. Four loci (ITS, *matK*, *rbcL* and *rpoC1*) were amplified from the field samples, while for market samples only ITS2, *rbcL* and *rpoC1* could be amplified. For amplification of the loci by PCR, the reaction mixture (20 µl) comprised 1 unit of *Pfu* DNA polymerase (Fermentas Inc., USA, #EP0502), 2 µl of 10X PCR buffer with 20 mM MgSO₄, 2 µl of 2 mM dNTPs, 2 µl each of forward and reverse primers (10 µM) and 20–30 ng of template DNA. The details of primers and the respective thermal cycles used for the amplification of ITS, ITS2, *matK*, *rbcL* and *rpoC1* are given in Online Resource 1, Table S3. The amplicons were electrophoresed on 1% TAE (Tris–acetic acid-EDTA buffer) agarose gel with 1 µM EtBr and visualized on UV transilluminator (Alpha Imager Pvt. Ltd, Bengaluru). The amplicons were cleaned using Exo-SAP method (Bell 2008). The clean amplified loci were sequenced using Sanger's method (Sanger et al. 1977) on ABI Prism 3700 DNA Analyzer, as described earlier (Singh et al. 2012).

DNA sequence analysis

Consensus sequences were generated using Codon Code Aligner V 5.0.1 (Codon Code Co., USA). The DNA sequences were submitted to GenBank and the accession numbers were obtained (Online Resource 1, Table S4). The sequences of each locus were aligned using BioEdit v 7.2.5 (Hall 1999). Intra- and inter-specific Kimura 2-parameter model (K2P) distances were calculated using distance matrices generated by MEGA 6.0 (Tamura et al. 2013). For determining the suitability of a locus as an identification tag (barcode) for the species, specificity of its sequence(s) for the species was checked by BLAST1 on NCBI (National Centre for Biotechnology Information) (Ross et al. 2008) and BOLD (Barcode of Life Database) identification tool (Ratnasingham and Hebert 2007) during

January 2017. ITS2 sequences were annotated and trimmed from the complete ITS sequence using the ITS2 database annotation tool (Keller et al. 2009). By BLAST1 method on NCBI, the query sequence was considered as species-specific if the first hit was with its own (already available or submitted by us). If the query sequence also matched with any other species with 100% query coverage and similarity, it was not considered species-specific. It was considered genus-specific if the first hit, with 100% similarity, was with any other species of the same genus. By BOLD identification, the sequence was considered species- or genus-specific if it was identified as the same species, as identified by us, or any other species of the same genus, respectively. The sequences of a locus that exhibited intra-specific variation were considered as specific to the species only if each sequence had the first hit with its own on NCBI. The botanical identities of the market samples were checked by three methods: Phylogenetic tree, BLAST1 and BOLD identifications. As for the first method, Neighbor joining (NJ) trees with 1000 bootstrap replicates based on ITS2, *rbcL* and *rpoC1* sequences from the herbal samples and the corresponding species/genus specific sequences generated from the vouchered specimens were constructed. In the phylogenetic analysis, if the sample sequence did not cluster along with the sequence(s) of the species it was sold as, the sample was not considered authentic and its identity was checked with BLAST1 and BOLD identification methods. By the former method, botanical identity of the herbal sample was considered correct, if the query sequence had its first hit with the species it is sold as. However, if the query sequence had first hit with any other species of the same or different genus, the botanical identity of the sample was considered accurately deciphered only if the similarity was 100% and the sequence with which it matched was of barcode quality as indicated by the title of the indented publication. BOLD identification method could be used only for *rbcL*.

Results

Amplification and sequencing success of the candidate barcode loci

Among the loci tested for the collected plants, *rpoC1* exhibited the highest (92%) amplification success rate, followed by *rbcL* (86%), ITS (82%), and *matK* (70%). The size of the amplicons of ITS, ITS2 (only from the herbal samples), *matK*, *rbcL* and *rpoC1* were 650–680 bp, 160–320 bp, 900–1000 bp, 550–650 bp and 420–460 bp, respectively. The sequencing success rates of the amplified ITS, *matK*, *rbcL* and *rpoC1* loci were 85.3%, 88.5%, 86% and 87%, respectively. For the herbal market samples,

amplification success rate of ITS2, *rbcL* and *rpoC1* were 57%, 71% and 32%, respectively. Of the amplicons, 75%, 100% and 100% of ITS2, *rbcL* and *rpoC1*, respectively, could be sequenced (Table 1). ITS and *matK* loci could not be amplified from any of the market samples. A total of 134 DNA sequences, comprising 31, 30, 35 and 38 of ITS, *matK*, *rbcL* and *rpoC1*, respectively, were generated from the plants of 14 species included in this study. Besides, 41 sequences comprising ITS2, *rbcL* and *rpoC1* were retrieved from 28 herbal samples procured from the markets.

Determination of intra- and inter-specific K2P distances

The intra-specific variations in the ITS sequences of *P. kurroa*, *R. serpentina* and *G. kurroo* were 0–0.003, 0.019 for *P. santalinus* and zero for rest of the investigated species (Online Resource 2). Likewise, *matK* sequences of *P. kurroa* and *S. chirayita* had intra-specific distances ranging from 0 to 0.003 and this value was 0–0.011 for *V. coerulea*, and zero for remaining species (Online Resource 3). Among the investigated species, only *rbcL* sequences of *S. chirayita* had intra-specific variation of 0.005 (Online Resource 4). The *rpoC1* sequences also varied among the accessions of only *R. serpentina* and *T. wallichiana* with the range being 0–0.003 for both species (Online Resource 5). Average inter-specific variations for ITS, *matK*, *rbcL*, and *rpoC1* were 0.423, 0.241, 0.096 and 0.149, respectively (Table 1).

Determination of species specificity

On the basis of BLAST1 and BOLD identification results, the ITS sequences of all 31 samples representing 12 of the 14-investigated species, were species-specific. ITS sequences retrieved from *D. deltoidea* and *T. wallichiana* matched with fungal ITS sequences, thus, were considered as not available (Table 2). However, ITS2 region of only ten of these twelve species were species-specific. ITS2 of

A. malaccensis and *P. pseudoginseng* also matched 100% with other species of their respective genera, and therefore, were only genus-specific (Table 2). The *matK* sequences, which could be retrieved from 10 species, were species-specific for seven species, while the sequences of *D. deltoidea*, *P. pseudoginseng* and *V. coerulea*, were genus specific. The *rbcL* and *rpoC1* sequences retrieved from all 14 species were specific to 11 and 7 species, respectively, while sequences of the remaining species for both loci were genus-specific (Table 2).

Deciphering the botanical identity of the market samples

Of the 28 herbal samples, presumably of 13 species based on the names as they were sold in market, ITS and *matK* sequences could not be obtained from any. The number of ITS2, *rbcL* and *rpoC1* sequences retrieved from these samples belonging to 7, 11 and 7 species were 12, 20 and 9, respectively. In the NJ tree based on ITS2 sequences of seven species, both samples of ‘Kutki’ (*P. kurroa*) and one of the two samples of ‘Sarpagandha’ (*R. serpentina*), of which ITS2 sequences were available in GenBank database, co-segregated with the sequences of their own species (Fig. 1). In the BLAST1 search, ITS2 sequence of one sample of ‘Kutki’ had the first hit with its own species, whereas the sequence of the other sample had first match with the respective sequence of *Berberis asiatica*, though with 96% similarity only (Online Resource 1, Table S5). ITS2 sequences of samples of other species segregated individually on independent branches or with other species (Fig. 1). Likewise, in BLAST1 search, their first matches were with species of genera other than the one they were expected to belong to (Online Resource 1, Table S5). The analysis of *rbcL* sequences with NJ tree, BLAST1 and BOLD identification results of 20 sequences belonging to 11 species, revealed that only one sample of *P. kurroa* had its correct match with its own species (Fig. 2, Online Resource 1, Table S5). On the contrary, *rbcL* sequences of all the three samples of *R. serpentina* matched correctly

Table 1 Amplification and sequencing success rates for the targeted barcode loci and average inter-specific divergence values

Locus	No. of species used for inter-specific divergence	Average inter-specific divergence (range)	Amplification success (%)	Success of sequencing of the amplicons (%)	Identification success rate using blast method (%)	Identification based on tree-based method (%)
ITS	12	0.423 (0.074–1.222)	82	85.3	85.7	100
ITS2	11	0.591 (0.314–1.853)	–	–	71.4	100
<i>matK</i>	10	0.241 (0.087–0.398)	70	91	70	100
<i>rbcL</i>	14	0.096 (0.025–0.195)	86	86	78.5	100
<i>rpoC1</i>	14	0.149 (0.024–0.382)	92	87	64.3	100

Table 2 BLAST1/BOLD based identification of the sequences generated in the present study

S. no.	Botanical name	ITS	ITS2	MatK	RbcL	RpoC1
1	<i>Aquilaria malaccensis</i>	C	C*	C	C	X
2	<i>Coptis teeta</i>	C	C	SNA	C	C
3	<i>Dioscorea deltoidea</i>	SNA	SNA	C*	C	C*
4	<i>Gentiana kurroo</i>	C	C	C	C*	C*
5	<i>Kaempferia galanga</i>	C	C	SNA	C	C
6	<i>Panax pseudoginseng</i>	C	C*	C*	C	C*
7	<i>Podophyllum hexandrum</i>	C	C	C	C*	C
8	<i>Picrorhiza kurroa</i>	C	C	C	C	C
9	<i>Pterocarpus santalinus</i>	C	C	C	C	X
10	<i>Rauvolfia serpentina</i>	C	C	C	C	C
11	<i>Saussurea costus</i>	C	C	SNA	C	C
12	<i>Swertia chirayita</i>	C	C	C	C	C
13	<i>Taxus wallichiana</i>	SNA	SNA	SNA	C*	C*
14	<i>Vanda coerulea</i>	C	C	C*	C	C*

C correct match, species-specific, C* matches with its own and other species of the same genus, thus, genus-specific, SNA sequences not generated, X sequence not even genus specific

with their own species in both NJ tree and BLAST1 analyses (Fig. 2, Online Resource 1, Table S5). However, by BOLD identification method, one of the three samples, was identified as *R. tetraphylla*. The *rbcL* sequence generated from the lone market sample of *T. wallichiana* co-segregated with that of genus-specific sequence (cf. Table 2) of its own species (Fig. 2). In BLAST1 search, it matched with another species, *T. fauna* (Online Resource 1, Table S5). Out of the nine sequences of *rpoC1* retrieved from seven species, only the sequence of *P. kurroa* (generated from one of its samples) and of *R. serpentina* (sequences from all three samples) matched correctly with their respective species in NJ trees (Fig. 3). In BLAST1 analysis, the results were almost the same, except for one sample of *R. serpentina*, which had its first hit with another species of the same genus, *R. tetraphylla* (Online Resource 1, Table S5). Intriguingly, for the samples which were not found to be authentic by phylogenetic analysis, more than one botanical identity was revealed based on the first match by BLAST and/or BOLD identification methods (Online Resource 1, Table S5).

Discussion

Though the development and large-scale production of chemically synthesized drugs during the last century have revolutionized health care across the world, a large number of people in developing countries still depend on herbal medicines for their primary health care. In India, 70% of the population depends on traditional medicines for their health care needs (Wachtel-Galor and Benzie 2011). The demand for most of the medicinal plants is met through

extensive and often destructive collections from wild as only few of these are cultivated. This could lead to extinction of some of the rare or endangered species. The regulatory measures in place to check their indiscriminate collections and illicit trade fail, if the plants are traded in fragmented forms, botanical identities of which are difficult to establish by taxonomic methods requiring whole plants. Another problem infesting this sector is adulteration or substitution for pecuniary benefits or because of lack of expertise of the plant collectors. The adulteration/substitution could be by (1) another easily available species, which could be less efficacious, totally ineffective or even sometimes harmful, (2) a part of the same species that is known to be lacking the therapeutically important biomolecule(s), especially in those species where efficacy of the herbal is restricted to a particular organ, and/or (3) by the plants grown/growing in eco-climatic conditions which are not suitable for building up of requisite levels of active molecules responsible for the therapeutic effect of the plant (Parveen et al. 2016). While the last two-mentioned substitutions can only be checked with biochemical profiling of the sample, botanical identity of a sample can be validated by DNA barcoding, a technique which can be applied even if a minute amount of tissue or its DNA is available (Parveen et al. 2012). The present paper describes results of such an attempt made for 14 important medicinal plant species. Among the four loci tested, *rbcL* and *rpoC1*, though could be retrieved from all the species, had species-specificity for 11 and 9 species, respectively. However, *rbcL* and *rpoC1* sequences for remaining 3 and 5 species, respectively, were genus-specific. These observations are in conformity with earlier investigations where ease of amplification and sequencing of these loci along with their

Fig. 1 Neighbour-joining tree of ITS2 sequences of vouchered specimens and the market samples. The sequences of vouchered specimens have been indicated by the botanical names, whereas market samples have been shown with their common names followed by the botanical names of species these are sold as (in bracket). The accession/sample number is given after the botanical names of the vouchered specimens and after the common name of the market samples. Numbers on each node indicate the bootstrap support

low discrimination rates at species level have been reported (Newmaster et al. 2013). Another locus from the chloroplast genome, *matK*, could be sequenced from 10 of the 14 species and was species-specific for only 7 species. The latter observation is slightly unexpected because this locus generally provides higher species discrimination than *rbcL* and *rpoC1* (Lahaye et al. 2008; CBOL Plant Working Group 2009; Parveen et al. 2012; Singh et al. 2012). In fact, this locus alone was once suggested as the universal barcode for plants (Lahaye et al. 2008). ITS, a locus from the nuclear genome, present in all eukaryotic organisms, poses problem in those plants which have symbiotic relations with other organisms or are contaminated (Hollingsworth et al. 2011), as was also encountered in the present study. Because of this reason, ITS sequences could be generated from 12 of 14 species. The sequences retrieved from the samples of the remaining two species, *D. deltoidea* and *T. wallichiana*, were of fungi. Nevertheless, ITS was species-specific for all the 12 species for which these were available, in consonance with earlier reports of high species discrimination power of this locus (Hollingsworth et al. 2011; Kool et al. 2012).

Contrary to the situation with the vouchered specimens, ITS and *matK* sequences could not be generated from herbal samples. Most of these samples were not freshly collected ones' due to the rarity of the species and some had been in storage for many years, as informed by the sellers. As these loci could be amplified and sequenced from corresponding vouchered specimens, failure with the market samples could possibly have been due to their highly degraded states. Similar problems with these loci have been encountered by others who attempted to use these loci for the identification of herbal samples (Chen et al. 2010; Kool et al. 2012; Kuzmina et al. 2012). To circumvent this, use of ITS2, rather than the complete ITS, has been suggested for identification of medicinal plants and their herbal samples (Chen et al. 2010; Newmaster et al. 2013). Therefore, in the present study ITS2 region annotated from the lab generated ITS sequences of 12 species, were used for identification. Similarly, sequencing of *rbcL* and *rpoC1* markers from the herbal samples did not meet with the similar success as with the vouchered specimens. Thus, only 20 sequences of *rbcL* marker generated from 11 species and 9 sequences of *rpoC1* marker

generated from 7 species could be obtained successfully. Such trends of relatively much lower amplification and sequencing success rates for herbal samples than those of the intact plants have also been reported in previous studies (Kuzmina et al. 2012; Newmaster et al. 2013; Kumar et al. 2015, 2016), possibly because of highly degraded states of the samples. Of the 28 herbals, sequences of ITS2, *rbcL* and *rpoC1* could be retrieved only from 22 samples. The analyses of these sequences by phylogenetic tree and BLAST1 methods revealed that only one sample of each of *P. kurroa* and *R. serpentina* were authentic. BLAST1 and BOLD identification methods, besides substantiating results obtained through analyses of phylogenetic trees, also revealed possible identities of some of the substitutes, though in some cases more than one botanical identity were revealed. The reason for this seemingly ambiguity could be that for the unknown sample, the sequence being tested is not species-specific as it has been revealed that none of the barcode loci provides 100% species discrimination across the plant kingdom, (Kress et al. 2005; Rubinoff et al. 2006). Therefore, the need for a multi-locus barcode was realized by many authors (Chase et al. 2007; Kress et al. 2005; Kress and Erickson 2007; Mishra et al. 2017; Parveen et al. 2017). Though the revealed identities of the samples which were not authentic are tentative, the alarming trend emerging from the results is that substitutions are not according to the well known Ayurvedic principle for rational substitution of drugs known as “Abhava Pratinidhi Dravyas”, propounded by Bhavmishra in sixteenth century A.D (Joshi et al. 2012; Giri 2013). This enunciates that in the case of unavailability ('Abhava') of the original drug ('Dravya'), its representative ('Pratinidhi') that possess similar 'Guna' (qualities) proven on the basis of pharmaco-therapeutically activity can be substituted (Giri 2013). Of the samples which were not found to be authentic, only in two cases, substitutions were according to this principle. Thus, *Thalictrum foliolosum*, sharing the common name of 'Mamira' with *C. teeta*, is an accepted substitute of the latter (Selvam 2012). One of the samples of 'Mamira' analyzed in the present study was identified as *T. foliolosum*, based on BLAST comparison of the sequences generated from ITS, *rbcL* and *rpoC1* markers. Likewise, *rbcL* sequence from one of the samples of *S. chirayita* ('Chiraiyata') matched 100% in identity and query coverage with *A. paniculata* ('Kalmegh'), earlier considered as an adulterant but now as an accepted substitute of *S. chirayita* (Girach et al. 1994; Joshi and Dhanwan 2005). Except these two examples of rational substitutions in case of *S. chirayita* and *C. teeta*, in other cases substitutions were totally by unrelated species. One of the major problems in the authentication of herbal materials using DNA barcoding, is the unavailability of authentic sequences in the GenBank associated with the

Fig. 2 Neighbour-joining tree of *rbcL* sequences of vouchered specimens and the market samples. The sequences of vouchered specimens have been indicated by the botanical names, whereas market samples have been shown with their common names followed by the botanical names of species these are sold as (in bracket). The accession/sample number is given after the botanical names of the vouchered specimens and after the common name of the market samples. Numbers on each node indicate the bootstrap support

Fig. 3 Neighbour-joining tree of *rpoC1* sequences of voucher specimens and the market samples. The sequences of voucher specimens have been indicated by the botanical names, whereas market samples have been shown with their common names followed by the botanical names of species these are sold as (in bracket). The accession/sample number is given after the botanical names of the voucher specimens and after the common name of the market samples. Numbers on each node indicate the bootstrap support

vouchered specimens deposited in the herbarium. The sequences available on NCBI might be from an incorrectly identified plant species and there is no way to verify the specific origin of that DNA (Newmaster et al. 2013). In the present study, to circumvent this problem, BLAST comparisons of the generated sequences were made both on NCBI and BOLD. Moreover, on NCBI only those hits were considered true positive which resulted with the sequences generated for DNA barcoding purpose, as was indicated by the title of the intended publication.

The present study provides and supports the applicability of DNA barcoding for identification and authentication of medicinal plants and their raw drugs. The resulting evidence from the comprehensive study based on 14 plant species, is an ample reflection of the extent of substitutions/adulterations that are prevalent in Indian markets in the samples of rare and endangered plants. In addition, fact that emerges from the present study is about the efficacy of even the authentic samples because of the inability of retrieving DNA sequences of many of the targeted loci from them. In the samples where even the DNA has degraded to a large extent, likelihood of biochemical integrity also becomes doubtful.

Acknowledgements The research work was financially supported by the grants provided to SBB by the Indian Council of Medical Research (ICMR), New Delhi and DST-Purse and Research and Development Grants, University of Delhi. The awards of Junior and Senior Research Fellowships to SM and AP by ICMR and University Grants Commission (UGC), New Delhi, respectively, are gratefully acknowledged.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Barnwell P, Blanchard AN, Bryant JA, Smirnov N, Weir A (1998) Isolation of DNA from the highly mucilaginous succulent plant *Sedum telephium*. *Plant Mol Biol Rep* 16:133–138
- Bell JR (2008) A simple way to treat PCR products prior to sequencing using ExoSAP-IT®. *Biotechniques* 44:834
- Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haider N et al (2005) Land plants and DNA barcodes: short-term and long-term goals. *Philos Trans R Soc Lond Biol Sci* 360:1889–1895
- Chase MW, Cowan RS, Hollingsworth PM, Van den Berg C, Madrinan S, Petersen G et al (2007) A proposal for a standardized protocol to barcode all land plants. *Taxon* 56:295–299
- Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haider N et al (2016) Correction to ‘Land plants and DNA barcodes: short-term and long-term goals’. *Philos Trans R Soc B* 371:20150063
- Chen S, Yao H, Han J, Liu C, Song J, Shi L et al (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* 5:e8613
- China Plant BOL Group (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc Natl Acad Sci USA* 108:19641–19646
- Dawa S, Hussain A, Tamchos T (2013) Medicinal and aromatic plants as a viable livelihood option for the people of Jammu and Kashmir. *Indian Hortic J* 3:40–42
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Fay MF, Swensen SM, Chase MW (1997) Taxonomic affinities of *Medusagynne oppositifolia* (Medusagynaceae). *Kew Bull* 52:111–120
- Ford CS, Ayres KL, Toomey N, Haider N, Stahl JVA, Kelly LJ, Wikstrom N, Hollingsworth PM, Duff RJ, Hoot SB, Cowan RS, Chase MW, Wilkinson MJ (2009) Selection of candidate coding DNA barcoding regions for use on land plants. *Bot J Linn Soc* 159:1–11
- Girach RD, Aminuddin, Khan SA (1994) *Andrographis paniculata*—a possible substitute for *Swertia chirata* in Southeastern India. *Int J Pharmacog* 32:95–97
- Giri CM (2013) Concept of Abhava Pratinidhi Dravyas, a rational substitution of drugs- a review. *Int J Adv Ayush* 2:148–161
- Haider N (2011) Identification of plant species using traditional and molecular-based methods. In: Davis RE (ed) *Wild plants: identification, uses and conservation*. Nova Science Publ, Inc., New York, pp 1–62
- Haider N (2015) The use of DNA based methods for detection of plant based adulterants in plant food products. In: Haynes A (ed) *Advances in food analysis research*. Nova Science Publishers, Inc. ISBN: 978-1-63483-795-8
- Haider N, Wilkinson NMJ (2011) A set of plastid DNA-specific universal primers for flowering plants. *Russ J Genet* 47:1204–1215
- Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucl Acids Symp* 41:95–98
- Hebert PDN, Cywinska A, Ball SL, de Waard JR (2003a) Biological identifications through DNA barcodes. *Proc Biol Sci* 270:313–321
- Hebert PDN, Ratnasingham S, de Waard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci* 270:S96–S99
- Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. *PLoS Biol* 2:e312
- Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, Chase MW et al (2009) CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci USA* 106:12794–12797
- Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. *PLoS One* 6:e19254
- Hussain S, Hore DK (2008) Collection and conservation of major medicinal plants of Arunachal Pradesh. *Indian For* 134:1663–1679
- Joshi P, Dhawan V (2005) *Swertia chirayita*—an overview. *Curr Sci* 89:635–640
- Joshi PR, Patel BR, Shukla VJ (2012) An overview of the causes of current practices in Pratinidhi Dravyas (substitution of drugs) in Ayurveda including newer techniques for their evaluation. *Ayu* 33:481–485
- Keller A, Schleicher T, Schultz J, Muller T, Dandekar T, Wolf M (2009) 5.8S-28S rRNA interaction and HMM-based ITS2 annotation. *Gene* 430:50–57

- Kool A, de Boer HJ, Kruger A, Rydberg A, Abbad A, Bjork L et al (2012) Molecular identification of commercialized medicinal plants in Southern Morocco. *PLoS ONE* 7:e39459
- Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One* 2:e508
- 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
- Kumar JUS, Krishna V, Seethapathy GS, Senthilkumar U, Raghupathy S, Ganeshaiah KN et al (2015) DNA barcoding to assess species adulteration in raw drug trade of “Bala” (genus: *Sida* L.) herbal products in South India. *Biochem Syst Ecol* 61:501–509
- Kumar JUS, Gogna N, Newmaster SG, Krishna V, Raghupathy S, Seethapathy GS et al (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* 130:1457–1470
- Kuzmina ML, Johnson KL, Barron HR, Hebert PDN (2012) Identification of the vascular plants of Churchill, Manitoba, using a DNA Barcode library. *BMC Ecol* 12:25
- Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G et al (2008) DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA* 105:2923–2928
- Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, Sundaresan V (2015) DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. *Plant Biotechnol J* 14:8–21
- Mishra P, Kumar A, Rodrigues V, Shukla AK, Sundaresan V (2016) Feasibility of nuclear ribosomal region ITS1 over ITS2 in barcoding taxonomically challenging genera of subtribe Cassiinae (Fabaceae). *Peer J* 4(4):e2638. <https://doi.org/10.7717/peerj.2638>
- Mishra P, Kumar A, Nagireddy A, Shukla AK, Sundaresan V (2017) Evaluation of single and multilocus DNA barcodes towards species delineation in complex tree genus *Terminalia*. *PLoS One* 12(8):e0182836
- Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S (2013) DNA barcoding detects contamination and substitution in North American herbal products. *BMC Med* 11:222
- Olmstead RG, Michaels HJ, Scott KM, Palmer JD (1992) Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. *Ann Missouri Bot Gard* 79:249–265
- Pang X, Song J, Zhu Y, Xie C, Chen S (2010) Using DNA barcoding to identify species within Euphorbiaceae. *Planta Med* 76:1784–1786
- Parveen I, Singh HK, Raghuvanshi S, Pradhan UC, Babbar SB (2012) DNA barcoding of endangered Indian *Paphiopedilum* species. *Mol Ecol Resour* 12:82–90
- Parveen I, Gafner S, Tehen N, Murch SJ, Abd Khan IA (2016) DNA barcoding for the identification of botanicals in herbal medicine and dietary supplements; strengths and limitations. *Planta Med*. <https://doi.org/10.1055/s-0042-111208>
- Parveen I, Singh HK, Malik S, Raghuvanshi S, Babbar SB (2017) Evaluating five different loci (*rbcL*, *rpoB*, *rpoC1*, *matK* and ITS) for DNA barcoding of Indian orchids. *Genome* 60:665–671
- Raclariu AC, Heinrich M, Ichim MC, de Boer H (2018) Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication. *Phytochem Anal* 29:123–128
- Ratnasingham S and Hebert PDN (2007) BOLD: The barcode of life data system (<http://www.barcodinglife.org>). *Mol Ecol Notes* 7: 355–364
- Ross HA, Murugan S, Li WL (2008) Testing the reliability of genetic methods of species identification via simulation. *Syst Biol* 57:216–230
- Rubinoff D, Cameron S, Will K (2006) Are plant DNA barcodes a search for the Holy Grail? *Trends Ecol Evol* 21:1–2
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467
- Sarwat M, Yamdagni MM (2014) DNA barcoding, microarrays and next generation sequencing: recent tools for genetic diversity estimation and authentication of medicinal plants. *Crit Rev Biotech* 36:191–203
- Saunders GW (2005) Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philos Trans R Soc Lond B Biol Sci* 360:1879–1888
- Selvam ABD (2012) Pharmacognosy of negative listed Plants. *Bot Surv Ind, MOEF, Kolkata*
- Selvaraj D, Shanmughanandhan D, Sarma RK, Joseph JC, Srinivasan RV, Ramalingam S (2012) DNA barcode ITS effectively distinguishes the medicinal plant *Boerhavia diffusa* from its adulterants. *Genom Proteom Bioinform* 10:364–367
- Singh HK, Parveen I, Raghuvanshi S, Babbar SB (2012) The loci recommended as universal barcodes for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species. *BMC Res Notes* 5:42
- Song J, Yao H, Li Y, Li X, Lin Y, Liu C et al (2009) Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. *J Ethnopharmacol* 124:434–439
- Sucher NJ, Carles MC (2008) Genome-based approaches to the authentication of medicinal plants. *Planta Med* 74:603–623
- Swetha VP, Parvathy VA, Sheeja TE, Sasikumar B (2017) Authentication of *Myristica fragrans* Houtt. using DNA barcoding. *Food Control* 73:1010–1015
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Wachtel S, Benzie IFF (2011) Herbal medicine: an introduction to its history, usage, regulation, current trends, and research needs. In: Benzie IFF, Wachtel-Galor S (eds) *Herbal medicine: biomolecular and clinical aspects*, 2nd edn. Taylor & Francis, Boca Raton (**Chapter 1**)
- Warude D, Chavan P, Joshi K, Patwardhan B (2003) DNA isolation from fresh, dry plant samples with highly acidic tissue extracts. *Plant Mol Biol Rep* 21:467
- White TJ, Burns T, Lee S, Taylor J (1990) Amplification and sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) *PCR protocols: a guide to methods and applications*. Academic Press, San Diego, pp 315–322
- Boldsystems.org (homepage on the internet). Ontario: Ministry of Research and Innovation. http://www.boldsystems.org/index.php/IDS_openIdEngine
- Nmpb.nic.in [homepage on the internet]. New Delhi: Government of India (Ministry of Ayush). Available from: <http://nmpb.nic.in/index.php>
- Yuan QJ, Zhang B, Jiang D, Zhang WJ, Lin TY, Wang NH et al (2015) Identification of species and *materia medica* within *Angelica* L. (Umbelliferae) based on phylogeny inferred from DNA barcodes. *Mol Ecol Resour* 15:358–371