Chapter 4
DNA Barcoding of Medicinal Plants

Swati Srivastava, Sanchita, Mili Bhargava, and Ashok Sharma

Abstract DNA barcoding is a useful technique for diversity analysis whereby a standardized region of DNA is used for the identification of a species or a taxonomic group of organisms. These standard regions used for identification are called the DNA barcode. These are small sequences of the entire genome. In plants, DNA barcoding has application in phylogenetic analysis, authentication, inter- and intra-specific diversity, classification into wild and cultivated genotypes, the study of phylogeographical patterns, and in the detection of adulteration. The barcode loci, i.e., the DNA regions used for the identification are able to discriminate the closely related species and identify new cryptic species as well. Depending on the taxon and complexity of the species, different barcode loci are used for the purpose. In animals, the universal DNA barcode, i.e., mitochondrial cytochrome c oxidase I (COI) gene is used for species discrimination. However, this gene cannot be used for plants due to its limited divergence. Thus, its use is limited only to some algae. Efforts are going on to find suitable universal barcode loci for plants. Since the last decades, matK, rbcL, trnH-psbA, ITS, trnL-F, 5S-rRNA, and 18S-rRNA candidate regions are being used as DNA barcodes in plants. The article provides an overview of the use of these candidate regions through different approaches which have gained importance due to the challenges in DNA barcoding of plants. The development of multilocus and tiered approaches along with the new frontier areas for application of this technique has been analyzed in detail.

Keywords DNA barcoding • Forensics • Herbal drugs • Medicinal plants • Multilocus • Single locus • Tiered

S. Srivastava • Sanchita • M. Bhargava • A. Sharma (✉)
Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow 226015, India
e-mail: ashoksharma@cimap.res.in

Abbreviations

BOLD   Barcode of Life Data Systems
CBOL   Consortium for the Barcode of Life
iBOL   International Barcode of Life Project
WHO   World Health Organization

4.1   Introduction

Medicinal herbs have been used in Indian medicinal systems since ancient times. Currently, authentication of medicinal plants is a big issue. Trading of medicinal plants and its products worldwide is estimated at around US$60 billion, and an annual turnover of Ayurvedic medicine in the international market is about Rs. 3500 crores (US$813 million) (Biswas and Biswas 2014). According to the World Health Organization's (WHO) guidelines, authenticity, purity, and safety are important aspects of standardization and evaluation of traditional medicines. Due to commercialization and increased demand of Ayurvedic herbs, safety, quality, and assurance are big issues (Chan 2003). Taxonomists are busy in naming and annotating the huge number of organisms constituting the biodiversity. A large variety of species are measured annually. However, there still remains an enormous diverseness to be explored. The current scenario of extinction and conservation rates of biodiversity is also a serious concern (Costello et al. 2013).

In the last decade, DNA sequences are being extensively used in the biological process analysis such as phylogenetic analysis of organism identification. Among various approaches, the DNA barcoding was proposed to overcome the problems faced in the traditional taxonomy (Hebert et al. 2003a, b). This approach has succeeded in the identification of already existing as well as unknown species. In this technique, a standard region of DNA known as "DNA barcode" is used for the biodiversity analyses. Different regions of DNA are used as markers for DNA barcoding. Two main characteristics of a good marker are its universality and high resolution (Hollingsworth et al. 2011). The universality of any region refers to the applicability of the chosen DNA barcode to a large number of organisms. High-resolution ability implies that the markers must discriminate the closely related species. For efficient discriminatory power, a marker must show high interspecific and low intraspecific divergence. This distinction between inter- and intraspecific distances is known as the "DNA barcoding gap." The DNA barcoding is a widely used technique for quick and accurate identification of species (Bhargava and Sharma 2013). COI (cytochrome oxidase I) is the universal barcode marker in animals (Hebert et al. 2003a). However for plants, it has remained elusive (Li et al. 2015; Kress and Erickson 2008). COI has shown a good success rate in animals but in plants, due to limited divergence, it cannot be used. There has been much debate about the regions to be used as barcodes for plants (Hollingsworth et al. 2011).

Presently, DNA barcodes work on the standard gene of any locus as well as in whole chloroplast genome in plants. Although different kinds of genome-based strategies are developed for the identification, DNA barcoding is the most powerful tool. Approximately 300,000 plant species are available worldwide. The identification and classification of such a vast range of plants may be a difficult task for taxonomists. DNA barcoding helps in a rapid and accurate identification of plant species (Costion et al. 2011). DNA-based methods are more suitable as compared to proteins and RNAs because DNA is available in all the tissues of the organisms is more stable, and remains unaffected by external factors. The species discrimination in plants is difficult because of a higher level of gene tree paraphyly (Fazekas et al. 2008). matK, rbcL, trnH-psbA, ITS, trnL-F, 5S-rRNA, and 18S-rRNA are majorly used markers for plants with regard to their discrimination capacity (Table 4.1). Cowan and Fay (2012) have described the major challenges associated with DNA barcoding of plants. However, the studies on plant barcoding are increasing consistently due to its capability of identifying the unknown samples. A general concept of the formation of DNA barcode has been explained in Fig. 4.1.

4.2 History of DNA Barcoding and Success Stories

Two international initiatives are operating for the DNA barcoding analyses, viz., the International Barcode of Life project (iBOL) and Consortium for the Barcode of Life (CBOL). iBOL is the biggest biodiversity genomics initiative ever undertaken. Its mission is to maintain and update the barcode reference library, Barcode of Life

Table 4.1 Major candidate regions used for DNA barcoding of plants

| Gene	Location/genome	Type	No. of GenBank accessions (plants)	No. of GenBank species (plants)
matK[a]	Plastid	Protein coding	~61,172	~33,089
rbcL[a]	Plastid	Protein coding	~86,806	~38,164
rpoB[a]	Plastid	Protein coding	~8328	~4399
rpoC1[a]	Plastid	Protein coding	~8650	~4787
trnH-psbA[a]	Plastid	IGS	~15,304	~4912
atpF-H[a]	Plastid	IGS	~370	~295
psbK-I[a]	Plastid	IGS	~3223	~1789
trnL-F	Plastid	Intron + IGS	~16,608	~7344
5S rRNA	Nuclear	Structural RNA	~9954	~2290
18S rRNA	Nuclear	Structural RNA	~40,291	~20,150
ITS	Nuclear	Transcribed spacers+5.8S rRNA	~179,520	~70,085

[a]CBOL proposed seven candidate barcodes
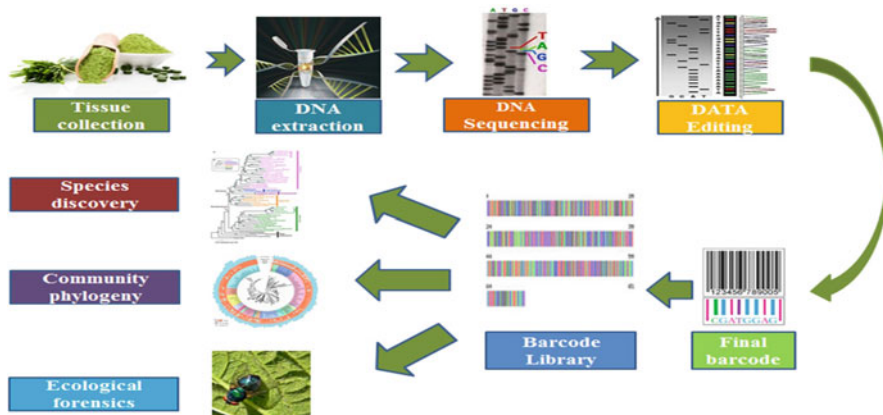
Fig. 4.1 Schematic representation of the formation of DNA barcodes

Data systems (BOLD), and further establishment as a robust resource for animal, plant, and fungal DNA barcodes (Ratnasingham and Hebert 2007). The work of the iBOL association is carried out by its constituent nodes, comprised of many counties grouped into separate operating teams. On the other hand, CBOL established in 2004 is functioning for DNA barcoding as a world methodology for the identification of plants and animals of earth's biodiversity. As far as the application of the technique is concerned, recently its usefulness has been explored in forensic botany to resolve the legal questions. Plant identifications at crime scenes are important in the criminal investigations. Every environment has a unique combination of pollens, suggesting the type of place where the crime took place. To overcome the problem related to ancient forensic botany, DNA barcoding could be a promising technique in several cases (Ward et al. 2005, 2009; Tsai et al. 2006, 2008; Ferri et al. 2009).

4.3 Status of DNA Barcoding of Medicinal Plants

Adulteration is a major problem in the herbal plant material market. Therefore, authentication and standardization is the prerequisite to minimize the unfair trade. According to the World Health Organization (WHO) total international seasoning, the drug market is calculated as US$62 billion and is anticipated to grow to the extent of US$5 trillion by the year 2050. The total available barcodes represent 363,584 sequences from 50,039 species. The criteria of DNA barcoding, i.e., minimum sequence length of 500 bp and more than three organisms per single species have been convincing by 13,761 species (Sarwat and Yamdagni 2015). However, most of these (98%) are animal species. In January 2009, iBOL started with the target to collect barcodes for 5 million species in first 5 years. The scientists from

25 countries have contributed to this initiative (Hajibabaei et al. 2006). The DNA barcoding project has the goal of the reference library development that might provide data even for very low taxonomic level with short and specific DNA fragments. The major efforts are underway for barcoding of medicinal and aromatic plants worldwide (Cowan and Fay 2012; Elliott and Jonathan 2014). However, very little work has been reported for barcoding of Indian medicinal plants. India carries 7–8 % of world biodiversity with excessive resources of medicinal plants (45,500 approx.). Out of those, 8,000 plant species are of medicative worth, and 960 species are considered in a trade. Out of that, 178 species have a yearly consumption of more than 100 metric tons (Aneesh et al. 2009; Efferth and Greten 2012). The demand of the medicinal plants at industrial level is higher due to its global growth within the herbal industries. Thus, the Indian market is a center of herbs with the calculable trade of US$140 million annually. The botanical and natural ingredients export worldwide was more or less US$33 billion throughout 2010, and it was expected to reach US$93 billion by 2015 according to the December 2011 bulletin of Market News Service. The export of Indian medicinal plants and their products is estimated to be about $0.2 billion. In addition to the international trade, there is a considerable volume of international exchange of medicinal plants in India with a turnover of $1.6–$1.8 billion (Mishra et al. 2015). Total world seasoning herbal market is of the scale of $60 billion yearly with India's contribution of 2.5 %. Thus, in spite of having an extensive heritage of Ayurvedic literature and a good variety of medicinal plant species, India is still struggling with the potential market demand (Mishra et al. 2015). For increasing the India's share in the global herbal market, the improvement in quality control, standardization, scientific ways of production, and analysis of business products is necessary. The standardized mass produce of herbal products tested scientifically would not only maintain the efficacy of the herbals but also offers a competing edge to other medicines. China is presently leading the efforts on DNA barcoding of medicinal plants and has developed the database of DNA barcodes (Lou et al. 2010). Some reports are also available for DNA barcoding of Indian medicinal plants (Parveen et al. 2012; Ghosh et al. 2013). Due to importance and demand of herbal raw materials and products, the herbal industry suffers from substitution and adulteration of medicinal plants with its closely related species. Adulteration and mixing cause major changes in formulation and are also considered as illegal practices. The efficacy of any drugs/herbal product decreases when the herb is adulterated, and sometimes it could be lethal if it is substituted with toxic adulterants. The correct formulation is important for the medicinal herb to be effective. The main source of income of herbalist is the trading of medicinal plants. The economic constraints might offer an incentive for herbalists to substitute rare ingredients with cheaper and a lot of pronto offered species. Due to the illegal overtrading of medicinal plants, many plant species have become endangered in India. Therefore, to avoid these practices, some identification tags are required to detect plant materials. DNA barcoding is a useful tool for the discrimination of raw materials of medicinal plants.

4.4   Different Approaches of DNA Barcoding in Plants

Due to the complexity of DNA barcoding in plants, may it be amplification, sequencing, or a significant "barcoding gap," the technique has demanded ample attention toward the improvement of the methodology of the identification process. Thus, approaches like combining multiple barcodes at the totally different taxonomic group or multiple combinations of barcodes in a tiered fashion, such as a particular combination of one taxonomic group followed by a more robust combination at the next level, have recently gained importance. These approaches are mentioned below to introduce to the readers with the current methods in DNA barcoding in plants.

4.4.1   Single-Locus Approach

Due to the differences in the efficiency of barcoding markers in discriminating plants of different families, individual markers have been comparatively evaluated in a number of families (Gao et al. 2010a, b; Hollingsworth et al. 2009; Li et al. 2012; Muellner et al. 2011; Pettengill and Neel 2010). matK is the nearest plant analogue to COI, the animal DNA barcode. It typically provides high resolution, leading to good species identification as a result of its speedily evolving coding fragment among the plastid genome (Lahaye et al. 2008). However, the disadvantage of this barcode marker is due to unavailability of universal primer sets for all taxa. It creates a problem in PCR resulting in low PCR amplification particularly in non-angiosperms (Kress and Ericsson 2007; CBOL Plant Working Group 2009). As compared to matK, the barcode marker rbcL (ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit) is easy to amplify and sequence. It is an important candidate for plant DNA barcoding even though its discriminatory power is not as good as matK. matK and rbcL have been suggested to be the core DNA barcodes for plants (Hollingsworth et al. 2009). Other than these, the plastid intergenic spacer trnH-psbA is also used as a supplementary DNA barcode. It has higher species discrimination success and variable intergenic spacers in plants (CBOL Plant Working Group 2009; Liu et al. 2012a, b). The main focus related to this locus includes the high frequency of mononucleotide repeats that cause simplex reads (Devey et al. 2009) and thus hamper the recovery of bidirectional sequences. Another common event encountered during this region is the microinversion. Microinversions in trnH-psbA have been studied in various angiosperms (Whitlock et al. 2010). Thus, the advanced design of trnH-psbA makes it tough to use as a barcode (Storchova and Olson 2007; Hao et al. 2010). However, the additional characters for species discrimination are provided by uncorrected microinversions (Jeanson et al. 2011). Although demonstrated a positive impact on species, discrimination by manually correcting inverted sequences has been demonstrated (Whitlock et al. 2010). A comprehensive analysis of the utility of trnH-psbA and its mixture has been studied (Pang et al. 2012). Many researchers have been concerned with the utilization of

nuclear internal transcribed spacer (ITS) region in the form of the standard barcode and have recommended one of the core barcodes for seed plants (Li et al. 2015). Subsequent to this study, part of this region, ITS2 was also suggested to be a novel barcode for both plants and animals (China Plant BOL Group 2011; Yao et al. 2010; Chen et al. 2010). On the basis of its performance in phylogenetic studies (Baldwin 1992), resolving power of ITS was not underestimated; however, the following three major problems were encountered in its use.

4.4.1.1 Sequencing

One of the main limitations for nuclear ITS is its recovery, since the amplification and sequencing are difficult for this region (Kress et al. 2005). An alternative to it is the use of ITS2 that is less complicated to work with due to the small length of target fragment which makes amplification and sequencing easier.

4.4.1.2 Paralogous Gene Copies

Nuclear ITS fragment is available in multiple copies in the cells. Concerted evolution of multiple copies leads to divergence co-occurring in the individuals (Alvarez and Wendel 2003; Bailey et al. 2003). This can lead to rendering the sequences unreadable resulting in messy sequences. This paralogy phenomenon may lead to misidentification of samples because depending upon the variant sequence, the species will be identified subsequently. However, the identification of region, compared to other markers, is not compromise due to the availability of paralogous copies (Hollingsworth et al. 2011).

4.4.1.3 Fungal Contamination

The fungal ITS regions represent similarity with their plant correspondent. The primers considered for amplification as well as sequencing are very similar. Thus, the fungal DNAs are amplified by accident in several cases, particularly for the plants containing fungal endophytes. This can lead to misidentification of samples. Therefore, no matter the quantity of primer sets obtainable for this explicit barcode region, amplification and sequencing have been hard for numerous samples (Gonzalez et al. 2009).

4.4.2 Alternate Regions as DNA Barcodes

Besides the core plastid markers (matK, rbcL), the supplementary trnH-psbA, and ITS regions, there are other plastid protein-coding genes (rpoB, rpoC1), plastid intergenic spacers (atpF-H, psbK-I), and low-copy number genes being tested for identification in several families (Pillon et al. 2013).

4.4.3 Multilocus Approach

The multilocus approach is an adequate method for DNA barcoding of plants with good discrimination success (Kress and Erickson 2007; Fazekas et al. 2008; Newmaster et al. 2008; CBOL Plant Working Group 2009). The practice of using multiple barcodes has emerged in view of the unsatisfactory performance by individual loci. The high discrimination-related results could be obtained through combining the universality, discriminatory power, and amplification success of each locus. Multilocus combinations also promote high clade support values in monophyly-based identification of samples as in the case of Nyssaceae (Wang et al. 2012). Any plant barcode can be a combination of two or more locus. One of them may be a conservative coding region like rbcL and the other, a rapidly evolving noncoding region. The noncoding trnL intron and trnL-F intergenic spacer (IGS) are recommended for situations involving extremely degraded tissue (Taberlet et al. 2007). In bryophytes, the power of this region has been tested (Quandt et al. 2004; Stech et al. 2013). Thus, trnL-F and trnL regions were further used successfully for distinguishing the mysterious aquatic fern gametophyte (Li et al. 2009). Even a project on two-locus DNA barcode for plants (matK + rbcL) has been proposed by the CBOL Plant Working Group (2009). In some cases, the combination of three loci has failed to improve the discrimination better than two-locus barcodes in few cases (Wang et al. 2012). To avoid expenses of using a three loci combination for large data sets, the two-locus barcode was accepted as the standard barcode for land plants (CBOL 2009). In the case of two-locus, the preserved coding loci align well with the taxa of a community sample to determine deep phylogenetic branches. The hypervariable region of the DNA barcodes can align with ease in the subclades of closely related species (Kress et al. 2009). The complementation of rbcL gene and the noncoding trnH-psbA spacer region has been demonstrated (Fazekas et al. 2008). In contrast to CBOL, they suggested the use of more than two regions because of the decreased discrimination identified in barcoding analyses with three or more regions. This concept would also be beneficial when some of the loci recovered are of bad quality. Another efficient work using rbcL in combination with trnL-F for ferns has also been demonstrated and shown great potential for species discrimination (de Groot et al. 2011). The composition vector (CV) approach (Qi et al. 2004) has been described as an efficient method for analyzing rRNA data sets. The changed CV methodology incorporates an adjustable weighted algorithmic program for the vector distance as per the magnitude relation of sequence length

found between a pair of taxa within the candidate genes. Recently, changed CV approach has been applied for studying huge multigene datasets for plant DNA barcoding (Li et al. 2012).

4.4.4 Tiered Approach

Combining barcode markers for discriminating species is robust and has high support values. A newer approach for combining the barcode, known as the tiered approach, is also evolving. Although vigorous efforts are going on to find suitable universal loci for plants, and there may be one in future, but relying on a single locus for plants will still be a bad choice. This is because of hybridization and introgression observed in some group of plants. Therefore, rather depending on maternally inherited genes, using a combination of both coding and noncoding regions in a stepwise manner will be the favorable and logical approach. It permits an unknown sample to be allotted at a taxon level, where a successful pair of primers can be targeted. Among a small group of taxa, the samples are aligned first followed by final assignment. In a specific taxonomic group, only a few studies have tested this methodology (Newmaster et al. 2006; Xiang et al. 2011). The first tier coding region, common in plants, has been used for differentiation at a definite taxonomic level, followed by a lot of variable second tier coding or noncoding region at the species level. Alignments at first tier (coding regions) would decrease the problem of aligning more divergent genera using noncoding regions at the second tier. So, under a common first tier sequence, the dataset will be properly organized to perform well at the next level of resolution. The method also preserves the efficiency of the previous multilocus approach since the complement regions for a group can still be used in this method. The rbcL has been considered as a primary tier barcode (Newmaster et al. 2006). Although it is the most identified plastid coding region in GenBank, covering a majority of groups and thus can work as a platform for comparison of different plastid genes. rbcL was analyzed to see how well it resolves congeneric species. This marker might be used for resolving congeneric species (85 % cases), so it should be used as the core first tier locus, followed by a choice of a secondary locus at the second tier. The method, therefore, provides flexibility in the choice of the next locus after a standard common region is used at a particular level. Similarly, this approach has been supported by Xiang et al. (2011) and also recommended that the use of matK at the generic level with further resolution at the second tier needs to be explored with a suitable second tier locus.

4.5 Bioinformatics Approaches

Bioinformatics play an important role in DNA barcoding analyses. The DNA barcoding processes depend on the availability of information in the form of data. If the data are available for query, we can use bioinformatics tools for the analysis of

Table 4.2 Different software and tools used for DNA barcoding

Categories of software and tools	Name of software and tools	References
Alignment-free and character based	BRONX (Barcode Recognition Obtained with Nucleotide eXposés)	Little (2011), Zhang et al. (2008), and Bertolazzi et al. (2009)
	BPSI (Back Propagation Species Identification)	
	BLOG (Barcoding with LOGic formulas)	
psbA-trnH based	PTIGS-Idit	Liu et al. (2011)
Distance based	TaxI	Steinke et al. (2005)
Character based	CAOS (Characteristic Attribute Organization System)	Sarkar et al. (2008)
Oligonucleotide frequency based	Oligonucleotide Frequency Barcode Generator (OFBG)	Tyagi et al. (2010)
Simultaneous sequence and structure alignment	4SALE	Seibel et al. (2006)
CBC detection	CBCAnalyzer	Wolf et al. (2005)
OTU clustering and annotation	jMOTU and taxonerator	Jones et al. (2011) and Kumar et al. (2011)
	CLOTU	
Identification of new barcode markers and associated PCR primers	ecoPrimers	Riaz et al. (2011)
Analysis of discrimination capacity of individual markers	TaxonGap	Slabbinck et al. (2008)
DNA barcoding based on Bayesian phylogenetics	Statistical Assignment Package (SAP)	Lou et al. (2010)
Analysis of resultant data	OTUbase	Beck et al. (2011), Brown et al. (2012) and Little (2010)
	SPIDER (SPecies IDentity and Evolution in R)	
	Barcode Quality Index (B)	

barcode data. After the collection of corresponding query sequences, sequence analysis and phylogenetic construction are performed. Sequence analysis basically involves the query and reference dataset sequence alignments. Some of the MSA programs ClustalW, T-Coffee, and MUSCLE, etc. are used for sequence analysis. In silico innovation approaches for DNA barcoding have been developed on the basis of compensatory base changes (CBCs) (Wolf et al. 2005), operational taxonomic units (OTUs) (Slabbinck et al. 2008), DNA metabarcoding (Riaz et al. 2011), locus-specific tools (Liu et al. 2011), tool for representing barcode symbology (Liu et al. 2012a, b), neural network techniques (Zhang et al. 2008), machine learning (Zhang et al. 2012a, b), data mining (Bertolazzi et al. 2009), composition vector (Kuksa and Pavlovic 2009), etc. The available software and tools analyzing the barcode data are given below (Table 4.2).

4.6 Limitations

4.6.1 The Absence of Universal Barcode and Selection of Appropriate Barcode Region

In DNA barcoding, the universality of the barcode is still a big problem. It is difficult to attain the universality of barcode due to the insufficient information of genetic variation in the less-studied taxonomic group. This problem is majorly found in plants as compared to animals. The differentiation and identification of species relying on interspecific variation among DNA sequences are due to the resolution capability of a barcode. Thus, there is a challenge in defining a good quality barcode consisting of a small and variable DNA sequence flanked by conserved regions (Hebert et al. 2003b; Moritz and Cicero 2004; Rubinoff et al. 2006a; Ficetola et al. 2010). The most important task of DNA barcoding is the identification of universal primers amplifying fragments with high resolution. However, it has been argued that a single short fragment will be sufficient to discriminate the organism at species level identification (Ficetola et al. 2010; Rubinoff et al. 2006a; Moritz and Cicero 2004). The single-locus DNA barcodes lack adequate variation in the closely connected taxonomic group, so for the identification of plants, no loci are available (Li et al. 2015).

4.6.2 Error Found in DNA Barcoding when Mitochondrial Sequences Are Used

DNA barcoding faced limitation due to the presence of the same copy of a gene of interest in the mitochondrial genome because of heteroplasmy in mtDNA, bacterial infection biasing, nuclear integration, and introgression in mtDNA. The duplication of a gene, i.e., if a portion of cytochrome oxidase I (COI) are duplicated in a given species, typical PCR may amplify these fragments. Thus, it will not be clear whether the paralogous copy had diverged from duplication of COI (Campbell and Barker 1999; Song et al. 2008). The heteroplasmy is the combination of more than one type of mitochondrial genome in a species. The overestimation of the quantity of distinctive species in barcoding results due to occurrence of co-amplification in divergent heteroplasmic copies of mtDNA (Rubinoff et al. 2006b; Song et al. 2008; Fišer Pečnikar and Buzan 2014; Magnacca and Brown 2010; Moulton et al. 2010; Valentini et al. 2009; Acs et al. 2010; Hurst and Jiggins 2005). The bacterial infection found in mtDNA due to the maternally inherited symbionts can cause linkage disequilibrium, and each individual becomes infected with such symbionts. These symbionts among closely connected species break the species barrier by conjugation followed by selective sweep leading to the identical mtDNA sequences in different species (Song et al. 2008; Whitworth et al. 2007). The nuclear integration of mtDNA creates error for barcoding. Nuclear mitochondrial pseudogenes (numts)

are a nonfunctional duplication of mtDNA in the nucleus and occur in the major clades of eukaryotes. The presence of numts in the nuclear region creates a problem in DNA barcode data library construction and species identification. The potential existence of COI numts causes a major problem to DNA barcoding (Bensasson et al. 2001; Richly and Leister 2004; Song et al. 2008; Zhang and Hewitt 1996). The introgression in mtDNA also creates a problem for barcoding. Introgression is the process of transfer of a gene from one species into the gene pool of other species through recurrent backcrossing of an interspecific hybrid with one among its parents. It causes confusion in species boundaries between evolutionary lineages (phylogenies) that might commonly be divergent (Rubinoff 2006). In meta-analysis of phylogenetic studies, it was found that over 20 % of the studies lineages present problem due to mtDNA introgression (Ballard and Whitlock 2004; Fišer Pečnikar and Buzan 2014; Rubinoff 2006; Valentini et al. 2009; Vences et al. 2005; Acs et al. 2010; Hurst and Jiggins 2005; Machado and Hey 2003). There are limitations of using mtDNA in infer species boundaries with the retention of ancestral polymorphism, male-biased gene flow, and selection on any mtDNA nucleotide (the whole genome is one linkage group). The introgression along with hybridization and paralogy results in the transfer of mtDNA gene copies to the nucleus (Hebert et al. 2004; Ballard and Whitlock, 2004; Bensasson et al. 2001). These factors in mtDNA create a problem for both animal and plant DNA barcoding.

4.6.3 Lack of Comprehensive Reference Database

DNA barcoding is affected due to incomplete a priori identification of specimen in the reference database. The conflictions are created in data assessment; different laboratories work on the same taxa and explain different nomenclatures of the same species through morphological identification (Becker et al. 2011; Collins and Cruickshank 2013). If the reference database is not comprehensive, it will create misidentification of the taxa (Meyer and Paulay 2005; Valentini et al. 2009). DNA barcoding faces limitations when the selected individual represents to every taxon within the reference database. The unknown specimen taken from undescribed biodiversity causes problems in the identification (Fišer Pečnikar and Buzan 2014; Rubinoff 2006). The reference sequences from taxonomically verified specimen lead to the validity of DNA barcoding. In the absence of the reference data, DNA barcoding will face limitations and challenges (Ajmal et al. 2014). DNA barcoding will also face difficulty when the query sequence lacks its target in the reference database. Therefore, the barcoding-based identification of the query at the species level fails (Nielsen and Matz 2006; Virgilio et al. 2010). The reference sequences are verified from voucher specimen that is documented by experienced taxonomists. Due to lack of reference database, there will be no authentic library for recently identified query sequences. As a result, there will be a large quantity of legacy data in the form of sequences that are available in GenBank. These will not be used as a

barcode. Thus, DNA barcoding does not improve the speed of cataloging the life on earth (Taylor and Harris 2012; Peterson et al. 2007).

4.6.4 Lack of Statistical Solution

DNA barcoding is a useful tool for the identification of unknown species. For this methodology, the threshold values providing a distinction between intraspecific and interspecific variation values are required. If the unknown sequence differs from the nearest reference sequence by a variation above the threshold, the organism containing the sequences will belong to a specific species, suggesting its classification needs additional investigation. The wide range of overlap between intra- and interspecific divergence values creates major problems. These overlaps seem comparatively restricted and far from the respective average values. Thus, only the mean values for intra- and interspecific comparisons of closely connected sibling species are required (Desalle 2006; Hebert et al. 2003a; Prendini 2005; Rubinoff 2006; Taylor and Harris 2012; Valentini et al. 2009; Vences et al. 2005; Casiraghi et al. 2010; Frézal and Leblois 2008). The use of a different threshold considering the tenfold rule (gap corresponds to a generic ten times the value of intraspecific divergence) has been proposed. This law has been extensively criticized (Meyer and Paulay 2005; Moritz and Cicero 2004; Matz and Nielsen 2005; Nielsen and Matz 2006; Valentini et al. 2009). To overcome this difficulty, currently it is predicted that the interspecific sequence divergence should increase to the threshold of 2 or 3 % dissimilarity. This threshold has been set on the basis of experimental proof observation of sequence variations among congeneric species (Hebert et al. 2003b). This approach might be simple to neglect the inconclusive or inaccurate results. Thus, there is a requirement of for statistical strategies when a sampled query sequence is the same as the specific database sequence to proof a species assignment of the query (Nielsen and Matz 2006). The strong assumptions based on the population genetics of the analyzed species revealed the statistical uncertainty in DNA barcoding (Nielsen and Matz 2006). The unrealistic assumption of excellent sequence identity at intraspecies level is abandoned. Thus, with not creating population genetic assumptions, the DNA barcoding is not possible (Acs et al. 2010; Hickerson et al. 2006). It has been observed that with robust population subdivision within species, the species assignment might fail due to the underlying demographics that have not been modeled capably. Another case is sequence sampled from a subpopulation with no gene flow with any of the population listed in the database. The DNA barcoding statistical methods which are used here do not categorize the query sequence as a member of the parental species, even though taxonomists would identify it as belonging to it. So, DNA barcoding might fail as a result of the unrecognition of taxonomical units corresponding to a population that is reproductively isolated and additionally if centered on a range of nucleotide changes (K) as a statistics within the hypothesis-based mostly approach. Nielsen and Matz (2006) have urged that a procedure that examines a number of nucleotide changes only between

a query sequence and its best match within the information match in the database is not optimal. During this series for DNA barcoding, two ways, K-test and Bayesian check, are developed from a perspective of applied mathematics genetics. The performance of K-test faces drawback if some species were missing from the information and such behavior might lead to incorrect assignment of queries derived from these "unrecorded species." On the other hand, Bayesian theorem is used as an advantage upon the K- test in terms of accuracy and ability to face the negligence with more than one sequence per species in the database. However, this methodology is also with difficulties of significant phylogenetic assumptions and species level assumptions which are not always correct. Therefore, the convenience of a full Bayesian theorem might not eliminate the necessity for illation procedures with controlled frequentist (hypothesis-based) properties (Nielsen and Matz 2006). Still, DNA barcoding faces the problem to check the clear hypothesis meaning alternative of inappropriate or suboptimal analytical technique because of confusion on the objectives of the study.

4.6.5   Limitation of Distance-Based and Tree-Building Method Used in DNA Barcoding

In some reports, it is noted that DNA barcoding fails in the form of taxonomic approach because it does not recover correct species tree (Hebert and Gregory 2005; Will and Rubinoff 2004; Rubinoff et al. 2006a). Some criticism has arisen due to distance-based and character-based methods. Some reports have mentioned that the distance-based method should not be used for DNA barcoding, as it is a phenetic measure and is not appropriate for species identification (Casiraghi et al. 2010; Desalle 2006; Meyer and Paulay 2005). In the distance-based method, NJ tree acts as a standard part of the procedure for DNA barcoding (Casiraghi et al. 2010; Collins and Cruickshank 2013). But, there is a good documentation about the poor performance of NJ trees on the basis of trial and error and theoretical (Collins and Cruickshank 2013; Little 2011; Meier et al. 2006; Virgilio et al. 2010; Zhang et al. 2012b). The inappropriate use of NJ trees for identification can decrease the effectiveness of DNA barcoding. This will ensue either mtDNA paraphyly or misidentification of species independently. The NJ trees do not seem to be resolved by exploitation the other tree inference ways (Desalle 2006; Desalle 2007; DeSalle et al. 2005; Lowenstein et al. 2009; Rubinoff et al. 2006a; Taylor and Harris 2012; Austerlitz et al. 2009; Collins and Cruickshank 2013; Kerr et al. 2009; Little 2011; Little and Stevenson 2007; Lowenstein et al. 2009; Srivathsan and Meier 2012; Virgilio et al. 2010; Zhang et al. 2012b; Collins et al. 2012; Collins and Cruickshank 2013; Will et al. 2005). On the other hand, the character-based methodology are used to identify the nucleotide combinations (Collins and Cruickshank 2013). The character-based methodology have failed to break into the most stream of DNA barcoding (Savolainen et al. 2005). However, currently, DNA barcoding via

tree-based approach did not stop at distance vs. character based approach. Avoiding any tree-building analysis due to its impression of inferring phylogenies and relationships with single-gene tree is well known as a problem of phylogenetics (DeSalle et al. 2005; Taylor and Harris 2012). Generally, the phylogenetic technique has been proposed as a data analysis in order to overcome the limitations of the threshold-based methodology in DNA barcoding. However, the application of these threshold-based approaches leads to some problem in a study on the relationship between DNA barcoding and molecular phylogeny. DNA barcoding is not a phylogenetic reconstruction. Still, these methods are being used along with the debate in phylogeny and identification in the area of DNA barcoding (Casiraghi et al. 2010; Moritz and Cicero 2004; Vogler and Monaghan 2007). The bootstrap resampling can further decrease the already low identification success rates associated with NJ trees (Brown et al. 2012; Collins and Cruickshank 2013; Fujita et al. 2012; Meyer and Paulay 2005; Monaghan et al. 2009; Puillandre et al. 2012; Virgilio et al. 2012; Zhang et al. 2012a). The use of bootstrap resampling in DNA barcoding studies creates confusion between species discovery and specimen identification. Bootstrapping in this situation also helps in addressing the problem with NJ trees, such as taxon-order bias and tied trees (Lowenstein et al. 2009; Meier et al. 2008). Use of bootstrap value as a cutoff for correct identification severely compromises the efficacy of a reference library and exacerbates the previously outlined weaknesses of using tree-based methods in general (Collins et al. 2012; Collins and Cruickshank 2013; Zhang et al. 2012b).

4.6.6 Limitation in Available Bioinformatics Tools and Algorithm

The biases occurred in methods used for the original cohort of DNA barcoding are being replicated by various studies and assisted by the analytical tools obtainable from the BOLD. A character-based tool, i.e., BLOG, has been made along with BOLD. But, presently it is available only on the Barcode of Life Data Portal (BDP) instead of various BOLD websites. The current popular methods could be a product of routine instead of wise selection. This means a systematic appraisal of taxa has not been capitalized by the barcoding movement. For DNA barcoding, easy-to-handle tools are required for species discrimination and identification. These tools use pairwise global alignment or alignment-free and automated selection of data partitions of an alignable group of samples (CBOL Plant Working Group 2009; Chu et al. 2009; Kress et al. 2009; Kuksa and Pavlovic 2009; Hollingsworth et al. 2011). Microinversions are common in noncoding regions leading to multiple groupings of samples.

4.6.7   Absence of Effective Bioinformatics Pipeline for DNA Barcoding

DNA barcoding is being used to recognize and identify the unknown species. Thus, despite its present limitations, the barcoding method provides a pipeline for the survey of biodiversity, a crucial task for prioritizing conservation efforts, given the present extinction crisis (Taylor and Harris 2012; Valentini et al. 2009). As mentioned earlier, there is a huge amount of sequence information stored in GenBank for which there are no voucher specimens, excluding this sequence from use as a barcode. During the DNA establishment of a barcoding reference library, there will be different unsampled taxa varying the depth of sample coverage for some markers. So, it is necessary to develop the bioinformatics framework having the access to select the sets of samples, directly comparable for a given set of markers. The integration of analytical approaches into a single easy-to-use workflow is required to provide comparable bioinformatics support for multi-marker barcoding in animals and plants (Hollingsworth et al. 2011; Bhargava and Sharma 2013).

4.7   Successful Uses of DNA Barcoding in Medicinal Plants

In many studies from 2003 to 2016, the results of DNA barcoding can provide accurate identification of many medicinal plant materials that are not morphologically distinguishable. DNA barcoding has found its applications in several areas like forensic science, biosecurity (Armstrong and Ball 2005), tracing of illegal trading of organisms (Galimberti et al. 2014), and pharmaceutical and herbal industries, among others (Gantait et al., 2014). When there is an insufficient morphological or anatomical data for the identification of a sample, a stretch of DNA sequence might be helpful in identifying a species. Samples with multiple fragments can now provide multiple species identification, giving a clear picture of habitat that offers a critical clue to the investigators (Ferri et al. 2015). DNA barcoding works with different identification fields and gives more accurate results of medicinal plant identification, i.e., DNA barcode identification with chemical analysis (Palhares et al. 2015) and next-generation sequencing (Shokralla et al. 2014). The plant materials are frequently encountered in criminal investigations but often overlooked as potential evidence. A forensic investigation that seeks to match evidence to a particular plant would require an updated database of samples. This requires the collection and genotyping of many samples from or near the crime scene. The law enforcement officers and attorneys are not very much familiar with the science of botany. So, the important plant-based evidence is often overlooked. Development of a robust DNA barcode database with highly authenticated sequence information will greatly contribute to the future of forensic botany (Ferri et al. 2009). Hallucinogenic compounds are pharmacological agents banned in most of the states or countries. They cause changes in perception, thought, emotion, and consciousness. Such kind

of plants producing hallucinogens has been detected using DNA barcoding technique in some of the forensic studies (Murphy and Bola 2013; Ogata et al. 2013). Various DNA barcodes available in medicinal plants till date are listed below (Table 4.3).

Table 4.3 List of available DNA barcodes of medicinal plants

S.N.	Species	Family	DNA barcode region	Reference
1.	*Acanthopanacis cortex*	Araliaceae	ITS2	Han et al. (2016a)
2.	*Acanthopanacis cortex*	Araliaceae	ITS2	Zhao et al. (2015)
3.	*Aconitum*	Ranunculaceae	psbA-trnH	He et al. (2010)
4.	*Acori Tatarinowii Rhizoma*	Araceae	ITS2	Han et al. (2016a)
5.	*Andrographis paniculata*	Acanthaceae	rbcL	Osathanunkul et al. (2016)
6.	*Andrographis paniculata*	Acanthaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
7.	*Angelica* spp.	Apiaceae	matK, rbcL,ITS, ITS2, psbA-trnH	Yuan et al. (2015) and Techen et al. (2014)
8.	*Astragalus* spp.	Fabaceae	matK, rbcL, ITS	Xiao et al. (2011) and Techen et al. (2014)
9.	*Boerhavia* spp.	Nyctaginaceae	ITS, ITS2,	Selvaraj et al. (2012) and Techen et al. (2014)
10.	*Brugmansia*, *Datura*	Solanaceae	ITS2	Wu et al. (2015)
11.	*Bupleuri radix*	Apiaceae	ITS2	Han et al. (2016a)
12.	*Butea superb*	Fabaceae	matK	Wiriyakarun et al. (2013)
13.	*Cassia* species	Fabaceae	rbcL	Sheth and Thaker (2015)
14.	*Centella asiatica*	Apiaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
15.	*Citrus* spp.	Rutaceae	matk	Penjor et al. (2013)
16.	*Cleome*	Cleomaceae	matK, rbcL, ITS1	Tamboli et al. (2016)
17.	*Clinacanthus nutans*	Acanthaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
18.	*Cosmos caudatus*	Asteraceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
19.	*Cymbidium*	Orchidaceae	ITS2	Sharma et al. (2012)
20.	*Cynanchum auriculatum*	Apocynaceae	trnL-F	Han et al. (2016b)
21.	*Cynanchum wilfordii*	Asclepiadaceae	trnL-F	Han et al. (2016b)
22.	*Dalbergiae Odoriferae* Lignum	Fabaceae	ITS2	Han et al. (2016a)
23.	*Dendrobium* spp.	Orchidaceae	psbA-trnH	Yao et al. (2009)

(continued)

Table 4.3 (continued)

S.N.	Species	Family	DNA barcode region	Reference
24.	*Dipsacus* spp.	Caprifoliaceae	ITS	Vickruck et al. (2011) and Techen et al. (2014)
25.	*Gentiana*	Gentianaceae	matK + ITS	Liu et al. (2016a)
26.	*Ginseng* genus	Araliaceae	matK, rbcL, ITS, psbA-trnH,	Dong et al. (2014),
			rpoB, rpoC1, ITS2	Wallace et al. (2012) and Liu et al. (2016b)
27.	*Ginseng radix*	Araliaceae	ITS2	Han et al. (2016b)
28.	*Hedyotis diffusa* Willd.	Rubiaceae	ITS	Sun et al. (2011)
29.	*Hypericum* spp.	Hypericaceae	ITS	Newmaster et al. (2013)
30.	*Illicium*	Schisandraceae	ITS + trnH-psbA	Zhang et al. (2015a)
31.	*Inulae flos*	Compositae	ITS2	Han et al. (2016b)
32.	*Justicia gendarussa*	Acanthaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
33.	*Lonicera* spp.	Caprifoliaceae	matK, rbcL, ITS, psbA-trnH, trnL-F	Techen et al. (2014)
34.	*Lonicerae japonicae* Flos	Caprifoliaceae	ITS2	Han et al. (2016b)
35.	*Meconopsis* spp.	Papaveraceae	ITS	Techen et al. (2014)
36.	*Mentha aquatica* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
37.	*Mentha spicata* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
38.	*Mucuna collettii*	Fabaceae	matK	Wiriyakarun et al. (2013)
39.	*Murraya koenigii*	Rutaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
40.	*Ochradenus* spp.	Resedaceae	ITS, rpoB, rpoC1	Techen et al. (2014)
41.	*Ocimum gratissimum*	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
42.	*Ocimum basilicum* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)

(continued)

Table 4.3 (continued)

S.N.	Species	Family	DNA barcode region	Reference
43.	*Ocimum tenuiflorum* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
44.	*Origanum heracleoticum* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
45.	*Origanum majorana* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
46.	*Orthosiphon stamineus*	Lamiaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
47.	*Paris* spp.	Melanthiaceae	ITS2	Zhu et al. (2010)
48.	*Persicaria odorata*	Polygonaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
49.	*Phyllanthus niruri*	Phyllanthaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
50.	*Phyllanthus* spp.	Phyllanthaceae	psbA-trnH	Srirama et al. (2010)
51.	*Pinelliae Tuber*, *Arisaematis Rhizoma*	Araceae	matK, rbcL	Moon et al. (2016)
52.	*Piper betel*	Piperaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
53.	*Piper sarmentosum*	Piperaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
54.	*Plectranthus asirensis*	Lamiaceae	rps16, rpoB	Al-Qurainy et al. (2014)
55.	*Polygonum multiflorum*	Polygonaceae	trnL-F	Han et al. (2016a)
56.	*Pueraria candollei*	Fabaceae	matK	Wiriyakarun et al. (2013)
57.	*Radix Astragali*	Fabaceae	ITS	Zheng et al. (2014)
58.	*Radix Rubi Parvifolii*	Gentianaceae	ITS2	Han et al. (2016b)
59.	*Rehmannia* spp.		ITS	Techen et al. (2014)
60.	*Rhodiola*	Crassulaceae	ITS	Zhang et al. (2015b)
61.	*Rhododendron*	Ericaceae	psbA-trnH	Chen et al. (2012)
62.	*Rhododendron* spp.	Ericaceae	matK, rbcL,ITS, ITS2, psbA-trnH	Yan et al. (2015), Chen et al. (2012) and Tsai et al. (2012)
63.	*Rhubarb*	Polygonaceae	matK	Xu et al. (2013)
64.	*Rosmarinus officinalis* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
65.	*Rubus* spp.	Rosaceae	ITS, psbA-trnH, trnL-F	Newmaster et al. (2013)

(continued)

Table 4.3 (continued)

S.N.	Species	Family	DNA barcode region	Reference
66.	*Ruta* spp.	Rutaceae	ITS, rpoB, rpoC1	Al-Qurainy et al. (2011)
67.	*Sabia* spp.	Sabiaceae	matK, rbcL, psbA-trnH	Techen et al. (2014)
68.	*Salvia divinorum*	Lamiaceae	trnL	Murphy and Bola (2013)
69.	*Salvia officinalis* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
70.	*Salvia rutilans*	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
71.	*Salvia sclarea*	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
72.	*Salvia uliginosa*	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
73.	*Sambucus chinensis*	Acanthaceae	ITS2, rpoC1, trnH-psbA	Aziz et al. (2015)
74.	*Isatis indigotica*	Cruciferae	ITS2	Chen et al. (2014)
75.	*Scutellaria baicalensis*	Lamiaceae	psbA-trnH	Guo et al. (2011)
76	*Scutellaria* spp.	Lamiaceae	matK, rbcL, psbA-trnH	Techen et al. (2014)
77.	*Senna* spp.	Fabaceae	psbA-trnH	Pansa Monkheang (2011)
78.	*Smilax* spp.	Smilacaceae	psbA-trnH	Techen et al. (2014)
79.	*Solanum* spp.	Solanaceae	matK, rbcL, ITS, psbA-trnH, trnL-F	Zhang et al. (2013) and Techen et al. (2014)
80.	*Thymus vulgaris* L.	Lamiaceae	rbcL, matK, trnH-psbA, rpoB	De Mattia et al. (2011)
81.	*Tulipa edulis*	Liliaceae	matK	Ma et al. (2014)
82.	*Uncaria*	Rubiaceae	ITS2	Zhang et al. (2015c)
83.	*Uyghur*	Apiaceae	ITS2	Fan et al. (2015)
84.	*Vitex* spp.	Lamiaceae	matK	Phoolcharoen and Sukrong (2013)

4.8 Conclusions

Over the past 12 years, DNA barcoding has been attracting a lot of interest all over the world. Researchers working in this field are busy in finding a more superior and desirable universal DNA barcode for an efficient conservation of the biodiversity. Since a major problem of barcoding lies in the case of plants, the research carried out so far in this area has been reviewed including the futuristic approaches. In the present chapter, various candidate markers used in plants and a number of barcoding reports have been summarized. Although the CBOL proposed seven candidate barcodes belonging to the plastid region, the proposed supplementary loci, i.e., nuclear-transcribed spacer regions ITS1 and ITS2, have a number of GenBank submissions of their respective sequences owing to its easy amplification due to high copy number. rbcL and matK (both plastid genes) come next followed by 18S rRNA (nuclear structural RNA), trnL-F (intron + IGS), and trnH-psbA (IGS), respectively. Since higher substitution rates are observed in plant nuclear genes than plastid genes, ITS is more in use and also acts as a supplementary marker. But once the choice of the locus is made, the approach of single-locus, multilocus, or tiered needs consideration. Based on the literature review, it can be inferred that multilocus and tiered approaches resulted in higher success rates than the single-locus approach if proper combinations of loci and selection of loci for each tier are done carefully.

References

Acs Z, Challis RJ, Bihari P, Blaxter M, Hayward A, Melika G, Csóka G, Pénzes Z, Pujade-Villar J, Nieves-Aldrey JL, Schönrogge K, Stone GN (2010) Phylogeny and DNA barcoding of inquiline oak gall wasps (Hymenoptera: Cynipidae) of the Western Palaearctic. Mol Phylogenet Evol 55:210–225. doi:10.1016/j.ympev.2009.12.004

Ajmal AM, Gyulai G, Hidvégi N, Kerti B, Al Hemaid FMA, Pandey AK, Lee J (2014) The changing epitome of species identification – DNA barcoding. Saudi J Biol Sci 21:204–231. doi:10.1016/j.sjbs.2014.03.003

Al-Qurainy F, Khan S, Tarroum M, Al-Hemaid FM, Ali MA (2011) Molecular authentication of the medicinal herb Rutagraveolens (Rutaceae) and an adulterant using nuclear and chloroplast DNA markers. Genet Mol Res 10:2806–2816. doi:10.4238/2011.November.10.3

Al-Qurainy F, Khan S, Nadeem M, Tarroum M, Al-Ameri A (2014) Selection of DNA barcoding loci and phylogenetic study of a medicinal and endemic plant, plectranthus asirensis J.R.I. Wood from Saudi Arabia. Genet Mol Res 13:6184–6190. doi:10.4238/2014.August.7.31

Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. Mol Phylogenet Evol 29:417–434

Aneesh T, Hisham M, Sekhar M, Madhu M, Deepa T (2009) International market scenario of traditional Indian herbal drugs-India declining. Int J Green Pharm 3:184–190. doi:10.4103/0973-8258.56271

Armstrong KF, Ball SL (2005) DNA barcodes for biosecurity: invasive species identification. Philos Trans R Soc B Biol Sci 360:1813–1823. doi:10.1098/rstb.2005.1713

Austerlitz F, David O, Schaeffer B, Bleakley K, Olteanu M, Leblois R, Veuille M, Laredo C (2009) DNA barcode analysis: a comparison of phylogenetic and statistical classification methods. BMC Bioinformatics 10(Suppl 14):S10. doi:10.1186/1471-2105-10-S14-S10

Aziz NAA, Ahmad MI, Naim DM (2015) Molecular DNA identification of medicinal plants used by traditional healers in Malaysia. Genet Mol Res 14:15937–15947. doi:10.4238/2015. December.7.5

Bailey CD, Carr TG, Harris SA, Hughes CE (2003) Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol Phylogenet Evol 29:435–455

Baldwin BG (1992) Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the Compositae. Mol Phylogenet Evol 1:3–16

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13:729–744

Beck D, Settles M, Foster JA (2011) OTUbase: an R infrastructure package for operational taxonomic unit data. Bioinformatics 27:1700–1701. doi:10.1093/bioinformatics/btr196

Becker S, Hanner R, Steinke D (2011) Five years of FISH-BOL: brief status report. Mitochondrial DNA 22(Suppl 1):3–9. doi:10.3109/19401736.2010.535528

Bensasson D, Zhang DX, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol Evol 16:314–321

Bertolazzi P, Felici G, Weitschek E (2009) Learning to classify species with barcodes. BMC Bioinforma 10(Suppl 14):S7. doi:10.1186/1471-2105-10-S14-S7

Bhargava M, Sharma A (2013) DNA barcoding in plants: evolution and applications of *in silico* approaches and resources. Mol Phylogenet Evol 67:631–641. doi:10.1016/j.ympev.2013.03.002

Biswas K, Biswas R (2014) DNA molecular markers based authentication of herbal drugs – a review. IJPRS 3:581–593

Brown SDJ, Collins RA, Boyer S, Lefort MC, Malumbres-Olarte J, Vink CJ, Cruickshank RH (2012) Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol Ecol Resour 12:562–565. doi:10.1111/j.1755-0998.2011.03108.x

Campbell NJ, Barker SC (1999) The novel mitochondrial gene arrangement of the cattle tick, Boophilus microplus: fivefold tandem repetition of a coding region. Mol Biol Evol 16:732–740

Casiraghi M, Labra M, Ferri E, Galimberti A, De Mattia F (2010) DNA barcoding: a six-question tour to improve users' awareness about the method. Brief Bioinform 11:440–453. doi:10.1093/bib/bbq003

CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci U S A 106:12794–12797. doi:10.1073/pnas.0905845106

Chan K (2003) Some aspects of toxic contaminants in herbal medicines. Chemosphere 52:1361–1371. doi:10.1016/S0045-6535(03)00471-5

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5, e8613. doi:10.1371/journal.pone.0008613

Chen K, Liu Y, Zhang L, Liu Z, Luo K, Chen S (2012) Species identification of Rhododendron (Ericaceae) using the chloroplast deoxyribonucleic acid PsbA-trnH genetic marker. Pharmacogn Mag 8:1–29. doi:10.4103/0973-1296.93311

Chen S, Zhu Z, Ma H, Yang J, Guo Q (2014) DNA barcodes for discriminating the medicinal plant Isatisindigotica Fort. (Cruciferae) and its adulterants. Biochem Syst Ecol 57:287–292. doi:10.1016/j.bse.2014.08.007

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 U S A 108:19641–19646. doi:10.1073/pnas.1104551108

Chu KH, Xu M, Li CP (2009) Rapid DNA barcoding analysis of large datasets using the composition vector method. BMC Bioinforma 10(Suppl 14):S8. doi:10.1186/1471-2105-10-S14-S8

Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. Mol Ecol Resour 13:969–975. doi:10.1111/1755-0998.12046

Collins RA, Armstrong KF, Meier R, Yi Y, Brown SDJ, Cruickshank RH, Keeling S, Johnston C (2012) Barcoding and border biosecurity: identifying cyprinid fishes in the aquarium trade. PLoS One 7, e28381. doi:10.1371/journal.pone.0028381

Costello MJ, May RM, Stork NE (2013) Can we name earth's species before they go extinct? Science 339:413. doi:10.1126/science.1237381

Costion C, Ford A, Cross H, Crayn D, Harrington M, Lowe A (2011) Plant DNA barcodes can accurately estimate species richness in poorly known floras. PLoS One 6, e26841. doi:10.1371/journal.pone.0026841

Cowan RS, Fay MF (2012) Challenges in the DNA barcoding of plant material. Methods Mol Biol Clifton NJ 862:23–33. doi:10.1007/978-1-61779-609-8_3

de Groot GA, During HJ, Maas JW, Schneider H, Vogel JC, Erkens RH (2011) Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. PLoS One 6(1), e16371. doi:10.1371/journal.pone.0016371

De Mattia F, Bruni I, Galimberti A, Cattaneo F, Casiraghi M, Labra M (2011) A comparative study of different DNA barcoding markers for the identification of some members of Lamiacaea. Food Res Int 44:693–702. doi:10.1016/j.foodres.2010.12.032

Desalle R (2006) Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff. Conserv Biol J Soc Conserv Biol 20:1545–1547. doi:10.1111/j.1523-1739.2006.00543.x

Desalle R (2007) Phenetic and DNA taxonomy; a comment on Waugh. BioEssays News Rev Mol Cell Dev Biol 29:1289–1290. doi:10.1002/bies.20667

DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos Trans R Soc B 360:1905–1916. doi:10.1098/rstb.2005.1722

Devey DS, Chase MW, Clarkson JJ (2009) A stuttering start to plant DNA barcoding: microsatellites present a previously overlooked problem in non-coding plastid regions. Taxon 58:7–15

Dong W, Liu H, Xu C, Zuo Y, Chen Z, Zhou S (2014) A chloroplast genomic strategy for designing taxon specific DNA mini-barcodes: a case study on ginsengs. BMC Genet 15:138. doi:10.1186/s12863-014-0138-z

Efferth T, Greten HJ (2012) Medicinal and aromatic plant research in the 21st century. Med Aromat Plant 1e:110. doi:10.4172/2167-0412.1000e110

Elliott TL, Jonathan Davies T (2014) Challenges to barcoding an entire flora. Mol Ecol Resour 14(5):883–891. doi:10.1111/1755-0998.12277

Fan C, Li X, Zhu J, Song J, Yao H (2015) Endangered Uyghur medicinal plant *Ferula* identification through the second internal transcribed spacer. Evid Based Complement Alternat Med 2015:1–6. doi:10.1155/2015/479879

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SC (2008) Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS One 7, e2802. doi:10.1371/journal.pone.0002802

Ferri G, Alù M, Corradini B, Beduschi G (2009) Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. Int J Legal Med 123:395–401

Ferri G, Corradini B, Ferrari F, Santunione AL, Palazzoli F, Alu M (2015) Forensic botany II, DNA barcode for land plants: which markers after the international agreement? Forensic Sci Int Genet 15:131–136. doi:10.1016/j.fsigen.2014.10.005

Ficetola GF, Coissac E, Zundel S, Riaz T, Shehzad W, Bessière J, Taberlet P, Pompanon F (2010) An *in silico* approach for the evaluation of DNA barcodes. BMC Genomics 11:434. doi:10.1186/1471-2164-11-434

Fišer Pečnikar Ž, Buzan EV (2014) 20 years since the introduction of DNA barcoding: from theory to application. J Appl Genet 55:43–52. doi:10.1007/s13353-013-0180-y

Frézal L, Leblois R (2008) Four years of DNA barcoding: current advances and prospects. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 8:727–736. doi:10.1016/j.meegid.2008.05.005

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol Evol 27:480–488. doi:10.1016/j.tree.2012.04.012

Galimberti A, Labra M, Sandionigi A, Bruno A, Mezzasalma V, De Mattia F (2014) DNA barcoding for minor crops and food traceability. Adv Agric 2014:1–8. doi:10.1155/2014/831875

Gantait S, Debnath S, Nasim Ali M (2014) Genomic profile of the plants with pharmaceutical value. 3 Biotech 4:563–578. doi:10.1007/s13205-014-0218-9

Gao T, Yao H, Song J, Zhu Y, Liu C, Chen S (2010a) Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. BMC Evol Biol 10:324. doi:10.1186/1471-2148-10-324

Gao T, Yao H, Song J, Liu C, Zhu Y, Ma X, Pang X, Xu H, Chen S (2010b) Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. J Ethnopharmacol 130:116–121. doi:10.1016/j.jep.2010.04.026

Ghosh S, Mahadani P, Sharma G (2013) Identification of ethnomedicinal plants (Rauvolfioideae: Apocynaceae) through DNA barcoding from northeast India. Pharmacogn Mag 9:255. doi:10.4103/0973-1296.113284

Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, Riéra B, Roger A, Thébaud C, Chave J (2009) Identification of amazonian trees with DNA barcodes. PLoS One 4, e7483. doi:10.1371/journal.pone.0007483

Guo X, Wang X, Su W, Zhang G, Zhou R (2011) DNA barcodes for discriminating the medicinal plant Scutellaria baicalensis (Lamiaceae) and its adulterants. Biol Pharm Bull 34:1198–1203

Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, Hebert PDN (2006) A minimalist barcode can identify a specimen whose DNA is degraded: BARCODING. Mol Ecol 6:959–964. doi:10.1111/j.1471-8286.2006.01470.x

Han EH, Cho K, Goo Y, Kim M, Shin YW, Kim YH, Lee SW (2016a) Development of molecular markers, based on chloroplast and ribosomal DNA regions, to discriminate three popular medicinal plant species, Cynanchumwilfordii, Cynanchumauriculatum, and Polygonummultiflorum. Mol Biol Rep 43:323–332. doi:10.1007/s11033-016-3959-1

Han J, Pang X, Liao B, Yao H, Song J, Chen S (2016b) An authenticity survey of herbal medicines from markets in China using DNA barcoding. Sci Rep 6:18723. doi:10.1038/srep18723

Hao DC, Chen SL, Xiao PG (2010) Sequence characteristics and divergent evolution of the chloroplast psbA-trnH noncoding region in gymnosperms. J Appl Genet 51:259–273

He J, Wong KL, Shaw PC, Wang H, Li DZ (2010) Identification of the medicinal plants in Aconitum L. by DNA barcoding technique. Planta Med 76:1622–1628. doi:10.1055/s-0029-1240967

Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst Biol 54:852–859. doi:10.1080/10635150500354886

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. Proc R Soc B Biol Sci 270:313–321. doi:10.1098/rspb.2002.2218

Hebert PDN, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc Biol Sci 270(Suppl 1):96–99. doi:10.1098/rsbl.2003.0025

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci U S A 101:14812–14817. doi:10.1073/pnas.0406166101

Hickerson MJ, Meyer CP, Moritz C (2006) DNA barcoding will often fail to discover new animal species over broad parameter space. Syst Biol 55:729–739. doi:10.1080/1063515060096989

Hollingsworth ML, Clark AA, Forrest LL, Richardson J, Pennington RT, Long DG, Cowan R, Chase MW, Gaudeul M, 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

Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. PLoS One 6:e19254

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc Biol Sci 272:1525–1534. doi:10.1098/rspb.2005.3056

Jeanson ML, Labat JN, Little DP (2011) DNA barcoding: a new tool for palm taxonomists? Annals Bot 108:1445–1451

Jones M, Ghoorah A, Blaxter M (2011) JMOTU and taxonerator: turning DNA barcode sequences into annotated operational taxonomic units. PLoS One 6, e19259. doi:10.1371/journal.pone.0019259

Kerr KC, Birks SM, Kalyakin MV, Red'kin YA, Koblik EA, Hebert P (2009) Filling the gap – COI barcode resolution in eastern Palearctic birds. Front Zool 6:29. doi:10.1186/1742-9994-6-29

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, Erickson DL (2008) DNA barcodes: genes, genomics, and bioinformatics. Proc Natl Acad Sci U S A 105:2761–2762

Kress JW, Wurdack JK, Zimmer AE, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci U S A 102:8369–8374. doi:10.1073/pnas.0503123102

Kress WJ, Erickson DL, Jones FA (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Natl Acad Sci U S A 106:18621–18626

Kuksa P, Pavlovic V (2009) Efficient alignment-free DNA barcode analytics. BMC Bioinforma 10(Suppl 14):S9. doi:10.1186/1471-2105-10-S14-S9

Kumar S, Carlsen T, Mevik BH, Enger P, Blaalid R, Shalchian-Tabrizi K, Kauserud H (2011) CLOTU: an online pipeline for processing and clustering of 454 amplicon reads into OTUs followed by taxonomic annotation. BMC Bioinforma 12:182. doi:10.1186/1471-2105-12-182

Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008) DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci U S A 105:2923–2928. doi:10.1073/pnas.0709936105

Li FW, Tan BC, Buchbender V, Moran RC, Rouhan G, Wang CN, Quandt D (2009) Identifying a mysterious aquatic fern gametophyte. Plant Syst Evol 281:77–86. doi:10.1007/s00606-009-0188-2

Li CP, Yu ZG, Han GS, Chu KH (2012) Analyzing multi-locus plant barcoding datasets with a composition vector method based on adjustable weighted distance. PLoS One 7:e42154

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S (2015) Plant DNA barcoding: from gene to genome. Biol Rev Camb Philos Soc 90:157–166. doi:10.1111/brv.12104

Little DP (2010) A unified index of sequence quality and contig overlap for DNA barcoding. Bioinformatics 26:2780–2781. doi:10.1093/bioinformatics/btq507

Little DP (2011) DNA barcode sequence identification incorporating taxonomic hierarchy and within taxon variability. PLoS One 6:e20552. doi:10.1371/journal.pone.0020552

Little DP, Stevenson DW (2007) A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. Cladistics 23:1–21. doi:10.1111/j.1096-0031.2006.00126.x

Liu C, Liang D, Gao T, Pang X, Song J, Yao H, Han J, Liu Z, Guan X, Jiang K, Li H, Chen S (2011) PTIGS-IdIt, a system for species identification by DNA sequences of the psbA-trnH intergenic spacer region. BMC Bioinforma 12(Suppl 13):S4. doi:10.1186/1471-2105-12-S13-S4

Liu Y, Zhang L, Liu Z, Luo K, Chen S, Chen K (2012a) Species identification of Rhododendron (Ericaceae) using the chloroplast deoxyribonucleic acid PsbA-trnH genetic marker. Pharmacogn Mag 8:29–36. doi:10.4103/0973-1296.93311

Liu C, Shi L, Xu X, Li H, Xing H, Liang D, Jiang K, Pang X, Song J, Chen S (2012b) DNA barcode goes two-dimensions: DNA QR code web server. PLoS One 7:e35146. doi:10.1371/journal.pone.0035146

Liu J, Yan HF, Ge XJ (2016a) The use of DNA barcoding on recently diverged species in the genus Gentiana (Gentianaceae) in China. PLoS One 11:e0153008. doi:10.1371/journal.pone.0153008

Liu Y, Wang X, Wang L, Chen X, Pang X, Han J (2016b) A nucleotide signature for the identification of American ginseng and its products. Front Plant Sci 7:1–319. doi:10.3389/fpls.2016.00319

Lou SK, Wong KL, Li M, But P, Tsui S, Shaw PC (2010) An integrated web medicinal materials DNA database: MMDBD (Medicinal Materials DNA Barcode Database). BMC Genomics 11:402. doi:10.1186/1471-2164-11-402

Lowenstein JH, Amato G, Kolokotronis SO (2009) The real maccoyii: identifying tuna sushi with DNA barcodes – contrasting characteristic attributes and genetic distances. PLoS One 4:e7866. doi:10.1371/journal.pone.0007866

Ma H, Zhu Z, Zhang X, Miao Y, Guo Q (2014) Species identification of the medicinal plant Tulipa edulis (Liliaceae) by DNA barcode marker. Biochem Syst Ecol 55:362–368. doi:10.1016/j. bse.2014.03.038

Machado CA, Hey J (2003) The causes of phylogenetic conflict in a classic Drosophila species group. Proc Biol Sci 270:1193–1202. doi:10.1098/rspb.2003.2333

Magnacca KN, Brown MJF (2010) Mitochondrial heteroplasmy and DNA barcoding in Hawaiian Hylaeus (Nesoprosopis) bees (Hymenoptera: Colletidae). BMC Evol Biol 10:174. doi:10.1186/1471-2148-10-174

Matz MV, Nielsen R (2005) A likelihood ratio test for species membership based on DNA sequence data. Philos Trans R Soc B 360:1969–1974. doi:10.1098/rstb.2005.1728

Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol 55:715–728. doi:10.1080/10635150600969864

Meier R, Zhang G, Ali F (2008) The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. Syst Biol 57:809–813. doi:10.1080/10635150802406343

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biol 3, e422. doi:10.1371/journal.pbio.0030422

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 14(1):8–21. doi:10.1111/pbi.12419

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58:298–311. doi:10.1093/sysbio/syp027

Moon BC, Kim WJ, Ji Y, Lee YM, Kang YM, Choi G (2016) Molecular identification of the traditional herbal medicines, Arisaematis Rhizoma and Pinelliae tuber and common adulterants via universal DNA barcode sequences. Genet Mol Res 15. doi:10.4238/gmr.15017064

Moritz C, Cicero C (2004) DNA barcoding: promise and pitfalls. PLoS Biol 2:e354. doi:10.1371/journal.pbio.0020354

Moulton MJ, Song H, Whiting MF (2010) Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from orthoptera (Arthropoda: Insecta). Mol Ecol Resour 10:615–627. doi:10.1111/j.1755-0998.2009.02823.x

Muellner AN, Schaefer H, Lahaye R (2011) Evaluation of candidate DNA barcoding loci for economically important timber species of the mahogany family (Meliaceae). Mol Ecol Resour 11:450–460

Murphy TM, Bola G (2013) DNA identification of Salvia divinorum samples. Forensic Sci Int Genet 7:189–193

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

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. Mol Ecol Res 8:480–490

Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S (2013) DNA barcoding detects contamination and substitution in north american herbal products. BMC Med 11:1–222. doi:10.1186/1741-7015-11-222

Nielsen R, Matz M (2006) Statistical approaches for DNA barcoding. Syst Biol 55:162–169. doi:10.1080/10635150500431239

Ogata J, Uchiyama N, Kikura-Hanajiri R, Goda Y (2013) DNA sequence analyses of blended herbal products including synthetic cannabinoids as designer drugs. Forensic Sci Int 227:33–41. doi:10.1016/j.forsciint.2012.09.006

Osathanunkul M, Suwannapoom C, Khamyong N, Pintakum D, Lamphun SN, Triwitayakorn K, Osathanunkul K, Madesis P (2016) Hybrid analysis (barcode-high resolution melting) for authentication of Thai herbal products, Andrographispaniculata (Burm.f.) Wall.ex Nees. Pharmacogn Mag 12:71–75. doi:10.4103/0973-1296.176112

Palhares RM, Gonçalves Drummond M, Dos Santos AFBB, Pereira Cosenza G, das GraçasLinsBrandão M, 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. doi:10.1371/journal.pone.0127866

Pang X, Luo H, Sun C (2012) Assessing the potential of candidate DNA barcodes for identifying non-flowering seed plants. Plant Biol (Stuttg) 14:839–844

Pansa M (2011) Species diversity, usages, molecular markers and barcode of medicinal Senna species (Fabaceae, Caesalpinioideae) in Thailand. J Med Plants Res 5:6173. doi:10.5897/JMPR11.1075

Parveen I, Singh HK, Raghuvanshi S, Pradhan UC, Babbar SB (2012) DNA barcoding of endangered Indian *Paphiopedilum* species: DNA barcoding of Indian *Paphiopedilum* species. Mol Ecol Resour 12:82–90. doi:10.1111/j.1755-0998.2011.03071.x

Penjor T, Yamamoto M, Uehara M, Ide M, Matsumoto N, Matsumoto R, Nagano Y (2013) Phylogenetic relationships of Citrus and its relatives based on matK gene sequences. PLoS One 8, e62574. doi:10.1371/journal.pone.0062574

Peterson AT, Moyle RG, Nyári AS, Robbins MB, Brumfield RT, Remsen JV (2007) The need for proper vouchering in phylogenetic studies of birds. Mol Phylogenet Evol 45:1042–1044. doi:10.1016/j.ympev.2007.08.019

Pettengill JB, Neel MC (2010) An evaluation of candidate plant DNA barcodes and assignment methods in diagnosing 29 species in the genus *Agalinis* (Orobanchaceae). Am J Bot 97:1391–1406

Phoolcharoen W, Sukrong S (2013) Molecular analysis of *Vitex* species using candidate DNA barcoding and PCR-RFLP of the matK gene for authentication of *Vitex glabrata*. Nat Prod Commun 8:125–128

Pillon Y, Johansen J, Sakishima T, Chamala S, Barbazuk WB, Roalson EH, Price DK, Stacy EA (2013) Potential use of low-copy nuclear genes in DNA barcoding: a comparison with plastid genes in two Hawaiian plant radiations. BMC Evol Biol 13:35. doi:10.1186/1471-2148-13-35

Prendini L (2005) Comment on "Identifying spiders through DNA barcodes". Can J Zool 83:498–504. doi:10.1139/z05-025

Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, automatic barcode gap discovery for primary species delimitation. Mol Ecol 21:1864–1877. doi:10.1111/j.1365-294X.2011.05239.x

Qi J, Wang B, Hao BL (2004) Whole proteome prokaryote phylogeny without sequence alignment: a k-string composition approach. J Mol Evol 58:1–11

Quandt D, Stech M (2004) Molecular evolution of the trnTUGU-trnFGAA region in Bryophytes. Plant Biol (Stuttg) 6:545–554

Ratnasingham S, Hebert PDN (2007) BOLD: the barcode of life data system (www.barcodinglife.org). Mol Ecol Notes 7:355–364

Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Res 39:e145. doi:10.1093/nar/gkr732

Richly E, Leister D (2004) NUMTs in sequenced eukaryotic genomes. Mol Biol Evol 21:1081–1084. doi:10.1093/molbev/msh110

Rubinoff D (2006) Utility of mitochondrial DNA barcodes in species conservation. Conserv Biol J Soc Conserv Biol 20:1026–1033

Rubinoff D, Cameron S, Will K (2006a) Are plant DNA barcodes a search for the Holy Grail? Trends Ecol Evol 21:1–2. doi:10.1016/j.tree.2005.10.019

Rubinoff D, Cameron S, Will K (2006b) A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. J Hered 97:581–594. doi:10.1093/jhered/esl036

Sarkar IN, Planet PJ, Desalle R (2008) caos software for use in character-based DNA barcoding. Mol Ecol Resour 8:1256–1259. doi:10.1111/j.1755-0998.2008.02235.x

Sarwat M, Yamdagni MM (2015) DNA barcoding, microarrays and next generation sequencing: recent tools for genetic diversity estimation and authentication of medicinal plants. Crit Rev Biotechnol 1–13. doi:10.3109/07388551.2014.947563

Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R (2005) Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philos Trans R Soc B 360:1805–1811. doi:10.1098/rstb.2005.1730

Seibel PN, Müller T, Dandekar T, Schultz J, Wolf M (2006) 4SALE–a tool for synchronous RNA sequence and secondary structure alignment and editing. BMC Bioinforma 7:498. doi:10.1186/1471-2105-7-498

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 Bioinforma 10:364–367. doi:10.1016/j.gpb.2012.03.002

Sharma SK, Dkhar J, Kumaria S, Tandon P, Rao SR (2012) Assessment of phylogenetic interrelationships in the genus Cymbidium (Orchidaceae) based on internal transcribed spacer region of rDNA. Gene 495:10–15. doi:10.1016/j.gene.2011.12.052

Sheth BP, Thaker VS (2015) Identification of a herbal powder by deoxyribonucleic acid barcoding and structural analyses. Pharmacogn Mag 11:570–574. doi:10.4103/0973-1296.172963

Shokralla S, Gibson JF, Nikbakht H, Janzen DH, Hallwachs W, Hajibabaei M (2014) Next-generation DNA barcoding: using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. Mol Ecol Resour 14:892–901. doi:10.1111/1755-0998.12236

Slabbinck B, Dawyndt P, Martens M, De Vos P, De Baets B (2008) TaxonGap: a visualization tool for intra- and inter-species variation among individual biomarkers. Bioinformatics 24:866–867. doi:10.1093/bioinformatics/btn031

Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc Natl Acad Sci U S A 105:13486–13491. doi:10.1073/pnas.0803076105

Srirama R, Senthilkumar U, Sreejayan N, Ravikanth G, Gurumurthy BR, Shivanna MB, Sanjappa M, Ganeshaiah KN, Uma Shaanker R (2010) Assessing species admixtures in raw drug trade of Phyllanthus, a hepato-protective plant using molecular tools. J Ethnopharmacol 130:208–215. doi:10.1016/j.jep.2010.04.042

Srivathsan A, Meier R (2012) On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics 28:190–194. doi:10.1111/j.1096-0031.2011.00370.x

Stech M, Veldman S, Larrain J, Muñoz J, Quandt D, Hassel K, Kruijer H (2013) Molecular species delimitation in the *Racomitrium canescens* complex (Grimmiaceae) and implications for DNA barcoding of species complexes in mosses. PLoS One 8(1), e53134. doi:10.1371/journal.pone.0053134

Steinke D, Vences M, Salzburger W, Meyer A (2005) TaxI: a software tool for DNA barcoding using distance methods. Philos Trans R Soc B 360:1975–1980. doi:10.1098/rstb.2005.1729

Storchová H, Olson MS (2007) The architecture of the chloroplast psbA-trnH non-coding region in angiosperms. Plant Syst Evol 268:235–256

Sun YL, Wang D, Yeon MH, Kim DH, Kim HG, Hong SK (2011) Molecular identification of medicinal herbs, Oldenlandia diffusa and Oldenlandia corymbosa based on nrDNA ITS region sequence. J Plant Biotechnol 38(4):301–307

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res 35, e14. doi:10.1093/nar/gkl938

Tamboli AS, Patil SM, Gholave AR, Kadam SK, Kotibhaskar SV, Yadav SR, Govindwar SP (2016) Phylogenetic analysis, genetic diversity and relationships between the recently segregated species of *Corynandra* and *Cleoserrata* from the genus *Cleome* using DNA barcoding and molecular markers. C R Biol 339:123–132. doi:10.1016/j.crvi.2016.02.005

Taylor HR, Harris WE (2012) An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. Mol Ecol Resour 12:377–388. doi:10.1111/j.1755-0998.2012.03119.x

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

Tsai LC, Yu YC, Hsieh HM, Wang JC, Linacre A, Lee JC (2006) Species identification using sequences of the trnL intron and the trnL-trnF IGS of chloroplast genome among popular plants in Taiwan. Forensic Sci Int 164:193–200

Tsai LC, Yu YC, Hsieh HM (2008) Bidens identification using the non coding regions of chloroplast genome and nuclear ribosomal DNA. Forensic Sci Int Genet 2:35–40

Tsai CC, Chen CH, Chou CH (2012) DNA barcodes reveal high levels of morphological plasticity among Rhododendron species (Ericaceae) in Taiwan. Biochem Syst Ecol 40:169–177. doi:10.1016/j.bse.2011.10.013

Tyagi A, Bag SK, Shukla V, Roy S, Tuli R (2010) Oligonucleotide frequencies of barcoding loci can discriminate species across kingdoms. PLoS One 5:e12330. doi:10.1371/journal.pone.0012330

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. Trends Ecol Evol 24:110–117. doi:10.1016/j.tree.2008.09.011

Vences M, Thomas M, Bonett RM, Vieites DR (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philos Trans R Soc B 360:1859–1868. doi:10.1098/rstb.2005.1717

Vickruck JL, Rehan SM, Sheffield CS, Richards MH (2011) Nesting biology and DNA barcode analysis of *Ceratina dupla* and *C. mikmaqi* and comparisons with *C. calcarata* (Hymenoptera: Apidae: Xylocopinae). Can Entomol 143:254–262. doi:10.4039/n11-006

Virgilio M, Backeljau T, Nevado B, De Meyer M (2010) Comparative performances of DNA barcoding across insect orders. BMC Bioinforma 11:206. doi:10.1186/1471-2105-11-206

Virgilio M, Jordaens K, Breman FC, Backeljau T, De Meyer M (2012) Identifying insects with incomplete DNA barcode libraries, African fruit flies (Diptera: Tephritidae) as a test case. PLoS One 7, e31581. doi:10.1371/journal.pone.0031581

Vogler AP, Monaghan MT (2007) Recent advances in DNA taxonomy. J Zool Syst Evol Res 45:1–10. doi:10.1111/j.1439-0469.2006.00384.x

Wallace LJ, Boilard SMAL, Eagle SHC, Spall JL, Shokralla S, Hajibabaei M (2012) DNA barcodes for everyday life: routine authentication of natural health products. Food Res Int 49:446–452. doi:10.1016/j.foodres.2012.07.048

Wang N, Jacques FMB, Milne RI, Zhang CQ, Yang JB (2012) DNA barcoding of Nyssaceae (Cornales) and taxonomic issues. Bot Stud 53:265–274

Ward J, Peakall R, Gilmore SR, Robertson J (2005) A molecular identification system for grasses: a novel technology for forensic botany. Forensic Sci Int 152:121–131

Ward J, Gilmore SR, Robertson J, Peakall R (2009) A grass molecular identification system for forensic botany: a critical evaluation of the strengths and limitations. J Forensic Sci 54(6):1254–1260. doi:10.1111/j.1556-4029.2009.01196

Whitlock BA, Hale AM, Groff PA (2010) Intraspecific inversions pose a challenge for the trnH-psbA plant DNA barcode. PLoS One 5:e11533. doi:10.1371/journal.pone.0011533

Whitworth TL, Dawson RD, Magalon H, Baudry E (2007) DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). Proc Biol Sci 274:1731–1739. doi:10.1098/rspb.2007.0062

Will KW, Rubinoff D (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20:47–55. doi:10.1111/j.1096-0031.2003.00008.x

Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. Syst Biol 54:844–851. doi:10.1080/10635150500354878

Wiriyakarun S, Yodpetch W, Komatsu K, Zhu S, Ruangrungsi N, Sukrong S (2013) Discrimination of the Thai rejuvenating herbs *Pueraria candollei* (white Kwao Khruea), *Butea superba* (red Kwao Khruea) and *Mucuna collettii* (black Kwao Khruea) using PCR-RFLP. J Nat Med 67:562–570. doi:10.1007/s11418-012-0716-1

Wolf M, Friedrich J, Dandekar T, Müller T (2005) CBCAnalyzer: inferring phylogenies based on compensatory base changes in RNA secondary structures. In Silico Biol 5:291–294

Wu Y, Xu L, Chen L, Wang B, Zhao R (2015) DNA molecular identification of Datura medicinal plants using ITS2 barcode sequence. Zhong Yao Cai 38:1852–1857

Xiang XG, Zhang JB, Lu AM, Li RQ (2011) Molecular identification of species in Juglandaceae: a tiered method. J Syst Evol 49:252–260

Xiao WL, Motley TJ, Unachukwu UJ, Lau CBS, Jiang B, Hong F, Leung PC, Wang QF, Livingston PO, Cassileth BR, Kennelly EJ (2011) Chemical and genetic assessment of variability in commercial Radix Astragali (Astragalus spp) by ion trap LC-MS and nuclear ribosomal DNA barcoding sequence analyses. J Agric Food Chem 59:1548–1556. doi:10.1021/jf1028174

Xu G, Wang X, Liu C, Li W, Wei S, Liu Y, Cheng X, Liu J, Donini P (2013) Authentication of official Da-huang by sequencing and multiplex allele-specific PCR of a short maturase K gene. Genome 56:109–113. doi:10.1139/gen-2012-0182

Yan LJ, Liu J, Möller M, Zhang L, Zhang XM, Li DZ, Gao LM (2015) DNA barcoding of Rhododendron (Ericaceae), the largest Chinese plant genus in biodiversity hotspots of the Himalaya-Hengduan Mountains. Mol Ecol Resour 15:932–944. doi:10.1111/1755-0998.12353

Yao H, Song JY, Ma XY, Liu C, Li Y, Xu HX, Han JP, Duan LS, Chen SL (2009) Identification of *Dendrobium* species by a candidate DNA barcode sequence: the chloroplast *psbA-trnH* intergenic region. Planta Med 75:667–669. doi:10.1055/s-0029-1185385

Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, Chen S (2010) Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS One 510:e13102

Yuan QJ, Zhang B, Jiang D, Zhang WJ, Lin TY, Wang NH, Chiou SJ, Huang LQ (2015) Identification of species and *materia medica* within *Angelica* L. (Umbelliferae) based on phylogeny inferred from DNA barcodes. Mol Ecol Resour 15:358–371. doi:10.1111/1755-0998.12296

Zhang DX, Hewitt GM (1996) Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. Mol Ecol 5:295–300

Zhang AB, Sikes DS, Muster C, Li SQ (2008) Inferring species membership using DNA sequences with back-propagation neural networks. Syst Biol 57:202–215. doi:10.1080/10635150802032982

Zhang A, Feng J, Ward RD, Wan P, Gao Q, Wu J, Zhao W (2012a) A new method for species identification via protein-coding and non-coding DNA barcodes by combining machine learning with bioinformatic methods. PLoS One 7, e30986. doi:10.1371/journal.pone.0030986

Zhang AB, Muster C, Liang HB, Zhu CD, Crozier R, Wan P, Feng J, Ward RD (2012b) A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. Mol Ecol 21:1848–1863. doi:10.1111/j.1365-294X.2011.05235.x

Zhang W, Fan X, Zhu S, Zhao H, Fu L (2013) Species-specific identification from incomplete sampling: applying DNA barcodes to monitoring invasive Solanum plants. PLoS One 8:e55927. doi:10.1371/journal.pone.0055927

Zhang J, Chen M, Dong X, Lin R, Fan J, Chen Z (2015a) Evaluation of four commonly used DNA barcoding loci for Chinese medicinal plants of the family Schisandraceae. PLoS One 10, e0125574. doi:10.1371/journal.pone.0125574

Zhang JQ, Meng SY, Wen J, Rao GY (2015b) DNA barcoding of *Rhodiola* (Crassulaceae): a case study on a group of recently diversified medicinal plants from the Qinghai-Tibetan Plateau. PLoS One 10, e0119921. doi:10.1371/journal.pone.0119921

Zhang ZL, Song MF, Guan YH, Li HT, Niu YF, Zhang LX, Ma XJ (2015c) DNA barcoding in medicinal plants: testing the potential of a proposed barcoding marker for identification of Uncaria species from China. Biochem Syst Ecol 60:8–14. doi:10.1016/j.bse.2015.02.017

Zhao S, Chen X, Song J, Pang X, Chen S (2015) Internal transcribed spacer 2 barcode: a good tool for identifying Acanthopanacis cortex. Front Plant Sci 6:840. doi:10.3389/fpls.2015.00840

Zheng S, Liu D, Ren W, Fu J, Huang L, Chen S (2014) Integrated analysis for identifying *Radix Astragali* and its adulterants based on DNA barcoding. Evid Based Complem Alternat Med 2014:1–11. doi:10.1155/2014/843923

Zhu Y, Chen S, Yao H, Tan R, Song J, Luo K, Lu J (2010) DNA barcoding the medicinal plants of the genus Paris. Yao XueXueBao 45:376–382