Utility of DNA Barcoding for Plant Biodiversity Conservation

Dhivya Selvaraj1, Jong-In Park1, Mi-Young Chung2, Yong-Gu Cho3, Sathishkumar Ramalingam4, Ill-Sup Nou1*

1Department of Horticulture, Sunchon National University, 413 Jungangro, Suncheon, Jeonnam 540-950, Republic of Korea
2Department of Agricultural Education, Sunchon National University, 413 Jungangro, Suncheon, Jeonnam 540-950, Republic of Korea
3Department of Crop Science, Chungbuk National University, 410 Seongbongro, Heungdokgu, Cheongju 361-763, Republic of Korea
4Plant Genetic Engineering Laboratory, Department of Biotechnology, Bharathiar University, Coimbatore 641-046, India

ABSTRACT DNA barcoding is a technique that provides rapid identification of species without using morphological cues. The method employs relatively small-standardized DNA fragments as tags to define or discover species. In plants, the mitochondrial genome evolves much more slowly than in animals. There is currently no consensus on which candidate markers comprise the best plant DNA barcoding region; however, DNA barcodes such as \textit{rbcL}, \textit{matK}, \textit{psbA-trnH} and \textit{psbA-trnH} have been proposed for the plant kingdom. And also very recently the chloroplast intergenic spacer (IGS) like \textit{trnE-trnT}, \textit{trnT-psbD}, \textit{ndhF-rpl32} and \textit{rpl14-rpl16} were also employed for discriminating the cultivar species. The region \textit{ITS2} showed better intra-species variation, followed by \textit{psbA-trnH}. Several analyses reveal that the \textit{ITS2} region is able to distinguish all tested species of the plant kingdom, but evaluations of DNA barcodes have to be conducted for more species covering many genera to confirm the above results. In this review we discussed the current view of DNA barcoding.

Keywords DNA barcode, PCR, \textit{matK}, \textit{rbcL}, \textit{atpB}

INTRODUCTION DNA barcoding is a technique, which provides quick identification of species without involving the morphological cues. It uses a relatively small-standardized DNA fragment as a tag, to define or discover a species. DNA barcoding uses minor differences of nucleotides in the particular gene loci of different organisms as a key to discriminate the organisms. The difference in base pairs of the gene is sequenced and deposited in the barcode database, which is termed as DNA barcodes. These genetic codes could be accessed through a digital library and used to identify the unknown species by any scientist around the world. Ideal DNA barcode should be normally a uniform short sequence of DNA (400–800 bp), able to be simply generated and characterized for all the living organisms (Savolainen \textit{et al.} 2005). Paul Hebert’s groups are the first one to design and use the short DNA sequences for biological identifications. DNA barcodes identifies the species across all forms of life including animals, plants and microbes at a rapid and reliable rate. The idea of barcoding has first emerged to describe the microorganisms in which the morphological keys were lacking. Now it is being applied successfully to animals. A massive online digital library of barcodes will be a standard to which the DNA barcode sequence of an unknown sample can be matched for the identification. The key process in DNA barcoding is identifying novel candidate gene, which can be used universally. It should allow the users efficiently to distinguish the known species and accelerate the discovery of species. DNA barcoding uses the information of one or a few regions in the genome to recognize all the species in a genus (Lahaye \textit{et al.} 2008). The effectiveness of DNA barcoding depends on the detection and description of new cryptic species (Handfield and Handfield 2006; Smith \textit{et al.} 2006; Anker \textit{et al.} 2007; Bucklin \textit{et al.} 2007; Gomez \textit{et al.} 2007; Pfenninger \textit{et al.} 2007; Tavares and Baker 2008) and sibling species (Hogg and Hebert 2004; Amaral \textit{et al.} 2007; van Velzen \textit{et al.} 2007).
In addition, it is important to distinguish, detect and trace the distribution of patented organisms in agro-biotechnology, to certify the source organism like truffles (Rastogi et al. 2007). DNA barcode has many scientific domains like ecology, biomedicine, epidemiology, evolutionary biology, biogeography, conservation biology and bio-industry. The low cost and time-effectiveness makes the process easier for enabling automated species identification in massive sampling campaigns (Rusch et al. 2007). Other novel molecular technologies like bioengineering (e.g. silicon-based microarrays, nylon membrane based macroarrays, etc.) are becoming cheaper and integrated into the second step of DNA barcoding (Summerbell et al. 2005). PCR-based techniques are highly applied in the fields of taxonomy, food and forensic (Teletchea et al. 2005).

DNA barcoding will open up new opportunities in DNA based investigations ranging from community phylogenetics (Webb et al. 2002) to ecological genomics (van Straalen and Roelofs 2012). Despite the lack of consent regarding a universal plant barcode, taxonomists, ecologists, evolutionary biologists and conservationists are already envisioning the application of genetic identifiers to a wide set of research and practical programs. The suitability of a locus for large-scale DNA barcoding can be easily investigated by comparing loci across a similar set of taxa under a selected set of reaction conditions. Thus, the statistics was taken into account between the ability to amplify a locus and the rate of divergence of that locus across a phylogenetic range of taxa. Additionally, the sequence alignment methodologies are available, which can be evaluated for the use of DNA barcodes with the following regards, the purpose of assurance limits to species assignment, the use of a part of sequences in database searches and strength of search algorithms of sequence length variation due to insertion/deletion events and the informative nature of these mutations.

DNA barcoding mirrors the distributions of intra- and inter-specific variation that are separated by a distance known as DNA barcoding gap (Hebert et al. 2003; Meyer and Paulay 2005). The Consortium of Barcode of Life coordinates DNA barcoding development and implementation universally. DNA barcoding is very essential for, (i) the molecular identification of already described species (Hebert et al. 2003) and (ii) the discovery of undescribed species (Valentini et al. 2006). DNA barcoding is the conjugation of molecularization the use of the variability in molecular markers as a discriminator, computerization, the transposition of the data using informatics and standardization. The ultimate aim of DNA barcoding is to discriminate the species using an automated system so that unexplored living organisms can be named before it gets extinct. Rapid progress in DNA sequencing and computational technologies made CBOL to build a universal organization for a living beings inventory: the BOLD system (Ratnasingham and Hebert 2007).

Many of the published work used a simple distance matrix analysis, a Neighbour-Joining (NJ) algorithm with Kimura-2-parameters (K2P). The identification and characterization of molecular entities are the main goal in DNA barcoding studies. An ideal DNA barcode should possess high inter-specific, low intra-specific sequence divergence. It should undergo universal amplification with standard primers and technically simple to analyze. It should be readily recoverable from the museum or herbarium samples and other degraded samples.

**Plant DNA barcoding**

Plant DNA barcoding has a huge role in the conservation biology especially in assessment of biodiversity hotspots and also to monitor the international trade in rare species. Plants have not been given much importance in the early stages of DNA barcoding due to inability of cytochrome oxidase (COX1) to work as a barcode in plants (Cho et al. 2004). The contest was set among the botanists to find a more suitable marker (Pennisi 2007). Many candidate gene regions have been recommended as possible barcodes for plants, but still there is no universal barcode (Kress et al. 2005). The lack of consensus region in plants as in the case of COX1 in animals as a universal barcode for plants has not been found until now, instead several groups have put forward different barcode candidates successfully for the smaller taxa. Several factors are considered in selecting a plant DNA barcode. These are (i) universal PCR condition, (ii) range of taxonomic diversity, (iii) power of species differentiation, and (iv) dry lab analysis and application. The various loci and intergenic spacers have been proposed
as plant barcodes. DNA barcoding must be realistic for a wide range of practitioners and the methodology must be accessible and easily carried out by multiple users. The strength of DNA barcoding was directly related to the data available in the barcode libraries which helps in building a very complete DNA barcoding database (Ekrem et al. 2007). These considerations required a standard range of PCR conditions along with a set of standard PCR primers per gene which serve as a robust barcode marker for the widest range of taxa and users. And also focused on the feasibility of barcoding plants from highly degraded samples (Taberlet et al. 2006) (e.g. permafrost samples) and other applied fields (e.g. processed food, customs, medicinal plants). They suggested the chloroplast trnL (UAA) intron or a shorter fragment of this intron (the P6 loop, 10-143 bp), in spite of relatively low resolution could be improved with highly conserved primers.

In plants, the mitochondrial genome evolves much more slowly than in animals. The mitochondrial gene COXI region was unsuitable for plant species distinction (Rubinoff 2006). The CBOL plant-working group (PWG) concluded that plant DNA barcoding would be multi-locus, with one ‘anchor’ (i.e. universal across the plant kingdom) and ‘identifiers’ to distinguish closely related species. Several combinations of DNA regions have been recently proposed by (Kress et al. 2005; Chase et al. 2007; Kress and Erickson 2007; Pennisi 2007). At present, there is still no consensus on which candidate markers are the best plant DNA barcoding region. The future combination will certainly contain noncoding intergenic spacers like trnH-psbA and plastidial coding sequences like matK (Chase et al. 2007).

Many applications can be implemented with the help of DNA barcodes like identification of different life stages thereby enabling discrimination between seed and seedlings and discovering the changes of metamorphosis. It is difficult to distinguish morphologically the young leaves, seeds and seed coats of certain species, which are related. DNA barcoding helps in distinguishing them at the molecular level. DNA barcoding can be employed in forensic analysis, helps in identifying adulterated products from original components. It performs as a biosecurity for trade in the controlled species like illegal import and export of economically valuable things. It can be employed to assess the taxonomically diversified species both native and invaded species. The DNA barcode project aims to develop a simple diagnostic tool based on strong taxonomic data that is collated in the DNA barcode reference library (Schindel and Miller 2005).

Potential plant barcode candidates

The plant DNA barcoding studies were initially restricted to the chloroplast genome to understand the variation of its gene sequences of coding (matK, rbcL, rpoC1 and psbA-trnH) and non-coding (ITS, psbA-trnH), which has been summarized by Chen et al. (2010). The studies proved that the chloroplast genome showed high degree of variation and discrimination ability (Lahaye et al. 2008). The slow rate of their substitution, limits the ability to identify plants among the species (Baldwin et al. 1995). The evidence suggests that species-lineage reconstructions using Chloroplast DNA (cpDNA) regions show significant errors, when hybridization and introgression events or linear sorting events occur (Doyle 1992).

matK

The matK gene of chloroplast codes for maturase like protein, which is involved in Group II intron splicing. The gene contains approximately 1500 base pairs, which was located within the intron of the trnK. The two exons of the trnK gene that flank the matK were lost, leaving the gene intact in the event of splicing (Wolfe et al. 1992). The gene contains high substitution rates within species and is emerging as one of the potential gene and it is being used in the studies of molecular systematics and evolution (Johnson and Soltis 1995). A homology search specifies 102 amino acids at the carboxyl terminus that are structurally related to portions of maturase-like polypeptide and this might be implicated in splicing of Group II introns (Mohr et al. 1993). The matK gene has been used effectively in phylogenetic studies for the families Saxifragaceae (Johnson and Soltis 1994), Polemoniaceae (Steele and Vilgalys 1994), Poaceae (Liang and Hilu 1996), Orchidaceae tribe Vandaeae (Jarrell et al. 1995), and Myrtaceae (Gadek et al. 1996). Among the available Chloroplast genes, matK has got high substitution rates, which makes it as one of the ideal barcode candidates. matK has been proven as DNA
barcode for the plant family Zingiberaceae (Selvaraj et al. 2008). In 2010, the Linnaean Society of London, performed large scale analysis of matK gene, compared the conventionally available primer to assess the existing primer match with previously generated matK sequences and to suggest possible modifications primarily to simplify the complications in barcoding.

**rbcL**

The *rbcL* gene in higher plant presents as a single copy per chloroplast genome, but many copies of the genome are present in each plastid, the actual *rbcL* copy number per chloroplast can be high. *rbcL* contains only exons, polypeptide with ~ 475 amino acids. Several chloroplast genes, including *rbcL*, the transcriptional (Bradley and Gatenby 1985; Gruissem and Zurawski 1985) and translational (Gatenby et al. 1989) identified sequences resemble as those found in prokaryotic organisms to the extent that chloroplast *rbcL* genes can be readily expressed in *Escherichia coli* (Gatenby et al. 1981). Transcriptional initiation rates from the *rbcL* promoter are not only influenced by the promoter sequence but also modified by the nearby *atpB* promoter. The two promoters are positioned 400 bp apart in opposite orientations, resulting in divergent transcription. They do not function independently, because RNA polymerase binding at the *rbcL* promoter interferes with binding and transcription of the *atpB* promoter, presumably by steric hindrance at the two RNA polymerase binding sites. Removals of one promoter or increasing the space between them eliminate this mutual interference (Hanley-Bowdoin and Chua 1987), which may be a control mechanism to regulate the different levels of expression in chloroplasts.

The substitutions between *rbcL* of the parent species are non-synonymous. Even single amino acid replacements in *rbcL* could result in differences in the CO2 and O2 specificity of ribulose 1, 5 bisphosphate carboxylase/oxygenase (RuBisCO) (Galmes et al. 2005). The speculation is that the amino acid difference in *rbcL* plays an important role in ecological adaptation. It provides all the catalytically essential residues of RuBisCO, a significant enzyme for both the reductive and oxidative photosynthetic carbon cycles. The sequence of *rbcL* has great phylogenetic importance because of its conserved nature (Albert et al. 1994), although substitutions occur in sites of known functional importance (Kellogg and Juliano 1997). A contingency analysis was performed to evaluate the *rbcL*+*matK* barcode among the major taxonomic groups of specifically, non-angiosperms vs. angiosperms and mono-vs. polytypic-genera (Burgess et al. 2011). By comparing the two DNA barcode markers *rbcL* and *matK*, *rbcL* shows better coverage than *matK* for the native flowering plants and conifers of Welsh flora (de Vere et al. 2012).

**psbA-trnH**

Intergenic cp DNA regions contain a wealth of information useful for inquiries in population genetics and low-level systematics (McCauley 1995; Provan et al. 2001). Portions of these non-coding regions whose functions were unknown but much of the variation may result from the spread of mutations not restrained by selection (Olmstead et al. 1992; Hamilton et al. 2003; Monde et al. 2000; Manuell et al. 2004). Population genetic studies have revealed that the end spacer of the intergenic spacer *psbA-trnH* nearest to *psbA* was highly conserved and showed a large inversion at the end nearest to *trnH*, which appears to be more variable. The intergenic spacers have been given more importance in ongoing DNA barcoding efforts (Kress and Erickson 2007; Shaw et al. 2005). The structure of *psbA-trnH* have a small spacer region with a length of about 200–500 bp in angiosperm and gymnosperm species, which is amplified using the universal primers developed by (Hamilton 1999, Storchova and Olson 2007). The high nucleotide variation has made it difficult to identify conserved regions among the highly diverged taxonomic groups. Surprisingly, short regions (6-30 bp) were conserved across angiosperms. Unfortunately, their brevity limits their potential for phylogenetic inference across highly diverged taxa. The *psbA-trnH* intergenic region contains two parts that differ in their evolutionary conservation, *psbA* 3’UTR, which is responsible for the regulation of gene expression, and *psbA-trnH*, which is a non-transcribed intergenic spacer that has no function because of its high variability across angiosperms. The *psbA* gene encodes the D1 reaction center protein of photosystem II and its expression depends on light intensity, the plant developmental stage and the physiological state of
the plant (He et al. 1998). This gene accumulates to high levels in chloroplasts, but is only present at low levels in amyloplasts (Deng and Grusissem 1987). The role of chloroplast psbA UTRs in regulation of gene expression has been examined rigorously for more than twenty years (Zurawski et al. 1982). The frequent inversions in a region of trnH-psbA are flanked by inverted repeats. There was an inversion in trnH-psbA that has been noted before in various plant lineages. The trnH-psbA region shows several traits, it is short length (often, 500bp), suspected ubiquity in plants, high interspecific sequence divergence and universal flanking primers that allow easy amplification and sequencing from both high molecular weight and it also effective in degraded DNA which suspects desirable and cloning and sequencing from both high molecular weight and it also effective in degraded DNA which suspects desirable

\[ \text{synthesized by} \]

Also it is established that atpB and rbcL are transcribed divergently in the chloroplast DNA.

ITS

The internal transcribed spacer (ITS) region belonging to the nuclear genome is a non-functional RNA sequence located between the 18S and 25S rRNA coding regions. The ITS1 is present between 18S and 5.8S rRNA, while ITS2 is between 5.8S and 25S rRNA (Gerbi 1985). The ITS was a transcriptional unit situated between the structural ribosomal RNA during rRNA maturation, the ITS spacers are excised and non-functional maturation product are rapidly degraded. Mutational studies of the ITS region of yeast (Saccharomyces cerevesiae) have shown that deletions of certain regions within ITS1 inhibited the production of mature small and large subunit rRNAs (Mustee et al. 1990; van Nues et al. 1994), whereas certain deletions or point mutations in ITS2 prevented or reduced processing of large subunit rRNAs (van der Sande et al. 1992). The length of ITS1 and ITS2 of all flowering plants varies, but is less than 300 bp for ITS1 (187 to 298) and nearly 250 bp for ITS2 (187 to 252). The total length of the ITS region is around 700 bp, including the 5.8S rRNA region, which has a constant length of 163 or 164 bp (Goldman et al. 1983; Stewart et al. 1983; Gonzalez et al. 1990). The length of the ITS region varies due to point mutations. The percentage of mutations of ITS1 at the intra-specific level is 5.1%, while it is 8.1% at the inter-specific level. The percentage of mutations of ITS2 at the intra-specific level is 6.3%, while it is 6.7% at the inter-specific level (Baldwin 1992).

The ITS region of the nuclear DNA (nrDNA) occurs as tandem repeats at multiple chromosomal loci (Rogers and Bendich 1987; Hamby and Zimmer 1992). The high copy number of the ITS region promotes detection, amplification, cloning and sequencing of nr DNA (Baldwin et al. 1995). The PCR efficiency of the ITS region is high when
compared to barcode candidates therefore, it can be further subjected to restriction digestion, which generates distinctive diagnostic bands that can be used to effectively differentiate and identify plants at their species level (Manhart and McCourt 1992).

Fortunately, comparison of chloroplast DNA with that of nuclear DNA has provided an effective means of identification that can significantly improve our understanding of the origin of polyploid species (Soltis and Kuzoff 1993). For this reason, we investigated the nuclear genome under the assumption that nr DNA regions must be evolutionarily conserved, phylogenetically interpretable, easily examined in the laboratory with universal primers and potentially useful characters for phylogenetic reconstruction (Baldwin 1992). DNA barcoding research is aimed at identification of unique candidate genes for identification of all plant species using both coding and non-coding regions (Kress and Erickson 2008). In the 3rd International Barcode of Life Conference, 2009, matK and rbcl were confirmed as Universal Plant DNA Barcodes. In the 4th International Barcode of Life Conference, 2011, matK, rbcl and ITS were proposed as barcode candidates for both plant and tree species. Earlier, quantitative measurements to compare barcode candidates were based on simple statistics calculations, PCR universality and sequence divergence. These analyses should allow direct comparison of putative DNA barcode markers.

From various studies, analysis and reports, it has been confirmed that the nuclear ITS2 and chloroplast psbA-trnH are capable of distinguishing closely related species. ITS2 proved to be better than psbA-trnH for the common medicinal plant species used in Chinese pharmacopoeia. ITS2 can also be used to group the plant samples into their right genus and has a relatively high accuracy for grouping samples into their exact species. The impressive practical value for DNA barcoding is to identify the plant specimens by individuals without enough taxonomic training and also it plays an important role in evolutionary studies (Coleman 2003, 2007). The ITS2 have also been proven to discriminate the plant species of Fabaceae (Gao et al. 2010a). Plant samples (1410) representing 893 species in 96 diverse genera from Rosaceae were evaluated using ITS2 as a barcode gene. ITS2 showed success rates of discriminating species of 78% and 100% at the species and genus levels, respectively (Pang et al. 2012). The ITS2 region is commonly used to authenticate Chinese herbal medicine (Coleman 2009; Song et al. 2009).

ITS2 has been shown to have a success rate of 80% for distinguishing between plant species of the Asteraceae (Gao et al. 2010b). ITS2 was also used to identify morphologically similar species such as *Swartzia grandifolia* and *S. longicarpa* for classification of the controversial species *Caranga rosea* and *C. sinica* of the plant Fabaceae (Gao et al. 2010a). Additionally, ITS2 has been shown to identify the system Eukaryota (Schultz et al. 2007), Pan-eukaryote (Coleman 2007) and as a universal barcode for plants and animals. There is also a report of using the chloroplast inter-gene region *psbA-trnH* to identify *Dendrobiun* species and it has been recommended as an ideal DNA barcode candidate (Song et al. 2009). DNA barcoding was employed to identify cryptic species distributed in the Velliangiri hills of India, adding value to both traditional and scientific knowledge (Newmaster and Raghupathy 2010). The same technique was applied to identify the Indian *Berberis* species (Roy et al. 2010) and Indian endangered species of *Paphiopedium* (Parveen et al. 2012). The ITS2 has also been used to distinguish the morphologically similar species of *Boerhavia* (Selvaraj et al. 2012). ITS2 in combination with rbcl has been used to identify contaminants in North American herbal products (Newmaster et al. 2013). And recently the chloroplast intergenic spacer (IGS) like *trnE-trnT, trnT- psbD, ndhF-rpl32* and *rpl14-rpl16* were applied for phylogenetic analysis and molecular evolution studies for Panax species (Kim et al. 2013). Additionally with coding regions of rbcl, matK, the IGS like *atpF-atpH, psbK-psbI* and *trnH-psbA* were used as DNA barcodes for discriminating the species of Korean Orchidaceae (Kim et al. 2013, in press).

DNA barcode candidate from the chloroplast genome like rbcl, matK, rpoCl, atpB, psbA-trnH (intergenic spacer) and nrITS (internal transcribed spacer) has been evaluated for the plant family of Apocynaceae. This study identifies the nrITS region to be better candidate than other regions tested. Especially the region ITS2 showed 4% intra-specific variation followed by *psbA-trnH*, 2%, whereas the other proposed candidates showed only 1% variation. The
parameters used to characterize the inter-specific divergence and intra-specific divergence based on within each genus and species. The identification efficiency was confirmed by two methods (a) BLAST1 method, and (b) Distance method at species level. ITS clearly distinguishes the 17 species of *Cynanchum* by using Minimum evolution method representing the out group *C. auriculatum_isolate* as shown in the Fig. 1. Henceforward, the present study concludes that ITS2 can be used to discriminate the plant species of Apocynaceae. However screening of more Apocynaceae species was required. Secondly the proposed DNA barcodes *rbcL*, *matK*, *psbA-trnH* and ITS2 were evaluated for the plant species of Zingiberaceae. The region ITS2 showed 4% intra-species variation followed by *psbA-trnH* (0.8%) shown in the Table 1. The analyses revealed, ITS2 region is able to distinguish all the tested species of Zingiberaceae. As Zingiberaceae is one of the larger families in the plant kingdom, the evaluations of the DNA barcodes have to be carried out for more number of species covering many genera in order to confirm the above results. Identification methods like BLAST and Distance were used for the gene ITS2 for the nine large genera, which were applied to species of Zingiberaceae (Table 2). These proposed barcode candidates were also applied to

![Phylogenetic tree constructed by Minimum Evolution method for the seventeen species of *Cynanchum* using ITS gene.](image)

**Table 1.** Analysis of interspecific divergence between congeneric species and intra-specific variation among the tested four loci for the plant family Zingiberaceae.

| Markers                          | *psbA-trnH* | ITS2     | *matK*    | *rbcL*    |
|----------------------------------|-------------|----------|-----------|-----------|
| All inter-specific distance      | 0.037 ± 0.112 | 0.143 ± 0.431 | 0.0157 ± 0.0308 | 0.0110 ± 0.0202 |
| Theta prime                      | 0.019 ± 0.049 | 0.122 ± 0.146 | 0.0175 ± 0.0212 | 0.0018 ± 0.0116 |
| Minimum inter-specific distance  | 0.013 ± 0.085 | 0.090 ± 0.208 | 0.003 ± 0.0260 | 0.0055 ± 0.0183 |
| All intra-specific distance      | 0.012 ± 0.0812 | 0.088 ± 0.180 | 0.0164 ± 0.079 | 0.0029 ± 0.0057 |
| Theta prime                      | 0.009 ± 0.273 | 0.023 ± 0.098 | 0.010 ± 0.0352 | 0.002 ± 0.0037 |
| Coalescent depth                 | 0.081 ± 0.202 | 0.078 ± 1.876 | 0.044 ± 0.799 | 0.001 ± 0.0373 |
differentiate and authenticate the medicinally important species of \textit{B. diffusa} from other species of \textit{Boerhavia} (\textit{B. repanda}, \textit{B. erecta} and \textit{B. verticillata}) using DNA barcodes. In this case ITS was able to distinguish \textit{B. diffusa} from other three species as shown in the Fig. 2. Hence, this study clearly proved that DNA barcoding could be used to authenticate the medicinal plant. More so, the information presented here is useful and applicable for identifying and discovering the new species as well as certifying the correct species of the plant lineages.

Fig. 2. Phylogenetic tree constructed by Minimum Evolution method for the fifteen species of \textit{Boerhavia} using ITS gene (Selvaraj \textit{et al.} 2012).

**Table 2.** Identification efficiency for nine large genera using ITS gene for the plant family Zingiberaceae.

| Genus      | Length (bp) | Number of species | % of Variation | Method          | % Success Identification |
|------------|-------------|-------------------|----------------|----------------|--------------------------|
| \textit{Afromum} | 605         | 53                | 0.3            | BLAST1 Distance | 89.0 92.2               |
| \textit{Boesenbergia} | 577         | 3                 | 0.09           | BLAST1 Distance | 93.0 89.6               |
| \textit{Curcuma} | 602         | 9                 | 0.09           | BLAST1 Distance | 93.0 89.6               |
| \textit{Globba} | 572         | 2                 | 0.06           | BLAST1 Distance | 90.0 97.6               |
| \textit{Haniffia} | 575         | 2                 | 0.0            | BLAST1 Distance | 100.0 100.0             |
| \textit{Hedychium} | 572        | 4                 | 0.03           | BLAST1 Distance | 98.0 97.0               |
| \textit{Kaempferia} | 689         | 3                 | 0.1            | BLAST1 Distance | 89.0 93.3               |
| \textit{Roscoea} | 569         | 5                 | 0.04           | BLAST1 Distance | 95.0 98.8               |
| \textit{Zingiber} | 450         | 21                | 0.21           | BLAST1 Distance | 81.2 88.3               |
ACKNOWLEDGMENTS

This research was supported by Golden Seed Project (Center for Horticultural Seed Development, No. 213003-04-1-CG100), Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Service (KFS).

REFERENCES

Albert VA, Backlund A, Bremer K, Chase MW, Manhart JR. 1994. Functional constraints and rbcL evidence for land plant phylogeny. Ann. Mo. Bot. Gard. 81: 534-567.

Amaral AR, Sequeira M, Coelho MM. 2007. A first approach to the usefulness of cytochrome c oxidase I barcodes in the identification of closely related delphinid cetacean species. Mar. Freshw. Res. 58: 505-510.

Anker A, Hurt C, Knowlton N. 2007. Revision of the Alpheus nuttingi (Schmitt) species complex (Crustacea: Decapoda: Alpheidae), with description of a new species from the tropical eastern Pacific. Zootaxa. 1577: 41-60.

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.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The ITS region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. Ann. Missouri Bot. Gard. 82: 247-277.

Bradley D, Gatenby AA. 1985. Mutational analysis of the maize chloroplast ATPase-P subunit gene promoter: The isolation of promoter mutants in E. coli and their characterization in a chloroplast in vitro transcription system. EMBO J. 4: 3641-3648.

Bucklin A,Wiebe PH, Smolenack SB, Copley NJ, Beaudet JG, Bonner KG, Farber-Lorda J, Pierson JJ. 2007. DNA barcodes for species identification of euhaploids (Euphausiacea, Crustacea). J. Plank Res. 29: 483-493.

Burgess KS, Fazekas AJ, Kesnakurthi RP, Graihan SW, Husband BC, Newmaster SG, Percy DM, Hajibabaei M, Barrett SCH. 2011. Discriminating plant species in a local temperate flora using the rbcL + matK DNA barcode. Methods Ecol. Evol. 2: 333-340.

Chase MW, Cowan RS, Hollingsworth PM, van der Berg C, Madrinan S, Petersen G, Seberg O, Jorgensen T, Cameron KM, Carine M, Pedersen N, Hedderon TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M. 2007. A proposal for a standardised protocol to barcode all land plants. Taxon. 56: 295-299.

Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesnakurthi RP, Haidar N, Savolainen V. 2005. Land plants and DNA barcodes: short-term and long-term goals. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360: 1889-1895.

Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang HX, Luo K, Li Y, Li XW, Jia XC, Lin YL, Leon C. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE. 5: e8613.

Cho Y, Mower JP, Qui YL, Palmer JD. 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proc. Natl. Acad. Sci. 101: 17741-17746.

Coleman AW. 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends Genet. 19: 370-375.

Coleman AW. 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucleic Acids Res. 35: 3322-3329.

Coleman AW. 2009. Is there a molecular key to the level of biological species in eukaryotes: A DNA guide. Mol. Phylogenet. Evol. 50: 197-203.

de Vere N, Rich TC, Ford CR, Trinder SA, Long C, Moore CW, Satterthwaite D, Davies H, Allainguillaume J, Ronca S, Tatarinova T, Garbett H, Walker K, Wilkinson MJ. 2012. DNA barcoding the native flowering plants and conifers of Wales. PLoS ONE 7: e37945.

Deng XW, Grutters W. 1987. Control of plastid gene expression during development: the limited role of transcriptional regulation. Cell 49: 379-387.

Doyle JJ. 1992. Gene trees and Species trees: Molecular Systematics as one character taxonomy. Sys. Bot. 17: 144-163.

Edwards D, Horn A, Taylor D, Savolainen V, Hawkins JA. 2008. DNA barcoding of a large genus, Aspalathus L. (Fabaceae). Taxon. 57: 1317-1327.

Ekrem T, Willassen E, Stur E. 2007. A comprehensive DNA sequence library is essential for identification with DNA barcodes. Mol. Phylogenet. Evol. 43: 530-542.
Utility of DNA Barcoding for Plant Biodiversity Conservation

Erickson DL, Spouge J, Resch A, Weigt LA, Kress WJ. 2008. DNA barcoding in land plants: developing standards to quantify and maximize success. Taxon. 57: 1304-1316.

Gadek PA, Wilson PG, Quinn CJ. 1996. Phylogenetic reconstruction in Myrtaceae using matK with particular to the position of Psiloxylon and Heteropyxis. Aust. Syst. Bot. 9: 283-290.

Galmes J, Flexas J, Keys AJ, Cifre J, Mitchell RAC. 2005. Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves. Plant Cell Environ. 28: 571-579.

Gao T, Yao H, Song J, Liu C, Zhu Y, Ma X, Pang X, Xu H, Chen S. 2010a. Identification of medicinal plants in the family Fabaceae using a potential DNA barcode ITS2. J. Ethnopharmacol. 130: 116 -121.

Gao T, Yao H, Song J, Zhu Y, Liu C, Chen S. 2010b. Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. BMC Evol. Biol. 10: 1-7.

Gatenby AA, Castleton JA, Saul MW. 1981. Expression in E. coli of maize and wheat chloroplast genes for large subunit of ribulose bisphosphate carboxylase. Nature 291: 117-121.

Gatenby AA, Rothstein SJ, Nomura N. 1989. Translational coupling of the maize chloroplast atpB and atpE genes. Proc. Natl. Acad. Sci. 86: 4066-4070.

Gerbi SA.1985. Evolution of ribosomal DNA. Mol. Biol. Evol. (ed. MacIntyre, R. J.), Plenum, New York, 419-517.

Goldman WE, Goldberg G, Bowman LH, Steinmetz D, Schlessinger D. 1983. Mouse rDNA: Sequences and evolutionary analysis of spacer and mature RNA regions. Mol. Cell Biol. 3: 1488-1500.

Gomez A, Wright PJ, Lunt DH, Cancino JM, Carvalho GR, Hughes RN. 2007. Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozaan taxon. Proc. R. Soc. B. Biol. Sci. 274: 199-207.

Gonalez IL, Chambers C, Gorski JL, Stambolian D, Schmickel RD, Sylvester JE. 1995. Sequence and structure correlation of human ribosomal transcribed spacers. J. Mol. Biol. 212: 27-35.

Gruissem W, Zurawski G. 1985. Analysis of promoter regions for the spinach chloroplast rbcL, atpB and psbA genes. EMBO J. 4: 3375-3383.

Hamby RK, Zimmer EA. 1992. Ribosomal RNA as a phylogenetic tool in plant systematics, p.50-91. In PS Soltis (ed.). Mol. Syst. Plants. Springer Press, Chapman & Hall, NY.

Hamilton MB. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8: 521-523.

Hamilton MB, Braverman JM, Soria-Hernanz DF. 2003. Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. Mol. Biol. Evol. 20: 1710-1721.

Handfield D, Handfield L. 2006. A new species of Plutia (Lepidoptera: Noctuidae) from North America. Can. Entomol. 138: 853-859.

Hanley-Bowdoin L, Chua NH. 1987. Chloroplast promoters. Trends Biochem. Sci. 12: 67-70.

He JX, Wen JQ, Chong K, Liang HG. 1998. Changes in transcript levels of chloroplast psbA and psbD genes during water stress in wheat leaves. Physiol. Plant. 102: 49-54.

Hebert PDN, Ratnasingham S, DeWaard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit I divergences among closely related species. Proc. Biol. Sci. 270: 96-99.

Hogg ID, Hebert PDN. 2004. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. Can. J. Zool. 82: 749-754.

Hollingsworth ML, Clark AA, Forrest LL, Richardson J, Pennington RT. 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.

Jarrell DC, Ciegg MT. 1995. Systematic implications of the chloroplast-encoded matK gene in the tribe Vandeae (Orchidaceae). Amer. J. Bot. 82: 137.

Johnson A, Soltis DE. 1994. matK DNA sequences and phylogenetic reconstruction in Saxifragaceae. Sys. Bot. 19: 143-156.

Johnson LA, Soltis DE. 1995. Phylogenetic inference in Saxifragaceae sensu stricio and Gilia (Polémonioaceae) using matK sequences. Ann. Mo. Bot. Gard. 82: 149-175.

Kellogg EA, Juliano ND. 1997. The structure and function of RUBISCO and their implications for systematic studies. Amer. J. Bot. 84: 413-428.

Kim JH, Jung JY, Choi HI, Kim NH, Park JY, Lee Y, Yang TJ. 2013. Diversity and evolution of major Panax species revealed by scanning the entire chloroplast intergenic spacer sequences. Genet. Resour. Crop Evol. 60: 41 -425.
Kim HM, Oh SH, Park CW. 2013. DNA barcoding of Orchidaceae in Korea. Mol. Ecol. Resour. (In Press).

Kress WJ, Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding \textit{rbcL} gene complements the non-coding \textit{trnH-psbA} spacer region. PLoS ONE. 2: e508.

Kress WJ, Erickson DL. 2008. DNA barcodes: Genes, genomics, and bioinformatics. Proc. Natl. Acad. Sci. 105: 2761-2762.

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

Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F. 2008. DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. 105: 2923-2928.

Liang H, Hilu KW. 1996. Application of the \textit{matK} gene sequences to grass systematics. Can. J. Bot. 74: 125-134.

Manhart JR, McCourt RM. 1992. Molecular data and species concepts in the algae. J. Phycol. 28: 730-737.

Manuell A, Beligni MV, Yamaguchi K, Mayfield SP. 2004. Regulation of chloroplast translation: interactions of RNA elements, RNA-binding proteins and the plastid ribosome. Biochem. Soc. Trans. 32: 601-605.

McCauley DE. 1995. The use of chloroplast DNA polymorphism in studies of gene flow in plants. Trends Ecol. Evol. 10: 190-202.

Meyer CP, Paulay G. 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. 3: 2229-2238.

Mohr G, Perlman PS, Lambowttz AM. 1993. Evolutionary relationships among group II intron encoded proteins and identification of a conserved domain that may be related to maturase function. Nucleic Acids Res. 21: 4991-4997.

Monde RA, Greene JC, Stern DB. 2000. The sequence and secondary structure of the 3’-UTR affect 3’-end maturation, RNA accumulation, and translation in tobacco chloroplasts. Plant Mol. Biol. 44: 529-542.

Musters W, Boon K, van der Sande CAFM, van Heerikhuizen H, Planta RJ. 1990. Function analysis of transcribed spacers of yeast ribosomal DNA. EMBO J. 9: 3989-3996.

Newmaster SG, Grgeric M, Shanmughanandhan D, Ramalingam S, Ragupathy S. 2013. DNA barcoding detects contamination and substitution in North American herbal products. BMC Med. 11: 222.

Newmaster SG, Ragupathy S. 2010. Ethnobotany genomics - discovery and innovation in a new era of exploratory research. J. Ethnobiol. Ethnomed. 6: 1746-4269.

Olmstead RG, Michaels HJ, Scott KM, Palmer JD. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of \textit{rbcL}. Ann. Missouri Bot. Gard. 79: 249-265.

Pang X, Liu C, Shi L, Liu R, Liang D, Li H, Chen SS, Chen S. 2012. Utility of the \textit{trnH-psbA} intergenic spacer region and its combinations as plant DNA barcodes: a meta-analysis. PLoS ONE 7: e48833.

Parveen I, Singh HK, Raghuvanshi S, Pradhan UC, Babbar SB. 2012. DNA barcoding of endangered Indian Paphiopedilum species. Mol. Ecol. Resour. 12: 82-90.

Pennisi E. 2007. Wanted: A barcode for plants. Science. 318: 190-191. doi: 10.1126/science.318.5848.190.

Pfenninger M, Nowak C, Kley C, Steinke D, Streit B. 2007. Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic \textit{Chironomus} (Diptera) species. Mol. Ecol. 16: 1957-1968.

Provan J, Powell W, Hollingsworth PM. 2001. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trends Ecol. Evol. 16: 142-147.

Rastogi G, Dharme MS, Walujkar S, Kumar A, Patole MS, Shouche YS. 2007. Species identification and authentication of tissues of animal origin using mitochondrial and nuclear markers. Meat Sci. 76: 666-674.

Ratnasingham S, Hebert PDN. 2007. BOLD: The Barcode of Life Data system. (www.barcodinglife.org). Mol. Ecol. Notes 7: 355-364.

Rogers SO, Bendich AJ. 1987. Ribosomal genes in plants: Variability in copy number and in the intergenic spacer. Plant Mol. Biol. 9: 509-520.

Roy S, Tyagi A, Shukla V, Kumar A, Singh UM, Chaudhary LB, Datt B, Bag SK, Singh PK, Nair NK, Husain T, Tuli R. 2010. Universal plant DNA barcode loci may not work in complex groups: A case study with Indian berberis species. PLoS ONE 5: e13674.

Rubinoff D. 2006. Utility of mitochondrial DNA barcodes in species conservation. Conserv. Biol. 20: 1026-1033.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K, Beeson K, Tran B, Smith H, Baden-Tillson H, Stewart C, Thorpe J, Freeman J, Andrews-Pfannkoch C, Venter JE, Li K, Kravitz S, Heidelberg JF, Utterback T, Rogers YH, Falcon Li, Souza V, Bonilla-Rosso G, Eguiarte LE, Karl DM, Sathyendranath S, Platt T, Bemingham E, Gallardo V, Tamayo-Castillo G, Ferrari MR, Strausberg RL, Nealon K, Friedman R, Frazier M, Venter JC. 2007. The Sorcerer II global ocean
Utility of DNA Barcoding for Plant Biodiversity Conservation

Sass C, Little DP, Stevenson DW, Specht CD. 2007. DNA barcoding in the Cycadales: Testing the potential of proposed barcoding markers for species identification of cycads. PLoS ONE 2: e1154.

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. London Ser. B. 360: 1805-1811.

Schindel DE, Miller SE. 2005. DNA barcoding a useful tool for taxonomists. Nature 435: 17.

Schultz J, Maisel S, Gerlach D, Muller T, Wolf M. 2007. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the eukaryote. RNA 11: 361-364.

Selvaraj D, Sarma RK and Sathishkumar R. 2008. Phylogenetic analysis of chloroplast matK gene from Zingiberaceae for plant DNA barcoding. Bioinform. 3: 24-27.

Selvaraj D, Shammuganandhan D, Sarma RK, Joseph J, Srinivasan RV, Ramalingam S. 2012. DNA Barcode ITS Effectively Distinguishes the Medicinal Boerhavia diffusa from Its Adulterants. Genom. Proteom. Bioinform. 10: 364-367.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast. Am. J. Bot. 92: 142-166.

Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PDN. 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). Proc. Natl. Acad. Sci. 103: 3657-3662.

Soltis PS, Kuzoff RK. 1993. ITS sequences variation within and among population of Lomatium grayi and L. laevigatum (Umbelliferae). Mol. Phylogenet. Evol. 2: 166-170.

Song J, Yao H, Li Y, Li X, Liu C, Han J, Xie C, Chen S. 2009. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. J. Ethnopharmacol. 124: 434-439.

Steele KP, Vilgalys R. 1994. Phylogenetic analysis of Polemoniaceae using nucleotide sequences of the plastid gene matK. Syst. Bot. 19: 126-142.

Stewart MA, Hall LMC, Maden BEH. 1983. Multiple heterogeneities in the transcribed spacers of ribosomal DNA from Xenopus laevis. Nucleic Acids Res. 10: 2851-2864.

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

Summerbell RC, Levesque CA, Seifert KA, Bovers M, Fell JW, Diaz MR, Boekhout T, de Hoog GS, Stalpers J, Crous PW. 2005. Microcoding: the second step in DNA barcoding. Philos. Trans. R. Soc. London Ser. B. 360: 1897-1903.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Verma T, Corthier G, Brochmann C, Willerslev E. 2006. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35: e14.

Tavares ES, Baker AJ. 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. BMC Evol. Biol. 8: 81.

Teletchea T, Maudet C, Hanni C. 2005. Food and forensic molecular identification: update and challenges. Trends Biotechnol. 23: 359-366.

Valentini A, Mattiucci S, Bondanelli P, Webb SC, Antonio A, Giannone M, Marlene Colom-Llavina M, Nascetti G. 2006. Genetic relationships among Anisakis species (Nematoda: Anisakidae) inferred from mitochondrial COX2 sequences, and comparison with allozyme data. J. Parasitology 92: 156-166.

van der Sande CAFM, Kwa M, van Nues RW, van Heerikhuizen H, Raue HA, Planta RJ. 1992. Functional analysis of internal transcribed spacer 2 of Saccharomyces cerevisiae ribosomal DNA. J. Mol. Biol. 223: 899-910.

van Nues RW, Rientjes JMJ, van der Sande CAFM, Zerp SF, Sluiter C, Venema J, Planta RJ, Raue HA. 1994. Separate structural elements within internal transcribed spacer 1 of Saccharomyces cerevisiae precursor ribosomal RNA directly the formation of 17S and 26S rRNA. Nucleic Acids Res. 22: 912-919.

van Straalen NM, Roelofs D. 2012. An Introduction to Ecological Genomics. Oxford University Press, Oxford.

van Velzen R, Freek T, Bakker L, Van Loon JJA. 2007. DNA barcoding reveals hidden species diversity in Cymothoe (Nymphalidae). Proc. Neth. Entomol. Soc. Meet. 18: 95-103.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. 2002.
Phylogenies and community ecology. Annu. Rev. Ecol. Syst. 33: 475-505.

Wolfe KH, Morden CW, Palmer JD. 1992. Function and evolution of a minimal plastid genome from a non-photosynthetic parasitic plants. Proc. Natl. Acad. Sci. 89: 10648-10652.

Zurawski G, Bohnert HJ, Whitfeld PR, Bottomley W. 1982. Nucleotide sequence of the gene for the MR 32,000 thylakoid membrane protein from *Spinacia oleracea* and *Nicotiana debneyi* predicts a totally conserved primary translation product of MR 38,950. Proc. Natl. Acad. Sci. 79: 7699-7703.