Int. j. adv. multidisc. res. stud. 2024; 4(2):305-308

Received: 15-01-2024
Accepted: 05-03-2024

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Utilizing DNA Barcoding for Enhanced Plant Species Identification: An In-Depth Review

[1] Prajapati Bhoomi K, [2] Solanki Hitesh A

[1] Ph.D. Research Scholar, Department of Botany, Bioinformatics and Climate Change Impacts Management, India
[2] Professor and Head, Department of Botany, Bioinformatics and Climate Change Impacts Management, India

Corresponding Author: **Prajapati Bhoomi K**

Abstract

Plant identification is crucial and routine taxonomic procedure in order to understand and conserve the biodiversity. Recently many tools and techniques are available for identification and conservation. One of them is DNA barcoding. DNA barcoding is a diagnostic technique for species identification using a short, standardized DNA region. DNA barcoding is an expedient sequence-based means of identifying species across all stages of development and even from trace amount of tissue. The aim DNA barcoding is to establish a shared community resource of DNA sequences that can be used for identification, discrimination or taxonomic classification of organisms. DNA barcoding may provide a rapid genetic screening tool to identifying selected noxious grass weeds at the vegetative growth. DNA barcoding is well established in animals but there is not yet any universally accepted barcodes for plants. Hence, working on monocotyledons is challenging and as well as interesting. There is currently no consensus on which candidate markers comprise the best plant DNA barcoding region; however, DNA barcodes such as *rbcL*, *matK*, *psbA-trnH* and *ITS* have been proposed for the plant kingdom. Present study provides an account of some DNA barcoding research that have been done previously on plants. Also presenting the future aspects for identifying the plant species through DNA barcoding.

Keywords: DNA-Barcoding, Taxonomy, Plants, Monocotyledons, Grass, rbcL, matK

Introduction

The Earth harbours an estimated 300,000 plant species, encompassing various categories such as flowering plants (87%), non-flowering plants (0.32%), ferns (4.4%), mosses (5%), and red and green algae (3.3%) (World Conservation Union, IUCN online). Biodiversity conservation is increasingly recognized as imperative from multidisciplinary perspectives (Qureshimatva *et al*., 2016) [17]. However, the rapid loss of species due to extinction poses a significant challenge, with many species potentially disappearing before they are even cataloged by scientists. Taxonomy, the discipline responsible for species classification, is undergoing significant challenges concerning its position within biology and its intersection with conservation biology (Prajapati, Maurya, and Solanki, 2020) [2].

Species identification, pivotal in taxonomic studies, involves the comparison of unknown specimens with known references or descriptions (Prajapati, Maurya, and Solanki, 2020) [2]. The development of reliable methods for discriminating among diverse species is a formidable task. DNA barcoding has emerged as a promising technique, leveraging genetic information to provide consistent and conclusive identification results (Bandyopadhyaya *et al*., 2013 as cited in Hebert *et al*., 2003 and Hebert *et al*., 2004) [1, 8, 9]. This diagnostic approach utilizes DNA sequences from specific genomic regions to aid in species identification, particularly beneficial in ecological and conservation studies where traditional taxonomic methods may be impractical (Lahaye *et al*., 2008) [12].

DNA barcoding is acclaimed as a transformative tool in taxonomy, facilitating efficient organization and analysis of specimen data for systematic research (DeSalle and Goldstein, 2019) [4]. The grass family, one of the largest among flowering plants, exemplifies the significance of DNA barcoding in elucidating evolutionary relationships (Nadai and Wilkinson, 2020) [7]. Over the past two decades, DNA barcoding has been instrumental in molecular systematics, enhancing our understanding of evolutionary patterns.

While considerable sequence data exist in public databases for established barcoding loci, further efforts are needed to augment these resources for broader applications of DNA barcoding (Crautlein *et al*., 2011) [20]. Continued expansion of public sequence libraries, coupled with advancements in sequencing technologies, promises to amplify the utility of DNA barcoding in species identification and classification.

DNA barcoding in taxonomy:

- The present study evaluated the potentiality of seven Consortium for Barcode of Life (CBOL) recommended standard DNA barcode regions in commercially important bamboo species of India. Among the analyzed barcode regions, multiple sequence alignment (MSA) of psbA-trnH barcode region showed species-specific nucleotide differences in the studied bamboo taxa. The major nucleotide changes observed were transitions/transversions as well as insertions/deletions of nucleotides. Even though species-specific mononucleotide differences could be identified for most of the studied bamboo taxa, a small amount of sequence similarities was found in some of the Dendrocalamus and Bambusa species, which were grouped together in tree-based analysis. In subtribe Melocanninae, Ochlandra travancorica, Melocanna baccifera and M. clarkei showed unique species-specifc psbA-trnH barcodes. Similarly, in the genus Oxytenanthera, unique species-specifc psbA-trnH barcodes were obtained for O. monadelpha and O. parvifolia. Thus psbA-trnH barcode region generated distinct species-specifc barcodes for commercial bamboo species in genera Bambusa, Dendrocalamus, Melocanna, Oxytenanthera as well as Ochlandra (Dev *et al.*, 2020) [5].
- Wang *et al.*, have developed a simple and rapid DNA-based molecular identification system for the Lemnaceae based on sequence polymorphisms. They compared the barcoding potential of the seven plastid-markers proposed by the CBOL (Consortium for the Barcode of Life) plant-working group to discriminate species within the land plants in 97 accessions representing 31 species from the family of Lemnaceae. A Lemnaceae-specific set of PCR and sequencing primers were designed for four plastid coding genes (rpoB, rpoC1, rbcL and matK) and three noncoding spacers (atpF-atpH, psbK-psbI and trnH-psbA) based on the Lemna minor chloroplast genome sequence. We assessed the ease of amplification and sequencing for these markers, examined the extent of the barcoding gap between intra and inter-specific variation by pairwise distances, evaluated successful identifications based on direct sequence comparison of the "best close match" and the construction of a phylogenetic tree.
- Lahaye *et al.*, 2008 undertook intensive field collections in two biodiversity hotspots (Mesoamerica and southern Africa). Using >1,600 samples, they compared eight potential barcodes. Going beyond previous plant studies, they assessed to what extent a ''DNA barcoding gap'' is present between intra- and interspecific variations, using multiple accessions per species. Given its adequate rate of variation, easy amplification, and alignment, identified a portion of the plastid matK gene as a universal DNA barcode for flowering plants. Critically, they further demonstrate the applicability of DNA barcoding for biodiversity inventories. In addition, analysing >1,000 species of Mesoamerican orchids, DNA barcoding with matK alone reveals cryptic species and proves useful in identifying species listed in Convention on International Trade of Endangered Species (CITES) appendixes.
- Miguez *et al.*, 2021 [13] performed a DNA barcoding study based on four DNA regions (two nuclear, two plastid). The results reveal that none of previous mainland Italian records corresponds to C. microcarpa, but to C. pendula, which also belong to section Rhynchocystis, and in other cases possibly to C. acutiformis, from a different species group (sect. Paludosae). The specimen from Montecristo Island was confirmed to be C. microcarpa.
- Wang *et al.*, 2014 have tested the utility of 18 chloroplast and nuclear genes as potential DNA barcodes for species identifications of introduced grasses present in Eastern Australia. Grasses examined (N = 417) included Nassella neesiana (Trin. & Rupr.) Barkworth (Chilean needle grass), Nassella trichotoma (Nees) Hack. ExArechav. (Serrated tussock), Eragrostis curvula (Schrad.) Nees (African love grass) and 26 other weed species collected from New South Wales (NSW), Australian Capital Territory (ACT) and other parts of Australia. Our preliminary results revealed three chloroplast genes, matK, ndhK and petL, which exhibit potentials as DNA barcodes for distinguishing and identifying weeds species of interest.
- In the present study, Rebijith *et al.*, 2013 employed CO-I for discriminating 142 individuals representing 32 species of aphids from India. Sequence analyses revealed that the intraspecific and interspecific distances ranged from zero to 3.8% and 2.31 to 18.9%, respectively. In addition, the study also showed for the first time the prevalence of three cryptic species, namely Brevicoryne brassicae (Linnaeus), Hyperomyzus carduellinus (Theobald) and Brachycaudus helichrysi (Kaltenbach) from India.
- Recent results, however, from the two chosen core plant DNA barcode regions rbcL and matK plus two supplementary regions trnH –psbA and internal transcribed spacer (ITS) (or ITS2) have demonstrated reasonable levels of species discrimination in both floristic and taxonomically focused studies. They have described sampling techniques, extraction protocols, and PCR methods for each of these two cores and two supplementary plant DNA barcode regions, with extensive notes supporting their implementation for both low- and high-throughput facilities (Fazekas *et al.*, 2012) [6].
- Muasya *et al.*, 2009 presented an analysis of 262 taxa representing 93 genera in 15 tribes, sequenced for the plastid rbcL and trnL-F (intron and intergenic spacer). Cyperaceae are monophyletic and resolved into two clades, here recognised as Mapanioideae and Cyperoideae, and the overall topology is similar to results from previous studies. Within Cyperoideae, Trilepideae are sister to rest of taxa whereas Cryptangieae, Bisboeckelerieae and Sclerieae are resolved within Schoeneae. The phylogenetic position of 40 species in 21 genera is presented in this study for the first time, elucidating their position in Abildgaardieae (Trachystylis), Cryptangieae (Didymiandrum, Exochogyne), Cypereae

(Androtrichum, Volkiella), Eleocharideae (Chillania), and Schoeneae (Calyptrocarya, Morelotia). More sampling effort (more taxa and the use of more rapidly evolving markers) is needed to resolve relationships in Fuireneae and Schoeneae.

- A chloroplast DNA (cpDNA) based cleaved amplified polymorphic sequence (CAPS) system of molecular species diagnosis that has the capacity to address the identification problem is presented using British grasses as a model. First PCRs were performed using primer pairs targeting 21 regions of the chloroplast genome in authenticated representative of the 117 grass species from British Isles, and universal amplification for all loci targeted was demonstrated. Another one, 54 restriction enzymes were applied on amplification products generated from all species. There were 10 locus-enzyme combinations that had the best diagnostic utility for 117 grass species. CAPS analysis on 16 representatives of three genera (Calamagrostis, Phleum and Agrostis) was then used to illustrate the potential utility of the pipeline for establishing a field-laboratory screen of species identity. CAPS DNA barcoding system developed here may have ecological, conservation and commercial applications (Haider, N. and Wilkinson, M.J., 2020) [7].
- The study has been demonstrated that arpF-atpH noncoding spaces could serve as a universal DNA barcoding marker for species-level identification of duckweeds. This marker will allow to identify unknown species or to exploit new species of duckweeds by reason of its reliable amplification, straightforward sequence alignment and rates of DNA variation between species and within species. DNA barcoding developed in this study are a significant contribution to the taxonomical structure in duckweeds compared with insensitive morphological classification (Wang *et al.*, 2010) [22].
- A global plant DNA barcode system is evaluated by comparing universal application and degree of sequence divergence for nine putative barcode loci, including coding and non-coding region, singly and in pair across a phylogenetically diverse set of 48 genera. No single locus could discriminate among species in a pair in more than 79% of genera, whereas discrimination increased to nearly 88% when non-coding trnH-psbA spacer was paired with one of three coding loci, including rbcL. Insilico trails were conducted in which DNA sequences from GenBank were used to further evaluate the discriminatory power of a subset of these loci. These trails supported the earlier observation that trnH-psbA coupled with rbcL can correctly identify and discriminate among related species. A combination of non-coding trnH-psbA spacer region and a portion of the coding rbcL gene is recommended as a two-locus global land plant barcode that provides the necessary universality and species discrimination (Kress and Erickson, 2007) [11].
- Birch *et al.*, 2017 have generated a tribe Poeae reference library and new sequence data for the official plastid barcoding (rbcL and matK) and associated (ITS) markers with comprehensive representation across the Australian continent. Using the ITS dataset, for the tribe Poeae in Australia they were able to correctly identify 97.6% of individuals to genera and 32.4% of individuals to species, based on the BCM distance-based method. The nearest neighbour method provided a higher percentage of correct specimen identifications, but suffered from a large number of incorrect identifications at the species rank. The TID method typically provided the lowest percentage of incorrect specimen determination with the cost of the stringency in the criteria applied being a lower percentage of correct identifications. A barcode gap that facilitated determination of species was identified for smaller genera of tribe Poeae including Briza, Catapodium, Cynosurus and Hookerochloa. Based on the ITS dataset and applying the liberal tree-based method to assess the maximum likehood phylogeny they were able to correctly identify 97.4% of individuals to genera and 28.5% of individuals to species. Tree-based method correctly identified almost all exotic species, including those in genera containing native species i.e. Poa, Festuca and Puccinellia.
- Genetic diversity analysis contributes to the conservation, protection and utilization of genetic resources towards efficient management of germplasm. The study shows, inter simple sequence repeats markers were used to analyse the genetic diversity of 32 genotypes of *Acmella paniculata* collected from 12 different states of Gujarat, India. The study reveals that there is no correlation between geographical distance diversity among the populations studied. The resultant data generated from ISSR molecular-based genetic diversity analysis of *A. paniculata* in Gujarat can provide a reference for the conservation and efficient management of the important Indian genetic resources. (S Patel *et al.*,2022) [15].
- Total 21 endangered plants samples were collected. Plants were selected on the basis of their conservation status listed in The Flora of Gujarat and synonym and accepted name on the basis of the plant list. Highly amplified sequences were obtained from all plant DNA samples. All plant species showed 99-100% match with BLAST algorithm nucleotide search of NCBI. Sequence alignment further process with clustalW for phylogenetic tree by MEGA X. The present study shows DNA barcode has been successfully amplified of all endangered species which shows individual barcode for individual plant which would be help in future identification, discrimination, closeness and evolutionary trend among their related species (N Purohit and H Solanki, 2019).

Conclusion

There is a considerable optimism about the use of DNA barcoding in providing quick and reliable information about species that can help in conservation strategies. Despite successful uses of genetic barcoding in classifying taxa, it is perhaps premature to suggest that DNA barcoding can offer replacement paradigm for assessing and understanding biodiversity. DNA barcodes have been also helpful in identifying cryptic diversity within species, especially when used in conjunction with other taxonomic criteria such as morphology and ecology. The role of DNA barcodes in identifying undescribed species ambiguous: Barcodes can help in speeding up biodiversity assessment but can provide information in the way that complex species can.

Future scope:

- More exploration for floristic accounts.
- Plant DNA barcodes can be used to assess species identification in conservation biodiversity hotspots.
- Using multiple barcoding regions can help to differentiate closely related species.
- DNA barcoding can help to identify and conserve critical grass species.

References

1. Bandyopadhyaya Shruti, Ramakrishnan M, Kulkarni Shruti, Rajanna L. DNA barcoding and its applications: A critical review. Cmrims J. Res. and Dev. 2013; 1(1):77-81.
2. Bhoomi Prajapati, Rupesh Maurya, Hitesh Solanki. Floristic account of Central Gujarat special reference to Ahmedabad: A review. International Journal of Engineering Development and Research (IJEDR). 2020; 8(1):535-538. ISSN:2321-9939
3. Birch JL, Walsh NG, Cantrill DJ, Holmes GD, Murphy DJ. Testing efficacy of distance and tree-based methods for DNA barcoding of grasses (Poaceae tribe Poeae) in Australia. PLoS One. 2017; 12(10):e0186259.
4. DeSalle R, Goldstein P. Review and interpretation of trends in DNA barcoding. Frontiers in Ecology and Evolution. 2019; 7:302.
5. Dev SA, Sijimol K, Prathibha PS, Sreekumar VB, Muralidharan EM. DNA barcoding as a valuable molecular tool for the certification of planting materials in bamboo. 3 Biotech. 2020; 10(2):1-12.
6. Fazekas AJ, Kuzmina ML, Newmaster SG, Hollingsworth PM. DNA barcoding methods for land plants. In DNA barcodes. Humana Press, Totowa, NJ, 2012, 223-252.
7. Haider N, Wilkinson MJ. Caps DNA barcoding for field laboratory identification of grass species (British grasses as a model). Agriculture (Pol'nohospodárstvo). 2020; 66(2):74-86.
8. Hebert PDN, Cywinska A, Ball SL, de Waard JR. Proc. R. Soc. London Ser. B. 2003; 270:313-321.
9. Hebert PDN, Stoeckle MY, Zemlak TS, *et al*., PLoS Biol. 2004; 2:312.
10. https://www.iucn.org/ (International Union for Conservation of Nature- IUCN)
11. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS one. 2007; 2(6):e508.
12. Lahaye R, Van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, *et al*. DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences. 2008; 105(8):2923-2928.
13. Miguez M, Bartolucci F, Jiménez-Mejías P, Martín-Bravo S. Re-evaluating the presence of Carex microcarpa (Cyperaceae) in Italy based on herbarium material and DNA barcoding. Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology, 2021, 1-7.
14. Muasya AM, Simpson DA, Verboom GA, Goetghebeur P, Naczi RF, Chase MW, *et al*. Phylogeny of Cyperaceae based on DNA sequence data: Current progress and future prospects. The Botanical Review. 2009; 75(1):2-21.
15. Patel S, Purohit N, Sapra P, Solanki H, Bishoyi AK. Genetic Diversity Analysis Revealed the Hot Spot of Acmella Paniculata Existing in Natural Populations of Gujarat, 2022.
16. Purohit NM. DNA barcoding of some selected endangered species, 2019.
17. Qureshimatva UM, Maurya RR, Gamit SB, Solanki HA. Check List of Flowering Plants in Panchmahal District, Gujarat. India. Forest Res. 2016; 5(176):2.
18. Rebijith KB, Asokan R, Kumar NK, Krishna V, Chaitanya BN, Ramamurthy VV. DNA barcoding and elucidation of cryptic aphid species (Hemiptera: Aphididae) in India. Bulletin of entomological research. 2013; 103(5):601-610.
19. Shah GL. Flora of Gujarat state, 1978.
20. von Cräutlein M, Korpelainen H, Pietiläinen M, Rikkinen J. DNA barcoding: A tool for improved taxon identification and detection of species diversity. Biodiversity and conservation. 2011; 20(2):373-389.
21. Wang A, Gopurenko D, Wu H, Stanton R, Lepschi BJ. DNA barcoding for identification of exotic grass species present in eastern Australia. In 19th Australasian Weeds Conference (2014) (pp. 444-447). Tasmanian Weed Society Inc, 2014.
22. Wang W, Wu Y, Yan Y, Ermakova M, Kerstetter R, Messing J. DNA barcoding of the Lemnaceae, a family of aquatic monocots. BMC Plant Biology. 2010; 10(1):1-11.