ISSN: 2348-1900
Plant Science Today
http://www.plantsciencetoday.online

Mini Review

DNA barcoding of plants: Selection of core markers for taxonomic groups

Ankush Ashok Saddhe and Kundan Kumar*

Department of Biological Sciences, Birla Institute of Technology & Science Pilani, K. K. Birla Goa Campus, Goa 403726, India

Article history
Received: 16 October 2017
Accepted: 17 November 2017
Published: 01 January 2018

Editor
K K Sabu

Publisher
Horizon e-Publishing Group

Correspondence
Kundan Kumar
✉ kundan@goa.bits-pilani.ac.in

Abstract

Plant identification is a crucial and routine taxonomic procedure in order to understand and conserve the biodiversity. Anthropogenic activity, pollution, deforestation, and exploitation of natural resources have been threatening to the plant biodiversity. Unfortunately, the major concern of traditional identification of plants is the gradual declined number of taxonomic expertise and lack of tools which accurately discriminate plant seeds, plant parts and seedling, and herbal adulterant. Presently, it is of utmost importance that plant biodiversity to be preserved. To overcome this issues the advent of molecular marker based technique which utilized short fragment of DNA and correctly assign plant taxa to their taxonomic group, called as DNA barcoding. First time, single marker based taxon identification successfully implemented to an animal taxa using mitochondrial cytochrome I (COI) gene fragment. However, Plant DNA barcoding is more complex and it often requires more than one set of DNA markers. In the present review, we have compiled the recent progress of plant DNA barcoding in various taxonomic groups and utility of plastids and nuclear DNA based markers for plant identification.

Keywords

DNA barcoding; *rbcL*; *matK*; ITS2

Citation

Saddhe A A, Kumar K. DNA barcoding of plants: Selection of core markers for taxonomic groups. Plant Science Today 2018;5(1):9-13. doi: https://dx.doi.org/10.14719/pst.2018.5.1.356

Introduction

Plant biodiversity is an essential and irreplaceable component of the ecosystem. In the present scenario, biodiversity hotspots are vulnerable due to habitat fragmentation, introduction of exotic species, overexploitation of species and anthropogenic activity. In order to identification, classification and conservation of plant species, present traditional taxonomic expertise is inadequate. Recently, the alternative revolutionary approach based on DNA marker was successfully introduced for an animal taxa using mitochondrial COI gene (1, 2). In contrast, plant DNA barcoding is a more complex and it often requires multiple loci. The Consortium for the Barcode of Life (CBOL) plant working group evaluated the efficacy of maturase K (*matK)* and ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) and recommended two-locus based approach with *trnH-psbA* intergenic spacer as a supplementary marker (3). China Plant Barcode of Life recommended the internal transcribed spacer (ITS) as additional candidate plant DNA barcode. Comparative studies of seven markers *trnH-psbA*, *matK*, *rbcL*, chloroplast RNA polymerase subunit (*rpoC1*), *ycf5*, ITS2, and ITS from medicinal plant species were performed (4). Authors recommended that ITS2 is the best potential marker which discriminated 92.7% plants at the species level in more than 6600 plant samples (5). However, most of plant taxonomists have suggested that a multi-locus approach may be essential to resolve plant species (6). Beside all these markers, several plastid regions such as *ycf1*, *atpF-H*, *psbK-psbI*, *ropC1*, *rpoB*, and

trnL-trnF were frequently evaluated as plant barcode. However, the application of DNA barcoding has been hindered owing to the difficulty in distinguishing closely related species, especially in recently diverged taxa. The plastid markers *rbcL* and *matK* loci exhibited poor resolution in species-rich genera and complex taxa of *Lysimachia*, *Ficus*, *Holcoglossum*, and *Curcuma* (7-10). However, DNA barcoding has significant impact on various research areas such as molecular phylogeny, population genetics, evolution and ecology, biosecurity and food product regulation (6, 11, 12). It helps to detect adulterant in food and medicinal product (6, 11). In recent years, identification and authentication of medicinal plants using DNA barcode markers have made significant progress (6, 11).

Here, we have discussed recent progress of plant DNA barcoding and evaluation of the potential new DNA candidate markers for plant identification. Most of the DNA barcoding works mainly focused on angiosperm, however very few reports are available on DNA barcoding of algae, bryophytes, pteridophytes and gymnosperms. Most commonly used DNA barcode markers utilized in plant identification is depicted in Fig 1. The complete list of DNA barcodes markers used for taxonomic identification is given in Table 1. CBOL recommended two marker based approach for plant identification but still in some group additional group specific markers need to be incorporated. We summarized current update of plant DNA barcoding according to groups such as algae, bryophytes, pteridophytes, gymnosperms and angiosperms.

DNA barcoding of algae

Algae are highly diverse group of organisms and classified into six major groups comprised of Chlorophyta (green algae), Rhodophyta (red algae), Phaeophyta (brown algae), Chrysophyta (golden algae), Bacillariophyta (diatoms), and Ulvophyceae (green algae). Their diversity is reflected at the morphological, structural, genetic, biochemical, physiological and ecological level (13). In addition, there is increased commercial importance of algae group such as ecological bioindicator, production of biofuel, food and fodder for animals (14).The algae taxonomy is a more tedious and difficult to identify microscopic and cryptic species. However, DNA barcoding opened the new alternative and confined ways to identify algal species regardless of life stage. Many DNA markers were evaluated including chloroplast (*rbcL*, *tufA* and 23S), mitochondrial (*COI*) and nuclear genes (18S rDNA, nuITS1 and nuITS2) (15-18). The protist working group of the CBOL recommended two step barcoding in which a universal barcode marker should be used first, followed by the use of a group-specific second barcode (19).

DNA barcoding of bryophytes

Bryophytes comprise three different phylogenetic lineages such as liverworts, hornworts, and mosses. They are the oldest land plants on earth and play an essential ecological role in various ecosystems. However, conservation strategies of bryophytes are always overlooked because of inadequate taxonomic expertise due to miniature size and small distinguish features. The development of new molecular identification tools for bryophytes would improve the ecological studies and help in investigating the impact of global climate change. Recently the closely related *Dicranum scoparium* species were collected from the high Arctic Archipelago of Svalbard resolved by combining five plastid regions (*rpoB*, *trnH-psbA*, *trnL-trnF*, *rps4-trnT*, *rps19-rpl2*) and the nuclear ribosomal ITS region (20). DNA barcoding of moss species diversity such as *Schistidium* species colonizing modern building surfaces showed morphological differences, and suggested cryptic taxa (21). Total 10 DNA barcode markers including proposed region (*atpF-atpH,* ITS2, *matK, psbK-psbI, rbcL, rpoB, rpoC1*, and *trnH-psbA*) and two popular phylogenetic markers (*rps4* and *trnL-trnF*) were tested in 49 moss species and 9 liverwort species (22).

DNA barcoding of pteridophytes

Pteridophytes comprised ferns and lycophytes which are seedless vascular land plants possessing distinct, free-living sporophyte (2n) and gametophyte (1n) generations (23). Japanese pteridophytes were resolved based on traditional as well as DNA barcode approach and the efficacy of two proposed plastid barcode markers such as *rbcL* and *trnH-psbA* were tested (23). The discriminatory power of the core DNA barcode (*rbcL* and *matK*), and supplementary proposed fern barcodes (*trnH-psbA* and *trnL-F*), were tested across two genera in the hyper diverse polypod clade *Deparia* (Woodsiaceae) and the *Cheilanthes marginata* group (24). Some of the pteridophytes have medicinal value in Chinese medicine and the same plants were tested using five chloroplast DNA barcode such as *psbA-trnH*, *rbcL*, *rpoB*, *rpoC1*, and *matK*) and found that *psbA-trnH* intergenic region was best candidate marker for pteridophytes authentication (25). Pteridophyte genus *Selaginella* is a non-seed bearing plant which was effectively resolved using ITS2 barcode (26). *Adiantum* L. genus was discriminated using morphological characteristic and six plastid markers such as *atpA*, *atpB*, *rbcL*, *trnL-F*, *rps4-trnS* and *matK* (27).

DNA barcoding of gymnosperm

Gymnosperms are seed bearing plants comprises an important four subclasses such as cycadidae, Gingoidae, Gnetidae and Pinidae, representing 12 families, 83 genera and about 990 species (28).

Table 1. List of DNA barcodes markers used in various plant division identification with the references cited.

Plant Division	DNA Barcode	References
Algae	COI, *rbcL*, *matK*, tufA, 23S, 18S rDNA, nuITS1 and nuITS2	Hall et al 2010; Buchheim et al 2011; Caisová et al 2011; Pawlowski et al 2012; Hadi et al 2016
Bryophytes	*rbcL*, *matK*, *rpoB*, *trnH-psbA*, *trnL-trnF*, *rps4-trnT*, *rps19-rpl2*, ITS, *atpF-atpH*, *psbK-psbI*, and *rpoC1*	Lang et al 2014; Hofbauer et al 2016
Pteridophytes	*rbcL*, *matK*, *trnH-psbA*, *trnL-trnF*, *rpoB*, *rpoC1*, *atpA*, *atpB*, *rps4-trnS*, and ITS2	Ebihara et al 2010; Ma et al 2010; Li et al 2011, Gu et al 2013; Wang et al 2017
Gymnosperm	*rbcL*, *matK*, *ndhJ*, *rpoB*, *accD*, *YCF5 and rpoC1*	Sass et al 2007; Li et al 2011
Angiosperm	*rbcL* , *matK*, *trnH-psbA*, ITS2, *trnL-trnF*, *rpoB*, *rpoC1*, *accD*, *YCF5* , *atpF-atpH*, *trnfM-trnT*, *trnD-psbM*, *petNtrnC*, *rps16*, *psaI*	CBOL 2009; Chen et al 2010; China Plant BOL Group, 2011; Saddhe et al 2016; Awad et al 2017; Saddhe et al 2017

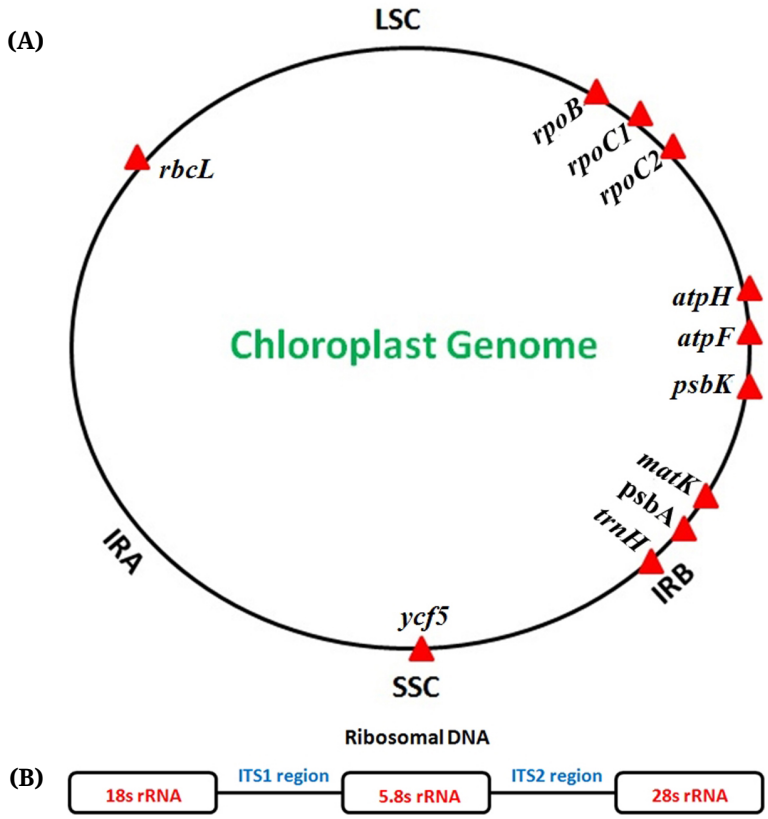

Fig. 1. Schematic representation of plastid (A) and nuclear (B) markers commonly used in plant DNA barcoding. Abbreviations used: LSC-large single copy region, SSC-small single-copy region, IR-large inverted repeat (IRA, IRB), *rbcL*-Ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit, matK- Maturase K, *rpoB* and *rpoC1* codes for chloroplast RNA polymerase subunit, *trnH-psb*- intergenic spacer, *atpF* and *atpH* encode ATP synthase subunits CFO I and CFO III respectively, *psbK* and *psbI* genes encode two polypeptides K and I, *ycf1* gene encodes Tic214 complex, ITS - Internal Transcribed Spacer.

Some gymnosperms are considered as 'living fossils' such as Cycads, *Ginkgo biloba*, *Metasequoia glyptostroboides* and *Glyptostrobus pensilis*. However, very few reports are available on gymnosperm DNA barcoding and assessment of potential DNA barcodes in this division. An ancient gymnosperm order Cycadales members were tested using universal DNA barcode markers such as *ndhJ*, *rpoB*, *matK*, *accD*, *YCF5* and *rpoC1* (29). Recently universality of 9 potential *matK* and 1 *rbcL* primers were assessed for barcoding gymnosperms (30).

DNA barcoding of angiosperm

Angiosperms are an economically important group of flowering plants including 416 families, about 13,164 genera and 295,383 known species (28). The efficacy of most of DNA barcode markers were evaluated using angiosperm plants as a case study. As CBOL recommended *rbcL* and *matK* as core barcode with few supporting markers such as ITS2, *trnH-psbA* was successfully implemented into angiosperm groups. Some inherent problems in plant taxa such as cryptic and closely related taxa, genotypic and phenotypic variability, and natural

hybridization which hide the success rate of DNA barcoding in some plant taxa (31). To overcome this issue, multiple and enormous DNA markers with different combinations were evaluated ranged from plastid coding (*rbcL*, *matK*) to non-coding regions (*trnH-psbA*), nuclear spacer (ITS) (31). The plastid and nuclear markers commonly used in plant DNA barcoding is shown (Fig. 1). The plastid marker *matK* can differentiate more than 90% of species in the Orchidaceae (Orchid family) but less than 49% in the Myristicaceae (nutmeg family) (32-33). The plastid markers such as *rbcL* and *matK* exhibited low resolution in species-rich genera and complex taxa such as *Lysimachia*, *Ficus*, *Holcoglossum*, and *Curcuma* (7-10). The lowest discriminatory power was observed in closely related groups of *Lysimachia* with *rbcL* (26.5-38.1%), followed by *matK* (55.9-60.8%) and combinations of core barcodes (*rbcL* + *matK*) had discrimination of 47.1-60.8% (10). Mangroves identification based on core DNA barcode exhibited *rbcL* 47.72%, *matK* locus assigned (72.09%), ITS2 (87.82%) and combinations of *matK* + ITS2 resolved (89.74%) species however *Avicennia* species required additional *atpF-atpH* marker (34-35). Identification of *Triticum* plants using chloroplast genome-wide analysis revealed combination of the intergenic region (*trnfM-trnT*) with either (*trnD-psbM*), cytochrome b6-f complex subunit 8 (*petN) with trnC*, (*matK-rps16*) or (*rbcL-psaI*) demonstrated a very high discrimination capacity (36).

Future Perspective

Besides the core DNA barcode *rbcL* and *matK*, plant barcoding needs some supplementary markers such as *trnH-psbA* and ITS. Moreover, in closely related and cryptic taxa DNA barcoding is always ambiguous and demands more group specific markers. However, DNA barcoding has significant impact on molecular phylogeny, population genetics, evolution and ecology, biosecurity and food product regulation. Recently developed tools such as metabarcoding coupled with high-throughput sequencing (HTS) are rapid, accurate, and cost-effective alternative to resolve cryptic taxa. Moreover, environmental DNA (eDNA) metabarcoding, which includes universal DNA barcodes and HTS to characterize biological communities from terrestrial and aquatic environmental samples can be effectively used.

Competing interests: The authors have declared that no competing interests exist.

Acknowledgements

Authors thank BITS Pilani K. K. Birla Goa Campus, for providing the necessary support. AAS is gratefully acknowledged to University Grant Commission (UGC), India for senior research fellowship.

References

1. Hebert PDN, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Biol Sci SerB. 2003; 270: S96-S99. https://doi.org/10.1098/rsbl.2003.0025
2. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc R Soc Biol Sci SerB. 2003; 270: 313-321. https://doi.org/10.1098/rspb.2002.2218
3. CBOL Plant Working Group. A DNA barcode for land plants. Proc Natl Acad Sci USA. 2009; 106: 12794-12797. https://doi.org/10.1073/pnas.0905845106
4. China Plant BOL Group, Li DZ, Gao LM, Li HT, Wang H, Ge XJ, et al. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci USA. 2011; 108: 19641-19646. https://doi.org/10.1073/pnas.1104551108
5. Chen S, Yao H, Han J, Liu C, Song J, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. Plos One. 2010; 5: e8613. https://doi.org/10.1371/journal.pone.0008613
6. Mohammed Abubakar B, MohdSalleh F, Shamsir Omar MS, Wagiran A. DNA Barcoding and Chromatography Fingerprints for the Authentication of Botanicals in Herbal Medicinal Products. Evid Based Complement Alternat Med. 2017.1352948.
7. Li HQ, Chen JY, Wang S, Xiong SZ. Evaluation of six candidate DNA barcoding loci in Ficus (Moraceae) of China. Mol Ecol Resour. 2012; 12: 783-790. https://doi.org/10.1111/j.1755-0998.2012.03147.x
8. Chen J, Zhao J, Erickson DL, Xia N, Kress WJ. Testing DNA barcodes in closely related species of *Curcuma* (Zingiberaceae) from Myanmar and China. Mol Ecol Resour. 2015; 15: 337-348. https://doi.org/10.1111/1755-0998.12319
9. Xiang XG, Hu H, Wang W, Jin XH. DNA barcoding of the recently evolved genus *Holcoglossum* (Orchidaceae: Aeridinae): a test of DNA barcode candidates. Mol Ecol Resour. 2011; 11: 1012-1021. https://doi.org/10.1111/j.1755-0998.2011.03044.x
10. Zhang CY, Wang FY, Yan HF, Hao G, Hu CM, et al. Testing DNA barcoding in closely related groups of *Lysimachia* L. (Myrsinaceae). Mol Ecol Resour. 2012; 12: 98-108. https://doi.org/10.1111/j.1755-0998.2011.03076.x
11. Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, Sundaresan V. DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant biotechnol J. 2016; 14: 8-21. https://doi.org/10.1111/pbi.12419
12. Janjua S, Fakhar-I-Abbas, William K, Malik IU, Mehr J. DNA Mini-barcoding for wildlife trade control: a case study on identification of highly processed animal materials. Mitochondrial DNA. 2017; 28: 544-6. https://doi.org/10.3109/24701394.2016.1155051
13. Manoylov KM. Taxonomic identification of algae (morphological and molecular): species concepts, methodologies, and their implications for ecological bioassessment. J phycol. 2014; 50: 409-24. https://doi.org/10.1111/jpy.12183
14. Subhadra B. Algal biorefinery-based industry: an approach to address fuel and food insecurity for a

carbon-smart world. J Sci Food Agric. 201; 91: 2-13. https://doi.org/10.1002/jsfa.4207

15. Hadi SIIA, Santana H, Brunale PPM, Gomes TG, Oliveira MD, Matthiensen A, et al. (2016) DNA barcoding green microalgae isolated from neotropical inland waters. Plos One 11(2): e0149284. https://doi.org/10.1371/journal.pone.0149284
16. Hall JD, Fucikova K, Lo C, Lewis LA, Karol KG. An assessment of proposed DNA barcodes in freshwater green algae. Cryptogam, Algol. 2010; 31(4):529-55.
17. Buchheim MA, Keller A, Koetschan C, Forster F, Merget B, Wolf M. Internal transcribed spacer 2 (nu ITS2 rRNA) sequence-structure phylogenetics: towards an automated reconstruction of the green algal tree of life. PLoS One. 2011; 6(2):e16931. https://doi.org/10.1371/journal.pone.0016931
18. Caisová L, Marin B, Melkonian M. A close-up view on ITS2 evolution and speciation-a case study in the Ulvophyceae (Chlorophyta, Viridiplantae). BMC Evol Biol. 2011; 11: 262. https://doi.org/10.1186/1471-2148-11-262
19. Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, Decelle J, Dunthorn M, Fiore-Donno AM. CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol. 2012; 10: e1001419. https://doi.org/10.1371/journal.pbio.1001419
20. Lang AS, Kruijer JD, Stech M. DNA barcoding of Arctic bryophytes: an example from the moss genus *Dicranum* (Dicranaceae, Bryophyta). Polar Biol. 2014; 37: 1157-69. https://doi.org/10.1007/s00300-014-1509-7
21. Hofbauer WK, Forrest LL, Hollingsworth PM, Hart ML. Preliminary insights from DNA barcoding into the diversity of mosses colonising modern building surfaces. Bryophyte Diversity Evol. 2016; 38:1-22. https://doi.org/10.11646/bde.38.1.1
22. Liu Y, YAN HF, Cao T, GE XJ. Evaluation of 10 plant barcodes in bryophyta (Mosses). J Syst Evol. 2010; 48: 36-46. https://doi.org/10.1111/j.1759-6831.2009.00063.x
23. Ebihara A, Nitta JH, Ito M. Molecular species identification with rich floristic sampling: DNA barcoding the pteridophyte flora of Japan. PLoS One 2010; 5: e15136. https://doi.org/10.1371/journal.pone.0015136
24. Li F-W, Kuo L-Y, Rothfels CJ, Ebihara A, Chiou W-L, Windham MD, et al. *rbcL* and *matK* earn two thumbs up as the core DNA barcode for Ferns. PLoS One. 2011; 6: e26597. https://doi.org/10.1371/journal.pone.0026597
25. Ma XY, Xie CX, Liu C, Song JY, Yao H, Luo K, Zhu YJ, Gao T, Pang XH, Qian J, Chen SL. Species identification of medicinal pteridophytes by a DNA barcode marker, the chloroplast psbA-trnH intergenic region. Biol Pharm Bull. 2010; 33: 1919-24. https://doi.org/10.1248/bpb.33.1919
26. Gu W, Song J, Cao Y, Sun Q, Yao H, Wu Q, Chao J, Zhou J, Xue W, Duan J. Application of the ITS2 region for barcoding medicinal plants of Selaginellaceae in Pteridophyta. PloS one. 2013; 8(6): e67818. https://doi.org/10.1371/journal.pone.0067818
27. Wang AH, Wang FG, Zhang WW, Ma XD, Li XW, Yi QF, Li DL, Duan L, Yan YH, Xing FW. Revision of series Gravesiana (*Adiantum* L.) based on morphological characteristics, spores and phylogenetic analyses. PloS one. 2017; 12(4): e0172729. https://doi.org/10.1371/journal.pone.0172729
28. Christenhusz MJ, Byng JW. The number of known plants species in the world and its annual increase. Phytotaxa. 2016; 261:201-17. https://doi.org/10.11646/phytotaxa.261.3.1
29. Sass C, Little DP, Stevenson DW, Specht CD. DNA barcoding in the Cycadales: testing the potential of proposed barcoding markers for species identification of Cycads. PloS One 2007; 2: e1154. https://doi.org/10.1371/journal.pone.0001154
30. Li Y, Gao LM, Poudel RC, Li DZ, Forrest A. High universality of matK primers for barcoding gymnosperms. J Syst Evol. 2011; 49: 169-75. https://doi.org/10.1111/j.1759-6831.2011.00128.x
31. Vijayan K, Tsou CH. DNA barcoding in plants: taxonomy in a new perspective. Curr. Sci. 2010:1530-41.
32. Kress J, 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: e508. https://doi.org/10.1371/journal.pone.0000508
33. Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J. Testing candidate plant barcode regions in the Myristicaceae. Mol Ecol Resour. 2008; 8: 480-490. https://doi.org/10.1111/j.1471-8286.2007.02002.x
34. Saddhe AA, Jamdade RA, Kumar K. Assessment of mangroves from Goa, west coast India using DNA barcode. SpringerPlus. 2016; 5:1554. https://doi.org/10.1186/s40064-016-3191-4
35. Saddhe AA, Jamdade RA, Kumar K. Evaluation of multilocus marker efficacy for delineating mangrove species of West Coast India. PLoS One. 2017; 12: e0183245. https://doi.org/10.1371/journal.pone.0183245
36. Awad M, Fahmy RM, Mosa KA, Helmy M, El-Feky FA. Identification of effective DNA barcodes for Triticum plants through chloroplast genome-wide analysis. Comput Biol Chem. 2017; 71: 20-31. https://doi.org/10.1016/j.compbiolchem.2017.09.003