bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Identification and validation of ITS2-specific universal primers for DNA barcoding in plants.

Ajay R Bhat[1], Subrahmanya Hegde[1], Supriya S Kammar[1], MB Muthamma[1], Gaurav Mudgal[2], Charukesi R[1]*, Thotegowdanapalya C Mohan[1]*.

[1]*Department of Biotechnology & Bioinformatics, JSS Academy of Higher Education & Research, Mysuru, Karnataka, India. 570015.*

[2]*University Institute of Biotechnology, Chandigarh University, Gharuan, Mohali, Punjab 140413, India.*

*Corresponding authors: mohantc@jssuni.edu.in & charukesi@jssuni.edu.in

Abstract

Accurate identification of plant species is fundamental for various purposes, including biodiversity conservation and effective utilization of plants. Molecular-based techniques, particularly DNA barcoding, have emerged as a rapid and accurate alternative to traditional taxonomic methods for plant species identification. DNA barcoding involves analysing specific DNA regions to classify and identify organisms. Among different barcode markers, internal transcribed spacer 2 (ITS2) has gained prominence due to its shorter length and ease of sequencing. In this study we present a comprehensive survey of the literature, identifying a universal set of primers for ITS2 and validating their universality through in silico and in vitro analyses. This manuscript provides valuable insights into establishing a standardized approach for DNA barcoding using ITS2, helping researchers select appropriate primers for accurate plant species identification and classification. Standardized universal primers alleviate the time-consuming primer selection process, promoting efficient and reliable DNA barcoding practices within the plant scientific community.

Keywords – ITS2, DNA barcoding, DNA markers, Species characterization

Introduction

Identification of plant species is essential for their usage, and conservation of plant biodiversity. In the case of closely related species, the conventional method of identification using taxonomic evidence is challenging. Therefore, for rapid and accurate identification of plant species, the usage of molecular-based techniques is becoming increasingly evident (de Boer *et al.*, 2022; Hebert *et al.*, 2003). DNA barcoding is one such molecular technique used to identify and classify organisms based on the analysis of a short specific DNA region. This approach has revolutionized the field of taxonomy, as it provides a quick and reliable method for species identification. Plant chloroplast DNA markers, such as matK, rbcL, and trnH-psbA, have been widely used as DNA barcodes for plants (Kress, 2017). In addition, recently many researchers have shown that the internal transcribed spacer 2 (ITS2) region of nuclear ribosomal DNA is highly effective and reliable for the differentiation of plants at species level which shows the highest discriminating power (97% and 74% at the genus and species level respectively)(Yao et al., 2010).

In 2011 China Plant BOL group did a comparative analysis of a large dataset from diverse plant species and recommended an internal transcribed spacer (ITS) (Figure I) as a core barcode for seed plants (D. Z. Li et al., 2011;X. Li et al., 2015). The internal transcribed spacer is as a spacer region situated between the genes encoding the smaller and larger subunit of rRNA. Specifically, ITS1 is positioned between the 18s rRNA and 5.8s rRNA genes, consisting of an average length of 600 base pairs. On the other hand, ITS2 is located between the 5.8s rRNA and 26s rRNA genes, with an average length of 400 base pairs (Han et al., 2013; Tripathi et al., 2013). In the past decade, between ITS1 and ITS2, the ITS2 region of the internal transcribed spacer has been widely used DNA barcode for plant species identification due to its shorter length and and the ease of sequencing (Wang et al., 2018). The ITS2 region provides several advantages over the whole ITS or ITS1 region. It exhibits greater interspecific variation, has a shorter length, and provides a higher success rate in both PCR amplification and sequencing (Gu et al., 2013)(Han et al., 2013). Additionally, another drawback associated with using the complete ITS or ITS1 is the risk of fungal contamination during PCR amplification. Due to the substantial similarity between the ITS regions of fungi and plants, there is an increased chance of amplifying fungal ITS.). Therefore, in the recent decade, ITS2 has been accepted as the ideal DNA barcode by the plant scientist community and largely used to identify and classify a wide range of plants from weeds to medicinal plants (Khan et al., 2019; Zhao et al., 2018; Zuo et al., 2011).

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure I

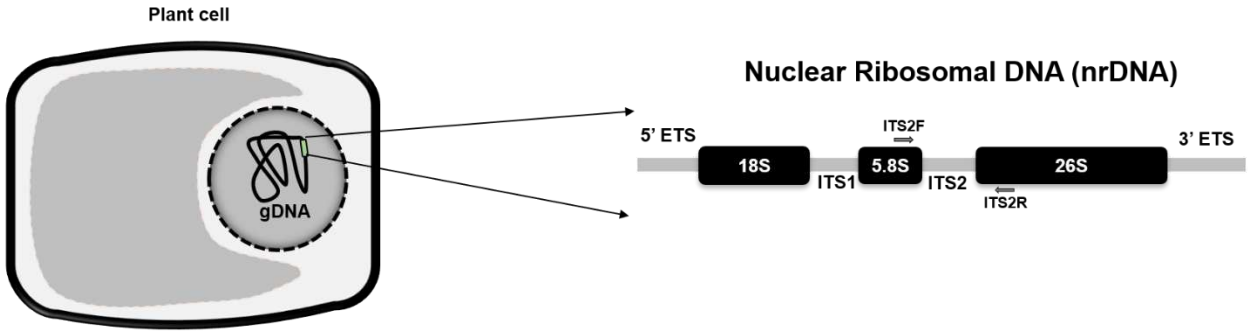


Fig. I Schematic representation of ITS region of the nuclear ribosomal RNA (nrRNA) genes.
gDNA – genomic DNA, ETS- External Transcribed spacer, ITS1 and 2 – Internal Transcribed spacer, ITS2F and ITS2R – Forwrard and Reverse primers to amplify ITS2 region

For an ideal DNA barcode, the availability of a universal primer is important for the ease of end users. Before and after the ITS2 was accepted as an ideal barcode several researchers used different ITS2 primers for the identification of a diverse range of plants. Currently, in the literature, there are more than 20 sets of primers being mentioned for ITS2 (Table II). For a novice, it is very time-consuming and creates ambiguity in selecting the correct primer for their experiment. Therefore, in this manuscript, we did a comprehensive literature survey and identified a set of universal primers and validated them using in silico and in vitro analysis.

Results and Discussion

Comprehensive analysis of ITS2 primers.

The development of an optimal DNA barcode is of utmost importance for its widespread applicability across a range of plant species. An ideal barcode should contain significant species-level genetic variability. It should exhibit good PCR amplification efficiency by using a universal primer, enabling the use of a single set of primers for barcoding across different species. Ideally, the chosen universal primers should amplify a short sequence of 600-700 bp for easy sequencing and Insilico analysis (Genes et al., 2008). In the recent past, the ITS2 marker has been specifically used as a core DNA barcode for plants. Nonetheless, to our knowledge, there is no such universal primer being adopted by the researchers for ITS2 marker.

To identify and suggest a universal set of primers for ITS2 marker we did an extensive survey of the literature published during the period of 1990 to 2023. We have referred more than 40 published research papers and found that there are 34 different sets of primers being used for the ITS2 marker (Table II). Out of 42 researchers, 18 researchers have used the primers set number 1 (Table II). The rest of the researchers have used different combinations of forward and reverse primers for ITS2 as shown in Table II. Our analysis indicates that novice researchers may find it challenging to identify and select an appropriate primer set for their experiments Therefore, to find a universal set of primer we have performed insilico analysis and invitro validation of the available set of primers.

In silico analysis of ITS2-specific universal primers.

Our literature analysis found that 18 researchers have used the ITS2 primer set number 1 (Table II) to perform barcoding in various plant species. Therefore, to analyse the universality of these primers, and to identify an ideal single set of universal primers for ITS2, we have compared the suitability of both reverse and forward primers across different species of plants. Utilising the ITS2 database (Merget et al., 2012) and the NCBI database we have obtained rRNA sequences of representative plants from the top 22 major plant families, accounting for over 150,000 accepted species of plants, including 21 angiosperm (flowering plant) families and 1 gymnosperm (non-flowering seed plant) family (Table

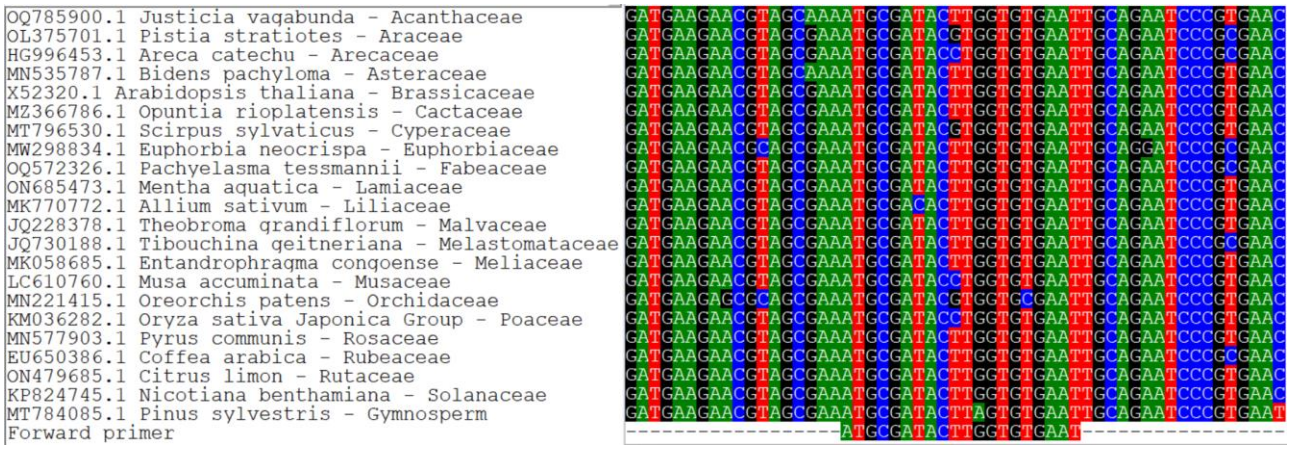
bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

I). These plant families represent a significant portion of the world's plant diversity (Christenhusz & Byng, 2016). The multiple sequence alignment of primer set number-1 against these rRNA sequences of various plant species using Bioedit tool (Hall et al., 2011) showed that both the forward (Figure IIa) and reverse (Figure IIb) primers have showed >90% and >95% sequence similarity respectively to the rest of the sequences. Moreover, the 3' end of the primer set-1 shows no dissimilarity which is important for stringent primer binding during the PCR. Our in-silico analysis shows that, the primer set-1 could be used as a universal primer for ITS2 marker to perform barcoding of plant species.

Table I – List of families used for validation by multiple sequence alignment.

Sl No.	Family
1	Acanthaceae
2	Araceae
3	Arecaceae
4	Asteraceae
5	Brassicaceae
6	Cactaceae
7	Cyperaceae
8	Euphorbiaceae
9	Fabaceae
10	Lamiaceae
11	Liliaceae
12	Malvaceae
13	Melastomataceae
14	Meliaceae
15	Musaceae
16	Orchidaceae
17	Poaceae
18	Rosaceae
19	Rubeaceae
20	Rutaceae
21	Solanaceae
22	Pinaceae

Figure IIa



bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure IIb

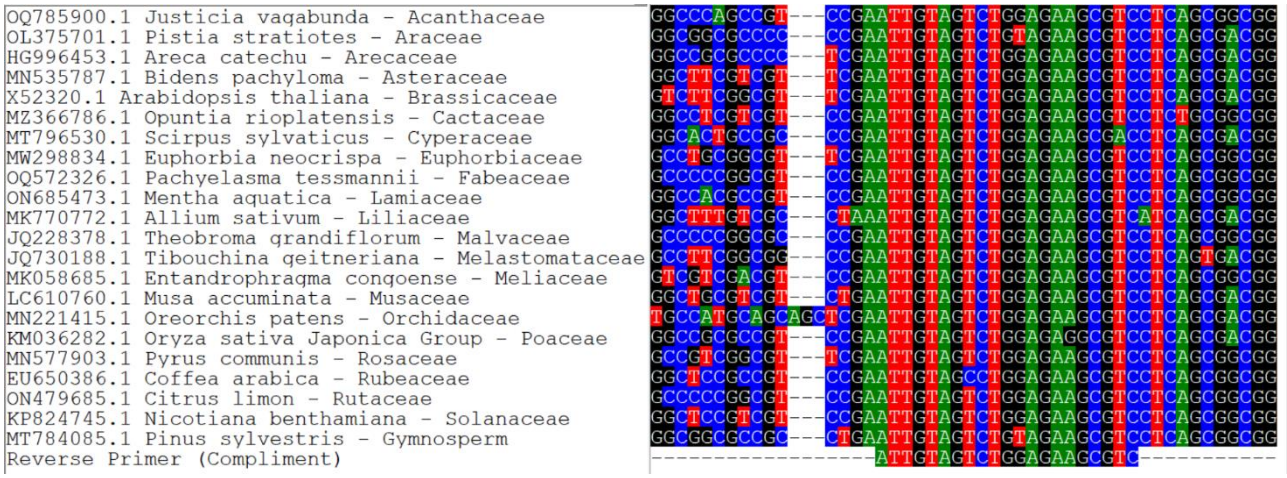


Fig. II (a) & (b) Multiple alignment of ITS2 specific forward primer (a) and reverse primer (b) with the rRNA sequences of 20 different family plant species

In vitro validation of ITS2 specific universal primers using PCR.

To experimentally prove the universality of the ITS2 specific universal primers, we have performed polymerase chain reaction (PCR) using primer set 1 for rice, banana and almond plants. In all the 3 plant species we have observed good amplification at approximately 500 bp of ITS2 region (Figure III). Our results confirm that researchers can utilize the primer set-1 as a universal primer to perform DNA barcoding in diverse plant species.

Figure III

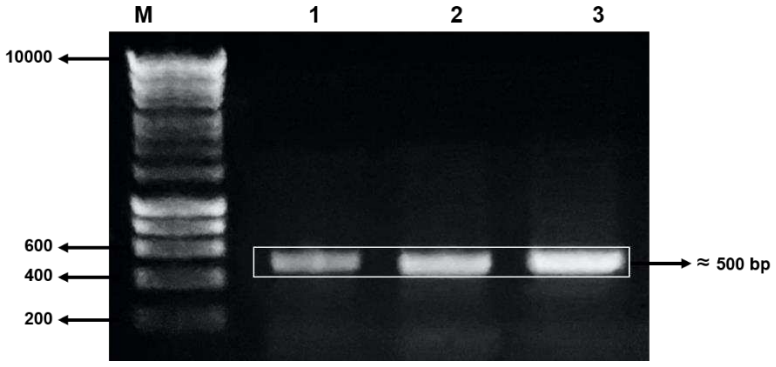


Fig. III PCR amplification of ITS 2 region using primer set-1. Total DNA extracted from Rice, Almond and banana plant were used as template. Amplicon length is ≈ 500 bp
M – 10 kb Marker, Lane 1-Rice, Lane 2-Almond and Lane 3-Banana.

Methods and Materials.

Finding ITS2 primers, Retrieval of ITS sequences, and in silico validation.
All the ITS2 specific primers were obtained from the available literature in Pubmed. The primers mentioned in the literature was collected and summarized as shown in Table II. The ITS2 sequences for the different plant species of 20 different families were downloaded in the FASTA format from the ITS2 database (http://its2.bioapps.biozentrum.uni-wuerzburg.de) and NCBI (National Centre for Biotechnology Information) database. The respective ITS2 specific complete rRNA sequences were obtained from GenBank. Multiple alignments of retrieved rRNA sequence and the primers were done using BIOEDIT tool.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table II – List of primers used by different researchers to amplify ITS2.

Primer set number.	Primer Sequence	Reference
1.	ITS2 F - ATGCGATACTTGGTGTGAAT ITS2 R - GACGCTTCTCCAGACTACAAT	(Mohanty et al., 2023) (Durán Escalante et al., 2023)(Song et al., 2023)(Cahyaningsih et al., 2022) (Qarni et al., 2022) (J. Chen et al., 2021) (Vasconcelos et al., 2021) (Frigerio et al., 2021) (LV et al., 2020)(Zhao et al., 2018) (Gong et al., 2018) (N. Yu et al., 2017) (Feng et al., 2016) (X. Yu et al., 2016)(Gu et al., 2013) (Tripathi et al., 2013) (Pang et al., 2012) (S. Chen et al., 2010)
2.	ITS2 F - GGGGCGGATATTGGCCTCCCCTTG ITS2 R - GACGCTTCTCCAGACTACAAT	(Saidon et al., 2023)
3.	ITS2 F - ATGCGATACTTGGTGTGAAT ITS2 R – TCCTCCGCTTATTGATATGC	(Rakotonirina et al., 2023)
4.	ITS2 F - ATGCGATACTTGGTGTGAAT ITS2 R - GACGCTTCTCCAGACTACAT	(Nasarodin et al., 2023)
5.	ITS2 F - ATGCGATACTTGGTGTGAATTATAGAAT ITS2 R - GACGCTTCTCCAGACTACAAT	(Hegde et al., 2022)
6.	ITS2 F - ATGCGATACTTGGTGTGAATTATAGAAT ITS2 R - GACGCTTCTCCAGACTACAAT	(Acharya et al., 2022)
7.	ITS2 F - AATGATACGGCGACCACCGAGATCTACAC ITS2 R - ACACTCTTTCCCTACACGACGCTCTTCCGATCT	(Fujii et al., 2022)
8.	ITS2 F- AACCTGCGGAAGGATCATTGTC ITS2 R- TGATATGCTTAAACTCAGCGGGTA	(Zhang et al., 2022)
9.	ITS2 F - CCCGTGAACCATCGAGTCTTT ITS2 R - GACGGCTCGCCTCTCAAC	(Xavier et al., 2021)
10.	ITS2 F- ATGCGATACTTGGTGTGAAT ITS2 R- TCCTCCGCTTATTGATATGC	(Dhivya et al., 2020)
11.	ITS2 F - GCGATACTTGGTGTGAAT ITS2 R – GACGCTTCCCAGACTAAAT	(Khan et al., 2019)
12.	ITS2 F - GCATCGATGAAGAACGCAGC ITS2 R - TCCTCCGCTTATTGATATGC	(Duan et al., 2019)
13.	ITS2 F - AGGAGAAGTCGTAACAAGGT ITS2 R - TCCTCCGCTTATTGATATGC	(Saifuldeen Ahmed Hasan, 2018)
14.	ITS2 F - GGAAGTAAAAGTCGTAACAAGG ITS2 R - TCCTCCGCTTATTGATATGC	(Zhu et al., 2018)
15.	ITS2 F - AGGAGAAGTCGTAACAAG ITS2 R - GTTTCTTTTCCTCCGCT	(Inglis et al., 2018)
16.	ITS2 F- GGAAGTAAAAGTCGTAACAAGG ITS2 R - TCCTCCGCTTATTGATATGC	(Liang et al., 2018)
17.	ITS2 F - ATGCGATACTTGGTGTGAATTATAGAAT ITS2 R - GACGCTTCTCCAGACTACAAT	(Q. Li et al., 2017)
18.	ITS2 F - ATGCGATACTTGGTGTGAAT	(Kim et al., 2016)

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	ITS2 R -TCCTCCGCTTATTGATATGC	
19.	ITS2 F - CGTAGCGAAATGCGATACTTGGTG ITS2 R - TCCTCCGCTTATTGATATGC	(Chao et al., 2014)
20.	ITS2 F - TCTCGCATCGATGAAGAACG ITS2 R - CCATGCTTAAACTCAGCGGGT	(He et al., 2014)
21.	ITS2 F - CCTTATCATTTAGAGGAAGGA ITS2 R- TCCTCCGCTTATTGATATGC	(Tripathi et al., 2013)
22.	ITS2 F - ATGCGATACTTGGTGTGAAT ITS2 R - GACGCTTCTCCAGACTACAAT	(Gu et al., 2013)
23.	ITS2 F - GTAGGTGAACCTGCAGAAGGATCA ITS2 R -CCATGCTTAAACTCAGCGGGT	(Sukrong et al., 2007)
24.	ITS2 F - GCATCGATGAAGAACGCAGC ITS2 R - TCCTCCGCTTATTGATATGC	(White et al., 1990)

Confirmation of universality of ITS2 specific primes using PCR

Genomic DNA isolation and quantification. Total genomic DNA (gDNA) was isolated from the fresh juvenile leaf tissues of the 3 plant species (Rice, Banana and Almond) using gDNA extraction kit (Qiagen DNeasy® Plant Mini Kit) by following manufacturer's protocol. The isolated gDNAs were quantified using a nanodrop spectrophotometer (Denovix DS-11) and quality assessed via agarose gel electrophoresis (Agarose concentration 0.8%) (Bio-Rad Laboratories, India). Total gDNA concentration adjusted to 100 ng µL–1 was used for PCR amplification with primer Set 1(Table II) ITS2 primers.

Primer selection and PCR amplification. DNA barcode primers for the ITS2 were synthesized at Sigma Aldrich Chemicals Pvt. Ltd., Bengaluru, India. The details of the ITS2 primer sequences (5' to 3') are, ITS2F - ATGCGATACTTGGTGTGAAT forward primer and ITS2R - GACGCTTCTCCAGACTACAAT reverse primer, PCR amplification was performed in a volume of 20µL, containing 100 ng of gDNA (1 µL) as a template, 10 µL 2 × PCR master mix (EmeraldAmp® GT PCR Master Mix, Takara Bio Inc.), primers (0.2µM), 0.5 µL each of forward and reverse primers), and 8 µL Milli–Q water. All PCR amplifications were performed in the thermal cycler (Bio-Rad Laboratories, India) following denaturation of 3 min at 95°C, 35 cycles of 30 sec at 95°C, 30 sec at 56°C of annealing, 40 sec at 72°C and a final extension of 10 min at 72°C. The PCR products were visualized in 1% agarose TAE gels, and the gel images were taken.

Conclusion:

The conventional taxonomic methods face considerable challenges, especially in differentiating closely related species. The advent of molecular techniques, notably DNA barcoding, has significantly advanced the field of taxonomy by providing a reliable, efficient, and quick method for species identification. The ITS2 region, identified as an ideal DNA barcode due to its shorter length and ease of sequencing, has gained widespread acceptance within the scientific community. One of the critical challenges in utilizing ITS2 for DNA barcoding has been the selection of appropriate universal primers. Major drawback in using ITS as a barcode is the occurrence of polymorphic ITS regions and pseudogenization is infrequently observed in few species (Add reference). The extent to which these issues might significantly affect a substantial portion of barcoding studies remains uncertain (Bailey et al., 2003)(Hollingsworth et al., 2011). In instances where this uncommon problem arises, Next Generation Sequencing (NGS) emerges as the viable solution (Halvarsson & Tydén, 2023). With an extensive literature review, our study has not only identified a reliable set of universal primers but has also validated their universality through in silico and in vitro analyses. Our results demonstrate the practicality and efficacy of a specific primer set (primer set-1) (Table I) for DNA barcoding across a diverse range of plant species. By providing a validated universal primer set, our research contributes to facilitating researchers, particularly novices, in streamlining their DNA barcoding experiments and contributing to plant biodiversity studies, conservation efforts, and broader ecological research. Our results contribute towards enhancing our understanding of plant diversity and ultimately support global biodiversity conservation initiatives.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgments:

This work is supported by K-FIST (L1) Vision Group on Science and Technology, Govt of Karnataka, SERB-Startup Research Grant, Science and Engineering Research Board, Department of Science and Technology, Government of India, Global Challenges Research Fund (GCRF), United Kingdom, in collaboration with the University of Warwick, United Kingdom and JSS Academy of Higher Education and Research, Mysuru. We thank the Centre of Excellence in Molecular Biology and Regenerative Medicine (CEMR) headed by Dr Subbarao, JSS Medical College for providing extended laboratory support.

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References –

Acharya, G. C., Mohanty, S., Dasgupta, M., Sahu, S., Singh, S., Koundinya, A. V. V., Kumari, M., Naresh, P., & Sahoo, M. R. (2022). Molecular Phylogeny, DNA Barcoding, and ITS2 Secondary Structure Predictions in the Medicinally Important Eryngium Genotypes of East Coast Region of India. *Genes*, *13*(9). https://doi.org/10.3390/genes13091678

Bailey, C. D., Carr, T. G., Harris, S. A., & Hughes, C. E. (2003). Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Molecular Phylogenetics and Evolution*, *29*(3), 435–455. https://doi.org/10.1016/j.ympev.2003.08.021

Cahyaningsih, R., Compton, L. J., Rahayu, S., Brehm, J. M., & Maxted, N. (2022). DNA Barcoding Medicinal Plant Species from Indonesia. *Plants*, *11*(10), 1–22. https://doi.org/10.3390/plants11101375

Chao, Z., Zeng, W., Liao, J., Liu, L., Liang, Z., & Li, X. (2014). DNA barcoding Chinese medicinal Bupleurum. *Phytomedicine*, *21*(13), 1767–1773. https://doi.org/10.1016/j.phymed.2014.09.001

Chen, J., Li, S., Wu, W., Xie, J., Cheng, X., Ye, Z., Yin, X., Liu, Y., & Huang, Z. (2021). Molecular Identification and Phylogenetic Analysis of the Traditional Chinese Medicinal Plant Kochia scoparia Using ITS2 Barcoding. *Interdisciplinary Sciences – Computational Life Sciences*, *13*(1), 128–139. https://doi.org/10.1007/s12539-021-00421-y

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., & Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*, *5*(1), 1–8. https://doi.org/10.1371/journal.pone.0008613

Christenhusz, M. J. M., & Byng, J. W. (2016). Phytotaxa. *Phytotaxa*, *261*(3), 201–217. https://doi.org/10.11646/phytotaxa.261.3.1

de Boer, H., Rydmark, M. O., Verstraete, B., & Gravendeel, B. (2022). Molecular identification of plants: from sequence to species. In *Molecular identification of plants: from sequence to species*. https://doi.org/10.3897/ab.e98875

Dhivya, S., Ashutosh, S., Gowtham, I., Baskar, V., Harini, A. B., Mukunthakumar, S., & Sathishkumar, R. (2020). Molecular identification and evolutionary relationships between the subspecies of Musa by DNA barcodes. *BMC Genomics*, *21*(1), 659. https://doi.org/10.1186/s12864-020-07036-5

Duan, H., Wang, W., Zeng, Y., Guo, M., & Zhou, Y. (2019). The screening and identification of DNA barcode sequences for Rehmannia. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-53752-8

Durán Escalante, K. C., Ortiz Díaz, J. J., Pinzón Esquivel, J. P., & Gálvez Mariscal, M. A. (2023). Utilidad de los códigos de barras de DNA en la identificación de plantas melíferas asociadas a la miel monofloral de Sabal yapa producida en el este de Yucatán, México. *Innotec*, *26*, 1–20. https://doi.org/10.26461/26.01

Feng, S., Jiang, M., Shi, Y., Jiao, K., Shen, C., Lu, J., Ying, Q., & Wang, H. (2016). Application of the ribosomal DNA ITS2 region of physalis (Solanaceae): DNA barcoding and phylogenetic study. *Frontiers in Plant Science*, *7*(2016JULY), 1–11. https://doi.org/10.3389/fpls.2016.01047

Frigerio, J., Agostinetto, G., Mezzasalma, V., De Mattia, F., Labra, M., & Bruno, A. (2021). Dna-based herbal teas' authentication: An its2 and psba-trnh multi-marker dna metabarcoding approach. *Plants*, *10*(10), 1–14. https://doi.org/10.3390/plants10102120

Fujii, T., Ueno, K., Shirako, T., Nakamura, M., & Minami, M. (2022). Identification of Lagopus muta japonica food plant resources in the Northern Japan Alps using DNA metabarcoding. *PLoS ONE*, *17*(3 March), 1–24. https://doi.org/10.1371/journal.pone.0252632

Genes, D. N. A., Kress, W. J., & Erickson, D. L. (2008). *DNA barcodes: Genes, genomics, and bioinformatics*. *105*(8), 2761–2762.

Gong, L., Qiu, X. H., Huang, J., Xu, W., Bai, J. Q., Zhang, J., Su, H., Xu, C. M., & Huang, Z. H. (2018). Constructing a DNA barcode reference library for southern herbs in China: A resource for authentication of southern Chinese medicine. *PLoS ONE*, *13*(7), 1–12. https://doi.org/10.1371/journal.pone.0201240

Gu, W., Song, J., Cao, Y., Sun, Q., Yao, H., Wu, Q., Chao, J., Zhou, J., Xue, W., & Duan, J. (2013). Application of the ITS2 Region for Barcoding Medicinal Plants of Selaginellaceae in Pteridophyta. *PLoS ONE*, *8*(6), 2–9. https://doi.org/10.1371/journal.pone.0067818

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hall, T., Biosciences, I., & Carlsbad, C. (2011). BioEdit: An important software for molecular biology. *GERF Bulletin of Biosciences*, *2*(June), 60–61.

Halvarsson, P., & Tydén, E. (2023). The complete ITS2 barcoding region for Strongylus vulgaris and Strongylus edentatus. *Veterinary Research Communications*, *47*(3), 1767–1771. https://doi.org/10.1007/s11259-022-10067-w

Han, J., Zhu, Y., Chen, X., Liao, B., Yao, H., Song, J., Chen, S., & Meng, F. (2013). The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. *BioMed Research International*, *2013*, 3–10. https://doi.org/10.1155/2013/741476

He, J. Y., Zhu, S., Komatsu, K., Goda, Y., & Cai, S. Q. (2014). Genetic polymorphism of medicinally-used Codonopsis species in an internal transcribed spacer sequence of nuclear ribosomal DNA and its application to authenticate Codonopsis Radix. *Journal of Natural Medicines*, *68*(1), 112–124. https://doi.org/10.1007/s11418-013-0780-1

Hebert, P. D. N., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, *270*(SUPPL. 1), 96–99. https://doi.org/10.1098/rsbl.2003.0025

Hegde, V., Roy, S. S., Renadevan, R., Kumar, K. B., & Patel, H. K. (2022). *Hill Region of India*.

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). *Choosing and Using a Plant DNA Barcode*. *6*(5). https://doi.org/10.1371/journal.pone.0019254

Inglis, P. W., Mata, L. R., Da Silva, M. J., Vieira, R. F., Alves, R. D. B. N., Silva, D. B., & Azevedo, V. C. R. (2018). DNA Barcoding for the Identification of Phyllanthus Taxa Used Medicinally in Brazil. *Planta Medica*, *84*(17), 1300–1310. https://doi.org/10.1055/a-0644-2688

Khan, S. A., Baeshen, M. N., Ramadan, H. A., & Baeshen, N. A. (2019). ITS2: An Ideal DNA Barcode for the Arid Medicinal Plant Rhazya Stricta. *Pharmaceutical Medicine*, *33*(1), 53–61. https://doi.org/10.1007/s40290-019-00266-3

Kim, W. J., Moon, B. C., Yang, S., Han, K. S., Choi, G., & Lee, A. Y. (2016). Rapid authentication of the herbal medicine plant species Aralia continentalis Kitag. and Angelica biserrata C.Q. Yuan and R.H. Shan using ITS2 sequences and multiplex-SCAR markers. *Molecules*, *21*(3). https://doi.org/10.3390/molecules21030270

Kress, W. J. (2017). Plant DNA barcodes: Applications today and in the future. *Journal of Systematics and Evolution*, *55*(4), 291–307. https://doi.org/10.1111/jse.12254

Li, D. Z., Gao, L. M., Li, H. T., Wang, H., Ge, X. J., Liu, J. Q., Chen, Z. D., Zhou, S. L., Chen, S. L., Yang, J. B., Fu, C. X., Zeng, C. X., Yan, H. F., Zhu, Y. J., Sun, Y. S., Chen, S. Y., Zhao, L., Wang, K., Yang, T., & Duan, G. W. (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(49), 19641–19646. https://doi.org/10.1073/pnas.1104551108

Li, Q., Xu, J., Han, L., Gao, C., Lu, J., Du, G., & Sun, X. (2017). Evaluation of ITS2 for intraspecific identification of Paeonia lactiflora cultivars. *Biotechnology Reports*, *15*(May), 101–106. https://doi.org/10.1016/j.btre.2017.07.003

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., & Chen, S. (2015). Plant DNA barcoding: from gene to genome. *Biological Reviews of the Cambridge Philosophical Society*, *90*(1), 157–166. https://doi.org/10.1111/brv.12104

Liang, L. J., Wang, E. H., Yang, Y. C., Xing, B. C., Ji, W., Liu, F., & Liang, Z. S. (2018). Study on hybrid characteristics of medicinally used cultivated codonopsis species using ribosomal Internal Transcribed Spacer (ITS) sequencing. *Molecules*, *23*(7), 1–14. https://doi.org/10.3390/molecules23071565

LV, Y. N., YANG, C. Y., SHI, L. C., ZHANG, Z. L., XU, A. S., ZHANG, L. X., LI, X. L., & LI, H. T. (2020). Identification of medicinal plants within the Apocynaceae family using ITS2 and psbA-trnH barcodes. *Chinese Journal of Natural Medicines*, *18*(8), 594–605. https://doi.org/10.1016/S1875-5364(20)30071-6

Merget, B., Koetschan, C., Hackl, T., Förster, F., Dandekar, T., Müller, T., Schultz, J., & Wolf, M. (2012). The ITS2 Database. *Journal of Visualized Experiments : JoVE*, *61*, 1–5. https://doi.org/10.3791/3806

Mohanty, S., Mishra, B. K., Dasgupta, M., Acharya, G. C., Singh, S., Naresh, P., Bhue, S., Dixit, A., Sarkar, A., & Sahoo, M. R. (2023). Deciphering phenotyping, DNA barcoding, and RNA secondary structure predictions in eggplant wild relatives provide insights for their future breeding strategies. *Scientific Reports*, *13*(1), 1–12. https://doi.org/10.1038/s41598-023-40797-z

Nasarodin, N. S. M., Akbar, A., & Wagiran, A. (2023). ITS2 secondary structure data improves authentication of

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Moringa oleifera tea products when using with DNA barcoding. *Plant Science Today*, *x*(x), 1–11. https://doi.org/10.14719/pst.2094

Pang, X., Luo, H., & Sun, C. (2012). Assessing the potential of candidate DNA barcodes for identifying non-flowering seed plants. *Plant Biology*, *14*(5), 839–844. https://doi.org/10.1111/j.1438-8677.2011.00554.x

Qarni, A., Muhammad, K., Wahab, A., Ali, A., Khizar, C., Ullah, I., Kazmi, A., Sultana, T., Hameed, A., Younas, M., & Rahimi, M. (2022). Molecular Characterization of Wild and Cultivated Strawberry (Fragaria × ananassa) through DNA Barcode Markers. *Genetics Research*, *2022*. https://doi.org/10.1155/2022/9249561

Rakotonirina, T. J., Viljoen, E., Rakotonirina, A. H., Leong Pock Tsy, J. M., & Radanielina, T. (2023). A DNA barcode reference library for CITES listed Malagasy Dalbergia species . *Ecology and Evolution*, *13*(3), 1–11. https://doi.org/10.1002/ece3.9887

Saidon, N. A., Wagiran, A., Samad, A. F. A., Mohd Salleh, F., Mohamed, F., Jani, J., & Linatoc, A. C. (2023). DNA Barcoding, Phylogenetic Analysis and Secondary Structure Predictions of Nepenthes ampullaria, Nepenthes gracilis and Nepenthes rafflesiana. *Genes*, *14*(3), 697. https://doi.org/10.3390/genes14030697

Saifuldeen Ahmed Hasan, S. A. H. (2018). DNA Barcoding for Differentiating the 11 Varieties of Musa Species. *International Journal of Agricultural Science and Research*, *8*(4), 121–130. https://doi.org/10.24247/ijasraug201816

Song, F., Deng, Y. F., Yan, H. F., Lin, Z. L., Delgado, A., Trinidad, H., Gonzales-Arce, P., Riva, S., Cano-Echevarría, A., Ramos, E., Aroni, Y. P., Rivera, S., Arakaki, M., & Ge, X. J. (2023). Flora diversity survey and establishment of a plant DNA barcode database of Lomas ecosystems in Peru. *Scientific Data*, *10*(1), 1–8. https://doi.org/10.1038/s41597-023-02206-y

Sukrong, S., Zhu, S., Ruangrungsi, N., Phadungcharoen, T., Palanuvej, C., & Komatsu, K. (2007). Molecular analysis of the genus Mitragyna existing in Thailand based on rDNA its sequences and its application to identify a narcotic species: Mitragyna speciosa. *Biological and Pharmaceutical Bulletin*, *30*(7), 1284–1288. https://doi.org/10.1248/bpb.30.1284

Tripathi, A. M., Tyagi, A., Kumar, A., Singh, A., Singh, S., Chaudhary, L. B., & Roy, S. (2013). The Internal Transcribed Spacer (ITS) Region and trnhH-psbA Are Suitable Candidate Loci for DNA Barcoding of Tropical Tree Species of India. *PLoS ONE*, *8*(2). https://doi.org/10.1371/journal.pone.0057934

Vasconcelos, S., Nunes, G. L., Dias, M. C., Lorena, J., Oliveira, R. R. M., Lima, T. G. L., Pires, E. S., Valadares, R. B. S., Alves, R., Watanabe, M. T. C., Zappi, D. C., Hiura, A. L., Pastore, M., Vasconcelos, L. V., Mota, N. F. O., Viana, P. L., Gil, A. S. B., Simões, A. O., Imperatriz-Fonseca, V. L., … Oliveira, G. (2021). Unraveling the plant diversity of the Amazonian canga through DNA barcoding. *Ecology and Evolution*, *11*(19), 13348–13362. https://doi.org/10.1002/ece3.8057

White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and Direct Sequencing of Fungal Ribosomal Rna Genes for Phylogenetics. In *PCR Protocols* (Issue 1). Academic Press, Inc. https://doi.org/10.1016/b978-0-12-372180-8.50042-1

Xavier, J. K. A. M., Maia, L., Figueiredo, P. L. B., Folador, A., Ramos, A. R., Andrade, E. H., Maia, J. G. S., Setzer, W. N., & da Silva, J. K. R. (2021). Essential oil composition and DNA barcode and identification of Aniba species (Lauraceae) growing in the Amazon Region. *Molecules*, *26*(7), 1–19. https://doi.org/10.3390/molecules26071914

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., Pang, X., Xu, H., Zhu, Y., Xiao, P., & Chen, S. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS ONE*, *5*(10). https://doi.org/10.1371/journal.pone.0013102

Yu, N., Wei, Y. L., Zhang, X., Zhu, N., Wang, Y. L., Zhu, Y., Zhang, H. P., Li, F. M., Yang, L., Sun, J. Q., & Sun, A. D. (2017). Barcode ITS2: A useful tool for identifying Trachelospermum jasminoides and a good monitor for medicine market. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-04674-w

Yu, X., Xie, Z., Wu, J., Tao, J., & Xu, X. (2016). DNA barcoding identification of kadsurae caulis and spatholobi caulis based on internal transcribed spacer 2 region and secondary structure prediction. *Pharmacognosy Magazine*, *12*(46), 165. https://doi.org/10.4103/0973-1296.182162

Zhang, J., Chi, X., Zhong, J., Fernie, A., Alseekh, S., Huang, L., & Qian, D. (2022). Extensive nrDNA ITS polymorphism in Lycium: Non-concerted evolution and the identification of pseudogenes. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.984579

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.05.574284; this version posted January 5, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Zhao, L. L., Feng, S. J., Tian, J. Y., Wei, A. Z., & Yang, T. X. (2018). Internal transcribed spacer 2 (ITS2) barcodes: A useful tool for identifying Chinese Zanthoxylum. *Applications in Plant Sciences*, *6*(6), 1–8. https://doi.org/10.1002/aps3.1157

Zhu, S., Li, Q., Chen, S., Wang, Y., Zhou, L., Zeng, C., & Dong, J. (2018). Phylogenetic analysis of uncaria species based on internal transcribed spacer (Its) region and its2 secondary structure. *Pharmaceutical Biology*, *56*(1), 548–558. https://doi.org/10.1080/13880209.2018.1499780

Zuo, Y., Chen, Z., Kondo, K., Funamoto, T., Wen, J., & Zhou, S. (2011). DNA barcoding of panax species. *Planta Medica*, *77*(2), 182–187. https://doi.org/10.1055/s-0030-1250166