Multi-locus DNA barcoding identifies matK as a suitable marker for species identification in Hibiscus L.

Journal:	<i>Genome</i>
Manuscript ID	gen-2015-0205.R2
Manuscript Type:	Note
Date Submitted by the Author:	23-May-2016
Complete List of Authors:	Sundar, Poovitha; SRM University Stalin, Nithaniyal; SRM University, Genetic engineering Raju, Balaji; SRM University, Genetic Engineering Madasamy, Parani; SRM University, Genetic Engineering
Keyword:	Hibiscus, barcoding, matK, ITS2, divergence

**Multi-locus DNA barcoding identifies *matK* as suitable marker for species identification
in *Hibiscus* L.**

Sundar Poovitha, Nithaniyal Stalin, Raju Balaji, Madasamy Parani*

Centre for DNA Barcoding, Department of Genetic Engineering, SRM University,
Kattankulathur, Tamil Nadu, India.

* Corresponding author

Madasamy Parani, Centre for DNA Barcoding, Department of Genetic Engineering, SRM
University, Kattankulathur, Chennai 603203, Tamil Nadu, India.

Tel.: 091-44-2741 7817; Fax: 091-44-2745 3622

E-mail address: parani.m@ktr.srmuniv.ac.in

Draft

Abstract

The genus *Hibiscus* L. includes several taxa of medicinal value and species used for the extraction of natural dyes. These applications require the use of authentic plant materials. DNA barcoding is a molecular method for species identification, which helps in reliable authentication by using one or more DNA barcode marker. In this study, we have collected 44 accessions, representing 16 species of *Hibiscus*, distributed in the southern peninsular India, to evaluate the discriminatory power of the two core barcodes, *rbcLa* and *matK* together with the suggested additional regions, *trnH-psbA* and ITS2. No intra-species divergence was observed among the accessions studied. Inter-species divergence was 0-9.6% with individual markers, which increased to 0-12.5% and 0.8-20.3% when using two-, and three-marker combination, respectively. Differentiation of all the species of *Hibiscus* was possible with the *matK* DNA barcode marker. Also, in two-marker combinations only those combinations with *matK* differentiated all the species. Though all the three-marker combinations showed 100% species differentiation, species resolution was consistently better when *matK* marker formed part of the combination. These results clearly showed that *matK* is more suitable when compared to *rbcLa*, *trnH-psbA* and ITS2 for species identification in *Hibiscus*.

Keywords: *Hibiscus*, barcoding, *matK*, ITS2, divergence

Introduction

Hibiscus L. is a large genus in Malvaceae with about 200 species distributed in the tropics and subtropics of the world. Linnaeus defined the genus *Hibiscus* but indicated that it consisted of at least two distinct groupings. Miller (1754) recognized these as the genera *Hibiscus* and *Ketmia* Moench. Fabricius (1759) accepted the two-genus concept but renamed *Hibiscus* as *Malvaviscus* Cav. However, in 1787, Cavanilles united them under *Malvaviscus* (Cavanilles 1787). Bakhuizen van den et al. (1966) again segregated it to *Hibiscus* and *Malvaviscus*, which are currently accepted. In 1824, De Candolle defined 11 sections under *Hibiscus* (De Candolle 1824). Subsequently, Sivarajan and Pradeep (1996) have revised the sections to 10 retaining only 4 sections from the original description, which was followed in this study. Currently there are 28 recognized *Hibiscus* species in India, with 20 of them occurring in the southern peninsular India. Most of them are herbaceous and shrubby species, growing along roadsides, wastelands and scrub jungles. *Hibiscus* species are known for their medicinal value related to the treatment of nervous disorders and fever (Nadkarni 1976). *Hibiscus sabdariffa* L. is reported to have antihypertensive effect by inhibiting angiotensin converting enzymes (Ojeda et al. 2010). *Hibiscus rosa-sinensis* L. is commonly used for cosmetics in Indian Ayurveda and in Chinese herbal medicine. Flowers of *H. rosa-sinensis* are used for treating hair loss and extracting natural dyes (Bose et al. 2012). This is the first report on DNA barcoding of *Hibiscus*, which will be useful for authentication of the raw material used for medicinal and other applications.

DNA barcoding uses standardized short sequences of DNA, called DNA barcodes, for the purpose of species identification. Various coding and non-coding regions of plastid, mitochondrial and nuclear genomes have been suggested as plant DNA barcodes (Kress and Erickson 2007; Hollingsworth et al. 2011). However, in 2009, the Plant Working Group of

the Consortium for the Barcode of Life (CBOL) recommended *rbcLa* and *matK* as core DNA barcodes for plants (CBOL Plant Working Group 2009). The non-coding chloroplast DNA region, *trnH-psbA*, was identified as useful independent marker, and as second tier marker in the 2-tier approach to DNA barcoding (Newmaster et al. 2006; Purushothaman et al. 2014). Yet the *trnH-psbA* marker often pose problem in sequencing due to homopolymer tails resulting in stutter peaks (Shinde et al. 2003; Devey et al. 2009). The recently developed non-coding nuclear DNA barcode, ITS2, is useful in plant DNA barcoding for its ability to discriminate closely related species (Chen et al. 2010, Gu et al. 2013; Liu et al. 2014). In the present study, we have used all these four DNA barcodes individually, and in two- and three-barcode combinations to identify a best barcode or barcode combination for species identification in *Hibiscus*.

Materials and methods

Sample collection

Specimens of 44 accessions belonging to 16 species of *Hibiscus* were collected from different parts of Tamil Nadu, Kerala, and Andhra Pradesh in the southern peninsular India (Table 1). The collection included three accessions from each species, except *H. hirtus* L. and *H. trionum* L., which were represented by one accession. All the specimens were identified by Dr. A. K. Pradeep, Department of Botany, University of Calicut, Kerala, India, who is an expert in Malvaceae. The voucher specimens were mounted on herbarium sheets, and deposited to the SRM University Herbarium.

DNA extraction

Genomic DNA from fresh 100 mg of leaves was isolated using cetyl trimethyl ammonium bromide (CTAB) method with minor modifications (Doyle and Doyle 1987). The

samples were ground with 500 µl of CTAB buffer (100 mM Tris pH 8.0, 20 mM EDTA pH 8.0, 1.4 M NaCl, 2% CTAB, 2% β-mercaptoethanol and 2% PVP). The samples were transferred to 1.5 ml centrifuge tubes and the suspension was incubated at 55°C for 30 minutes. After cooling to room temperature, 500 µl of chloroform was added, mixed well, and centrifuged at 10,000 rpm for 10 minutes. The aqueous phase was transferred to fresh tubes, and an equal volume of ice-cold isopropanol was added to precipitate the DNA. The samples were centrifuged at 10,000 rpm for 10 minutes. The pellet was washed twice with 70% ethanol, air dried, and dissolved in 100 µl of TE buffer pH 8.0 (10 mM Tris, 1 mM EDTA). The DNA was checked on 0.8% agarose gel, and quantified.

PCR amplification and sequencing

The primers used for PCR amplification of the barcode markers include *rbcLa*: *rbcLa*F (ATGTCACCACAAACAGAGACTAAAGC), *rbcLa*Jf634R (GAAACGGTCTCTCCA ACGCAT) (Kress et al. 2005; Fazekas et al. 2008), *matK* : 3F_KIM (CGTACAGTAC TTTTGTGTTTACGAG), 1R_KIM (ACCCAGTCCATCTGGAAATCTTGGTTC) (Ki-Joong Kim, School of Life Sciences and Biotechnology, Korea University, Korea, unpublished), *trnH-psbA*: *psbA*3'F (GTTATGCATGAACGTAATGCTC), *trnH*R (CGCGCATGGTGGATTCAACAATCC) (Kress et al. 2005) and ITS2: S2F (ATGCGATACTTGGTGTGAAT), S3R (GACGCTTCTCCAGACTACAAT) (Chen et al. 2010). The primers were synthesized by Bioserve India Pvt Ltd, India. Polymerase chain reaction (30 µl) was performed in a thermal cycler (Eppendorf, Germany). The reaction mixture consisted of 20–50 ng of genomic DNA, 1× PCR Buffer with 1.5M MgCl₂, 200 µM dNTPs, 5 pmol primers and 1.0 U of *Taq* DNA polymerase (Genet Bio, Korea). Amplification involved initial denaturation at 95°C for 5 minutes followed by 35 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for

30 seconds and extension at 72°C for 1 minute, with a final extension at 72°C for 5 minutes. The amplicons were checked on 1% agarose gels, and purification was done using EZ-10 Spin Column PCR Purification Kit (Bio Basic Inc. Ontario, Canada). Samples were bidirectionally sequenced using 3130xl Genetic analyzer (Applied Biosystems, CA, USA).

Data Analysis

The sequences were edited manually using Sequence Scanner Software v. 1.0 (Applied Biosystems, CA, USA) and full length sequences were assembled. Sequences were submitted to Barcode of Life Data Systems (BOLD: www.boldsystems.org; Ratnasingham and Hebert 2007). Database search for species identification was done using Basic Local Alignment Search Tool (BLAST) against non-redundant nucleotide database at NCBI (www.blast.ncbi.nlm.nih.gov/Blast.cgi). Intra and inter-species pairwise divergences were calculated using TaxonDNA v. 1.6.2 (Meier et al. 2006). Divergence was calculated as the percentage of mismatched nucleotides over the total number of aligned nucleotides. Genetic distances were calculated by Kimura 2 Parameter distance model (Kimura 1980), and phylogenetic trees were constructed by Neighbor-Joining (NJ) method using ClustalW in MEGA v. 5.1. Bootstrap support was analyzed with 1,000 replications. Degree of species resolution from the phylogenetic trees was determined as described by Kim et al. 2014.

Results and discussion

Genomic DNA was successfully extracted from all 44 accessions represented by 16 species of the genus *Hibiscus*. Good PCR amplification efficiency and sequencing success rate are important for using short DNA regions as DNA barcodes for species identification (Kress and Erickson 2007; Ford et al. 2009; Hollingsworth et al. 2009). The ability to

perform bidirectional sequencing with little requirement for manual editing of the trace files is another important criterion for a successful DNA barcode marker. Though *rbcLa* was reported to give consistently good amplification and sequencing success rates, the results from other markers were highly variable (Chen et al. 2014; Krawczyk et al. 2014; Zhang et al. 2015). Larger size of the marker (~800 bp), lack of universal primers, and problems in sequencing were reported as drawbacks of *matK* (Kress et al. 2007; Wang et al 2012; Zhang et al, 2012). However, in the present study, *matK* was amplified and bidirectionally sequenced from all the 44 accessions. The sequence quality was good and open reading frames were found to be intact for the coding markers *rbcLa* and *matK*. As expected, there was no size variation in the *rbcLa* marker. Size variation was highest in *trnH-psbA* (491 - 703 bp), followed by *matK* (804 - 846 bp), and ITS2 (456 - 478 bp).

Intra- and inter-species divergences were calculated using the four markers individually, and in two-, and three-marker combinations. Intra-species divergence was zero in species for which multiple accessions were analyzed. Inter-species divergence calculated based on individual markers and marker combinations are given in Table 2. Among the four markers, *matK* was the only marker, which could differentiate all the species and the inter-species divergence ranged between 0.3 and 6.5%. In multi-locus DNA barcoding, *matK*+ITS2 (divergence: 0.9 to 12.5%) and *matK*+ITS2+*trnH-psbA* (divergence: 2.6 to 20.3%) were found to be the most suitable two-, and three-marker combinations for species differentiation. In our earlier study on *Sida* L., (Malvaceae), species discrimination power of *matK* was not better than *trnH-psbA* or ITS2 (Vassou et al. 2015). In *Gossypium* L., (Malvaceae), *matK* was found to be useful in species differentiation only when it was combined with ITS2 (Ashfaq et al. 2013). Though *matK*, in general, has less power for species differentiation (Kress et al., 2007; Zhang et al., 2015), it has been reported to identify

80% of the species in different genera of Fabaceae with 100% identification in *Vigna* Savi (Gao et al. 2011).

Phylogenetic trees were constructed using the data from four markers individually, and in two-, and three-marker combinations (Fig.1, Fig.S1, Fig.2, Fig.S2, Fig.S3). The 16 species of *Hibiscus* included in the present study included 9 out of the 10 sections represented in this genus. All the sections, except *Trichospermum* and *Bombicella*, formed monophyletic clades in the trees constructed using the data from single markers. *Hibiscus lunariifolius* of section *Trichospermum* and *H. platanifolius* of *Spatula* were clustered within one clade. Species of these two sections were reported to be closely related based on shared morphological characters (Sivarajan and Pradeep 1996). Two-and three-marker combinations also did not resolve these two sections. The six species of section *Furcaria* were found to be more difficult to differentiate using individual markers. Morphologically, section *Furcaria* could be clearly distinguished from other sections based on distinct morphological features such as 10-costate calyx, bifurcate involucellar bracteoles, aculei on stems, and nectar glands. However, species delimitation within this section was difficult due to overlapping morphological characters (Sivarajan and Pradeep 1996). In our study, the section *Furcaria* formed a clade in which *matK* differentiated all the species but *rbcLa* and ITS2 differentiated only one and two species, respectively. Unexpectedly, the resolving power of *trnH-psbA* was lower than *rbcLa* since it did not differentiate any species in this clade (Fig.S1). In two-marker combinations also only those combinations, which included *matK* differentiated all species of this section. However, all the three-marker combinations differentiated this section very well (Fig.2, Fig. S3). The degree of species resolution for individual barcode regions ranged between 56 and 100%, which was significantly enhanced in multi-region combinations. Three of the two-marker combinations and all the three-marker combinations showed 100% species resolution (Table 2). Inclusion of *matK* as a member of multi-locus

combination consistently showed better species resolution. Phylogenetic trees constructed using *matK* marker alone or in combination with *rbcLa* or *trnH-psbA* showed better species resolution in Lamiaceae (Theodoridis et al. 2012). Species resolution in the recently evolved genus *Holcoglossum* in Orchidaceae was found to be the best with *matK* marker, which was further improved when combined with ITS marker (Xiang et al. 2011).

Conclusion

The present study showed that *matK* either alone or in combination with ITS2 or *trnH-psbA* is best suited for species identification in *Hibiscus*. Therefore, DNA barcoding using these markers can be successfully used for authentication of the plant materials derived from this genus.

Acknowledgement

We acknowledge Dr. A. K. Pradeep (Department of Botany, University of Calicut, Kerala, India) who helped in collection of some species of *Hibiscus* and taxonomic identification of all specimens. We also thank Mr. K. Devanathan, Mr. N. Harshavardhan, Dr. M. Udaya Kumar for helping in collecting the species. Funding from SRM-DBT Partnership Platform for Contemporary Research Services and Skill Development in Advanced Life SciencesTechnologies (Order No. BT/PR12987/INF22 / 205 / 2015) is acknowledged.

References

Ashfaq, M., Asif, M., Anjum, Z.I., and Zafar, Y. 2013. Evaluating the capacity of plant DNA barcodes to discriminate species of cotton (*Gossypium*: Malvaceae). *Mol Ecol Resour.* **13**(4):573-82. doi: 10.1111/1755-0998.12089

- Bakhuzien van den Brink, R.C., Van Borssum Waalkes, J.W., and Van Steenis, C.G.G.J. 1966. Proposals on Hibiscus and Malvaviscus. *Taxon* **15**: 43.
- Bose, S., and Nag, S. 2012. Isolation of Natural Dyes from the Flower of Hibiscus rosa-sinensis. *Am. J PharmTech Res* **12**(3): 761-770. <http://www.ajptr.com/archive/volume-2/june-2012-issue-3/article-240.html> [accessed 30 November 2015]
- Candolle, A.P.de. 1824. Malvaceae. *Prodromus Systematis naturalis regni vegetabilis*, 1. Paris
- Cavanilles, A.J. 1785-1790. *Monodelphiae classis dissertations decem*. Paris & Madrid.
- CBOL, 2009. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. U. S. A.* **106**(31):12794-12797. doi: 10.1073/pnas.0905845106
- Chen, J., Zhao, J., Erickson, D.L., Xia, N., and Kress, W.H. 2014. Testing DNA barcodes in closely related species of Curcuma (Zingiberaceae) from Myanmar and China. *Mol Ecol Resour* **1**: 1-12. doi: 10.1111/1755-0998.12319
- Chen, S.L., Yao, H., Han, J.P., Liu, C., Song, J.Y., Shi, L.C., et al., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* **5**(1):e8613. doi: 10.1371/journal.pone.0008613
- Devey, D.S., Chase, M.W., and Clarkson, J.J.A. 2009. A stuttering start to plant DNA barcoding: microsatellites present a previously overlooked problem in non-coding plastid regions. *Taxon* **58**(1): 7-15. URL:<http://www.jstor.org/stable/27756818>. [accessed 30 November 2015]
- Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15.

- Fabricius, P.C. 1759. *Enumeratio methodical plantarum*, ed.1. Helmstedt
- Fazekas, A.J., Burgess, K.S., Kesanakurti, P.R., Graham, S.W., Newmaster, S.G., Husband, B.C., et al., 2008. Multiple multi locus DNA barcode from the plastid genome discriminate plant species equally well. *PLoS One* **3**(7):e2802 doi: 10.1371/journal.pone.0002802.
- Ford, C.S., Ayres, K.L., Toomey, N., Haider, N., Stahl, J.V.A., Kelly, L. J., Wikstrom, N., et al. 2009. Selection of candidate coding DNA barcoding regions for use on land plants. *Bot. J. Linn. Soc* **159**(1): 1-11. doi:10.1111/j.1095-8339.2008.00938.x
- Gao, T., Zhiying, S., Yao, H., Kingyuan, S., Yingjie, Z., Ma, X., and Shilin C. 2011. Identification of fabaceae plants using the DNA barcode matK. *Planta Med.* **77**: 92-94 doi: <http://dx.doi.org/10.1055/s-0030-1250050>
- Gu, W., Song, J., Cao, Y., Sun, Q., Yao, H., Wu, Q., et al., 2013. Application of the ITS2 region for barcoding medicinal plants of Selaginellaceae in Pteridophyta. *PLoS One* **8**(6): e67818 doi: 10.1371/journal.pone.0067818
- Hollingsworth, P.M., Graham, S.W., and Little, D.P. 2011. Choosing and Using a Plant DNA Barcode. *PLoS ONE* **6**(5): e19254. doi:10.1371/journal.pone.0019254
- Kim, H.M., Oh, S.H., Bhandari, G.S., Kim, C.S., and Park, C.W. 2014. DNA barcoding of Orchidaceae in Korea. *Mol Ecol Resour* **14**: 499–507. doi: 10.1111/1755-0998.12207
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**(2): 111–120
- Krawczyk, M., Szczecinska, M., and Sawicki, J. 2014. Evaluation of 11 single-locus and seven multilocus DNA barcodes in *Lamium* L. (Lamiaceae) *Mol Ecol Resour* **14**: 272–285. doi: 10.1111/1755-0998.12175

- Kress, W.J., and Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One* **2**(6): e508. doi: 10.1371/journal.pone.0000508
- Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., and Janzen, D.H. 2005. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 8369–8374 doi: 10.1073/pnas.0503123102
- Liu, J., Shi, L., Han, J., Li, G., Lu, H., Hou, J., Zhou, X., Meng, F., and Downie, S.R. 2014. Identification of species in the Angiosperm family Apiaceae using DNA barcodes. *Mol. Ecol. Resour.* **14**(6): 1231–1238. doi: 10.1111/1755-0998.12262
- Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* **55**: 715–728. doi: 10.1080/10635150600969864
- Miller, P., 1754. *Gardener's Dictionary*, Abridg. ed. 4. London.
- Nadkarni, K.M. 1976. *Indian Materia Medica*. Popular Prakashan Pvt. Ltd., Mumbai.
- Newmaster, S.G., Fazekas, A.J., and Ragupathy, S. 2006. DNA barcoding in the land plants: evaluation of *rbcL* in a multigene tiered approach. *Can. J. Bot.* **84**: 335–341. doi: 10.1139/B06-047
- Ojeda, D., Jimenez-Ferrer, E., Zamilpa, A., Herrera-Arellano, A., Tortoriello, J., and Alvarez, L. 2010. Inhibition of angiotensin convertin enzyme (ACE) activity by the anthocyanins delphinidin- and cyaniding-3-O-sambubiosides from *Hibiscus sabdariffa*. *J of Ethnopharmacol.* **127**: 7–10. doi: 10.1016/j.jep.2009.09.059.

- Purushothaman, N., Newmaster, S.G., Ragupathy, S., Stalin, N., Suresh, D., Arunraj, D.R., Gnanasekaran, G., Vassou, S.L., Narasimhan, D., and Parani, M. 2014. A tiered barcode authentication tool to differentiate medicinal *Cassia* species in India. *Genet. Mol. Res.* **13**(2):2959–2968 doi: 10.4238/2014
- Shinde, D., Lai, Y., Sun, F., and Arnheim, N. 2003. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)_n and (A/T)_n microsatellites. *Nucleic Acids Res.* **31**(3): 974-980. doi:10.1093/nar/gkg178
- Sivarajan, V.V., and Pradeep, A.K. 1996. *Malvaceae of Southern Peninsula India: A Taxonomic Monograph*. Daya Publishing House, New Delhi.
- Theodoridis, S., Stefanaki, A., Tezcan, M., Aki, C., Kokkini, S., and Vlachonasios, K.E. 2012. DNA barcoding in native plants of the Labiatae (Lamiaceae) family from Chios Island (Greece) and the adjacent Cesme-Karaburun Peninsula (Turkey). *Mol Ecol Resour.* **12**(4): 620–633. doi:10.1111/j.1755-0998.2012.03129.x
- Vassou, S.L., Kusuma, G., and Parani, M. 2015. DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market samples of *Sida cordifolia*. *Gene.* **559**(1):86-93 doi: 10.1016/j.gene.2015.01.025
- Wang, N., Jacques, F.M.B., Milne, R.I., Zhang, C.Q., and Yang, J.B. 2012. DNA barcoding of Nyssaceae (Cornales) and taxonomic issues. *Bot Stud.* **53**: 265-274. URL: <http://ir.xtbg.org.cn/handle/353005/2965> [Accessed on 30 November 2015]
- Xiang, X.G., Hu, H., Wang, W., and Jin, X.H. 2011. DNA barcoding of the recently evolved genus *Holcoglossum* (Orchidaceae: Aseridinae): a test of DNA barcode candidates. *Mol Ecol Resour.* **11**: 1012–1021. doi: 10.1111/j.1755-0998.2011.03044.x.

Zhang, C.Y., Wang, F.Y., Yan, H.F., et al. 2012 Testing DNA barcoding in closely related groups of *Lysimachia* L. (Myrsinaceae). *Mol Ecol Resour*, **12**(1), 98–108. doi: 10.1111/j.1755-0998.2011.03076.x

Zhang, Z., Song, M., Guan, Y., Li, Hai., Niu, Y., Zhang, L., and Ma, X. 2015. DNA barcoding in medicinal plants: Testing the potential of a proposed barcoding marker for identification of *Uncaria* species from China. *Biochem Syst Ecol* **60**: 8-14.
<http://dx.doi.org/10.1016/j.bse.2015.02.017>

Draft

Figure captions

Figure 1: Phylogenetic trees constructed using neighbour joining method for 16 species of *Hibiscus* based on individual DNA barcodes a) *matK* and b) ITS2

Figure 2: Phylogenetic trees of *Hibiscus* species constructed using neighbour joining method for the best two-marker (a) and three-marker combination (b)

Draft

Supplement data figure captions

Figure S1: Phylogenetic trees constructed using neighbour joining method for the species of *Hibiscus* based on individual DNA barcodes a) *rbcLa*, b) *trnH-psbA*

Figure S2: Phylogenetic trees constructed using neighbour joining method for the species of *Hibiscus* based on combination of DNA barcodes a) *rbcLa+matK*, b) *rbcLa+trnH-psbA*, c) *rbcLa+ITS2*, d) *matK+trnH-psbA*, e) *trnH-psbA+ITS2*

Figure S3: Phylogenetic trees constructed using neighbour joining method for the species of *Hibiscus* based on combination of DNA barcodes a) *rbcLa+matK+trnH-psbA*, b) *rbcLa+matK+ITS2*, c) *rbcLa+ITS2+trnH-psbA*

Draft

Table 1: Details of the species of *Hibiscus* collected for the present study

S. No.	Section	Species Name	Voucher ID	No. of accessions	Place of Collection	BOLD Accession No.	GPS Co-ordinates
1	Azanzae DC.	<i>Hibiscus tiliaceus</i> L.	B0512A, B, C	3	Tamil Nadu	SRM000984A, B, C	12°41'N 79°58'E, 12°49'N 80°02'E, 13°00'N 80°14'E
		<i>Hibiscus hirtus</i> L.	B0936A	1	Kerala	SRM000973A	11°08'N 75°53'E
2	Bombicella DC.	<i>Hibiscus micranthus</i> L.	B0940A, B, C	3	Tamil Nadu	SRM000976A, B, C	13°00'N 80°14'E, 09°11'N 77°52'E, 12°41'N 79°58'E
		<i>Hibiscus acetosella</i> Welw. Ex Hiern	B0974A, B, C	3	Tamil Nadu	SRM000970A, B, C	12°55'N 80°06'E, 12°41'N 79°58'E, 12°50'N 80°03'E
		<i>Hibiscus cannabinus</i> L.	B0615A, B, C	3	Andhra Pradesh	SRM000971A, B, C	13°14'N 79°06'E, 13°37'N 79° 24'E, 14°26'N 79°59'E
		<i>Hibiscus hispidissimus</i> Griffith	B0776A, B, C	3	Kerala	SRM000972A, B, C	11°08'N 75°53'E, 11°35'N 76°06'E, 11°00'N 75°59'E
		<i>Hibiscus radiatus</i> Cav.	B0708A, B, C	3	Tamil Nadu	SRM000980A, B, C	12°55'N 80°06'E, 12°47'N 80°01'E, 11°39'N 78°17'E
		<i>Hibiscus sabdariffa</i> L.	B0778A, B, C	3	Tamil Nadu	SRM000982A, B, C	12°55'N 80°06'E, 12°41'N 79°58'E, 12°50'N 80°03'E
3	Furcaria DC.	<i>Hibiscus surattensis</i> L.	B0947A, B, C	3	Tamil Nadu	SRM000983A, B, C	12°41'N 79°58'E, 13°00'N 80°14'E, 09°11'N 77°52'E
	Lilibiscus						
4	Hochr.	<i>Hibiscus rosa-sinensis</i> L.	B0511A, B, C	3	Tamil Nadu	SRM000981A, B, C	12°47'N 80°01'E, 12°49'N 80°02'E, 12°55'N 80°06'E
	Solandra						
5	Hochr.	<i>Hibiscus lobatus</i> (Murr.) Kuntze	B0665A, B, C	3	Kerala	SRM000974A, B, C	11°35'N 76°06'E, 11°08'N 75°53'E, 11°00'N 75°59'E
		<i>Hibiscus platanifolius</i> (Willd.)					
6	Spatula Hochr.	Sweet	B0944A, B, C	3	Kerala	SRM000979A, B, C	11°35'N 76°06'E, 11°08'N 75°53'E, 11°00'N 75°59'E
	Trichospermum	<i>Hibiscus lunariifolius</i> Willd.	B0938A, B, C	3	Kerala	SRM000975A, B, C	11°35'N 76°06'E, 11°08'N 75°53'E, 11°00'N 75°59'E
7	Hochr.	<i>Hibiscus panduriformis</i> Burm. f.	B0943A, B, C	3	Tamil Nadu	SRM000978A, B, C	12°41'N 79°58'E, 09°11'N 77°52'E, 11°39'N 78°17'E
8	Trionum DC.	<i>Hibiscus trionum</i> L.	B0949A	1	Kerala	SRM000985A	11°35'N 76°06'E
9	Venusti Ulbr.	<i>Hibiscus mutabilis</i> L.	B0941A, B, C	3	Tamil Nadu	SRM000977A, B, C	12°50'N 80°03'E, 11°39'N 78°17'E, 09°11'N 77°52'E

Table 2: Inter-species divergence in *Hibiscus* based on individual barcodes and two-barcode combinations

S.No.	Barcode marker	Inter-species divergence (%)	Species Resolution (%)
1	<i>rbcL</i>	0-2.8	56
2	<i>matK</i>	0.3-6.5	100
3	<i>trnH-psbA</i>	0-9.6	56
4	<i>ITS2</i>	0-8.5	75
5	<i>rbcL+matK</i>	0.4-9	100
6	<i>rbcL+trnH-psbA</i>	0-9.6	88
7	<i>rbcL+ITS2</i>	0-8.5	88
8	<i>matK+trnH-psbA</i>	0.8-11.2	100
9	<i>matK+ITS2</i>	0.9-12.5	100
10	<i>trnH-psbA+ITS2</i>	0-9.6	75
11	<i>matK+trnH-psbA+ITS2</i>	2.6-20.3	100
12	<i>rbcL+matK+ITS2</i>	1.7-15.8	100
13	<i>rbcL+matK+trnH-psbA</i>	1.5-17.2	100
14	<i>rbcL+ITS2+trnH-psbA</i>	0.8-16.2	100

(a)

(b)

(c)

(d)

(e)

(a)

(b)

(c)