Int.J.Curr.Microbiol.App.Sci (2018) *7*(12): 117-124

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 7 Number 12 (2018)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2018.712.015

DNA Barcoding of *Psoralea corylfolia*, *Mucuna pruriens* and *Clitoria ternatea* for Species Identification

K.V. Deepak, V.A. Mohanlal*, J. Johnny Subakar Ivin and R. Anandan

Department of Genetics and Plant Breeding, Faculty of Agriculture, Annamalai University, Annamalai Nagar - 608 002, Tamil Nadu, India

**Corresponding author*

Keywords

DNA barcode, *P. corylifoilia*, *M. pruriens*, *C. ternatea*, *trnH-psbA*, Phylogeny

Article Info

Accepted:
04 November 2018
Available Online:
10 December 2018

A B S T R A C T

DNA barcoding is a useful tool for species identification and phylogenetic construction. We tested the genus/species identification and phylogenetic construction in three tropical medicinal plants such as *P. corylifoilia*, *M. pruriens* and *C. ternatea* using a universal chloroplast DNA barcode *trnH-psbA*. The *trnH-psbA* primer pairs were used to amplify 250 bp DNA fragment in all three genera. The barcode sequences were submitted on NCBI-BLAST search and showed the nucleotide sequence similarity (homology) of 96% with *P. corylifolia*, 100% with *M. pruriens* and with 77% *C. ternatea*. The phylogenetic tree constructed using Neighbour joining method with the sequence of *trnH-psbA* revealed that grouping of the entire three different genera within the family Fabaceae.

Introduction

Logistic assumption that genetic variation between the species is always larger than the variation existing within the species facilitated the evolvement of a molecular profiling technique known as DNA barcode. The limitations inherent in morphology-based identification systems and the dwindling pool of taxonomists signal the need for a new approach to species recognition. DNA barcoding seeks to advance both species identification and discovery through the analysis of patterns of sequence divergence in a standardized gene region (Kress *et al.*, 2009). There is no single universal DNA barcode marker for plants, and each marker has its own advantages and disadvantages. DNA barcoding technique is being successfully exploited in resolving taxonomic ambiguity (Pettengill and Neel, 2010), assessing biodiversity (Lahaye *et al.*, 2008), establishing phylogenetic relationship (Zimmermann *et al.*, 2013), determining the sex of the plants at juvenile stage (Gonzalez *et al.*, 2009) and determining the purity of the herbal samples (Newmaster *et al.*, 2013) besides addressing the issues related to forensic analyses (Ferri *et al.*, 2009). Plant working group of the Consortium for the Barcode of Life (CBOL) has recommended many candidate plant barcodes, including the nuclear internal transcribed spacer (*ITS*) regions, chloroplast intergenic spacers (*trnH–*

psbA) and chloroplast coding regions such as ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (*rbcL*), maturase K (*matK*) as core DNA barcodes for species discrimination in plants (CBOL Plant Working Group, 2009). With this background the present study was pre-set with the objective of to establish DNA barcodes of *Psoralea corylfolia*, *Mucuna pruriens*, *Clitoria ternatea* using *trnH-psbA* marker and to elucidate the phylogenetic relationships with genera from the family Fabaceae.

Materials and Methods

DNA extraction

Cetyl Trimethyl Ammonium Bromide (CTAB) extraction protocol was used to isolate the total genomic DNA as described by Doyle and Doyle (1990).Leaf tissue samples (2 grams) were collected from above mentioned medicinal plants and grinded in pestle and mortar by adding 450µl of pre heated (65˚C) CTAB buffer. Extracted samples were incubated in the water bath for 30 minutes at 65˚C. After incubation, around 450µl of chloroform: Isoamyl alcohol (24:1) was added in to the tubes and inverted twice to mix. Then the tubes were kept in centrifuge for 10 minutes at 12000 rpm. Then the aqueous layer was transferred in to the new Eppendorf tubes. An amount of equal volume of isopropanol (stored at -20˚C) was added to each sample and inverted once in mix and kept at overnight at 4˚C. The samples were centrifuged at 5500 rpm for15 minutes on the next day. The supernatant was discarded from each sample and the pellets settled in the bottom were air dried for 30 minutes. A quantity of 100 µl of TE buffer was added into each sample and stored it overnight at 4˚C. RNase (3µl) was added into each sample to exclude the RNA contamination on the following day. An amount of 200µl chloroform: Isoamyl alcohol (24:1) was added to the tubes and centrifuged at 5000 rpm for 5 minutes and the supernatant was taken into the fresh tubes. To which the twice the volume of absolute ethanol and 1/10th volume of 3M sodium acetate was added and kept the samples at 4˚C for overnight. Centrifuge the tubes at 5500 rpm for 15 minutes. The supernatant was discarded and 200µl of 70% ethanol was added and centrifuged at 5000 rpm for 5 minutes. The supernatant was discarded and the pellet was air dried for 30 minutes. The pellet was resuspended by using 100µl TE buffer (10mM Tris; 0.1mM EDTA; pH 8.0) and kept at -20˚C for long term use.

DNA amplification

The DNA barcode marker *trnH-psbA* (*trnH*-F 5'-CGCGCATGGTGGATTCACAATCC-3' and *trnH*-R 5'-GTTATGCATGAACGTAATGCTC-3') was used to explore the taxonomic/phylogenetic relationship of above mentioned medicinal plants. Total volume for amplification was 20 µl consisted of 3 µl DNA sample, 1 µl of each primer and PCR master mix. For PCR reactions, the initial temperature set for denaturation of double stranded DNA was 94ºC. Then the primers annealing temperature was set at 55ºC (depending on the primer base composition and length). Extension temperature was set at 72ºC. The cycle was repeated 35 times for the amplification of the desired DNA sequence. The PCR reactions were carried out on Master cycler nexus gradient (Eppendorf AG).Sequencing was done by outsourcing (Eurofins Genomics India Pvt. Ltd, Bangalore).

Data analysis

The species identification success rate was calculated using genetic distance and BLAST methods at http://blast.ncbi.nlm.nih.gov/Blast.cgi, using all *trnH-psbA* sequences as query sequences.

Graphical display had shown where the query is similar to the sequence, hit list, sequence accession number and the name, description, the bit score, the e- value (the expectation value), the alignment and length. The three sequences were aligned and the conserved sequence in the generated sequence was identified using the Software MEGA 6.0. The sequences were used to construct the phylogenetic tree and the inter-specific distances were calculated using K2P (Kimura 2 parameter) model as recommended by the Consortium of Barcode of Life (CBOL, http:// www.barcoding.si.edu/protocols.html) using MEGA 6.0 software.

Results and Discussion

The DNA barcode was used for precisely identifying unknown species, with no taxonomic expertise (Hebert *et al.,* 2003). The nuclear and chloroplast genomes are the major targets for plant species authentication and phylogenetic studies. In the chloroplast genome, the *trnH-psbA* spacer suitable for identification at the species level because it evolves so quickly that provides enough character to analyse evolution below family level (Barthet, 2006; Kondo *et al.,* 2007; Hollingsworth, 2011). Species identification in Fabaceae family is important in order to distinguish invasive species from endangered species or species that has an economic significance. The species, *P. corylifolia*, *M. pruriens* and *C. ternatea* were characterised using the universal barcoding primer pairs of *trnH-psbA* (Fig. 1).

The sequences were subjected to BLAST analyze to discriminate the species. The BLAST analysis of *trnH-psbA* sequences of *P. corylifolia* showed 96 per cent maximum identity and 100 percent query coverage. The accession number is GU396699.1 and GQ434957.1. The sequences of *M. pruriens* covered 100 per cent query and the maximum identity is 100 per cent in KX606897.1 and KX606896.1. The NCBI Gen Bank accession numbers KY806280.1 and KJ468096.1 showed 100 per cent and 77 percent of respective maximum identity of *C. ternatea* sequences. But the query coverage for the accessions were 10 and 43 respectively (Table 1).

The multiple sequence alignment was performed between the three species and already known or deposited sequences in NCBI Gen Bank to observe the nucleotide variations among them. As the results revealed by different figures, the figures 2 to 7 showed the quality and variations of the sequences of species and accessions. The variations of nucleotides are showed in different colours.

The conserved sequence in the generated sequence was identified. A total of nine sequences were aligned and the conserved regions across the total length of 360 base pairs was identified and represented as "*" symbol. The results showed similar nucleotides for three species at 35 bp (A), 46 bp (T), 49 bp (T) and 58 bp (A). The low number of conserved regions was attributed to the query coverage of *C. ternatea* which caused the anomaly during multiple sequence alignment.

The genetic similarity was determined for the *P. corylifolia*, *M. pruriens*, *C. ternatea* and six accessions based on *trnH-psbA* sequences. The phylogenetic tree generated had shown II clusters (Fig. 8). Cluster I was further divided into 2 sub-clusters (Ia and Ib). Sub-cluster Ia had *Mucuna pruriens* and it was closely associated with *P. corylifolia*. *M. pruriens* (n=11) and *P. corylifolia* (n=11) are same in their chromosome number. Based on the chromosome number, two species may be grouped under one cluster. Interestingly, the two accessions KX606897.1 and KX606896.1 had fallen within the *M. pruriens*.

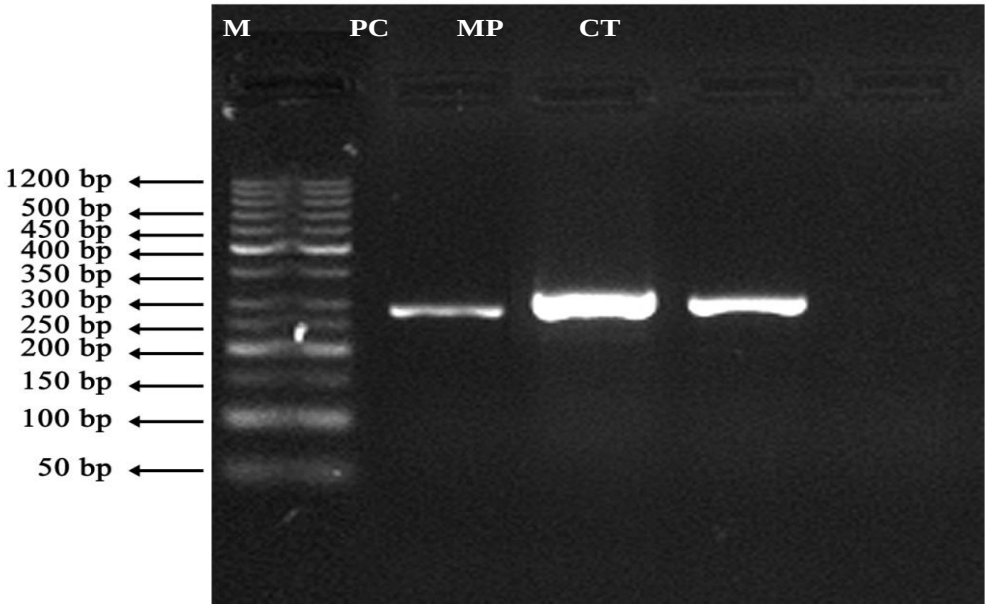
Table.1 NCBI-BLAST similarity results of three species of *trnH-psbA* sequences

S. No.	Species	Sequence length (bp)	Query coverage (%)	Maximum identity (%)	BLAST "E" value	Accession No.
1	*Psoralea corylifolia*	250	100	96	$2e^{-111}$	GU396699.1
			97	96	$4e^{-108}$	GQ434957.1
2	*Mucuna pruriens*	355	100	100	0.00	KX606897.1
			100	100	0.00	KX606896.1
3	*Clitoria ternatea*	298	10	100	$4e^{-05}$	KY806280.1
			43	77	$4e^{-05}$	KJ468096.1

Table.2 The distance among the accession based on the *trnH-psbA* sequences, as derived using MEGA 6.0

S. No.	Species	1	2	3	4	5	6	7	8	9
1.	*Psoralea corylifolia*	0.000								
2.	GU396699.1	1.816	0.000							
3.	GQ434957.1	1.816	0.034	0.000						
4.	*Mucuna pruriens*	0.400	1.622	1.622	0.000					
5.	KX606897.1	0.400	1.622	1.622	0.000	0.000				
6.	KX606896.1	0.400	1.622	1.622	0.000	0.000	0.000			
7.	*Clitoria ternatea*	1.713	1.856	1.483	2.663	2.663	2.663	0.000		
8.	KY806280.1	1.729	1.962	1.962	2.499	2.499	2.499	0.034	0.000	
9.	KJ468096.1	1.729	1.962	1.962	2.499	2.499	2.499	0.034	0.000	0.000

Fig.1 Agarose Gel Electrophoresis of *trnH-psbA* based PCR product for sequencing analysis



M – Ladder (50-1200 bp) PC – *Psoralea corylifolia*, MP – *Mucuna pruriens* and CT – *Clitoria ternatea*

Fig.2 Multiple sequence alignment (1-60 Nucleotides)

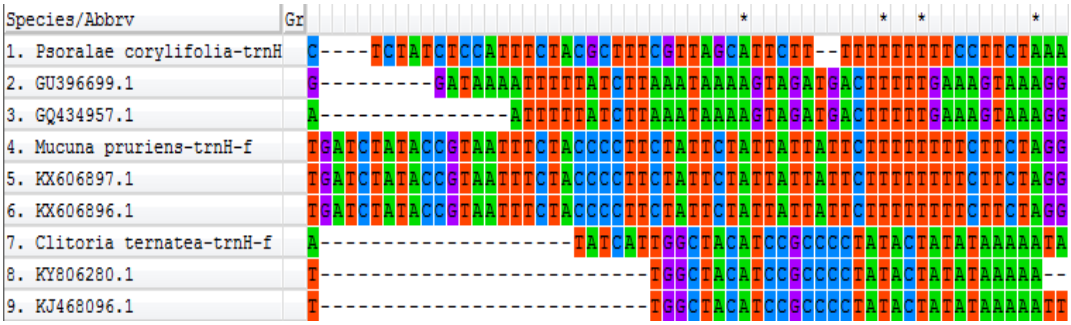


Fig.3 Multiple sequence alignment (61-120 Nucleotides)

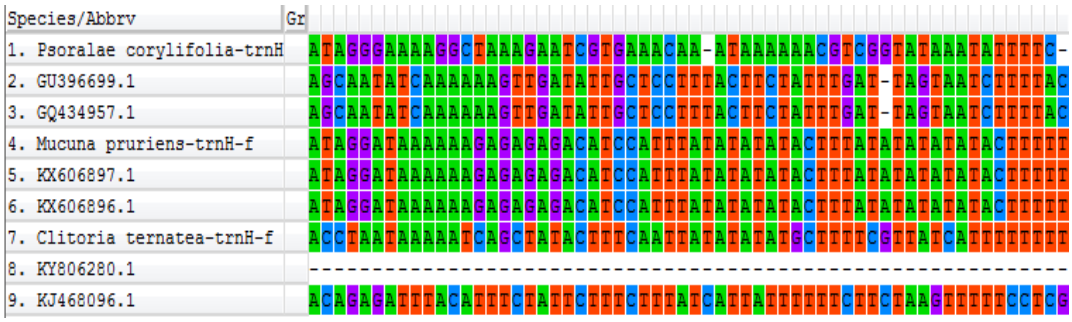


Fig.4 Multiple sequence alignment (121-180 Nucleotides)

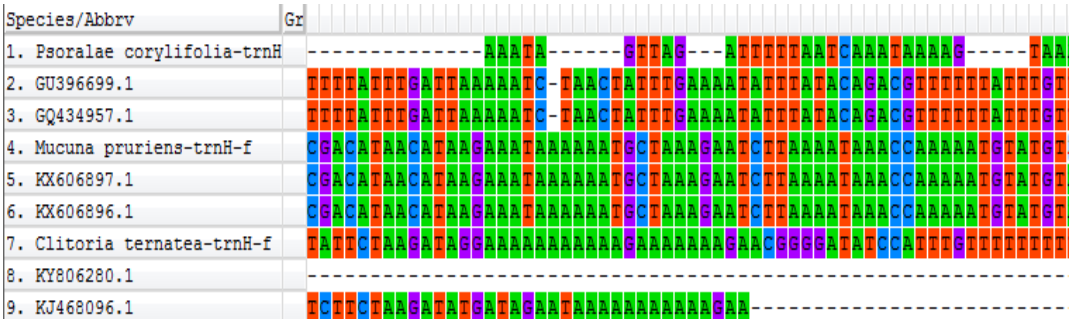


Fig.5 Multiple sequence alignment (181-240 Nucleotides)

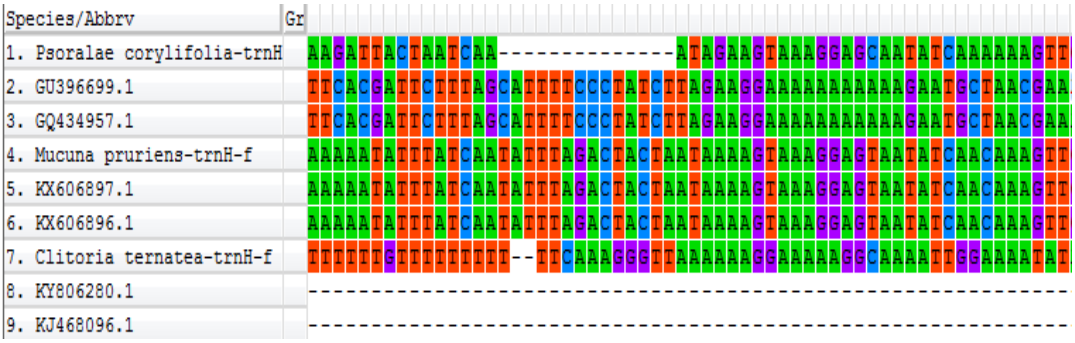


Fig.6 Multiple sequence alignment (241-300 Nucleotides)

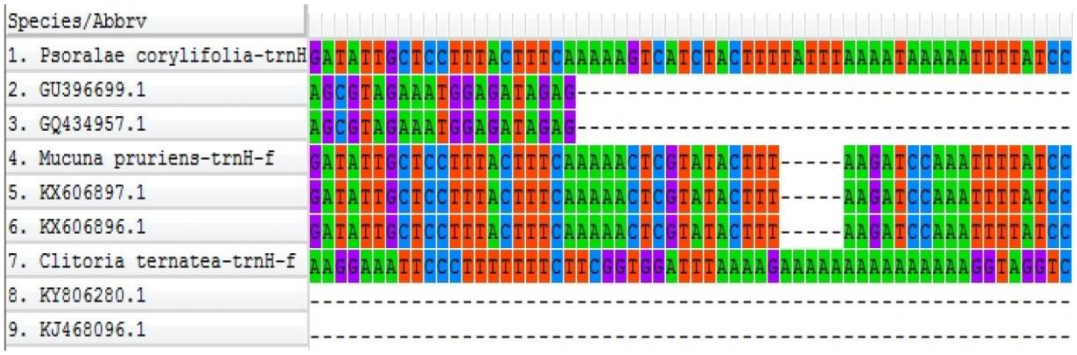


Fig.7 Multiple sequence alignment (301-360Nucleotides)

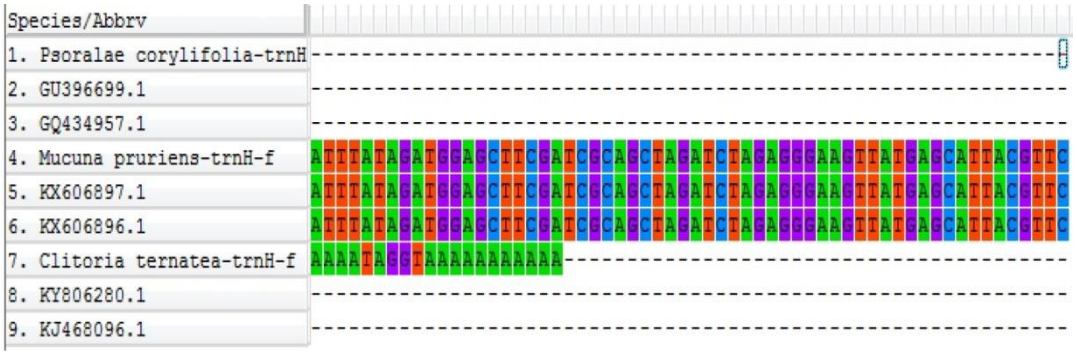
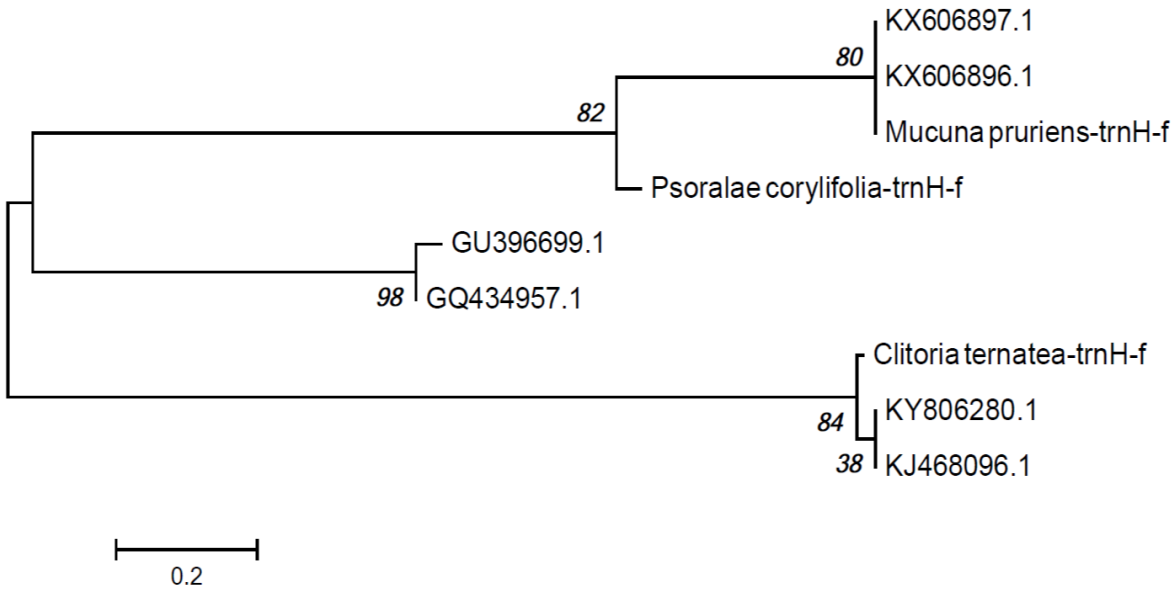


Fig.8 The phylogenetic tree of *Psoralea corylifolia*, *Mucuna pruriens*, *Clitoria ternatea* and different accessions



This revealed that *M. pruriens* and the other two accessions have the same genome content. Cluster II had *C.ternatea* and two accessions KY806280.1 and KJ468096.1. *M. pruriens* and *C. ternatea* have been grouped in separate clusters, showing the accurate diversity among the species. Similar studies were carried out by Jaheer *et al.,* (2015) who analysed the sequence of *trnH-psbA* and resolved monophyletic tree for annual species represented by *M. pruriens* varieties with *M. bracteata* emerging as an ancestor.

Genetic diversity assessment is also important of species that are endemic, rarely found or endangered, because it helps in plant conservation (Tallei *et al.,* 2014). The interspecific distances were calculated using K2P (Kimura 2 parameter). The range of pairwise dissimilarity was 0.034 to 2.663 with an average of 1.543. Among the three species, maximum genetic distance was observed between *C. ternatea* and *M. pruriens* (2.663) and minimum between KY806280.1 and *C. ternatea* (0.034) (Table 2). The dissimilarity between *C. ternatea* and *M. pruriens* had shown maximum genetic distance (2.663) indicating diverging origin of the two species. Thus, *trnH-psbA* could discriminate these three species of different genera. Similar study was conducted by Wong *et al.,* (2013) to discriminate between genera *Gentiana* and *Podophyllum.* But one of the challenges in the selection of this barcode is to distinguish very closely related or newly developed species.

A species cannot be precisely identified using DNA barcodes if the variation within barcode between species is low or related species still retain ancestral polymorphism or they have a history of hybridization (Dick and Kress, 2009). Ideally, a DNA barcode should be routinely used as a primer pair, which can also be used for bidirectional sequencing and provides the maximum level of discrimination between species (Janzen, 2009).

Nucleotide sequence submission to NCBI

Deposition of sequences was carried out with the submission tool BankIt to GenBank. NCBI have provided a GenBank accession number for the nucleotide sequence of *Clitoria ternatea.* The accession number for *Clitoria ternatea* is MH535465.

References

Barthet, M.M. 2006. Expression and function of the chloroplast-encoded gene *matK.* Ph. D thesis, Virginia Polytechnic Institute and State University, Virginia, pp. 213.

Dick, C.W., and Kress, W. J. 2009. Dissecting tropical plant diversity with forest plots and a molecular toolkit. *Bio. Science,* 59(9): 745-755.

Doyle, J.J. and Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. *Focus*. 12: 13-15.

Ferri, G., Alu, M., Corradini, B. and Beduschi, G. 2009. Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. *Int. J. Leg. Med.* 123: 395-401.

Gonzalez, M.A., Baraloto, C., Engel, J., Mori, S.A., Petronelli, P., Riera, B., Roger, A. and Thebaud, C. 2009. Identification of Amazonian trees with DNA barcodes. *PLoS One.* 4: e7483.

Hebert, P.D.N., Ratnasingham, S. and De Waard, J.R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. Biol. Sci.* 270: 96-99.

Hollingsworth, P.M. 2011. Refining the DNA barcode for land plants. *Proc Natl Acad Sci USA*.108: 19451-19452.

Jaheer, M., Chopra, R., Kunder, K.R., Bhat, D., Rashmi, K.V. and Sathyanarayana, N.2015. Cytogenetic and ITS-psbA-trnH sequence analysis for phylogenetic

inference in Mucuna sp. of India. *Tropical plant biology*. 8(3-4): 108-116.

Janzen, D.H. 2009. A DNA barcode for land plants. *PNAS.*, 106(31): 12794-12797.

Kondo, K., Shiba, M., Yamaji, H., Morota, T., Zhengmin, C., Huixia, P. and Shoyama, Y. 2007. Species identification of licorice using nrDNA and cpDNA genetic markers. *Biological and Pharmaceutical Bulletin*. 30(8): 1497-1502.

Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., Perez, R. and Sanjur, O. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Natl Acad Sci., U.S.A. 106(44): 18621-18626.

Lahaye, R., Van Der Bank, M., Bogarin, D., Warner, J., Pupulin. F. and Gigot, G. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences.* 105(8): 2923-8.

Newmaster, S. G., Grguric, M., Shanmughanandhan, D., Ramalingam, S. and Ragupathy, S. 2013. DNA barcoding detects contamination and substitution in North American herbal products. *BMC medicine.* 11(1): 222.

Pettengill, J.B. and Neel, M.C. 2010. An evaluation of candidate plant DNA barcodes and assignment methods in diagnosing 29 species in the genus *Agalinis* (Orobanchaceae). *Am. J. Bot.*, 97: 1391-1406.

Tallei, T. E. and kolondam, B.J. 2015. DNA barcoding of Sangihe Nutmeg (*Myristicafragrans*) using matK gene. *HAYATI Journal of Biosciences*, 22(1): 41-47.

Zimmermann, T., Bocksberger, G., Bruggemann, W. and Berberich, T. 2013. Phylogenetic relationship and molecular taxonomy of African grasses of the genus *Panicum* inferred from four chloroplast DNA-barcodes and nuclear gene sequences. *J. Plant. Res*. 126: 363-371.

How to cite this article:

Deepak, K.V., V.A. Mohanlal, J. Johnny Subakar Ivin and Anandan, R. 2018. DNA Barcoding of *Psoralea corylfolia*, *Mucuna pruriens* and *Clitoria ternatea* for Species Identification. *Int.J.Curr.Microbiol.App.Sci.* 7(12): 117-124. doi: https://doi.org/10.20546/ijcmas.2018.712.015