Available online at www.scholarsresearchlibrary.com

Scholars Research Library

Der Pharmacia Lettre, 2016, 8 (4):304-309
(http://scholarsresearchlibrary.com/archive.html)

ISSN 0975-5071
USA CODEN: DPLEB4

DNA barcoding for molecular characterization of a medicinal plant – *Citrullus colocynthis*

Helen Mary L.[1], Florida Tilton[2], Jerrine Joseph*[3], Sudarsanam D.[1], Anitha C. P.[4], Suresh A.[3], Rajasekar T.[3], Ravisankar N.[3] and M. Muthuswamy[5]

[1]*Department of Advanced Zoology and Biotechnology, Loyola College, Chennai - 600034,*
[2]*Biozone Research Technologies Pvt. Ltd., Chennai*
[3]*Centre for Drug Discovery and Development, Sathyabama University, Chennai-600119,*
[4]*NSS College KR Puram, Cherthala, Kerala 688556.*
[5]*Thiruvallur University College of Arts and Science, Kallakuruchi-606202*

ABSTRACT

The aim of the investigation is to achieve molecular characterization of the medicinal plant Citrullus colocynthis through short orthologous DNA marker. The molecular characterization of Citrullus colocynthis is performed by following standard protocols for DNA barcoding specific to Angiogerms based on trnH-psbA intergenic spacer region. The method initially begins with plant sample collection followed by Doyle and Doyle method of DNA isolation, PCR amplification and DNA sequencing. The short sequence obtained was also analyzed insilico using Blastn and Portable software for Molecular Evolutionary Genetics Analysis (MEGA version 6). The DNA sequence of Citrullus colocynthis trnH-psbA intergenic spacer region was 181 base pair long. The Blastn sequence analysis found an array of 13 homologous eudicot species. The evolutionary divergence analysis based on maximumlikelihood (ML) phylogenetic tree and pairwise distance Matrix revealed closely related species and Ancestral data.

Keywords: *Citrullus colocynthis*, DNA barcoding, Medicinal Plant.

INTRODUCTION

In 2003, Paul Hebert's research group published a finding called Biological identification through DNA Barcodes", that drew the attention of many taxonomists, geneticists and evolutionary biologists. DNA barcoding [1] is a technique that performs molecular level characterization of a new or unknown species using its sample. The collected sample of the species when subjected to standard protocols of DNA barcoding [2] obtains a short DNA sequence from the barcode region of the genome. This serves as a basic molecular data to understand evolutionary and Phylogenic relationship of the species [3].

Citrullus colocynthis belonging to the family Cucurbitaceae which belongs to major group of Angiosperms (Flowering plants) [4). Its morphological data is shown in figure 1 and 2. The Species of Cucurbitaceae contains 126 plant genera. The Cucurbitaceae are mostly prostrate or climbing herbaceous annuals comprising about 90 genera and 700 species that are further characterized by commonly having 5-angled stems and coiled tendrils. The leaves are alternate and usually palmately 5-lobed or divided; stipules are absent. The flowers are actinomorphic and nearly always unisexual.

Figure 1: It show the image of *Citrullus colocynthis*

Figure 2: cross section of *Citrullus colocynthis*

Even though the phenotype recognition of species are clearly understood and listed, its genotypic relationship is not identified at molecular level so far. Hence this sequence based identification system called DNA barcoding was used.

The Consortium for the Barcode of Life (CBOL) plant working group proposed that the *trnH-psbA*region remains the leading candidate as a significant marker for DNA Barcoding in plant species [5]. The chloroplast *trnH-psbA*spacer has been proposed as a barcode for plants, either alone or in conjunction with other sequences [6] *trnH-psbA*intergenic spacer is the most variable among angiosperm chloroplast genome [7]. These markers were proposed for DNA barcodes because of their potential for identity of putative regulatory elements, as they harbor sufficient sequence diversity, individually or in combination, to distinguish among species level phylogenetics [8]. The *trnH-psbA* spacer, although short (≈450-bp), is the most variable plastid region in angiosperms [9] and it is easily amplified across a broad range of land plants. Hence DNA barcode *trnH-psbA* intergenic spacer region is used as the DNA Marker to study Citrullus colocynthis [10].

MATERIALS AND METHODS

Plant collection

Healthy, ailment free leaves of *Citrullus colocynthis* were collected from Tirunelveli area of Kanyakumari District of Tamil Nadu, sponge air-dried leaves was then powdered using liquid nitrogen and processed for DNA isolation.

Isolation of Genomic DNA

Plant genomic DNA was extracted using the CTAB based DNA isolation method [11]. To 0.5g of pounded plant sample, preheated 1.5 ml of CTAB buffer (2% hexadecyl trimethyl ammonium bromide (CTAB), 1.4M NaCl, 0.2% mercaptoethanol, 20mM EDTA, 100mM Tris-HCl, pH8.0) were added. After an incubation at 60°C for 30mins, equal volume of Chloroform:isoamyl alcohol (24:1) were added and centrifuged at 6000g for 10mins. To the supernatant 2/3 volumes of cold isopropanol was mixed gently to precipitate nucleic acids. Nucleic pellet was washed with 70% ethanol and resuspended in TE buffer (10mM TrisHCl, 1mM EDTA, pH 7.4).

The DNA sample was electrophoretically separated using Agarose gel Electrophoresis. In the procedure Agarose was weighed and transferred to a conical flask along with 50 ml of 1X TAE and was melted to a clear solution by heating. It was allowed to cool and then 2.5µl of ethidium bromide stock solution was added. A gel casting tray was placed in a leveling table and the melted agarose was poured. After the gel gets solidified, the comb was taken out carefully and the casted gel was placed in an electrophoresis tank and 1X TAE buffer was added until the gel was completely submerged. The DNA sample was mixed with the gel loading buffer and loaded into the well. The samples were then electrophoresed at 50V until the gel loading buffer reached 2/3rd of the gel. Finally this gel containing the DNA was then viewed under UV Trans-illuminator.

The nucleic acid was quantitatively and qualitatively determined using spectrophotometric method. A solution of nucleic acids strongly absorbs UV with an absorbance maximum of 260nm and proteins at 280nm which is linearly related with the concentration of DNA and RNA and the amount of contamination in the solution. At wavelength of 260nm and 280nm the instrument is set at zero absorbance with T.E buffer or sterile water as blank. 5 or 7µl of the sample is taken in a quartz cuvette and made up to 3ml with TE buffer or sterile water and the Absorbance of the solution with the sample was read. The concentration of DNA in the sample was calculated using the given formula:

Concentration of dsDNA = A_{260} x 50µg x dilution factor

Purity of the DNA is signified by the ratio; A_{260} : A_{280} ratio = A_{260} / $A_{280.}$ The ratio of 1.8 signifies: pure DNA, 1.7 – 1.9 signifies: fairly pure DNA(acceptable ratio for PCR), less than 1.8: presence of proteins and greater than 1.8: presence of organic solvent. Internal region of 250bp of *trnH-psbA*intergenic spacer was amplified using PCR technique. Amplification was carried out in a 20µl reaction set up containing 0.3µM of each primer, 0.2mM deoxynucleotide triphosphates (dATP, dCTP, dGTP and dTTP), 100ng of template DNA sample and 1U of Prime TaqDNA polymerase. The reaction tubes were subjected to thermal cycling reactions consisting of an initial denaturation (5 min at 94°C) followed by 32 cycles of denaturation (1 min at 94°C), annealing (46 s at 47°C for *matK*gene) (46 s at 55°C for *trnH-psbA*intergenic spacer), and extension (1 min at 72°C), with a final extension (10 min at 72°C). The PCR product was purified (QIAquick PCR purification kit, Qiagen, Madrid, Spain) and DNA sequencing was performed by Sanger‟s method.

Sequence Analysis

The nucleotide sequence for *trnH-psbA*integenic spacer obtained was subjected to sequence alignment using BLAST [12] tool. The Number of hits with homologous sequences is inferred based on Similarity and alignment.

Phylogenetic tree

Pairwise base identities actually represents multiple mutations and that this proportion increases with increasing overall sequence divergence. Pairwise distance is calculated using maximum-likelihood estimators of substitution rates. Portable software Molecular Evolutionary Genetics Analysis (MEGA) version 6, was used to construct maximum likelihood (ML) tree for the obtained sequences to identify it's inter and intra species relationships. The Distance Matrix Explorer, an action menu of MEGA5 was also used to compute the pairwise difference between the obtained target sequences to its maximum aligned sequences.

Felsenstein's maximum-likelihood approach for inferring phylogeny from DNA sequences assumes that the rate of nucleotide substitution is constant over different nucleotide sites [13]. To infer a tree from genetic data using likelihood principle a minimize change conditional on a specific mutation model is created. The mutation model can take into account that not all substitution events are observed, because recent events might hide ancient events. Maximum Likelihood [14] is therefore undercounting the number of changes and so might have a shorter tree than the true tree. An alternative to likelihood is an approach based on evolutionary distances between a pair of sequences using pairwise distance method. Pairwise methods evaluate all pairs of sequences and transform the differences into a distance. This essentially is a data reduction from a possibly many state difference to a single number. Combining these distances to estimate a tree must be less powerful than the full likelihood approach. In addition, an identical distance can be generated from different sequence pairs. Distance methods have still their merit because once the distance matrix is calculated the tree building can be very fast and under many circumstances are the trees generated with such methods often are identical to the likelihood tree.

RESULTS

The results of PCR amplicons of trnH-psbAintergenic spacer region in *Citrullus colocynthis* is shown in figure 3. The Sequence based analysis result obtained through MEGA sotware is shown in Figure 4,5 and 6

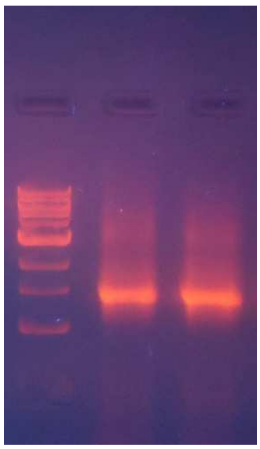

Figure 3: It shows the PCR Amplicons of trnH-psb A interenic spacer region in *Citrullus colocynthis*

[Lane 1: 1Kb Ladder (10,000 bp, 8000 bp, 6000 bp, 5000 bp, 4000 bp, 3000 bp, 2500 bp, 2000 bp, 1500 bp, 1000 bp, 750 bp, 500 bp, 250 bp. Lane 2 and 3 : PCRAmplicon of trnH-psbA gene]

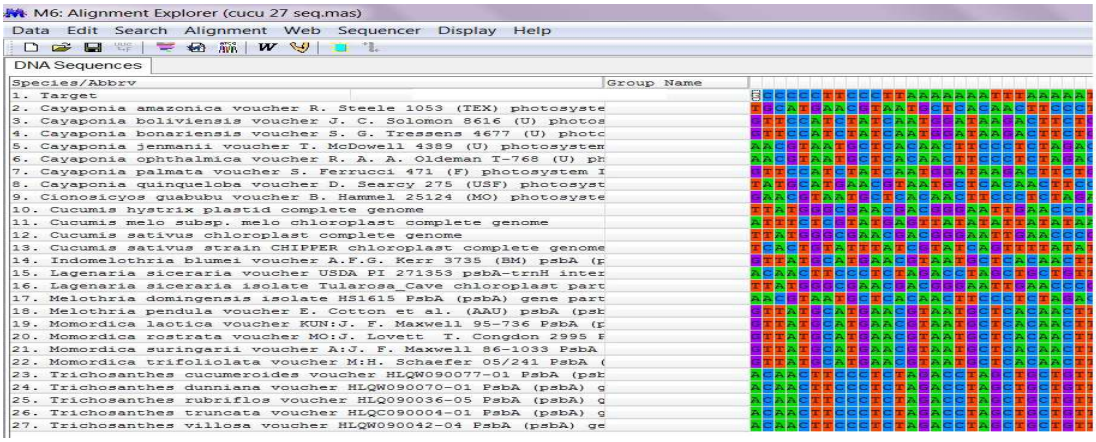

Figure 4: The figure shows the Blast hits and alignment of trnH-psbA spacer gene of *Citrulluscolocynthis* with its homologues in MEGA software

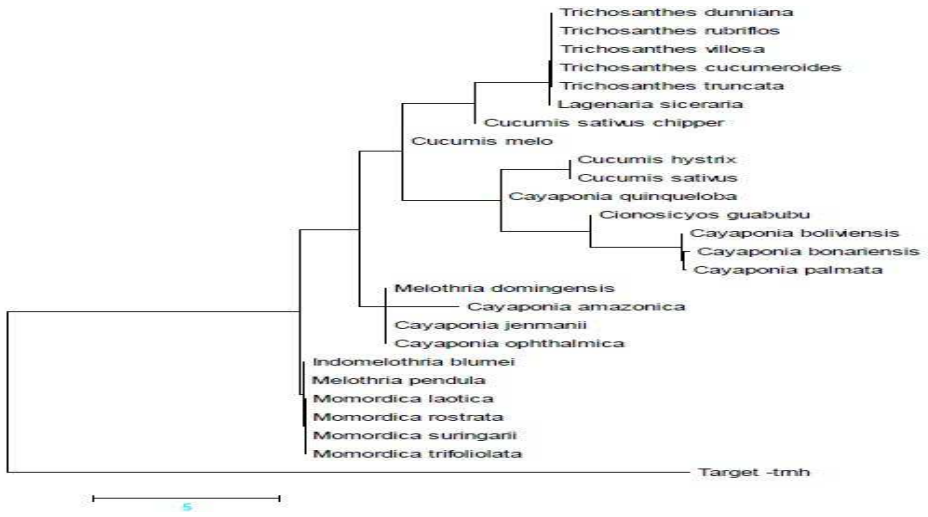

Figure 6: Molecular evolutionary genetics of target *Citrulluscolocynthis*: trnH-psbAintergenic spacer region with sequence producing significant alignment from the database spacer region and significant homologues from nr database

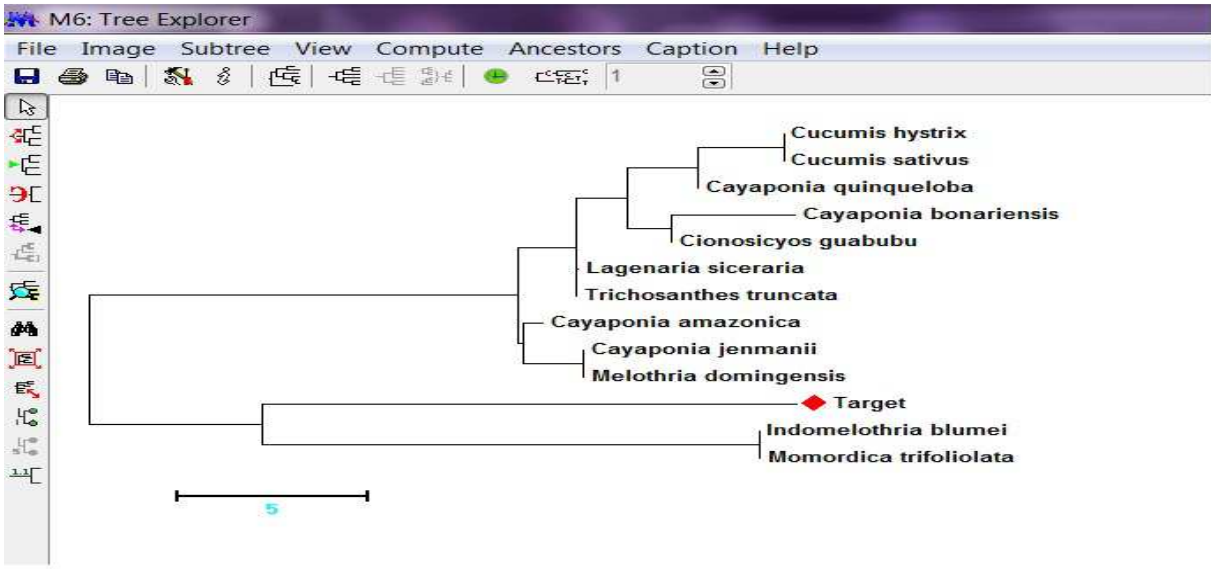

Figure 5: Maximum Likelihood tree of target *Citrullus colocynthis*: trnH-psbAintergenic

File Display Average Caption Help

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Target -trnh													
2. Cayaponia amazonica	4.516												
3. Cayaponia bonariensis	3.503	5.856											
4. Cayaponia jenmanii	2.919	2.752	5.715										
5. Cayaponia quinqueloba	4.277	2.932	2.855	3.288									
6. Cionosicyos guabubu	3.327	4.640	3.014	2.482	3.594								
7. Cucumis hystrix	3.070	3.726	3.071	4.339	2.387	3.476							
8. Cucumis sativus	3.070	3.726	3.071	4.339	2.387	3.476	0.000						
9. Indomelothria blumei	3.127	3.054	3.449	3.801	2.932	4.635	3.480	3.480					
10. Lagenaria siceraria	3.240	3.600	3.923	4.098	3.977	5.410	5.797	5.797	4.205				
11. Melothria domingensis	2.827	2.909	5.675	0.017	3.198	2.269	4.429	4.429	3.842	4.054			
12. Momordica trifoliolata	3.023	3.096	3.440	4.994	3.419	4.493	3.415	3.415	0.035	4.245	5.062		
13. Trichosanthes truncata	2.891	3.448	3.829	4.023	4.055	5.534	5.983	5.983	4.256	0.046	3.977	4.356	

Figure 7: Estimates of evolutionary divergence between Target trnH-psbA spacer region and its 13 homologues

DISCUSSION

The main reason for DNA Barcoding in this species is to gain Biological identification, determine evolutionary and phylogenetic relationship and finally to preserve its molecular data in an electronic library of barcodes. The results of Genomic DNA isolated by Doyle and Doyle method was isolated from the collected plant samples and the Amplicons obtained after PCR was 250bp for *trnH-psbA intergenic* spacer. The quantitative determination using spectrophotometric method found the purity of the DNA to be 1.81 and the DNA sequence was obtained by PCR amplification and sequencing.

The sequence was submitted to NCBI Genbank and the accession number generated was KJ572577. The input target DNA sequence of trnH-psbAintergenic spacer of *Citrullus colocynthis* obtained was used for MEGA software. The BLASTresult revealed 26 major hits for Eudicot with significant E-values. Thus *Citrullus colocynthis is homologous to a* wide spectrum of *Eudicots* like *Momordica* Sp, *Hamamelis* Sp, *Sycopsis* Sp, *Trichocladus* Sp, *Cayaponia* Sp, *Cucumic* Sp, Citrullus Sp, *Indomelothria* Sp, *Melothria* Sp, *Trichosanthus* Sp, *Lagenaria* Sp, *Benincasa* Sp and *Cionosicyos* Sp.

Exemplifies the Maximum likelihood (ML) tree constructed for 26 hits and the target *Citrullus colocynthis* (*trnH-psbA*intergenic spacer) using MEGA5 software[15). Branch lengths and substitution rate parameters are then optimized for each model to fit the data. Species having identical branch length are filtered and a set of 13 species are selected for phylogenetic inference. From the ML phylogenetic tree of 13 selected species, it is inferred that the recent ancestor is common for clade containing Target *Citrullus colocynthis*, *Momordica trifoliolata*. The distant recent ancestor also has a clade containing *Cayaponia jenmanii*, *Melothria domingensis*, *Cayaponia amazonica*, *Trichosanthes truncata*, *Lagenaria siceraria*, *Cayaponia bonariensis*, *Cayaponia quinqueloba*, *Cucumis sativus* and *Cucumis hystrix*. Overall the phylogenetic tree reveals that the two recent ancestors appear to have diverged from a single ancestor.

Represent the pairwise Distance Matrix for the 13 nucleotide sequences obtained from ML tree data using MEGA6. This grid displays the pairwise distances between taxa or within groups in the form of a lower or upper triangular matrix. The analysis involved 13 nucleotide sequences obtained from ML tree data. Codon positions included were: 1st+2nd+3rd+Noncoding and all positions containing gaps and missing data were eliminated.From the matrix it is inferred that the Target *Citrullus colocynthis was found to* have less variation with an Evolutionary divergence of 2.827 and 2.891 with *Melothria domingensis* and *Trichosanthes* truncata respectively. And it has *large variation with* Cayaponi quinqueloba and *Cayaponia amazonica* with an evolutionary divergence of 4.277 and 4.516 respectively.

Melothria domingensis (Cucurbitaceae), is an endangered Caribbean endemic, is a Cayaponia. It is a small group of yellow- or white-flowered climbers with small to medium-sized fruits [16]. The species medicinal pharmacological property includes Anti inflammatory (Anti-inflammatory activity of flavonoids from *Cayaponiatayuya* roots [17]. The Tricosanthus species has effective abortifacient, immunosuppressive and anti-tumor and anti HIV medicinal value [18]. Trichosanthin, a ribosome inactivating protein having Anti HIV property is derived from Trichosanthus Species [19]

CONCLUSION

DNA barcoding has enhanced evolutionary investigations using right genetic markers for the desired organism. In this study, molecular characterization of the collected plant was carefully performed and its evolutionary relationship was achieved. The results verified using advanced bioinformatics tool has demonstrated the functional properties of homologous plants evolved during divergence. Disclosure of evolutionary divergence of target and different plant species with *trnH-psbA*intergenic spacer has lead to indispensable knowledge of biodiversity. The plant is hence believed to possess Antiinflammatory, Anti HIV and other thereupeutic properties which can be further studied to achieve medicinal criterion of the plant

REFERENCES

[1] P.D. Hebert, A. Cywinska, S.L. Ball, J.R. deWaard, *Proc Biol Sci*. **2003**, 270(1512), 313-21.
[2] M.W. Chase, R.S. Cowan, P.M. Hollingsworth, C. van den Berg, S. Madrinan *et al*.. **2007**, *Taxon*, 56,295–299.
[3] D. Andreas, B.F. Baxevanis Francis Ouellette, 2001. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Second Edition.p1-450.
[4] H. Schaefer, S.S. Renner, *Taxon*, **2011**, 60 (1), 122–138.
[5] W.J. Kress, K.J. Wurdack, E.A. Zimmer, L.A. Weigt, D.H. Janzen, *PNAS*, **2005**, 102(23), 8369–8374.
[6] A. Barbara, M. Whitlock, M.H. Amanda, A.G. Paul, *PLoS ONE*, **2010**, 5(7), 11533.
[7] D.H. Kress, *PNAS*, **2009**, 106(31),12794 –12797
[8] C. Sass, Little, D.W .Stevenson, C.D. Specht, *PLos ONE*, **2007,** 11, e1154.
[9] H. Storchova, M.S. Olson, *Plant Systematics and Evolution*, **2007**, 268(1-4), p235-236.
[10] C.S. Ford *et al*.. *Bot J Linn Soc*, **2009**, 159, 1–11.
[11] J.J. Doyle, J.L. Doyle, *Phytochem*. Bull, **1987**, 19, 11-15.
[12] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, *Nucleic Acids Res*, **1997**, 25, 3389-3402.
[13] J. Felsenstein, *J. Mol. Evol*. **1981**,17,368-376.
[14] Z. Yang, Mol Biol Evol, **1993**,10(6),1396-401.
[15] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, *Molecular Biology and Evolution*, **2011**, 28(10), 2731-9.
[16] H. Schaefer, M.H. Nee, *Phyto Keys*, **2012,** 18, 45–60.
[17] S. Aquila, R.M Giner, M.C. Recio, E.D. Spegazzini, J.L. Rios, *Journal of Ethnopharmacology*, **2009,** 121(2), 333–337.
[18] M.X. Li, H.W. Yeung; L.P. Pan, S.I. Chan, *Nucleic Acids Res*, 25, **1991**, 19(22), 6309–6312.
[19] P.C. Shaw, K.M. Lee, K.B. Wong, *Toxicon*, **2005**, 45 (6), 683-9.