Suitability of ITS2, *nad1* and *ycf1b* as DNA barcodes for the Ten Medicinal Plants of the Philippines

Levi Letlet H. Larcia II[1], Joseph Christian M. Manzano[1], Kyle Maleen O. Sagulili[2], Jerica Margarita G. Ibanez[2], Ciara Christianne Y. Lim[2], Joanne Marie M. Del Rosario[2], Paul Benedic U. Salvador[1], Laarni G. Corales[1] and Leslie Michelle M. Dalmacio[1]

[1]Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila

[2]Department of Physical Sciences and Mathematics, College of Arts and Sciences, University of the Philippines Manila

ABSTRACT

The Department of Health had listed ten herbal medicines commonly used in the Philippines for different ailments referred to as "Sampung Halamang Gamot" or 10HG[1]. Ensuring proper identification of these herbal medicines is important to prevent numerous adverse reactions. DNA barcoding, a taxonomic method that uses a short genetic marker in an organism's DNA to identify it as belonging to a particular species[2], can be used to augment the current methods of identification. ITS2, *nad1* and *ycf1b* are three of the most commonly used genetic loci for DNA barcoding[3,4]. This study determined the suitability of these genetic markers as DNA barcodes for the identification of 10HG by comparing their PCR success rate, sequence quality, and discriminatory power. For PCR success rate, the 10 HG gave 100%, 80% and 40% amplification using ITS2, nad1, and *ycf1b*, respectively. With regard to the sequence quality, high quality sequences were given by *A. sativum, P.guajava, Q. indica,* and *E.microphylla* using ITS2 while high quality sequences were given by *C. alata*, *M. charantia*, *P. guajava*, and *V. negundo* using *ycf1b.* Finally, for the discriminatory power, six out of the 10 HG gave species level discrimination using ITS2 and two out of the ten for *ycf1b. Nad1* data for sequence quality and discriminatory power was inconclusive. From the data gathered, it can be inferred that ITS2 is the most suitable DNA barcode for the 10HG. The utilization of ITS2 as a standard DNA barcode for 10HG is supported by the study of Chen et al.[5] wherein they concluded that the ITS2 region can be potentially used as a standard DNA barcode to identify medicinal plants and their closely related species. For the suitability of *nad1* and *ycf1b* as 10HG DNA barcodes, however, further studies are recommended to make a conclusion.

INTRODUCTION

The ten herbal medicines commonly used in the Philippines for different ailments include *akapulko (Cassia alata)* for treating ringworm and other fungal infection*, ampalaya (Momordica charantia)* for the treatment of non-insulin dependent diabetes mellitus*, bawang (Allium sativum)* for reducing blood cholesterol*, bayabas (Psidium guajava)* for use as antiseptic*, lagundi (Vitex negundo)* for the relief of asthma and cough*, niyog-niyogan (Quisqualis indica L.)* for use as anthelmintic *, sambong (Blumea balsamifera)* for use as diuretic*, tsaang gubat (Ehretia microphylla* Lam.*)* for the treatment of colic and diarrhea*, ulasimang bato (Peperomia pellucida)* for the treatment of gouty arthritis and *yerba buena (Clinopodium douglasii)* for use as analgesic[1,6,7] (Principe and Jose, 2002; Philippine Herbal Medicine, 2019; DOH-BFAD, 2005). DNA barcoding, a taxonomic method that uses a short genetic marker in an organism's DNA to identify it as belonging to a particular species.[2] may be used to augment the current methods of identifying plant materials that merely relies on the plant's physical characteristics. ITS2, *nad1* and *ycf1b* are three of the most commonly used genetic loci for DNA barcoding.[3,8] This study determined the suitability of ITS2, *nad1* and *ycf1b* as DNA barcodes for the identification of 10HG by comparing the PCR success rate, sequence quality, and resolution power of these markers.

METHODOLOGY

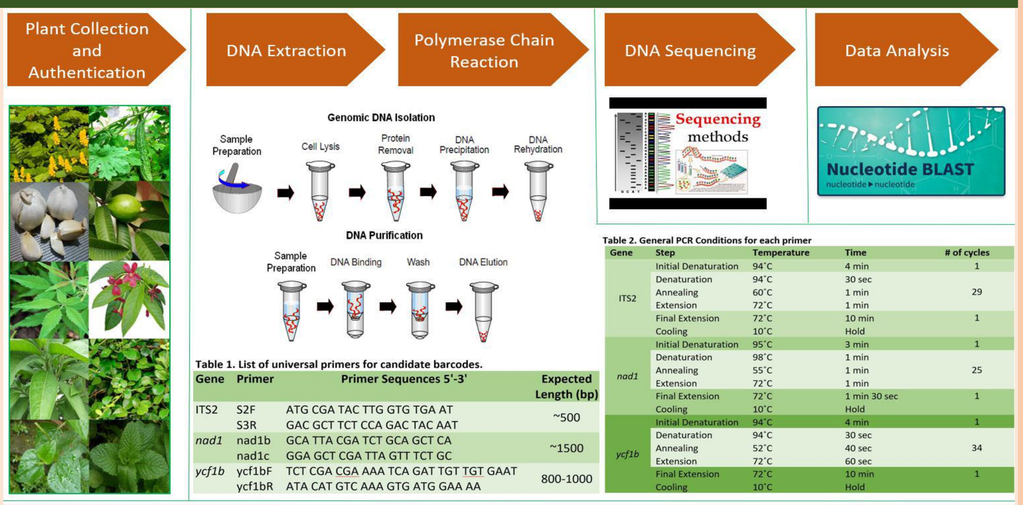

Table 1. List of universal primers for candidate barcodes.

Gene	Primer	Primer Sequences 5'-3'	Expected Length (bp)
ITS2	S2F	ATG CGA TAC TTG GTG TGA AT	~500
	S3R	GAC GCT TCT CCA GAC TAC AAT	
nad1	nad1b	GCA TTA CGA TCT GCA GCT CA	~1500
	nad1c	GGA GCT CGA TTA GTT TCT GC	
ycf1b	ycf1bF	TCT CGA CGA AAA TCA GAT TGT TGT GAAT	800-1000
	ycf1bR	ATA CAT GTC AAA GTG ATG GAA AA	

Table 2. General PCR Conditions for each primer

Gene	Step	Temperature	Time	# of cycles
ITS2	Initial Denaturation	94°C	4 min	1
	Denaturation	94°C	30 sec	29
	Annealing	60°C	1 min	
	Extension	72°C	1 min	
	Final Extension	72°C	10 min	1
	Cooling	10°C	Hold	
nad1	Initial Denaturation	95°C	3 min	1
	Denaturation	98°C	1 min	25
	Annealing	55°C	1 min	
	Extension	72°C	1 min	
	Final Extension	72°C	1 min 30 sec	1
	Cooling	10°C	Hold	
ycf1b	Initial Denaturation	94°C	4 min	1
	Denaturation	94°C	30 sec	34
	Annealing	52°C	40 sec	
	Extension	72°C	60 sec	
	Final Extension	72°C	10 min	1
	Cooling	10°C	Hold	

RESULTS AND DISCUSSION

Figure 1. Representative plant DNA Isolates from 10HG on agarose gel.

Table 3. Summary of PCR Amplification Success rate for each primer used

Plant	PCR Amplification Success Rate		
	ITS2	*nad1*	*ycf1b*
Cassia alata	✓	✓	✓
Momordica charantia	✓	×	✓
Allium sativum	✓	✓	×
Psidium guajava	✓	✓	✓
Vitex negundo	✓	✓	✓
Quisqualis indica	✓	×	×
Blumea balsamifera	✓	✓	×
Ehretia microphylla Lam.	✓	✓	×
Peperomia pellucida	✓	✓	×
Clinopodium douglasii	✓	✓	×

✓indicates successful PCR amplification
×indicates failed PCR amplification

Table 4. Summary of 10HG Sequence Quality based on QV Scores

Plant	Sequence Quality		
	ITS2	*nad1*	*ycf1b*
Cassia alata	*	*	✓
Momordica charantia	×	*	✓
Allium sativum	✓	*	×
Psidium guajava	✓	*	✓
Vitex negundo	×	*	✓
Quisqualis indica	✓	*	*
Blumea balsamifera	*	*	*
Ehretia microphylla Lam.	✓	*	*
Peperomia pellucida	*	*	×
Clinopodium douglasii	×	*	*

✓ indicates high quality sequences
× indicates low quality sequences
* indicates absence of data.

Table 5. Summary of species-level discriminatory power of candidate DNA barcodes for the 10HG

Plant	Discriminatory Power		
	ITS2	*nad1*	*ycf1b*
Cassia alata	✓	*	×
Momordica charantia	✓	*	×
Allium sativum	✓	*	*
Psidium guajava	✓	*	✓
Vitex negundo	×	*	✓
Quisqualis indica	✓	*	*
Blumea balsamifera	×	*	*
Ehretia microphylla Lam.	×	*	*
Peperomia pellucida	✓	*	*
Clinopodium douglasii	×	*	*

✓ indicates with species level discrimatory power
× indicates without species level discriminatory power
* indicates absence of data

DNA Extraction was successful for all the 10HG samples. Figure 1 shows representative bands greater than 10,000 bp indicating the presence of genomic DNA.

Table 3 summarizes the data on the success rate of PCR Amplification for each primer. From the data, it can be seen that the 10 HG were all successfully amplified using ITS2. On the other hand, eight plants for *nad1* and 4 plants for *ycf1b* were successfully amplified. The ITS2 region is short [9] resulting to its easy amplification as compared to *nad1* and *ycf1b* which are composed of approximately 1500 and 1000 base pairs respectively. In addition, a combination of a conserved secondary structure with a variable sequence appears to be a major benefit of utilizing ITS2[9]. *Cassia alata, Psidium guajava*, and *Vitex negundo* were the plants found to be easily amplified by the 3 candidate barcodes.

The sequence quality produced by the 10HG for each primer is tabulated (Table 4). For ITS2, high quality sequences were given by *A. sativum, P.guajava, Q. indica, and E.microphylla.* For *ycf1b,* high quality sequences were given by *C. alata*, *M. charantia*, *P. guajava*, and *V. negundo.* For *nad1*, QV Score data was not available. The sequence quality was based on QV Score. For each sample sequence, the number of bases with the QV Score ≥20 was observed. A QV Score of greater than or equal to 20 means that the chance of a base miscall is 1 in 100. Only ITS2 and ycf1b have data on QV Score.

Based on the summary given in Table 5, ITS2 is the most promising candidate DNA barcode for the 10HG. Six out of the 10 HG gave species level discrimination using ITS2 and two out of the ten for *ycf1b.* No data for discriminatory power can be attributed to *nad1* due to absence of sequencing data. The species-level discriminatory power of candidate barcodes was observed by generating phylogenetic trees using the query sequence and the top 5 hits using BLAST. If the sample sequence was able to cluster with the similar species and did not cluster with the different species, then it can be said that the candidate DNA barcode was able to discriminate between species of the same genus. ITS2 was able to discriminate species of *C. alata*, *M. charantia*, *A. sativum*, *P. guajava*, and *P.pellucida* while *ycf1b* was able to discriminate species of *P. guajava* and *V. negundo*.

CONCLUSION

From the data on PCR amplification success rate, sequence quality and discriminatory power it can be inferred that ITS2 is the most suitable DNA barcode for the 10HG. The utilization of ITS2 as a standard DNA barcode for 10HG is supported by the study of Chen et al. (2010) wherein they concluded that the ITS2 region can be potentially used as a standard DNA barcode to identify medicinal plants and their closely related species. For the suitability of *nad1* and *ycf1b* as 10HG DNA barcodes, however, the data gathered is not enough to make a conclusion.

REFERENCES

[1] Philippine Herbal Medicine. 2019. http://www.philippineherbalmedicine.org/doh_herbs.htm

[2] CBOL, n.d. Consortium for the Barcode of Life, n.d., 'DNA Barcoding' *Consortium for the Barcode of Life* [Online] Available from <http://barcoding.si.edu/DNABarCoding.htm

[3] Yao H, Song J, Liu C, Luo K, Han J, Li Y, et al. (2010) Use of ITS2 Region as the Universal DNA Barcode for Plants and Animals. PLoS ONE 5(10): e13102. https://doi.org/10.1371/journal.pone.0013102)

[4] Dong, Wenpan, Chao Xu, Changhao Li, Jiahui Sun, Yunjuan Zuo, Shuo Shi, Tao Cheng, Junjie Guo & Shiliang Zhou. 2015. *ycf1*, the most promising plastid DNA barcode of land plants. Scientific Reports 5 :8348. DOI:10.1038/srep08348.

[5] 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), e8613. https://doi.org/10.1371/journal.pone.0008613

[6]Principe, E.B. and Jose, A.S. 2002. Propagation Management Of Herbal and Medicinal Plants. RESEARCH INFORMATION SERIES ON ECOSYSTEMS Volume 14 No. 2 May – August 2002

[7]DOH-BFAD. (2005) Philippine Pharmacopeia I with Supplement. Muntinlupa City: Department of Health, Bureau of Food and Drugs

[8] Lin, CP and Danforth, BN. 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Mol Phylogenet Evol. 30(3):686-702.

[9]Han, Jianping, Yingjie Zhu, Xiaochen Chen, Baoshen Liao, Hui Yao, Jingyuan Song, Shilin Chen, and Fanyun Meng. (2013). The Short ITS2 Sequence Serves as an Efficient Taxonomic Sequence Tag in Comparison with the Full-Length ITS. Biomed Res Int. 2013; 2013: 741476. doi: 10.1155/2013/741476. PMID: 23484151.