INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
20–1135/2021/25–1–227–230
DOI: 10.17957/IJAB/15.1660
http://www.fspublishers.org

Short Communication

Molecular Authentication of Commercially Cultivated Coffee (*Coffea* spp.) in the Philippines using DNA Barcodes

Arleen C. Panaligan[1,2*], Miriam D. Baltazar[2,3] and Grecebio Jonathan D. Alejandro[1,4]
[1]*The Graduate School, University of Santo Tomas, España, 1015 Manila, Philippines*
[2]*Department of Biological Sciences, Cavite State University, Indang, 4122 Cavite, Philippines*
[3]*National Coffee Research, Development and Extension Center, Cavite State University, Indang, 4122 Cavite, Philippines*
[4]*College of Science and Research Centre for the Natural and Applied Sciences, University of Santo Tomas, España, 1015 Manila, Philippines*
[*]For correspondence: acpanaligan@cvsu.edu.ph
Received 08 August 2020; Accepted 19 September 2020; Published 10 December 2020

Abstract

Accurate identification of commercially cultivated coffee species is necessary since the cup quality may be attributed to the kind of species used. Hence, DNA barcoding was performed using nuclear ribosomal *internal transcribe spacer* (*ITS*) and *maturase K* (*matK*). Both markers had 100% amplification and sequencing success rates. Although *ITS* had lower resolution in *Coffea* species, it efficiently discriminated *Coffea liberica.* The *matK* barcode discriminated all the species. Findings revealed that *matK* is an efficient barcode over *ITS* for commercially cultivated *Coffea* species by generating the highest rate of both universality and discrimination power. DNA barcoding as a method of authentication will benefit the coffee industry and coffee growers for large-scale plantations. © 2021 Friends Science Publishers

Keywords: *Coffea* spp.; DNA barcode; *ITS; matK*

Introduction

Coffee (*Coffea* L.) belongs to the Rubiaceae family, comprising of 124 species (Davis *et al*. 2019). Among these species, *C. arabica* L., *C. canephora* Pierre ex A. Froehner and *C. liberica* W. Bull ex Hiern are commercially cultivated in the Philippines. *C. liberica* has two known varieties, namely, excelsa and liberica. Generally, the fruits of var. liberica are bigger, more tapered at the base, have thicker and leatherier pericarp than that of var. excelsa (Bridson 1988).

Planting materials such as seedlings need accurate identification since the market value of coffee depends on the cup quality which may be attributed to the species. Consumers tend to choose *C. arabica* (locally known as Arabica) because of its rich aroma. *C. canephora* (locally known as Robusta) is more bitter and has higher caffeine content than *C. arabica* (Lecolier *et al.* 2009), while *C. liberica* (locally known as Kapeng Barako) has a strong flavor.

Morphology-based identification is the usual method of identifying plants including coffee. However, it is difficult to discriminate *Coffea* species at the seedling stage. Generally, distinct characteristics of *Coffea* species can be observed at maturity but they still possess overlapping characters (Panaligan *et al*. 2020). Another technique that can be used for species identification is the utilization of short DNA sequences as barcodes. This technique is not dependent on environment and life stages (Hebert *et al*. 2003). The *matK* marker was used in this study because of its greater plant species discrimination (CBOL Plant Working Group 2009). The nuclear ribosomal *internal transcribe spacer* (*ITS*) was added to *matK* as suggested by the China Plant BOL Group (2011). The *ITS* and *matK* were used to assess possible barcodes in authenticating commercially cultivated *Coffea* species in the Philippines. Hence, this study aims to evaluate the PCR success rate, sequencing success rate, and discriminatory power of *ITS* and *matK*.

Materials and Methods

Sample collection and amplification of the DNA barcodes

Young coffee leaf samples were collected and stored in bags with silica gel for DNA extraction. Herbarium vouchers were made and deposited at the University of Santo Tomas Herbarium (USTH).

To cite this paper: Panaligan AC, MD Baltazar, GJD Alejandro (2021). Molecular authentication of commercially cultivated coffee (*Coffea* spp.) in the Philippines using DNA barcodes. *Intl J Agric Biol* 25:227–230

DNA samples were extracted following the protocol of Dneasy Plant Minikit (Qiagen, Hilden, Germany). Using universal primer pairs (Table 1), the *ITS* and *matK* regions were amplified with a total volume of 25 μL per reaction. The PCR mixture contained 19.45 μL water, 2.5 μL 10x reaction buffer, 0.5 μL 50 m*M* $MgCl_2$, 0.4 μL 10 m*M* dNTP, 0.5 μL 10 μ*M* forward and reverse primers, 0.15 μL 5u/ µL *Taq* DNA polymerase (Vivantis) and 1.0 μL DNA. The PCR amplification was performed using a T100 Thermal Cycler (Bio-Rad) as follows: initial denaturation at 97°C for 90 s followed by 35 cycles of 95°C for 30 s, 55°C for 20 s (*ITS*) or 50°C for 20 s (*matK*), 72°C for 1 min, followed by final extension at 72°C for 10 min (Li *et al.* 2012). The PCR products were purified using QIA-quick Purification Kit (Qiagen, Germany) and sent to Macrogen Inc., Seoul, South Korea for bidirectional DNA sequencing.

Sequence analyses

Consensus sequences were edited and assembled using Codon Code Aligner v. 4.1.1 (Codon Code Co., Centerville, MA, USA). Multiple Sequence Alignment was performed using MEGA 7 (Kumar *et al*. 2016). Neighbor-Joining (NJ) trees were constructed in MEGA 7 using 1000 bootstrap replicates. Pairwise distances of these markers were computed using Kimura–2–parameter (K2P) model (Kimura 1980) with MEGA 7 software (Kumar *et al*. 2016). Wilcoxon two–sample test was performed using S.P.S.S. software (version 20.0; S.P.S.S. Inc., Chicago, U.S.A.).

Results

A total of 48 sequences were newly generated in this study from *ITS* and *matK* and were deposited in GenBank (Table 2). The *ITS* and *matK* regions were 100% amplified and sequenced (Table S1). Interspecific distances of *ITS* and *matK* were significantly higher than their intraspecific distances ($P < 0.001$) (Table S2). There were two members of *C. canephora* that interclustered with *C. arabica* in the Neighbor-Joining tree of *ITS* (Fig. 1A), while *matK* discriminated all the *Coffea* species (Fig. 1B).

Table 1: Universal primers of the two candidate barcodes

Barcode	Primer	Primer sequence (5'–3')	Reference
ITS	*ITS5*	GGAAGTAAAAGTCGTAACAAGG	White *et al*. (1990)
	ITS4	TCCTCCGCTTATTGATATGC	
matK	*3F_Kim f*	CGTACAGTACTTTTGTGTTTACGAG	CBOL- PWG 2009
	1R_Kim r	ACCCAGTCCATCTGGAAATCTTGGTTC	

ITS = internal transcribe spacer; matK = maturase K

Table 2: List of *Coffea* species used in the study and their accessions

Species	Place of Origin/Collection Place	Code	USTH Accession	GenBank Accession	
				ITS	*matK*
C. arabica	Indang, Cavite	A1	014856	MK611791	MK722268
C. arabica	Indang, Cavite	A3	014857	MK611792	MK722267
C. arabica	Ampasit, Benguet	A31-1	014858	MK615726	MK722270
C. arabica	Ampasit, Benguet	A31-2	014859	MK615727	MK722269
C. arabica	BSU, Benguet	A311-1	014860	MK615728	MK722266
C. arabica	BSU, Benguet	A311-2	014861	MK615729	MK722265
C. arabica	Mascarenes/Nicaragua	–	–	DQ153609	AB973195
C. canephora	Indang, Cavite	C1	014862	MK615730	MK722259
C. canephora	Indang, Cavite	C2	014863	MK615731	MK855097
C. canephora	Alfonso, Cavite	C6	014864	MK615732	MK722261
C. canephora	Alfonso, Cavite	C7	014865	MK615733	MK722260
C. canephora	Indang, Cavite	C8	014866	MK615734	MK722264
C. canephora	Indang, Cavite	C9	014867	MK615735	MK722263
C. canephora	Indang, Cavite	C10	014868	MK615736	MK722262
C. canephora	NOMIARC, Bukidnon	C512-2	014869	MK615737	MK722258
C. canephora	NOMIARC, Bukidnon	C513-2	014870	MK615738	MK855098
C. canephora	Cameroon/Vietnam	–	–	DQ153593	AB973198
C. canephora	Mexico/Indonesia	–	–	MF417755	AB973197
C. liberica	Indang, Cavite	L1	014871	MK615739	MK722250
C. liberica	Indang, Cavite	L3	014872	MK615740	MK722249
C. liberica	Alfonso, Cavite	L6	014873	MK615741	MK722253
C. liberica	Silang, Cavite	L7	014874	MK615742	MK722252
C. liberica	Alfonso, Cavite	L8	014875	MK615743	MK722251
C. liberica	Indang, Cavite	LE1	014876	MK615744	MK722255
C. liberica	Indang, Cavite	LE3	014877	MK615745	MK722254
C. liberica	Alfonso, Cavite	LE7	014878	MK615746	MK722257
C. liberica	Silang, Cavite	LE8	014879	MK615747	MK722256
C. liberica var. dewevrei	Central African Republic	–	–	DQ153603	–
C. liberica var. liberica	Congo/ not indicated	–	–	DQ153610	AM412465

USTH = University of Santo Tomas Herbarium; *ITS = internal transcribe spacer; matK = maturase K*

Fig. 1: Neighbor-Joining tree inferred using Kimura two-parameter distances and 1000 bootstrap replicates of (**A**) *ITS* and (**B**) *matK*. Numbers on nodes are bootstrap support (BS) values. Outgroups were obtained from GenBank. Note: A = *C. arabica*; C = *C. canephora*; L = *C. liberica* (liberica); LE = *C. liberica* (excelsa)

Discussion

The high amplification and sequencing success rates of *ITS* and *matK* indicate that these DNA barcodes are universal for *Coffea* species. The universality of *ITS* and *matK* are congruent with the findings of Huang *et al.* (2019) and Amin *et al.* (2020). Having significantly higher interspecific than intraspecific distances suggests that these two candidate barcodes have the potential to discriminate *Coffea* at the species level. The lower discriminatory power of *ITS* as compared to *matK* has previously been reported by Chen *et al.* (2020). In this study, *matK* had higher discriminatory power than *ITS* (Fig. 1), which is incongruent with the finding of Huang *et al*. (2019). The different results of these

studies imply that the efficiency of a DNA barcode varies in different genera and species. As an example, *ITS* was efficient in discriminating *C. liberica* but was not able to discriminate *C. canephora*. Some species would be better resolved by other DNA regions, as exemplified by *matK*, which discriminated all the *Coffea* species (Fig. 1B).

In the Philippines, there are four known coffee types and these are 'arabica', 'robusta', 'liberica', and 'excelsa' corresponding to *C. arabica*, *C. canephora*, *C. liberica* and *C. excelsa*, respectively. However, Panaligan *et al*. (2020) reported that liberica and excelsa (locally known as Kapeng Barako) belong to a single species that is *C. liberica*. In this study, the NJ trees for *ITS* and *matK* showed that liberica and excelsa grouped in a single clade (Fig. 1), thus supporting the result of the previous study that liberica and excelsa belong to the same species (Panaligan *et al*. 2020).

Conclusion

Although *ITS* had lower discriminatory power in *Coffea*, it efficiently discriminated *C. liberica*. The *matK* region was able to discriminate the three *Coffea* species, indicating that this DNA barcode is efficient for authentication of commercially cultivated coffee at the species level.

Acknowledgements

The authors would like to thank the University of Santo Tomas - RCNAS and Cavite State University for providing the laboratory facilities. The first author is grateful to DOST-PCIEERRD and NMNH, Paris, France for the sponsorship to training; CHED-FDP and DOST-PCAARRD for the financial grant.

Author Contributions

AP conducted the experiment and wrote the manuscript. MB and GJA edited the manuscript.

References

Amin S, S Ghosh, B Biswas, M Arifuzzaman, AK Azad, Z Siddiki (2020). Molecular identification of four medicinal plants using DNA barcoding approach from Chittagong, Bangladesh. *J Adv Biotechnol Exp Ther* 3:268–272

Bridson DM (1988). *Coffea*, pp:706–709. *Flora of Tropical East Africa, Rubiaceae*. Balkema, Rotterdam, The Netherlands

CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proc Natl Acad Sci USA* 6:12974–12797

Chen Q, X Wu, D Zhang (2020). Comparison of the abilities of universal, super, and specific DNA barcodes to discriminate among the original species of *Fritillariae cirrhosae* bulbus and its adulterants. *PLoS One* 15; Article e0229181

China Plant BOL Group (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (*ITS*) should be incorporated into the core barcode for seed plants. *Proc Natl Acad Sci* 108:19641–19646

Davis AP, H Chadburn, J Moat, R O'Sullivan, S Hargreaves, EN Lughadha (2019). High extinction risk for wild coffee species and implications for coffee sector sustainability. *Sci Adv* 5; Article eaav3473

Hebert PD, A Cywinska, S Ball, JRD Waard (2003). Biological identifications through DNA barcodes. *Proc Roy Soc Lond Ser B Biol Sci* 270:313–321

Huang Y, X Chen, J Chen, Z Li (2019). DNA barcoding and phylogenetic relationships in *Omphalogramma* (Primulaceae) from the Hengduan Mountain region of China. *Phytotaxa* 402:97–106

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120

Kumar S, G Stecher, K Tamura (2016). *MEGA 7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. BMC Ecol* 33:1870–1874

Lecolier A, P Besse, A Charrier, TN Tchakaloff, M Noirot (2009). Unraveling the origin of *Coffea arabica* 'Buorbon pointy' from La Reunion: A historical and scientific perspective. *Euphytica* 168:1–11

Li M, K Au, H Lam, L Cheng, R Jiang, P But, P Shaw (2012). Identification of Baiying (Herba Solani Lyrati) commodity and its toxic substitute Xungufeng (Herba Aristolochiae Mollissimae) using DNA barcoding and chemical profiling techniques. *Food Chem* 135:1653–1658

Panaligan AC, MD Baltazar, GJD Alejandro (2020). Genetic polymorphism of registered and popularly cultivated coffee (*Coffea* spp.) in the Philippines using inter-simple sequence repeats markers. *Biodiversitas* 21:4228–4233

White TJ, T Bruns, S Lee, J Taylor (1990). *Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. PCR Protocols: A Guide to Methods and Applications*, pp:315–322. Academic Press, Inc., New York, USA