EISSN: 2707-0425

East African Journal of Science, Technology and Innovation, Vol. 4 (2): March 2023
This article is licensed under a Creative Commons license, Attribution 4.0 International (CC BY NC SA 4.0)

Potential application of DNA barcoding technology for identification of medicinal plant species traded in Nairobi County, Kenya

[1]*MWAURA A, [2]OMWOYO O, [2]KAMAU J

[1]Centre for Biodiversity Department, National Museums of Kenya, P.O Box 40658 – 00100, Nairobi.
[2]Department of Plant Sciences, Kenyatta University, P.O Box: 43844-00100. Nairobi*
***Corresponding author**: mwauran@gmail.com

Abstract

There exists a long history of herbal products use but recently there are concerns on authenticity and safety of some these products. Majority of the herbal products traded in Kenyan markets are processed plant parts, which have been dried or in powder form and are extremely difficult to identify taxonomically. This study was carried out in selected markets in Nairobi County, the capital city of Kenya to assess authenticity of herbal products traded in the region using deoxyribonucleic acid (DNA) barcoding technology by applying internal transcribed spacer (*ITS)* and ribulose -1,5-bisphosphate carboxylase Large subunit (rbcL) markers. One hundred and twelve herbal products were purchased from practicing herbalists and analyzed using *ITS* and rbcL DNA barcoding markers. DNA was extracted from acquired samples, amplified and sequences generated. BLAST search with edited generated sequences against the NCBI database was done to recover the expected species. Our findings showed that 35% of the herbal products were wrongly labeled and contained different species compared to those listed on the labels. Out of 92 samples with positive sequences, 58 matched their expected species after the BLAST search. Among the authentic species identified are *O. lanceolata, M. oleifera, W. ugandensis, O. europaea, Acacia species, and C. spinarum.* However, 32 species did not match the expected species as listed on labels and were concluded to have been substituted. Among them were *R. prinoides, T. abyssinica, U. massaica* and *A. indica,* which were substituted with their close relatives, while *R. tridentate, Z. usambarense* and *Prunus africana* were substituted with unrelated species with no reported medical benefits. *ITS* marker had remarkable identification efficiency at 98% to species level compared to rbcL at 89%. We recommend further studies to establish at what stage substitution takes place, as well as utilization of DNA barcoding technology for medicinal plants authentication and verification.

Keywords: *authentic; DNA barcoding; herbal products; identification; Kenya; markets; traded*

Cite as: Mwaura *et al.,* (2023) Potential application of DNA barcoding technology for identification of medicinal plant species traded in Nairobi County, Kenya. *East African Journal of Science, Technology and Innovation 4(2).* http://doi.org/10.3329/bjar.v42i3.34505

Received: 07/09/2022
Accepted: 14/03/2023
Published:29/03/2023

Introduction

Trade in medicinal plant products is a flourishing commercial enterprise globally due to increasing demand of traditional therapy, increase in costs for healthcare and inaccessibility to health facilities especially in rural Africa (Newmaster *et al.,* 2013). It has been documented that at least 65% of people from developing countries use botanical herbal extracts and remedies for various disease treatments. In some countries such as Uganda and Ethiopia, the usage is as high as 85% (Mahomoodally, 2013; Tugume and Nyakoojo, 2019), whereas in Kenya its estimated to be over

70 % (Mbuni *et al.,* 2020). Other significant benefits that make people prefer the herbal product is their affordability, ease of access and personalized health care. Majority of the patients use herbal products to treat diseases, relief pain as well as treatment of long term illness such as HIV/AIDs, cancer, hypertension, asthma among others (Ezekwesili-Ofili and Okaka, 2019).

Several challenges associated with identification of herbal products have been reported. They include identity confusion resulting from use of folk names on labels, accidental adulteration and misidentification caused mostly by mix up with closely related plant species (Han *et al.,* 2016). Previous studies have shown discrepancies between species listed on product labels and those analyzed in the laboratory, thus raising concerns on herbal product quality and safety (Palhares *et al.,* 2021). The inability of consumers to visually confirm the identity of the already processed medicinal products further complicates the problem. Moreover, there lacks a systematic approach to assess the safety and effectiveness of these remedies in most developing countries (Han *et al.,* 2016).

Majority of studies carried out on herbal medicine in Kenya has mainly focused on utilization of plants within specific regions or communities for treatment of common diseases while other studies reported on the ethnobotanical survey results on species utilization. Despite reported increased demand and use of herbal products, limited documented data is available on validation of medicinal products traded in Kenyan markets (Okumu *et al.,* 2017). Lack of established government regulations and policies on harvesting, processing and trade of herbal products lead to increased safety concerns, erosion of local indigenous knowledge on herbal medicine, loss of critical biodiversity as well as encourage trade malpractices (Chebii *et al.,* 2020). Therefore, proper identification of plant species from where herbal products are extracted from is vital towards safeguarding safety, quality and efficacy of herbal products.

Previous studies have reported species substitutions and mislabeling of medicinal herbs sold in markets. Herbal products adulteration may be due to improper identification by unskilled collectors, mix up of local dialects, as well as intentional and unintentional misidentification of closely related plant species. Substandard and counterfeit herbal products are a potential risk; therefore efficient biomonitoring techniques are a necessity in authentic identification of listed species to safeguard the health of consumers (Xin *et al.,* 2018).

There is therefore an urgent need for reliable methods that can be used to authenticate herbal products traded in Kenyan markets. Research to determine efficacy, active drug constituents, dosage and safety of herbal medicine has been recommended in Kenya (Chebii *et al.,* 2020), and this can only be achieved if proper identification of potential species has been achieved. Furthermore, DNA technology has not been tested in medicinal plant extracts but has been used to document and generate plant DNA barcodes library using vouchered specimens (Gill *et al.,* 2019). There is therefore limited data on authenticity of herbal products in Kenyan markets, a gap that this study seeks to fill.

Traditionally, herbal products were identified using local indigenous knowledge and local names were assigned on the labels. In most instances, the dried medicinal plants and products on sale lack morphological and taxonomic characteristics that would aid in identification hence the proposal to use the novel DNA technology as well as test two of the recommended DNA barcoding markers in Kenyan products. Species identification using DNA barcoding and metabarcoding techniques to identify plant extracts including those comprising of mixed species has been recommended (Urumarudappa *et al.,* 2020; Yu *et al.,* 2021) and adopted in some countries for instance in China.

The purpose of this study was therefore to apply DNA barcoding technique in identification of herbal products traded in selected markets in Nairobi County alongside the use of taxonomic methods; to test identification efficiency of two

of the recommended barcoding markers, *rbcL* and *ITS* in herbal products traded; and to validate herbal products identified with what was in the product label. Verification of species at DNA level offers more dependable results because DNA is a stable macromolecule present in all tissue and not affected by external factors (Yu *et al.,* 2021).

Barcoding requires only a minute tissue sample from a plant or animal for identification and can be used in cases where the specimen is processed or degraded or when it consists of only non-characteristics parts of the organism. The standardized barcode for plants is a fragment of the plastid gene ribulose 1, 5-biphosphate carboxylase gene (rbcL) combined with a fragment of the maturase (matK) gene. The choice for the rbcL in this study is grounded on easy amplification recovery while *ITS* has high discriminatory ability for closely related species (Khan *et al.,* 2019). The rbcL barcode region is easy to amplify in most of the plant and its sequencing success rate is also high although it is not the most variable one. It has provided valuable barcode dataset besides having lower discriminatory power compared to matK. Internal Transcribed Spacer (*ITS*2) on the other hand has shown high discrimination ability among medicinal plant species and their close relatives and is easy to amplify and sequence (Chen *et al.,* 2010; Nazar *et al*., 2022). The *ITS*2 marker has been used in previous studies and has emerged as a core DNA barcode marker. The two-marker combination of *ITS* and rbcL have been cited to give an identification of over 70% accuracy (Al-Juhani and Khalik, 2021).

Materials and methods

Description of the study area

The study was carried out in Nairobi County which is the capital city of Kenya. Geographically, Nairobi is located in the Central region of Kenya, at an altitude of 1670 m above sea level. The County has a population of 4,397,073 of whom 2,192,452 are male and 2,204,376 are females (KNBS, 2019). The Nairobi county is 703.9 km^2 and has a population density of 6,247 per square kilometer and therefore hosts a large clientele for herbal products (National Coordinating Agency for Population and Development, 2008). The age structure of Nairobi County as per 2019 census is: 0-14 years 1,336,249 persons; 15-64 years 3,002,314 persons; and those above 65 years are 58,122. The study area was purposively selected to include diverse locations where herbal medicine trade is popular within the Nairobi County. Samples were obtained from several markets including Eastleigh, Gikomba, Jogoo road, Machakos Country Bus, Kibera, Ngara, Kamukunji, Kawangware and Pangani markets that are within the high populated locations and fall within ten kilometers radius from the city center (Figure 1).

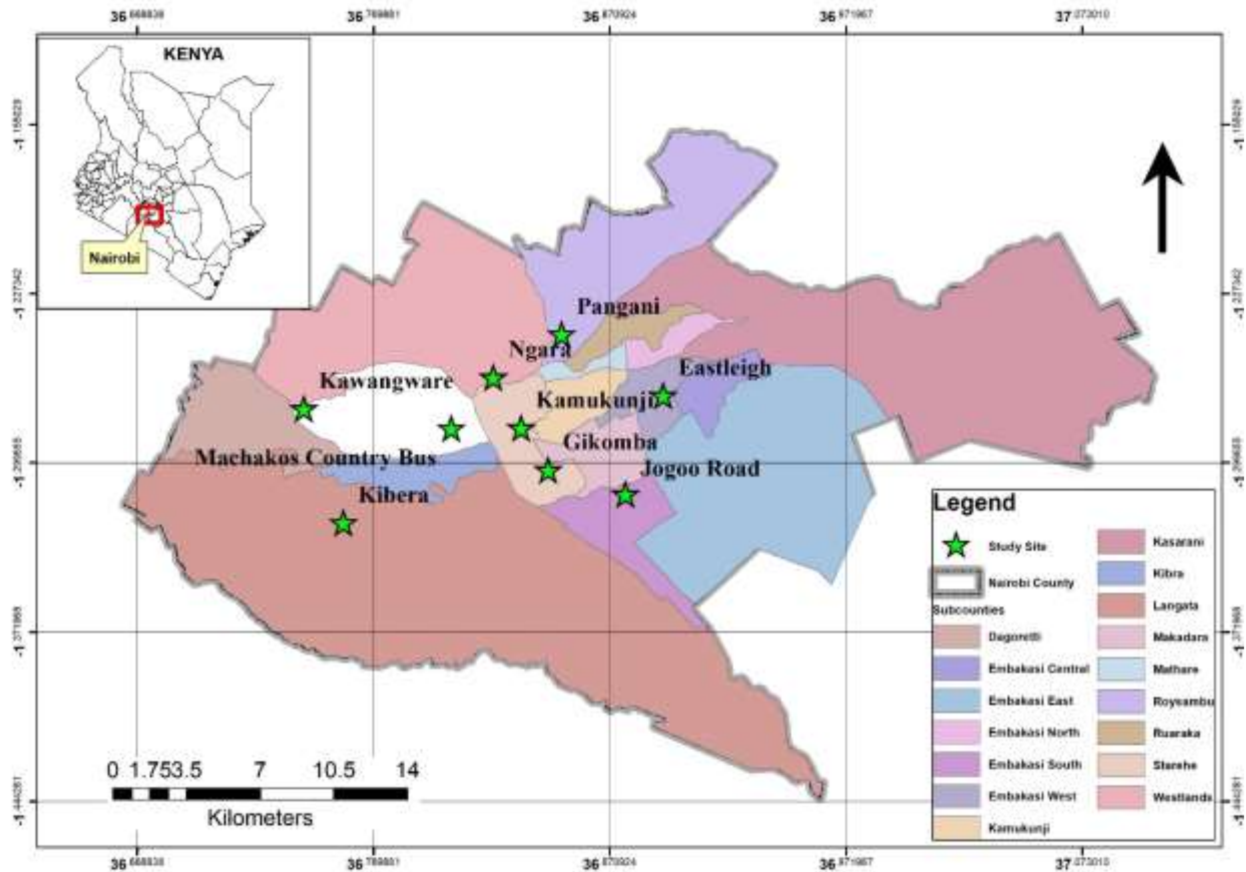

Figure 1. A map showing position of Nairobi County in Kenya and selected sampling area (markets)

The selected study markets are from 6 sub-counties which are known for open air markets for herbal medicines. Kawangware was the furthest and the most populous center in Dagoreti North sub-county while Kamukunji, Starehe, Mathare and Makandara have close proximity to Central Business District and are business hubs within the city and covers some central, western and eastern areas of Nairobi County. Kibra on the other hand is among the smallest sub counties in Nairobi despite being one of the mostly densely populated. Machakos country bus is one of the largest bus stop and hosts travelers from most counties in Kenya and borders major retail markets like Muthurwa, Gikomba, Kamukunji, Wakulima and Marikiti markets.

Sample collection

Since the total number of traders involved with herbal products is not documented, herbalists were identified during market days and randomly selected to participate in the study (Mwaura *et al.,* 2020). Samples were collected from 30 herbalists who consented and were engaged in sale of herbal plant products in the selected open-air markets in Nairobi County. Only single species herbal products were considered for this study. In total 112 samples whose folk/common or scientific names of the ingredient were given were considered for molecular analysis in this study.

A sample weighing approximate 100g of the herbal product(s) was collected from each of the selected herbalists, packaged in self-sealing bags, labeled and transported to the Molecular Genetics laboratory at National Museums of Kenya for further molecular identification and authentication. The vernacular/folk names of the 112 samples collected were translated to scientific names using published literature, parataxonomists and taxonomists from East Africa Herbarium at National Museums of

Kenya and resulted to 55 species. The 55 species were allocated species codes from NB1-NB55 during the analysis as listed in suplimentary Table 1.

Based on their label names, a list of 55 species was generated for voucher specimen collection from the wild for use as reference samples. Herbarium voucher specimens for the listed plant species were collected from neighboring forests in respective geographical areas using standard botanical procedures. Specimens of fresh leaf samples were individually collected from identified species in the wild, labeled and pressed in the field and assigned laboratory identification numbers for further analysis and archival. Further identification and confirmation were done using the relevant taxonomic keys at the East African Herbarium (Beentje *et al.,* 1994; Agnew, 2013). These were reference samples which were later quarantined, accessioned and stored at East African herbarium (EA) at National Museums of Kenya. Sequences generated from these samples and other available barcodes in GenBank were used as reference points to identify the sampled market herbal products.

DNA extraction, amplification and sequencing

Total genomic DNA extraction and amplification was conducted at the Molecular Genetics Laboratory at National Museums of Kenya. DNA was extracted from 0.1g of the dry leaf material from reference plants and market herbal samples using the DNeasy® plant mini kit (Qiagen®) according to the manufacturers' instructions with slight modifications where the water bath digestion time was increased to at least 6 hours for market herbal samples, elution buffer (buffer AE) was heated to 65 ^{o}c before use and elution time increased to 20 minutes. The DNA was eluted in 100ul elution buffer. The quality and concentration of the genomic DNA was determined using Spectro-biophotometer Model X200 and electrophoresis on a 1% agarose gel. The gel was prepared in 1 × Tris Acetate EDTA (TAE) buffer and stained with 0.5 mg/mL ethidium bromide. The samples run for 30 minutes at 80 volts against a 100 bp DNA ladder (Promega Corporation, Madison, WI, USA). The total genomic DNA extracted was stored at -20°C and was later used as template for polymerase chain reaction (PCR).

The amplification of the target loci sequences was performed using lyophilized Bioneer AccuPower® HotStart PCR PreMix (cat no K-5050, Bioneer Corporation, Daejeon, South Korea) in a 20 µL reaction mixture that contained 2.0 µL of 10ng genomic DNA, 1 µL each of forward and reverse primers (10 pM); and 16 µL deionized nuclease free PCR water.

PCR reaction was performed using Nexus Eppendorf master cycler for 35 cycles. The barcode primers used for amplification were rbcL: rbcLa_F - ATGTCACCACAAACAGAGACTAAAGC (Kress & Erickson, 2007); and RbcLR590 – AGTCCACCGCGTAGACATTCAT (de Vere *et al.,* 2012); and *ITS*: *ITS* 2F - ATGCGATACTTGGTGTGAAT, (Chen *et al.,* 2010) and *ITS*3 R – GACGCTTCTCCAGACTACAAT (Chen *et al.,* 2010). The cycling conditions included initial denaturation step for 5 min at 94°C, 35 cycles of denaturation 30 sec at 94 °C, annealing at 56 °C for 30 sec, extension for 40 sec at 72 °C and final extension at 72 °C for 10 min. Success of amplification was confirmed by loading 3 µL of amplified product on 1.5 % agarose gel in 1 × TAE buffer and against a 100 bp DNA ladder under same conditions mentioned above.

The positive amplified fragments (amplicons) with only single bands were purified using ZYMO DNA clean and concentrator kit (Catalogue no. ZR D4034), following the manufacturers' protocol. The concentration of purified product was measured using a Spectro-biophotometer and amplicons with a concentration of 20 ng/µl at 260/280 wavelength were sent to Macrogen Inc. Netherlands for bidirectional sequencing using the same primers. The amplification primers were used as the sequencing primers. The rbcL and *ITS* amplicons were sequenced according to Sanger sequencing method.

Table 1. A summary of market herbal samples whose DNA identified species varied from species listed on labels

S. No	Sample ID	Expected species scientific name based on label	Detected using *ITS* Marker	Detected using rbcL marker
1	EST1	*Acacia nilotica, Schumach. & Thonn*	*Urtica dioica,*	*Urtica dioica,*
2	JGR1	*Acacia nilotica, Schumach. & Thonn*	*Urtica dioica*	*Urtica dioica*
3	KWG3	*Acacia nubica, Benth.*	*Urtica dioica*	*Urtica dioica*
4	NGR2	*Acokanthera schimperi, (A. DC.) Benth. & Hook. f. ex Schweinf.*	*Urtica dioica*	*Urtica dioica*
5	EST2	*Azadirachta indica, A.Juss.*	*Melia azedarach*	*Melia azedarach*
6	JGR3	*Azadirachta indica, A.Juss.*	*Melia azedarach*	*Melia azedarach*
7	MCB4	*Azadirachta indica, A.Juss.*	*Melia azedarach*	*Melia azedarach*
8	NGR4	*Azadirachta indica, A.Juss.*	*Melia azedarach*	*Melia azedarach*
9	PNN2	*Azadirachta indica, A.Juss.*	*Melia azedarach*	*Melia azedarach*
10	JGR6	*Prunus Africana, (Hook.fil.) Kalkm.*	*Urtica dioica*	*Urtica dioica*
11	EST3	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
12	GIK3	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
13	JGR7	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus sp*
14	KBR14	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
15	KKJ9	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
16	KWG12	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
17	MCB13	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
18	NGR6	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus staddo*
19	PNN4	*Rhamnus prinoides, L'Hér*	*Rhamnus staddo*	*Rhamnus sp*
20	KBR15	*Rhoicissus tridentata, (L .f.) Wild & R. B. Drumm.*	*Croton dichogamus*	*Croton sp*
21	NGR8	*Rhoicissus tridentata, (L .f.) Wild & R. B. Drumm.*	*Croton dichogamus*	
22	EST6	*Toddalia asiatica, (L.) Lam.*	*Urtica dioica*	*Urtica dioica*
23	KWG13	*Toddalia asiatica, (L.) Lam.*	*Urtica dioica*	*Urtica dioica*
24	MCB16	*Toddalia asiatica, (L.) Lam.*	*Urtica dioica*	*Urtica dioica*
25	PNN5	*Toddalia asiatica, (L.) Lam.*	*Urtica dioica*	*Urtica dioica*
26	EST7	*Turraea abyssinica, Hochst.*	*Turraea mombassana*	*Turraea mombassana*
27	KBR17	*Turraea abyssinica, Hochst.*	*Turraea mombassana*	*Turraea mombassana*
28	NGR10	*Turraea abyssinica, Hochst.*	*Turraea mombassana*	*Turraea mombassana*
29	MCB17	*Urtica massaica, Mildbr.*	*Urtica dioica*	*Urtica dioica*
30	KBR19	*Zanthoxylum usambarense, (Engl.)*	*Warbugia ugandensis*	*Warbugia ugandensis*
31	KWG15	*Zanthoxylum usambarense, (Engl.)*	*Warbugia ugandensis*	*Warbugia ugandensis*
32	MCB20	*Zanthoxylum usambarense, (Engl.)*	*Warbugia ugandensis*	*Warbugia ugandensis*

Data analysis

The forward and reverse sequences generated were assembled and analyzed per sample per gene using Geneious software R11 (Kearse *et al.*, 2012) using "highest quality" setting and

manual trimming to generate barcodes. TranslatorX was used for alignments of protein coding regions for rbcL sequences. The generated barcode sequences from voucher specimens were deposited to GenBank. To identify market samples, BLAST algorithm (http://blast.ncbi.nlm.nih.gov/blast .cgi) was done using consensus sequences generated from raw reads against known specimens in GenBank sequences. Only BLAST hits with a minimum identification similarity (query coverage) of 97 %, E-value of 0 and identification identity of 97% and above were considered. The query sequences were detected by selecting the highest maximum score with an E value of 0. If the BLAST hit of the query sequence was to species level, it was scored as identified and ambiguous for those identified to genus level. Query sequences identified to species level were scored as authentic if the species identified was similar to the listed or not authentic if species identified was different from the one on the label. Phylogenetic tree analysis using the neighbor-joining method was used for further confirmation of BLAST hits with ambiguity with available *ITS* and rbcL reference sequences in GenBank.

Results

Herbal plant species utilization and extracted DNA quality

The 112 collected medicinal plant samples from the market matched to 55 species and belonged to 32 families after translating from vernacular and common names to botanical names. Among the 32 families, Fabaceae, Apocynaceae, Meliaceae and Rhamnaceae were the 4 most utilized plant families as recorded. They were represented by 9 and 4 species for Fabaceae and Apocynaceae respectively while Meliaceae and Rhamnaceae had 3 species each. The species that were most utilised by the herbalist were *Warburgia ugandensis*, Engl. and *Rhamnus prinoides*, L'Hér at 30 %, followed by *Acacia nilotica*, Schumach. and Thonn at 23 % and *Azadirachta indica*, *A. Juss*, and *Rapanea melanophloeos* (L.) Mez at 17 % respectively.

The barcode sequences for *ITS* and rbcL genes derived from the voucher specimen collected were used as the reference barcode library to authenticate the market samples together with available barcodes from NCBI and BOLD systems. Over 300 reference barcodes were generated and deposited in GenBank. Genomic DNA extraction from the dried samples was challenging with only 82% (92) of the 112 samples yielding detectable DNA. The quality of DNA as determined using spectrobiophotometer from OD260/280 ratio and agarose gel electrophoresis was very variable from 2µg to 89µg per microliter. The PCR amplification was successful in all the 92 dry samples collected for both *ITS* and rbcL markers after dilutions to standardize the concentrations. Some of the amplified amplicons did not produce quality sequences that could be edited for further analysis.

Molecular identification of herbal products

The *ITS*2 barcode had the most success during amplification and produced quality sequences in all the 92 samples analyzed (100%). Recovered *ITS*2 barcodes ranged from 279–438 base pair (bp). The *ITS* barcode had 90 out of 92 (98%) samples identification to species level and (2%) to genus level (Figure 2). The rbcL marker amplified and produced quality sequences for 84 samples (91%) as shown in Figure 2.

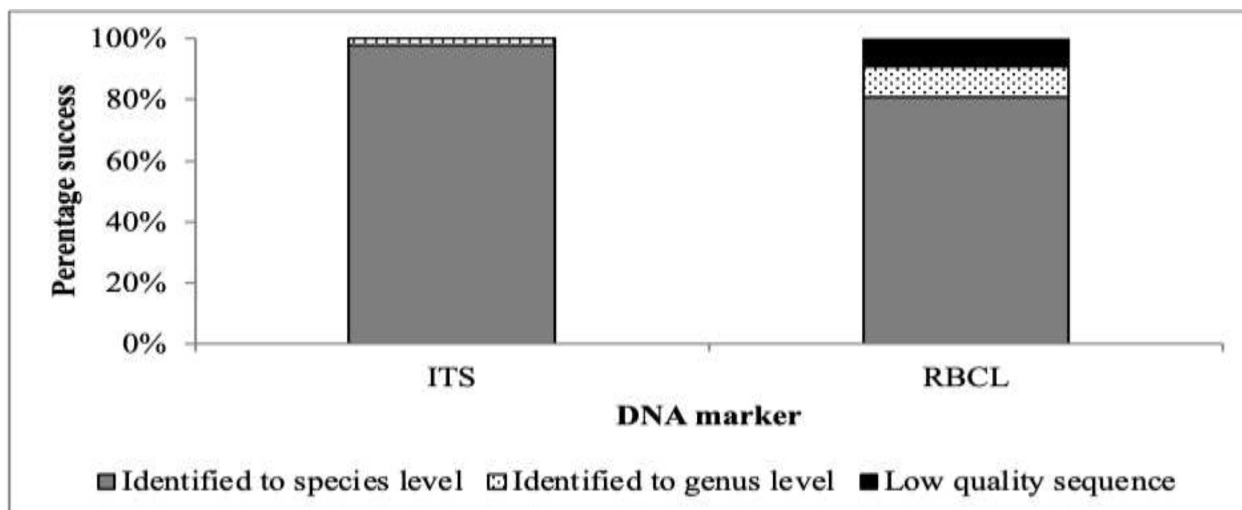

Figure 2. Barcoding markers sequencing and identification success

The rbcL sequenced barcodes ranged from 418–529 bp. Out of the 84 amplified samples, 75 of them (89%) were identified to species level and 11 % to genus level. The 2 markers barcodes recovered for market samples that were not correctly labeled are in Table 1.

Fifty-eight (58) out of the 92 samples sequenced (63%) matched their predicted species hence authentic based on MEGABLAST with *ITS*2 barcode. 32 *ITS*2 sequences (35%) matched other species not listed on labels (Figure 3). The remaining 2 sequences (2%) could only be identified to genus level based on our previous threshold elaborated in the methods section. For the rbcL marker, 47 out of the 92 samples sequenced (56%) were authentic while 28 (33%) matched other species not listed on the labels (Figure 4).

Thirty-two (32) out of the 92 analyzed samples were identified to species level using *ITS* barcode marker and none of them belong to the expected species. This was similar with rbcL marker except for 3 samples (KBR15, JGR7 and PNN4) that were identified to genus level while NGR8 did not produce any sequence (Table 1). Except for A*cacia nubica,* all the other expected species had reference sequences for one or the two barcode markers available in GenBank. Some of the species were substituted with their close relatives such as *Rhamnus prinoides* was substituted with its close relative *Rhamnus staddo* while *Turraea abyssinica* was substituted with *Turraea mombassana*. *Urtica massaica* was substituted with its close relative *Urtica dioica and Azadirachta indica* with its look-alike *Melia azedarach*. Some of the *Acacia nilotica* samples were correctly labled while others were substituted with *Urtica dioica*. Some other species were substituted with unrelated species. For instance, the *Acacia nubica*, *Acokanthera schimperi, Prunus africana* and *Toddalia asiatica* were substituted with *Urtica dioica*, a species that is not closely related to them. *Urtica dioica* is the most utilized species in substitution of the unauthentic samples as detected in this study. Summarized data of substituted samples is in Table 1.

The neighbor joining phylogenetic trees clustered non-authentic species in different clades from the expected species. For instance, samples EST6, KWG13, MCB16 and PNN5 (Table 1) were sold as *Toddalia asiatica* yet their generated sequences from *ITS* and rbcL markers clustered in the same group with *Urtica dioica* and this matched their blast top hits with a similarity between 98.7 to 100%.

Sample JGR6 which was expected to be *Prunus africana* and believed to be a remedy for cancer was grouped together with *Urtica dioica* and had a top blast hit with *U. dioica* with a 100% similarity. Other notable result was that all herbal products on sale as Muarobaini/Neem (*Azadirachta indica)* were not authentic and belonged to its close relative *Melia azedarach* as

all generated sequences clustered in same the group and distantly from *A. indica.*

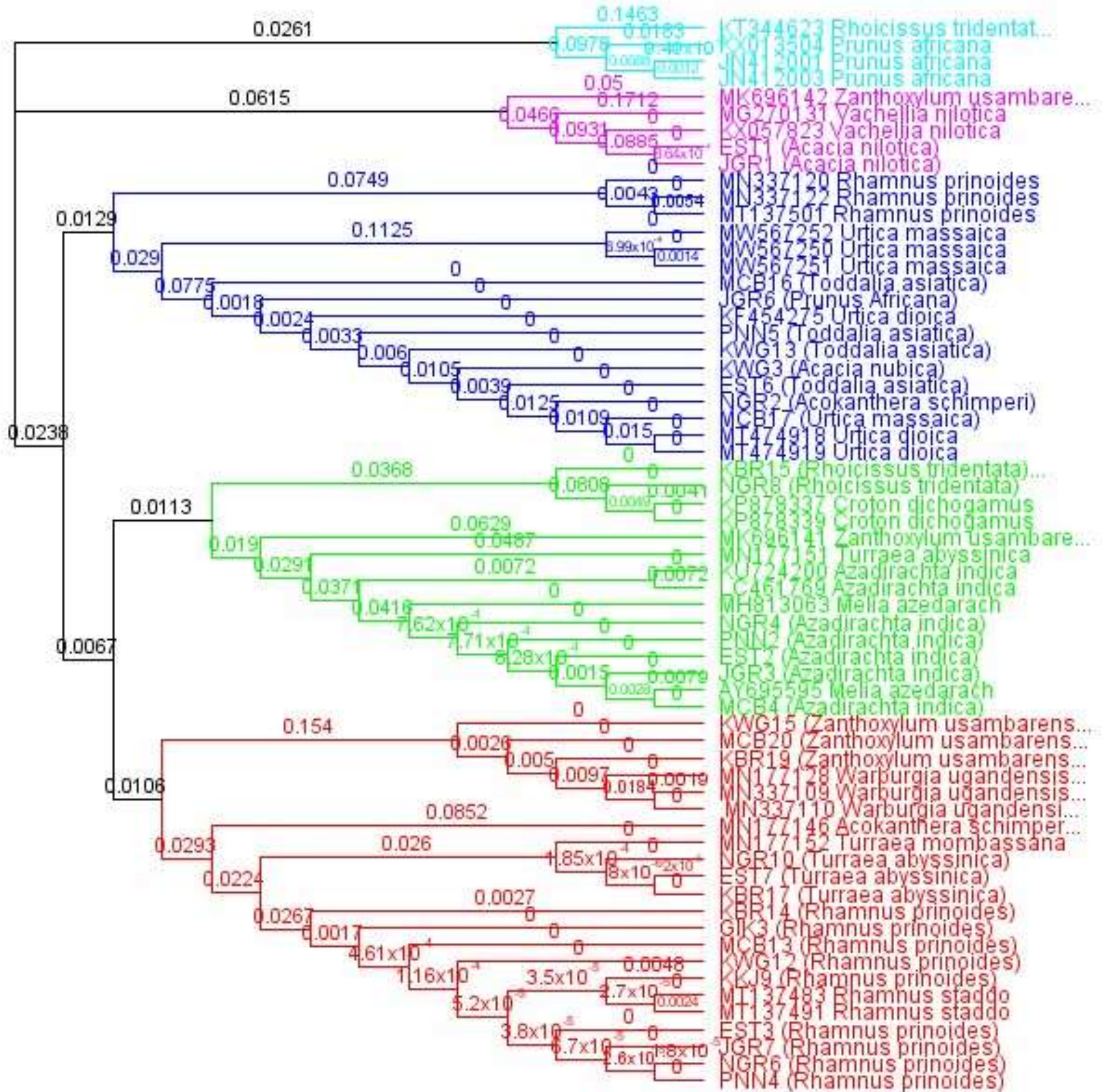

Figure 3. Phylogenetic tree constructed using ITS sequences generated from markets samples and reference sequences from GenBank. Scientific names in () indicate expected species on labels of market samples while those others are reference sequences of best hits

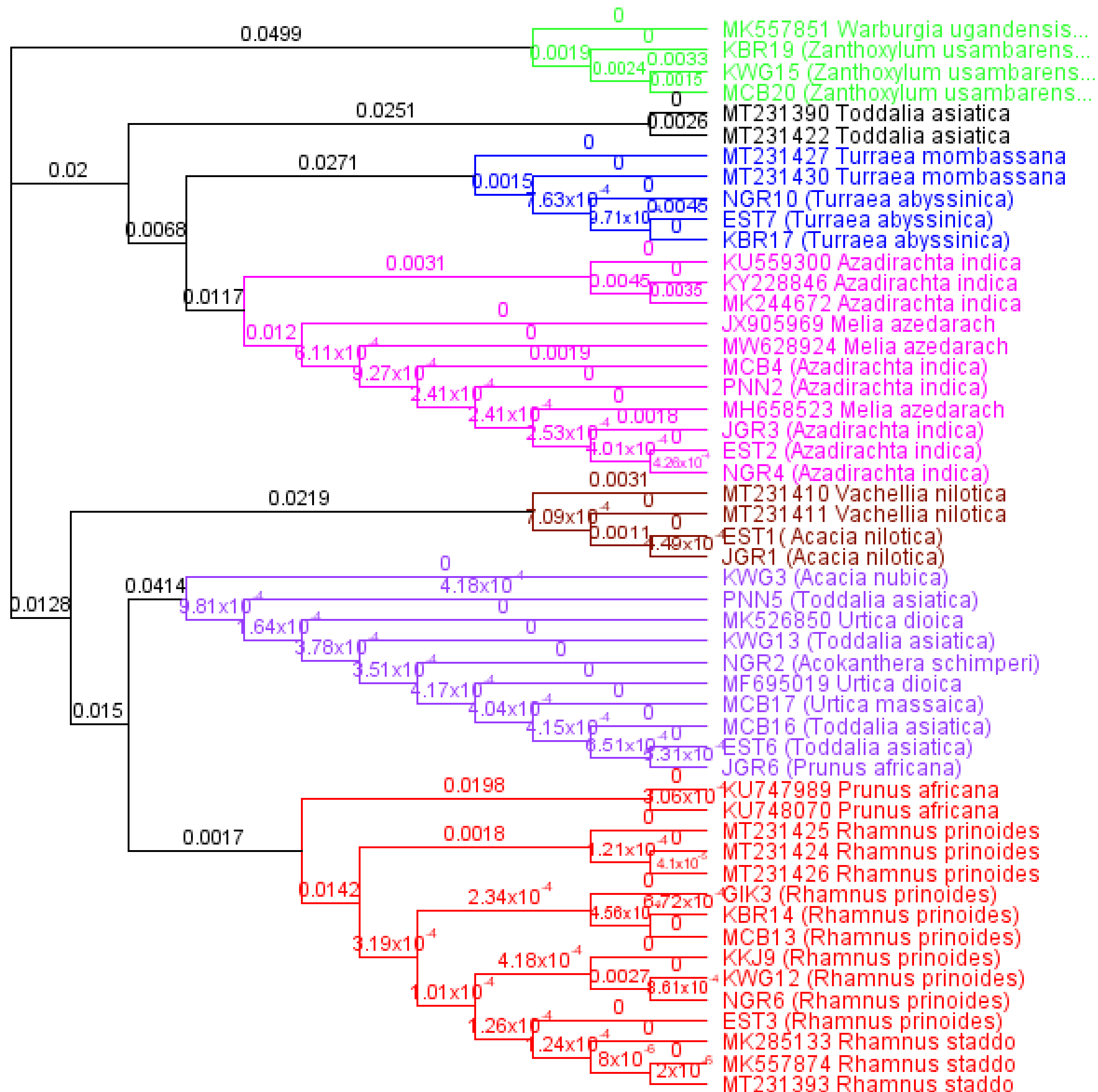

Figure 4. Phylogenetic tree constructed using rbcL sequences generated from markets samples and reference sequences from GenBank. Scientific names in () indicate expected species on labels of market samples while those others are reference sequences of best hit

Thus, the trees constructed using generated sequences using both *ITS* and rbcL markers confirmed species identification done using BLAST since the identified species could not cluster with the expected species (Figure 3 and Figure 4).

Discussion

A total of 55 species belonging to 32 families were recorded as the most traded in Nairobi County. Fabaceae family is the most traded family followed by Apocynaceae, Meliaceae and Rhamnus. The four families have been reported in previous studies carried out in Central

province in Kenya as preferred by herbalists trading in Thika and Nairobi towns (Njoroge, 2012). Similarly, Fabaceae has also been reported to be the most traded plant family in herbal medicine in Madhupur forest area of Bangladesh (Islam *et al.,* 2014); in India (Silambarasan *et al.,* 2017); Tigray region of Ethiopia (Kidane *et al.,* 2018); and in Tanzania (Hilonga *et al.,* 2019). Utilization of Rhamnaceae as ethnomedicinal plants has been reported in other parts of Kenya, Pakistan and other parts of the world (Shinwari, 2010; Mailu *et al.,* 2020) similar to the findings of this study. Apocynaceae and Fabaceae families have been reported among the top ten families that are being utilized in herbal medicine in China (Yu *et al.,* 2021) and these results are in line with current study results. Meliaceae and Apocynaceae families have been reported in studies carried out in Suba district in Kenya as among the most utilized families by herbalists treating sexually transmitted diseases (Nagata *et al.,* 2011).

Molecular identification depends mainly on PCR amplification and sequencing. Genomic DNA extraction from the dried market samples was challenging unlike in fresh samples with a success rate of 82%. In this study, *ITS* marker has shown the highest species identification efficiency at 98% to species level and 2% to genus level. The rbcL was second with 89% identification to species level and 11% to genus level. Previous studies have also recorded *ITS* as an efficient marker for identification of medicinal products to species level (Nazar *et al.,* 2022). Such studies where *ITS* was utilized are in Iran where it was used to distinguish *Ziziphora* species; in molecular authentication of the Chinese medicinal species of Liguisticum, usually sold as dried sliced material; and in Saudi Arabia where it was reported as an ideal marker for identifying arid medicinal plants (Khan *et al.,* 2019; Liu *et al.,* 2019; Sheidai *et al.,* 2019). Thus, this study results concur with recommendations by Chen *et al.,* (2010) that *ITS*2 should be the gold standard barcode for plants identification.

Combination of *ITS* and *rbcL* in this study resulted to identification of majority of the samples to species level. Previous studies have established that the combination of different DNA barcode markers is necessary for adequate species verification (Tnah *et al.,* 2019). In Canada, the species from native temperate flora of Ontario were tested using *rbcL*, matK and *trnHPsbA* barcodes that had 95.3 % identification success (Burgess *et al.,* 2011).

The BLAST results were confirmed by the phylogenetic tree constructed to assess the closely related species and expected species relationship. The neighbor joining trees clustered the generated sequences closely to reference sequences in NCBI database. Some of the species with more than one individual used in constructing the phylogenetic trees either by rbcL or *ITS* demonstrated the presence of intraspecific variations; for example, *Acacia nilotica and Zanthoxylum usambarense* as shown in Figure 3 and Figure 4 respectively. From this study, over 30% of the market samples analyzed were not authentic since the phylogenetic trees placed them in different genetically distinct groups from expected reference species concurring with the top hits produced by BLAST.

Both BLAST and Phylogenetic tree results indicated that some market samples did not contain the species listed on the label. Some of the listed species are replaced with unrelated species for instance *Rhoicissus tridentata* was substituted with *Croton dichogamus* and *Urtica dioica,* while *Zanthoxylum usambarense* was substituted with *Warbugia ugandensis*. *Prunus africana* (Rosaceae), a species commonly used to treat cancer and other medical conditions was substituted with U*rtica dioica* (Urticaceae) which is unrelated. The results are in contrast to the belief that medicinal plants are generally substituted or adulterated with closely related species (Keshari, 2021). This affects the efficacy of the intended herbal remedies and the resultant effects to consumers may never be established. Such adulteration and substitution has been reported in several other studies for instance in North America and United Arab Emirates where most of the tested herbal products were contaminated or substituted (Mosa *et al.,* 2018; Newmaster *et al.,* 2013).

The other listed species were substituted with their close relatives for instance *Rhamnus prinoides* was substituted with its close relative *Rhamnus staddo* while *Turraea abyssinica* was substituted with *Turraea mombassana*. *Urtica massaica* was substituted with its close relative *Urtica dioica*, as confirmed by both BLAST and Phylogenetic tree where reference sequences from GenBank for both species were included. This scenario was similar for *Azadirachta indica* where the market samples clustered with its look alike *Melia azedarach*. This would be expected for closely related species with similar vernacular names that are normally hard to distinguish during wild collection. These results are similar to those recorded in Malaysia by Tnah *et al.,* (2019) where herbal products on labels were found to be substituted with herbal plant species with similar vernacular names.

The *Acacia nubica*, *Acokanthera schimperi, Prunus africana* and *Toddalia asiatica* were substituted with *Urtica dioica*, a species that is not closely related to them. *Urtica dioica* is the most utilized species in substitution of the unauthentic samples. The reason for such a prolific substitution may be due to the abundance of the *Urtica dioica* species in the wild in most parts of the country and normally utilized as vegetable and medicinal products. The *ITS* and rbcL barcoding of the market products was necessary in uncovering these contaminated/substituted products. These results underscore the importance of implementation of quality control procedures by regulatory agencies to safeguard the safety of the consumers.

Conclusion

This study has made a first attempt to use DNA barcoding technology to taxonomically identify species traded as herbal products in Nairobi, Kenya. The generated reference barcodes from vouchered specimens have been deposited in GenBank for medicinal plants of economic importance. Development of such DNA library for medicinal plants of Kenya will aid in authentication of herbal products being traded in the region and this will be an asset to both consumers and traders due to the existence of increased demand of herbal medicine. Our findings have shown that DNA can be recovered in ground and powdered herbal products and ITS marker was superior in identifying medicinal plants to species level compared to rbcL. Since *ITS*2 DNA barcode marker is reliable and easily amplified, it can be used by regulatory agencies to verify economically important plants in the region to enhance safety and efficacy. Further studies are recommended to ascertain at what point the adulteration occurs and DNA barcoding technology can be used as a reliable tool for identification and reduce the chances of undesirable consequences due to the use of misidentified plant materials.

Acknowledgments

We acknowledge support by Mathias Mbale and Henry Saitabau in parataxonomy and voucher specimen collection, G. Mwangi is acknowledged for development of the maps. This work was partly funded by the Government of Kenya through the National Research Fund (NRF) grants for post graduates.

References

Agnew, A. D. (2013). *Upland Kenya wild flowers and ferns : a flora of the flowers, ferns, grasses, and sedges of highland Kenya*. Nairobi Kenya: Nature Kenya-The East Africa Natural History Society, 530 pp.

Al-Juhani, W. S., & Khalik, K. N. A. (2021). Identification and molecular study of medicinal Plectranthus species (Lamiaceae) from Saudi Arabia using plastid DNA regions and ITS2 of the nrDNA gene. *Journal of King Saud University - Science*, *33*(5), 101452. https://doi.org/10.1016/j.jksus.2021.101452

Beentje, H., Adamson, J., & Bhanderi, D. (1994). *Kenya trees, shrubs, and lianas*. Online publication

Burgess, K. S., Fazekas, A. J., Kesanakurti, P. R., & Graham, S. W. (2011). *Discriminating plant species in a local temperate flora using the rbcL + matK DNA barcode*. (August). https://doi.org/10.1111/j.2041-210X.2011.00092.x

Chebii, W. K., Muthee, J. K., & Kiemo, K. (2020). The governance of traditional medicine

and herbal remedies in the selected local markets of Western Kenya. *Journal of ethnobiology and ethnomedicine, 16(1)*, 1-24.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., … 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

de Vere, N., Rich, T. C. G., Ford, C. R., Trinder, S. A., Long, C., Moore, C. W., … Wilkinson, M. J. (2012). DNA Barcoding the Native Flowering Plants and Conifers of Wales. *PLoS ONE*, *7*(6), e37945. https://doi.org/10.1371/journal.pone.0037945

Ezekwesili-Ofili, J. O., & Okaka, A. N. C. (2019). Herbal Medicines in African Traditional Medicine. In *Herbal Medicine*. https://doi.org/10.5772/intechopen.80348

Gill, B. A., Musili, P. M., Kurukura, S., Hassan, A. A., Goheen, J. R., Kress, W. J., … Kartzinel, T. R. (2019). Plant DNA-barcode library and community phylogeny for a semi-arid East African savanna. *Molecular Ecology Resources*, *19*(4), 838–846. https://doi.org/10.1111/1755-0998.13001

Han, J., Pang, X., Liao, B., Yao, H., Song, J., & Chen, S. (2016). An authenticity survey of herbal medicines from markets in China using DNA barcoding. *Scientific Reports*, *6*(1), 18723. https://doi.org/10.1038/srep18723

Hilonga, S., Otieno, J. N., Ghorbani, A., Pereus, D., Kocyan, A., & de Boer, H. (2019). Trade of wild-harvested medicinal plant species in local markets of Tanzania and its implications for conservation. *South African Journal of Botany*, *122*, 214–224. https://doi.org/10.1016/j.sajb.2018.08.012

Islam, M. K., Saha, S., Mahmud, I., Mohamad, K., Awang, K., Jamal Uddin, S., … Shilpi, J. A. (2014). An ethnobotanical study of medicinal plants used by tribal and native people of Madhupur forest area, Bangladesh. *Journal of Ethnopharmacology*, *151*(2), 921–930. https://doi.org/10.1016/j.jep.2013.11.056

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)*, *28*(12), 1647–1649. https://doi.org/10.1093/BIOINFORMATICS/BTS199

Kenya National Bureau of Statistics (KNBS). (2019). 2019 Kenya Population and Housing Census Volume 1: Population by County and Sub-County. In *2019 Kenya Population and Housing Census*. Retrieved from https://www.knbs.or.ke/?wpdmpro=2019-kenya-population-and-housing-census-volume-i-population-by-county-and-sub-county

Keshari, P. (2021). Controversy, Adulteration and Substitution: Burning Problems in Ayurveda Practices. *Natural Medicinal Plants*. https://doi.org/10.5772/INTECHOPEN.98220

Khan, S. A., Baeshen, M. N., Ramadan, H. A., & Baeshen, N. A. (2019). ITS2: An Ideal DNA Barcode for the Arid Medicinal Plant Rhazya Stricta. *Pharmaceutical Medicine*, *33*(1), 53–61. https://doi.org/10.1007/S40290-019-00266-3

Kidane, L., Gebremedhin, G., & Beyene, T. (2018). Ethnobotanical study of medicinal plants in Ganta Afeshum District, Eastern Zone of Tigray, Northern Ethiopia. *Journal of Ethnobiology and Ethnomedicine*, *14*(1), 1–19. https://doi.org/10.1186/s13002-018-0266-z

Kress, W. J., & 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. https://doi.org/10.1371/journal.pone.0000508

Liu, M., Li, X. W., Liao, B. S., Luo, L., & Ren, Y. Y. (2019). Species identification of poisonous medicinal plant using DNA

barcoding. *Chinese Journal of Natural Medicines*, *17*(8), 585–590. https://doi.org/10.1016/S1875-5364(19)30060-3

Mahomoodally, M. F. (2013). Traditional medicines in Africa: An appraisal of ten potent African medicinal plants. *Evidence-Based Complementary and Alternative Medicine*, *2013*. https://doi.org/10.1155/2013/617459

Mailu, J. K., Nguta, J. M., Mbaria, J. M., & Okumu, M. O. (2020). Medicinal plants used in managing diseases of the respiratory system among the Luo community: An appraisal of Kisumu East Sub-County, Kenya. *Chinese Medicine (United Kingdom)*, *15*(1). https://doi.org/10.1186/s13020-020-00374-2

Mbuni, Y. M., Wang, S., Mwangi, B. N., Mbari, N. J., Musili, P. M., Walter, N. O., … Wang, Q. (2020). Medicinal plants and their traditional uses in local communities around cherangani hills, western Kenya. *Plants*, *9*(3), 1–16. https://doi.org/10.3390/plants9030331

Mosa, K. A., Soliman, S., El-Keblawy, A., Ali, M. A., Hassan, H. A., Tamim, A. A. B., & Al-Ali, M. M. (2018). Using DNA Barcoding to Detect Adulteration in Different Herbal Plant- Based Products in the United Arab Emirates: Proof of Concept and Validation. *Recent Patents on Food, Nutrition & Agriculture*, *9*(1). https://doi.org/10.2174/2212798410666180409101714

Mwaura, A., Kamau, J., & Ombori, O. (2020). An ethnobotanical study of medicinal plants commonly traded in Kajiado, Narok and Nairobi counties, Kenya. *East African Journal of Science, Technology and Innovation*, *1*(3), 2020. https://doi.org/10.37425/EAJSTI.V1I3.153

Nagata, J. M., Jew, A. R., Kimeu, J. M., Salmen, C. R., Bukusi, E. A., & Cohen, C. R. (2011). Medical pluralism on Mfangano Island: Use of medicinal plants among persons living with HIV/AIDS in Suba District, Kenya. *Journal of Ethnopharmacology*, *135*, 501–509. https://doi.org/10.1016/j.jep.2011.03.051

National Coordinating Agency for Population & Development. (2008). Seeking Solutions for Traditional Herbal Medicine: Kenya Develops a National Policy. In *Policy Brief No. 1*.

Nazar, N., Howard, C., Slater, A., & Sgamma, T. (2022). Challenges in Medicinal and Aromatic Plants DNA Barcoding—Lessons from the Lamiaceae. *Plants*, *11*(1). https://doi.org/10.3390/plants11010137

Newmaster, S. G., Grguric, M., Shanmughanandhan, D., Ramalingam, S., & Ragupathy, S. (2013). DNA barcoding detects contamination and substitution in North American herbal products. *BMC Medicine*, *11*(1), 222. https://doi.org/10.1186/1741-7015-11-222

Njoroge, G. N. (2012). Traditional medicinal plants in two urban areas in Kenya (Thika and Nairobi): Diversity of traded species and conservation concerns. *Ethnobotany Research and Applications*, *10*(0), 329–338. https://doi.org/10.17348/era.10.0.329-338

Okumu, M. O., Ochola, F. O., Onyango, A. O., Mbaria, J. M., Gakuya, D. W., Kanja, L. W., … Onyango, M. A. (2017). The legislative and regulatory framework governing herbal medicine use and practice in Kenya: a review. *Pan African Medical Journal*, *28*. https://doi.org/10.11604/pamj.2017.28.232.12585

Palhares, Rafael M., Baratto, L. C., Scopel, M., Mügge, F. L. B., & Brandão, M. G. L. (2021). Medicinal Plants and Herbal Products From Brazil: How Can We Improve Quality? *Frontiers in Pharmacology*, *11*(January), 1–4. https://doi.org/10.3389/fphar.2020.606623

Sheidai, M., Tabaripour, R., Talebi, S. M., Noormohammadi, Z., & Koohdar, F. (2019). Adulteration in medicinally important plant species of Ziziphora in Iran market: DNA barcoding approach. *Industrial Crops and Products*, *130*, 627–633. https://doi.org/10.1016/j.indcrop.2019.01.025

Shinwari, Z. K. (2010). Medicinal plants research in Pakistan. *Journal of Medicinal Plants Research*, *4*(3), 161–176. Retrieved from http://www.academicjournals.org/JMPR

Silambarasan, R., Sureshkumar, J., Krupa, J., Amalraj, S., & Ayyanar, M. (2017). Traditional herbal medicines practiced by the ethnic people in Sathyamangalam forests of Western Ghats, India. *European Journal of Integrative Medicine*, *16*(August), 61–72. https://doi.org/10.1016/j.eujim.2017.10.010

Tnah, L. H., Lee, S. L., Tan, A. L., Lee, C. T., Ng, K. K. S., Ng, C. H., & Nurul Farhanah, Z. (2019). DNA barcode database of common herbal plants in the tropics: a resource for herbal product authentication. *Food Control*, *95*, 318–326. https://doi.org/10.1016/j.foodcont.2018.08.022

Tugume, P., & Nyakoojo, C. (2019). Ethno-pharmacological survey of herbal remedies used in the treatment of paediatric diseases in Buhunga parish, Rukungiri District, Uganda. *BMC Complementary and Alternative Medicine*, *19*(1), 1–10. https://doi.org/10.1186/s12906-019-2763-6

Urumarudappa, S. K. J., Tungphatthong, C., Prombutara, P., & Sukrong, S. (2020). DNA metabarcoding to unravel plant species composition in selected herbal medicines on the National List of Essential Medicines (NLEM) of Thailand. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-75305-0

Xin, T., Xu, Z., Jia, J., Leon, C., Hu, S., Lin, Y., … Newmaster, S. G. (2018). Biomonitoring for traditional herbal medicinal products using DNA metabarcoding and single molecule, real-time sequencing. *Acta Pharmaceutica Sinica B*, *8*(3), 488–497. https://doi.org/10.1016/j.apsb.2017.10.001

Yu, J., Wu, X., Liu, C., Newmaster, S., Ragupathy, S., & Kress, W. J. (2021). Progress in the use of DNA barcodes in the identification and classification of medicinal plants. *Ecotoxicology and Environmental Safety*, *208*. https://doi.org/10.1016/j.ecoenv.2020.111691