Global Journal of Science Frontier Research: G
Bio-Tech & Genetics
Volume 15 Issue 1 Version 1.0 Year 2015
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN:2249-4626&Print ISSN:0975-5896

DNA Barcoding, Identification and Validation of the Genus *Phyllanthus* in Nigeria using Plastid rbcL and matK Genetic Markers

By Daniel Azubuike Awomukwu, Bio Louis Nyananyo, Chiedozie Joel Uka, Paula Spies & Bulelani Londoloza Sizani

Federal University, Nigeria

Abstract- DNA of five species of *Phyllanthus* in Nigeria namely, *P. amarus* Schum and Thonn, *P. urinaria* Linn., *P. odontadenius* Mull-Arg., *P. niruroides* Mull-Arg. and *P. muellerianus* (O. Ktze) Excel belonging to the family of Phyllanthaceae were extracted, purified, PCR amplified and sequenced using plastid Ribulose -1,5 bisphosphate carboxylase large chain (rbcL a) and Maturase K (matK) genetic marker to identify unknown *Phyllanthus* species. The plastid region revealed that the *Phyllanthus* species were able to be amplified optimally for sequencing. The results of the nucleotide sequences were further compared on Basic Local Alignment Sequence Tool (BLAST) on Gen Bank and BoldSystems for validation. Results revealed that the closely related species, *P. niruroides* Mull_Arg. and *P. odontadenius* Mull-Arg. had no DNA record to separate them on both GenBank and BoldSystems using both rbcL and matK gene regions while *P. amarus* Schum and Thonn and *P. urinaria* Linn. were clearly compatible with other works. *P. muellerianus* (O. Ktze) Excel was only compatible with other works using matK gene region but none for rbcL.

Keywords: phyllanthus species, DNA extraction, PCR amplification, nucleotide sequences, BLAST, bold systems, rbcL, matk, barcoding.

GJSFR-G Classification : FOR Code: 069999

Strictly as per the compliance and regulations of :

DNA Barcoding, Identification and Validation of the Genus *Phyllanthus* in Nigeria using Plastid rbcL and matK Genetic Markers

Daniel Azubuike Awomukwu [α], Bio Louis Nyananyo [σ], Chiedozie Joel Uka [ρ], Paula Spies [ω] & Bulelani Londoloza Sizani [¥]

Abstract- DNA of five species of *Phyllanthus* in Nigeria namely, *P. amarus* Schum and Thonn, *P. urinaria* Linn., *P. odontadenius* Mull-Arg., *P. niruroides* Mull-Arg. and *P. muellerianus* (O. Ktze) Excel belonging to the family of Phyllanthaceae were extracted, purified, PCR amplified and sequenced using plastid Ribulose -1,5 bisphosphate carboxylase large chain (rbcL a) and Maturase K (matK) genetic marker to identify unknown *Phyllanthus* species. The plastid region revealed that the *Phyllanthus* species were able to be amplified optimally for sequencing. The results of the nucleotide sequences were further compared on Basic Local Alignment Sequence Tool (BLAST) on GenBank and BoldSystems for validation. Results revealed that the closely related species, *P. niruroides* Mull_Arg. and *P. odontadenius* Mull-Arg. had no DNA record to separate them on both GenBank and BoldSystems using both rbcL and matK gene regions while *P. amarus* Schum and Thonn and *P. urinaria* Linn. were clearly compatible with other works. *P. muellerianus* (O. Ktze) Excel was only compatible with other works using matK gene region but none for rbcL. BLAST validation also revealed that *P. amarus* is synonymous with *P. niruri* due to overall similarity they share on both rbcL and matK genome and needs to be subsumed. Results of the nucleotide sequences and fragment analysis were published on BoldSystems for barcoding as a standard coding marker translation matrix.

keywords: Phyllanthus species, DNA extraction, PCR amplification, nucleotide sequences, BLAST, BoldSystems, rbcL, matK, Barcoding.

I. Introduction

One of the most reliable methods for identification of medicinal plants involves morphological and genetic analysis. Molecular techniques have been also introduced for DNA fingerprinting (Sucher and Charles, 2008). Analysis of the DNA that is present in all organisms is a suitable method for identifying plant materials because the genetic composition is unique for each individual organism. DNA extracted from the leaves, stems or roots of plants all carry the same genetic information without being affected by physiological conditions and environmental factors. Polymerase Chain Reaction (PCR) in combination with Sequencing and DNA barcoding has been widely used for DNA fingerprinting (Weisling *et al*., 2005). Species-species regions in nuclear DNA, mitochondrial DNA and chloroplast DNA have been used for identification of individual species. DNA fingerprinting is an important tool for molecular characterization of various groups of plants. It offers a faster and more precise way of determining relationships among closely related species than that of morphological investigation (Rahman, 2007). This is because morphological characteristics are subject to environmental influence and extensive studies of mature plants are often necessary for taxonomic classification.

The first proposed DNA-barcoding regions for universal plant DNA-barcoding suggested by the PWG, were the multicopy nuclear Internal Transcribed Spacer (ITS), the rbcLa subunit and matK (Stoeckle *et al*., 2004). Since then, several barcoding regions have been investigated, tested and proposed for different groups. Gene regions that are popular in phylogenetic studies have been investigated for possible candidate regions to be used in barcoding. Loci that are popular in plant systematics are rbcL, the trnL-F intergenic spacer, matK, ndhF and atpB. Two of these regions, rbcL and atpB, are used in phylogenetic studies to distinguish at genus level and above. Even though the characteristic of a suitable barcoding region is that there should be distinction at the species level, rbcL and atpB have been considered as barcoding regions (Blaxter, 2004). The regions matK and ndhF have enough variation to be used in phylogeny on interspecific level, but unfortunately only when the sequenced length is more than 1000bp does it provide enough variation for discrimination (Kress *et al*., 2005). The CBOL plant working group proposed the use of rbcL and matK as universal plant barcoding regions

Author α: Department of Biological Sciences, Federal University, Otuoke, Bayelsa State, Nigeria. e-mail: xdanny18@yahoo.com
Author σ: Department of Plant Science and Biotechnology, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria. e-mail: bionyananyo@yahoo.com
Author ρ: Department of Botany, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria. e-mail: dozemancentury@yahoo.com
Author ω: Department of Genetics, University of the Free State, Bloemfontein, Free State, South Africa. e-mail: spiesp@ufs.ac.za
Author ¥: Department of Genetics, University of the Free State, Bloemfontein, Free State, South Africa. e-mail: nhonhophela@gmail.com

(CBOL Plant Working Group, 2009). The SciVerse Scopus bibliography database (accessed on October 2012) have 283 citations to this article and the common conclusion of many of these studies is that a universal barcode system can still not be agreed upon due to lack of universality, sequence quality and lack of discriminatory power (CBOL Plant Working Group, 2009).

The molecular data of *Phyllanthus*, the third largest genus in the family, has received considerable attention as an important character in inferring systematic relationships (Wurdack *et al*., 2004; Samuel *et al*., 2005; Kathriarachchi *et al*., 2005). The potential to apply barcoding in plant taxonomy were first explored during an exploratory workshop in 2003 (held at the "Cold Spring Harbor Banbury Conference Center" from 9 – 12 March; accessed on http://www.barcodeoflife.org/content/about/what-cbol) and it was predicted that barcoding will in future be utilized in species identification, conservation biology and mapping the extent of species by linking maps to barcodes. It was also predicted that the cost of barcoding a sample would decrease to such a degree that it would be affordable to be used by science teachers and "backyard naturalists" (Stoeckle *et al*., 2003). A DNA-barcode is "a short DNA-sequence that identifies a species" (Stoeckle *et al*., 2003), by comparing the sequence of an unknown specimen to barcodes in a sequence database of known species (Kress & Erickson, 2007). The main use of these sequences is for identification and not for phylogenetic reconstruction (Kress & Erickson, 2007) or as only criterion in describing new species (Stoeckle *et al*., 2003). It differs from molecular phylogeny in that the main goal is not to determine patterns of relationship but to identify an unknown sample in terms of a preexisting classification.

Phyllanthus has a remarkable diversity of growth forms (annual and perennial herbaceous, arborescent, climbing, floating aquatic, pachycaulous, and phyllocladous), floral morphology (Bancilhon, 1971), and chromosome numbers (Webster and Ellis, 1962). The diversity of pollen types (Ko¨ hler, 1965, 1967; Punt, 1967, 1987; Webster and Carpenter, 2002; Sagun and Van der Ham, 2003) rivals that of any genus of flowering plants. The vast majority of *Phyllanthus* species, however, share a distinctive vegetative specialization known as ''phyllanthoid branching'' (Webster, 1956) with leaves on the main axes reduced to scales called ''cataphylls'' and those on lateral (plagiotropic), deciduous, floriferous axes developing normally.

Phyllanthus has a long history of use in tropical countries in indigenous medicine for the treatment of liver ailment. They were examined during several researches. The pharmacognistic importance of some of these *Phyllanthus* species found in Nigeria has been elucidated by Awomukwu *et al*. (2014). These medicinal plants have been underutilized in orthodox medicine but have confirmed to be used worldwide in the pharmaceutical, food, cosmetics and perfume industries. Different authors in different groups have studied the taxonomic significance and scientific implication of the morphological, anatomical and epidermal features of *Phyllanthus* species in Nigeria. They include Edeoga *et al*. (2007) and Uka *et al*. (2014). Awomukwu *et al*. (2015) has successfully delineated the indigenous species occurring in Nigeria based on its nuclear ribosomal Internal Transcribed Spacer (ITS 4-5) genetic marker in which the DNA molecular data sequences were able to identify, validate and classify the various species. None of such works have been done on barcoding using the concatenation of rbcL and matK gene regions. Series of documented descriptions of the morphological, anatomical, ethnobotanical and phytochemical characteristics of *Phyllanthus* exist; there is confusion in recognizing individual species of these plants in Nigeria. The aim of this work is to provide reliable genetic information in order to aid easy recognition and to authentically barcode the *Phyllanthus* species for pharmacognistic research with reference to modern day genetic search tools.

II. Materials and Methods

a) Collection of plant materials

Mature plants of the five species *P. amarus, P. urinaria, P. niruroides, P. odontadenius* and *P. muellerianus.* were collected from different locations of Nigeria (bounding box coordinates: upper left – 6.3333, 7; lower right – 4.75, 6.8333) by various investigators as in Table 1. Only healthy, fresh and succulent parts of the plants were collected. The five specimens were identified and authenticated at the Herbaria of the Department of Plant Science and Biotechnology, Michael Okpara University of Agriculture, Umudike, Umuahia, Abia State and the Department of Botany, Nnamdi Azikiwe University, Awka, Anambra State. Herbarium specimens were also studied at the various institutions as well making reference to the Flora of West Tropical Africa by Hutchinson and Dalziel (1963). The accessions were deposited at the Herbarium of the Department of Plant Science and Biotechnology, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria with their sample and process identity numbers for further research. The same specimen records were also submitted on public data portal at Boldsystems. Further laboratory analysis was carried out at the Molecular Genetic Lab, University of the Free States, Bloemfontein, South Africa.

Table 1 : Collection sites, identity numbers and collection dates of the five *Phyllanthus* species studied

SPECIES	SITE OF COLLECTION	CORDINATE/ ELEVATION	SAMPLE ID	PROCESS ID	DATE OF COLLECTION
P. amarus	Along the school fence, Abia State Polytechnic, Aba.	N5.42; E6.33; 25.0m	AWOM UPH PA 010	PHSN 001-14	April 15, 2014
P. urinaria	Field around National Root Crop Research Institute, Umudike.	N4.75; E6.83; 20.0m	AWOM UPH PN 050	PHSN 003-14	April 14, 2014
P. odontadenius	Road side along National Root Crop Research Institute, Umuahia.	N5.42; E7.50; 25.0m	AWOM UPH PO 040	PHSN 006-14	April 15, 2014
P. niruroides	Science Village, Nnamdi Azikiwe University, Awka.	N6.33; E7.00; 22.0m	AWOM UPH PU 030	PHSN 005-14	April 15, 2014
P. muellerianus	Near the Herbarium Building, Nnamdi Azikiwe University, Awka.	N6.33; E7.00; 23.0m	AWOM UPH PM 020	PHSN 002-14	April 15, 2014

b) DNA extraction and PCR amplification

Total genomic DNAs were extracted from young, dry leaves of *Phyllanthus* specimens using the CTAB method by Doyle & Doyle (1987), because of high amount of polysaccharides present; the CTAB concentration was adapted to 3% (Doyle & Doyle 1990). The DNA pellet was then diluted into 50 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). An additional phenol/chloroform purification step was performed to remove proteins and potentially interfering secondary metabolites as follows: 150 µl of TE was added to bring the initial volume to 200 µl. Equal volume of phenol was then added, vortex gently and spin for 2 min at 13 000 rpm at room temperature. It was followed by addition of 100 µl of Sevag (chloroform-isoamyl alcohol 96:4), centrifugation at 13 000 rpm at room temperature of the collected aqueous steps (repeated twice). DNA was precipitated by incubating with 2/3 100% EtOH-3M NaOAc (25:1) at -20°C for 24 hours, then centrifuged for 2 min at 13 000 rpm at 4°C. The DNA pellet was washed with 70% EtOH, centrifuged at 13 000 rpm for 5 min, air dried in the dark and finally re-suspend in 50 µl TE.

The plastid regions, rbcL and matK (Table 2) were amplified using either direct PCR (Finnzymes Direct PCR kit) or iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories, CA) according to the manufacturer's protocol. iProof GC buffer (Bio-Rad) was used when nuclear phylogenetic gene amplification with HF buffer (Bio-Rad) did not provide satisfactory results. A finally concentration of 5% DMSO was added in 20 µl PCR reaction. In few cases direct PCR was done on dry leaves or on fresh leaves stored at -20°C.

The PCR amplifications were performed using a G storm PCR system 9700 (Perkin-Elmer) with the following thermal cycle conditions unless stated otherwise: DNA was initially denatured at 98°C for 5 minutes followed by 40 cycles at 98°C for 5 seconds, primer annealing at 55.8°C (for rbcL) and 55.5°C (for matK) for 5 seconds and elongation at 72°C for 20 seconds with a final 1 minute elongation at 72°C. The PCR fragment lengths were determined on a 1% agarose gel (see plate 2 and 3)

c) DNA Sequencing

PCR products were sequenced directly after 1:5 diluting with dH_2O. Amplified regions were sequenced in both directions with an automated sequencer 3730 Genetic Analyser Applied Biosystems v1.1/3.1 Cycle Sequencing Kit, according to the protocol provided with few modifications. Briefly: the component and volumes for the sequencing PCR reactions were: 1 µl of 5x sequencing buffer, 0.5 µl premix (Applied Biosystems), 3 µl of 10 µM primer, 3 µl dH_2O, 5% DMSO and 2 µl PCR product were used. Cycle sequencing steps were as follows: initial denaturation at 96°C for 1 min, following by 25x cycles of 96°C for 10 seconds (with a ramp seed of $3°Cs^{-1}$), 48°C for 15 seconds, 60°C for 4 min; 72°C for 1 min.

Table 2 : Nucleotide sequence data of the primers used for the amplification of the regions.

Region	Primer name	Primer sequence 5'-3'	Tm (°C)	MW	GC(%)	Length
matK	matK-472F	CCCRTYCATCTGGAAATCTTG	57.8	7583.5	42.86	21
	matK-1248R	GCTRTRATAATGAGAAAGATT	52.5	8033.3	23.81	21
rbcL	rbcLa-F	ATGTCACCACAAACAGAGACT	57.2	7950.3	42.86	21
	rbcLa-R	GTAAAATGTAGTCCACCRCG	52.8	6087.0	40.00	20

d) Cleanup

Extension products were adjusted to 20 µl by adding 10 µl of dH_2O, precipitated with 5 µl of 125 mM EDTA, and 60 µl of absolute EtOH for 15 min and centrifuged for 15 min at 20 000 g. For pellet purification, 60 µl of 70% EtOH was added after

removal of supernatant, centrifuged for 5 min at 20 000 g at 4°C. Tubes were then dried at 55°C for 5 min.

e) Sequence Alignment and Data Analysis

The data sequences of the rbcL and matK regions were aligned respectively, followed by manual adjustment and trimming at the ambiguous ends. The software programme Geneious Pro 7.0.4 (Biomatters, Ltd., http://www.geneious.com) with the default alignment parameters was used to view, assemble and edit the sequence trace files. Consensus sequences were aligned with the MUSCLE plug-in in Geneious R7. The post-trimmed lengths were at least 80% of the original read length. Sequences which covered more than 70% overlap between the forward and reverse sequences were considered. Other statistical values for each gene region such as the composition, GC content, % pairwise residue, % identical alignment, the minimum, maximum and mean lengths for the sequences as well as the pairwise number were also obtained in Geneious R7. Consensus sequences were compared on GenBank nucleotide Basic Local Alignment Search Tool (BLAST) in the National Centre for Biotechnology Information website (www.ncbi.nlm.nih.gov/Blast.cgi) and Boldsystems (http://www.boldsystems.org) for species identity validation.

f) DNA Barcoding Method

The project was registered under the workbench portal on the website (http://www.boldsystems.org) which is the software for DNA barcoding. Information on the title, code, description, campaign and markers were provided. This required integration of a field information management system (FIMS), laboratory information management system (LIMS), sequence analysis tools, workflow tracking to connect field data and laboratory data, database submission tools and pipeline automation for scaling up to eco-system scale projects. The genetic information were obtained from the concatenation of the rbcL and matK chloroplast genes. Then after, the details of the specimen data (such as species name, voucher data, collection record, identifier and primer information), images, sequences and trace files will be uploaded by following the formal submission guidelines and instructions on the website. The deposited sequences will automatically be tested for incongruence by the website software and validated by the database administrators before publication.

III. Results

Fig 1. shows the map of the various areas in which the samples were collected in Nigeria. From the information below, it could be gathered southern zone of Nigeria have abundance of these species. Virtually, all the states within southern zones of the country served as a sample area for the collection of the species. This does not actually apply that the species are endemic to this area or that they do not exist in other areas of the country. The map presented herein shows the states in the country that fall within the bounding box coordinates: upper left – 6.3333, 7; lower right – 4.75, 6.8333. The states included: Abia, Akwa Ibom, Anambra, Bayelsa, Cross River, Ebonyi, Enugu, Imo and Rivers. Mature plants of these five species, *P. amarus, P. urinaria, P. niruroides, P. odontadenius* and *P. muellerianus.* exist within these sites and location as mentioned in Table 1 above.

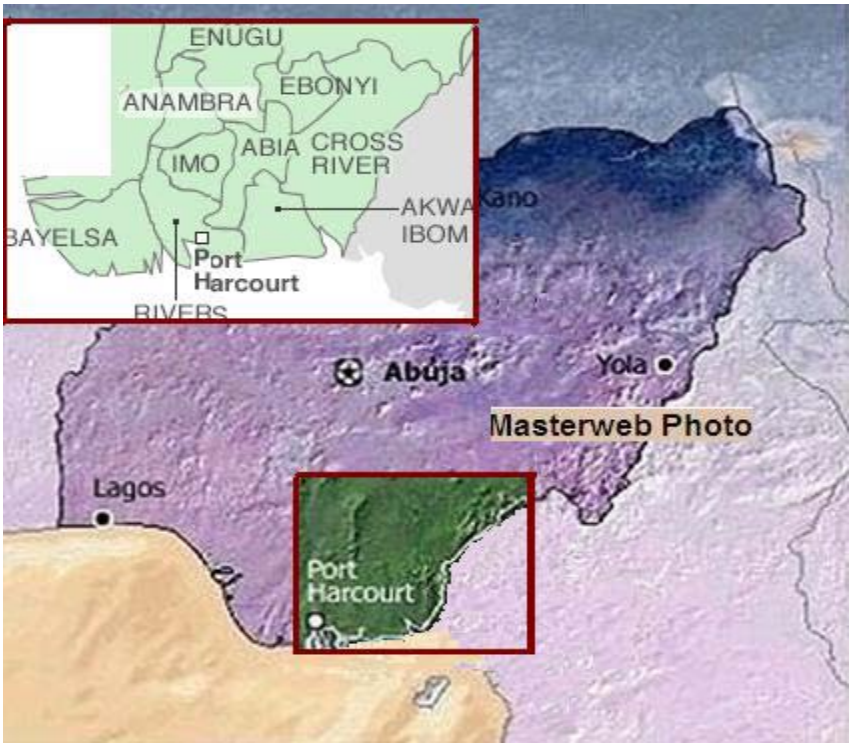

Figure 1 : Map of Nigeria showing the collection sites of the species of *Phyllanthus* used in the study


2b.i


2b.ii


2c.i


2c.ii


2d.i


2d.ii

Figure 2 : Photographs of the morphology of the genus *Phyllanthus* studied.

(a.i) *P. amarus*, showing the dorsal surface of the herb and the elliptic-oblong shape of the leaves.

(a.ii) *P. amarus*, exposing the greenish fruits and the pentatepalous flowers on the ventral surface of the leaves.

(b.i) *P. urinaria*, showing the dorsal surface of the herb and the mucronate leaf apices .

(b.ii) *P. urinaria*, exposing the reddish-green friuts, stipules and buds on the ventral surface of the leaves.

(c.i) *P. odontadenius*, showing the dorsal surface of the herb and the oblong shape of the leaves.

(c.ii) *P. odontadenius*, exposing the whitish-green tepals and reddish laterally free stipules.

(d.i) *P. niruroides*, showing the dorsal surface of the herb and the mucronate leaf apices.

(d.ii) *P. niruroides*, exposing the greenish fruits and tepals and reddish laterally free stipules.

(e.i) *P. muellerianus*, showing the dorsal surface of the shrub branch and the ovate-elliptic shape of the leaves.

(e.ii) *P. muellerianus*, exposing the spiny stipules of the leaves and thorny branches.

In this study, DNA quality was assayed by gel electrophoresis and intense bands were seen on 0.8% agarose. Only duplicate samples of each species were selected for sequencing. The thick band showed the PCR amplified rbcL (Plate 1) and matK (plate 2) regions which ranged from 500bp – 600bp in the respective samples. The leaf materials used were mainly dry leaves stored in silica gel. The samples contained both old and young leaves. The older and harder leaves had a higher level of secondary compounds which could act as PCR inhibitory components. The young leaves of *Phyllanthus* yielded a better quality DNA than the older, outer leaves of the plants.

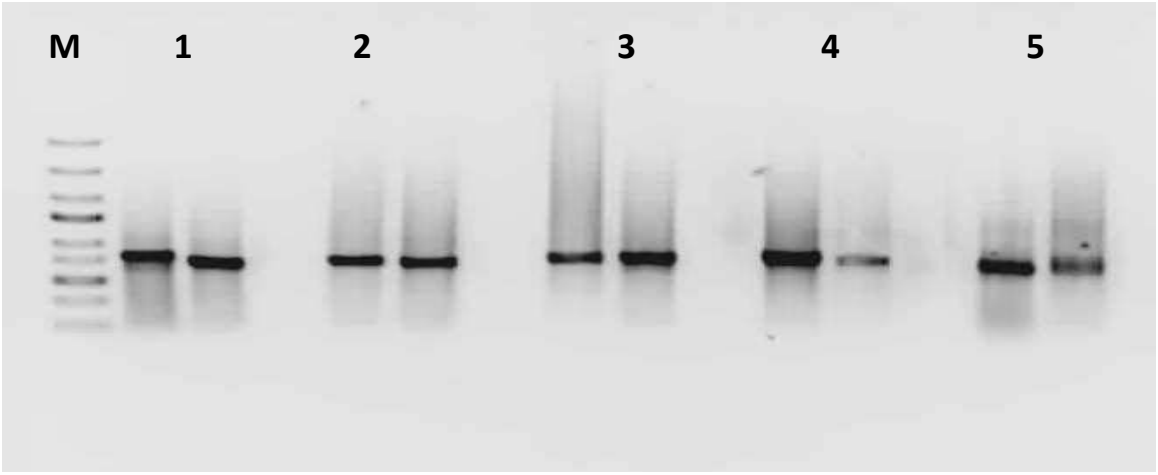

Plate 1 : PCR profile of the amplified rbcL gene region of the leaf of the five *Phyllanthus* species investigated. DNA fragment size range from 500bp – 600bp. Each species is represented by a duplicate of the samples. 1 – *P. amarus*, 2 – *P. muellerianus*, 3 – *P. niruroides*, 4 – *P. odontadenius*, 5 – *P. urinaria*

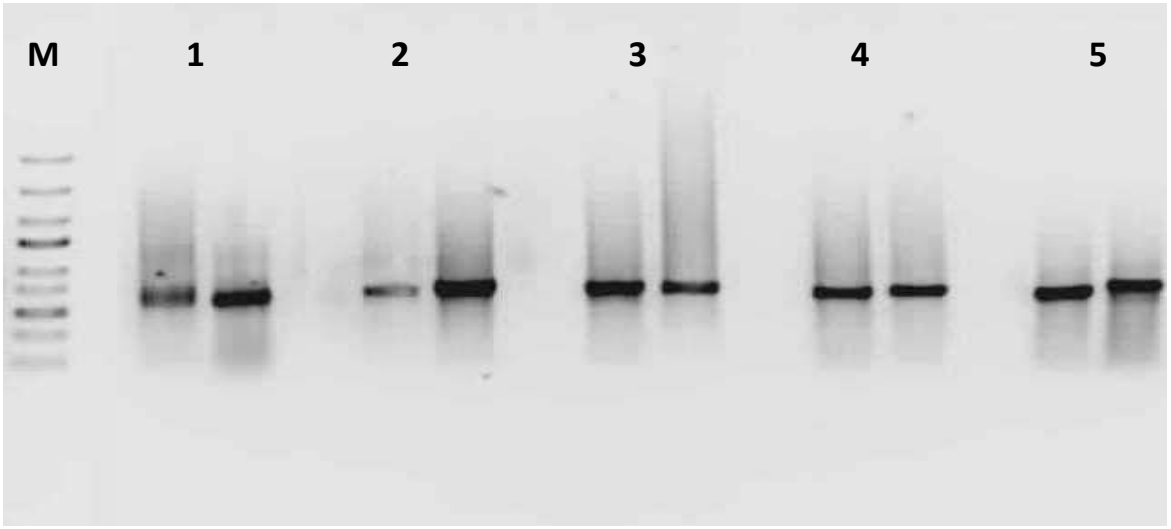

Plate 2 : PCR profile of the amplified matK gene region of the leaf of the five *Phyllanthus* species investigated. DNA fragment size range from 500bp – 600bp. Each species is represented by a duplicate of the samples. 1 – *P. amarus*, 2 – *P. muellerianus*, 3 – *P. niruroides*, 4 – *P. odontadenius*, 5 – *P. urinaria*

When visualized on an agarose gel, the DNA from the young leaves had less or no smears. Even by adapting the extraction protocol by adding recommended chemicals (such as PVP and SDS) at different stages in the process, the secondary compounds (e.g. polysaccharides and polyphenolic compounds) could not be removed completely. The polysaccharides co-precipitate with the DNA (Kumari *et al.*, 2012), resulting in a viscous solution. Some leaf samples, where the direct PCR approach was followed, seemed to contain PCR inhibitory components, and were excluded from the study after 3-4 amplification attempts. For future studies on *Phyllanthus* it is suggested that the young leaves are used for DNA extraction, or alternatively that the method of Kumari *et al.* (2012) be implemented as an extraction protocol.

The aligned rbcL matrix consisted of 475 base pairs (bp), 95.7% pairwise residue, 90.1% identical alignment, 475 bp maximum sequence and 475 bp minimum sequences while the aligned matK matrix consisted of 481 base pairs (bp), 88.3% pairwise residue, 74.4% identical alignment, 469 bp maximum sequences and 360 bp minimum sequences (Table 3).

Table 3 : Matrix and statistics of the rbcL and matK data sets in the study

Sampling Information	rbcL	matK	Description
Sequence Length (bp)	475	481	Residue length of sequence.
Number of species sequences	5	5	Number of sequences in an alignment or set of sequences.
Pairwise residue (%)	95.7	88.3	Percentage of parwise residues that is identical in the alignment, including gap vs non- gap residue, but excluding gap vs gap residue.
Identical alignment (%)	90.1	74.4	Percentage of columns in the alignment for which all sequences are identical.
Maximum sequence (bp)	475	469	Maximum length of non-reference sequences in a sequence list of or alignment.
Minimum sequence (bp)	475	360	Minimum length of non-reference sequences in sequence list or alignment.

The information on the composition of the various nucleotide bases of the sequences obtained from the samples as amplified from the primer gene markers were analyzed and summarized in the Table 4 and Table 5. The total number of sequence length per species was also recorded. The record was generated from BoldSystems during the analysis and submission.

Table 4 : Summary of the composition of the nucleotide bases from rbcL sequence data

TAXA	NUCLEOTIDE COMPOSITION	TOTAL SEQUENCE LENGTH (bp)
P. amarus	A (169), G (117), C (133), T (154)	573
P. muellerianus	A (150), G (141), C (132), T (166)	589
P. urinaria	A (169), G (118), C (130), T (156)	573
P. odontadenius	A (156), G (133), C (117), T (169)	546
P. niruroides	A (169), G (118), C (131), T (152)	570

Table 5 : Summary of the composition of the nucleotide bases from matK sequence data

TAXA	NUCLEOTIDE COMPOSITION	TOTAL SEQUENCE LENGTH (bp)
P. amarus	A (257), G (118), C (98), T (200)	673
P. muellerianus	A (163), G (88), C (90), T (203)	544
P. urinaria	A (151), G (88), C (93), T (186)	518
P. odontadenius	A (178), G (93), C (86), T (172)	529
P. niruroides	A (153), G (88), C (81), T (188)	518

During the research in BOLD, the different sequences were compared and the differences were represented in a form that resembles barcodes in which each colour represents one of the nucleotide bases of DNA [Adenine (A), Thymine (T), Cytosine (C) or Guanine (G)]. Like a barcode for products in a shop, each species of the *Phyllanthus* investigated had unique nucleotide barcoding pattern in BOLD and were identifiable from this unique pattern. Other *Phyllanthus* barcodes submitted on BOLD were compared with our accessions and only *P. amarus* and *P. urinaria* had previous records of other authors. *P. muellerianus, P. niruroides* and *P. odontadenius* had no previous submitted sequence records on BOLD and therefore were the first time to be registered and submitted on the website (http://www.boldsystems.org). Table 6 and 7 displays the information on the link of the accessions published on BOLD with the illustrative barcodes being developed in the linked website for each taxon.

Table 6 : Information and links of the *Phyllanthus* species accessions submitted on BOLD Systems for rbcL gene region

TAXA	PROCESS ID	SEQUENCE ID	LINK/REFERENCE
P. amarus	PHSN 001-14	PHSN 001-14.rbcLa	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN001-14
P. muellerianus	PHSN 002-14	PHSN 002-14.rbcLa	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN002-14
P. urinaria	PHSN 003-14	PHSN 003-14.rbcLa	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN003-14
P. odontadenius	PHSN 006-14	PHSN 006-14.rbcLa	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN006-14
P. niruroides	PHSN 005-14	PHSN 005-14.rbcLa	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN005-14

Table 7 : Information and links of the *Phyllanthus* species accessions submitted on BOLD Systems for matK gene region

TAXA	PROCESS ID	SEQUENCE ID	LINK/REFERENCE
P. amarus	PHSN 001-14	PHSN 001-14.matK	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN001-14
P. muellerianus	PHSN 002-14	PHSN 002-14.matK	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN002-14
P. urinaria	PHSN 003-14	PHSN 003-14.matK	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN003-14
P. odontadenius	PHSN 006-14	PHSN 006-14.matK	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN006-14
P. niruroides	PHSN 005-14	PHSN 005-14.matK	http://www.boldsystems.org/index.php/Public_RecordView?processid=PHSN005-14

IV. Discussion

Barcoding, identification and validation of *Phyllanthus* species using the concatenation of the plastid Ribulose bisphosphate carboxylase large chain (rbcL) and Maturase K (matK) chloroplast genes sequences are seen as a promising tool for the authentication plant species and ensuring better quality herbs and pharmaceuticals. This is otherwise known as DNA fingerprinting apart from identifying alterations in the genotypes of plant species, is also used for the betterment of drug-yield by tissue culturing and of which *Phyllanthus* is certainly one of the most useful herbs in the world. Barcoding results showed that rbcL and matK barcode regions can be used to identify the various species. Barcoding can, with 100% certainty, be used to distinguish between natural growing *Phyllanthus* species in Nigeria. The majority of *Phyllanthus* species may be easily and correctly identified based on morphological traits, but in the vegetative state identification by an inexperienced person become problematic, especially since some of the leaf-characteristics overlap between species. One such example is the nature of the leaf apices, that are characteristic of *P. urinaria* and *P. niruroides*, but that may not only also be absent in both species but also sometimes occur in other species. The leaf apices of both these species are mucronate shaped and as well have reddish, free stipules. These two species, as well as many of the other species, can very easily be misidentified, especially at the seedling stage. Also, it is very important to note that *P. urinaria* have been recorded to be used as fish poison in India, Indo-China, Malaysia and Indonesia (Awomukwu *et al*., 2014; Burkill, 1994) while *P. niruroides* has recorded no toxic effect but medicinal. For this reason it is necessary to identify the appropriate DNA-barcoding regions in *Phyllanthus*, to explore the most appropriate analysis method and to create a proper database for the genus.

From our research study, it was observed that some researchers have published *P. amarus* as *P. niruri* in BoldSystems. Most published sequences were actually mined from GenBank. This might further create confusion in validation and proper nomenclature of the species if not revealed early enough. Also, the results from BLAST on GenBank and BoldSystems showed that *P. amarus* and *P. niruri* may have been ignorantly misquoted or misidentified as separate taxa by different researchers in various parts of the world as regards to its taxonomic relevance. Moreover, from the morphological point of view, *P. amarus* shared similar characteristics with *P. niruri* than any other species in the genus (www.tropilab.com/black-cat.html). It therefore proposed that *P. niruri* should be subsumed to *P. amarus* due to its overall 100% similarity in the DNA sequence data based on the rbcL and matK gene regions. These species are not separate taxa but rather one species of the same genus. An authentic barcode data is highly essential in taxonomic identification and validation. This will help clear confusion encountered in morphological identification of species when plant species are ambiguously identified. The close morphological similarities of members of the genus *Phyllanthus* in Nigeria often times lead to confusion in species identification and collection. Existing taxonomic on the genus in the country is based on morphological characters documented in Hutchinson and Dalziel (1963), which reported some indeterminate species in which *P. amarus* and *P. niruri* were enlisted as separate species according the West African Flora. Reassessment of the members of this genus is therefore worthwhile. We have begun this by subjecting species within the southern region of the country to DNA characterization and barcoding. Other *Phyllanthus* barcodes submitted on BOLD were compared with our accessions and only *P. amarus* and *P. urinaria* had previous records of other authors. *P. muellerianus, P. niruroides* and *P. odontadenius* had no previous submitted sequence records on BOLD and therefore were the first time to be registered and submitted on the website (http://www.boldsystems.org).

DNA of interest can be stored as germplasm, which is then used for future cultivation (Henry, 2001).

In addition, germplasm can be used for the conservation of selected plant species, which are endangered. Although morphology, anatomy and phytochemistry can be used as useful tools in biosystematics but factors such as soil, climate and adaptability dictate the viability of a particular species and subsequently its contents but such discrepancies cannot be tolerated by DNA barcoding and fingerprinting technique. DNA barcoding has been found to be useful in identifying species, which may have morphological or anatomical difficulties.

It is important to reemphasize the fact that current information is lacking on members of the genus *Phyllanthus* in Nigeria. The efficiency of a two-barcode should be investigated but due to the size of the genus, it is likely that a three or more gene barcode will be necessary. The concatenation of the rbcL and matK chloroplast genes had always been the desirable choice of locus in BoldSystems (Paul *et al*., 2003). Currently in Nigeria, there is need for taxonomists to employ the system of molecular identification of every plant in order to establish a standard genetic library. This will help to curtail the confusion created by ambiguous morphological identification. Moreover, it will also aid to review obsolete literature in the field of taxonomy because recent discoveries will eradicate wrong information on overlapping species identification, nomenclature and classification.

V. Conclusion

The rbcL and matK chloroplast regions are very good barcoding candidate to distinguish the genus *Phyllanthus* in Nigeria because the genus has remained an interesting and challenging one. In addition, the gene markers, rbcL and matK have proved useful in identifying unknown species in the genus. With this type of analysis, identification will never be a problem because it can effectively be used to query sample belonging to any one of the properly delimited species, i.e. *P. muellerianus*, *P. odontadenius, P. niruroides*, *P. urinaria* or *P. amarus* and should be equally effective as an identification tool as the tree-based method. Other genera in the Nigerian Flora is worthy of investigation as only few records exist at the moment. Knowledge of this information will help guide other scientists in the pharmaceutical industries and herbal practitioners who may wish to investigate these plants for further research.

References Références Referencias

1. Awomukwu, D. A., Nyananyo, B. L., Onukwube, N. D., Uka, C. J., Okeke, C. U. and Ikpeama, A. I. (2014). Comparative Phytochemical Constituents and Pharmacognistic Importance of the Vegetative Organs of Some *Phyllanthus* species in South Eastern Nigeria. *International Journal of Modern Botany,* 4(2):29-39.
2. Awomukwu, D. A., Nyananyo, B. L., Spies. P. and Sizani, B. L. (2015). Identification, Validation and Classification of the genus *Phyllanthus* in Nigeria using ITS Genetic Marker and the Taxonomic Implication. *International Journal of Genetics and Genomics* 3(1): 1-7.
3. Bancilhon, L. (1971). Contribution a´ l'e´tude taxonomique du genre *Phyllanthus* (Euphorbiace´es). *Boissiera* 18: 9–81.
4. Blaxter, M. L. (2004). The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359: 669–679.
5. CBOL Plant Working Group. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, USA, 106: 12794-12797
6. Doyle, J. J. and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19: 11115.
7. Doyle, J. J. and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12: 13–15.
8. Edeoga, H. O., Omosun, G., Osuagwu, G., Awomukwu, D. A. (2007). Anatomical Features of Vegetative Organs of *Phyllanthus* species . *Journal of Environment and Ecology*, 25(3):548-553
9. Geneious 7.0.4 (Biomatters, Ltd.,http://www.geneious.com)
10. Genbank (http://www.ncbi.nlm.nih.gov/Blast.cgi)
11. Henry, R. J. (2001). *Plant genotyping: The DNA fingerprinting of plants*, CABI Publishing, New York. http://www.boldsystems.org/
12. Hutchinson, J. and Dalziel, M. D. (1963). *Flora of West Tropical Africa*, vol. 2. Crown Agents, London, UK.
13. Kathriarachchi, H., Hoffmann, P., Samuel, R., Wurdack, K. J. and Chase, M. W. (2005). Molecular phylogenetics of Phyllanthaceae inferred from five genes (plastid atpB, matK, 30ndhF, rbcL, and nuclear PHYC). *Molecular Phylogenetics and Evolution*, 36: 112–134.
14. 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: e508.
15. Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A. & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* 102: 8369.
16. Ko¨ hler, E. (1965). Die Pollenmorphologie der biovulaten Euphorbiaceae und ihre Bedeutung fu¨ r die Taxonomie. *Grana Palynologica*, 6: 26–120.

17. Ko¨ hler, E. (1967). U¨ ber Beziehungen zwischen Pollenmorphologie and Polyploidiestufen im Verwandtschaftsbereich der Gattung *Phyllanthus* (Euphorbiaceae). *Feddes Repertorium Specierum Novarum Regni Vegetabilis,* 74: 159–165.
18. Kumari, V., Bansal, A., Aminedi, R., Taneja, D. & Das, N. (2012). Simplified extraction of good quality genomic DNA from a variety of plant materials. *African Journal of Biotechnology* 11: 6420–6427.
19. Paul, D. N., Herbert, A. C., Shelly, L. B. and Jeremy, R. D. (2003) Biological Identification through DNA barcodes. *Proceedings of the Royal Society B*. 270: 313-321.
20. Punt, W. (1967). Pollen morphology of the genus *Phyllanthus* (Euphorbiaceae). *Review of Palaeobotany and Palynology,* 3: 141–150.
21. Punt, W. (1987). A survey of pollen morphology in Euphorbiaceae with special reference to *Phyllanthus*. *Botanical Journal of the Linnean Society,* 94: 127–142.
22. Rahman, M. O. (2007). DNA fingerprinting in *Utricularia* L. Section Utricularia as revealed by PCR based assay. *International Journal of Botany*, 3(1): 56-68.
23. Sagun, V. G. and Van der Ham, R. W. J. M. (2003). Pollen morphology of the Flueggeinae (Euphorbiaceae, Phyllanthoideae). *Grana*, 42: 193–219.
24. Samuel, R., Kathriarachchi, H., Hoffmann, P., Barfuss, M. H. J., Wurdack, K. J., Davis, C. C. and Chase, M. W., (2005). Molecular phylogenetics of Phyllanthaceae: evidence from plastid matK and nuclear PHYC sequences. *American Journal of Botany,* 92: 132–141.
25. Stoeckle, M., Janzen, D., Hallwachs, W., Hanken, J. & Baker, J. (2003). *Taxonomy, DNA, and the Barcode of Life*. Draft Conference Report, Banbury Center, Cold Spring Harbor Laboratory, New York, NY.
26. Stoeckle, M., Waggoner, P. & Ausubel, J. (2004). *Consortium for the Barcode of Life, Inaugural Meeting*. Draft, Smithsonian Institution, Washington DC.
27. Sucher, N. J. and Charles, M. C. (2008). Genome-Based approaches to the authentication of medicinal plants. *Planta Med*., 74(6): 603-623.
28. Uka, C. J., Okeke, C. U., Awomukwu, D. A., Aziagba, B. and Muoka, R. (2014). Taxonomic Significance of Foliar Epidermis of some *Phyllanthus* species in South Eastern Nigeria. *IOSR Journal of Pharmacy and Biological Sciences* (IOSR-JPBS), 9(4): 1-6.
29. Webster, G. L. (1956). A monographic study of the West Indian species of *Phyllanthus*. *Journal of the Arnold Arboretum* 37: 91–122, 217–268, 340–359.
30. Webster, G. L. and Carpenter, K. J. (2002). Pollen morphology and phylogenetic relationships in neotropical *Phyllanthus* (Euphorbiaceae). *Botanical Journal of the Linnean Society,* 138: 325–338.
31. Webster, G. L. and Ellis, J. R. (1962). Cytotaxonomic studies in the Euphorbiaceae subtribe Phyllanthinae. *American Journal of Botany*, 49: 14–18.
32. Weisling, K., Nybom, H., Wolff, K. and Kahl, G. (2005). *DNA fingerprinting in plants: Principles, methods and applications*. 2nd ed. Boca Raton: CRC Publisher.
33. Wurdack, K. J., Hoffmann, P., Samuel, R., De Bruijn, A. Y., Van Der Bank, M. and Chase, M. W. (2004). Molecular phylogenetic analysis of Phyllanthaceae (Phyllanthoideae pro parte, Euphorbiaceae sensu lato) using plastid rbcL DNA sequences. *American Journal of Botany,* 91:1882–1900. www.tropilab.com/black-cat.html