TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2017) 41: 455-463
© TÜBİTAK
doi:10.3906/bot-1610-9

Evaluating the utility of *rbcL* in identifying Nigerian species of the genus *Afzelia* Sm. (Fabaceae; Caesalpinioideae)

Aramide OSHINGBOYE*, George NODZA, Temitope ONUMINYA, Oluwatoyin OGUNDIPE
Molecular Systematics Laboratory, Department of Botany, University of Lagos, Akoka, Yaba Lagos, Nigeria

Received: 04.10.2016 • **Accepted/Published Online:** 07.07.2017 • **Final Version:** 28.09.2017

Abstract: The genus *Afzelia* Sm. (Fabaceae) represents one of the most useful tree species in tropical Africa. These trees have huge socioeconomic and ecological importance, but are under serious threat from incessant exploitation. This has caused a decline of the species in the wild; several species are now registered as threatened and endangered in the IUCN red list. However, conserving this species for posterity has been hampered by the complexity of morphological characters that exist among *Afzelia* species. In our study, we evaluated the utility of *rbcL* as a single barcode region, exploring its performance to accurately identify Nigerian *Afzelia* species and elucidate the relationships that exist between species. Our results revealed an overlap between intraspecific distance and interspecific divergence with an overall mean distance of 0.002. Likewise, identification based on all species barcode, best match, and best close match analysis of TaxonDNA depicted 46.15% and 57.14% of correct and ambiguous identification, respectively. Correct identification according to all species barcodes accurately identified 64.3% of data. Phylogenetic analysis based on UPGMA, NJ, MP, and BI analysis supports the monophyly of *Afzelia* with support values at 98%, 100%, 93.41%, and 1, respectively. It also revealed noncongruence among species of *A. africana* Pers. and *A. quanzensis* Welw., which suggests an unclear relationship between *A. africana* and its sister species. This implies *rbcL* is not an ideal plant barcode region for successful identification of Nigerian *Afzelia* species.

Key words: *Afzelia*, conservation, DNA barcodes, IUCN, plant identification

1. Introduction

The genus *Afzelia* Sm., named after Adam Afzelius, a Swedish botanist 1750–1837, (Orwa et al., 2009), belongs to the family Fabaceae. They are large deciduous trees, growing at maturity to a height of 18–35 m, distributed in the sub-Saharan and humid regions of Africa. *Afzelia* comprises about eleven species, seven of which occur in tropical Africa and four in Southeast Asia, closely related to *Intsia* Thouars (Gérard and Louppe, 2011). According to Kadiri and Olowokudejo (2008), there are six species in West Africa, four of which are well distributed in Nigeria: *Afzelia africana* Pers., *A. bella* Harms, *A. bipindensis* Harms, and *A. pachyloba* Harms (Ariwaodo and Harry-Asobara, 2015). Members of this genus have played numerous key roles in the sustainable livelihood of communities within this region: they are a good source of medicinal remedies (Burkill, 1985; Kadiri and Olowokudejo, 2008; Nacoulma et al., 2011), as well as for construction of buildings, canoes, and pestles and mortars, and used as firewood, for charcoal production, and dye preparation from the heartwood. The leaves of *A. africana* are important fodder and nitrogen sources for animals during the dry season in the agro-sylvopastoral system (Ouédraogo-Koné et al., 2008), and is the most sought after fodder tree in the entire sub-Saharan and humid region of Africa. Wood from this species is a good source of export trade. For example, Ghana exported 9000 m^3 of sawn *Afzelia* wood (as 'papao'), and 7000 m^3 in 2005 and 2006, respectively, at an average price of US$ 780 per m^3 (Gérard and Louppe, 2011). Despite the usefulness of these species, active conservation measures are hampered by incessant exploitation. This is moving apace due to the international marketability of the sawn wood (Donkpegan et al., 2015) and the uncertainty of species boundaries within the *Afzelia*, which could be sympatric or parapatric. The species are often not easily distinguishable by forest managers and logging companies. Even morphological characters used in distinguishing species are seldom useful in their identification (Donkpegan et al., 2014). This poses a great concern for conservation efforts, as most species are substantially decimated in many countries of the subhumid zone (Ouédraogo-Koné, 2008). Undoubtedly, these species are threatened by overexploitation, logging for timber and utensils, and high rate of deforestation.

* Correspondence: dlparamide2012@gmail.com

The seedlings are sensitive to fire, browsing, and drought (which retards their distribution and regeneration). Increasing cattle populations also threaten the species as areas are deforested for grazing (Volker and Ross, 2011) and herdsmen regularly lop *A. africana* to feed their cattle. As a result of these threats, *Afzelia* species are considered endangered (*A. africana*) (IUCN, 2011) and vulnerable (*A. pachyloba* and *A. bipidensis*) (Donkpegan et al., 2015).

There have been previous studies on the species such as comparative foliar epidermal morphology (Kadiri and Olowokudejo, 2008), impact of land use types on population structure and extent of bark and foliage harvest (Nacoulma, 2011), potentials of vegetable oil production (Igwenyi, 2011), regeneration and seedling survival constraints (Ouédraogo et al., 2006), characterization of nutritive value, chemical composition (Ouédraogo-Koné et al., 2008), medicinal properties (Akinpelu et al., 2008), and ethnobotanical usefulness of the species (Burkil, 1985). However, information on the genetic diversity of *Afzelia* is limited, especially within tropical Africa with the exception of Donkpegan et al. (2015). Therefore the need to evaluate species identification using the genetic marker *rbcL* to aid conservation management is imperative. In this study, we adopted the DNA barcoding techniques as the current and most reliable means of conservation. This technique has been in practice for over a decade; the idea has been gaining popularity among the scientific community to successfully and accurately identify species based on a globally accepted barcode region. For identification of animals, DNA barcoding has yielded great success and has proven a 98% success based on the mt*coxI* region (Hebert et al., 2003, 2004; Hogg et al., 2004; Barret et al., 2005). However, this approach has suffered a setback in plant identification as a single valid barcode region has not yet been found. This has prompted the barcoding community to discuss and pursue the need to identify an ideal universal barcode region for all land plants (Chase et al., 2005; Kress et al., 2005; Cowan et al., 2006; Newmaster et al., 2006; Kress and Erickson, 2007). Based on this, the CBOL (barcode of life) working group proposed the two loci *rbcL+matK* as a core barcode region. Since then, most studies have been on the combined locus; the variability and utility of a single barcode region remain unexplored. In our study we present the utility of the *rbcL* region (one of the proposed regions for plant barcoding) in its performance to accurately identify Nigerian *Afzelia* species and elucidate the relationships that exists among species.

2. Materials and methods

2.1. Sampling

Samples were collected from the wild and identified using keys from floras such as Hutchinson and Dalziel (1954), Keay (1989), and others obtained from the Forestry Herbarium Ibadan (FHI). Details of the source of samples are given in Table 1.

2.2. DNA extraction

Genomic DNA was extracted following the modified 2X CTAB protocol described by Doyle and Doyle (1987). This was isolated from approximately 0.0180 g of herbarium leaf samples, which were ground with 2 tungsten carbide beads in a 2-mL Eppendorf tube using a bead beater at 30 Hz for 45 s. Upon completion, the beads were removed and 800 µL of prewarmed 2X CTAB buffer was pipetted into the tubes and incubated for 1 h with occasional inversion at 5–10 min intervals. Afterwards, they were centrifuged at 13,000 rpm for 3 min and the aqueous upper phase was carefully transferred into a clean 1.5-mL Eppendorf tube; an equal volume of SEVAG was then added and the resulting mixture was mixed well to obtain an emulsion and centrifuged at 13,000 rpm for 5 min. The upper layer was again carefully transferred into another clean 1.5-mL Eppendorf tube and the chloroform extraction phase was repeated while the aqueous phase was transferred into a clean 1.5-mL screw cap tube. Upon centrifugation, 0.08 volumes of cold 7.5 M ammonium acetate and 0.54 volumes (using the combined volume of aqueous phase and added AmAc) of cold propan-2-ol were added, mixed, and kept in the freezer for 1 week to precipitate the DNA. Afterwards, the samples were centrifuged at 13,000 rpm for 15 min to form a pellet. The liquid was carefully poured

Table 1. Details of sample source.

S/N	Species	Collector	Locality
1.	*Afzelia africana*	B.O. Daramola 3978 LUH	Borgu Nigeria
2.	*A. bella*	Jones & Onochie 16716 FHI	Sapoba
3.	*A. bipindensis*	O.O Oyebanji 6011 LUH	Epe Lagoon
4.	*A. bracteata*	Hutchison & Dalziel 890 FHI	Liberia
5.	*A. pachyloba*	Ogba Kennedy 16116 LUH	Sapoba

out and 700 µL of cold 70% ethanol was then added to the tubes, which were mixed and allowed to stand for a few minutes. The tubes were centrifuged, the liquid was pipetted off, and the ethanol phase was repeated to remove the remaining contaminants. The dry pellets were dried in a CentriVap at 35 °C and samples were resuspended in 100 µL of TE buffer overnight. Resultant DNA samples were stored at –20 °C for subsequent use.

A polymerase chain reaction (PCR) was performed in 50-µL reaction mixtures containing 25 µL of biomix, 1 µL of BSA, 2 µL of DMSO, 1.75 µL of 10 µM of each primer, 17.5 µL of Millipore H_2O, and 1 µL of 30–50 ng template DNA. Primer sequences used were obtained from the protocols proposed by the CBOL plant working group with a forward sequence rbcL 1F "ATGTCACCACAAACAGAAAC" and reverse rbcL 724R "TCGCATGTACCTGCAGTAGC". Amplifications were run on a Veriti 96-well thermal cycler with profile initial denaturing at 96 °C for 50 s, 30 cycles of denaturation at 96 °C for 50 s, annealing at 53 °C for 50 s, and extension at 72 °C for 2 min, and a final extension at 72 °C for 7 min. Each PCR product was run on 1% agarose gel stained with ethidium bromide and successful products were sent to Source Bioscience (UK) for bidirectional sequencing using the same primers used in PCR.

2.3. Barcoding analysis

This utilized sequence data based on the *rbcL* region and available *rbcL Afzelia* sequences downloaded from GenBank. Raw sequences were assembled and edited using SeqMan II, one of the programs of the Lasergene software package (DNASTAR, Inc.), and verified by a Blastn search on GenBank. The edited sequences were aligned using the multiple alignment Clustal W algorithm (Thompson et al., 1994) as implemented in MegAlign (Lasergene, DNASTAR, Inc.), with further manual adjustment. All aligned sequences were trimmed to be of equal lengths to allow comparison within the dataset. Genetic distances were calculated using the Kimura two parameters (K2P) distance method as implemented in MEGA 5.0 (Tamura et al., 2011) and intra- and interspecific genetic distances for species that were represented by more than one individual were calculated. To assess the barcoding gap, the level of divergence between and within species was calculated by pairwise analysis of sequences. Minimum interspecific distance was plotted against the maximum intraspecific distance as recommended by CBOL. To estimate the utility and reliability of barcode regions for accurate species discrimination, the "all species barcodes", "best match", and "best close match" functions of TAXONDNA were used (Meier et al., 2006). This sequence comparison approach was performed on all the barcode data using uncorrected pairwise distance and a minimum 300-bp sequence overlap.

2.4. Tree-based analysis

Several tree-based methods were used to assess species-specific clusters. Neighbor joining (NJ) and UPGMA were conducted in MEGA 5.0 with the K2P model of nucleotide substitutions and node support of 1000 bootstrap replicates. Maximum parsimony was also conducted in PAUP with 100 random sequence addition and other default parameters; node support was obtained from heuristic searches of 1000 bootstrap replicates. Bayesian inference was conducted in Mr Bayes v3.1.2. MrModelTest was run to specify the best fit model nucleotide substitution with the Akaike Information criterion; the GTR model was chosen for the dataset (Nyalander, 2004). Two independent runs with four Markov chains (one cold and three heated) run for 1,000,000 generations and each tree sampled every 1000th generation were performed. The first 250,000 samples trees were discarded as 'burn in' while the remaining trees were used to build a 50% majority rule consensus tree with posterior probability for nodes. In tree-based analysis, genus identification was considered successful when the unknown sample formed a monophyletic group together with all members of a single genus, with bootstrap support of >70%. An equal strategy was applied for species-level identification (de Groot et al., 2011).

3. Results

3.1. PCR amplifications and DNA sequencing

One of the challenges of isolating DNA from herbarium specimens is the ability to get good quality DNA usable for further molecular analysis; however, in our study good quality DNA of high molecular weight was isolated and used for amplifications as given in Table 2. Amplifications were successful with distinct bands and all samples were successfully sequenced with length of 701–768 bp with no indels. *Afzelia* species sequences available from GenBank were downloaded and included for the analysis as given in the Appendix. Our dataset comprised 14 sequences in 5 species characterized by 663 constant sites, 83 variable sites, and 54 parsimony informative sites.

3.2. Intraspecific and interspecific divergence

Estimation of genetic distance between and among species revealed a zero intraspecific distance within species while an overall mean distance of 0.002 was obtained between species. Our results reveal that the maximum intraspecific distance was equivalent to the minimum interspecific distance (Table 3). Distribution of interspecific divergence between congeneric species and intraspecific variation revealed an overlap between both distances and hence no barcoding gap within species (Figure 1).

3.3. Resolving power

Results of the best match and best close match of sequences performed on TAXONDNA are shown in Table 4 while Table 5 has results of identification based on all species barcodes.

Table 2. Spectrophotometric check of extracted DNA.

S/no	Sample name	Code	A260/A280	Conc. (ng/µL)
1.	*Afzelia bipidensis*	H001	1.76	302.2
2.	*A. bella*	H009	1.88	284
3.	*A. pachyloba*	H010	1.66	323.5
4.	*A. pachyloba*	H011	1.42	526.7
5.	*A. africana*	H012	1.86	99
6.	*A. africana*	H012i	1.96	108
7.	*A. africana*	C025	1.85	405.2

Table 3. Analysis of interspecific and intraspecific divergence in *rbcL* sequences in *Afzelia*.

Species	Intraspecific distance (%)	Interspecific distance (%)
A. africana	0.000	0.001–0.004
A. bella	0.000	0.002–0.006
A. pachyloba	0.000	0.002–0.006
A. quanzensis	0.001	0.001–0.005

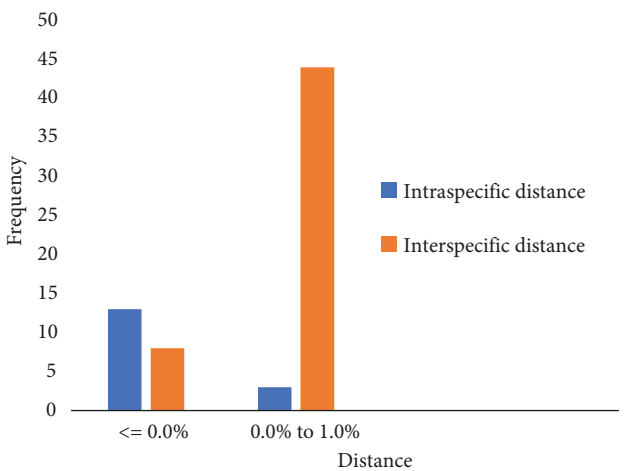

Figure 1. Distribution of interspecific divergence between congeneric species and intraspecific variation.

3.4. Tree-based analysis

Results obtained from tree-based analyses support the monophyly of the genus *Afzelia* with high support values (98%–100%). UPGMA analysis depicted a cluster of all species with their conspecifics at within 45%–98% bootstrap support with *A. africana* clustering with *A. quanzensis* at 42% support value (Figure 2). In contrast, NJ analysis (Figure 3) revealed a cluster of *A. bella* and *A. pachyloba* at distinct nodes with high support values while relationships among other *Afzelia* species are not well resolved. The maximum parsimony and Bayesian inference exhibited a congruent tree and a congruent topology (Figures 4–6), also elucidating a similar relationship among *Afzelia* species.

4. Discussion

One of the important criteria for an ideal DNA barcode region is efficient recovery of good quality sequence data (CBOL Plant Working Group, 2009); based on this, our results revealed a 100% amplification and sequencing success corroborating previous studies on the *rbcL* loci due to its applicability across plants, easy and unambiguous alignment, high primer universality, and high sequence quality (Dong et al., 2014). Based on the intraspecific and interspecific divergence, our results revealed between 0.000 and 0.001 intraspecific distance among species and their conspecifics and a mean interspecific distance of 0.002 ranging from 0.001 to 0.006; this depicts an overlap between intra- and interspecific divergence. Although the *rbcL* region has been used successfully in some studies (Meier et al., 2008; Bafeel et al., 2011, 2012; Dong et al., 2014) to accurately identify species, this region could not accurately identify our understudied species. This could be due to the low variability of the *rbcL* region (Gao et al., 2010). Hence *rbcL* is not a suitable barcode region to accurately identify *Afzelia* species.

Table 4. Identification success based on the best match and best close match.

Barcodes	No of sequences & species	Best match (%)			Best close match (%)			No match
		Correct	Ambiguous	Incorrect	Correct	Ambiguous	Incorrect	
rbcL	14 (5)	46.15	53.84	0	46.15	53.84	0	0

Table 5. Identification success based on all species barcodes.

Analysis	Values
Sequences	14
Sequences without valid species name	0
Sequences with at least one sequence with an overlap of 300 base pairs	14
Correct identification according to "All Species Barcodes"	9 (64.28%)
Ambiguous according to "All Species Barcodes"	4 (28.57%)
Incorrect identification according to "All Species Barcodes"	1 (7.14%)
Species with no match closer than 3%	0 (0%)

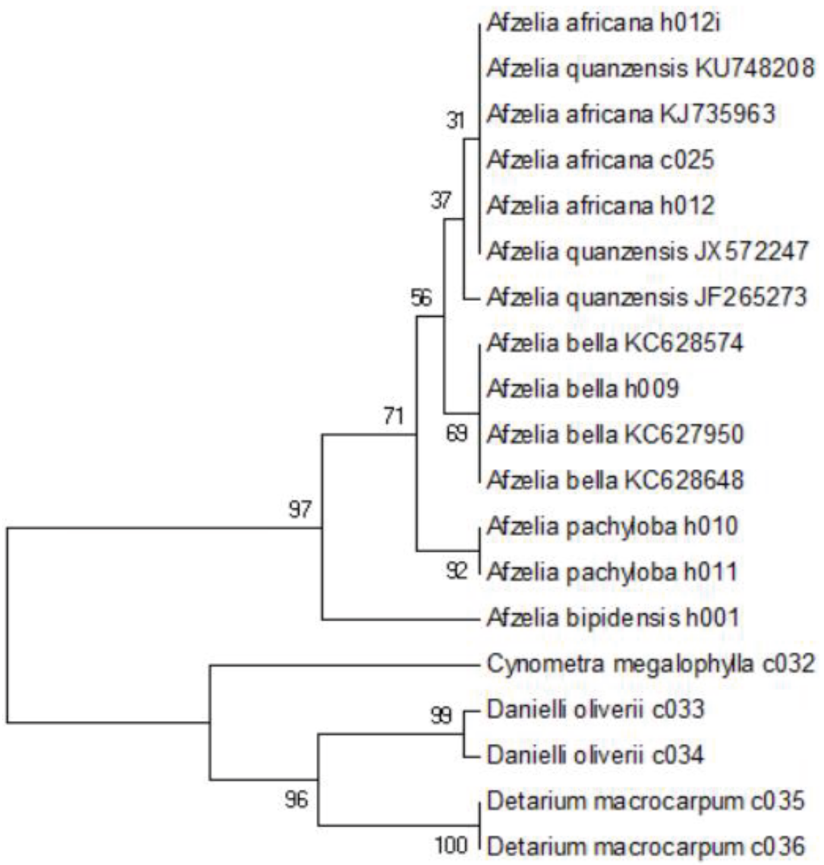

Figure 2. UPGMA cladogram based on K2P model of substitution; numbers at nodes indicate bootstrap support values.

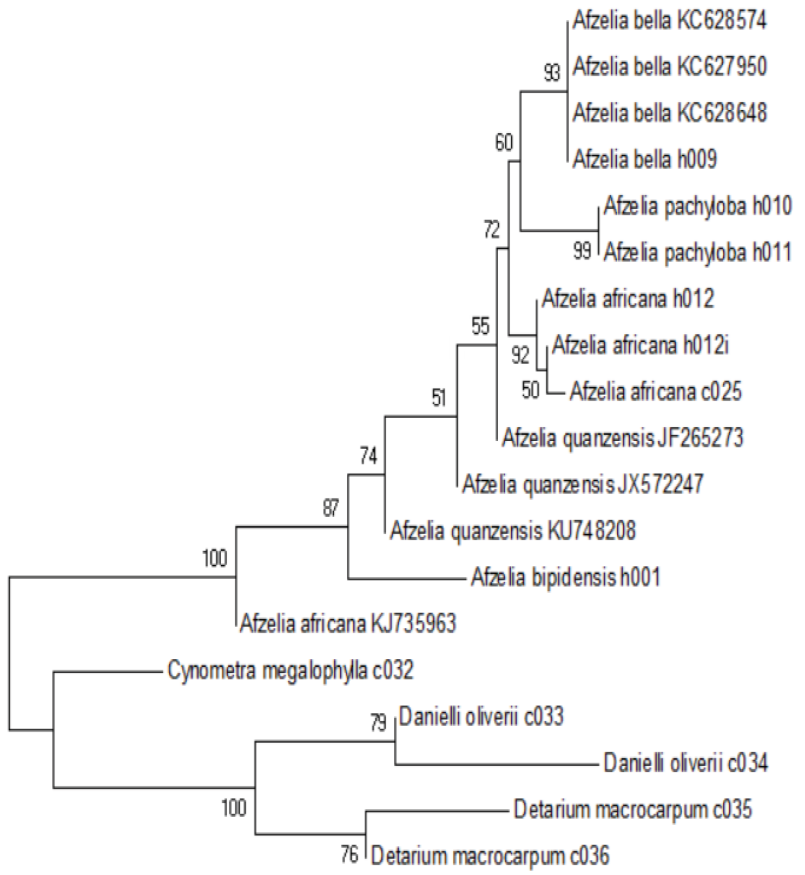

Figure 3. Neighbor joining cladogram based on K2P model of substitution; numbers at nodes indicate bootstrap support values.

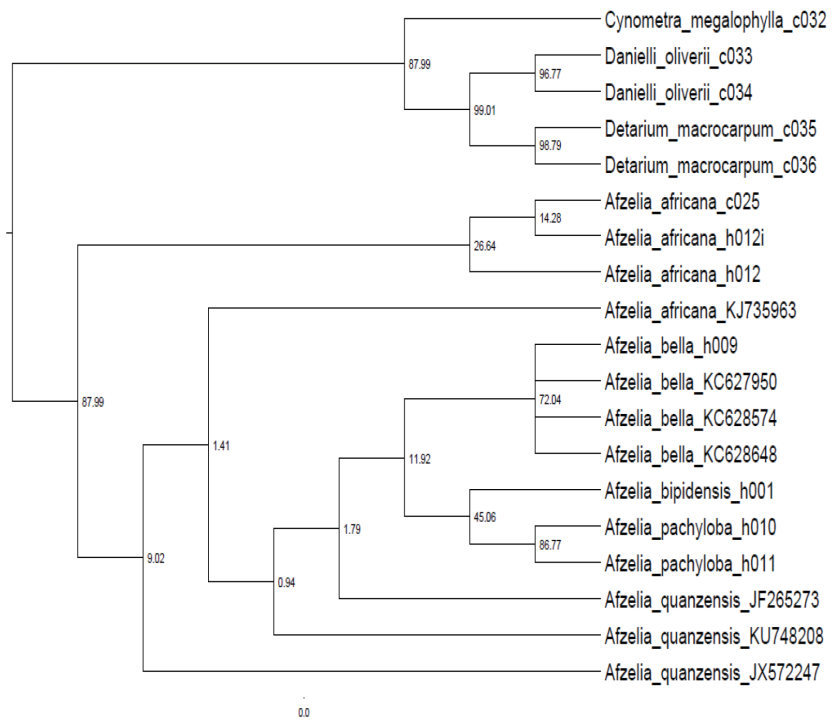

Figure 4. Maximum parsimony phylogram of a strict consensus tree; numbers at nodes indicate bootstrap support values.

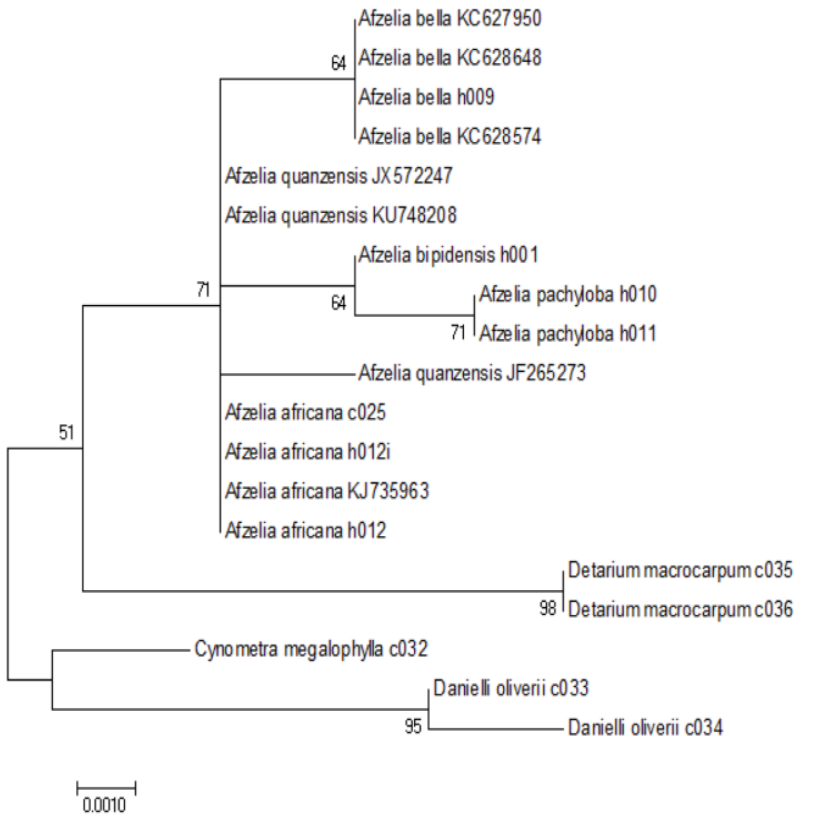

Figure 5. Maximum parsimony phylogram of the original tree; numbers at nodes indicate bootstrap support values.

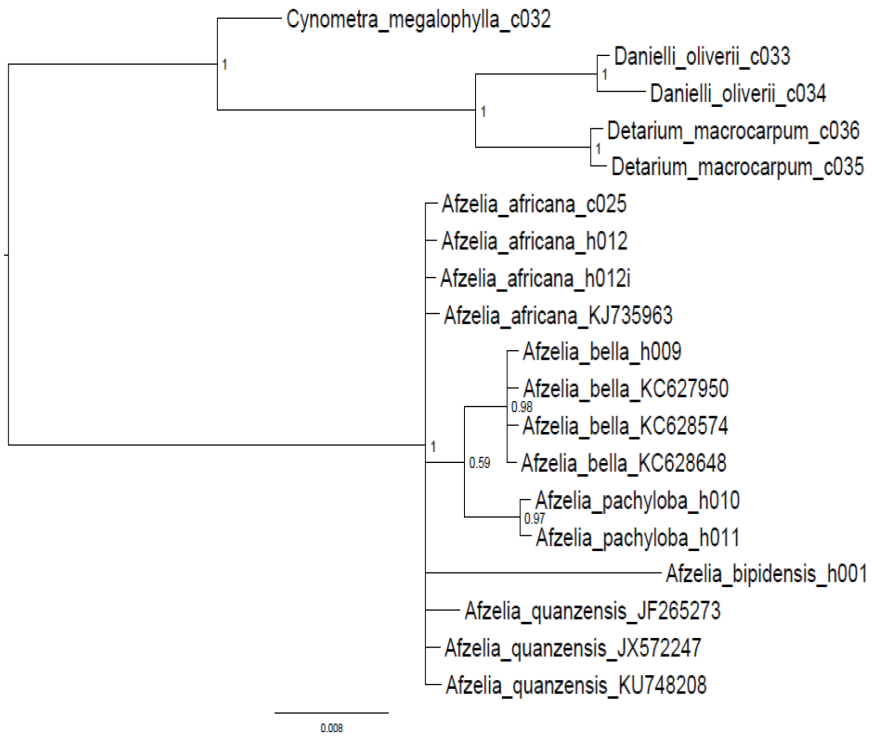

Figure 6. Fifty percent majority consensus phylogram inferred by Bayesian analysis based on GTR model of substitution; numbers at nodes indicate posterior probability value.

The results of TAXONDNA analysis depicted a higher percentage (57.14%) of ambiguous identification, owing to the near same genetic distance between *Afzelia africana* and *A. quanzensis* and likewise absence of conspecific sequences to *A. bipidensis*. However, *A. pachyloba* and *A. bella* were successfully distinguished due to their distances within the set 3% threshold frequency for an accurate identification. Likewise, the reliability of the *rbcL* region was further confirmed in all species barcode results, with a 64.28% correct identification; *A. quanzensis* was ambiguous while *A. bipidensis* was incorrectly identified due to the absence of a conspecific sequence. This further revealed the inaccurate resolving power of *rbcL* to discriminate species based on its low levels of species recovery (Burgess et al., 2011).

The tree-based criterion corroborated the previous results with clustering of species and their conspecifics at distinct nodes with high support values according to de Groot et al. (2011) except for *Afzelia africana* and *A. quanzensis*, which revealed noncluster of species at a distinct node or cluster at a distinct node but with a low support value (UPGMA). A phylogenetic result supports the monophyly of the genus *Afzelia* and a congruent topology for MP and BI analysis. All members of the genus were clustered and differentiated from the other closely related genera *Cynometra* L. and *Detarium* Juss., members of tribe Detarieae. Although relationships between *Afzelia africana* and *A. quanzensis* were not well resolved, a rapidly evolving gene would help to elucidate the relationships existing between species and reveal a wider range of distances between and among species.

Afzelia species are one of the most valuable economically important tree species and are mostly exploited for their international wood market. This poses threats of varying degree on *Afzelia* species. Due to this, the reliable identification and discrimination of species is pertinent and crucial for the control of this tradable plant resource, contributing to the protection and maintenance of *Afzelia* species diversity. In conclusion, our study highlighted the utility of the *rbcL* gene region in identifying Nigerian *Afzelia* species as a single barcode region, which can be utilized by forest managers and logging companies to help efficient and accurate identification of the Nigerian *Afzelia* species. We suggest assessing the *rbcL* region and other barcode regions to accurately identify all the African *Afzelia* species. In addition, other highly evolving gene regions should be evaluated to shed more light on the relationships among *Afzelia* species so as to provide adequate information for the conservation and utilization of these species.

References

Ariwaodo JO, Harry-Asobara JL (2015). Preliminary investigation on flowering and fruiting pattern in a plantation grown *Afzelia africana* Sm stand in Umuahia, Nigeria. American Journal of Plant Sciences 6: 219-227.

Bafeel SO, Ariff IA, Bakir MA, Khan HA, Al Farhan AH, Al Homaidan AA, Ahmad A, Thomas J (2011). Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant Omics J 4: 195-198.

Bafeel SO, Ariff IA, Bakir MA, Al Homaidan AA, Al Farhan AH, Khan HA (2012). DNA barcoding of arid wild plants using rbcL gene sequences. Genet Mol Res 11: 1934-1941.

Barrett SCH, Hebert PDN (2005). Identifying spiders through DNA barcodes. C J Zoolog 83: 481-491.

Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, Percy DM, Hajibabaei M, Barret SCH (2011). Discriminating plant species in a local temperate flora using the rbcL+matK DNA barcode. Methods in Ecology and Evolution 2: 333-340.

Burkill H (1985). The Useful Plants of West Tropical Africa. Volume 1. London, UK: Royal Botanic Gardens Kew.

CBOL Plant Working Group (2009). A DNA barcode for land plants. Proc Natl Acad Sci USA 106: 12794-12797.

Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, Savolainen V (2005). Land plants and DNA barcodes: short-term and long-term goals. Philos Trans R Soc Lond B Biol Sci 360: 1889-1895.

Cowan RS, Chase MW, Kress WJ, Savolainen V (2006). 300,000 species to identify : problems, progress, and prospects in DNA barcoding of land plants. Taxon 55: 611-616.

Donkpegan ASL, Hardy OJ, Lejeune P, Oumorou M, Dainou K, Doucet J (2014). Un complexe d'espèce d'Afzelia des forêts africaines d'intérêt économique et écologique (synthèse bibliographique). Biotech Agron Société et Environ 18: 233-246.

Donkpegan ASL, Doucet J, Dainou K, Hardy OJ (2015). Microsatellite development and flow cytometry in the African tree genus *Afzelia* (Fabaceae, Caesalpinioideae) reveal a polyploid complex. Appl Plant Sci 3: 1400097.

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11-15.

Gerard J, Louppe D (2011). *Afzelia bipindensis* Harms. In: Lemmens RHMJ, Louppe D, Oteng-Amoako AA editors. Plant Resources of Tropical Africa (PROTA). Wageningen, the Netherlands: CTA Publishing.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. Proc Biol Sci 270: 313-321.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004). Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci USA 101: 14812-14817.

Hogg ID, Hebert PDN (2004). Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. Can J Zoolog 82: 749-754.

International Union for Conservation of Nature (IUCN) (2001). Red list of threatened species. Website: http://www.iucnredlist.org. Assessed 26 March, 2015.

Kadiri AB, Olowokudejo JD (2008). Comparative foliar epidermal morphology of the West African species of the genus *Afzelia* Smith (Leguminosae: Caesalpinioideae). Gayana Botánica 65: 84-92.

Kress WJ, Wurdack KJ, Zimmer E, Weigt L, Janzen DH (2005). Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA 102: 8369-8374.

Kress WJ, Erickson DL (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.

Meier R, Kwong S, Gaurav V, Peter KLN (2006). DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. Syst Biol 55: 715-728.

Meier R, Guanyang Z, Farhan A (2008). The use of mean instead of smallest interspecific distances exaggerates the size of the 'barcoding gap' and leads to misidentification. Syst Biol 57: 809-813.

Newmaster SG, Fazekas AJ, Ragupathy S (2006). In a multigene tiered approach. Can J Botany 84: 335-341.

Nylander JAA (2004). MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Rudolf M, Shiyang K, Vaidya G, Ng PKL (2006). DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol 55: 715-728.

Swofford DL (2002). PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0b10 [computer program]. Sunderland (Massachusetts): Sinauer.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M Kumar S (2011). MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680.