Assessing the effectiveness of molecular data for species identification and diversity studies of Nigerian herpetofauna

Lotanna M. Nneji, Adeniyi C. Adeola, Agboola O. Okeyoyin, Ojo C. Oladipo, Yohanna Saidu, John Y. Usongo and Adiaha A. A. Ugwumba

State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China; National Park Service Headquarter, Abuja, Federal Capital Territory, Nigeria; Old Oyo National Park, Oyo, Nigeria; Gashaka Gumti National Park, Serti, Nigeria; Department of Zoology, Modibbo Adama University of Technology, Yola, Nigeria; Department of Zoology, University of Ibadan, Ibadan, Nigeria

ABSTRACT
We evaluated the effectiveness of molecular data (mitochondrial cytochrome c oxidase subunit I – COI and 16S ribosomal RNA – 16S) for species identification and diversity studies of Nigerian herpetofauna. We collected 189 specimens consisting of 91 amphibians and 98 reptiles from the largest and under-studied national park in Nigeria – Gashaka Gumti National Park (GGNP), eastern Nigeria. The resulting COI and 16S data allowed us to compare our sequences with others in global databases for species identification. During query searches, we observed poor representation of COI sequences for African herpetofauna in the GenBank and BOLD databases. Neighbour-joining (NJ) trees constructed based on COI and 16S datasets for all specimens sequenced for this study clustered most individuals of the same species. On the other hand, morphologically unidentified amphibian (Leptopelis sp and Ptychadena sp) and reptile (Hemidactylus sp and Lycophidion sp) species formed strongly supported lineages in the NJ trees, and were clearly separated from their sister species. Further, the NJ tree-based analyses recovered at least two well supported divergent lineages within Agama agama and Trachylepis maculilabris. Our study confirms the efficiency of integrating morphological and DNA barcoding approaches in identification and diversity studies of Nigerian herpetofauna. Finally, we recommend development of comprehensive and reliable DNA barcode reference libraries for global amphibians and reptiles.

1. Introduction
Biodiversity inventory and documentation are one of the leading concepts in conservation biology. However, this is hinged closely to accurate species identification. In last two decades, molecular data (mitochondrial DNA) has gained popularity as a complementary tool for species identification. This involves the use of mitochondrial cytochrome c oxidase subunit I (COI) gene fragment and in some cases, 16S ribosomal RNA (16S) as an alternative gene marker in systematics (Hebert et al. 2003; Vences et al. 2005, 2012, Rockney et al. 2011; Nagy et al. 2012). This method has proven fast and reliable in identification of species across life stages and unveiling cryptic diversities (Smith et al. 2008; Crawford et al. 2011; Nagy et al. 2012).

Despite the accrued global importance of mitochondrial DNA (COI and 16S) in taxonomy, fewer to no studies have extended this approach in identification of herpetofauna from Nigeria. Nigeria is a country in West Africa with high diversity of amphibians and reptiles (collectively referred to as herpetofauna). Herpetofauna play vital roles in ecosystems as predators, food for certain animal species, herbivores and commensal taxa (Wyman 1998; Hopkins 2007; Bohm et al. 2013). Anthropogenic activities, such as deforestation, bush burning etc., pose threat to its herpetofauna (Nneji et al. 2018). This calls for concerted approaches to document herpetofaunal diversity for initiation of conservation plans. Our study, therefore, aims to evaluate the effectiveness of mitochondrial DNA sequences (COI and 16S genes) in identification of Nigerian herpetofauna and detection of possible hidden species diversity.

2. Materials and methods
2.1. Permits, sample collection and morphological identification
Sample collection followed ethical approval from National Park Service, Nigeria. Herpetofauna were collected from Gashaka
Gumti National Park (GGNP) in eastern Nigeria (Supplementary Appendix 1) on two occasions: 21 August–25 September 2017 and 12–16 February 2018. GGNP is located between latitude 06°58′ 08″ 05′ N and longitude 11°10′ N–12°13′ E (Dunn 1999). We collected 189 specimens including 91 amphibian and 98 reptile specimens (Supplementary Appendix 2) with the aid of visual encounter surveys and opportunistic observations. Tissue samples (toe clips) were collected and preserved in 95% ethanol (Sigma-Aldrich, Germany). Subsequently, some animals were released at the site of collection. About 115 voucher specimens were euthanized humanely, tissues (liver/muscle) sampled and subsequently stored in 95% ethanol. Thereafter, specimens were fixed with 4% formalin (Sigma-Aldrich, Germany) and transferred to 70% ethanol. Voucher specimens were deposited in the museums of Department of Zoology, University of Ibadan and GGNP, Nigeria. Ethanol preserved-tissues were kept frozen for further molecular studies.

Preliminary identification of specimens was based on morphology. Further, we referred to taxonomic papers and books on African amphibians and reptiles (Romer 1953; Schiötz 1963; Walker 1967; Dunger 1967a, 1967b; 1968; Walker 1968; Perret and Amiet 1971; Hoogmoed 1974; Perret 1977; Hughes 1983; Butler and Reid 1990; Reid et al. 1990; Schiötz 1999; Rödel 2000; Mediannikov et al. 2012) for clarification of species-level assignments. Specimens with uncertain species-level identification were referred to its genus. Species nomenclatures were in accordance with the Amphibian Web (2018) and The Reptile Database (Uetz et al. 2018).

2.2. Genomic DNA extraction, PCR amplification and sequencing

Total genomic DNA was isolated using standard phenol-chloroform method (Sambrook and Russell 2001). Primer sets used for Polymerase Chain Reaction (PCR) amplification and Sanger sequencing (Bosuut and Milinkovitch, 2000; Che et al. 2012) are shown in Supplementary Appendix 3. The amplification was performed in 25 μl volume reaction that contained 1.5 μl of genomic working DNA, 18.5 μl of PCR water, 2.5 μl of PCR buffer, 2 μl of dNTP, 1 μl of each of the forward and reverse primers (10 pm/μl) and 0.30 μl of rTaq polymerase. The PCR cycle profile consisted of 5 min initial denaturation at 94°C, followed by 35 cycles of 1 min at 94°C, annealing for 45 s at 46°C (COI) or 55°C (16S), extension for 1 min at 72°C; final extension for 10 min at 72°C. The PCR products were checked on 1.2% agarose gel. DNA sequencing of the purified PCR products was performed in both directions using the primers described previously with an automated DNA sequencer (ABI 3730, DNA Analyzer; Applied Biosystems, Foster City, CA).

2.3. Genetic analyses

The nucleotide sequences were viewed and confirmed by eye using SeqManTMII (DNASTAR Lasergene 7) and aligned in MEGA 6.0 using ClustalW (Tamura et al. 2013) with default parameters. The aligned COI sequences were translated to amino acids to check for premature stop codons and to confirm that the open reading frame was maintained in the protein-coding locus.

We attempted to use variety of methods for analyzing our molecular data, including BLAST searches for sequence similarity, the Barcode Index Number (BIN) system, neighbour-joining (NJ) trees and percentage sequence divergence. To confirm the identity of amplified sequences, we conducted BLAST searches on GenBank. However, the unavailability of COI sequences in the database limited its use for species identification. For instance, BLAST searches in GenBank (November 2018) for ‘anurans, Africa, COI’ and ‘reptiles, Africa, COI’ revealed 263 (128 of which were from Xenopus) and 351 (182 of which were from Acontias meleagris) COI records for amphibians and reptiles, respectively. Thus, we failed to use BIN system for our COI data for the species identification. Given the large availability of 16S sequences for African herpetofauna (2905 for anuran and 3022 reptile sequences), we used the data for molecular identification of specimens. After the BLAST search, we downloaded all the related 16S sequences in the Genbank and conducted NJ tree searches for each of the morphologically identified species as implemented in MEGA 6.0 based on the Kimura 2 parameter distance (K2P) (Kimura, 1980). Our identification using 16S NJ trees were based on criteria set by Deichmann et al. (2017) as follows: (1) specimens were identified as the same species if our sequences clustered with same individuals from GenBank; (2) specimens were treated as geographic variation of the same species if specimens were morphologically identified as the same but in different clades (population divergence); (3) specimens were treated as different species if they do not fit morphological description of that species and also showed substantial genetic variation; and (4) specimens were treated as different species even when identified morphologically as the same species but in different clades, sister to other species. Following our specimen identification, we used MEGA v. 6.0 to create NJ trees based on the K2P distance (K2P) and estimated inter- and intraspecific sequence divergences for 16S and COI DNA barcode data for all specimens sequenced for this study.

3. Results and discussion

We provide in Supplementary Appendix 4, the sequence information of 16S and COI gene fragments. Two species of reptiles (Trachylepis affinis and Boaedon lineatus) failed to amplify for COI fragment even after repeated attempts. The amplified COI sequences were without gaps or stop codons. We deposited nucleotide sequences in GenBank with accession number MH708899-MH709081 (16S sequences) and MH700787-MH700946 (COI sequences; Supplementary Appendix 2). The associated 16S NJ trees of our specimens alongside others in GenBank can be found in Supplementary Appendix 5–6.

Using GenBank BLAST tool to confirm identity of the amplified sequences, 81 (83.50%) 16S sequences showed sequence similarity of ≥97% corresponding to 10 morphologically identified amphibian species (66.70%). For reptiles, 74 (80.40%) 16S sequences belonging to 11 species (78.60%)
showed sequence similarity of $\geq$97%. COI sequences yielded low sequence query successes due to poor representation of COI sequences of African herpetofauna in global databases. Given to this, we failed to use BIN system in BOLD database for the species identification. Our study, therefore, shows that GenBank’s BLAST tool for comparing DNA sequence similarity supports the use of 16S gene marker for confirming identities of amplified sequences for Nigerian herpetofauna due to its wide availability in GenBank. It is obvious that the use of molecular data in species identification would not be effective if most African herpetofauna are not well documented in sequence libraries. Thus, efforts in building comprehensive DNA barcode libraries for Nigerian as well as African herpetofauna are recommended.

Anurans morphologically identified as *Hoplobatrachus occipitalis*, *Ammirana galamensis*, *Sclerophrys maculata*, *Sclerophrys regularis*, *Ptychadena bibroni*, and *Xenopus fischbergi* formed 16S clade with others of same species in the GenBank (Supplementary Appendix 5). Following this, we maintain the species names as valid for the specimens. Our NJ tree analysis showed that morphologically unidentified *Sclerophrys* sp clustered with *S. maculata* from GenBank including our newly collected *S. maculata* from GGNP. Therefore, we refer our specimen (*Sclerophrys* sp) as *S. maculata* pending further morphological and genetic identification.

*Hyperolius* sp formed a sub-clade and sister clade within *P. longirostris* group that comprised other individuals from Cote D’Ivoire (Supplementary Appendix 5). For this reason, we refer our specimen as *P. longirostris* and highlight possibility of population divergence within this species group. Individuals identified as *Ptychadena pumilio* formed a clade within *P. pumilio* complex comprising *Ptychadena taenioscelis* and *P. pumilio*. Our data, consistent with Deichmann et al. (2017), suggest that *P. pumilio* is a species complex in need for further revision, and pending this, we identify our species as *P. pumilio*.

Our NJ tree analyses showed that reptile species morphologically identified as *Agama agama*, *Agama parafricana*, *Psammophis lineatus* and *Varanus niloticus* formed 16S clade with others of same species in the GenBank, and thus, we maintain species names as valid (Supplementary Appendix 6). On the other hand, individuals morphologically identified as *Boaedon lineatus* clustered with *Boaedon paralineatus* in the GenBank, but are likely part of *Boaedon lineatus* complex. However, we refer our specimen as ’B. paralineatus’ pending further work is done on this group. Individuals identified as *Boaedon capensis* clustered with *Boaedon perisilvestris* from...
Congo and Chad. The type locality of *B. perisilvestris* is Congo, thus, we refer our individuals as *B. perisilvestris*.

Further, individuals morphologically identified as *Chamaeleo gracilis* formed a distinct sub-clade in a clade comprising *C. gracilis*, *Chamaeleo necasi* and *Chamaeleo laevigatus*. Although we refer our specimen as *C. gracilis*, we note the need for a detailed comparison of *C. gracilis*, *C. necasi* and *C. laevigatus* including the type material, to determine true identity and species identification of these individuals.

Further, individuals morphologically identified as *Trachylepis affinis*, *Trachylepis perrotetti*, *Trachylepis maculilabris* and *Trachylepis quinquetiaeniata* showed significant structures (sub-clades) within clades comprising respective same species from GenBank (Supplementary Appendix 6). We, therefore, retain species names assigned to the individuals as valid and highlight the need for further investigation of processes driving the intraspecific population divergence. We also note that individuals morphologically identified as *Hemidactylus* sp and *Lycophidion* sp formed distinct 16S clades. Thus, we refer these individuals as *Hemidactylus* sp and *Lycophidion* sp, and highlight the need for detailed genetic and molecular comparisons to ascertain if these unknown individuals represent new species or unidentified species.

In sum, our study precludes the use of either morphology or molecular-based approach alone in identification of herpetofauna. Combined use of both approaches would enable accurate species identification.

From the NJ tree analyses (Figures 1–2) for all specimens sequenced for this study, individuals of the same species clustered together. The clustering pattern is in accordance with morphological and genetic identification, enabling the differentiation of individual species. Most of the species clusters in 16S and COI NJ trees were strongly supported (Figures 1–2). Likewise, morphologically unidentified amphibian (*Leptopelis* sp and *Ptychadena* sp) and reptile (*Hemidactylus* sp and *Lycophidion* sp) species formed strongly supported lineages in the NJ trees, and were clearly separated from their sister species. Further, our NJ tree-based analyses provide evidence of at least two strongly supported divergent lineages within *A. agama* (Figures 1–2). In contrast to 16S NJ tree, our COI phenetic analysis revealed strongly supported divergent lineages within *T. maculilabris* (Figure 2). This finding suggests possible hidden diversities within *A. agama* and *T. maculilabris*. Using morphology alone, this diversity remained undetected due to morphological similarity within individuals belonging to these groups.

The K2P interspecific divergence obtained for amphibians was in a range of 5.70 to 31.70% in 16S gene and 19.60 to 32.80% in COI gene (Supplementary Appendix 7). We observed high intraspecific divergences for COI (0.409%) and 16S (0.4175%) genes in *K. cassinoides* (Supplementary Appendix 8). In reptiles, we obtained K2P interspecific divergence ranged from 4.50 to 36.80% in 16S and 8.20 to 44.20% in COI respectively (Supplementary Appendix 9). However, the highest intraspecific divergences for COI (1.90%) and 16S (0.47%) genes were observed in *A. agama* (Supplementary Appendix 10).

Finally, our study illustrates the relevance of molecular data in species identification and unveiling hidden diversity within Nigerian herpetofauna. Our results largely support the integration of morphological and molecular data in taxonomy. In an ongoing effort to build comprehensive regional sequence libraries, this study adds new sequences to the global databases that would aid in identification of herpetofauna. Further, we recommend validation of sequences deposited in global databases with respect to its voucher specimen for accurate species identification.

Acknowledgements

We thank the management of Gashaka Gumti National Park, Taraba, Nigeria for providing technical assistance during our field survey.
Disclosure statement

The authors declare that they have no competing interests.

Funding

This study was funded by Rufford Small Grant for Nature Conservation (22507-1) and IDEA WILD Nature Conservation Grant.

References

Amphibian Web. 2018. Amphibians of Nigeria. https://amphibiaweb.org/cgi/amphib_query?rel=–isocc=like&orderbyaw=Order&where=–isocc=Nigeria. [Accessed 2018 May 17].

Bohm M, Collen B, Baillie J, Bowles P, Chanson J. 2013. The conservation status of the world’s reptiles. Biol Conserv. 157:372–385.

Bossuyt F, Milinkovitch MC. 2000. Convergent adaptive radiations in Madagascan and Asian ranid frogs reveals covariation between larval and adult traits. PNAS. 97:6585–6590.

Butler JA, Reid JC. 1990. Records of snakes from Nigeria. Nigerian Field. 50:19–40.

Che J, Chen HM, Yang JX, Jin QJ, Jiang K, Yuan ZY, Murphy RW, Zhang YP. 2012. Universal COI primers for DNA barcoding amphibians. Mol Ecol Resour. 12:247–258.

Crawford AJ, Alonso R, Jaramillo CA, Sucre S, Ibenez R. 2011. DNA barcoding identifies a third invasive species of ELEFTHERODACTYLUS (Anura: ELEUTHERODACTYLIDAE) in Panama City, Panama. ZooTaxa. 2890:65–67.

Deichmann JL, Mulcahy DG, Vanthomme H, Tobi E, Wynn AH, Zikmus GT, Dunger GT. 1967a. The lizards and snakes of Nigeria Part 1: the chame-

Deichmann JL, Mulcahy DG, Vanthomme H, Tobi E, Wynn AH, Zikmus GT, Dunger GT. 1967b. The lizards and snakes of Nigeria Part 3: the monitor

Deichmann JL, Mulcahy DG, Vanthomme H, Tobi E, Wynn AH, Zikmus GT, Dunger GT. 1968. The lizards and snakes of Nigeria Part 4: the geckos of

Dunger GT. 1967a. The lizards and snakes of Nigeria Part 1: the chame-

Dunger GT. 1967b. The lizards and snakes of Nigeria Part 3: the monitor

Dunger GT. 1968. The lizards and snakes of Nigeria Part 4: the geckos of

Dunga A. 1999. Gashaka Gumi National Park: A Guidebook. Lagos, Nigeria: Nigerian Conservation Foundation/WWF-UK.

Hebert PDN, Ratnasingham S, deWard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B Biol Sci. 270:96–99.

Hoogmoed MS. 1974. Ghanese lizards of the genus MABuya (Scincidae, Saurops, Reptilia). Zool Verh. 138:1–68.

Hopkins WA. 2007. Amphibians as models for studying environmental change. Icar. 48:270–277.

Hughes B. 1983. African snake faunas. Bonn Zool Beitr. 34:311–355.

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

Mediannikov O, Trape S, Trape JF. 2012. A molecular study of the genus Agama (Squamata: Agamidae) in West Africa, with description of two new species and a review of the taxonomy, geographic distribution, and ecology of currently recognized species. Russ J Herpetol. 192:115–142.

Nagy ZT, Sonet G, Flaw F, Vences M. 2012. First large-scale DNA barcoding assessment of reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI primers. PLoS ONE. 7:e34506.

Nazarov R, Poyarkov NA, Orlov NL, Phung TM, Nguyen TT, Hoang DM, Ziegler T. 2012. Two new cryptic species of the Cyrtodactylus irregula-

Nheji LI, Adeola AC, Yan F, Okeyoyin AO, Oladipo OC, Olagunju TE, Omotoso O, Oladipo SE, Iyiola OA, Auta T, et al. 2018. Cryptic lineages of Nigerian Agama (Squamata: Agamidae) within the West African radiation. Russ J Herpetol. 25:97–112.

Perret JL. 1977. Les Hylarana (Amphibiens, Ranidé) du Cameroun [The Hylarana (Amphibians, Ranidae) of Cameroon]. Rev Suisse Zool. 84:1–866. French

Perret JL, Amiet JL. 1971. Remarques sur les Bufo (Amphibiens, Anoures) du Cameroun [Remarks on the Bufo (Amphibians, Anura) of Cameroon]. Ann Fac Sci Cameroun 55. French. 5:47.

Reid JC, Owens A, Laney R. 1990. Records of frogs and toads from Akwa Ibom State. Nigerian Field. 55:113–128.

Rockney HJ, Ofori-Boateng C, Porcino N, Leaché AD. 2015. A comparison of DNA barcoding markers in West African frogs. Afr J Herpetol. 64:135–147.

Rödel MO. 2000. Herpetofauna of West Africa, Volume I: Amphibians of the West African savanna. Frankfurt am Main, Germany: Edition Chimaira.

Romer JD. 1953. Reptiles and amphibians collected in the Port Harcourt area of Nigeria. Cop.2:121–123.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Press.

Schietz A. 1963. The amphibians of Nigeria. Videnskabelig Meddelelser Før Dansk Naturhistorisk Forening. 125:1–92.

Schietz A. 1999. Treefrogs of Africa. Frankfurt am Main, Germany: Edition Chimaira.

Smith MA, Poyarkov NA, Hebert PDN. 2008. DNA barcoding: CO1 DNA barcoding amphibians: take the chance, meet the challenge. Mol Ecol Resour. 8:235–246.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol. 30:2725–2729.

Uetz P, Freed P, Hosek J. 2018. The Reptile Database. http://reptile-data-base.reptarium.cz/search?search=–Nigeria&submit=Search. [Accessed 2018 May 17].

Vences M, Nagy ZT, Sonet G, Verheyen E. 2012. DNA barcoding amphibians and reptiles. Meth Mol Biol. 858:79–107.

Vences M, Thomas M, van der Meijden A, Chiari Y, Vieites DR. 2005. Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. Front Zool. 2:5.

Walker RB. 1967. Rana galamensis, an elusive frog. Nigerian Field. 32:22–26.

Walker RB. 1968. The amphibians of Zaria, in the northern guinea savanna, Nigeria. Cop. 1:164–167.

Wyman RL. 1998. Experimental assessment of salamanders as predators of detrital food webs: effects on invertebrates, decomposition and the carbon cycle. Biodivers Conserv. 7:641–650.