Molecular characterization of a rabies virus isolated from trade dogs in Plateau State, Nigeria

GSN Kia¹,²,³, Y Huang¹, M Zhou¹, Z Zhou¹, CW Gnanadurai¹, CM Leysona¹, JU Umoh², HM Kazeem⁴, DO Ehizibolo⁵, JKP Kwaga²,³, CI Nwosu⁵ & ZF Fu¹*

¹. Department of Pathology, College of Veterinary Medicine, University of Georgia Athens, GA 30602, USA
². Department of Veterinary Public Health and Preventive Medicine, Faculty of Veterinary Medicine, Ahmadu Bello University Zaria, Nigeria
³. African Centre of Excellence in Neglected Tropical Diseases and Forensic Biotechnology, Ahmadu Bello University Zaria, Nigeria
⁴. Department of Veterinary Microbiology, Faculty of Veterinary Medicine, Ahmadu Bello University Zaria, Nigeria
⁵. National Veterinary Research Institute Vom, Jos Plateau State, Nigeria

*Correspondence: Tel.: +1 7065427021; E-mail: zhenfu@uga.edu

Abstract
To have a better understanding of the prevalence of rabies virus (RABV) in the dog trading markets in Plateau State of Nigeria, a total of 532 brain samples collected between 2010 and 2011 were subjected to rabies virus antigen detection and molecular characterization. RABV antigen was detected in 92 out of 532 (17.3%) brain samples from two major commercial dog markets in Plateau State of Nigeria by direct fluorescence antibody assay, indicating the possibility of a high prevalence of RABV in the dog trading markets. The complete genomic sequence was obtained from one of these RABV positive samples. The overall organization of this virus (DRV-NG11) was typical of all the other wild-type RABV. Phylogenetic analysis using either the complete or partial genomic sequence of DRV-NG11 demonstrates that this isolate is most closely related to viruses previously shown to circulate in Nigeria, belonging to the Africa 2 Clade. To our knowledge, this is the first report on the molecular characterization of the complete genome of RABV from trade dogs, which provides a better understanding of the molecular epidemiology, pathogenesis and rabies control in this country.

Keywords: Genome, Nigeria, Rabies virus, Street strain, Trade dogs

Introduction
Rabies is an ancient zoonosis and remains a public health threat, causing more than 70,000 human deaths each year worldwide (Knobel et al., 2005). More than 95% of human deaths occur in Asia and Africa where canine rabies is endemic (Fu, 1997). Rabies virus (RABV), the causative agent of rabies, belongs to the genus Lyssavirus in the Rhabdoviridae family (Jackson, 2002). Human infections are mostly due to the canine biotype of RABV (Rupprecht et al., 2008; Nel, 2013). The Lyssavirus genus to date consists of 14 recognized species all capable of causing fatal encephalitic rabies. The prototype species of this genus is Rabies lyssavirus and the rest are known as the rabies-related lyssaviruses (Coertse et al., 2017). Its genome is a single-stranded negative-sense RNA of approximately 12 kb in length and encodes five structural proteins: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA-dependent RNA polymerase (L).
Based on sequence analysis, RABVs in Africa were delineated into four clades as Africa 1, 2, 3, and 4 (Kissi et al., 1995; David et al., 2007). The Africa 4 Clade includes the isolates from North Africa such as Egypt and Israel, while the Africa 3 clade is common in southern Africa and has been adapted to the mongoose (David et al., 2007). The Africa 2 clade has a wide distribution in western and central Africa, with minimal overlap with the Africa 1 clade in Central African Republic (CAR) and Nigeria (Talbi et al., 2009). The Africa 1 clade is distributed in North, Central and South Africa and was further differentiated into 1a and 1b (Kissi et al., 1995).

In Nigeria, rabies was officially reported as far back as 1912 and 1925 in humans and animals, respectively (Boulger & Hardy, 1960), and it is still a serious public health threat nowadays. Rabies outbreaks occur frequently in Nigeria because of the low immunization coverage in dogs. A clinical human rabies surveillance study in the University of Nigeria Teaching Hospital from 2004 to 2013 showed that all the reported human rabies cases were transmitted by dogs, especially the stray dogs with unknown rabies vaccination history, and the outcome was 100% fatal (Eke et al., 2015). The National Veterinary Research Institute (NVRI) had confirmed 4809 animal rabies cases in Nigeria within the last 77 years (1938–2005). Because of the poor rabies reporting system, the number of real rabies cases in Nigeria could be much higher than that reported (Ajayi et al., 2006).

As a central state of Nigeria, Plateau State has flourishing commercial dog markets (captured stray dogs are sold), dogs from different states and other neighboring countries (Niger and Chad) are sold to traders from different states of Nigeria (Sabo et al., 2008). The occurrence of apparently healthy dogs with rabies virus in the brain tissues, yet not showing clinical signs of rabies may be more widespread than previously envisaged as rabies antigens were also found in the brain tissues of apparently healthy dogs slaughtered for human consumption in Borno, Benue, Plateau, Katsina and Sokoto, Cross River and Lagos State (Ajayi et al., 2006; Akombo, 2009; Sabo, 2009, Garba et al., 2010; Odeh et al., 2013; Isek, 2013; Hambolu et al., 2013). The need to safe guard the health of the people by reducing the rabies risk of exposure warrants continued surveillance.

In an earlier report of the first complete genome sequence of RABV (DRV-NG11) isolated from Nigeria, Zhou et al. (2013) compared its N gene sequences to others obtained from the GenBank. Based on this phylogenetic comparison, DRV-NG11 was found to be similar to another Nigerian RABV (EU038098) with high homology (99.4%) belonging to the Africa 2 lineage (Kissi et al., 1995; Talbi et al., 2009). Thus in this study, findings on the status of rabies in the trade dogs slaughtered in Plateau State, molecular epidemiology and characterization of the complete sequences of the 5 proteins (N, P, M, G and L) of DRV-NG11 are presented. Understanding of the phylogenetic relationship of DRV-NG11 with other global or African isolates, pathogenicity of the DRV-NG11 and molecular baseline sequence information presented in this report will shed more light on the mechanism of rabies pathogenesis which is still unclear despite the remarkable progress that has been made (Jackson, 2011). The information from this study will also be relevant to vaccine production.

Materials and Methods

Study area

Dawaki (Latitude 9°27’N; Longitude 9° 30’E) and Amper (Latitude 9° 20’N; Longitude 9° 45’E) are the two major dog markets in Kanke Local Government area (LGA) of Plateau State. The market days are Wednesdays for Dawaki and Saturdays for Amper. A minimum of 40 dogs were slaughtered on each of these market days. Dawaki is known all over Nigeria by the dog eating population as a major point for mass purchase of dogs for resale elsewhere (Sabo et al., 2008).

Specimen collection and RABV antigen detection

A total of 532 dog brain tissue samples were collected from two trading markets in Plateau State from 2010 to 2011. Collection of brain tissues was carried out according to the method of Barrat (1996). The technique allows samples to be taken from the medulla oblongata, the base of the cerebellum and the hippocampus (Ammons horn and the cerebral cortex). Brain tissues collected were put into individual sterile labeled tubes and immediately stored at -20 °C in a small Thermocool freezer while on the field and later transferred to -80°C freezers in the National Veterinary Research Institute (NVRI) Vom, Plateau State Nigeria. Later, the samples were shipped on dry ice to the College of Veterinary Medicine, University of Georgia, Athens (UGA) USA, where they were stored at -80°C until use.

The brain samples were subjected to the direct fluorescent antibody test (DFAT) for the presence of RABV antigens. Briefly, touch impressions of brain tissue were fixed in cold acetone; air dried and stained with fluorescence isothiocyanate-labeled anti-rabies antibodies (Fujirebio, Malvern, PA).
RABV-infected and normal mouse brains were used as positive and negative controls, respectively.

**Total RNA extraction and reverse transcription polymerase chain reaction (RT-PCR)**

To confirm the DFA results, RT-PCR was performed. Total RNA was extracted with Trizol® reagent (Invitrogen, San Diego, CA) and reverse transcription was performed with random primer pd(N)₆ using Superscript II (Invitrogen) according to the manufacturer’s instructions. RABV N gene was amplified with primers N127 (ATGTAACCCCTCACAATGG) and N8m (CAGTCTCTCGCCATCTC). For full-length genome amplification, 10 overlapping fragments were amplified by PCR using the Pfu polymerase kit (Invitrogen) and primers used were described previously (Zhou et al., 2013). All the amplicons were subjected to gel extraction and then cloned into pCR-Blunt II vector (Invitrogen). All the recombinant plasmids were identified by restriction enzyme digestion, and then 3-5 positive recombinant plasmids were selected for sequencing using a BigDye terminator cycle sequencing kit on an ABI Prism 3730 sequencer.

**Sequence alignment and analysis**

The forward and reverse sequences were assembled and manually checked using the ATGC program version 4 (Genetyx Co., Tokyo, Japan). The CLUSTALX version 1.83 package was used for the multiple alignments of the sequences. Phylogenetic trees were constructed using Molecular Evolutionary Genetics Analysis (MEGA), version 4 using the neighbor-joining (NJ) algorithm with the Kimura two-parameter model. The reliability of the phylogeny groupings was evaluated using bootstrapping with 1000 replicates. DNAStar was used to translate the gene sequences and to determine the percentage identities. Genedoc and BioEdit software were used for variability analyses and to predict the functional regions, respectively.

**Table 1:** Sequences of transcription, initiation and termination signals of all five genes and intergenic regions of the DRV-NG11 virus

<table>
<thead>
<tr>
<th>Gene or region</th>
<th>Gene or region</th>
<th>Gene or region</th>
<th>Gene or region</th>
</tr>
</thead>
<tbody>
<tr>
<td>3′Leader/N</td>
<td>3′Leader/N</td>
<td>3′Leader/N</td>
<td>3′Leader/N</td>
</tr>
<tr>
<td>N/P</td>
<td>N/P</td>
<td>N/P</td>
<td>N/P</td>
</tr>
<tr>
<td>P/M</td>
<td>P/M</td>
<td>P/M</td>
<td>P/M</td>
</tr>
<tr>
<td>M/G</td>
<td>M/G</td>
<td>M/G</td>
<td>M/G</td>
</tr>
<tr>
<td>G/L</td>
<td>G/L</td>
<td>G/L</td>
<td>G/L</td>
</tr>
<tr>
<td>L/5′trailer</td>
<td>L/5′trailer</td>
<td>L/5′trailer</td>
<td>L/5′trailer</td>
</tr>
</tbody>
</table>

**Results**

**Specimen detection, genome characterization and phylogenetic analysis of DRV-NG11**

In this study, 92 (17.3%) out of the 532 brain samples collected from the dog trading markets in Plateau State were positive for RABV by DFAT. Out of these 92 samples 4 yielded rabies-specific bands of expected amplicon size (1.3kb). The N genes of those positive samples were sequenced; the sequencing result showed that all the samples had 99.8% homology (data not shown). The complete genome of the RABV from one of the positive samples termed DRV-NG11 encoding the five structural genes in the order of 3′-N-P-M-G-L-5′, was further sequenced and characterised. The complete genome of DRV-NG11 has 11,923 nucleotides (nt) with the lengths of the coding sequences as follows: 1353 nt (N), 894 nt (P), 609 nt (M), 1575 nt (G), and 6384 nt (L). The length of the G-L non-coding region is 424 nt. The sequences of the transcription initiation sites and the termination signals of all five genes and the intergenic regions between these genes did not differ from other lyssaviruses as shown in Table 1. The phylogenetic analysis of DRV-NG11 with other global or African isolates, based on the entire genome showed that DRV-NG11 is genotype 1 *Lyssavirus* and belongs to the clade I (Figure 1A). The phylogenetic analysis based on the N gene presence DRV-NG11 to be closely related to other RABVs identified in Nigeria belonging to the Africa 2 (Figure 1B). The variants found in Nigeria are distinguishable from those of North (Africa 1a), East (Africa 1b) and South Africa (Africa 3) (Figure 1B).

**The characteristics of DRV-NG11 individual structural proteins**

The N protein of DRV-NG11 has 450 amino acid residues and shows only 2 unique substitutions (Ser⁸¹ and Ala¹²⁷), while Phe⁸¹ and Val¹²⁷ are present in other African isolates. The three recognized antigenic sites (antigenic site I, IV and RNA binding domain) are conserved in DRV-NG11 and other African isolates. In P protein, four amino acid
Figure 1: Neighbor-joining phylogenetic trees constructed from DRV-NG11 isolate (in black box) and 29 whole genome sequences from all over the world (A) and 36 full-length nucleoprotein gene sequences from Africa (B). Bootstrap values are indicated at the nodes (only those with values >90% are indicated), and the scale bar indicates nucleoprotein substitutions per site.

Substitutions Asp\(^{51}\), Gln\(^{56}\), Val\(^{174}\), and Arg\(^{273}\) are unique in DRV-NG11. The motifs for interaction with the dynein light chain and binding are encoded as KSTQT and FSKKYKF in DRV-NG11 (Figure 2). In M protein, the late domain motif is encoded as PPEY in DRV-NG11 and all the other African isolates. At position 58 is Glu in DRV-NG11 and all the other African isolates (Figure 2). It was found that DRV-NG11 G protein is similar to all the other RABV G proteins, consisting of a signal peptide sequence (the first 19 residues), an ectodomain (aa1-439), a transmembrane domain (aa440-461), and a cytoplasmic domain (aa462-505). The antigenic site GI (aa231), GII (aa34-42 and 198-200) and G III (aa330-338) are conserved in all strains (Figure 2). Arginine at G333 (Arg\(^{333}\)) is conserved in DRV-NG11, while a Glutamine (Gln\(^{333}\)) mutation occurred at this position in attenuated strains CTN181 and HEP-Flury (Figure 2). DRV-NG11 G has two N-glycan sites (Asn\(^{17}\) and Asn\(^{119}\)), while another two N-glycan sites on Asn\(^{158}\) (PV and HEP-Flury) or Asn\(^{247}\) (PV, CTN181, SADB19 and ERA) exist only in the vaccine strains (Figure 2). Asp\(^{31}\), Phe\(^{33}\), Arg\(^{37}\), Gly\(^{38}\) and Lys\(^{39}\) are important sites which correspond to the amino acid at position 190, 192, 196, 197 and 198 in RABV G protein, all of these sites are conserved and retained in DRV-NG11 (Figure 2). The GHP motif (aa372-374), putative leucine-zipper motif and LNSPL motif (aa39-43) in domain I of L protein are well conserved in DRV-N11. The RNA polymerase pentapeptide QGDNQ (aa727-731) motif in domain III and GXGXG motif (aa1704-1708) in domain VI involving in mRNA capping are all also conserved in DRV-NG11 as well as in other strains.

Discussion

It is estimated that 24,000 humans die of rabies in Africa each year, almost all of whom died from infection with dog RABVs (Talbi et al., 2009). Plateau State has the largest commercial dog market for human consumption in western Africa, where the dogs are experiencing a serious rabies epidemic (Okoh, 2000). Stray or domestic dogs are collected from 9 northern states of Nigeria, Niger as well as Chad Republics and then traded in these commercial markets for human consumption (Sabo, 2008). After subjecting these 92 DFAT positive brain samples to RT-PCR, RABV RNA was detected in only four (4.35%) of the samples. This result was contrary to what was expected considering the high sensitivity of PCR. The
**Figure 2:** Multiple alignment of mature G protein amino acid residues 1-505 for African isolations and vaccine strains. Antigenic site GI (amino acid 231), GII (amino acids 34–42 and 198–200), GIII (amino acids 330–338) are indicated with dotted lines. The transmembrane domain is indicated with solid line. The N-linked glycosylation sites are indicated with triangles. The ligand for acetylcholine receptor is highlighted.
low RT-PCR detection of 4.35% might have been caused by contamination with some RNAase at some level of sample processing unknown to the researchers. It may also be due to poor RNA quality and integrity probably compromised during transportation. However, the use of other short sequence primers targeting the N gene in a nested RT-PCR may have increased the sensitivity of the PCR (Dzikwi, 2008). However, the 17.3% positive rate by DFAT may indicate that the dogs sold in the markets are highly potential RABV reservoirs of infection and the dog trading may contribute to the endemicity of rabies in Nigeria. It also presents risks for those who are involved in the dog trading chain, including hunters, handlers, slaugtherers, as well as consumers. The complete genome of DRV-NG11 was found to be similar to other street RABVs published to date (Yu et al., 2012; Zhang & Fu, 2012; Wunner, 2007). The transcription, initiation and termination sequences are consistent with all lyssavirus genotypes (Marston et al., 2007). Previous observations that viruses from the same geographical area tend to be grouped together agree with the findings from this study (Kissi et al., 1995). The variants found in Nigeria are distinguishable from those of North (Africa 1a), East (Africa 1b) and South Africa (Africa 3) (Figure 1B) (Kissi et al., 1995).

The N protein of any RABV is the most conserved gene (Szano et al., 2008). Ser81 and Ala427, were the only two amino acid substitutions, unique in DRV-NG11 strains. The three recognized antigenic sites (antigenic site I, IV and RNA binding domain) of the N protein (Goto et al., 2000) are conserved in DRV-NG11 and other African isolates. Ser389, which is considered to be related to casein-type phosphorylation site and regulation of viral RNA transcription and replication, is highly conserved in all other African strains (Yang et al., 1999).

In P protein, four amino acid substitutions Asp51, Gln56, Val124, and Arg273 are unique in DRV-NG11. Correspondently, Glu51, His56, Ala124, Lys273 are conserved in other African strains. The serum at positions 162, 210 and 271 involved in phosphorylation of P protein are all highly conserved in DRV-NG11 and all the other African strains (Gupta et al., 2000). Moreover, the first 19 amino acids in P protein, where the L protein is known to bind, are conserved in all strains. The motifs for interaction with the dynein light chain (LC8) (K/P XTQT, aa144-148) and binding N-RNA (aa209-215) are encoded as KSTQT and FSKKYYKF in DRV-NG11, respectively (Poisson et al., 2001; Liu et al., 2004).

In M protein, the late domain motif reported to bind class I WW-domain-containing E3-ubiquitin ligases to reduce virus budding (Jayakar et al., 2000) is encoded as PPEY in DRV-NG11 and all the other African isolates. The important residue at position 58 involving regulation of viral RNA synthesis is Glu in DRV-NG11 and all the other African isolates (Finke and Conzelmann, 2003). G is the only viral protein exposed on the surface of the virion and plays an essential role in RABV pathogenicity (Lentz et al., 1983; Thoulouze et al., 1998). DRV-NG11 was shown to be highly virulent and caused 100% percent mortality in experimental dogs after intramuscular administration in a lower dose (Gnanadurai et al., 2015). To identify the differences of functional regions between DRV-NG11 and attenuated RABVs, their G protein sequences were aligned. It is found that DRV-NG11 G protein is similar to all the other RABV G proteins, consisting a signal peptide sequence (the first 19 residues), an ectodomain (aa1-439), a transmembrane domain (aa440-461), and a cytoplasmic domain (aa462-505). However, the well-defined virulence determinant residue Arginine at G333 (Arg313) is conserved in DRV-NG11, while a Glutamine (Gln333) mutation occurred at this position in attenuated strains CTN181 and HEP-Flury (Figure 2). N-linked glycosylation of G protein is another important factor in biology and pathogenesis of neurotropic RABV and the number of N-linked glycosylation sites has been demonstrated to affect its pathogenicity (Yamada et al., 2013). In this study, DRV-NG11 G has two N-glycan sites (Asn137 and Asn199), while another two N-glycan sites on Asn158 (PV and HEP-Flury) or Asn247 (PV, CTN181, SADB19 and ERA) exist only in the vaccine strains (Figure 2). Asn37 N-glycan has been suggested to be a non-core glycosylation site, while Asn379 N-glycan is critical for viral production and fusion activity, and additional one or two N-glycan sites on Asn379, Asn158 or Asn247 could enhance virus production (Yamada et al., 2013). This may be the reason why Asn379 N-glycan is the unique N-linked glycosylation site existing in all RABVs (Badrane et al., 2001) and the additional N-glycan sites are only observed in vaccine strains, not in DRV-NG11 (Figure 2). In addition, the acetylcholine receptor was considered as a host-cell receptor for the highly neurotropic RABVs. The amino acid sequence (aa189-214) of RABV G protein has higher similarity with that in loop 2 of snake venom curare-mimetic neurotoxins, potent ligands of the acetylcholine receptor (Lentz et al., 1984). The important sites Asp91, Phe93, Arg97, Gly38 and Lys39 of
neurotoxins corresponds to the amino acid at position 190, 192, 196, 197 and 198 in RABV G protein, all of these sites are conserved and retained in DRV-NG11 (Figure 2). L protein is an essential component of RNP, responsible for the enzymatic activities involved in viral transcription and replication. As established previously for the *Rhabdovirus* L protein, it is organized into six well conserved domains that contain some invariant motifs (Poch et al., 1990). Many specific motifs within these six conserved domains are also retained in DRV-NG11. The invariant GHP motif (aa372–374), putative leucine-zipper motif and LNSPL motif (aa39-43) (Marston et al., 2007) in domain I are well conserved. The pentapeptide GQDGNQ (aa727–731) motif in domain III is retained in NG11 as well as in other strains. This motif is critical for RNA polymerase activities and template recognition or phosphodiester bond formation (Schnell & Conzelmann, 1995). Moreover, the GXGXG motif (aa1704–1708) in domain VI involving in mRNA capping (Ferron et al., 2002) is also conserved in DRV-NG11.

In conclusion, the status of rabies in the two biggest commercial dog markets in Plateau State of Nigeria, indicates that dogs sold in the markets are highly potential RABV reservoirs of infection. Furthermore, the molecular characteristics of the whole genome were determined for one of the isolates, DRV-NG11. Our studies indicate that DRV-NG11 is a typical RABV in its genome organization and in the functional domains of each individual structural protein. The phylogenetic analysis showed that DRV-NG11 is genotype 1 lyssavirus and belongs to the clade I. Phylogenetic analysis based on the N gene has shown that DRV-NG11 cluster with Africa 2 clade and is closely related to those previously isolated from Nigeria and other neighboring countries ( Chad and Niger republic), indicating that the transportation and transnational relocation of trade dogs from different states or countries may contribute to the high prevalence of rabies in Plateau state, Nigeria.

To our knowledge, this is the first characterization of a complete genome sequence of RABV isolated from Nigeria, which will advance further studies in rabies molecular epidemiology, pathogenesis and rabies control in this area.

**References**


Isek TI (2013). Epidemiological Studies of Canine Rabies in Ogoja, Cross River State. MSc thesis, Department of Veterinary Public Health and Preventive Medicine, Faculty of Veterinary Medicine, Ahmadu Bello University, Zaria, Nigeria. Pp 15-56.


