EMERGENCE OF A NEW GENETIC LINEAGE OF NEWCASTLE DISEASE VIRUS IN WEST AND CENTRAL AFRICA - IMPLICATIONS FOR DIAGNOSIS AND CONTROL

G. Cattoli, A. Fusaro, I. Monne, S. Molia, A. Le Menach, B. Maregeya, A. Nchare, I. Bangana, A. Garba Maina, Koffi J.-N. N'Goran, et al.

Title: EMERGENCE OF A NEW GENETIC LINEAGE OF NEWCASTLE DISEASE VIRUS IN WEST AND CENTRAL AFRICA – IMPLICATIONS FOR DIAGNOSIS AND CONTROL

Authors: G. Cattoli, A. Fusaro, I. Monne, S. Molia, A. Le Menach, B. Maregeya, A. Nchare, I. Bangana, A. Garba Maina, Koffi J.-N. N'Goran, H. Thiam, O.E.M.A. Bezeid, A. Salviato, R. Nisi, C. Terregino, I. Capua

PII: S0378-1135(09)00499-4
DOI: doi:10.1016/j.vetmic.2009.09.063
Reference: VETMIC 4623

To appear in: *VETMIC*

Received date: 22-4-2009
Revised date: 24-8-2009
Accepted date: 16-9-2009

Please cite this article as: Cattoli, G., Fusaro, A., Monne, I., Molia, S., Le Menach, A., Maregeya, B., Nchare, A., Bangana, I., Maina, A.G., N'Goran, K.J.-N., Thiam, H., Bezeid, O.E.M.A., Salviato, A., Nisi, R., Terregino, C., Capua, I., EMERGENCE OF A NEW GENETIC LINEAGE OF NEWCASTLE DISEASE VIRUS IN WEST AND CENTRAL AFRICA – IMPLICATIONS FOR DIAGNOSIS AND CONTROL, *Veterinary Microbiology* (2008), doi:10.1016/j.vetmic.2009.09.063

EMERGENCE OF A NEW GENETIC LINEAGE OF NEWCASTLE DISEASE VIRUS IN WEST AND CENTRAL AFRICA – IMPLICATIONS FOR DIAGNOSIS AND CONTROL

Cattoli G[1*], Fusaro A[1], Monne I[1], Molia S[2], Le Menach A[3], Maregeya B[4], Nchare A[5], Bangana I[6], Garba Maina A[7], N'Goran Koffi J-N[8], Thiam H[9], Bezeid O E M A[10], Salviato A[1], Nisi R[1], Terregino C[1], Capua I[1]

Affiliations

1 OIE/FAO Reference Laboratory for avian influenza and Newcastle disease, Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020 Legnaro (PD), Italy.

2. CIRAD, UPR AGIRs, BP1813, Bamako, Mali.

3. Food and Agriculture Organisation (FAO), AGAH Division, Viale delle Terme di Caracalla, Rome, Italy.

4. Food and Agriculture Organisation (FAO) Burundi, Avenue du Peuple Murundi, Bujumbura-Burundi.

5. Ministère de l'Elevage, des Pêches et des Industries/DSV, BP 2228 Yaoundé, Cameroun.

6. Direction de la Santé Animale / Ministère des Ressources Animales, B.P 12091 Niamey, Niger.

7. ECTAD Unit, Food and Agriculture Organisation (FAO), United Nations House, Central Area District, Abuja, Nigeria.

8. Label Ivoire Elevage (LIVEL), 06 BP 117 Abidjan 06, Côte d'Ivoire.

9. Consultant, 01 BP 2266 Ouagadougou 01, Burkina Faso.

10.Centre National d'Elevage et de Recherche Vétérinaires (CNERV), BP 167, Nouakchott, Islamic Republic of Mauritania

*Corresponding author:

Giovanni Cattoli

OIE/FAO Reference Laboratory for avian influenza and Newcastle disease,
Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10,
35020 Legnaro (PD), Italy. Tel +39 0498084389 Fax +39 0498084360
gcattoli@izsvenezie.it

Abstract

Newcastle disease (ND) is an OIE listed disease caused by virulent avian paramyxovirus type 1 (APMV-1) strains, which affect many species of birds and may cause severe economic losses in the poultry sector. The disease has been officially and unofficially reported in many African countries and still remains the main poultry disease in commercial and rural chickens of Africa. Unfortunately, virological and epidemiological information concerning ND strains circulating in the Western and Central regions of Africa is extremely scarce. In the present study, sequence analysis, pathotyping and detailed genetic characterization of virulent ND strains detected in rural poultry in West and Central Africa revealed the circulation of a new genetic lineage, distinguishable from the lineages described in the Eastern and Southern parts of the continent. Several mismatches were observed in the segment of the matrix gene targeted by the primers and probe designed for the molecular detection of APMV-1, which were responsible for the false negative results in the diagnostic test conducted. Furthermore, deduced amino acid sequences of the two major antigens eliciting a protective immune response (F and HN glycoprotein) revealed protein similarities < 90% if compared to some common vaccine strains. Distinct mutations located in the neutralizing epitopes were revealed, indicating the need for detailed assessment of the efficacy of the current vaccines and vaccination practices in Africa. The present investigation provides important information on the epidemiology, diagnosis and control of NDV in Africa and highlights the importance of supporting surveillance in developing countries for transboundary animal diseases.

Key words: Newcastle Disease, poultry, Africa, genetic characterization

Introduction

Newcastle disease (ND) is a viral disease affecting many species of birds and causing severe economic losses in the poultry sector. The aetiological agent is a single stranded, non-segmented, negative-sense RNA virus belonging to the order *Mononegavirales*, family *Paramyxoviridae*, subfamily *Paramyxovirinae*, genus *Avulavirus* (Lamb et al., 2005). This genus contains the 9 serogroups of avian paramyxoviruses (APMV-1 to -9) described so far. According to their virulence in poultry, APMV-1 isolates can be classified as lentogenic, mesogenic or velogenic (i.e. of low, moderate or high virulence for poultry, respectively). In the present paper the term ND virus (NDV) refers to the virulent APMV-1 strains that cause ND.

The virulence of an APMV-1 isolate can be determined in the laboratory by an *in vivo* test performed by the experimental inoculation of SPF day-old chicks with the isolate (intracerebral pathogenicity index, ICPI) or it can be predicted molecularly by the analysis of the amino acid sequence at the cleavage site of the fusion glycoprotein (F protein) (Alexander, 2009). Both these methods are officially recognized for APMV-1 pathotyping by the World Organisation for Animal Health – OIE - (OIE, 2008). The ICPI is considered the most sensitive “in vivo” test (Terregino & Capua, 2009) and an index value of 0.7 or greater classifies the isolate as virulent (i.e. velogenic or mesogenic) thus, confirming a diagnosis of ND in the infected hosts. Alternatively, the virulence of ND strains can be assessed by the determination of the presence or not of multiple basic amino acids (e.g. lysine – K; or arginine – R) located at the C-terminus of the F1 protein and phenylalanine (F) at the N-terminus of the F2 protein, which correspond to the cleavage site of the precursor F0 glycoprotein (positions 112 to 117) (OIE, 2008).

In addition to the gene encoding the structural glycoprotein F, the APMV-1 genome of about 15 kb is composed of 5 other genes encoding 5 structural proteins (nucleoprotein, NP; matrix protein, M; phosphoprotein, P; RNA polymerase, L and haemagglutin-neuraminidase, HN). Two additional proteins are encoded by the RNA editing of the P protein, namely proteins V and W. The cleavability of protein F is certainly the main determinant for viral virulence, but other proteins such

as HN and V are also believed to influence pathogenicity (Huang et al., 2004; de Leeuw et al., 2005; Mebatsion et al., 2001).

As for many other RNA viruses, genetic variability of APMV-1 is significant. Recent analyses have revealed the existence of two main distinct genetic clades, termed class I and II (Czeglédi et al., 2006). Class I includes almost exclusively low virulence strains recovered from live bird markets in US or wild waterfowl worldwide. Class II is comprised of the vast majority of viruses of high and low virulence isolated from poultry and wild birds (Czeglédi et al., 2006; Kim et al., 2007). Restriction enzyme and partial sequence analysis of the gene encoding for the F protein (F gene) classified class II APMV-1 isolates in eight main distinct groups (I to VIII) with several sublineages within them (Herczeg et al., 1999). In another study, based on the partial F gene sequence of a large number of APMV-1 collected worldwide (Aldous et al., 2003), the viruses were grouped in 6 distinct genetic lineages (1 to 6) with several sublineages within them. The majority of the NDV isolates responsible of the recent outbreaks in Europe, Asia and Southern Africa are placed in lineages 3, 4 and 5 (Alexander et al., 1999; Herczeg et al., 1999; Herczeg et al., 2001; Cattoli et al., 2001; Aldous et al 2003; Abolnik et al., 2004). As an OIE listed disease the notification of ND outbreaks by the OIE member countries is mandatory. APMV-1 is distributed worldwide and ND outbreaks have been notified regularly in all continents since the first description of the disease in 1926 (Alexander, 2009). In Africa, outbreaks in poultry caused by NDV have been described in the Eastern and Southern part of the continent (Herczeg et al., 1999; Otim et al., 2004; Abolnick et al, 2004). The disease has been officially and unofficially reported in other parts of Africa and, according to the United Nations-Food and Agricultural Organization (UN-FAO), ND still remains the main poultry disease in commercial and rural chickens of West Africa (Bebay, 2006). Unfortunately, virological and epidemiological information concerning ND strains circulating in the Western and Central regions of Africa is extremely scarce and while historically ND has been reported to be widespread (Lancaster & Alexander, 1975) in recent years there has been only a single report (Snoeck et al., 2009). In that investigation, NDVs were detected by PCR in non-

commercial farms in Niger, Nigeria and Burkina Faso but viral isolates were not obtained. For this reason, characterization of the detected ND was restricted to the sequence analysis of partial F gene sequences. The authors suggested that NDVs responsible for the outbreaks in the three countries were phylogenetically distinguishable from NDVs circulating in Eastern and Southern Africa and were tentatively grouped in new sublineages within lineage 5, namely 5e, 5f and 5h (Snoeck et al., 2009).

In the present investigation we perform a full molecular characterization of ND viruses causing ND outbreaks in poultry in 7 Western and Central African countries. We also discuss the diagnostic implications of the genetic variability detected.

Materials and methods

Samples included in this study

Virological investigations using real time RT-PCR (rRT-PCR) for APMV-1 and virus isolation (VI) attempts were performed on 50 samples collected from rural chickens in Nigeria, Niger, Ivory Coast, Burkina Faso, Mauritania, Cameroon and Burundi between 2006 and 2008. Samples consisted of internal organs (such as trachea, lungs and intestines) and swabs (cloacal and tracheal) collected during surveillance programs (3,614 samples) or in cases of suspected disease (387 samples) and submitted to the OIE/FAO Reference Laboratory for Newcastle disease and Avian Influenza at the Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe) in Italy within the framework of diagnostic activities and international collaboration projects to support developing countries.

Laboratory testing

For molecular testing, viral RNA was extracted from the samples using the Nucleospin RNA II Kit (Machery-Nagel, Duren, Germany) and the detection of APMV-1 RNA was performed using a rRT-PCR protocol targeting the M gene (Wise et al, 2004). For virus isolation (VI), samples were prepared and inoculated in SPF embryonated chicken eggs according to the international regulations (OIE, 2008). The allantoic fluid of eggs showing embryo mortality was collected and tested for

haemagglutinating activity. Haemagglutinating agents were identified by means of haemagglutination inhibition (HI) tests using standard protocols (OIE, 2008). The virulence of the APMV-1 isolates for chickens was assessed by intracerebral pathogenicity index (ICPI) tests in day-old SPF chickens and by genetic sequencing of the fusion protein cleavage site (Terregino & Capua, 2009).

Nucleotide sequencing and genetic pathotyping

Viral RNA was extracted from the allantoic fluid of APMV-1 positive SPF eggs as described above. For genetic pathotyping a 266-bp hypervariable region of the F gene encompassing the cleavage site (positions 4652 to 4917, with reference to NDV strain La Sota, GenBank accession number AF077761) was RT-PCR targeted (Cattoli & Monne, 2009). Briefly, the APMV-1 F gene segment was amplified with primers NOH-For (5' TACACCTCATCCCAGACAGG 3') and NOH-Rev (5' AGTCGGAGGATGTTGGCAGC 3') using the Qiagen one-step RT-PCR kit (Qiagen, Hilden, Germany). Amplicons were sequenced using the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystem, Foster City, CA - USA). The products were cleaned using PERFORMA DTR Ultra 96-Well kit (Edge BioSystems, Gaithersburg, MD - USA) and sequenced in a 16-capillary ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA - USA).

Phylogenetic analysis

To determine the phylogenetic relationships between West and Central Africa viruses and APMV-1 viruses previously isolated in Africa and other parts of the world, the sequences of the 266-bp hypervariable region of the F gene were compared to the corresponding region of 332 representative viruses available in GenBank (http://www.ncbi.nlm.nih.gov/), for which the genotype was known. Previous studies have demonstrated that sequences of 250 base pairs give meaningful phylogenetic analyses, comparable with those obtained with much longer sequences (Seal et al., 1995; Lomniczi et al., 1998; Snoeck et al., 2008).

The lineage-based nomenclature illustrated in a previous study (Aldous et al., 2003) is adopted throughout this paper.

To confirm the robustness of the genetic groupings obtained by the phylogenetic analysis of the partial F gene sequences and to detect mutations potentially associated with antigenic and virulence changes, the complete sequences of the fusion (F), haemagglutinin-neuraminidase (HN) and matrix (M) genes of selected African APMV-1 isolates were obtained. For this purpose, RNAs extracted from infective allantoic fluid were reverse transcribed with the SuperScript III Reverse Transcriptase kit (Invitrogen, Carlsbad, CA - USA) and PCR amplification was performed using gene specific primers (Table S1). Using the full length genome positions from the La Sota vaccine strain complete genome (accession number AF077761), the nucleotide regions sequenced were as follows: the complete coding region for the genes F (positions 4544 to 6205; n = 16 isolates), HN (positions 6412 to 8127; n = 17 isolates), M (positions 3290 to 4384; n = 18 isolates). Sequence alignments were constructed and used to develop phylogenetic trees using Bayesian methods available in the computer program MrBayes version 3.1.2 (Ronquist & Huelsenbeck, 2003). Specifically, we executed a Markov chain Monte Carlo search for one million generations using two runs and four chains (temperature=0.05) and represented the results of the search as a 50% majority rule consensus tree. In each case, the best-fit model of nucleotide substitution was identified using the Akaike information criterion (Posada & Buckley, 2004) as implemented in Modeltest version 3.7 (Posada & Crandall, 2001) (parameter values available from authors on request).

Nucleotide diversity was calculated using the Maximum Composite Likelihood method implemented in the MEGA4 software package (Tamura et al., 2007).

Nucleotide sequence accession numbers

GeneBank accession numbers of the sequences analysed in this study are FJ772445 to FJ772495

Results

Detection of APMV-1 isolates in the samples

The rRT-PCR protocol detected the presence of APMV-1 RNA in 43 of 50 of the clinical samples included in the investigation (Table 1). Forty-one of the samples positive by rRT-PCR were also

positive by virus isolation. One sample from Niger and one from Nigeria, both collected in 2008 were positive by rRT-PCR but gave a negative result by virus isolation. However, 7/16 of the samples collected in 2006 (3 in Niger and 4 in Mauritania) were negative by the rRT-PCR protocol used but positive by virus isolation in eggs. The M gene rRT-PCR test was carried out on the positive allantoic fluids obtained in the virus isolation tests from these samples in which APMV-1 viruses had been confirmed by HA and HI tests, but the rRT-PCR results were also negative. When the rRT-PCR-negative/HI-positive allantoic fluids were tested by the RT-PCR protocol used for molecular pathotyping targeting the F gene positive results were obtained, thus confirming the HA/HI test results and enabling the pathotyping of the isolates.

Phylogenetic analysis of partial F gene sequences

For genetic sequencing, 28 of the 48 APMV-1 isolates were selected, based on the year of isolation, country of origin and outbreak location. The 266 bp region of the fusion gene of these isolates was sequenced and compared phylogenetically to the corresponding region of 332 representative viruses belonging to the six lineages characterized by Aldous et al., (2003). The resulting phylogenetic tree is presented in figure 1.

From the topology of the Bayesian tree, presented in figure 1, it can be seen that the two NDV isolates from Burundi were placed in sublineage 5b, a heterogeneous sublineage which is composed of isolates mainly originating from Europe, Africa, Asia and America (Aldous et al., 2003). The sequences of these viruses revealed the highest similarity at nucleotide (nt) level (97.6%) with an isolate from South Africa ZA/71/BA/94 (accession number AF532751). In contrast, the other viruses analysed clustered together in a distinct group showing a mean nt similarity of 88.7% with the sequences of lineage 5 viruses and, more specifically, 89.4% with sublineage 5b viruses. Since the mean nt diversity observed between this new group and lineage 5 (11.3%) falls within the range of nt diversity (10.3% – 43.2%) calculated for the existing 6 lineages, as shown in table 2, we considered this group represented a new lineage, provisionally named lineage 7. At present 33 isolates make up lineage 7, all originating from West Africa. Twenty-six isolates were sequenced in

the present study and the sequences of seven other viruses were obtained in a previous study (Snoeck et al., 2009). Lineage 7 can be further subdivided into four, statistically supported, sublineages (7a to 7d) with posterior probabilities >91% (figure 2).

Sublineage 7a contained 8 isolates, of which 7 were from samples obtained in 2006 in Mauritania and one from a 2008 Ivory Coast sample. The maximum nt divergence within this group was 6.4%.

Sublineage 7b was the most heterogeneous sublineage in terms of country of origin of the viruses included in this study. This lineage consisted of 14 viruses originating from four different countries, namely Niger(n=4), Nigeria (n=4), Cameroon (n=2) and Burkina Faso (n=4) obtained from samples collected in 2006 and 2008. The maximum nt divergence within this group was 4.7%.

Sublineage 7c consisted exclusively of three isolates from Burkina Faso obtained in 2006, the partial sequences of which were recently published in GenBank (Snoeck et al., 2009). The maximum divergence within this group was 1.5%.

The isolates in sublineage 7d were grouped in two further distinguishable clusters: the first (group I) contained five isolates exclusively from Niger, collected in 2006 and 2007 and the second cluster (group II) was composed of three viruses obtained in 2006, two from Niger and one from Nigeria. The maximum divergence within this group was 7.7%.

Pathotyping of APMV-1 isolates

For the APMV-1 isolates selected, the virulence was determined by molecular sequencing and ICPI tests . The deduced amino acids at the F0 protein protease cleavage site revealed that all 28 of the selected isolates had motifs that are typical of velogenic strains (OIE, 2008). Three possessed the motif ^{112}RRRKR/F^{117} and the other 25 had the motif ^{112}RRQKR/F^{117} . In order to reduce the use of experimental animals, ICPI tests were limited to representative isolates for each country that possessed the different deduced amino acid motifs at the cleavage site of the F protein (Table 3). As shown in table 3, the ICPI values were similar for the two cleavage site motifs and confirmed the virulence of these viruses. The lineage 5b isolate from Burundi produced the highest ICPI value of 1.97. Viruses placed in the new lineage 7 gave ICPI values ranging from 1.7 to 1.87.

Phylogenetic analyses of complete F, HN and M gene sequences

For lineage 7, the topology of the phylogenetic tree created using 266-bp fusion gene fragments was confirmed by analyses of the complete coding regions of the F, M and HN genes of 16, 18 and 17 representative viruses, respectively (figure 3). By this analysis it was possible to confirm the results obtained for three of the four sublineages, 7a,7b and 7d. For sublineage 7c containing viruses that were not available for sequencing in the present study, only partial F gene sequences had been deposited in GenBank (Snoeck et al., 2009) and therefore sublineage confirmation was not possible. Sequences of F, HN and M genes show comparable ratios of non-synonymous to synonymous substitutions per site (K_a/K_s=0.086, 0.109 and 0.053 respectively) indicating the absence of positive selection.

Complete nucleotide sequences of the HN genes of viruses belonging to lineage 7 showed the highest nucleotide diversity (12%) compared to the complete sequences of F and M genes of the same viruses for which 11.5% and 10% diversity was recorded respectively.

However, in the M gene several mismatches were observed in the region targeted by the primers and probe designed for the rRT-PCR detection of APMV-1 (Wise et al; 2004) (Figure 4). Additional rRT-PCR testing (data not shown) allowed us to establish that the presence of at least 3 mismatches in the probe was presumably responsible for the false negative results in the diagnostic molecular test.

Analysis of amino acid sequences

At the F protein cleavage site, sublineages 7a, 7b and 7d-group I possessed four basic amino acids (RRQKR/F), as do most of the viruses belonging to lineage 5. Sublineages 7c and 7d-group II showed five basic amino acids (RRRKR/F). The amino acid arginine at position 114 was codified by two different codons (CGA for sublineage 7c and CGG for sublineage 7d) suggesting that this point mutation occurred independently in the two groups of viruses.

Potential N-glycosylation sites at positions 85, 191, 366, 447, 471, 541, characteristic of APMV-1, were conserved in the F glycoprotein of all the NDV isolates belonging to lineage 7. The HN

glycoprotein sequence of lineage 7 viruses contained the five potential N-glycosylation sites (residues 119, 341, 433, 481, 508) that are conserved among APMV-1 viruses. One isolate from Nigeria and one from Niger contained an additional site at position 538. However, a previous study has shown that two of these sites (at positions 508 and 538) are not truly glycosylated (McGinnes *et al*., 1995).

Residues within the HN protein, which is essential for receptor binding, neuraminidase (NA) and haemagglutination (HA) activity of APMV-1 viruses, (Seal, 2004) were conserved in all the isolates analysed.

For 13 of the 14 viruses for which the complete HN and F sequences were obtained at least one amino acid substitution was detected within the known virus neutralizing epitopes of the HN and F proteins (table 4). Specifically, mutations E347D, Y350H and D569G were detected in three neutralizing epitopes of the HN protein (Iorio, et al., 1991), while amino acid substitutions K78R, D170S and L343P were detected within three epitopes located in the F protein (Yusoff et al., 1989).

Comparison of the nt and amino acids sequences available for some common vaccine strains, such as La Sota (Genbank accession number AF077761) and V4 (Genbank accession number AY225110), revealed HN protein similarity ranging between 85.5% and 87.7% with the La Sota strain and between 85.1% and 87.4% with the V4 strain . For the F protein, similarity ranged from 86.4% to 88.8% with the La Sota strain and from 86.2% to 88.5% with the V4 strain, while for M protein similarity ranged from 87% to 89.5% with the La Sota strain and from 88.8% to 92.2% with V4 strain.

The HN proteins of each of the analysed viruses had a length of 571 amino acids, which is the length most commonly observed for the virulent NDV isolates and possessed a cysteine at position 123, which is required for the formation of disulphide-linked HN homodimers. A previous study demonstrated that the amino acid cysteine at this position has an influence on both attachment activity and fusion promotion activity and increased the virulence and the pathogenicity of NDV viruses (Romer-Oberdorfer et al., 2006).

Discussion

During the period 2005-2008, 38 African countries notified ND outbreaks to the World Organisation for Animal Health (OIE-WAHID interface, Available at: http://www.oie.int/wahis/public.php. Accessed on 2nd February 2009). Nineteen of the 38 countries are in West and Central Africa. APMV-1 is clearly widespread in Africa but very limited information is available on the characteristics of the strains circulating in this continent, with the exception of those seen in South Africa (Herczeg et al, 1999; Aldous et al., 2003; Abolnik et al., 2004) and Uganda (Otim et al., 2004). In West and Central Africa, virological data concerning APMV-1 were almost completely absent. Very recently, Snoeck et al (2009) detected NDVs belonging to lineage 3 and 4 in Nigeria. But other APMV-1 viruses detected in Nigeria, Burkina Faso and Niger were tentatively placed in lineage 5 by genetic analysis of a partial fragment of the gene F. In this study Snoeck et al (2009) suggested these viruses formed novel sublineages, namely 5g, 5f and 5h or even new genetic lineages. However, isolates were not obtained and the conclusions were considered preliminary by the authors since data were obtained on partial sequence (354 nt) of the F gene only.

In our present study the results of the phylogenetic analyses indicate that the majority of the West and Central African strains belong to a novel lineage, here provisionally named lineage 7, which includes the viruses previously assigned to 5g, 5f and 5h sublineages by Snoeck et al., (2009). These findings are supported by the analysis of the complete sequences of the gene F, M and HN. Furthermore, our study indicate that viruses of this newly described lineage are circulating in a vast area of Africa, and are probably endemic in the region between Cameroon in the Central part of Africa and Mauritania in the extreme west. Of the samples from West and Central African countries that were examined in the present study only those from Burundi failed to yield virus of lineage 7.

We detected a high degree of genetic variability within lineage 7, which was similar in magnitude to that described for lineage 5 (Aldous et al., 2003).

The present study also demonstrated the important impact genetic variability of viruses may have on the diagnosis of this infection, highlighting the limit of the advanced molecular diagnostic tests on viral disease diagnosis. Kim et al., (2007) recently demonstrated that the broad genetic diversity between class I, which includes the genotype 6 described earlier (Aldous et al., 2003), and class II APMV-1 may impair the performances of rapid molecular tests, including those having conserved genes as target for PCR amplification. The results presented in our study on the sequence analysis of the matrix gene demonstrated that genetic variability within class II viruses, may be equally responsible for false negative molecular test results. A broader number of strains are necessary to assess whether this problem is restricted to lineage 7 viruses or extends to other lineages.

In a comprehensive study Czeglédy et al., (2006) concluded that the origin of the most recent virulent genotypes was not clear, but it appears that the diversification of the recent epizootic viruses and lineages, has accelerated in the last few decades and suggested that vaccination could have played a role in the generation and spread of the most recent virulent viruses. It has been recently demonstrated (Miller et al., 2007) that genetic diversity in APMV-1 strains influences the efficacy of vaccination control, not necessarily in terms of clinical protection, but mainly in term of virus shedding and subsequent spread of the infection. Miller et al., (2007) showed that virus shedding in chickens experimentally challenged with ND strains correlated with HN and F protein similarities between the challenge and vaccine strains. HN and F protein similarities less than or equal to 90%, as those observed in our study, were significantly related to higher virus shedding. In addition to this, the distinct mutations detected in the African strains analysed in the present study, located in the neutralizing epitopes of the two major antigens (F and HN glycoprotein) targeted by a protective immune response (Seal et al., 2000), present the scientific rationale for deeper investigations on the antigenic properties of these isolates and for detailed assessments of the efficacy of the current vaccines and vaccination practices in Africa.

To conclude, the present investigation provides important information on the epidemiology, diagnosis and control of NDV in Africa and highlights the importance of supporting surveillance in

developing countries for transboundary animal diseases. It provides evidence on the need to facilitate data and sequence sharing and to monitor the efficacy of validated diagnostic tests continuously. As demonstrated by the present and previous studies (Xing et al., 2008), neglecting traditional assays such as virus isolation may pose serious risks for the correct diagnosis of contagious diseases. As predicted by Aldous et al. (2003), novel genetic lineages and sublineages are emerging as a consequence of the continuous evolution of APMV-1. However, well defined criteria for lineage or sublineage designation are currently not existing for this virus. This study provide additional genetic data which might be useful for the definition of general criteria for lineage designation, similarly to what has been done for the highly pathogenic avian influenza virus A/H5N1 (WHO/OIE/FAO H5N1 Evolution Working Group, 2008).

Acknowledgments

The Food and Agricultural Organization of the United Nations partially funded this study and facilitated the collection and submission of samples (LoA/PR 37772, Project EPIAAF, Project OSRO/RAF/602). The Italian Ministry of Health (RC IZSVe 26/2007) and the European Commission (EPIZONE) also partially funded this study. The Ministry of Agriculture of the Islamic Republic of Mauritania, Niger, Nigeria, Cote d'Ivoire, Burkina Faso, Burundi, Cameroun, are also gratefully acknowledged. We are grateful to Dr G Koch (CVI-Lelystad, NL) for his technical assistance on NDV sequencing and to Dr DJ Alexander for the critical revision of this manuscript.

References

Abolnik, C., Horner, R. F., Bisschop, S. P., Parker, M. E., Romito, M., Viljoen, G. J.; 2004. A phylogenetic study of South African Newcastle disease virus strains isolated between 1990 and 2002 suggests epidemiological origins in the Far East. Arch Virol. 149(3):603-19.

Aldous, E. W., Mynn, J. K., Banks, J., Alexander, D. J.; 2003. A molecular epidemiological study of avian paramyxovirus type 1 (Newcastle disease virus) isolates by phylogenetic analysis of a partial nucleotide sequence of the fusion protein gene. Avian Pathol. 32(3):239-56.

Alexander, DJ.; 2009. Ecology and epidemiology of Newcastle disease. In: I. Capua & D.J. Alexander (Eds.). Avian Influenza and Newcastle Disease, a field and laboratory manual. Springer-Verlag, Italy, pp. 19- 26.

Alexander, D. J., Banks, J., Collins, M. S., Manvell, R. J., Frost, K. M., Speidel, E. C., Aldous, E. W.; 1999. Antigenic and genetic characterisation of Newcastle disease viruses isolated from outbreaks in domestic fowl and turkeys in Great Britain during 1997. Vet Rec. 9;145(15):417-21.

Bebay, C. E. 2006. Première évaluation de la structure et de l'importance du secteur avicole commercial et familial en Afrique de l'Ouest. Synthèse des rapports nationaux. FAO/ECTAD/AGAP (Ed.).

Cattoli, G., Manvell, R. J., Tisato, E., Banks, J., Capua, I.; 2001. Characterization of Newcastle disease viruses isolated in Italy in 2000. Avian Pathol. 30(5):465-9.

Cattoli G. & Monne I.; 2009. Molecular diagnosis of Newcastle disease virus. In: I. Capua & D.J. Alexander (Eds.). Avian Influenza and Newcastle Disease, a field and laboratory manual. Springer-Verlag, Italy, pp 127-132.

Czeglédi, A., Ujvári, D., Somogyi, E., Wehmann, E., Werner, O., Lomniczi, B.; 2006. Third genome size category of avian paramyxovirus serotype 1 (Newcastle disease virus) and evolutionary implications. Virus Res. 120(1-2):36-48.

De Leeuw, O. S., Koch, G., Hartog, L., Ravenshorst, N., Peeters, B. P.; 2005. Virulence of Newcastle disease virus is determined by the cleavage site of the fusion protein and by both the

383 stem region and globular head of the haemagglutinin-neuraminidase protein. J Gen Virol. 86 (Pt
384 6):1759-69.
385 Herczeg, J., Pascucci, S., Massi, P., Luini, M., Selli, L., Capua, I., Lomniczi, B.; 2001. A
386 longitudinal study of velogenic Newcastle disease virus genotypes isolated in Italy between 1960
387 and 2000.Avian Pathol. 30(2):163-8.
388 Herczeg, J., Wehmann, E., Bragg, R. R., Travassos Dias, P. M., Hadjiev, G., Werner, O., Lomniczi,
389 B; 1999. Two novel genetic groups (VIIb and VIII) responsible for recent Newcastle disease
390 outbreaks in Southern Africa, one (VIIb) of which reached Southern Europe. Arch
391 Virol.;144(11):2087-99.
392 Huang, Z., Panda, A., Elankumaran, S., Govindarajan, D., Rockemann, D. D., Samal, S. K.; 2004.
393 The hemagglutinin-neuraminidase protein of Newcastle disease virus determines tropism and
394 virulence. J Virol. 78(8):4176-84.
395 Iorio, R. M., Syddall, R. J., Sheehan, J. P., Bratt, M. A., Glickman, R. L., Riel, A. M.; 1991.
396 Neutralization map of the hemagglutinin-neuraminidase glycoprotein of Newcastle disease virus:
397 domains recognized by monoclonal antibodies that prevent receptor recognition. J
398 Virol.;65(9):4999-5006.
399 Kim, L. M., King, D. J., Curry, P. E., Suarez, D. L., Swayne, D. E., Stallknecht, D. E., Slemons, R.
400 D., Pedersen, J. C., Senne, D. A., Winker, K., Afonso, C. L.; 2007. Phylogenetic diversity among
401 low-virulence newcastle disease viruses from waterfowl and shorebirds and comparison of
402 genotype distributions to those of poultry-origin isolates. J Virol. 81(22):12641-53.
403 Lamb, R., Collins, P. L., Kolakofsky, D., Melero, J. A., Nagai, Y., Oldstone, M. B. A., Pringle C.
404 R., Rima, B. K. 2005. The negative sense single stranded RNA viruses, p.607-738. In C. M.
405 Fauquet, M. A. Mayo, Maniloff J., Desselberg U and L. A. Ball (Eds), virus taxonomy. Eight report
406 of the International Committee on Taxonomy of viruses. Elsevier Academic Press, San Diego, CA.
407 Lancaster, J.E., Alexander, D.J. 1975. Newcastle disease: Virus and Spread. Monograph No. 11
408 Canada Department of Agriculture : Ottawa. pp 79.

Lomniczi, B., Wehmann, E., Herczeg, J., Ballagi-Pordány, A., Kaleta, E. F., Werner, O., Meulemans, G., Jorgensen, P. H., Manté, A. P., Gielkens, A. L., Capua, I., Damoser, J.; 1998. Newcastle disease outbreaks in recent years in western Europe were caused by an old (VI) and a novel genotype (VII). Arch Virol. 143(1):49-64.

McGinnes L. W., Morrison T. G.;1995. The role of individual oligosaccharide chains in the activities of the HN glycoprotein of Newcastle disease virus. Virology. 212(2):398-410.

Mebatsion, T., Verstegen, S., De Vaan, L. T., Römer-Oberdörfer, A., Schrier, C. C.; 2001. A recombinant newcastle disease virus with low-level V protein expression is immunogenic and lacks pathogenicity for chicken embryos. J Virol. 75(1):420-8.

Miller, P. J., King, D. J., Afonso, C. L., Suarez, D. L.; 2007. Antigenic differences among Newcastle disease virus strains of different genotypes used in vaccine formulation affect viral shedding after a virulent challenge. Vaccine, 25(41):7238-46.

OIE 2008. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2008. Available on the website: http://www.oie.int/eng/normes/mmanual/A_summry.htm

Otim, M. O., Christensen, H., Jørgensen, P. H., Handberg, K. J., Bisgaard, M.; 2004. Molecular characterization and phylogenetic study of Newcastle disease virus isolates from recent outbreaks in eastern Uganda. J Clin Microbiol. 42(6):2802-5.

Posada, D., Buckley, T. R.; 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Syst Biol. 53(5):793-808.

Posada, D., Crandall, K. A.; 2001. Selecting the best-fit model of nucleotide substitution. Syst Biol 50:580-601.

Römer-Oberdörfer, A., Veits, J., Werner, O., Mettenleiter, T. C. 2006. Enhancement of pathogenicity of Newcastle disease virus by alteration of specific amino acid residues in the surface glycoproteins F and HN. Avian Dis.;50(2):259-63.

Ronquist, F., Huelsenbeck, J. P.; 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics;19(12):1572-4.

Seal, B. S, King, D. J, Bennett, J. D.; 1995. Characterization of Newcastle disease virus isolates by reverse transcription PCR coupled to direct nucleotide sequencing and development of sequence database for pathotype prediction and molecular epidemiological analysis. J Clin Microbiol. 33(10):2624-30.

Seal, B. S., King, D. J., Sellers, H. S.; 2000. The avian response to Newcastle disease virus. Dev Comp Immunol. 24(2-3):257-68. Review.

Seal, B. S.; 2004. Nucleotide and predicted amino acid sequence analysis of the fusion protein and hemagglutinin-neuraminidase protein genes among Newcastle disease virus isolates. Phylogenetic relationships among the Paramyxovirinae based on attachment glycoprotein sequences. Funct Integr Genomics. 4(4):246-57.

Snoeck, C. J., Ducatez, M. F., Owoade, A. A., Faleke, O. O., Alkali, B. R., Tahita, M. C., Tarnagda, Z., Ouedraogo, J. B., Maikano, I., Mbah, P. O., Kremer, J. R., Muller, C. P.; 2009. Newcastle disease virus in West Africa: new virulent strains identified in non-commercial farms. Arch Virol. 154(1):47-54.

Tamura, K., Dudley, J., Nei, M., Kumar, S.; 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596-1599.

Terregino, C., Capua, I.; 2009. Conventional diagnosis of Newcastle disease virus infection. In: I. Capua & D.J. Alexander (Eds.). Avian Influenza and Newcastle Disease, a field and laboratory manual. Springer- Verlag, Italy, pp. 123- 126.

Wise, M. G., Suarez, D. L., Seal, B. S., Pedersen, J. C., Senne, D. A., King, D. J., Kapczynski, D. R., Spackman, E.; 2004. Development of a real-time reverse-transcription PCR for detection of Newcastle disease virus RNA in clinical samples. J Clin Microbiol.; 42(1):329-38.

World Health Organization/World Organization for Animal Health/Food and Agriculture Organization H5N1 Evolution Working Group; 2008. Toward a unified nomenclature system for

highly pathogenic avian influenza virus (H5N1). Emerg Infect Dis. [serial on internet]. Available from: http://www.cdc.gov/EID/content/14/7/e1.htm

Xing, Z., Cardona, C., Dao, P., Crossley, B., Hietala, S., Boyce, W.; 2008. Inability of real-time reverse transcriptase PCR assay to detect subtype H7 avian influenza viruses isolated from wild birds. J Clin Microbiol. 46(5):1844-6.

Yusoff, K., Nesbit, M., McCartney, H., Meulemans, G., Alexander, D.J., Collin, M. S., Emmerson P. T., Samson, A. C. ;1989. Location of neutralizing epitopes on the fusion protein of Newcastle disease virus strain Beaudette C. J Gen Virol. 70 (Pt 11):3105-9.

468 **Table 1**. Detection of APMV-1 viruses in samples by rRT-PCR and virus isolation.
469

Country of origin	Year of isolation	Number of samples positive for APMV-1 by:	
		rRT-PCR	Virus Isolation
Burkina Faso	2008	4	4
Burundi	2008	2	2
Ivory Coast	2008	1	1
Niger	2006	3	6
	2008	11	10
Nigeria	2006	3	3
	2008	5	4
Cameroon	2008	11	11
Mauritania	2006	3	7
Totals		43/50	48 /50

470
471

Table 2 - The mean number of base substitutions per site for all sequence pairs between lineages is shown . Analyses were conducted using the Maximum Composite Likelihood method in MEGA4 software package.

LINEAGE	1	2	3	4	5	6
2	0.126					
3	0.140	0.168				
4	0.145	0.172	0.116			
5	0.159	0.193	0.127	0.103		
6	0.400	0.417	0.442	0.432	0.420	
7	0.203	0.212	0.158	0.140	0.113	0.457

476 **Table 3.** Pathotyping results of representative African APMV-1 isolates
477

Isolate	Cleavage site amino acid motif	Lineage	ICPI[a]
Ivory Coast/2601/2008	^{112}RRQKR/F^{117}	7a	1.7
Mauritania/1532-1/2006	^{112}RRQKR/F^{117}	7a	1.85
Burkina Faso/2415-580/2008	^{112}RRQKR/F^{117}	7b	1.8
Burkina Faso/2415-361/2008	^{112}RRQKR/F^{117}	7b	1.8
Niger/2602-468/2008	^{112}RRQKR/F^{117}	7b	1.7
Nigeria/3724-6/2008	^{112}RRQKR/F^{117}	7b	1.8
Cameroon/3490-147/2008	^{112}RRQKR/F^{117}	7b	1.87
Niger/1377-14/2006	^{112}RRQKR/F^{117}	7d-I	1.78
Niger/1377-8/2006	^{112}RRQKR/F^{117}	7d-I	1.8
Niger/1377-9/2006	^{112}RRRKR/F^{117}	7d-II	1.7
Burundi/4132-12/2008	^{112}RRQKR/F^{117}	5b	1.97

478 [a]Intracerebral pathogenicity in day-old chicks.
479

479 **Table 4 -** Amino acid substitutions located within the neutralizing epitopes of HN and F proteins
480

Gene	Substitution	Virus
HN	E347D	NDV_ck_2415_580_Burkina Faso_2008
		NDV_ck_2415_577_Burkina Faso_2008
	Y350H	NDV_ck_3490_147_Cameroon_2008
		NDV_ck_3490_149_Cameroon_2008
		NDV_ck_2602_605_Niger_2008
		NDV_ck_2602_625_Niger_2008
	D569G	NDV_ck_1377_1_Niger_2006
F	K78R	NDV_ck_1377_1_Niger_2006
		NDV_avian_3724_6_Nigeria_2008
		NDV_ck_2601_Ivory Coast_2008
	D170S	NDV_ck_2601_Ivory Coast_2008
		NDV_avian_1532_14_Mauritania_2006
	L343P	NDV_ck_2415_361_Burkina Faso_2008

481
482

Figure 1 – A radial phylogram constructed by Bayesian analysis of 359 NDV isolates belonging to the six lineages identified by Aldous et al., (Aldous et al., 2003). The region analysed was a 264 base pair fragment (109 to 374) at the 3' end of the fusion protein gene. Different colours are used to differentiate the viruses from distinct lineages. The new lineage identified in this study is black coloured. * indicates the isolates from Burundi sequenced in this study. The scale indicates the number of substitutions per site. Posterior probabilities of the clades are showed above branches.

Figure 2 – A radial phylogram constructed by Bayesian analysis of 33 NDV isolates (26 sequenced in this study) belonging to lineage 7. The region analysed was a 266 base pair fragment (109 to 374) at the 3' end of the fusion protein gene. The circles indicate the four sublineages (7a to 7d) identified in this study. The scale indicates the number of substitutions per site. Posterior probabilities of clades are indicated above branches.

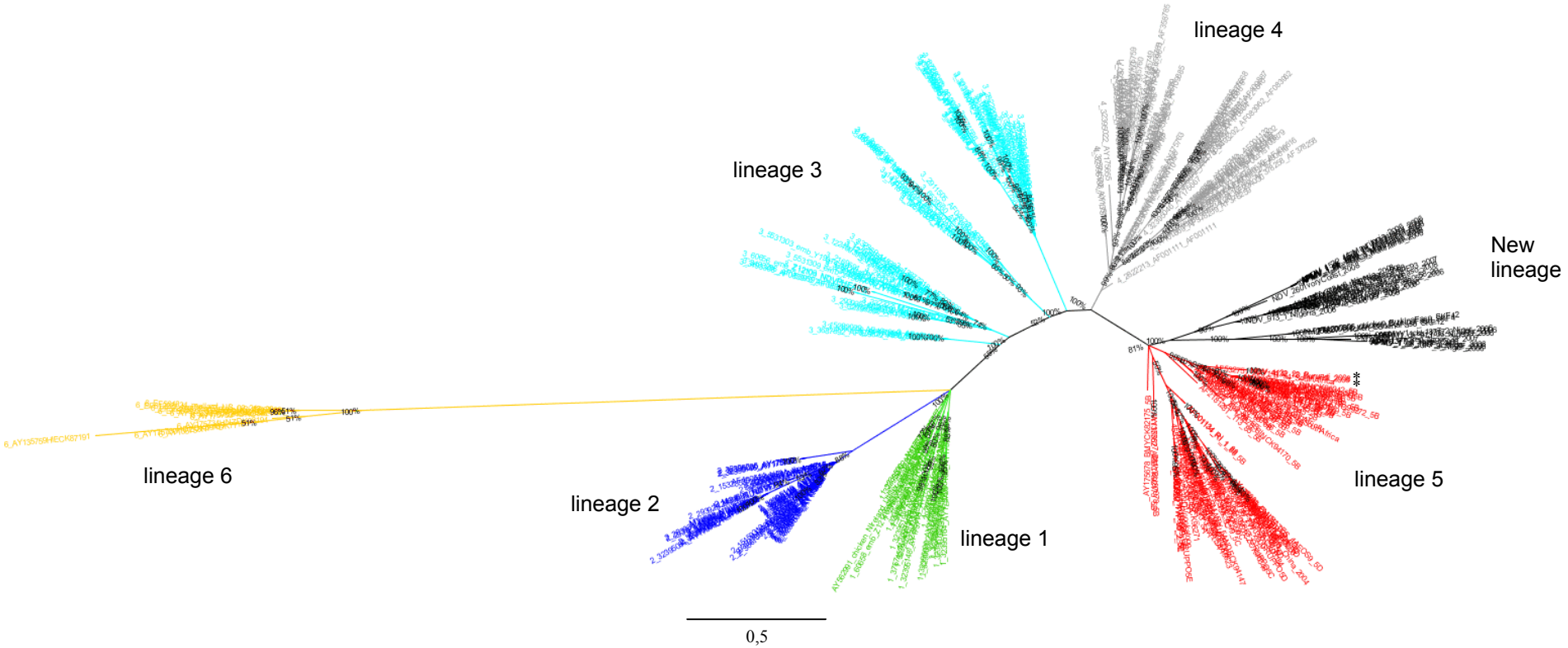
Figure 3 – Phylogenetic trees constructed by Bayesian analysis of the complete coding region of HN, M and F genes of viruses analysed in this study. DE/R49/99 and La Sota were selected as representative strains of class I and class II respectively. Blue, red and yellow rectangles indicates the sublineages 7a, 7b and 7d, respectively. The scale indicates the number of substitutions per site. Posterior probabilities of clades are indicated above branches.

Figure 4 – Nucleotide sequence alignment of the matrix gene from representative samples tested positive (+) or negative (-) by rRT-PCR. Nucleotide sequence differences between isolates are shown, while identical nucleotides are indicated by a dot. The numbering of nucleotides is relative to the complete genome sequence of the La Sota strain (reference sequence). Primer annealing regions are indicated by two arrows, while probe annealing region is marked by a rectangular.

508 **Supplemental material**

509 **Table S1**: Sequences of primers used for the amplification and sequencing of the F, M and HN gene

510



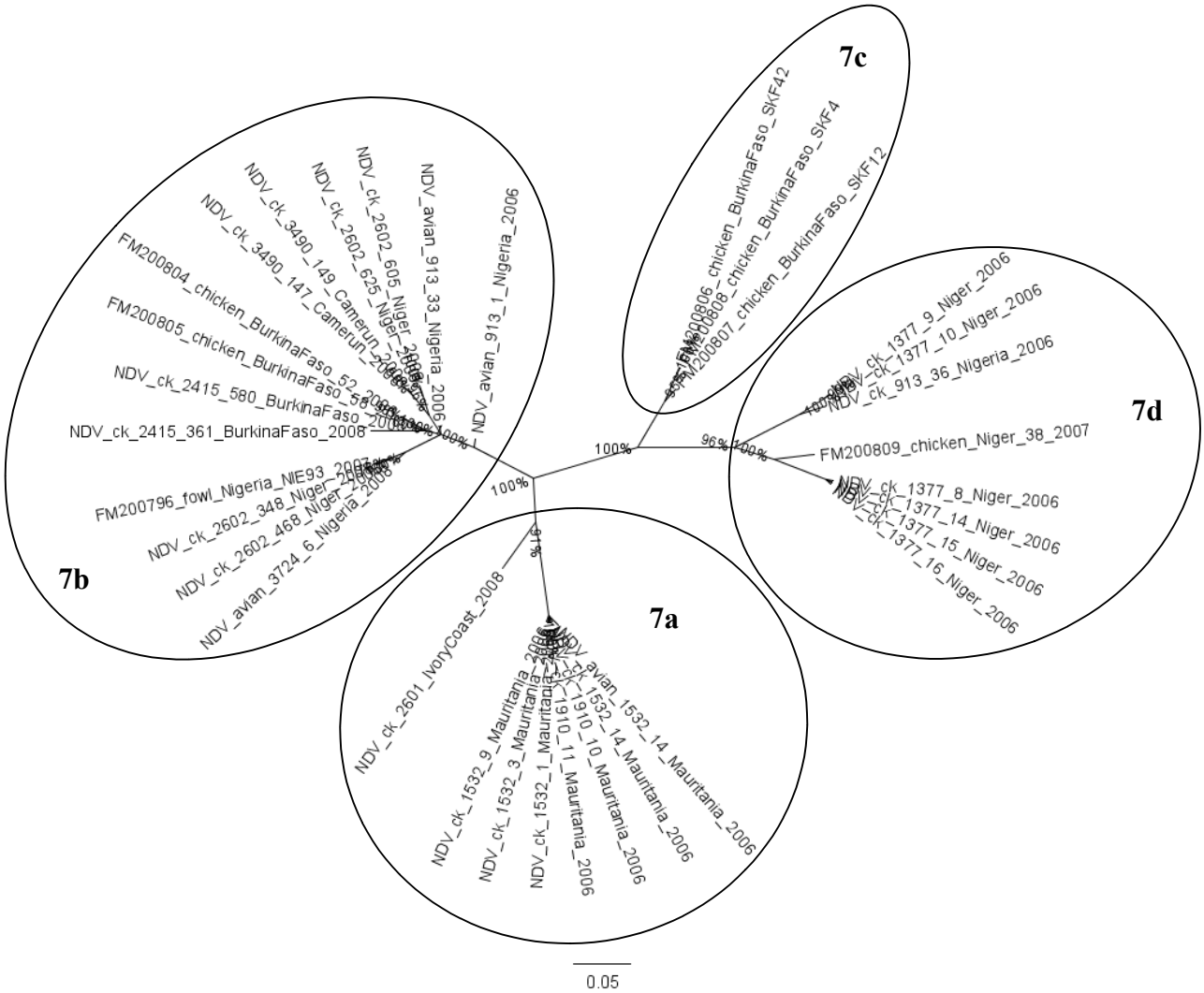
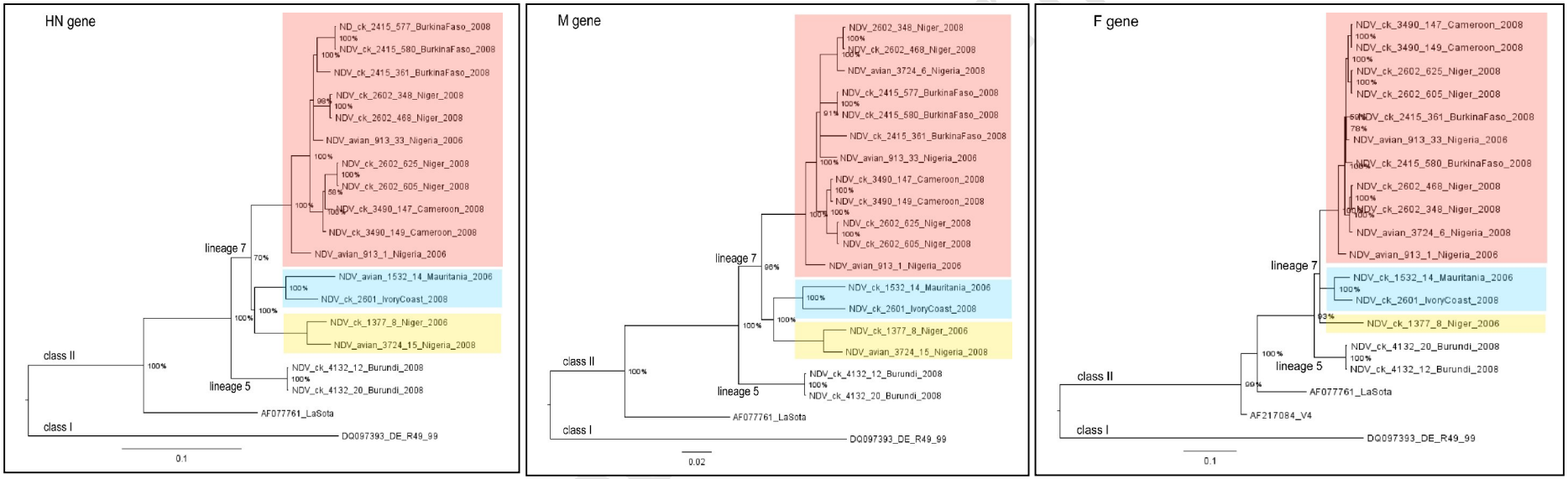


Figure 3



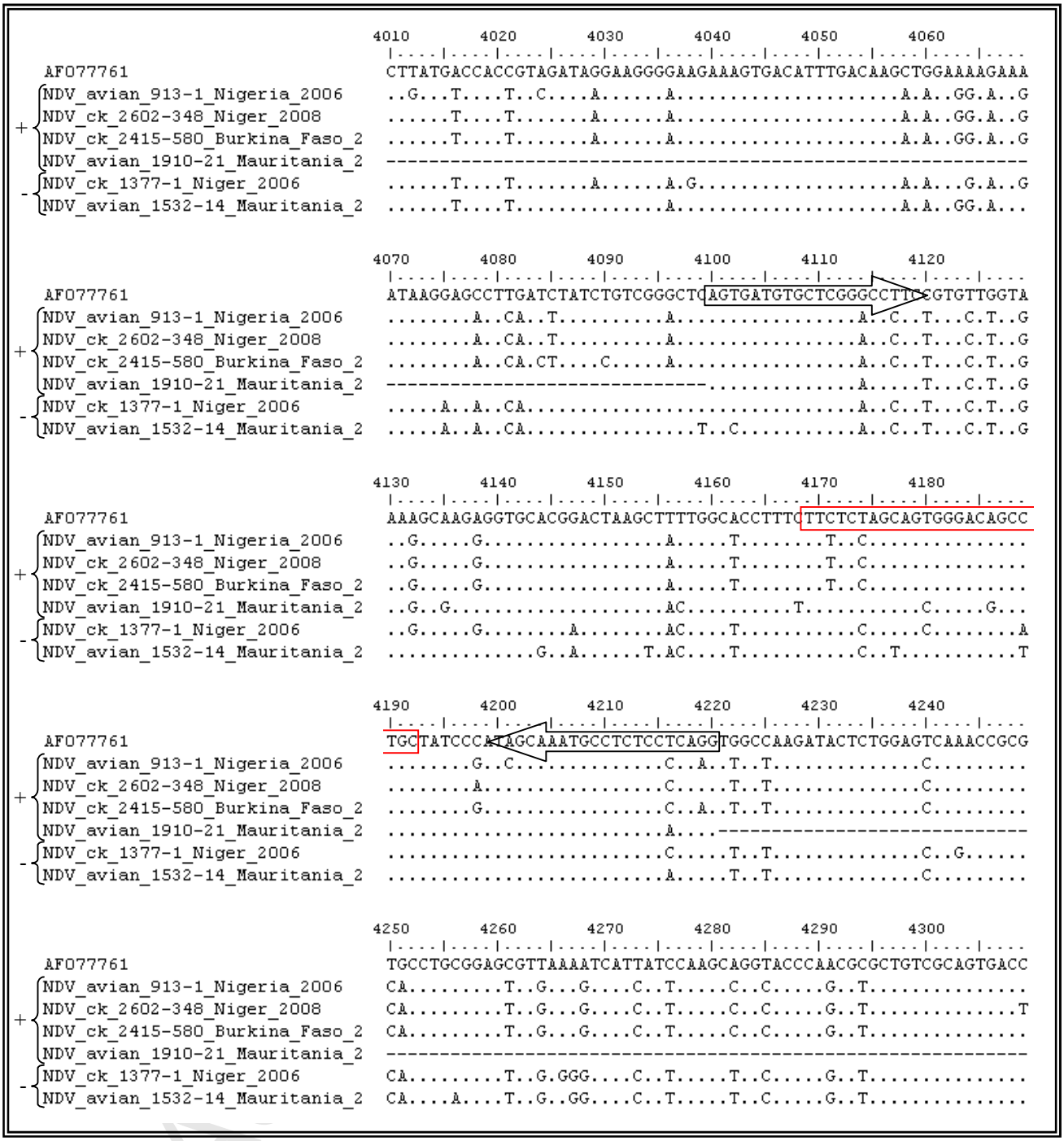


Table S1: Sequences of primers used for the amplification and sequencing of the F, M and HN gene

NAME		SEQUENCE	GENE
M-2996-F	forward	TGAAAACGACGGCCAGTCYAAGCTCCTRAGYAAGCTRGA	M
M-3910-R	reverse	CAGGAAACAGTATGACCCATCAATAGTGACATTGA	M
M-3293-F	forward	TGAAAACGACGGCCAGTAARATGGACTCATCYAGRAC	M
MF-3710-F	forward	TGAAAACGACGGCCAGTCAAAGCTGTADGGTTGTG	M
MF-4650-R	reverse	CAGGAAACAGTATGACCAAGAGGCCTGCCRTCAA	M
F-4514-F	forward	TGAAAACGACGGCCAGTGTAGAAGADTYTGGATCC	F
F-5218-R	reverse	CAGGAAACAGTATGACCGAATACYGTAGTCAAYTCRG	F
F-5443-R	reverse	CAGGAAACAGTATGACCAGGTGGCACGCATATTATT	F
F-5100-F	forward	TGAAAACGACGGCCAGTATGCAGCARTTTGTYAAT	F
F-4544-R	reverse	CAGGAAACAGTATGACCTARGTAATRAGAGCRGATG	F
F-5664-F	forward	TGAAAACGACGGCCAGTAGACYGAAGGCGCACTYAC	F
F-7979-R	reverse	CAGGAAACAGTATGACCAGRGCCACYTGCTTRTATA	F
F-5757-F	forward	TGAAAACGACGGCCAGTAGATRACAACATGTAGRTG	F
F-6449-R	reverse	CAGGAAACAGTATGACCGGCTAACYGCRCGGTCCAT	F
F-START	forward	TGAAAACGACGGCCAGTATGGGCYCYARAYCTTCTAC	F
F-460-R	reverse	CAGGAAACAGTATGACCAGCAATGCTCTCYTTAAG	F
F-586-R	reverse	CAGGAAACAGTATGACCTTTAYACAGTCCAATTC	F
F-656-F	forward	TGAAAACGACGGCCAGTTACYTAACTGARYTGACTAC	F
F-1258-R	reverse	CAGGAAACAGTATGACCACATTGCATGAWTGTCTRTC	F
HN-6398-F	forward	TGAAAACGACGGCCAGTTCMCAACATCCGTTCTACCGCATC	HN
HN-7040-R	reverse	CAGGAAACAGTATGACCGAATGYGAGTGATCTCTGCA	HN
HN-6731-F	forward	TGAAAACGACGGCCAGTTTATGAAYGCAATAACNTC	HN
HN-6979-F	forward	TGAAAACGACGGCCAGTATGAGYRCTACCCAYTACTG	HN
HN-7549-R	reverse	CAGGAAACAGTATGACCATAGATAAGATGGCYTGCTG	HN
HN-7326-F	forward	TGAAAACGACGGCCAGTGGTGGCAAAYTACCCAGGAG	HN
HN-8072-R	reverse	CAGGAAACAGTATGACCAGGGTATTGGATATTTCRGCAATGC	HN
HN-7110-F	forward	TGAAAACGACGGCCAGTATCGRAAGTCYTGCAGTGTG	HN
HN-8106-R	reverse	CAGGAAACAGTATGACCATCTCAACTAGTAAVGGRACGATYC	HN
HN-8015-F	forward	TGAAAACGACGGCCAGTCATACACRACATCRACATG	HN
HN-8509-R	reverse	CAGGAAACAGTATGACCGGTARCCCAGTYAATTTCCA	HN