Whole Genomes of Avian orthoavulavirus 1 (Newcastle Disease Virus Genotypes VIg or new genotype XXI.) in Wild Birds in Kazakhstan

Vitaliy Strochkov[1,2], Yerbol Burashev[1], Nurlan Sandybayev[1,3], Gang Xie[4*], Tracy H. Erkkila[4], Helen Cui[4], Jeanne M. Fair[4*]

[1]Research Institute for Biological Safety Problems, Ministry of Education and Science of Republic of Kazakhstan-Science Committee, Kordaiskiy Rayon, Kazakhstan;[2]LLP Scientific Diagnostic Center "Animal Expert Group", Almaty, Republic of Kazakhstan;[3]Kazakh National Agrarian University, Almaty, Republic of Kazakhstan; [4]Los Alamos National Laboratory, Biosecurity & Public Health, Los Alamos, USA

ABSTRACT

Background: The remarkable diversity and mobility of Newcastle disease viruses (NDV) includes virulent viruses of genotype VI. These viruses are often referred to as pigeon paramyxoviruses 1 because they are normally isolated and cause clinical disease in birds from the Columbidae family. Genotype VI viruses occasionally infect, and may also cause clinical disease in poultry. Thus, the evolution, current spread and detection of NDV are relevant to avian health.

Methods: Nucleotide sequencing and phylogenetic studies of three Kazakhstan isolations were performed to characterize the complete fusion (F)-protein gene as well as whole genome sequence. Sequence data were compared with 106 Fusion genes representing different NDV genotypes and sub-genotypes, as well as 225 fusion genes and 37 whole genome of class VI NDV strains from different regions of the world at different time periods. Phylogenetic trees were constructed to determine evolutionary relationships among these strains. We analysed fusion (F) protein gene and whole genome sequences, including the cleavage site.

Results: The complete genome of these 3 isolates contained 15142bp, 15085bp and 15102bp in length, similar to those of Newcastle disease virus (NDV) strains in genotypes VIg, with the gene order 3'-NP-P-M-F-HN-L-5'. The cleavage site of the fusion protein was 112KRQKR116-F117, a feature generally associated with virulent NDV strains. Phylogenetic analysis, based on genomic sequences, SNP and fusion gene sequences, revealed that three isolates should be classified as class II genotype VIg NDVs. Phylogenetic analysis revealed that the NDV from various sites in Kazakhstan was highly similar genetically and that it clustered together with NDV of genotype VIg. Based on our data analysis, VIg isolates shared highest sequence identity with Russian and Ukraine isolates of the VIg subgenotype, suggests the possible spread of velogenic NDV in this region through cross-border live bird trade.

Conclusion: Our study provides baseline information on the genetic characteristics of NDV circulating in Kazakhstan and we propose that the evolutionary and epidemiological study of virulent NDV could help to provide accurate molecular data about variants circulating in this region, thus aiding in the design of more efficient recombinant vaccines.

Keywords: Newcastle disease virus; Avian orthoavulavirus 1; Wild birds; Kazakhstan

INTRODUCTION

Newcastle disease (ND) is a highly contagious viral infection in domesticated and wild birds occurring worldwide. ND, alongside with avian influenza, is considered one of the two most devastating diseases in wild birds and poultry industry. It has become endemic in many areas, such as in Kazakhstan since its first recorded outbreak in 1980, resulting severe economic consequences. ND is caused by virulent strains of Newcastle disease virus (NDV), commonly known as avian paramyxovirus 1, which was previously designated as Avian avulavirus 1. The International Committee on Taxonomy of Viruses has recently created three genera, named Orthoavulavirus, Metaavulavirus, and Paraavulavirus, within a

new subfamily Avulavirinae of the family Paramyxoviridae. Viruses of genus Avian orthoavulavirus 1(AOAV-1) (formerly designated as Avian avulavirus 1 (AAvV-1)) was commonly knownas Avian paramyxoviruses 1 (APMV-1) or Newcastle disease viruses (NDV) - ICTV 2019 (https://talk.ictvonline.org/taxonomy/) [1-3].

Of these, AOAV-1, synonymous with Newcastle disease virus (NDV), is the most widely characterized and studied due to the economic importance of Newcastle disease (ND) caused by virulent strains of the virus. NDV has a single stranded, non-segmented, negative-sense RNA genome consisting of six genes in order of 3'to 5': nucleocapsid (NP), phosphoprotein (P), matrix (M), fusion (F), hemagglutinin-neuraminidase (HN), and polymerase (L), coding for these six structural proteins and at least one additional V protein [4,5].

Newcastle disease viruses (NDV) are genetically grouped into two divergent classes that are further classified in genotypes and sub-genotypes [6-8]. Transmission between host species appears to be an important 82 mechanism contributing to the maintenance, propagation, and spread of NDV.

Bi-directional viral exchange between domestic and wild bird populations has also been indicated by routes such as viral diffusion. Previous research has shown the presence of live vaccine strains intended for domestic poultry in wild birds [9,10]. It has been known that NDV genotype VIg can be transmitted from pigeons and doves to domesticated chicken especially at ecological contact surfaces [11,12], and that some of them dramatically gain virulence upon a few passages in chicken [13,14].

Their occurrence in pigeon population economically threatens the poultry subsector in the region. Kazakhstan is on the main pathways of transcontinental migratory routes of many wild birds. Genomic analysis of circulating NDV strains in Kazakhstan can potentially provide insights to NDV genetic evolution and pathogenic characteristics. Currently, only limited studies on the genetic variability of NDV strains in Kazakhstan have been conducted, and few NDV whole genome sequences are available.

Newcastle disease viruses are constantly evolving and different genetic groups undergo simultaneous evolutionary changes in different geographical locations making the available genetic makeup information outdated [5,15]. These evolutionary changes present challenges to prompt diagnosis. Some currently validated methods are target-oriented and might fail to detect new viral genetic variants [16-18].

Lack of comprehensive NDV genetic information further hinders understanding of Newcastle disease evolution and epidemiology. Furthermore, mixed viral infection is not uncommon in animals, and in the case of poultry, they are quite frequent [19]. Efficient and accurate detection of these pathogens is important for the creation of appropriate strategies for disease control and disease

106 surveillance in live birds, households and commercial poultry farms [20]. Full-length viral genome sequencing allows for de novo, high-quality genetic characterization of NDV. In this study, we selected three NDV isolates from Kazakhstan region and performed whole genome sequencing. Partial sequences of these three NDV strains have been previously generated and deposited into GenBank (NDV 15 submission # 2294838NDV 7, Submission # 2314571,

NDV 10 Submission # 2314574).

MATERIALS AND METHODS

Sample collection

A large number of tracheal and cloacal samples were collected for wild bird viral surveillance monitoring in 2011 and 2014, in the Shakpak Ornithological Station in the Zambyl Oblast and other regions in the Republic of Kazakhstan. The three selected NDV isolates for this study were from 2014 collection of tracheal swabs [21]. Fulica atra/KZ/WKO/4/2014 was collected from the Eurasian coot (Fulica atra) in the West KZ Oblast; Acridotheres/KZ/Zhambyl/17/2014 from common myna (Acridotheres tristis) and Columba livia/KZ/Zhambyl/32/2014 from rock dove (Columba livia), both in Zambyl Oblast.

Viral RNA extraction

This study used the previously prepared virus isolation from specific-pathogen-free embryonated chicken eggs [21]. Lyophilized samples were diluted in the original volume with water (1 ml). The suspension was used for RNA extraction using the QIAamp Viral RNA Mini Kit (Qiagen, UK), according to the manufacturer's instructions. RNA Elution was performed using water, 2 times, 40μl in each case.

Genome sequencing

A set of primers was designed to target NDV genotype VI. A subset of full NDV genome sequences from GenBank were selected as input to the NCBI Primer-BLAST. Considerations to optimize the primers included primer length, GC content and melting temperature. The primers were commercially synthesized, and diluted to a working concentration of 10 pmol. The PCR was performed using a SuperScript III One-Step RT-PCR System Kit (ThermoFisher Scientific). 24 pairs of primers were used for the VI genotype (Table S1) thus preparing a master mix for 25 reactions. Purification of PCR products was carried out using the AMPure kit (Beckman Culter), according to the manufacturer's instructions. The quality of the PCR products was checked on electrophoretic analysis. PCR products for sequencing were obtained with the BigDye® Terminator v3.1 Cycle Sequencing Kit. Purification of the sequencing reaction was carried out using the Clean Seq kit. Sequencing was performed on a 16 capillary 3130Xl Genetic Analyzer sequencer (Applied Biosystems/Hitachi). Genomic assembly, genome annotation, comparative genomics, and phylogenetic analysis were carried out using the CLC Genomic Workbench 12 program.

Table S1: NDV VI sequencing primers.

Name	Sequence	Product, bp
F5	AACAGAGAATCTGTGAGTTACGATA	811
R815	CTAAGAAAATGCGGACTGCCA	
F649	AGCAAAGGCCATGACTGCAT	714
R1362	CCAGCCTGTTGAGTGGGAAT	
F1199	TGGAGACTCGGAGTGGAGTA	753
R1975	CCCTGGGCCGTAATTATGCT	

F1780	AGCGCAGAGCCCAAAGTATT	825
R2604	CCTTCTGTGATAATGCCTCCAT	
F2444	TGGACAACGGAAGGAGTCAC	598
R3041	TTCGATCGACTTGGCTGCAT	
F2833	CTTATGTGACGCAAGGGGGT	868
R3700	AGGTGCCTGCATTACTGAGA	
F3504	ATGAAGAGGTCACTGTCGGT	667
R4170	GGTGCCAGTAGCTTAGTCCG	
F4021	GTCCACTGTAGACAAGAAGGG	823
R4843	GGAATCACCAAGGGGGTTGA	
F4591	GTTGATCGCCCGAATCATGC	773
R5363	GGCCTGTGATCAAGCCACTA	
F5225	ACTTCCCCTGCCCTAACTCA	883
R6107	CCAGGCTAAGTGCACCGAAA	
F5915	ATCGTGACAGGCAACCTTGAT	797
R6711	ATTTAGCAGTGCCAGCGGAG	
F6552	GCATAGCATGAGGGCCAGTA	772
R7323	CACTCCCGGGTAATTTGCCA	
F7225	ATGGTGCACGGAAGGTTAGG	738
R7962	GGTTACACGACTGCGGGATA	
F7869	GAGAGGGGTCTTCGGGACAA	804
R8672	CCAAACACCTGGGATGGAGTA	
F8563	TGAATCTGCCACTCCTGACA	727
R9289	TTCAAGCAGCTGAACAGCAC	
F9141	CTACGGCGACACATCTCAGAA	900
R10040	GCCCATTAACCTTCACCTCCT	
F9838	TAGGCGGAACCTTCTCTCTGA	796
R10633	TTGTACCATGCAGGCAACCC	
F10521	GTGCTAGAGGGGGCATTGAG	832
R11352	ATATTTGGGCTTGCGACAGTCT	
F11125	TCAGCTGGGAGGACTGAGTAA	873
R11997	GAAGCAGCCCTCCTCTCTTG	
F11856	ACCTTTGACAGGGGGTAGGA	810
R12665	ACTGGCCAATCAACTTCCCA	
F12464	TTCGAGCTACTTGGGTTGGC	810
R13273	TGGCTTCACAGCATCTGACA	
F13132	GCTTCAGCTGATTTCCCGGT	881
R14012	ATACAGACCTGTAGGGCACCA	
F13897	AGCGGAAGTGCCGTGTAAAG	614
R14492	GTCCCCCGGCTTCAATCAA	
F14300	GAGGTGGTGAGGATGGCAAA	884
R15183	AGATTTGGTGAATGACAGGACT	

Phylogenetic analysis

Whole genome tree: The three newly sequenced NDV genome

sequences combined with 37 complete genomes of genotype VI NDV downloaded from GenBank were used for complete genome-based analysis. Phylogenetic trees were constructed by Maximum Likelihood Phylogeny 1.3 on CLC Genomics Workbench 12, using the 'Neighbour Joining' method and the Jukes Cantor model with Gamma distribution parameter 1.0 and 100 bootstrap replications to assign confidence levels to branches.

SNP tree analysis: The same 37 available complete genomes of NDV in genotype VI were used for SNP based analysis. The SNP tree was generated by using PhaME-1 version 1(https://github.com/mshakya/PhaME-1) with default settings. The phylogenetic trees were inferred with RAxML-Best tree and bootstrap valve mapped on Bipartitions tree. F gene tree analysis: Reference NDV genomes listed in the Table S1 of Ogali, et al. were selected and downloaded from GenBank (n=106) [22]. To represent all NDV genotypes. At least three sequences representing each genotype were used in the analysis. All known NDV isolated from KZ deposited in GenBank are downloaded. There are total 76 NDV strains isolated in Kazakhstan, 106 previously described NDV sequences representing different NDV genotype groups were used for comparison. Multiple alignment and comparison of three study sequences and downloaded GenBank sequences were performed using MUSCLE v3.8.31 [23]. Phylogenetic and molecular evolutionary analyses were conducted using MEGA (Molecular Evolutionary Genetics Analysis) version 7.0 [24]. We constructed phylogenetic trees using the maximum likelihood (ML) method and estimated the tree using the Kimura 2-parameter model of nucleotide substitution with gamma-distributed rate variation among sites. We employed a bootstrap resampling process (100 replications) to assess the robustness of individual nodes of phylogeny. For building VI group subtree, all sequences belong to VI group in the Table S1 of Dimitrov, et al. were downloaded from GenBank and curated resulting in a large datasets (n =225) [8]. Together with three sequences obtained in the current study, all sequences were aligned and constructed a phylogenetic tree according methods described above (Table 1).

Table 1: Strains isolated from the samples.

Name	Sampling place	Date	Bird species
Fulica atra/KZ/WKO/4/2014	Aktobe region, Martuk district, Bulak Lake	Apr. 2014	Eurasian coot
Columbalivia/KZ/Zhambyl/32/2014	Zhambyl, Shakpak ornithological station	Oct. 2014	Rock dove
Acridotheres/KZ/Zhambyl/17/2014	Zhambyl, Shakpak ornithological station	Oct. 2014	Common myna

Genbank accession number

The nucleotide sequences determined in this study are available in the GenBank under accession numbers

Acridotheres/KZ/Zhambyl/17/2014 (NDV-15)   MT081320

Columbalivia /KZ/Zhambyl/32/2014 (NDV-7)   MT081321

Fulica atra/KZ/WKO/4/2014 (NDV-10)   MT081322

RESULTS

Geographic sampling location

Three isolates were selected from positive samples from a previous study [21]. These three isolates that had been previously sequenced its fusion gene, were selected from the Research Institute of Biological Safety Problems (RIBSP) collection for whole genome sequencing and analysis. These samples collection sites are shown in Figure 1.
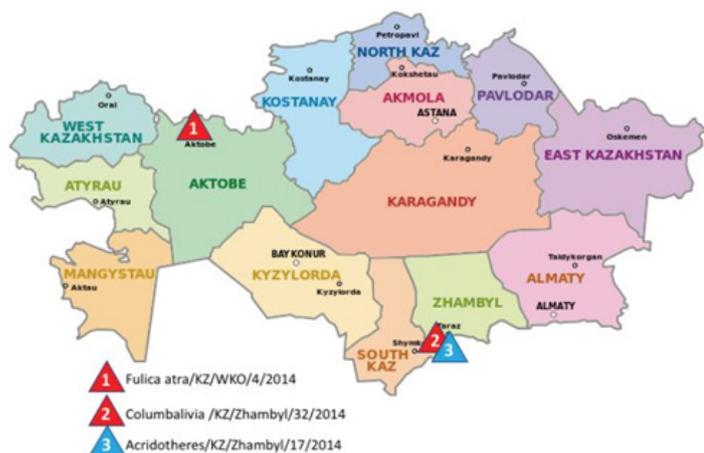

Figure 1: Map showing sampling locations of the three strains in this project. The samples are annotated with number and color (representing strain relationship).

Genome length

Whole genome assembly reveals that genome sizes for three Kazakhstan ND strains Fulica atra/KZ/WKO/4/2014, Columbalivia/KZ/Zhambyl/32/2014, and Acridotheres/KZ/Zhambyl/17/2014 are 15142bp, 15085bp and 15102bp respectively. The length variation occurs at the upstream UTR region of nucleocapsid protein (NP) gene (Table 1).

Coding regions

Similar to other NDV strains, the genome of these three isolates contained six genes in the order of 3'-NP-P-M-F-HN-L-5' as expected. The protein coding capacity of the genome of these three isolates was the same at approximately 90.6%. The genomic features and protein characteristics of three isolates are summarized in Table 2.

Table 2: Genetic characteristics of three Kazakh NDV strains.

Gene	Gene length (nt)	ORF (nt)	3' UTR (nt)	5' UTR (nt)	Intergenic region (nt)	Amino acid length (aa)
NP	1752	1470	66	216	2	489
P	1451	1188	83	180	1	395
M	1241	1095	34	112	1	364
F	1792	1662	46	84	31	553
HN	2002	1716	91	195	47	571
L	6703	6615	11	77		2204

Non-coding regions

The AOAV-1 genome started from the 3′-leader region, and the genes were transcribed into separate mRNAs through a start–stop–restart mechanism guided by the conserved gene start (GS) and gene end (GE) of the genes. Located between GS and GE were the non-coding intergenic sequences (IGS), which ranged from 1 nt to 47 nt. The open reading frame of each gene overhanged with the 3' and 5' untranslated regions (UTRs) on their respective ends. The sequences of the GS and GE as well as of the IGS in 3 strains are summarized in Table 2. The GS for the five NP, P, M, F and HN genes was ACGGGTAGAA, whereas that for the L gene was ACGGGTAGGA. The GE for NP, M and L was TTAGAAAAAA, whereas that for P, HN and F gene was TAAGAAAAAA. The lengths of intergenic sequence (IGS) of P-M and M-F were 1 nt, while the IGS of NP-P, F-HN and HN-L were 2 nt, 31 nt and 47 nt, respectively (Table 3).

Table 3: Amino acid sequence comparison of Fulica atra/KZ/WKO/4/2014 and Acridotheres/KZ/Zhambyl/17/2014.

NDV gene	Percent identity	Number of amino acid substitutions
Nucleocapsid Protein (NP)	98.8	6
Phosphoprotein (P)	95.1	19
Matrix Protein (M)	99.7	1
Fusion Protein (F)	98.4	9
Hemagglutin Neuraminidase (HN)	96.8	18
Large Polymerase Protein (L)	98.4	17

Comparative study

A comparative analysis of the complete genomes of the strain Fulica atra/KZ/WKO/4/2014 and Acridotheres/KZ/Zhambyl/17/2014 reveal 99% coverage and 96.4% identity between them. The genome comparison reveals the Columbalivia/KZ/Zhambyl/32/2014 is identical to Fulica atra/KZ/WKO/4/2014, except for one nucleotide substitution in the hemagglutinin-neuraminidase (HN) gene and 57bp longer at the upstream UTP region of NP gene in Fulica atra/KZ/WKO/4/2014.

Megablast against GenBank revealed these 3 strains having 97% identity with strains isolated in Russia in 2009–2011 and 95% identity with a strain isolated in Ukraine in 2013. The pairwise comparison among 3 strains, NDV/Altai/pigeon/770/2011, Pi/Rus/Kemerovo/0267/09, and pigeon/Ukraine/Doneck/3/968/2007 reveal that NP and M genes were the most conserved with sequence identities that ranged from 97.34% to 99.73% (nucleotide) and 1 to 13 (amino acid changes) in all strains. In contrast, the P-gene was the most variable with sequence identities that ranged from 91.92% to 96.46% (nucleotide) and 14 to 32 (amino acid changes), depending on the strain. Analysis of sequence identities using available complete genome sequences from the GenBank showed that Altai/pigeon/777/2010, Pi/Rus/Kemerovo/0267/09, and Altai/pigeon/777/2010 were most closely related to NDV- 10(97.18%), NDV-15 (97.23%), and NDV-7 (97.18%) respectively.

F gene comparison

Analysis of the primary sequences of the fusion proteins (F protein) revealed that all strains exhibited the cleavage site sequence 112KRQKR116- F117, which is typical for pathogenic strains of NDV. Analysis of the functional domains of the F-gene sequence of three strains showed that the deduced amino acid sequences were mostly similar with the consensus amino acid sequences derived from 37 NDV strains in genotype VI; except that the F

Table 4: Amino acid substitutions in the functional domains of the F protein.

Strains	Signal peptide (1-31)			Fusion peptide (117-141)	Hra (143-185)	HRc (471-500)				Transmembrane domain (501-521)
	2	13	22	125	146	476	485	490	499	509
Consensus	G	L	I	V	Q	N	N	K	S	V
Fulica atra/KZ/WKO/4/2014	S	P	T	V	R	S	N	K	P	V
Columbalivia/KZ/Zhambyl/32/2014	S	P	T	V	R	S	N	K	P	V
Acridotheres/KZ/Zhambyl/17/2014	S	L	I	V	Q	S	N	R	S	A
Pi/Rus/Kemerovo/0267/2009	S	L	I	V	Q	S	S	K	S	A
Altai/pigeon/770/2011	S	L	I	I	Q	S	N	K	P	V

protein of Fulica atra/KZ/WKO/4/2014 and Acridotheres/KZ/Zhambyl/17/2014 had 0/1 at the transmembrane domain, 3/1 at the signal peptide, and 0/0 at the fusion peptide respectively. Analysis of the three heptad repeat regions (HR) showed a total of 1/0 substitutions in HRa (143–185 amino acids), and 2/2 substitutions in HRc (471–500 amino acids) respectively (Table 4).

Phylogenetic analysis

We illustrate the relationship between all 106 Kazakh isolates deposited in Genbank plus 3 Kazakh strains obtained in this study and other 76 NDVs available in GenBank, using a phylogenetic tree of F gene sequence of the Kazakh isolates obtained in the study and the other GenBank sequences of NDV (Figure 2).
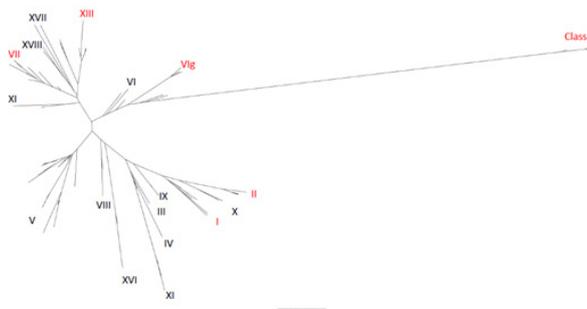

Figure 2: Molecular Phylogenetic analysis of partial fusion (F) gene nucleotide sequences of NDVs (n=76+106+3) by Maximum Likelihood method. Clads having sequences isolated from KZ are red color labeled. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model (Kimura). The tree with the highest log likelihood (-20419.91) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 231 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1638 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Huson, et al.) [24].

From the 241 tree, all three sequences were grouped together with other sub-genotype VIg sequences from KZ. Zoom-in subgenotype

VIg clad, Columbalivia/KZ/Zhambyl/32/2014 and Fulica atra/KZ/WKO/4/2014 were more closely identical (Figures 2 and 3) to previous KZ strains (MK693035.1_AAvV 1/rock_dove/Southern_Kazakhstan/6444/2014; MK693038.1_AAvV-1/golden_eagle/Southern_Kazakhstan/6451/2014; KT965728.1 columbalivia/KZ/Zhambyl/32/2014; KT965731.1 buteo rufinus/KZ/Zhambyl/41/2014;KT965730.1 anas acuta/KZ/EKO/1/2014) sequences.
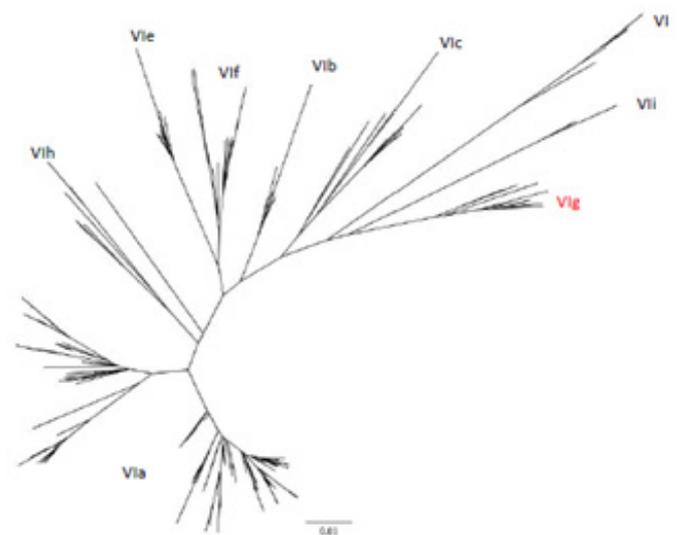

Figure 3: Molecular Phylogenetic analysis of partial fusion (F) gene nucleotide sequences of NDVs (n=225+3) by Maximum Likelihood method. Clads having sequences isolated from KZ are red color labeled. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model (Kimura). The tree with the highest log likelihood (-20419.91) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 231 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1638 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Huson, et al.) [24].

While Acridotheres/KZ/Zhambyl/17/2014 was more closely identical to previous KZ strains (JN806237.1 AAvV- 1/pigeon/Almaty/1065/05; KT965732.1 columbalivia/KZ/Zhambyl/27/2014; and KT965729.1 acridotheres/KZ/Zhambyl/17/2014) sequences.

This study VIg sequences together with previous KZ strains formed a distinct clade branching with 100% bootstrap value at the defining node from other subgenotypes of genotype VIg including genotype VIg viruses from Europe and Africa, such as JF824032.1 Pi/Rus/Vladimir/687/05, JQ039385.1 dove/Nigeria/VRD07163/2007 and JF824013.1 Pi/Rus/Kemerovo/0267/09.

By checking with the latest NDV classification and nomenclature system proposed by international consortium, F gene of Columbalivia/KZ/Zhambyl/32/2014/Fulica atra/KZ/WKO/4/2014 (NDV-7 and NDV-10) and Acridotheres/KZ/Zhambyl/17/2014 (NDV-15) are identified as XXI.1.1_VI_g_735_KT965728_pigeon_Kazakhstan_Zhambyl_32_2014 and XXI.1.1_VI_g_734_KT965727_pigeon_Kazakhstan_EKO_15_2014 respectively in the Supplemental Dataset S2 (class ii) [25].

Both sequences belong to new genotype XXI. In addition, a phylogenetic analysis of three Kazakh strains and 37 NDV genotype VI strains was conducted on the basis of the complete genome sequence using the neighbour-joining algorithm with bootstrap 500 replicates on CLC Genomic Workbench v12. All 3 isolates were grouped under class II NDVs in genotype VIg (Figure 4).

Figure 4: Phylogenetic tree of the nucleotide 588 sequences of complete genomes of the Newcastle disease virus (N=37+3). The highlighted sequences are from the three isolates targeted by this project.
3 KZ strains phylogenetically close to other VIg strains, such as KT962979 Altai/pigeon/777/2010, KJ920204 NDV/Altai/pigeon/770/2011, and JF827027 Pi/Rus/Kemerovo/0267/09

isolates. Phylogenetic analysis using whole genome SNP yielded the same tree topology and phylogenetic groupings (Figure 5).
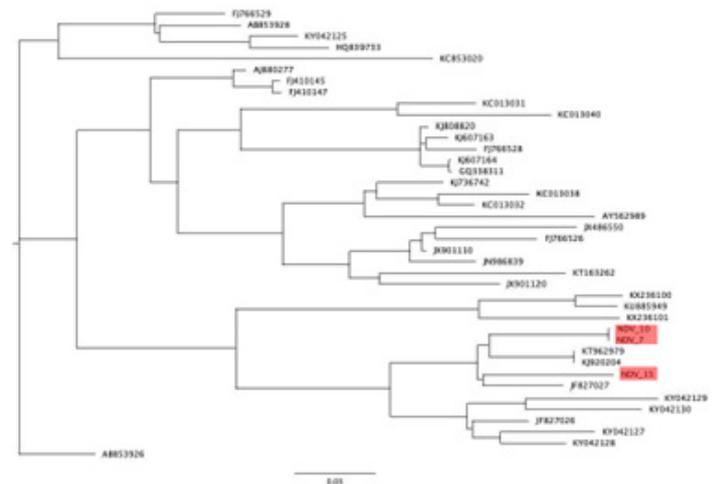

Figure 5: SNP tree of the nucleotide sequences of complete genomes of the Newcastle disease virus (N=76+3). The highlighted sequences are from the three isolates targeted by this project.

DISCUSSION

This study focused on molecular and genomic characterization of three NDV isolates from wild birds from a larger study in Kazakhstan. Phylogenetic analysis using 3+37 complete NDV genome sequences in genotype VI revealed that three Kazakh strains belong to subgenotype VIg. SNP Phylogenetic analysis using the same input yielded the same phylogenetic groupings. Subgenotype analyses using curated VI gene sequences showed that all three strains belong to the subgenotype VIg. This topology of the F-gene phylogenetic tree indicated the viruses from three strains isolated from KZ grouped together and clustered within rest of sub-genotype VIg. These three viruses were closely related to viruses isolated from pigeons in Russia, Ukraine, and Kazakhstan during 2005–2014 [26-28].

These results were consistent with the results of the complete genome analysis, as well as updated unified phylogenetic classification system based on the complete fusion gene sequences of isolates representing NDV class II [25]. The new phylogenetic classification system has shown these three NDV isolates belonging to new genotype XXI Dimitrov, et al. together with other XXI sequences from Ukraine, Russia, Kazakhstan, as well as distanced regions from Egypt, Nigeria, and Pakistan [25]. As these kind of close genetic relationships among these isolates from different regions, it could be of epidemiological significance and certainly suggest a recent common ancestry during their evolution [29,30]. Multiple virulent subgenotype class I, class II-I, II, VIg, VII and XIII NDV virus have been showing circulating in KZ. It has been known that class II genotypes I, V, VI and VII have been isolated in many countries and avian species, suggesting viral dispersion across geographic areas and among diverse hosts [8,31]. Viruses from sub-genotype VIg also have been known being isolated from pigeons and a dove in Nigeria, Russia, and Ukraine during 2005–2011 Van Borm, et al., Pchelkina, et al., Yurchenko, et al. and Iran during 2012 to 2014 Mayahi, et al., as well as in different West and Central African countries (Snoeck, Adeyanju et al. 2013, Hicks, et al. [26-29,32,33]. An interesting observation from this study is that two of NDV virus samples from different bird species and oblasts (Aktobe oblast in Western KZ and Zhambyl oblast in
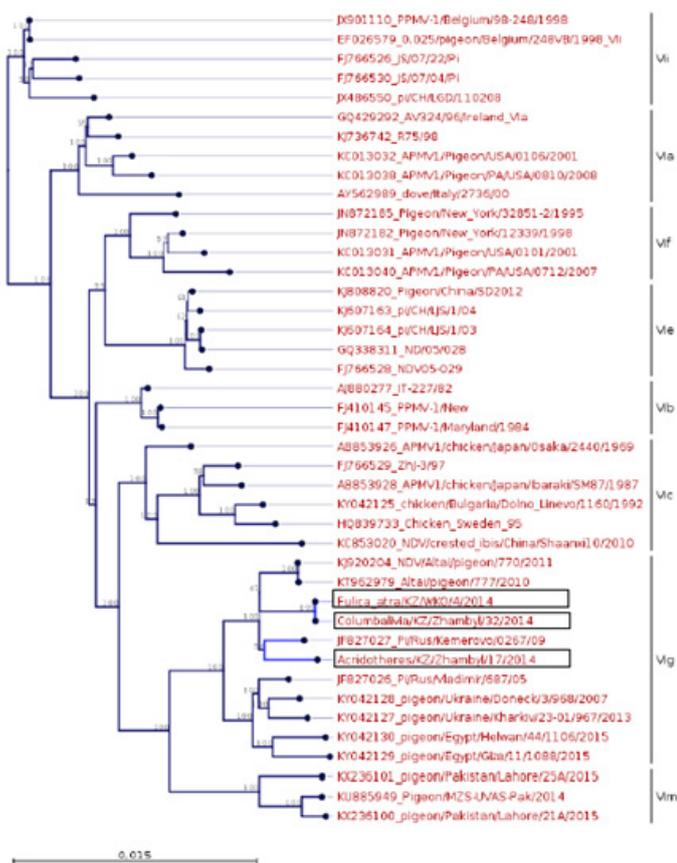
South KZ) were identified as belonging to the same strain. These two specimens with nearly identical genomes originated from birds of different species separated by about 1500km. This illustrates how the virus can move rapidly between different species of wild birds over long distances. Such transitions from an established host into a new recipient host have been inferred to occur, as evidenced by host-adapted lineages: Columbiformes-adapted genotype VI viruses, (Hicks, et al. [33]. A possible explanation for the spread of these viruses over long distances is contact between columbid birds during competitive flights, exhibitions, or due to such birds 'intensive international trade [34]. Other possibilities are the international trade between countries of live birds from other species or avian products, either through legal or illegal import and export routes. Due to the close proximity of multiple bird species and the frequent turnover of naive hosts, these live bird markets can represent a distinct ecological niche that could generate minimal cross-species transmission for AOAV-1 [33].

In addition, NDV of genotype VI have been isolated from birds from non-Columbidae species kept in captivity and from wild birds, including perdrixes, faisans, swans, falcons, blackbirds, cockatoos, budgerigars, raptors, pertridges, crested ibises, waterfowl, starlings, pintails, gannets, and buzzards [35-38]. Genotype VI viruses have previously been reported to circulate in apparently healthy pigeons Teske, et al., Byarugaba, et al., Wang, et al. and recent work on the VI-restricted model of genotype Hicks, et al., and a previously published phylodynamic study of genotype VI Chong, et al., showed that global migration of VI genotype could be driven by pigeon transport, resulting in viral dispersal from Europe to other regions, including Africa and East Asia [33,39-41].

Because live animal movement (i.e., trade and wild bird migration) are the most likely drivers of global AOAV-1 spread, inanimate, vaccine and non-avian animal contamination may also serve as mechanisms for viral migration Yates, et al., Alexander, et al., Alexander, et al., Jorgensen, et al., Li, et al., Sharma, et al., Pedersen, et al., Zhao, et al., these could explain why major poultry trade partners are also reflected in the global diffusion network [42-49].

Specifically, Europe, which is a net exporter of poultry to the rest of the world Bahl, et al., was supported as an important source of class II AOAV-1 [50]. The appearance of fecal and oral swabs of genotype VI viruses indicates viral replication, which could result in virus transmission and possible outbreaks in poultry, as seen previously Hicks, et al., is occurring in pigeons [33].

Although sporadic, NDV outbreaks in poultry caused by viruses of genotype VI have been reported, and the potential of the virus to cause clinical disease in poultry must not be under estimated, as NDV can be transmitted to poultry via wild birds; therefore, studying the strain variance in wild birds will give an indication of how to design diagnostic tests for poultry and enable improved readiness for any future outbreaks [43,51-54]. Especially considering countries in Africa, Europe and the Middle East are particularly underrepresented, continued surveillance and sampling among regions with low representation would improve future analyses and may be showing more important in the global diffusion of AOAV-1 than what have reported.

CONCLUSION

The results of this study also could be helpful for developing efficient recombinant vaccines that are homologous to viruses that are circulating in the country.

(http://www.oie.int/wahis_2/public/wahid.php/Diseaseinformation/Diseasetimelines).

REFERENCES

1. Afonso CL, Amarasinghe GK, Bányai K, Bào Y, Basler CF, Bavari S, et al. Taxonomy of the order Mononegavirales: Update 2016. Arch Virol. 2016;161(8): 2351-2360.

2. Goraichuk I, Sharma P, Stegniy B, Muzyka D, Pantin-Jackwood MJ, Gerilovych A, et al. Complete Genome Sequence of an Avian Paramyxovirus Representative of Putative New Serotype 13. Genome Announc. 2016;4(4): e00729-16.

3. Thampaisarn R, Bui VN, Trinh DQ, Nagai M, Mizutani T, Omatsu T, et al. Characterization of avian paramyxovirus serotype 14, a novel serotype, isolated from a duck fecal sample in Japan. Virus Res. 2017;228: 46-57.

4. Chambers P, Millar NS, Bingham RW, Emmerson PT. Molecular cloning of complementary DNA to Newcastle disease virus, and nucleotide sequence analysis of the junction between the genes encoding the haemagglutinin-neuraminidase and the large protein. J Gen Virol. 1986;67(Pt 3): 475-486.

5. Miller PJ, Decanini EL, Afonso CL. Newcastle disease: Evolution of genotypes and the related diagnostic challenges. Infect Genet Evol. 2010;10(1): 26-35.

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

7. Diel DG, da Silva LH, Liu H, Wang Z, Miller PJ, Afonso CL. Genetic diversity of avian paramyxovirus type 1: Proposal for a unified nomenclature and classification system of Newcastle disease virus genotypes. Infect Genet Evol. 2012;12(8): 1770-1779.

8. Dimitrov KM, Ramey AM, Qiu X, Bahl J, Afonso CL. Temporal, geographic, and host distribution of avian paramyxovirus 1 (Newcastle disease virus). Infect Genet Evol. 2016;39: 22-34.

9. Garcia SC, Lopez NR, Morales R, Olvera MA, Marquez MA, Merino R, et al. Molecular epidemiology of Newcastle disease in Mexico and the potential spillover of viruses from poultry into wild bird species. Appl Environ Microbiol. 2013;79(16): 4985-4992.

10. Ayala AJ, Dimitrov KM, Becker CR, Goraichuk IV, Arns CW, Bolotin VI, et al. Presence of vaccine-derived newcastle disease viruses in wild birds. PLoS One. 2016;11(9): e0162484.

11. Zhang P, Xie G, Liu X, Ai L, Chen Y, Meng X, et al. High Genetic Diversity of Newcastle Disease Virus in Wild and Domestic Birds in Northeastern China from 2013 to 2015 Reveals Potential Epidemic Trends. Appl Environ Microbiol. 2015;82(5): 1530-1536.

12. Akhtar S, Muneer MA, Muhammad K, Tipu MY, Rabbani M, Ul-Rahman A, et al. Genetic characterization and phylogeny of pigeon paramyxovirus isolate (PPMV-1) from Pakistan. Springerplus. 2016;5(1): 1295.

13. Dortmans JC, Rottier PJ, Koch G, Peeters BP. Passaging of a Newcastle disease virus pigeon variant in chickens results in selection of viruses with mutations in the polymerase complex enhancing virus replication

and virulence. J Gen Virol. 2011;92(Pt 2): 336-345.

14. Meng C, Qiu X, Yu X, Li C, Sun Y, Chen Z, et al. Evolution of newcastle disease virus quasispecies diversity and enhanced virulence after passage through chicken air Sacs. J Virol. 2016;90(4): 2052-2063.

15. Miller PJ, Kim LM, Ip HS, Afonso CL. Evolutionary dynamics of Newcastle disease virus. Virol. 2009;391(1): 64-72.

16. Kim LM, Afonso CL, Suarez DL. Effect of probe-site mismatches on detection of virulent Newcastle disease viruses using a fusion-gene real-time reverse transcription polymerase chain reaction test. J Vet Diagn Invest. 2006;18(6):519-528.

17. Rue CA, Susta L, Brown CC, Pasick JM, Swafford SR, Wolf PC, et al. Evolutionary changes affecting rapid identification of 2008 Newcastle disease viruses isolated from double-crested cormorants. J Clin Microbiol. 2010;48(7): 2440-2448.

18. Ahmed Khan T, Rue CA, Rehmani SF, Ahmed A, Wasilenko JL, Miller PJ. Phylogenetic and biological characterization of newcastle disease virus isolates from Pakistan. J Clin Microbiol. 2010;48(5): 1892-1894.

19. Allison AB, Gottdenker NL, Stallknecht DE. Wintering of neurotropic velogenic Newcastle disease virus and West Nile virus in double-crested cormorants (Phalacrocorax auritus) from the Florida Keys. Avian Dis. 2005;49(2): 292-297.

20. Dimitrov KM, Sharma P, Volkening JD, Goraichuk IV, Wajid A, et al. A robust and cost-effective approach to sequence and analyze complete genomes of small RNA viruses. Virol J. 2017;14(1): 72.

21. Orynbayev MB, Fereidouni S, Sansyzbai AR, Seidakhmetova BA, Strochkov VM, Nametov AM, et al. Genetic diversity of avian avulavirus 1 (Newcastle disease virus genotypes VIg and VIIb) circulating in wild birds in Kazakhstan. Arch Virol. 2018;163(7):1949-1954.

22. Ogali TN, Wamuyu LW, Lichoti JK, Mungube EO, Agwanda B, Ommeh SC. Molecular Characterization of Newcastle Disease Virus from Backyard Poultry Farms and Live Bird Markets in Kenya. Inter J Microbiol. 2018; 2368597.

23. Edgar RC. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5(1): 113.

24. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. Genome Res. 2007;17(3): 377-386.

25. Dimitrov KM, Abolnik C, Afonso CL, Albina E, Bahl J, Berg M, et al. Updated unified phylogenetic classification system and revised nomenclature for Newcastle disease virus. Infect Genet Evol. 2019;74: 103917.

26. Van Borm S, Obishakin E, Joannis T, Lambrecht B, vandenBerg T. Further evidence for the widespread co-circulation of lineages 4b and 7 velogenic Newcastle disease viruses in rural Nigeria. Avian Pathol. 2012;41(4): 377-382.

27. Pchelkina IP, Manin TB, Kolosov SN, Starov SK, Andriyasov AV, Chvala IA, et al. Characteristics of pigeon paramyxovirus serotype-1 isolates (PPMV-1) from the Russian Federation from 2001 to 2009. Avian Dis. 2013;57(1): 2-7.

28. Yurchenko KS, Sivay MV, Glushchenko AV, Alkhovsky SV, Shchetinin AM, Shchelkanov MY, et al. Complete Genome Sequence of a Newcastle Disease Virus Isolated from a Rock Dove (Columba livia) in the Russian Federation. Genome Announc. 2015;3(1): e01514-14.

29. Snoeck CJ, Adeyanju AT, Owoade AA, Couacy-Hymann E, Alkali BR, Ottosso U, et al. Genetic Diversity of Newcastle Disease Virus in Wild Birds and Pigeons in West Africa. 2013;79(24): 7867-7874.

30. Snoeck CJ, Owoade AA, Couacy-Hymann E, Alkali BR, Okwen MP, Adeyanju AT, et al. High genetic diversity of Newcastle disease virus in poultry in West and Central Africa: cocirculation of genotype XIV and newly defined genotypes XVII and XVIII. J Clin Microbiol. 2013;51(7): 2250-2260.

31. Ramey AM, Reeves AB, Ogawa H, Ip HS, Imai K, Bui VN, et al. Genetic diversity and mutation of avian paramyxovirus serotype 1 (Newcastle disease virus) in wild birds and evidence for intercontinental spread. Arch Virol. 2013;158(12): 2495-2503.

32. Mayahi V, Esmaelizad M. Molecular evolution and epidemiological links study of Newcastle disease virus isolates from 1995 to 2016 in Iran. Arch Virol. 2017;162(12): 3727-3743.

33. Hicks JT, Dimitrov KM, Afonso CL, Ramey AM, Bahl J. Global phylodynamic analysis of avian paramyxovirus-1 provides evidence of inter-host transmission and intercontinental spatial diffusion. BMC Evol Biol. 2019;19(1): 108.

34. Aldous EW, Fuller CM, Ridgeon JH, Irvine RM, Alexander DJ, Brown IJ. The evolution of pigeon paramyxovirus type 1 (PPMV-1) in Great Britain: A molecular epidemiological study. Transbound Emerg Dis. 2014;61(2): 134-139.

35. Aldous EW, Fuller CM, Mynn JK, Alexander DJ. A molecular epidemiological investigation of isolates of the variant avian paramyxovirus type 1 virus (PPMV-1) responsible for the 1978 to present panzootic in pigeons. Avian Pathol 2004;33(2): 258-269.

36. Irvine RM, Aldous EW, Manvell RJ, Cox WJ, Ceeraz V, et al. Outbreak of Newcastle disease due to pigeon paramyxovirus type 1 in grey partridges (Perdix perdix) in Scotland in October 2006. Vet Rec. 2009;165(18): 531-535.

37. Krapez U, Steyer AF, Slavec B, Barlic-Maganja D, Dovc A, Racnik J, et al. Molecular characterization of avian paramyxovirus type 1 (Newcastle disease) viruses isolated from pigeons between 2000 and 2008 in Slovenia. Avian Dis. 2010;54(3): 1075-1080.

38. Teske L, Ryll M, Rautenschlein S. Epidemiological investigations on the role of clinically healthy racing pigeons as a reservoir for avian paramyxovirus-1 and avian influenza virus. Avian Pathol. 2013;42(6): 557-565.

39. Byarugaba DK, Mugimba KK, Omony JB, Okitwi M, Wanyana A, Otim MO, et al. High pathogenicity and low genetic evolution of avian paramyxovirus type I (Newcastle disease virus) isolated from live bird markets in Uganda. Virol J. 2014;11: 173.

40. Wang J, Liu H, Liu W, Zheng D, Zhao Y, Li Y, et al. Genomic characterizations of six pigeon paramyxovirus Type 1 viruses isolated from live bird markets in China during 2011 to 2013. PLoS One. 2015;10(4): e0124261.

41. Chong YL, Lam TT, Kim O, Lu H, Dunn P, Poss M. Successful establishment and global dispersal of genotype VI avian paramyxovirus serotype 1 after cross species transmission. Infect Genet Evol. 17: 260-268.

42. Yates VJ, Fry DE, Henderson BW. Isolation of Newcastle disease virus from a calf. J Am Vet Med Assoc. 1952;120(900): 149-150.

43. Alexander DJ, Wilson GW, Russell PH, Lister SA, Parsons G. Newcastle disease outbreaks in fowl in Great Britain during 1984. Vet Rec. 1985;117(17): 429-434.

44. Alexander DJ. Newcastle disease and other avian paramyxoviruses. Rev Sci Tech. 2000;19(2): 443-462.

45. Jørgensen PH, Handberg KJ, Ahrens P, Manvell RJ, Frost KM, Alexander DJ. Similarity of avian paramyxovirus serotype 1 isolates of low virulence for chickens obtained from contaminated poultry vaccines and from poultry flocks. The Vet Rec. 2000;146(23): 665-668.

46. Li X, Chai T, Wang Z, Song C, Cao H, Liu J, et al. Occurrence and transmission of Newcastle disease virus aerosol originating from infected chickens under experimental conditions. Vet Microbiol. 2009;136(3): 226-232.

47. Sharma B, Pokhriyal M, Rai GK, Saxena M, Ratta B, Chaurasia M, et al. Isolation of Newcastle disease virus from a non-avian host (sheep) and its implications. Arch Virol. 2012;157(8): 1565-1577.

48. Pedersen JC, Hines NL, Killian ML, Predgen AS, Schmitt BJ. Isolation of a virulent Newcastle disease virus from confiscated LaSota vaccine. Avian Dis. 2013;57(2): 316-320.

49. Zhao P, Sun L, Sun X, Li S, Zhang W, Pulscher LA, et al. Newcastle disease virus from domestic mink, China, 2014. Vet Microbiol. 2017;198: 104-107.

50. Bahl J, Pham TT, Hill NJ, Hussein ITM, Ma EJ, Easterday BC, et al. Ecosystem Interactions Underlie the Spread of Avian Influenza A

Viruses with Pandemic Potential. PLoS Pathog. 2016;12(5): e1005620.

51. Alexander DJ, Parsons G, Marshall R. Infection of fowls with Newcastle disease virus by food contaminated with pigeon faeces. Vet Rec. 1984;115(23): 601-602.

52. Abolnik C, Gerdes GH, Kitching J, Swanepoel S, Romito M, Bisschop SPR. Characterization of pigeon paramyxoviruses (Newcastle disease virus) isolated in South Africa from 2001 to 2006. Onderstepoort J Vet Res. 2008;75(2): 147-152.

53. Alexander DJ. Newcastle disease in the European Union 2000 to 2009. Avian Pathol. 2011;40(6): 547-558.

54. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980;16(2): 111-120.