Gene 579 (2016) 34–40

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

# Genetic characterization and evolutionary analysis of Newcastle disease virus isolated from domestic duck in South Korea

Satish Gaikwad, Ji-Ye Kim, Hyun-Jeong Lee, Suk Chan Jung, Kang-Seuk Choi *

*OIE reference laboratory for Newcastle disease, Avian Diseases Division, Animal and Plant Quarantine Agency, Anyang, Gyeonggi 430-757, Republic of Korea*

ARTICLE INFO

*Article history:*
Received 26 June 2015
Received in revised form 17 December 2015
Accepted 18 December 2015
Available online 22 December 2015

*Keywords:*
Newcastle disease virus
Phylogeography
Avirulent strain
Domestic duck origin

ABSTRACT

Domestic ducks are considered a potential reservoir of Newcastle disease virus. In the study, a Newcastle disease virus (NDV) isolated from a domestic duck during surveillance in South Korea was characterized. The complete genome of the NDV isolate was sequenced, and the phylogenetic relationship to reference strains was studied. Phylogenetic analysis revealed that the strain clustered in genotype I of Class II ND viruses, has highly phylogenetic similarity to NDV strains isolated from waterfowl in China, but was distant from the viruses isolated in chickens and vaccine strains used in South Korea. Pathogenicity experiment in chickens revealed it to be a lentogenic virus. The deduced amino acid sequence of the cleavage site of the fusion (F) protein confirmed that the isolate contained the avirulent motif $^{112}$GKQGRL$^{117}$ at the cleavage site and caused no apparent disease in chickens and ducks. With phylogeographic analysis based on fusion gene, we estimate the origin of an ancestral virus of the isolate and its sister strain located in China around 1998. It highlights the need of continuous surveillance to enhance current understanding of the molecular epidemiology and evolution of the pathogenic strains.



## 1. Introduction

Newcastle disease virus (NDV) is avian paramyxovirus type 1 of the genus Avulavirus in the family Paramyxoviridae. NDV has a broad host range being able to infect over 240 species of birds (Umali et al., 2014). Of them, gallinaceous birds such as chickens are highly susceptible to NDV (Cappelle et al., 2014). NDV has a single-stranded, negative-sense RNA genome of 15,186 to 15,198 nt in length (Gogoi et al., 2015). The virus genome encodes structural proteins including nucleoprotein (N), phosphoprotein (P), matrix (M), fusion (F), hemagglutinin–neuraminidase (HN) and RNA-dependent RNA polymerase (L). The replication of the RNA genome is controlled by the rule of six, which requires the genome length to be multiple of six for the proper packaging of the RNA genome (El Najjar et al., 2014).

There are variations in virulence for NDV isolates. NDV isolates are pathotyped into at least three groups based on their clinical presentations in chickens. The least virulent are termed lentogenic, moderately virulent strains are mesogenic and most virulent are velogenic. Currently, intracerebral pathogenicity index (ICPI) score is the internationally recognized score for classifying the virulence of NDV strains, which can be supported by determining the cleavage site of fusion protein. The cleavage site of F protein is main determinant of NDV virulence, as strains with an F protein cleavage site with at least 3 arginine or lysine residues between positions 113 and 116 and a phenylalanine residue at the position 117 are considered virulent. Newcastle disease, defined as an infection of virulent NDV can cause significant economic losses in poultry industry (Miller et al., 2015).

NDV isolates are grouped into two distinct genetic classes (class I and II) based on genome length and nucleotide sequence diversity (Courtney et al., 2013). According to recent new genotype classification criteria proposed by Diel et al., Class I and II viruses consist of single serotype and at least fifteen genotypes (I to XV), respectively, on the basis of the full sequence of the F protein gene. Class I NDVs with a genome size of 15,198 nt long are frequently isolated from waterfowl and live-bird markets (LBMs), and most of them are avirulent (Cai et al., 2011). Class II viruses are typically found in wild birds and poultry species. Most virulent NDVs belong to class II virus. The genotypes that are considered “early” (I to IV and IX) contain 15,186 nucleotides while the other genotypes that emerged “late” contain 15,192 nucleotides (Diel et al., 2012).

NDVs have frequently been isolated from many bird species, both wild and domestic. In particular, ducks may play a significant role in virus transmission and maintenance in nature as a reservoir. In this study, we present the molecular characterization and phylogenetic

*Abbreviations:* NDV, Newcastle disease virus; ICPI, intracerebral pathogenicity index; LBMs, live-bird markets; DK13, KR/duck/13/07; SPF, specific pathogen free; $EID_{50}$, 50% egg infective dose; MCMC, Markov chain Monte Carlo; ESS, Effective sample size; CDS, coding sequence; TMRCA, time to most recent common ancestor; HPD, Highest posterior density.

* Corresponding author at: Avian Diseases Division, Animal and Plant Quarantine Agency, 175 Anyang-ro, Manan-gu, Anyang, Gyeonggi 430-757, Republic of Korea.
*E-mail address:* kchoi0608@korea.kr (K.-S. Choi).

http://dx.doi.org/10.1016/j.gene.2015.12.040
0378-1119/

analysis of an NDV isolate from domestic duck and investigated its relationship with NDVs isolated during the same period.

## 2. Materials and methods

### 2.1. Virus

An NDV isolate KR/duck/13/07 (DK13), a wild type virus was used in the study. The virus was isolated from cloacal/fecal swabs collected from apparently healthy domestic ducks during the winter season of 2006/2007 at the farm located in Gyeongbuk province, South Korea. The virus was propagated using egg inoculation in specific pathogen free (SPF) chicken embryonated eggs and aliquots were stored at −70 °C until use. The infective titer was expressed as 50% egg infective dose ($EID_{50}$). Virus virulence was determined by ICPI test using day-old SPF chicks following standard protocol (Kim et al., 2012).

### 2.2. Genome sequencing

Viral genomic RNA was extracted from infected allantoic fluid using TRIzol reagent (Invitrogen, USA) as per manufacture's protocol. cDNA synthesis was performed using Thermoscript first strand synthesis kit (Invitrogen, USA).

The complete genome sequence of KR/duck/13/07 was sequenced using reference primers (primers available on request) to cover full-length genome of the isolate. The PCR reaction was carried out in a Mastercycler (Eppendorf, Germany) using DreamTaq polymerase (Thermo Scientific, USA). The amplified products were purified using a QIAquick gel extraction kit as per the manufacturer's instructions. The overlapping PCR products were cloned in the TOPOTA vector (Invitrogen, USA). Three clones for each overlapping fragment were sequenced in both template directions and consensus sequence was determined. The leader and trailer sequences were generated by rapid amplification of cDNA ends (RACE), as previously described (Samuel et al., 2013). The sequence editing and assembly were done using VectorNTI and compared with published strains by BLAST. Reference strains were picked based on BLAST search results as well as from representatives for different lineages. Sequences were aligned using multiple sequence alignment algorithm implemented in muscle software (Edgar, 2004). Analysis of best-fit nucleotide substitution model and phylogenetic analyses were conducted in MEGA software version 6.0 (Tamura et al., 2013).

### 2.3. Genome annotation

Domains of individual proteins and features of the genome are annotated as per protein blast and Uniprot analysis (Altschul et al., 1997; O'Donovan and Apweiler, 2011). Sliding window analysis on similarity of complete genome sequences was performed in software Simplot. RNA structure prediction of potential RNA editing site in phosphoprotein was done by RNAfold algorithm available at RNAfold server (Hofacker, 2003).

### 2.4. Time scaled phylogeny and phylogeographic reconstruction

Complete genome sequences of previously characterized Newcastle disease virus strains were compared to assign a phylogenetic position to the sequenced isolate. The evolutionary history was inferred using the Maximum likelihood method with statistical analysis based on 100 bootstrapping replicates. The divergence of nucleotide sequences was calculated using MEGA6 (Tamura et al., 2013).

Bayesian time scaled phylogenetic analysis and Bayesian phylogeographic analysis were co-estimated using fusion gene sequence datasets. The fusion gene of NDV is most commonly sequenced gene during surveillance. The fusion sequence data incorporated 72 isolates throughout the globe. These isolates were annotated according to their location. Molecular clock model (uncorrelated lognormal distribution, uncorrelated exponential distribution and random) were tested with different demographic models (Nonparametric Bayesian Skyline plot, the parametric constant exponential growth, Speciation: Birth death) and the best model was selected by means of Bayes factor test using marginal likelihood values in Tracer v 1.5. A Bayesian Maximum clade credibility tree was constructed by using Bayesian Markov chain Monte Carlo (MCMC) analysis. An input file for BEAST analysis was prepared using Bayesian evolutionary analysis utility software v.1.8.1, in which sequences were annotated with year of collection. Three independent chains were run to get output of 10,000 trees with effective sample sizes (ESS) more than 200. Three individual runs were combined using LogCombiner V1.8.1 in BEAST software package. The posterior trees were summarized using TreeAnnotator after 10% burnin of samples. Phylogenetic MCC tree was visualized in FigTree software.

## 3. Results and discussion

### 3.1. The virulence of duck NDV

The ICPI value of isolate DK13 measured in day-old chicks as per the OIE protocol was 0.2. As per classification, the virus was assigned to lentogenic pathotype (Fan et al., 2015; Heiden et al., 2014). NDV strain DK13 contains the F protein cleavage site motif sequence $^{112}$E-R-Q-E-R-L$^{117}$, which is the major determinant of virulence for NDV strains (Mehrabanpour et al., 2014; Sun et al., 2015). This low proportion of basic amino acids and the presence of a leucine at position 117 (L117) are characteristic of avirulent strains, as previously described for other NDV strains (Abolnik and Wandrag, 2014; Kim et al., 2014; Wen et al., 2013; Wu et al., 2010; Zhu et al., 2014).

### 3.2. Genome analysis

The complete genome sequence of strain DK13 was submitted to GenBank with accession No.KT186351. The full genome length of NDV strain DK13 was 15,186 nucleotides in length, which is characteristic of early genotypes of NDV (Czeglédi et al., 2006). The 3′ leader and 5′ trailer sequences of NDV strain DK13 were 55 and 114 nucleotides respectively. These termini were highly conserved. The genome encodes genes in the order of 3′ leader-N-P-M-F-H-N-L-5′ trailer. Features of viral genes are summarized in Table 1. The pairwise comparison of the

**Table 1**
Genomic features of Newcastle disease virus strain DK13.

| Gene | Gene start sequence | Start position | Gene end sequence | Intergenic sequences |
|---|---|---|---|---|
| NP | ACGGGTAGAA | 56 | TTAGAAAAAA | GT |
| P | ACGGGTAGAA | 1804 | TTAAGAAAAAA | T |
| M | ACGGGTAGAA | 3256 | TTAGAAAAAA | C |
| F | ACGGGTAGAA | 4498 | TTAAGAAAAAA | CTGCCGGATGTAGATGACCAAAGGATAATAT |
| HN | ACGGGTAGAA | 6321 | TTAAGAAAAAA | TGTGAGCGGTAGTGGGACACAAGGCAAAACAACTCCTAGAGAATATT |
| L | ACGGGTAGGA | 8370 | TTAGAAAAAA | - |

NP: nucleoprotein, P: Phosphoprotein, M: matrix protein, F: fusion protein, HN: Haemagglutination neuraminidase protein, L: Large RNA polymerase protein.