

Characterization and full genome sequencing of a velogenic Newcastle disease virus (NDV) strain Ck/IR/Beh/2011 belonging to subgenotype VII(L)

H. GOUDARZI¹, S. VAN BORM², M. BASHASHATI¹, F. SABOURI³, M. ABDOSHAH⁴,
A. NOURI¹, M. BANANI¹, M. M. EBRAHIMI⁵, A. MOLOUKI^{1*}

¹Department of Avian Disease Research and Diagnostics, Razi Vaccine and Serum Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Alborz, Iran; ²Molecular Platform, Veterinary and Agrochemical Research Centre, Ukkel, Belgium; ³Department of Poultry Diseases, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran; ⁴Department of Quality Control, Razi Vaccine and Serum Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Alborz, Iran; ⁵Department of Avian Vaccines Production, Razi Vaccine and Serum Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Alborz, Iran

Received June 18, 2018; accepted August 28, 2018

Summary. – Avian avulavirus 1, better known as Newcastle disease virus (NDV), causes substantial loss to the poultry industry in many developing countries. In this study we have characterized and fully sequenced the genome of a velogenic NDV strain named Beh (Ck/IR/Beh/2011) that has been used in our lab for a number of challenge and immunological studies over the last few years. This strain was isolated from poultry in the city of Behshahr, Mazandaran Province, Iran after an outbreak reported in the region in 2011. The intracerebral pathogenicity index (ICPI) was 1.8 in one-day-old chicks, characteristic of a velogenic NDV strain. Later, the virus was purified using a sucrose gradient centrifugation and used for next-generation sequencing (NGS). The results showed that the genome length was 15192 bp, similar to those of class II velogenic strains. In addition, the phylogenetic analysis based on the complete F gene showed that the NDV strain Beh has an F protein cleavage site ¹¹²RRQKR↓F¹¹⁷ and belongs to the newly identified subgenotype VII(L). Based on the biological and genetic characterization, NDV strain Beh is now the best documented reference isolate representing the novel subgenotype VII(L) in Iran.

Keywords: NDV; NGS; velogenic strain, subgenotype VII(L); phylogenetic analysis

Introduction

Newcastle disease virus (NDV), or avian avulavirus 1, is a paramyxovirus for which three genome size categories of 15186, 15192 and 15198 bp have been documented (Czegledi *et al.*, 2006; Molouki and Peeters, 2016). The genome encodes six structural proteins and in analogy to other paramyxoviruses the transcription and replication of the virus is controlled by its ribonucleoprotein complex (RNP).

The two glycoproteins of NDV have been strongly associated with its virulence, although it has been shown that NDV's virulence may be a multigenetic factor (Dortmans *et al.*, 2011). The cleavage site on the fusion (F) protein has been used as a major indicator for differentiating lentogenic strains from mesogenic and velogenic strains. Moreover, the F gene itself has been used in phylogenetic analysis to study the relationship between different strains. Currently, a unified classification system based on the complete F gene is being used (Diel *et al.*, 2012).

After many decades of control measures, including the use of vaccines against NDV, the virus is still causing significant problems to poultry farmers worldwide as sometimes the vaccines do not provide complete protection against the virulent strains (Dimitrov *et al.*, 2016). Iran is among countries most

*Corresponding author. E-mail: a.molouki@rvsri.ac.ir; phone: +98-263457 00 38 (ext 2434).

Abbreviations: NDV = Newcastle disease virus; NGS = next-generation sequencing; IVA = iterative virus assembler; ICPI = intracerebral pathogenicity index

affected as it has been dealing with endemic NDV for a few decades and unfortunately the problem is still ongoing (Bozorgmehri-Fard and Keyvanfar, 1979; Dimitrov *et al.*, 2016).

The NDV strain Beh (Ck/IR/Beh/2011) was isolated from a Newcastle disease (ND) affected poultry farm in Northern Iran in 2011. Over the years the virus has been used in numerous challenge and immunological studies in our laboratory. Therefore, in this article we report its full biological and genetic characterization including the method we used to sequence its entire genome. New technical developments such as next-generation sequencing (NGS) have helped researchers understand the biology and genetic variation of the different strains much more efficiently and effectively. The number of complete NDV genome sequences submitted to GenBank has exponentially increased in recent years thanks to the new innovations in rapid sequencing and analysis of NGS data (Dimitrov *et al.*, 2017). In this article, we show how the NDV strain Beh was isolated, purified, sequenced and characterized. Moreover, due to the discovery of the novel subgenotype VII(L) in this region (Sabouri *et al.*, 2018), it was important to determine the origin of strain Beh and therefore, a phylogenetic analysis was included to study its relationship to other Iranian isolates.

Materials and Methods

Sample collection and preparation. The virus was isolated from commercial broiler poultry in the city of Behshahr, Mazandaran Province, Iran after an outbreak with a mortality rate of 95%, reported in a farm in 2011. The supernatant of homogenized tissue (trachea and spleen) was inoculated into 9-day-old embryonated SPF eggs for 2–3 days. After harvest, the allantoic fluid was subjected to HA and HI test according to OIE guidelines (OIE, 2012).

Intracerebral pathogenicity index (ICPI). The ICPI was performed based on OIE manual of diagnostic tests and vaccines for terrestrial animals (OIE, 2012). Briefly, fresh infective allantoic fluid with HA titer $>2^4$ was diluted 1:10 in sterile normal saline. Ten SPF one-day-old chickens were inoculated with 50 μ l of the diluted virus and examined every day for 8 days. The chickens were scored: 0 if normal, 1 if sick, and 2 if dead. The ICPI was the mean score per bird per monitoring over the 8-day period.

Virus purification. The allantoic fluid was subjected to sucrose gradient purification as previously described (Molouki, 2011). Briefly, the allantoic fluid was clarified using a floor centrifuge (Beckman, USA) and JA14 fixed angle rotor at 6,000 x g for 20 min at 4°C. The fluid was then further centrifuged using JA20 fixed angle rotor at 40,000 x g for 2 h at 4°C and the pellet was resuspended in 1 ml ice-cold PBS. Next, a gradient of different sucrose concentrations (60%, 50%, 40%, 30% and 20%) was made in an ultra-centrifuge tube and incubated for at least 1 h at 4°C to stabilize. After incubation, 1 ml of the resuspended virus was added to the top of the gradient and centrifuged at 161,000 x g for 4 h at

4°C in an ultra-centrifuge (Sorvall, USA). The virus band, found in the lower middle of the gradient, was carefully pipetted into a fresh ultra-centrifuge tube and overlaid with ice-cold PBS. The virus was centrifuged at 161,000 x g for 4 h at 4°C and the pellet was resuspended in 0.1–1 ml of ice-cold PBS. The HA was determined, and the purified virus was aliquoted and stored at -80°C.

SDS-PAGE and western blot. SDS-PAGE and western blot analyses were performed as previously described (Molouki *et al.*, 2010). Briefly, a 15% SDS-PAGE gel was made, and the purified virus sample was loaded. The proteins were then electro transferred onto an Immobilon-P membrane (Millipore Corp., USA) and the membrane was blocked with a casein buffer (Pierce, USA). Membrane was incubated with a chicken serum containing anti-NDV strain Beh polyclonal antibody which was collected from an infected chicken. The blot was then incubated with a secondary antibody (AP-conjugated goat anti-chicken, Abcam, USA) for 1 h at room temperature and developed with an alkaline phosphatase buffer containing NBT/BCIP (Thermo, USA).

Next-generation sequencing. The purified virus was subjected to RNA extraction using a High pure viral RNA kit (Roche, Germany). Random hexamers were used to synthesize double-stranded cDNA and library preparation was performed using Nextera XT kit (Illumina, USA) following the manufacturer's instructions. After quality control using Bioanalyzer (Agilent, USA) the library was sequenced using a HiSeq v4 kit (set to a read length of 100 bp) in a multiplex run using standard Nextera XT indices on a HiSeq sequencing platform (Illumina). Raw data (10 million paired end reads) were verified using FastQC v0.11.3 and trimmed using Trim Galore! (q = 30, l = 50, paired; https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). A taxonomic classification of the trimmed data using the KAIJU web server (<http://kaiju.binf.ku.dk>) (Menzel *et al.*, 2016) 2016 showed that the majority (>70 %) of reads were NDV specific. A random paired subset (2 x 50,000 reads) was used for *de novo* assembly using IVA v1.0.0 (Hunt *et al.*, 2015) resulting in four NDV specific contigs that were combined using GenBank Acc. No. JF340367 as a reference for scaffolding. The complete genome sequence of NDV strain Beh was submitted to GenBank (Acc. No. MF417546) and the ORF prediction and assignment of proteins was completed relative to other strains with length 15192 bp, e.g. JF340367.

Phylogenetic analysis. Phylogenetic analysis was performed and interpreted according to Diel (Diel *et al.*, 2012). Briefly, the sequence of the entire F gene from the NGS results was used in MEGA6 (Tamura *et al.*, 2013) for further analysis. The phylogenetic tree was constructed using Maximum likelihood method with the optimum nucleotide method (GTR+G+I) and 500 bootstrap replicates. Furthermore, estimation of mean inter-population evolutionary distances was computed using Maximum Composite Likelihood model with the gamma distribution model. Codon positions included were the 1st, 2nd, 3rd, and non-coding. All positions containing gaps and missing data were eliminated. Interpopulation distances exceeding 10% and 3–10% were used to assign genotypes and sub-genotypes, respectively (Diel *et al.*, 2012).

Results

The isolate Beh is a highly virulent NDV strain

The intracerebral pathogenicity index (ICPI) was performed according to OIE guidelines (OIE, 2012) and it was found to be 1.81 in one-day-old SPF chicks, characteristic of velogenic strains. The allantoic fluid extracted from SPF eggs was also subjected to purification using a sucrose gradient procedure. The ultra-purified band was run on SDS-PAGE to visualize the protein pattern of the paramyxovirus (Fig. 1). Western blot was also performed using blood serum collected from the infected chicken and the protein bands on the membrane were visualized using an AP-conjugated goat anti-chicken antibody (Fig. 1). The preliminary NGS analysis using KAIJU showed that the cDNA prepared from the ultra-purified virus mainly belonged to an avian avulavirus 1 and had very low percentage of other contaminants including chicken cDNA (Fig. 2). The sequencing results showed that the F gene contained the cleavage site ¹¹²RRQKR↓F¹¹⁷, characteristic of most virulent strains (Diel *et al.*, 2012).

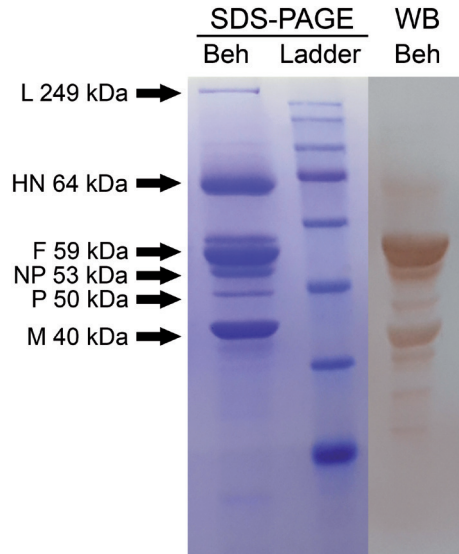


Fig. 1

SDS-PAGE and Western blot analysis

The 6 protein bands on SDS-PAGE are characteristic of structural proteins of paramyxoviruses. Western blot results shows that the serum antibody reacts with the ultra-purified virus.

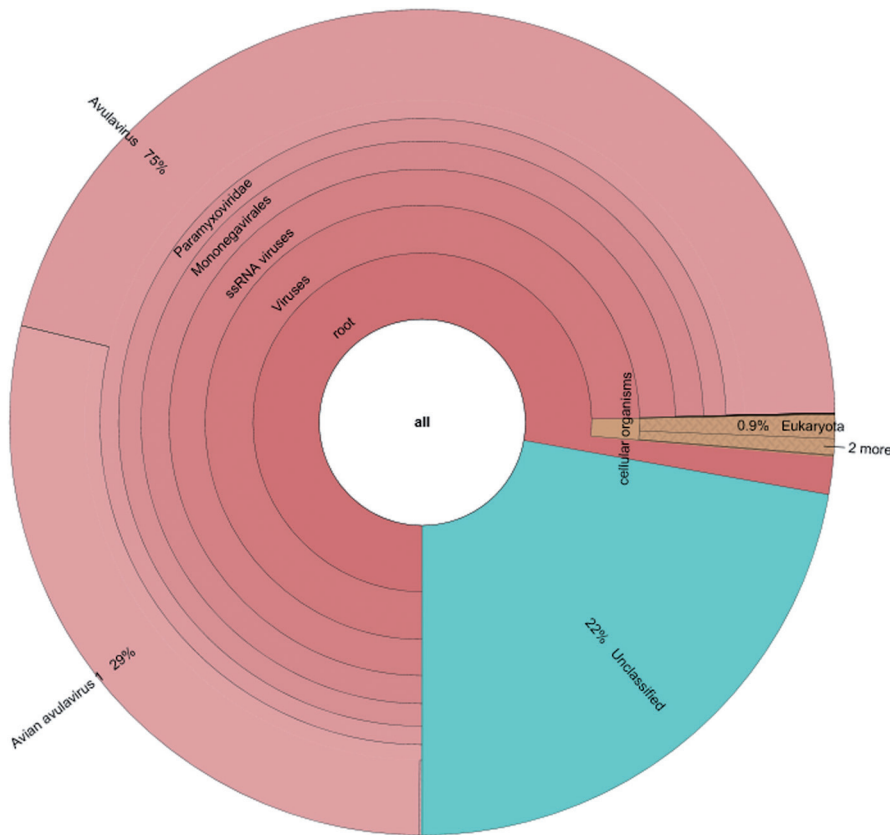


Fig. 2

Preliminary NGS analysis using KAIJU

NGS showed that the ultra-purified virus mainly belongs to an avian avulavirus 1 and has very low percentage of other contaminants.

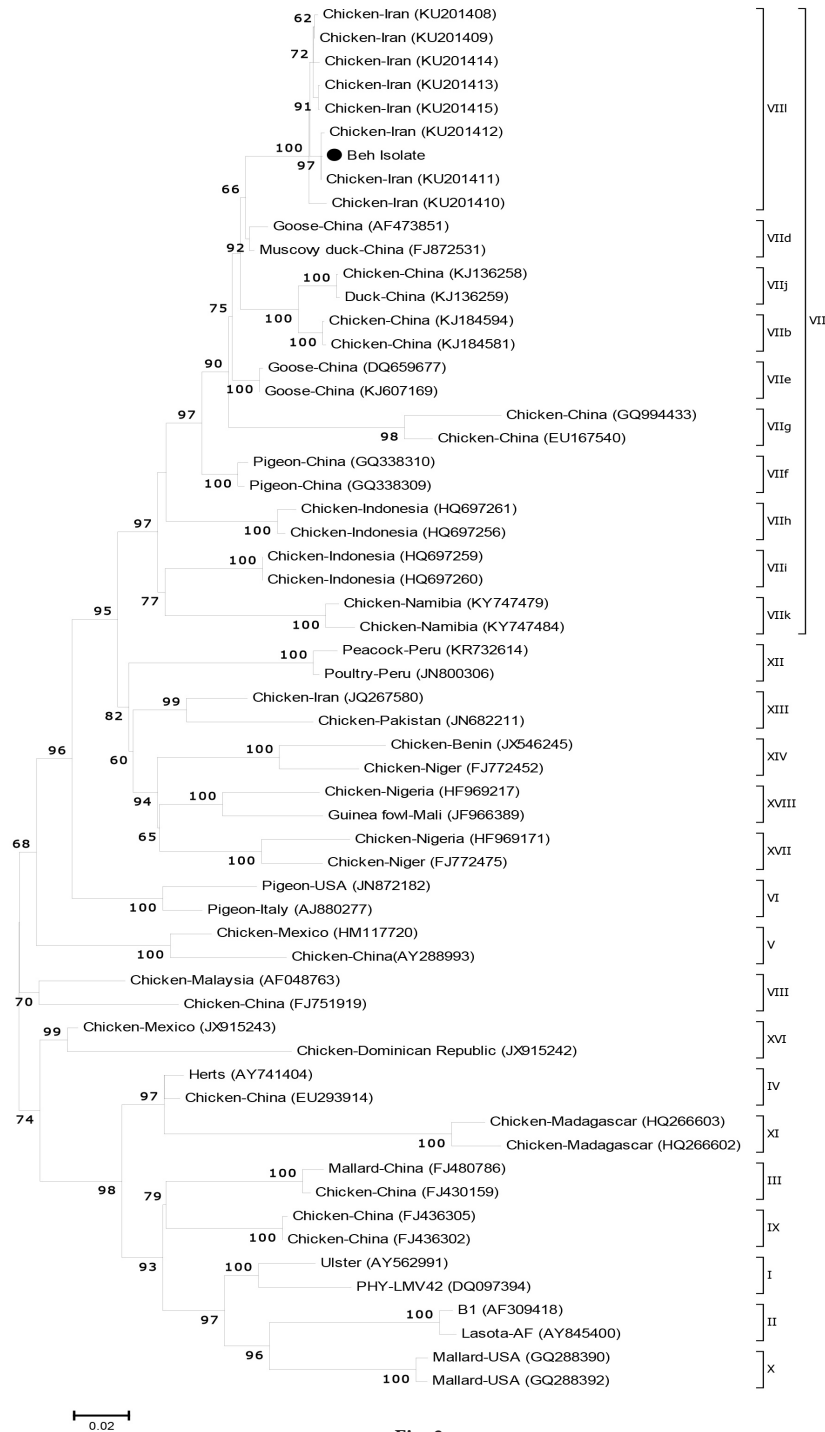


Fig. 3

Molecular phylogenetic analysis by maximum Likelihood method

NDV strain Beh clusters with other newly identified subgenotype VII(L) NDV. Most of these strains are unique to Iran. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model (Nei and Kumar, 2000). The tree with the highest log likelihood (-15592.8292) 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 by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.6639)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 0.0000% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 59 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1659 positions in the final dataset.

Complete genome characterization

The complete genome sequence of isolate Beh was determined to be 15192 bp, confirming the genome organization and the rule of six described for paramyxoviruses (Kolakovsky, 2005, 2016). Moreover, phylogenetic analysis based on the complete F gene of NDV showed that the strain Beh falls under the newly identified subgenotype VII(L) cluster (Fig. 3).

Discussion

Previously, it was believed that most Iranian strains belong to subgenotype VIId. Converse to popular belief, in a recent study by Sabouri and colleagues (Sabouri *et al.*, 2018) it was shown that some Iranian NDV isolates from 2011 to 2013 form a new cluster designated NDV subgenotype VII(L). When compared, strain Beh also clustered to the same group. It looks like this subgenotype contains only, or mostly, Iranian virulent NDV strains that have ICPI of approximately 1.8 and genome length of 15192 bp.

According to the criteria suggested by Diel and colleagues (Diel *et al.*, 2012), the difference 3–10% in the mean inter-population evolutionary distances were proposed as a cut-off value for assigning a novel subgenotype. Since distance value of this new cluster was in the mentioned range, it could be concluded that subgenotype VII(L) is a completely separate group.

Other genotypes, including subgenotype XIIIa and XIIIb have also been reported in Iran (Miller *et al.*, 2015). Therefore, more study is required to determine the exact dominant genotypes and subgenotypes currently circulating in Iran before a final conclusion can be made. Perhaps, the new samples and their use in vaccine challenge studies can help us determine why traditional NDV vaccines are not entirely able to protect birds and poultry from the new deadly outbreaks.

The number of NDV complete genomes submitted to GenBank has increased rapidly in recent years thanks to the fast sequencing capabilities of NGS systems available in the market. Previously, using conventional Sanger sequencing, one should have sequentially amplified and sequenced several overlapping fragments spanning the entire NDV genome. In addition, rapid amplification of cDNA ends (RACE) used to be required to reveal the correct sequence of the 3' and 5' ends which usually demands primers designed for other strains, thus not revealing the exact sequence of the cDNA ends. NGS has sped up this time-consuming process as well as reducing the cost and number of reagents used.

We managed to inoculate, isolate, purify and perform the molecular identification within days and the NGS sequencing results were analyzed and the final complete genome

sequence was obtained in less than a week. To our knowledge, the genome of NDV strain Beh is the first complete genome of any NDV virus obtained and reported in Iran. This, however, is somehow disappointing for a country that has been dealing with severe endemic NDV for years if not decades. On the other hand, it is good news that NGS technology is able to help local scientists in rapid, cost-effective and very accurate analysis of new NDV strains. Nonetheless, based on the biological and genetic characterization shown in this study, NDV strain Beh is the best documented isolate representing the novel subgenotype VII(L) in Iran, which can be used as a reference isolate for future studies.

Acknowledgment. Hossein Goudarzi was awarded a research grant No. 12-18-18-010-96006-960279 by Razi Vaccine and Serum Research Institute.

References

- Bozorgmehri-Fard MH, Keyvanfar H (1979): Isolation of Newcastle disease virus from teals (*Anas crecca*) in Iran. *J. Wild. Dis.* 15, 335–337. <https://doi.org/10.7589/0090-3558-15.2.335>
- Czegledi A, Ujvari 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, 36–48. <https://doi.org/10.1016/j.virusres.2005.11.009>
- Diel DG, da Silva LH, Liu H, Wang Z, Miller PJ, Afonso CL (2012): Genetic diversity of avian paramyxovirus type 1: proposal for a unified nomenclature and classification system of Newcastle disease virus genotypes. *Infect. Genet. Evol.* 12, 1770–1779. <https://doi.org/10.1016/j.meegid.2012.07.012>
- Dimitrov KM, Afonso CL, Yu Q, Miller PJ (2016): Newcastle disease vaccines-A solved problem or a continuous challenge? *Vet. Microbiol.* 206, 126–136. <https://doi.org/10.1016/j.vetmic.2016.12.019>
- Dimitrov KM, Sharma P, Volkening JD, Goraichuk IV, Wajid A, Rehmani SF, Basharat A, Shittu I, Joannis TM, Miller PJ, Afonso CL (2017): A robust and cost-effective approach to sequence and analyze complete genomes of small RNA viruses. *Viol. J.* 14, 72. <https://doi.org/10.1186/s12985-017-0741-5>
- Dortmans J, Koch G, Rottier P, Peeters B (2011): Virulence of new-castle disease virus: what is known so far? *Vet. Res.* 42, 122. <https://doi.org/10.1186/1297-9716-42-122>
- Hunt M, Gall A, Ong SH, Brener J, Ferns B, Goulder P, Nastouli E, Keane JA, Kellam P, Otto TD (2015): IVA: accurate de novo assembly of RNA virus genomes. *Bioinformatics* 31, 2374–2376. <https://doi.org/10.1093/bioinformatics/btv120>
- Kolakovsky D (2016): Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: A review. *Virology* 498, 94–98. <https://doi.org/10.1016/j.virol.2016.08.018>
- Kolakovsky D, Roux L, Garcin D, Ruigrok RW (2005): Paramyxovirus mRNA editing, the „rule of six“ and error catastrophe:

- a hypothesis. *J. Gen. Virol.* 86, 1869–1877. <https://doi.org/10.1099/vir.0.80986-0>
- Menzel P, Ng KL, Krogh A (2016): Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7, 11257. <https://doi.org/10.1038/ncomms11257>
- Miller PJ, Haddas R, Simanov L, Lublin A, Rehmani SF, Wajid A, Bibi T, Khan TA, Yaqub T, Setiyaningsih S, Afonso CL (2015): Identification of new sub-genotypes of virulent Newcastle disease virus with potential panzootic features. *Infect. Genet. Evol.* 29, 216–229. <https://doi.org/10.1016/j.meegid.2014.10.032>
- Molouki A (2011): Newcastle disease virus infection promoted Bax-mediated apoptosis assisted by the BH3 domain of matrix protein. Universiti Putra Malaysia, Kuala Lumpur, Malaysia.
- Molouki A, Hsu YT, Jahanshiri F, Rosli R, Yusoff K (2010): Newcastle disease virus infection promotes Bax redistribution to mitochondria and cell death in HeLa cells. *Intervirology*. 53, 87–94. <https://doi.org/10.1159/000264198>
- Molouki A, Peeters B (2016): Rescue of recombinant Newcastle disease virus: current cloning strategies and RNA polymerase provision systems. *Arch. Virol.* 162, 1–12. <https://doi.org/10.1007/s00705-016-3065-7>
- Nei M, Kumar S (2000): *Molecular evolution and phylogenetics*. Oxford University Press, New York, Oxford.
- OIE (2012): *Newcastle disease: Manual of diagnostic tests and vaccines for terrestrial animals*, Vol. 1. World Organisation for Animal Health, Paris, France, pp. 555–573.
- Sabouri F, Vasfi Marandi M, Bashashati M (2018): Characterization of a novel VIII sub-genotype of Newcastle disease virus circulating in Iran. *Avian Pathol.* 47, 90–99. <https://doi.org/10.1080/03079457.2017.1376735>
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013): MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. <https://doi.org/10.1093/molbev/mst197>