The Journal of Veterinary Medical Science

NOTE

Virology

Identification of Newcastle disease virus P-gene editing using next-generation sequencing

Xi CHEN[1)#], Yanqing JIA[2)#], Shanhui REN[1)], Siqi CHEN[1)], Xiangwei WANG[1)], Xiaolong GAO[1)], Chongyang WANG[1)], Fathalrhman E.A. ADAM[3)], Xinglong WANG[1)*] and Zengqi YANG[1)*]

[1)]College of Veterinary Medicine, Northwest A&F University, Yangling, Shaanxi Province 712100, P.R. China
[2)]Department of Animal Engineering, Yangling Vocational & Technical College, Yangling, Shaanxi Province 712100, P.R. China
[3)]Department of Preventive Medicine and Public Health, Faculty of Veterinary Science, University of Nyala, P.O. Box: 155 Nyala, Sudan

ABSTRACT. Avian paramyxoviruses 1 has the ability to edit its P gene to generate three amino-coterminal proteins (P, V and W), but its kinetic change is unclear. In this study, next-generation sequencing (NGS) was used to analyze the P-gene editing of Newcastle disease virus (NDV). Transcriptome analysis of chicken embryonic tissues and bursa of fabricius showed the P-gene editing frequencies were 45.46–52.70%. To investigate the rules of P-gene editing along time, the ratio of PVW was determined by PCR based deep sequencing at multiple time points in cells infected with velogenic and lentogenic strain respectively. The results confirmed similar editing frequencies with transcriptome data and the PVW ratios were stable along time among different NDVs, but had a greater V-gene transcript on velogenic strain infection ($P<0.001$), which were different from previous reports. Also, it was shown that the number of inserted G residues in P-derived transcripts was not limited to +9G, and +10G transcripts were identified. These results confirmed the NDV P-gene editing frequencies and provided a novel point of view on NDV P-gene editing with NDV virulence.

KEY WORDS: next-generation sequencing, P-gene editing frequencies, transcriptome analysis, V-gene transcript

J. Vet. Med. Sci.
82(8): 1231–1235, 2020
doi: 10.1292/jvms.18-0707

Received: 20 December 2018
Accepted: 13 June 2020
Advanced Epub:
6 July 2020

Newcastle disease virus (NDV) belongs to the genus *Avulavirus,* family *Paramyxovirinae*, which is a causative agent of Newcastle disease (ND). It can cause significant financial losses with high mortality worldwide [4]. The virus has a negative-sense RNA virus genome of 15 kb encoding 6 structural proteins, namely NP, P, M, F, HN and L. Additionally, two non-structural proteins (V and W) are derived from P gene by inserting one or two non-template G residue(s) into the conserved site 5′-AAAAAGGG [10]. P protein is one of the subunits of viral RNA-dependent RNA polymerase (vRdRp), which involved in the replication and transcription of NDV [5]. V protein has multiple functions in viral replication, viral virulence and anti-viral host immune responses [6, 16], while the shortest protein, W, is still unknown in its function.

Previously, Sanger sequencing (direct sequencing) was widely used in RNA-editing of several viruses. P-gene editing frequencies of Mumps Virus (MuV), Nipah Virus (NiV), Measles Virus (MeV) and Sendai Virus (SeV) were reported as 37.13, 66.34–94.25, 42.87 and 31.70%, based on conventional methodologies including RT-PCR, cloning and Sanger sequencing [2, 3, 8, 14]. V and W derived from NDV P-gene editing were firstly detected and described in 1991 and 1993, respectively [17, 18]. And the editing ratio of PVW was 68.25, 29.66 and 2.12% according to the consequences of selected clones by Sanger sequencing [11, 15].

Since entering of gene era, NGS has become a powerful method to make up the deficiencies of the traditional Sanger sequencing. This technology can indiscriminately detect all the genetic information in the samples such as gene mutation, polymorphism of noncoding RNA and microRNA [9, 12]. Accordingly, it could also be applied in studying gene editing frequency of *paramyxoviruses*. To gain a comprehensive understanding of P-gene editing kinetics of NDV. NGS was used to investigate the

*Correspondence to: Yang, Z.: yzq8162@163.com, Wang, X.: wxlong@nwsuaf.edu.cn
#These authors contributed equally to this work.

(Supplementary material: refer to PMC https://www.ncbi.nlm.nih.gov/pmc/journals/2350/)

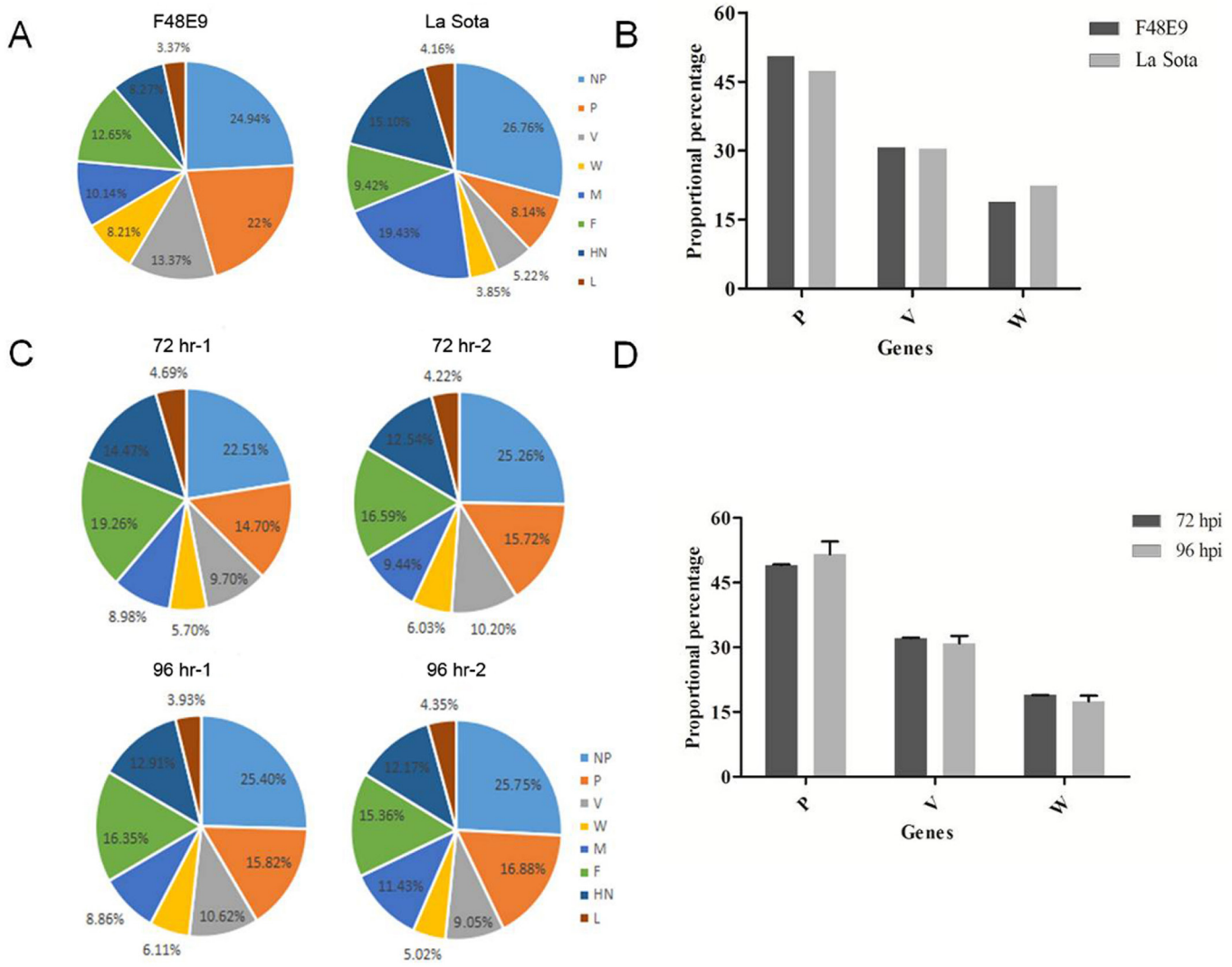

Fig. 1. Transcriptional analysis of viral genes from transcriptome sequencing. A. Transcriptional analysis of viral genes from chicken embryonic tissues at 36 hr post F48E9 or La Sota infection. B. The percentages of P, V and W transcripts in Newcastle disease virus infected embryonic tissues. C. Transcriptional analysis of viral genes from chicken bursa fabricius at 72 hr and 96 hr post F48E9 infection. D. The percentages of P, V and W transcripts in F48E9 infected chicken bursa fabricius.

editing frequency of NDVs infecting chicken embryos, chicken tissues, and the chicken fibroblast (DF-1) cell lines respectively.

F48E9 (velogenic strain) or La Sota (lentogenic strain) were independently injected into the allantoic cavity of 10-day-old chicken embryos with a titer of 10^4 PFU and visceral tissues were collected at 36 hr and then sent to Novogene (Wuhan, P. R. China) for transcriptome sequencing. The raw reads were acquired and underwent quality trimming, and, then, the Q20, Q30, and GC contents were calculated. Clean reads with Q Score >30 were selected and aligned with the corresponding paired reads. Finally, sequences with target lengths were used for the differential expression analysis. In total, 2.3×10^6 clean reads related to F48E9 and 1.3×10^4 clean reads associated with La Sota were obtained. Of which the percents of NP, P, M, F, HN, L, V and W deriving from F48E9 were 24.94%, 22.01, 10.14, 12.65, 8.27, 3.37, 13.37 and 8.21%. Correspondingly, the transcripts of genes from La Sota took up 26.76, 8.14, 19.43, 9.42, 15.11, 4.16, 5.22, and 3.85% (Fig. 1A). The proportions of PVW transcripts of F48E9 were 50.48, 30.68 and 18.84%. And the PVW of La Sota took up 47.30, 30.33 and 22.37% (Fig. 1B).

To prove the reliability of the results, the lymph tissue and bursa of fabricius were collected from the 4-week-old SPF chicken infected with the F48E9. And its transcriptome was examined at 72 and 96 hr after infection. A similar phenomenon was observed with the PVW percentage of 48.84–49.20, 31.92–32.23, 18.87–18.94% at 72 hr and 48.6–54.54, 29.24–30.63, 16.22–18.77% at 96 hr, which also confirmed that the proportions of PVW were not changed much over time (Fig. 1C). Besides, the transcripts of NP, P, M, F, HN, L, V and W were 23.89, 15.21, 9.21, 17.93, 13.51, 4.46, 9.95 and 11.73% at 72 hr post F48E9 infection in bursa of fabricius, and that were 25.58, 16.35, 10.15, 15.86, 12.54, 4.14, 9.84 and 5.57% at 96 hr, respectively (Fig. 1D). Both transcriptome analysis of tissue demonstrated the P-gene editing frequencies were 45.46–52.70%, which differed from the data previous (32.71%).

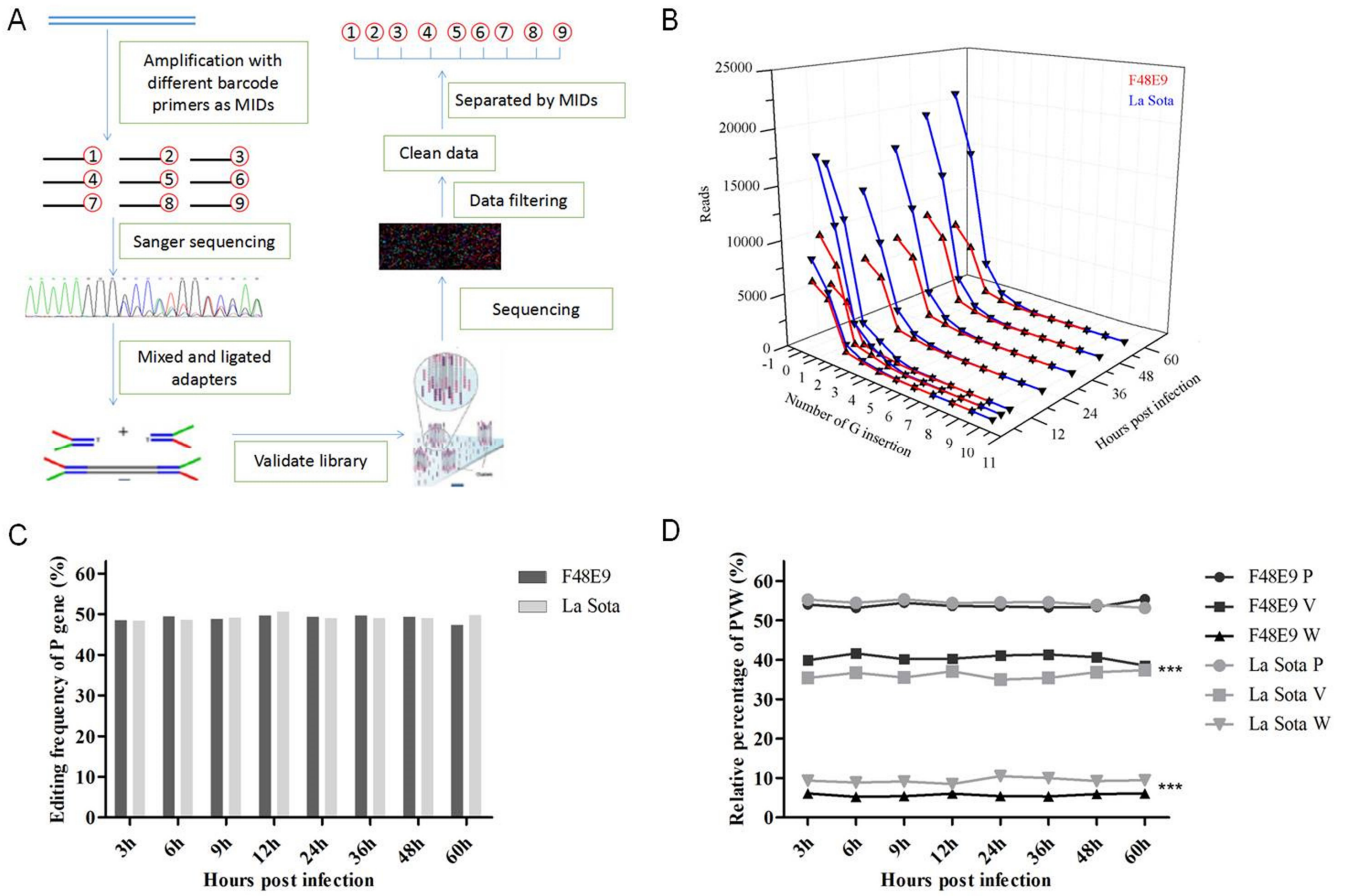

Fig. 2. The P-gene editing frequencies of different time points post NDV infection. A. Flow diagram of high-throughput sequence. B. The number of G insertion during Newcastle disease virus infection. Origin 8.0 was used to create 3D picture, red line represents F48E9 infection, and blue represent La Sota infection. C. The relative percentages P-gene editing frequencies of F48E9 and LaSota. D. The dynamic changes of P, V and W of F48E9 and LaSota. ***P<0.001 between groups.

PCR based deep sequencing was also used to analyze P-gene editing. Briefly, cDNA derived from DF-1 cells of 3, 6, 9, 12, 24, 48, 60 hr infected by NDVs were synthesized and used for amplification of the fragments covering P-gene editing site (AAAAAGGG) with primers fusing with different bar-code sequences (Supplementary Table 1). The PCR products were purified with DNA Gel Extraction Kit (TIANGEN, Beijing, P. R. China) and sent to the company for sequencing (Fig. 2A). With the similar protocol, a total of 4.63×10^5 reads of clean data were obtained for F48E9 and 1.23×10^5 clean reads for La Sota. Reads with overlap regions were merged to obtain the complete sequences using Pandaseq (version 2.7), and the merged sequences were identified, separated and classified into eight groups based on bar-code sequences. The PCR based deep sequencing results showed that the proportions of PVW transcripts of F48E9 were 53.14–55.37, 38.53–41.64 and 5.21–6.10%. And the PVW percents of La Sota were 53.17–55.34, 34.94–37.37 and 8.5–10.48% (Fig. 2B), the editing frequencies of which were also parallel with the transcriptome data and stable along time (Fig. 2C). In addition, F48E9 strain has a mean value of about 4.33% greater V-gene transcript than La Sota in every time point, but has a mean value of about 3.68% lower W-gene transcript than La Sota in every time point (P<0.001) (Fig. 2D).

Since the insertion of non-template G residues in editing site A_nG_n in the P-gene can generate different proteins, To know how many G can be inserted in and whether the number of G is changeable along the infection time is necessary. According to the NGS reads obtained in PCR based deep sequencing, a maximum of +10 G insertion was detected (Fig. 2B and Supplementary Tables 2 and 3). As the ORF did not shift when +0 G, +3 G, +6 G or +9 G were inserted in, these sequences were all defined as P Gene. Similar, V gene was with +1 G, +4 G, +7 G or +10 G insertion and W has +2 G, + 5G, or +8 G insertion. However, +0 G (51.52% for La Sota and 51.37% for F48E9), +1 G (34.25 and 39.34%) and +2 G (9.01 and 5.55%) were the majority. The rest insertion only took a small proportion (5.22 and 3.74%). This means that most of the transcripts have 0, 1 or 2 G insertion and only a few with 3 or more G. A time-related variation of G insertion was also monitored. To La Sota, +0–10 G insertion were found in each hr, except 6 hr (+0–9 G) and 60 hr (+0–8 G) (Fig. 2B and Supplementary Tables 2 and 3). Whereas, +0–9 G insertion happened at 60 hr and +9 G is the maximum insertion for F48E9. And most of time, the maximum insertion number of G were +6–8 G (Fig. 2B and Supplementary Tables 2 and 3). These findings were similar to previous studies, in which the main insertion numbers of G

were 1 or 2 for NDVs except that the maximum number of inserted G residues could reach as much as 10 [11, 15].

V and W derived from NDV P gene editing were firstly detected and described in 1991 and 1993, respectively [17, 18]. Since then, the related reports about NDV P-gene editing were limited. In 2001, Mebatsion *et al*. used a recombinant NDV virus to infect BSR-T7/5 cells for 24–36 hr. After RT-PCR with P-gene specific primers and constructing to vector, they selected a total of 41 independent colonies around the editing locus for Sanger sequencing, of which 28 (68.30%) colonies of the sequenced plasmids encoded P protein, and 12 (29.33%) encoded the V protein, only one plasmid (2.47%) encoded W protein [11]. Hence, they confirmed that the PVW ratio of NDV was 68.30, 29.34 and 2.46%. In 2016, Qiu *et al*. used three different toxic strains of NDV to infect DF-1 cells for 4, 6, 8, 10, 12, 24 hr respectively [15]. By the similar method, 1.24×10^3 clones were successfully sequenced and the P-gene editing frequencies was 34.13–38.70% at 4, 24 hr in accordance with Mebatsion *et al*. (32.71%), but differ from 34 to 60.4% along with time and different NDV strains [11, 15]. Besides, they also found the number of inserted G residues was not limited to one or two: +3G, +4G, +5G and even +9G transcripts were identified. In our study, we monitored all the information of samples with the indiscriminate technology NGS. Distinguishing from the previous report (32.71%), either the transcriptome analysis of tissues or PCR based deep sequencing of cells all showed the P-gene editing frequencies were 45.46–52.70%. Besides, we monitor the time-related variation of G insertion and the inserted G residues can reach as much as 10.

V protein has been confirmed to play an important role in anti-viral host immune responses, however, little knowledge about protein W has been reported. In our study velogenic NDV induced a higher level of protein V and lower level of protein W expression at each time than lentogenic strain. The result was similar with the previous report, in which, with traditional sequencing methods, Qiu *et al*. (2016) proved different NDV strains could also have regular PVW proportions along time [15]. As the similar regulation was founded in different reports, it seemed that the infection processing of different viruses was mediated by the concerted action of V and W proteins.

Interestingly, according to the transcriptome analysis of chicken embryonic tissues infected by F48E9 and La Sota strain respectively, ratios of M, F, or HN mRNAs were different between the two strains. These differences also could contribute to the difference in virulence. As we know that NDV infection is initiated by receptor recognition and binding to the host cell surface, which is followed by fusion mediated by the HN and F. These two glycoproteins on the NDV surface make major contributions to virulence, contagiousness, host range, and tissue tropism. Numerous studies have demonstrated that the M protein of NDV is a nucleocytoplasmic shuttling protein [1]. In addition to participating in the assembly and budding of progeny virions at the cell membrane later in infection [7, 13], the M protein is localized in the nucleus early in infection, which may inhibit host cell transcription [1].

In summary, NGS results indicated that P-gene editing frequencies were similar at different infection time or different toxic strains of NDV. But the PVW ratios differ from NDVs, as more V-gene transcript but less W-gene transcript in velogenic NDV strain than lentogenic NDV strain was observed. Those results suggested that V-gene transcriptional level may be related in viral virulent and their role in antagonism of host immune response, which provided a novel insight for further research on NDV virulence and growth property.

The GraphPad Prism 6.0 was used to analyze the experimental data (San Diego, CA, USA). The differences between F48E9 and La Sota infection were compared by *t*-test. Statistical significances were represented as $*P<0.05$, $**P<0.01$ and $***P<0.001$ between groups.

The experimental use of chicken embryos and chickens in this study were operated on strictly according to the Committee for the Ethics of Animal Care and Experiments in Northwest A&F University.

CONFLICT OF INTEREST. The authors declare that they have no conflict of interest.

ACKNOWLEGMENT. This study was funded by the National Natural Science Foundation of China to ZQY (No. 31572538) and XLW (No.31672581).

REFERENCES

1. Audsley, M. D., Jans, D. A. and Moseley, G. W. 2016. Roles of nuclear trafficking in infection by cytoplasmic negative-strand RNA viruses: paramyxoviruses and beyond. *J. Gen. Virol.* **97**: 2463–2481. [Medline] [CrossRef]
2. Bankamp, B., Lopareva, E. N., Kremer, J. R., Tian, Y., Clemens, M. S., Patel, R., Fowlkes, A. L., Kessler, J. R., Muller, C. P., Bellini, W. J. and Rota, P. A. 2008. Genetic variability and mRNA editing frequencies of the phosphoprotein genes of wild-type measles viruses. *Virus Res.* **135**: 298–306. [Medline] [CrossRef]
3. Curran, J., Boeck, R. and Kolakofsky, D. 1991. The Sendai virus P gene expresses both an essential protein and an inhibitor of RNA synthesis by shuffling modules via mRNA editing. *EMBO J.* **10**: 3079–3085. [Medline] [CrossRef]
4. Dortmans, J. C F M., Peeters, B. P. H. and Koch, G. 2012. Newcastle disease virus outbreaks: vaccine mismatch or inadequate application? *Vet. Microbiol.* **160**: 17–22. [Medline] [CrossRef]
5. Hausmann, S., Garcin, D., Delenda, C. and Kolakofsky, D. 1999. The versatility of paramyxovirus RNA polymerase stuttering. *J. Virol.* **73**: 5568–5576. [Medline] [CrossRef]
6. Huang, Z., Krishnamurthy, S., Panda, A. and Samal, S. K. 2003. Newcastle disease virus V protein is associated with viral pathogenesis and functions as an alpha interferon antagonist. *J. Virol.* **77**: 8676–8685. [Medline] [CrossRef]
7. Ke, Z., Strauss, J. D., Hampton, C. M., Brindley, M. A., Dillard, R. S., Leon, F., Lamb, K. M., Plemper, R. K. and Wright, E. R. 2018. Promotion of virus assembly and organization by the measles virus matrix protein. *Nat. Commun.* **9**: 1736. [Medline] [CrossRef]
8. Kulkarni, S., Volchkova, V., Basler, C. F., Palese, P., Volchkov, V. E. and Shaw, M. L. 2009. Nipah virus edits its P gene at high frequency to express

the V and W proteins. *J. Virol.* **83**: 3982–3987. [Medline] [CrossRef]
9. Mardis, E. R. 2008. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**: 133–141. [Medline] [CrossRef]
10. Mebatsion, T., de Vaan, L. T. C., de Haas, N., Römer-Oberdörfer, A. and Braber, M. 2003. Identification of a mutation in editing of defective Newcastle disease virus recombinants that modulates P-gene mRNA editing and restores virus replication and pathogenicity in chicken embryos. *J. Virol.* **77**: 9259–9265. [Medline] [CrossRef]
11. Mebatsion, T., Verstegen, S., De Vaan, L. T. C., Römer-Oberdörfer, A. and 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**: 420–428. [Medline] [CrossRef]
12. Ong, M., Carreira, S., Goodall, J., Mateo, J., Figueiredo, I., Rodrigues, D. N., Perkins, G., Seed, G., Yap, T. A., Attard, G. and de Bono, J. S. 2014. Validation and utilisation of high-coverage next-generation sequencing to deliver the pharmacological audit trail. *Br. J. Cancer* **111**: 828–836. [Medline] [CrossRef]
13. Pantua, H. D., Mcginnes, L. W., Peeples, M. E. and Morrison, T. G. 2007. Requirements for the assembly and release of newcastle disease virus-like particles. *J. Virol.* **81**: 1537–1537. [Medline] [CrossRef]
14. Paterson, R. G. and Lamb, R. A. 1990. RNA editing by G-nucleotide insertion in mumps virus P-gene mRNA transcripts. *J. Virol.* **64**: 4137–4145. [Medline] [CrossRef]
15. Qiu, X., Fu, Q., Meng, C., Yu, S., Zhan, Y., Dong, L., Ren, T., Sun, Y., Tan, L., Song, C., Han, X. and Ding, C. 2016. Kinetic analysis of RNA editing of Newcastle disease virus P gene in the early period of infection. *Acta Virol.* **60**: 71–77. [Medline] [CrossRef]
16. Rodriguez, K. R. and Horvath, C. M. 2014. Paramyxovirus V protein interaction with the antiviral sensor LGP2 disrupts MDA5 signaling enhancement but is not relevant to LGP2-mediated RLR signaling inhibition. *J. Virol.* **88**: 8180–8188. [Medline] [CrossRef]
17. Samson, A. C. R., Levesley, I. and Russell, P. H. 1991. The 36K polypeptide synthesized in Newcastle disease virus-infected cells possesses properties predicted for the hypothesized 'V' protein. *J. Gen. Virol.* **72**: 1709–1713. [Medline] [CrossRef]
18. Steward, M., Vipond, I. B., Millar, N. S. and Emmerson, P. T. 1993. RNA editing in Newcastle disease virus. *J. Gen. Virol.* **74**: 2539–2547. [Medline] [CrossRef]