ISSN 1682-8356
ansinet.org/ijps

INTERNATIONAL JOURNAL OF POULTRY SCIENCE

ANSInet
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan
Mob: +92 300 3008585, Fax: +92 41 8815544
E-mail: editorijps@gmail.com

International Journal of Poultry Science 8 (10): 985-994, 2009
ISSN 1682-8356
© Asian Network for Scientific Information, 2009

Sequence Analysis of HN and F Genes of a less Virulent Newcastle Disease Virus Isolated from Unvaccinated Village Chicken

P. Ponnusamy, J. John Kirubaharan and Andrew Chandramohan
Department of Veterinary Microbiology, Madras Veterinary College, Chennai-600 007, India

Abstract: The HN and F genes of the D58 isolate have coding senquence of 577 and 553 amino acids respectively. The HN gene has stalk (1-143) and globular head (123-571) regions with transmembrane domain at positions 25-45, sialic acid residue at 234-239, glycosylation sites at 119, 341, 433, 481 and 538 and cysteine residues at 172, 186, 196, 238, 247, 251, 344, 455, 461, 465, 531 and 542. The glycosylation site at 508 specific for virulent strains and cysteine residue at 123 were absent in D58 isolate. The F gene has transmembrane domains at positions 15-25 and 501-523 and glycosylation sites at 85, 191, 366, 447, 471 and 541. The fusion protein cleavage site has amino acids GRQGRL and fusion peptide starts with leucine. The HN and F genes have 26 and 23 B cell epitopes respectively. The D58 isolate was grouped with less virulent strains in phylogenetic analysis.

Key words: Newcastle disease, Avian paramyxovirus 1, HN gene, F gene, virulence, phylogenetic analysis, B cell epitopes

INTRODUCTION

Newcastle disease-ND (Doyle, 1927) is considered as the most important infection in the Asian countries leading to substantial production losses. Avian paramyxovirus 1 (APMV 1) with an Intracerebral Pathogenicity Index (ICPI) value of >0.7 has been reported to cause ND with respiratory distress and diarrhoea with increased morbidity and mortality (Alexander, 2003). The disease has been reported to be present in endaemic form with frequent outbreaks in chickens. Newcastle Disease Virus (NDV) isolates have been classified into 5 pathotypes namely viscerotropic velogenic, neurotropic velogenic, mesogenic, lentogenic and apathogenic. The lentogenic and apathogenic strains are commonly used as seed virus in live and inactivated vaccines. The Intra Cerebral Pathogenicity Index (ICPI), Intra Venous Pathogenicity Index (IVPI), Mean Death Time (MDT) and deduced amino acid sequences at the F Protein Cleavage Site (FPCS) form the basis for grouping of NDV isolates in to these pathotypes. The genome of NDV-single stranded negative sense RNA has been reported to be made up of approximately 15,186 nucleotides comprising of genes coding for Nucleoprotein (NP), Phosphoprotein (P), Matrix protein (M), Haemagglutinin Neuraminidase (HN), Fusion (F) and Large protein (L). Of the different proteins, Fusion (F) protein has been reported to mediate the fusion of viral envelope with cell membrane and considered as the major determinant of virulence (Peeters *et al*., 2001). Hence, NDV isolates were differentiated into virulent and less virulent strains by determining the amino acid sequence at the FPCS located at amino acid positions 112-119 of the F gene. A Pair of basic amino acids lysine (K) and/or arginine (R) at positions 116 and 115 with a phenylalanine (F) at position 117 and arginine at position 113 has been reported to determine the virulence for chickens (Aldous and Alexander, 2001). However, in the recent past it has been reported that apart from the fusion protein, the HN protein has also been reported to play crucial role in the process of infection (Huang *et al*., 2003) particularly the globular head region of the HN gene (Olav *et al*., 2005). The commonly carried out test to assess the pathotype of NDV namely ICPI involves injection of virus into the brain of day old chicken and hence not allowed on most occasions due to ethical reasons. Though the Terrestrial Manual of OIE (2004), has approved the determination of amino acid sequence of the FPCS as one of the methods for identifying the pathotype of NDV, the report of field NDV isolates with amino acid sequence at FPCS typical of low virulent type and ICPI values typical of highly virulent strains (Lei-Tao Tan *et al*., 2008) and the virulence of NDV attributed to more than one gene warrants confirmation of pathogenicity of NDV isolates by analyzing the entire sequence of coding sequences (cds) of HN and F genes and not with FPCS alone. Hence, this research was carried out. Further, the European Union in its commission decision 93/152/EEC restrains preparation of live or inactivated vaccine from NDV strain/isolate having ICPI value of >0.4. This also necessitates the need to generate sequence data of the coding sequence (cds) of HN gene and to compare it with the cds of F gene to determine the virulence of NDV strains/isolates particularly for those strains/isolates of NDV used in live or inactivated vaccine production. In the present work, analysis of coding sequences of HN and F genes of a NDV isolate-D58, that is being used as vaccine strain in south India

to prime the chickens is being discussed. This vaccine virus was reported to have ICPI value of 0.14 and amino acids ^{112}G-R-Q-G-R-L^{117} at the FPCS (Ananth *et al.*, 2008) typical for low virulent NDV. Sequence data on HN and F genes for this vaccine virus is not available.

MATERIALS AND METHODS

Virus: The thermostabilized low Intracerebral Pathogenicity Index (ICPI-0.14) isolate of NDV (D58) obtained from apparently normal village chicken was used in this study (Ananth *et al.*, 2008). This virus was maintained in the laboratory by propagating it in the Embryonated Chicken Eggs (ECE) as per the standard procedures mentioned in the Terrestrial Manual of OIE (2004). The virus was cultivated in the allantoic cavity of embryonated chicken eggs and the infected Amnioallantoic Fluid (AAF) collected after 96 hrs of infection was used for RNA extraction.

Viral RNA extraction and RT-PCR: The RNA was extracted from plaque purified D58 isolate of NDV using TRIZOL® LS reagent as per the manufacturer's instructions with minor modifications at RNA pellet washing stage. The RNA pellet was washed thrice with 75% ethanol by centrifuging at 7,500 g for 10 min instead of one time as suggested by the manufacturer and the RNA pellet was subsequently air-dried. Reverse transcription was carried out using Thermoscript RT at 50°C for 50 min using sense gene specific sense primers specific for HN and F genes. The cDNA of HN gene was amplified by XL PCR using sense (5'-GTA GGC TAG CAA GAG AGG CCG CCC CTC AAT-3') and anti sense primers (5'-CGA GCC CGG GCC GGC ATT CGG TTT GAT TCT TG-3') (Peeters *et al*, 2001). The XL PCR was carried out with an initial denaturation at 94°C for 2 min followed by first 10 cycles of 94°C for 30 sec, 58°C for 30 sec, 72°C for 90 sec, which was followed by second 20 cycles of 94°C for 30 sec, 58°C for 40 sec and 72°C for 90 sec with an elongation cycle of 5 sec for each cycle and a final extension of 72°C for 10 min. The cDNA of F gene was also amplified by XL PCR using sense primer (5'-GCG TCG ACA TGG GCY CYA RAY CTT CTA CCA-3') and anti sense primer (5'-TAG TCG ACT CAB ATT YTT GTA GTG GCY CTC ATC T-3') (Li Yu *et al.*, 2001). The XL PCR was carried out for F gene with an initial denaturation at 94°C for 2 min, follwed by first 10 cycles of 94°C for 15 sec, 52°C for 30 sec, 72°C for 60 sec, which was followed by second 20 cycles of 94°C for 15 sec, 52°C for 30 sec and 72°C for 60 sec with an elongation cycle of five seconds for each cycle and a final extension of 72°C for 10 min.

DNA sequencing and phylogenetic analysis: The HN and F gene amplicons were purified by using AuPrep PCR purification kit (Life Technologies, India-Cat.# PP28-104LT) following manufacturer's instruction. The sequencing of purified HN and F genes of ND vaccine viruses under study was carried out by big dye termination chemistry method in an automated sequencer (ABI Prism, version 3, Applied Biosystems, USA) through gene walking. The multiple sequence alignment of HN and F genes were carried out using the Bio-Edit software (North Carolina State University, USA) to generate sequence analysis data. The parameters used for sequence analysis were multiple alignment (Clustal W), sequence identify plotter and sequence identity matrix at both nucleotide and amino acids levels. The phylogenetic tree was developed using Neighbour Joining (NJ) algorithm using bootstrap values and distance in Mega 3.1 software (1993-2005). The B cell epitopes of HN and F genes were predicted in a computer algorithm using web server based software http://bioinfo.bug.ac.il/bsu/immunology/epitope_pred/index.htm following the algorithm developed by Bates and Sternberg (1999), Bates *et al.* (2001) and Conteras-Moreira and Bates (2002). The transmembrane domains were predicted using DAS trasmembrane domain prediction server - http://www.sbc.su.se/~miklos/DAS/maindas.html (Cserzo *et al.*, 1997).

RESULTS

RT-PCR: The HN and F genes of the D58 virus were amplified by RT-PCR. The product size for HN and F genes were approximately 1892 and 1662 bp in length respectively. The PCR products of HN and F were purified and sequenced. The complete coding sequence (cds) of HN and F gene sequences generated for D58 isolate of NDV was submitted to GenBank and the accession numbers EU305607 and EU330230 were obtained for HN and F genes, respectively.

HN gene sequences: The HN gene of D58 isolate was found to have a coding sequence (cds) comprising of 1731 nucleotides coding for 577 amino acids with two major areas namely stalk and globular head comprising 1-143 amino acids and 125-577 amino acids respectively. The HN gene was also found to have three transmembrane domains at positions 24-47, 25-45 and 557-563, a sialic acid binding site at position 234-239 and 12 Cysteine (C) residues at position 172, 186, 196, 238, 247, 251, 344, 455, 461, 465, 531 and 542. The cysteine residue reported at position 123 in certain strains like Queensland V4, D26, Ulster 2C, I_2, Mukteswar, Herts 33, Italien and ITA 45 has been replaced by tryptophan (W) in D58 isolate. The HN gene of D58 isolate was also found to have five glycosylation sites at positions 119, 341, 433, 481 and 538. Another glycosylation site reported at position 508, which was reported to be commonly present in virulent strains, is absent in D58 isolate. These details are provided in Table 1. The sequence identity matrix for HN gene between D58 and other strains of NDV are provided in

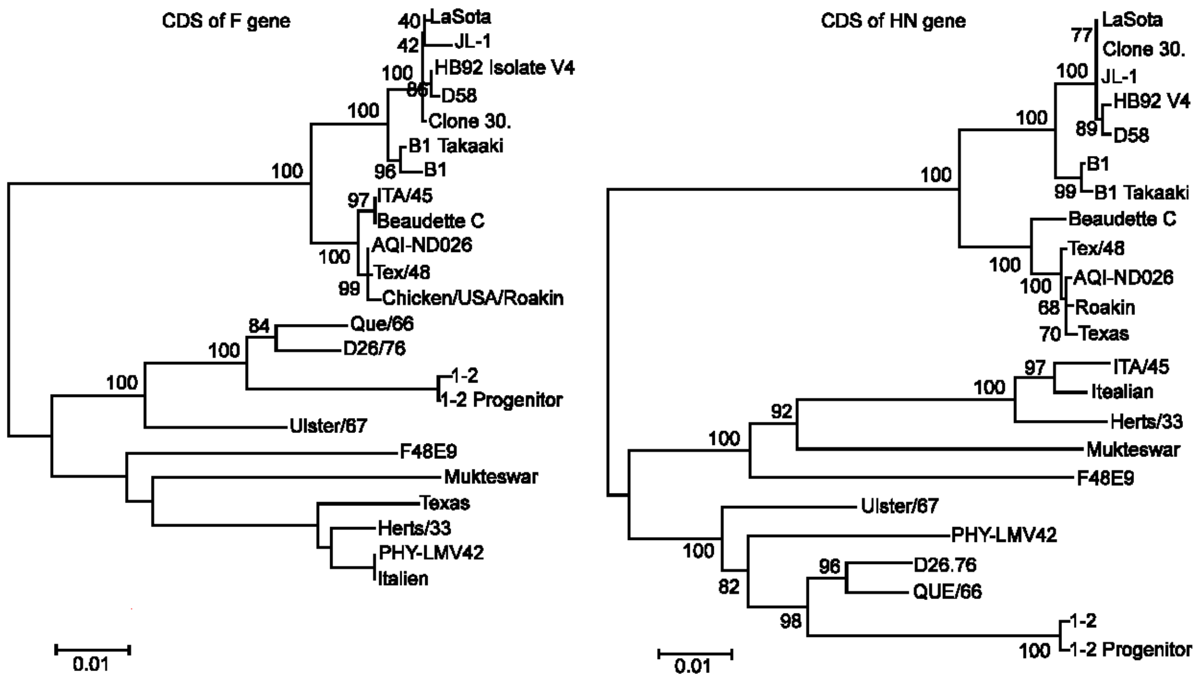

Fig. 1: Phylogenetic tree with distance and bootstrap value

Table 2. The B cell epitopes predicted for HN gene are provided in Table 3. The phylogenic tree constructed by neighbour joining algorithm with bootstrap values and distance is provided as Fig. 1.

F gene sequences: The D58 isolate was found to have cds comprising of 1659 nucleotides coding for 553 amino acids just like any other NDV strains/isolates. The isolate was found to have the - amino acids "^{112}G-R-Q-G-R-L^{117}" at the fusion protein cleavage site, amino acids "^{117}LIG AIIGGVALGVATAAQITAAAALI^{142}" at fusion peptide, 3 heptad repeat regions-HRa, Hrb and HRc at positions 143-185, 268-299 and 471-500, transmembrane domains between positions 14-27, 15-25, 118-131, 120-128, 266-269, 429-432, 499-525 and 501-523 and glycosylation sites at positions 85, 191, 366, 447, 471 and 541. These details are provided in Table 4. The sequence identity matrix for F gene between D58 and other strains of NDV are provided in Table 5. The B cell epitopes predicted for F gene are provided in Table 3. The phylogenic tree constructed by neighbour joining algorithm with bootstrap values and distance is provided as Fig. 1.

DISCUSSION

DNA Sequence analysis of the HN gene: The HN gene of NDV was reported to code for the multifunctional haemagglutinin and neuraminidase proteins, which has a molecular weight of 74 kDa (Sakaguchi *et al*., 1996). The Haemagglutinin and neuraminidase proteins were reported to be responsible for agglutination of chicken erythrocytes (Scheid and Choppin, 1974) and for hydrolysing the ketosidic bond between virus and substituted neuraminic acids on host receptors respectively, which allow the fusion protein of the virus to come into contact with host cells membrane (Lamb and Kolakofsky, 1996) (Deng *et al*., 1997) thus helping virus-cell and cell-cell fusion (Scheid and Choppin, 1974). In addition to these functions the HN gene also has a role in virulence determination (Olav *et al*., 2005). The HN gene of ND viruses was reported to have different lengths of cds comprising of 571, 577, 581 or 616 amino acids. The largest of all 616 amino acid cds of HN gene has been reported in apathogenic strains like Queensland V4 and Ulster 2C in the form of HN_0 (precursor protein) from which 45 residues cleave from the Carboxy (C) terminus to form HN protein (Sakaguchi *et al*., 1989). The NDV of high virulence alone were reported have 571 amino acids where as some virulent and less virulent strains were reported to have cds comprising of 577 amino acids (Oberdorfer *et al*., 2003). The amino acid sequence of HN gene of D58 virus was compared with 23 other standard strains (both virulent and avirulent) and has been observed that the variation between less virulent strains like D58, Clone 30, HB92 V4, LaSota and B1 were upto 3% and between less virulent and virulent strains variations range from 3-10%. The stalk region of less virulent strains were found to be conserved, where as the stalk region varies significantly between less virulent and virulent strains particularly at

Table 1: Sequence analysis of HN gene

Sl No	Feature	Position	Amino acid sequence (consensus)			
			Less virulent	Virulent	Apathogenic	D58
1	Transmembrane domain					
		24-47*	IFRIAILFLTVVTLAIS VASLLYS	IFRIAILLLTVVTLATS VASLVYS	VFRIAILLLTVVTLAIS AAALAYS	IFRIAILFLTVVTLAIS VASLLYS
		25-45	FRIAILFLTVVTLAISVASLL	FRIAILLLTVVTLATSVASLV	FRIAILLLTVVTLAISAAALA	FRIAILFLTVVTLAISVASLL
		557-563*	RIVPLLV	RIVPLLV	RIVPLLV	RIVPLLV
2	Sialic acid binding site	234-239	NRKSCS	NRKSCS	NRKSCS	NRKSCS
3	Cysteine residues	123, 172, 186, 196, 238, 247, 251, 344, 455, 461, 465, 531 and 542	Tryptophan at position123	Tryptophan at position123 in Texas and Cysteine at position 123 in Herts 33	Cysteine at position123	Tryptophan at position123
4	Glycosylation sites	119, 341, 433, 481,508 and 538	At position 508 is absent	At position 508 is present in Herts 33	At position 508 is absent	At position 508 is absent

Sequences in boxes indicate variation with consensus sequence of less virulent strains . * These transmembrane domains were not considered because of low cut off value (17)

Table 2: Sequence identity matrix for CDS of HN gene

	D58	Clone 30	HB92-V4	La Sota	B1	B1 Takaaki	JL-1	QUE/66	D2676	Ulster 67	I-2 Progenitor	I-2	PHY-LMV42	F48E9	AQI-ND026	MUKTESWAR	Tex/48	HERTS /33	Italien	ITA/45	Roakin	Beaudette C	Texas
D58		99.40	99.60	99.40	98.90	98.40	99.40	93.70	94.40	94.20	92.70	92.50	94.20	90.40	97.00	89.40	97.00	91.10	90.80	90.30	96.80	96.50	96.80
Clone 30	99.80		99.40	100.00	99.40	98.90	100.00	94.20	94.90	94.80	93.20	93.00	94.80	90.90	97.50	89.90	97.50	91.60	91.30	90.80	97.40	97.00	97.40
HB92-V4	99.80	99.70		99.40	98.90	98.40	99.40	93.70	94.40	94.20	92.70	92.50	94.20	90.40	97.00	89.40	97.00	91.10	90.80	90.30	96.80	96.50	96.80
La Sota	99.80	100.00	99.70		99.40	98.90	100.00	94.20	94.90	94.80	93.20	93.00	94.80	90.90	97.50	89.90	97.50	91.60	91.30	90.80	97.40	97.00	97.40
B1	98.90	99.10	98.90	99.10		99.40	99.40	94.10	94.80	94.40	92.70	92.50	94.20	90.60	97.40	89.70	97.40	91.50	91.10	90.60	97.20	96.80	97.20
B1 Takaaki	98.70	98.90	98.70	98.90	99.80		98.90	93.90	94.60	94.20	92.50	92.30	94.10	90.40	96.80	89.60	96.80	91.30	91.00	90.40	96.70	96.70	96.70
JL-1	99.80	100.00	99.70	100.00	99.10	98.90		94.20	94.90	94.80	93.20	93.00	94.80	90.90	97.50	89.90	97.50	91.60	91.30	90.80	97.40	97.00	97.40
QUE/66	89.90	90.00	89.80	90.00	90.40	90.30	90.00		98.60	97.90	96.30	96.10	98.00	93.40	93.50	93.00	93.50	93.20	93.70	93.20	93.40	93.90	93.40
D2676	89.90	90.00	89.90	90.00	90.40	90.40	90.00	98.20		98.60	96.70	96.50	98.40	94.10	94.20	93.50	94.20	93.70	94.20	93.70	94.10	94.60	94.10
Ulster 67	90.10	90.30	90.10	90.30	90.50	90.50	90.30	95.60	95.70		96.00	95.80	97.90	93.50	93.90	93.40	93.90	93.40	94.10	93.50	93.70	94.20	93.70
I-2 Proenitor	88.30	88.40	88.40	88.40	88.70	88.60	88.40	95.30	95.10	93.50		99.40	95.80	91.80	92.00	91.50	92.00	91.50	91.80	91.30	91.80	92.30	91.80
I-2	88.30	88.40	88.40	88.40	88.70	88.60	88.40	95.30	95.10	93.50	99.70		95.60	91.60	91.80	91.30	91.80	91.30	91.60	91.10	91.60	92.20	91.60
PHY-LMV42	89.60	89.70	89.50	89.70	90.00	90.00	89.70	95.60	95.30	95.00	93.00	93.00		93.50	93.40	93.20	93.40	93.40	93.90	93.40	93.20	93.70	93.20
F48E9	87.40	87.60	87.40	87.60	87.50	87.40	87.60	89.70	89.70	90.50	87.80	87.80	89.50		89.70	92.50	89.70	93.00	93.50	93.20	89.70	90.10	89.60
AQI-ND026	96.50	96.70	96.60	96.70	96.70	96.50	96.70	90.30	90.30	91.00	88.50	88.50	89.80	87.70		89.10	99.80	91.10	90.60	90.10	99.80	98.00	99.80
Mukteswar	87.70	87.80	87.60	87.80	88.00	88.00	87.80	90.90	90.70	91.90	89.10	89.10	90.50	91.00	88.10		89.10	92.80	93.50	93.20	88.90	89.40	88.90
Tex/48	96.60	96.70	96.70	96.70	96.80	96.60	96.70	90.20	90.30	90.90	88.40	88.40	89.70	87.80	99.80	88.10		91.10	90.60	90.10	99.60	98.00	99.60
HERTS/33	87.70	87.80	87.70	87.80	88.00	87.90	87.80	90.20	90.30	90.70	88.40	88.40	89.70	90.20	88.70	92.00	88.80		97.90	97.00	90.90	91.60	90.90
Italien	87.90	88.00	87.80	88.00	88.00	87.90	88.00	90.70	90.70	91.40	88.90	88.90	89.90	90.70	88.50	92.50	88.50	97.90		98.70	90.40	91.00	90.40
ITA/45	87.60	87.70	87.60	87.70	87.70	87.60	87.70	90.40	90.50	91.20	88.60	88.60	89.70	90.30	88.20	92.10	88.20	97.10	98.80		90.10	90.40	89.90
Roakin	96.50	96.70	96.60	96.70	96.70	96.50	96.70	90.10	90.20	90.80	88.40	88.40	89.70	87.70	99.80	88.10	99.80	88.70	88.50	88.20		97.90	99.60
Beaudette C	96.60	96.70	96.70	96.70	96.80	96.70	96.70	90.20	90.30	90.90	88.50	88.50	89.70	87.70	99.00	88.20	99.00	88.80	88.50	88.20	99.00		97.90
Texas	96.50	96.60	96.50	96.60	96.70	96.50	96.60	90.10	90.10	90.80	88.30	88.30	89.60	87.70	99.80	88.00	99.70	88.60	88.40	88.10	99.80	98.90	

Amino acid differences are indicated in bold letters and nucleotide differences are indicated in regular font

Table 3: B cell epitopes of HN and F genes

HN GENE				F GENE			
Epitope	Amino acid position		Amino acid sequence	Epitope	Amino acid position		Amino acid sequence
Number	From	To		Number	From	To	
1	23	48	LIFRIAILFLTVVTLAISVASLLYSM	1	13	30	MMLIIRVALVLSCICPAN
2	555	566	EFRIVPLLVEIL	2	488	528	LDKVNVKLTSTSALITYIVLTIISLVFGILSLILACYLMYK
3	183	213	THYCYTHNVILSGCRDHSHSYQYLALGVLRT	3	315	332	ASALVPKVVTQVGSVIEE
4	452	481	SIPCQASARCPNPCVTGVYTDPYPLIFYRN	4	59	69	GSIIVKLLPNL
5	4	11	AVSQVALE	5	165	184	VHEVTDGLSQLAVAVGKMQQ
6	526	537	YTTSTCFKVVKT	6	335	354	TSYCIETDLDLYCTRIVTFP
7	384	395	EDPVLTVPPNTV	7	413	437	GEAVSLIDKQSCNVLSLGGITLRLS
8	539	547	KTYCLSIAE	8	271	293	GNPILYDSQTQLLGIQVTLPSVG
9	77	99	NQDVVDRIYKQVALESPLALLKT	9	396	410	MTTCRCVNPPGIISQ
10	370	381	QQAILSIKVSTS	10	34	47	GRPLAAAGIVVTGD
11	404	430	ILTVGTSHFLYQRGSSYFSPALLYPMT	11	117	145	LIGAIIGGVALGVATAAQITAAAALIQAK
12	236	254	KSCSVSATPLGCDMLCSKV	12	356	365	SPGIYSCLSG
13	331	338	EGKYVIYK	13	196	241	ELDCIKIAQQVGVELNLYLTELTTVFGPQITSPALNKLTIQALYNL
14	309	319	DSRVWFSVYGG	14	73	83	KEACAKAPLDA
15	296	302	VANYPGV	15	247	253	DYLLTKL
16	283	291	EKDLDVTTL	16	302	309	YLETLSVS
17	147	166	DVTSFYPSAFQEHLNFIPAP	17	387	394	KGSVIANC
18	493	508	DGVQARLNPASAVFDS	18	451	458	QDSQVIIT
19	216	224	TGRVFFSTL	19	260	267	LSSLIGSG
20	53	62	PSDLVGIPTR	20	89	97	TTLLTPLGD
21	138	145	KELIVDDA	21	533	539	QKTLLWL
22	437	445	TLHSPYTFN	22	49	55	AVNIYTS
23	108	115	TSLSYQIN	23	151	157	ILRLKES
24	263	270	TSLSYQIN				
25	125	131	APIHDPD				
26	514	520	ITRVGSS				

A-Alanine, R-Arginine, Asparagine, D-Aspartic acid, B-Asparagine/ Aspartic acid, C-ysteine
Q-Glutamine, E-Glutamic acid, Z-Glutamine/ Glutamic acid, G-Glycine, H-Histidine, I-oleucine
L-Leucine, K-Lycine, M-Methionine, F-Phenyl alanine, P-Proline, S-Serine
T-Threonine, W-Tryptophan, Y-Tyrosine, V-Valine

he transmembrane domain. In the present study we have identified transmembrane domains at positions 24-47, 25-45 and 557-563 using online bioinformatic webserver DAS-transmembrane prediction server and observed that of the three transmembrane domains predicted in this study, the cutoff value for transmembrane domain predicted between position 25-45 was 2.2. Whereas for the other two predicted transmembrane domains a loose cutoff value of 1.7. Hence, these 2 domains were not considered. Interestingly the transmembrane domain at 25-45 have also been confirmed by crystallography for the Kansas strain of NDV (Crennell *et al.*, 2000). Significant variations are observed at the transmembrane domain between less virulent, apathogenic and virluent strains. Variations at this domain are significant since this domain is involved in promoting fusion activity in collaboration with F protein, which is essential for the virus entry and establishment of infection (McGinnes *et al.*, 2002; Gravel and Morrison, 2003). Transmembrane domain has been defined as a transmembrane alpha helix of a transmembrane protein that can be folded independently from the rest of the protein. The transmembrane domain(s) of less virulent strains like D58, Clone 30, HBR92-V4, LaSota, B1 were found to be identical. However, similarities noticed between apathogenic strains like D26, Ulster 67, I2 and virulent strains at this domain need to further studied. The sialic acid binding site located at 234-239 was found to be highly conserved in all strains irrespective of their virulence except in Mukteswar and Roakin, where the amino acid Isoleucine (I) has been replaced by Valine (V) in Mukteswar at position 235 and Glutamine (Q) by Histidine (H) in Roakin at position 236. The sialic acid binding site is essential for attachment of virus to cells. Failure of virus to attach to cells leads to aborted infection. Since, the changes reported are not observed in other velogenic strains like Texas or Herts 33, it need to be further studied whether the change in these amino acids in these strains actually result in structural changes at the sialic acid binding site. The D58 isolate was found to have 12 Cysteine (C) residues at positions 172, 186, 196, 238, 247, 251, 344, 455, 461, 465, 531 and 542. The cysteine residue reported at position 123 in certain strains like Queensland V4, D26, Ulster 2C, I_2, Mukteswar, Herts 33, Italien and ITA 45 has been replaced by tryptophan (W) in D58 isolate. However, replacement of cysteine residue at position 123 by tryptophan has been reported to be not unusual (Sakaguchi *et al.*, 1989) (Tan *et al.*, 1995) and in number of other strains like Clone 30, HB92-V4, LaSota, B1 *etc.*, the Cysteine (C) residue at position 123 has been replaced by tryptophan (W). These residues were reported to be essential for formation of intramolecular disulphide bonds (Crennell *et al.*, 2000) to stabilize HN molecules and the cysteine at position 123 stabilizes the oligomeric structure of HN gene and not involved in the actual functioning of HN gene. Hence, the absence of cysteine residue at position 123 in D58 is not unusual. The D58 isolate was also found to have five glycosylation sites at positions 119, 341, 433, 481 and 538. Another glycosylation site reported at position 508, which was reported to be commonly present only in virulent strains, is absent in D58 isolate. It has been reported that the glycosylation sites at positions 119, 341, 433, 481 and 538 are conserved among ND viruses and not the one at position 508 (Crennell *et al*, 2000; Pitt *et al.*, 2000). Further, the glycosylation site at position 508 was commonly reported only in virulent strains. Hence absence of glycosylation site at 508 adds further proof to the low virulence of D58.The HN gene of D58 isolate was found to have 26 B cell epitopes with only 6 epitopes in the stalk region and 20 epitopes in the globular head region (Table 2). In the present study B cell epitopes were predicted using bioinformatics tool. In an earlier study, where monoclonal antibodies have been used, seven overlapping antigenic sites, which induce neutralizing antibodies against HN were found only in globular head and not in stalk region (Iorio *et al.*, 1989) (Sakaguchi *et al.*, 1989). Our finding to a good extent correlates with the previous study, as most of the epitopes are concentrated in the globular head region. This finding may not be considered as inconsistent with the earlier finding, despite six epitopes located at the stalk region, since the epitopes located at this region may not be immunodominant or they may not be located prominently in the three dimensional structure of HN protein. This hypothesis has to be confirmed by synthesizing the epitopes in stalk region as custom peptide and check for its reactivity with recombinant HN protein and whole virus protein. In the phylogenetic tree constructed by Neighbour Joining (NJ) algorithm with bootstrap value along with distance, the D58 isolate was located in the group comprising less virulent strains clearly establishing that D58 isolate is an isolate of low virulence (Fig. 1). Further, from the phylogenetic tree, one could appreciate the fact that grouping is distinct with groups comprising of strains according to their virulence. This indicates that variation in HN gene has relevance to the virulence of the virus. We could not come across any mismatch in this record.

DNA sequencing and analysis of F gene: The amino acid sequence at FPCS and fusion peptide of D58 isolate was also found to have the amino acid sequence similar to that of low virulent strains like Clone 30, V4, LaSota, B1 and JL1. In all the virulent strains, the amino acid phenylalanine (F) replaced Leucine (L) at position 117. Leucine was observed in less virulent strains. In the same way the amino acid Glycine (G) at position 124 has been replaced by Serine (S). Interestingly the amino acid Alanine (A) at position at 139 observed in less

Table 4: Sequence analysis of F gene

SI No	Feature	Position	Amino acid sequence (consensus)			
			Less virulent	Virulent	Apathogenic	D58
1	Transmembrane domains	14-27*	MLTIRVALVLSCIC	MLIIRIALILSCIR	MLTVRVALELSCVC	MLIIRVALVLSCIC
		15-25	LTIRVALVLSC	LIIRIALILSC	LTVRVALELSC	LIIRVALVLSC
		118-131*	IGAIIGGVALGVAT	IGAIIGSVALGVAT	IGAIIGGAALGVAT	IGAIIGGVALGVAT
		120-128*	AIIGGVALG	AIIGSVALG	AIIGGAALG	AIIGGVALG
		266-269*	SGLI	SGLI	SGLI	SGLI
		429-432*	LGGI	LDGI	LLGI	LGGI
		499-525*	SALITYIVLTIISLVFGILSLILACYL	SALITYIVLTVISLVFGVLSLVLACYL	SALITYIVLTVISLVCGILSLVLACYL	SALITYIVLTIISLVFGILSLILACYL
		501-523	LITYIVLTIISLVFGILSLILAC	LITYIVLTVISLVFGVLSLVLAC	LITYIVLTVISLVCGILSLVLAC	LITYIVLTIISLVFGILSLILAC
3	Fusion protein cleavage site (FPCS)	112-117	GRQGRL	RRQRRF	GKQGRL	GRQGRL
4	Fusion peptide	117-142	LIGAIIGGVALGVATAAQITAAAALI	FIGAIIGSVALGVATAAQITAASALI	LIGAIIGGAALGVATAAQITAASALI	LIGAIIGGVALGVATAAQITAAAALI
5	Hetpad repeats					
	HRa	143-185	QAKQNAANILRLKESIAATNEAVHEVTD GLSQLAVAVGKMQQF	QANQNAANILRLKESIAATNEAVHEVTD GLSQLAVAVGKMQQF	QANQNAANILRLKESIAATNEAVHEVTD GLSQLAVAVGKMQQF	QAKQNAANILRLKESIAATNEAVHEVTD GLSQLAVAVGKMQQF
	HRb	268-299	LITGNPILYDSQTQLLGIQVTLPSV GNLNNMR	LITGNPILYDSQTQILGIQVTLPSV GNLNNMR	LITGNPILYDSQTQLLGIQVTLPSV GNLNNMR	LITGNPILYDSQTQLLGIQVTLPSV GNLNNMR
	HRc	471-500	NNSISNALNKLEESNRKLDK VNVKLTSTSA	NNSISNALNKLEESNSKLDK VNVKLTSTSA	NNSISNALDKLEESNSKLDK VNVKLTSTSA	NNSISNALNKLEESNRKLDK VNVKLTSTSA
6	Glycosylation sites	85, 191, 366, 447, 471 and 541	Present	Present	Present	Present

Sequences in boxes indicate variation with consensus sequence of less virulent strains. * These transmembrane domains were not considered because of low cut off value (17)

Table 5: Sequence identity matrix for CDS of F gene

	D58	Clone 30	HB92-V4	LaSota	B1	B1 Takaaki	JL-1	QUE/66	D2676	Ulster 67	I-2 Progenitor	I-2	PHY-LMV42	F48E9	AQI-ND026	MUKTESWAR	Tex/48	HERTS/33	Italien	ITA/45	Roakin	Beaudette C	Texas
D58		99.20	99.60	99.40	98.90	99.20	98.50	92.20	93.30	93.10	92.50	92.40	92.20	91.60	97.40	90.20	97.20	92.20	92.20	97.40	97.20	97.40	9240
Clone 30	99.60		99.60	99.80	99.20	99.60	98.90	92.40	93.40	93.30	92.70	92.50	92.00	92.00	97.60	90.40	97.40	92.00	92.00	97.60	97.40	97.60	9220
HB92-V4	99.80	99.80		99.80	99.20	99.60	98.90	92.40	93.40	93.30	92.70	92.50	92.00	92.00	97.60	90.40	97.40	92.00	92.00	97.60	97.40	97.60	9220
LaSota	99.60	99.80	99.80		99.40	99.80	99.00	92.20	93.30	93.10	92.50	92.40	91.80	91.80	97.80	90.20	97.60	91.80	91.80	97.80	97.60	97.80	9200
B1	98.70	98.90	98.80	98.90		99.60	98.50	91.60	92.70	92.50	92.00	91.80	91.30	91.30	97.20	89.60	97.10	91.30	91.30	97.20	97.10	97.20	9150
B1 Takaaki	98.90	99.10	99.00	99.10	99.60		98.90	92.00	93.10	92.90	92.40	92.20	91.60	91.60	97.60	90.00	97.40	91.60	91.60	97.60	97.40	97.60	9180
JL-1	99.30	99.50	99.40	99.50	98.50	98.70		91.30	92.40	92.20	92.00	91.80	91.50	92.20	96.90	89.60	96.70	91.50	91.50	96.90	96.70	96.90	9160
QUE/66	89.40	89.50	89.50	89.40	89.80	89.90	89.10		98.00	96.30	97.10	96.90	93.40	93.10	92.00	92.90	91.80	93.60	93.40	92.20	91.80	92.20	9340
D2676	89.50	89.70	89.70	89.60	89.90	90.00	89.20	98.00		97.40	97.60	97.40	94.90	94.50	93.10	94.00	92.90	95.10	94.90	93.30	92.90	93.30	9470
Ulster 67	90.40	90.60	90.50	90.40	90.70	90.70	90.10	95.40	95.40		95.80	95.60	94.00	93.40	92.90	92.70	92.70	94.00	94.00	93.10	92.70	93.10	9400
I-2 Proenitor	88.30	88.50	88.40	88.30	88.60	88.70	88.10	96.00	95.50	93.50		99.40	93.30	93.40	92.70	92.70	92.50	93.40	93.30	92.90	92.50	92.90	9360
I-2	88.20	88.30	88.30	88.20	88.50	88.60	88.00	95.90	95.40	93.30	99.70		93.10	93.30	92.50	92.50	92.40	93.30	93.10	92.70	92.40	92.70	9340
PHY-LMV42	89.30	89.40	89.30	89.20	89.40	89.40	89.10	91.80	92.10	92.20	90.40	90.30		95.10	92.20	94.50	92.00	98.50	100.00	92.40	92.00	92.40	9830
F48E9	88.90	89.10	89.10	89.00	89.10	89.20	89.10	91.20	91.30	91.50	90.00	89.80	92.90		92.40	93.80	92.20	94.90	95.10	92.40	92.20	92.40	9470
AQI-ND026	97.30	97.40	97.40	97.40	97.50	97.80	97.10	90.30	90.30	91.10	89.20	89.10	89.90	89.70		90.90	99.80	92.20	92.20	99.80	99.80	99.80	9240
Mukteswar	88.50	88.70	88.60	88.60	88.70	88.80	88.30	90.90	91.30	91.10	89.80	89.70	92.90	91.80	89.10		90.70	95.10	94.50	91.10	90.70	91.10	9470
Tex/48	97.20	97.40	97.30	97.40	97.50	97.70	97.00	90.20	90.30	91.00	89.10	89.00	89.80	89.60	99.90	89.00		92.00	92.00	99.60	99.60	99.60	9220
HERTS/33	89.30	89.40	89.30	89.20	89.40	89.40	89.10	91.80	92.10	92.10	90.40	90.30	98.70	93.00	89.90	93.10	89.80		98.50	92.40	92.00	92.40	9870
Italien	89.30	89.40	89.30	89.20	89.40	89.40	89.10	91.80	92.10	92.20	90.40	90.30	100.00	92.90	89.90	92.90	89.80	98.70		92.40	92.00	92.40	9830
ITA/45	97.10	97.30	97.20	97.30	97.40	97.70	96.90	90.50	90.60	91.20	89.40	89.30	90.10	89.80	99.60	89.20	99.50	90.10	90.10		99.60	100.00	9250
Roakin	97.10	97.20	97.20	97.20	97.40	97.60	96.90	90.10	90.10	90.90	89.10	88.90	89.70	89.60	99.80	88.90	99.70	89.70	89.70	99.40		99.60	9220
Beaudette C	97.10	97.30	97.20	97.30	97.40	97.70	96.90	90.50	90.60	91.20	89.40	89.30	90.10	89.80	99.60	89.20	99.50	90.10	90.10	100.00	99.40		9250
Texas	88.70	88.80	88.70	88.60	88.90	88.90	88.50	91.20	91.50	91.80	90.60	90.40	97.70	91.90	89.40	92.10	89.40	97.60	97.70	89.70	89.30	89.70	

virulent strains has been replaced by Serine (S) in both apathogenic and virulent strains. It has been reported that fusion peptide participate in the fusion of viral envelope with host cell membrane. Hence, the presence of amino acid Serine (S) in both virulent and apathogenic strains and Alanine (A) in less virulent strain needed to be studied further. We could not come across any reference elaborating reasons for this variation and are also not proposing any hypothesis since the data available to us in this regard are very limited. The heptad repeat regions were reported to form a triple stranded coil containing three alpha helices and believed to play an important role in fusion (Buckland *et al.*, 1992; Reitter *et al.*, 1995). The HRa region was found to be conserved among less virulent strains and the amino acid sequence of HRa region of D58 isolate of NDV was similar to this. Whereas in virlent and avirulent strains of NDV irrespective of virulence, lysine (K) at position 145 found in less virulent strains have been replaced with asparagine (N). The amino acid sequence of HRb region of D58 isolate of NDV was found to be similar to the less virulent strains and apathogenic strains. Whereas in virulent strains at position 282, amino acid Lysine (L) has been replaced by Isoleucine (I). Further, in deference to the structure of Paramyxovirus F protein, the HRb also contain leucine zipper-repeating leucine residue at every 7th amino acid at HRb. The HRc region of D58 isolate also matches with that of less virulent strains. Whereas, in apathogenic strains at positions 478 and 485, amino acids asparagine (N) and arginine (R) have been replaced by aspartic acid (D) and serine (S) respectively. In virulent strains only at position 485, arginine (R) has been replaced by Serine (S). Since the HRs were reported to play a role in fusion the changes in these positions need to be studied further with advanced bioinformatics tools, to find out the degree to which they alter the fusion activity of respective NDV strains. For F gene also we have identified transmembrane domain using online bioinformatic webserver DAS-transmembrane prediction server and observed that possible transmembrane domains are located between positions 14-27, 15-25, 118-131, 120-128, 266-269, 429-432, 499-525 and 501-523 for the D58 isolate. However, only the transmembrane domains at positions 15-25 and 501-523 had cutoff value of 2.2 and all other had a loose cutoff of 1.7. Hence, the domains with loose cutoff values were not considered. In earlier reports (Yusoff and Tan, 2001) (Tan *et al.*, 1995), a single transmemebrane domain at positions 501-529 alone has been reported. We are reporting one more domain at position 15-25. Since we have used bioinformatic tools to predict transmembrane domains, these needs to be confirmed by crystallographic studies. However, we did not come across any reference pertaining to crystallographic studies with F gene of NDV. No amino acid variation was observed between D58 isolate and less virulent strains in the transmembrane domain at positions 501-523. However, a single amino acid variation at position 16 has been observed between D58 isolate and less virulent strains. At this position the amino acid found in most of less virulent strains Threonine (T) has been replaced by Isoleucine (I) in D58 isolate. Major amino acid variations were also observed at these 2 transmembrane domains between less virulent, apathogenic and virulent strains. Based on phylogenetic tree constructed by Neighbour Joining (NJ) algorithm with bootstrap value along with distance using cds of F gene (Fig. 1), the D58 isolate was placed in the group comprising of less virulent strains like LaSota, JL1, HB92-V4 and Clone 30. In the same way based on FPCS also D58 isolate was found in the group comprising of less virulent strains like LaSota, JL1, HB92-V4 and Clone 30. The F gene of D58 isolate was found to have 23 B cell epitopes (Table 3). We could not come across any other reference regarding this.

To conclude, the above data generated with HN and F genes of D58 isolate of NDV and comparing it with sequences available in GenBank prove that this isolate of NDV used as vaccine virus is a less virulent strain. Further, we also hypothesize that though virulence of NDV is multigenic, we did not come across any NDV strain/isolate belonging to the group comprising of less virulent strains as per the phylogenetic tree drawn using F gene being placed in the group comprising of virulent strains in the phylogenetic tree drawn using HN gene or vice-versa. Hence, we conclude that grouping of NDV into virulent and avirulent is still possible with FPCS sequences. The word of caution here comes from a recent report on isolation of virulent NDV with FPCS amino acid sequences typical of less virulent strains (Lei-Tao Tan *et al.*, 2008). Hence, it would be ideal to predict pathotype after analyzing the sequence of HN and F genes instead of predicting with FPCS sequences alone.

REFERENCES

Aldous, E.W. and D.J. Alexander, 2001. Detection and differentiation of Newcastle disease Virus (avian paramyxovirus type 1). Avian Pathol., 30: 117-128.

Alexander, D.J., 2003. Newcastle Disease, Other Paramyxo Viruses and Pneumovirus Infections. 11th Edn. In: Saif, Y.M., H.J. Barnes, J.R. Ghison, A.M. Fadly, Mc Douglad and D.E. Sweyne (Eds.). Diseases of poultry. Iowa State press, Amcs, IA, pp: 63-100.

Ananth, R., J. John Kirubaharan, M.L.M. Priyadarshini and A. Albert, 2008. Isolation of Newcastle disease viruses of high virulence in unvaccinated healthy village chickens in south India. Int. J. Poult. Sci., 7: 368-373.

Bates, P.A. and M.J.E. Sternberg, 1999. Model Building by Comparison at CASP3: Using Expert Knowledge and Computer Automation Proteins. Structure, Function Gen., 3: 47-54.

Bates, P.A., L.A. Kelley, R.M. MacCallum and M.J.E. Sternberg, 2001. Enhancement of Protein Modelling by Human Intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM Proteins. Structure. Function Gen., 5: 39-46.

Buckland, R., E. Malvoisin and P. Beauverger, 1992. A leucine zipper structure present in the measles virus fusion protein is not required for its tetramerization but is essential for fusion. J. Gen. Virol., 73: 1703-1707.

Contreras-Moreira, B. and P.A. Bates, 2002. Domain Fishing: A first step in protein comparative modelling. Bioinformatics, 18: 1141-1142.

Crennell, S., T. Takimoto, A. Portner and G. Taylor, 2000. Crystal structure of the multifunctional Paramyxovirus hemagglutinin-neuraminidase. Nature Struct. Biol., 7: 1068-1074.

Cserzo, M., E. Wallin, I. Simon, Von Heijne and A. Elofsson, 1997. Prediction of transmembrane alpha-helices in procariotic membrane proteins: The Dense Alignment Surface method. Prot. Eng., 10: 673-676.

Deng, R., A.M. Mirza, P.J. Mahon and R.M. Iorio, 1997. Functional chimeric HN glycoproteins derived from Newcastle disease virus and human parainfluenza virus-3. Arch. Virol., 13: 115-130.

Doyle, T.M., 1927. A hitherto unrelated disease of fowls due to a filter passing virus. J. Comp. Pathol., 40: 144.

Gravel, A.K. and T.G. Morrison, 2003. Interacting domains of HN and F genes of Newcastle disease viruses. J. Virol., 77: 11040-11049.

Huang, Z., S. Krishnamurthy, A. Panda and S.K. Samal, 2003. Newcastle disease virus V protein is associated with viral pathogenesis and functions as an alpha interferon antagonist. J. Virol., 77: 8676-8685.

Iorio, R.M., R.J. Syddall, R.L. Glickman, A.M. Riel, J.P. Sheehan and M.A. Bratt, 1989. Identification of amino acid residues important to the neuraminidase activity of the HN glycoprotein of Newcastle disease virus. Virol., 156: 12-14.

Lamb and D. Kolakofsky, 1996. Paramyxoviridae the Viruses and Their Replication. 3rd Edn. Fields Virology. In: Fields, B.N., D.N. Knife and P.M. Howley (Eds.). Philadelphia: Lippincott-raven, pp: 1177-1203.

Lei- Tao Tan, Huai-Ying Xu, You-Ling Wang, Zhuo-Ming Qin, Lei Sun and Zhi-Zhong Cui, 2008. Molecular characterization of 3 new virulent newcastle disease virus variants isolated in china. J. Clin. Microbiol., 46: 750-753.

Li Yu, Zhillang Wang, Yihal Jiang, Leo Chang and Jimmy Kwank, 2001. Characterization of newly emerging Newcastle disease virus isolates from the People's Republic of China and Taiwan. J. Clin. Microbiol., 39: 3512-3519.

McGinnes, L,W., A.K. Gravel and T.G. Morrison, 2002. Newcastle Disease Virus HN Protein alters the Conformation of the F Protein at Cell Surfaces. J. Virol., 76: 12622-12633.

Oberdorfer, A., O. Werner, J. Veits, T. Mebatsion and T.C. Mettenleiter, 2003. Contribution of the length of the F protein cleavage site to Newcastle disease virus pathogenicity. J. Gen. Virol., 84: 3121-3129.

OIE, 2004. Manual of Diagnostic tests and Vaccines for Terrestrial Animals 2004. 5th Edition. OIE Publications. Part 2; Section; 2.1; Chapter 2.1.15.

Olav, S., O.S. Deleeuw, Gueskoch, Leo Hantog, Niek Ravenshost and B.P. Peeters, 2005. Virulence of Newcastle disease virus is determined by the cleavage site of the fusion protein and by both the stem region and globular head of the haemagglutnin-neuraminidase protein. J. Gen. Virol., 86: 1759-1769.

Peeters, B.P., O.S. Deleeuw, I. Verstigen, G. Koch and A.L. Gielkens, 2001. Generation of a recombinant chimeric Newcastle disease virus vaccine that allows serological differentiation between vaccinated and infected animals. Vaccine, 19: 1616-1627.

Pitt, J.J., E. Da silva and J. Gorman, 2000. Determination of the disulfide bond arrangement of Newcastle disease virus Haemagglutinin-neuraminidase Correlation with a beta-sheet propeller structural fold predicted for Paramyxoviridae attachment proteins. J. Biol. Chem., 275: 6469-6478.

Reitter, J., T. Sergel and T.G. Morrison, 1995. Mutational analysis of the leucine zipper motif in the Newcastle disease virus fusion protein. J. Virol., 69: 5995-6004.

Sakaguchi, M., H. Nakamura, K. Sonoda, F. Hamada and K. Hirai, 1996. Protection of chickens from Newcastle diseases by vaccination with a linear plasmid DNA expressing the F protein of Newcastle disease virus. Vaccine, 14: 747-752.

Sakaguchi, T., T. Toyoda, B. Gotoh, N.M. Inocencio, K. Kuma, T. Miyata and Y. Nagai, 1989. Newcastle disease virus evolution I Multiple lineages defined by sequence variability of the hemagglutinin-neuraminidase gene. Virol., 169: 260-272.

Scheid, A. and P.W. Choppin, 1974. The haemagglutinin and neuraminidase protein of a paramyxovirus: Interaction with neuramic acid in affinity chromatography. Virol., 62: 125-133.

Tan, W.S., C.H. Lau, B.K. Ng, A.L. Ibrahim and K. Yusoff, 1995. Nucleotide sequence of the Haemagglutinin-Neuraminidase (HN) gene of a Malaysian heat resistant viscerotropic-velogenic Newcastle disease virus. DNA Sequence, 6: 47-50.

Yusoff, K. and W.S. Tan, 2001. Newcastle disease virus macromolecules and opportunities. Avian Pathol., 30: 439-455.