The Pharma Innovation Journal 2023; 12(7): 159-164

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2023; 12(7): 159-164

www.thepharmajournal.com
Received: 24-05-2023
Accepted: 30-06-2023

K Anbu Kumar
Assistant Professor, Bioinformatics Centre & Department of Veterinary Microbiology, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai, Tamil Nadu, India

P Devendran
Professor and Head, Bioinformatics Centre, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai, Tamil Nadu, India

PL Sujatha
University Deputy Librarian, Department of Library Sciences, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai, Tamil Nadu, India

Corresponding Author:
K Anbu Kumar
Assistant Professor, Bioinformatics Centre & Department of Veterinary Microbiology, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai, Tamil Nadu, India

Phylogenetic characterization of up-to-date complete F gene sequences of Newcastle disease viruses from India reveals increase in the genotypic diversity

K Anbu Kumar, P Devendran and PL Sujatha

Abstract
Newcastle disease virus (NDV) causes a high disease burden to the global poultry production as well as infection in other avian species. Understanding the molecular epidemiology of NDVs in a region is essential in developing and deploying appropriate vaccine strains to control the disease burden. This study utilized 83 publicly available full length F gene sequences of NDVs from India, including all avian species. The recently updated unified phylogenetic classification system and revised nomenclature for Newcastle disease virus was used to delineate the Indian NDVs. The analysis of the dataset revealed the presence of genotypes I, II, III, IV, VI, VII, XIII and XXI in India. Among these, genotypes VI, VII and XXI are identified for the first time based on the updated unified phylogenetic classification system. The study reaffirmed the predominance of NDV genotypes XIII and II in chicken in India, with XIII being the highest circulating genotype (51.8%) among the analysed sequences followed by genotype II at 31.3%. All the currently used vaccine strains in India group under the genotype II, except the vaccine strain Avinew which aligned to the genotype I. The laboratory strain Mukteswar, originally isolated in the 1940s, grouped to the genotype III. The genotypes VI and XXI were exclusively from pigeons. The study highlights the high diversity of circulating NDV in India among chicken and other avian species.

Keywords: Newcastle disease virus, molecular diversity, updated unified phylogenetic classification system

1. Introduction
Newcastle disease virus (NDV) is classified in the family *Paramyxoviridae*, genus *Orthoavulavirus* and Avian *Orthoavulavirus* -1. It is also commonly referred to as Avian paramyxoviruses 1 (APMV-1) [1]. NDV is a contagious viral avian disease affecting many domestic and wild bird species; it is transmissible to humans, but rarely causes mild symptoms. NDV has different strains with varying degrees of virulence to chicken and other avian species, ranging from mild to very virulent, referred to as avirulent, lentogenic, mesogenic and velogenic (neurotropic or viscerotropic). The organs of enteric system and neurological system are mainly affected by the disease. Significant genetic diversity of the NDV is observed due to its ubiquitous presence in the farmed poultry around the world and in other avian species such as pigeons, goose, fowls, pheasants, etc., and emergence of new variants is often recorded. In addition, live attenuated NDV vaccines are widely used worldwide, contributing to the genetic flux in the NDV circulating virus. Though all APMV-1 are considered as one serotype, based on the complete genome size they are classified into class I and class II, and each class is further subdivided into many genotypes. Many schemes for classification of the genotypes were proposed to understand the evolution and genetic diversity of the NDV, mainly focusing on the Hemagglutinin-Neuraminidase (HN) gene and Fusion protein (F) gene. In 2012, A "unified and objective NDV classification system" [2] was proposed which utilized the complete F (fusion) gene coding sequences and stipulated a set of objective criteria for classification of NDV. In this system the class I NDV were considered as one genotype and further sub grouped into three sub-genotypes. The class II NDV were classified into 15 genotypes (I to XV) and included many sub-genotypes. Additional three genotypes (XVI to XVIII) were classified later based on the same scheme. Subsequently in 2019, a consortium of 40 scientists from 29 laboratories around the world and of all OIE reference laboratories for NDV revised the system of classification [3] to include 20 genotypes with a new objective set of criteria for classification and also identification of new genotypes or sub-genotypes. The earlier genotype of XV was disregarded as an genotype and reclassified as a recombinant group. A dataset of curated, up-to-date, complete F gene class I and class II

NDV sequences were publicly provided for the purpose of classification of new sequences. Further, a pilot dataset, analysis methodology and rooting guidelines for phylogenetic analysis that allow rapid genotype identification of new complete F gene sequences were also provided. The study presented here utilized all the publicly available full length (complete) F gene sequences of NDV/APMV-1/Avian *Orthoavulavirus* -1 from India [4, 5, 6, 7, 8, 9] (83 sequences) to understand the comprehensive genetic diversity and molecular epidemiology of NDV in India.

2. Materials and Methods

2.1 Collection of sequences

Sequences from the GenBank, that were >99% of the full-length (1662 nucleotides) of the F gene of NDV from India were retrieved from the GenBank and included for analysis. The total number of such sequences were 83 and are provided in Table 1. The sequences were aligned with a pruned set of 37 complete F gene sequences, representing all the 20 genotypes and sub-genotypes, from the updated unified phylogenetic classification system and revised nomenclature for NDV, 2019.

Table 1: List of complete Indian F gene sequences used in the study with GenBank Accession number

S. No	Accession No	S. No	Accession No	S. No	Accession No
1	KM056349.1	31	KY828161.1	61	KM056351.1
2	KM056346.1	32	KY828156.1	62	MT409243
3	KM056348.1	33	KX345397.1	63	HQ902590.1
4	KM056345.1	34	KY828155.1	64	KT445901.1
5	KM056347.1	35	MK796809.1	65	KJ636208.1
6	KM056344.1	36	KY828157.1	66	KC987036.1
7	KP089979.1	37	MF422123.1	67	KM056358.1
8	MF422124.1	38	KX589265.1	68	EU330230.1
9	KM056352.1	39	KT734766.1	69	HM357251.1
10	KM056350.1	40	KR072665.1	70	MT409241.1
11	KX372710.1	41	KY774445.1	71	MK796810.1
12	KX372709.1	42	KJ577585.1	72	MT409242.1
13	KX372711.1	43	KF727980.1	73	MT409237.1
14	MF422127.1	44	MZ546197.1	74	KJ563940.1
15	MF422126.1	45	OR185447.1	75	KM056354.1
16	KX242342.1	46	KU885390.1	76	MT409240.1
17	MF422128.1	47	KX710209.1	77	MT409239.1
18	MF422129.1	48	KX710210.1	78	KJ563939.1
19	MF422125.1	49	FJ665433.2	79	KJ563938.1
20	KX061544.1	50	GU187941.1	80	KJ563937.1
21	KF740478.1	51	KM056353.1	81	KJ563936.1
22	KT734767.1	52	KT987209.1	82	MK796808.1
23	KT734765.1	53	KT901462.1	83	KM056357.1
24	KX372708.1	54	FJ986192.2		
25	KX372707.1	55	MH392214.1		
26	MT409244.1	56	EF201805.1		
27	MT409238.1	57	KM056356.1		
28	KY828160.1	58	JX316216.1		
29	KY828159.1	59	KJ769262.1		
30	KY828158.1	60	KM056355.1		

Table 2: List of reference sequences used for the study with GenBank Accession number

S. No	Accession No	Genotype
1	AY935490	I.1
2	HG326605	I.2
3	AF077761	II
4	JN872151	II
5	GU182327	III
6	MH996904	III
7	AY741404	IV
8	MH996900	IV
9	JN872194	V.1
10	EU518684	V.2
11	MG840654.1	VI.2
12	FJ865434	VI.1
13	KC542905	VII.1
14	GQ338309	VII.1.2
15	KU862293	VII.2
16	FJ751918	VIII
17	FJ436303	IX
18	KX857716	X
19	JX518884	XI
20	KU594618	XII.1
21	MF278927	XII.2
22	MF409241	XIII.1.1
23	JQ267579	XIII.1.2
24	GU182323	XIII.2.1
25	GU182331	XIII.2.1
26	KF113338	XIII.2.1
27	JQ039386	XIV.1
28	HF969210	XIV.2
29	JX186997	XVI
30	HF969176	XVIIa
31	HF969194	XVII b
32	JX518885	XVIII.1
33	JX518886	XVIII.2
34	KC433530	XIX
35	AF458016	XX
36	KY042136	XXI.1.1
37	KY042141	XXI.1.2

2.2. Alignment of sequences

The 83 F gene sequences from India and the 37 reference F genes were directly analysed using Multiple Alignment with Fast Fourier Transformation (MAFFT v7.221.3) available at the EMBL-EBI server (MAFFT < Multiple Sequence Alignment < EMBL-EBI).

The 120 F gene sequence were also analysed using Molecular Evolutionary Genetics Analysis version 11 (MEGA11) software. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model with 1000 bootstrap replicates as a test of phylogeny. The tree with the highest log likelihood (-24737.41) was selected (Figure, 1). The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood

value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

The output phylogenetic trees generated was saved in the newick format and visualized using ITO hosted on the EMBL-EBI server (iTOL: Interactive Tree of Life (embl.de).

2.3. Minimum spanning tree visualization

The minimum spanning tree visualization was constructed using GrapeTree SA version. Precalculated maximum likelihood phylogenetic tree in Newick format (with 120 NDV F gene sequences) were uploaded into the GrapeTree SA along with the metadata on host species, region of India and year of collection or characterization.

3. Results

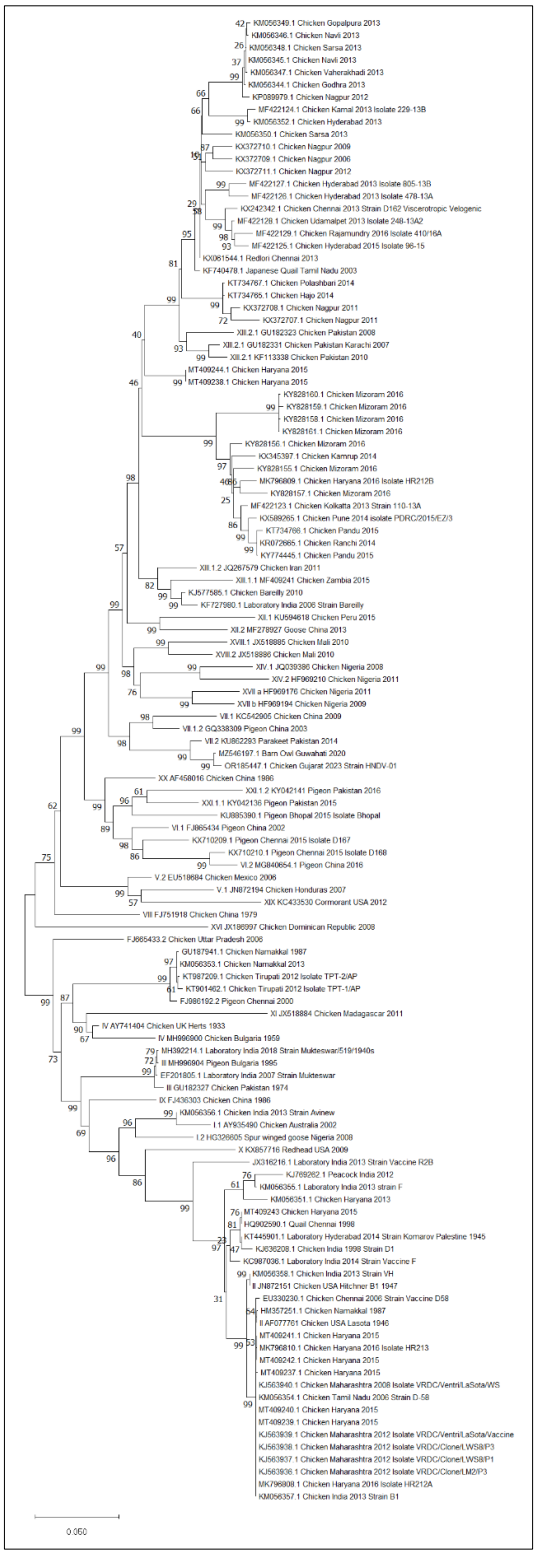

Fig 1: Maximum Likelihood Phylogenetic Tree of Indian complete NDV F gene sequences with the reference F gene sequences with the genotype I to XXI indicated before the GenBank accession numbers. The percentage of trees in which the associated taxa clustered together (1000 bootstraps) is shown next to the branches. The GenBank accession number, host species (or laboratory strain), region and year of collection are indicated in the tree

The phylogenetic analysis revealed the grouping of the full length (complete) F gene sequences with genotypes I, II, III, IV, VI, VII, XIII and XXI from the updated unified phylogenetic classification system and revised nomenclature for NDV, 2019. 43 out of the 83 sequences (51.8%) grouped with the genotype XIII. Of these two sequences were classified as sub-genotype XIII.1 and the remaining 41 sequences were classified as sub-genotype XIII.2. In addition, the XIII. 2 sequences were also grouped into two major clades with 27 and 14 sequences respectively. Twenty three sequences were delineated as genotype II (31.3%), all of which fell into one clade. Overall, the genotypes I.1, I.2, II, III, IV, IX, X, XI delineated into one branch and genotypes V, VI, VII, VIII, XII, XIII, XIV, XVI, XVII, XVIII, XIX, XX, XXI delineated into another branch. The genotypes V, VIII, IX, X, XII, XIV, XVI, XVII, XVIII, XIX and XX were not detected among the analysed sequences in this study.

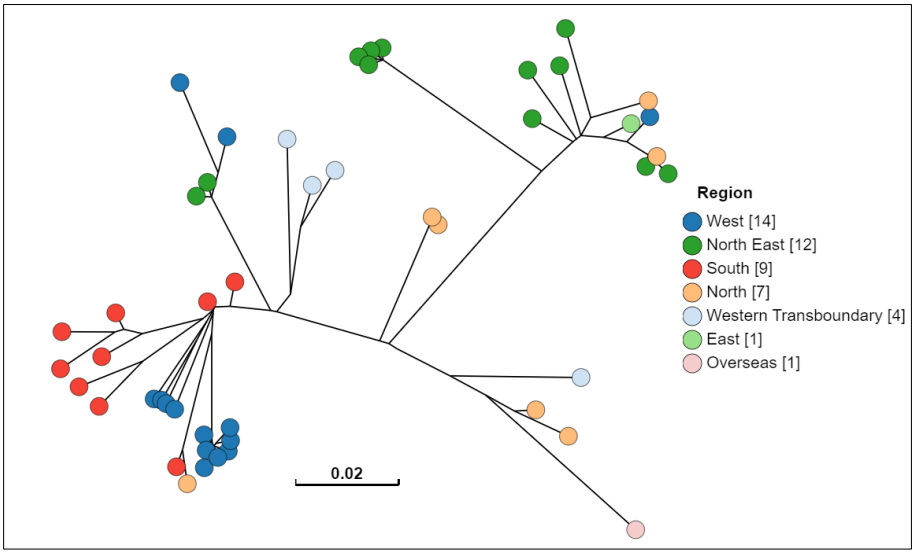

Fig 2: Minimum spanning tree of the sub-clades within the genotype XIII of the Indian complete NDV F gene sequences and reference sequences with the region of India / world where sample was collected depicted in colour. The genetic distance is indicated in the scale. The western transboundary samples were from Pakistan or Iran.

The minimum spanning tree analysis (Figure. 2) reveals spatial (geographical) clustering of the sequences from south (red), west (blue) and north east (dark green) states of India.

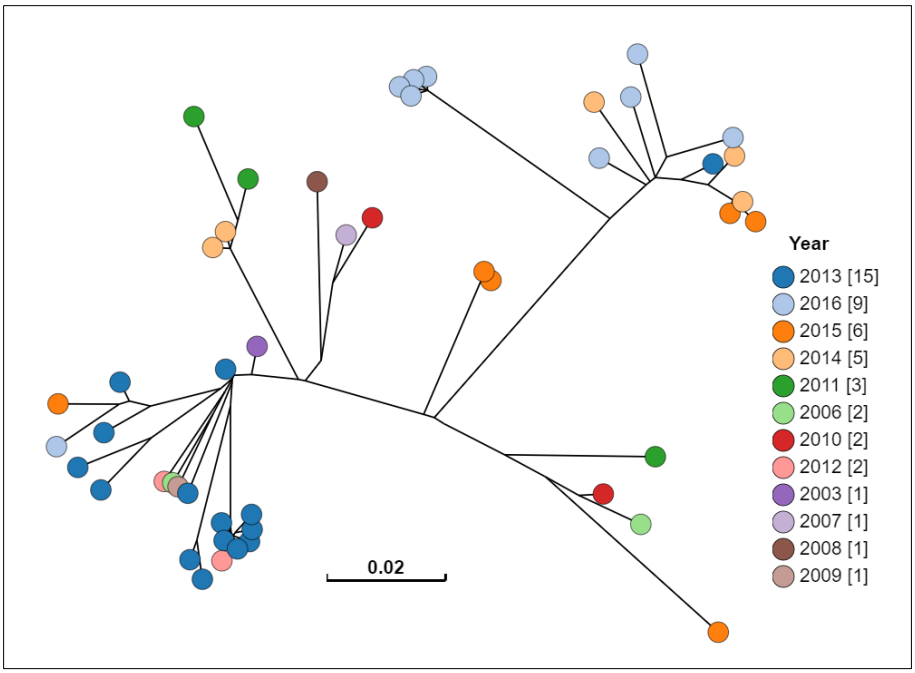

Fig 3: Minimum spanning tree of the sub-clades within the genotype XIII of the Indian complete NDV F gene sequences and reference sequences with the year of collection depicted in colour. The genetic distance is indicated in the scale

The minimum spanning tree analysis with the genotype XIII (Figure. 3) reveals temporal (year of collection) clustering with all the sequences collected in 2013 in one clade and all the sequences from 2016 in another clade consisting of two sub-clades.

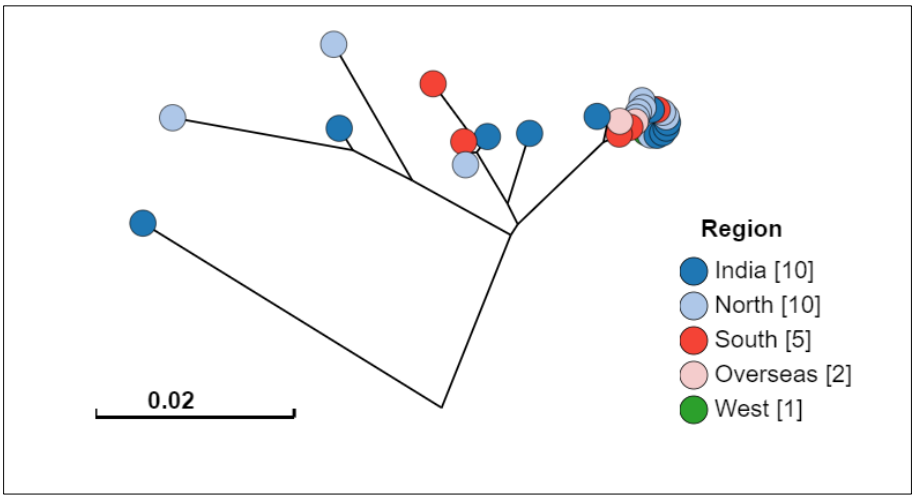

Fig 4: Minimum spanning tree of the sub-clades within the genotype II of the Indian complete NDV F gene sequences and reference sequences with the region of India / world where sample was collected depicted in colour. The genetic distance is indicated in the scale

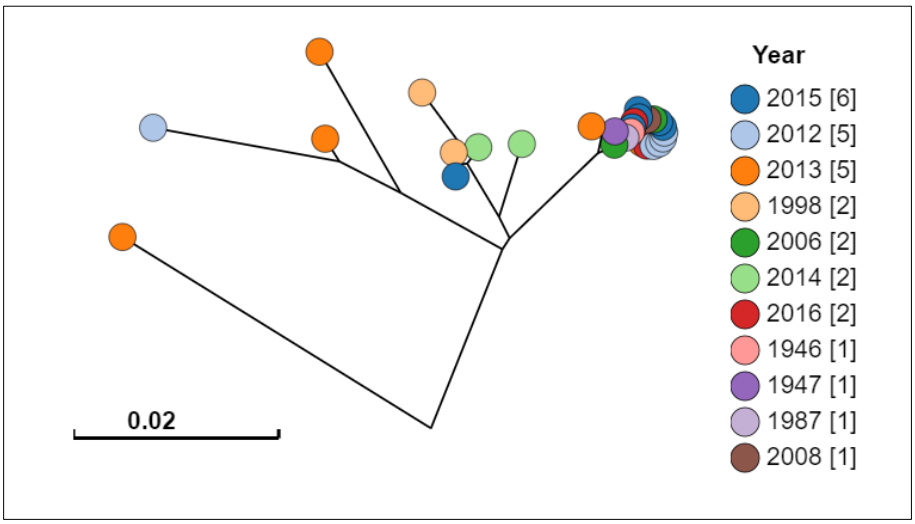

Fig 5: Minimum spanning tree of the sub-clades within the genotype II of the Indian complete NDV F gene sequences and reference sequences with the year of collection depicted in colour. The genetic distance is indicated in the scale

The minimum spanning analysis within the genotype II (Figure. 4 and Figure. 5) reveals that there is no geographic or temporal clustering within the sequences.

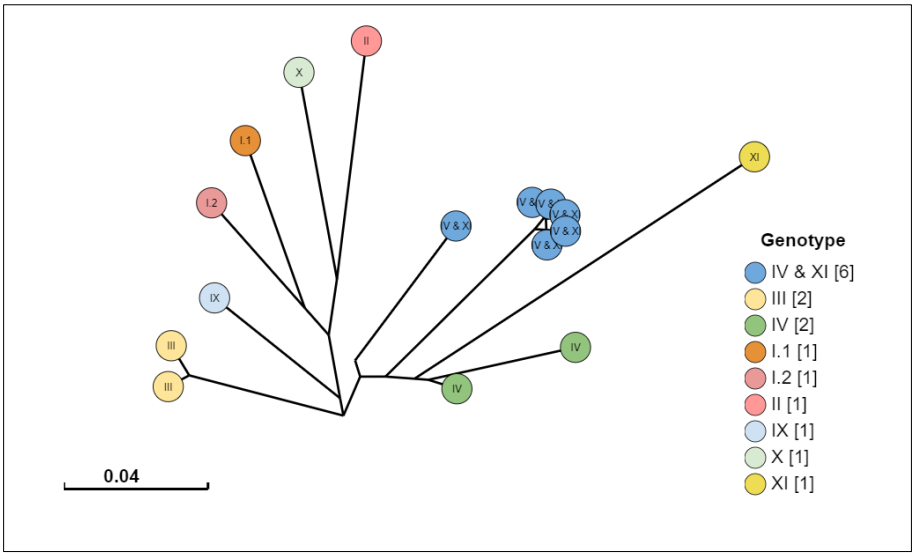

Fig 6: Minimum spanning tree of the genotype IV, XI along with genotypes I, II, IX and X of the Indian complete NDV F gene sequences and reference sequences with the genotype depicted in colour. The genetic distance is indicated in the scale

The minimum spanning tree of the sequences from genotypes IV, XI, I, II, IX and X, shows the alignment of six sequences, closely related to genotypes IV and XI but branching into a sub-clade.

4. Discussion

The phylogenetic analysis of the full length NDV F gene sequences from India has ascertained the predominance of genotype XIII, especially sub-genotype XIII. 2 (41/83) in circulation among the analysed sequences. Only two sequences aligned with the reference genotype XIII.1. The sub-genotype XIII. 2 displayed higher diversity (Figure 2 and Figure 3) and was further delineated into two distinct major clades. Graphical analysis of minimum spanning trees revealed geographical and temporal clustering of the sequences within sub-genotype XIII. 2. The genotype II was the second most frequent (29/83) in the analysed dataset. However, compared to sub-genotype XIII. 2, the diversity within genotype II was lower and also did not show any temporal or spatial clustering. All vaccine strains currently used in India, except the Avinew strain, fell in the genotype II, including Lasota, F strain, VH strain, R2B, D58, etc. The Avinew strain grouped with the genotype I. The laboratory strain Mukteswar aligned with genotype III. However, the genotype I, II and III and significantly closely related to each other than to genotype XIII. Five sequences (FJ986192, GU187941, KM056353, KT987209, KT901462) from south India, which were earlier thought to be of genotype IV, formed a distinct clade (with 99% bootstrap support) in between clades of genotype IV and XI (Figure 1 and Figure 6). Two sequences, one each from barn owl and chicken, grouped with genotype VII. Genotype VI and XXI were found only in pigeons. One sequence FJ665433 did not align with any of the genotypes in the current scheme with 99% bootstrap support.

The genotype XIII, especially XIII. 2, is the major genotype in the epidemiology of NDV infection in India. All of the sequences of XIII. 2 genotype, except one, were derived from chicken. The remaining one sequence was from Japanese quail. The earliest sequence of the genotype XIII. 2 is from 2003, whereas the genotype II sequences dating from 1940's are present in the analysed dataset. Interestingly, in spite of the widespread and decades long of use of live attenuated NDV vaccine strains of genotype II, the genotype XIII. 2 is predominant. Some of the reference strains for genotypes for XIII. 2 used in this study from Pakistan were found to group with Indian sequences from west (Figure. 2), indicating potential transboundary spread of the NDV. On the same note, sequence KU885390 collected from Bhopal in 2015 aligned with contemporary sequences of genotype XXI from Pakistan, suggesting a potential transboundary spread.

The emerging epidemiological picture from the current study warrants the further investigation of reason for the higher prevalence of genotype XIII in India, and the potential role of molecular variation in the HN gene or other genes among the circulating NDV. The potential need for shift to vaccines based on genotype XIII also arises.

5. References

1. Amarasinghe GK, Arechiga Ceballos NG, Banyard AC, Basler CF, Bavari S. Taxonomy of the order Mononegavirales: Update 2018. Arch Virol. 2018;163:2283-2294.
2. Diel DG, da Silva LH, Liu H, Wang Z, Miller PJ. Genetic diversity of avian paramyxovirus type 1: Proposal for a unified nomenclature and classification system of Newcastle disease virus genotypes. Infect Genet Evol. 2012;12(8):1770-1779.
3. Dimitrov KM, Abolnik C, Afonso CL, Albina E, Bahl J. Updated unified phylogenetic classification system and revised nomenclature for Newcastle disease virus. Infect Genet Evol. 2019;74:103917.
4. Tirumurugaan KG, Kapgate S, Vinupriya MK, Vijayarani K, Kumanan K, Elankumaran S, *et al*. Genotypic and pathotypic characterization of Newcastle disease viruses from India. PLoS One. 2011;6(12):e28414.
5. Dey S, Chellappa MM, Gaikwad S, Kataria JM, Vakharia VN. Genotype characterization of commonly used Newcastle disease virus vaccine strains of India. PLoS One. 2014 Jun 4;9(6):e98869.
6. Jakhesara SJ, Prasad VV, Pal JK, Jhala MK, Prajapati KS. Pathotypic and sequence characterization of Newcastle disease viruses from vaccinated chickens reveals circulation of genotype II, IV and XIII and in India. Transbound Emerg Dis. 2016;63(5):523-539.
7. Gowthaman V, Ganesan V, Gopala Krishna Murthy TR, Nair S. Molecular phylogenetics of Newcastle disease viruses isolated from vaccinated flocks during outbreaks in Southern India reveals circulation of a novel sub-genotype. Transbound Emerg Dis. 2019;66(1):363-372.
8. Maan S, Mor SK, Jindal N, Joshi VG, Ravishankar C, Singh VK, *et al*. Complete Genome Sequences of Newcastle Disease Virus Isolates from Backyard Chickens in Northern India. Microbiol Resour Announc. 2019;8(27):e00467-19.
9. Joshi VG, Chaudhary D, Bansal N, Singh R, Maan S. Prevalence of Newcastle Disease Virus in Commercial and Backyard Poultry in Haryana, India. Front Vet Sci. 2021;8:725232.
10. Bansal N, Singh R, Chaudhary D, Mahajan NK, Joshi VG. Prevalence of Newcastle Disease Virus in Wild and Migratory Birds in Haryana, India. Avian Dis. 2022;66(2):141-147.