RESEARCH ARTICLE Agricultural Science Digest, Volume 45 Issue 2: 337-349 (April 2025)

In silico Prediction of Molecular Docking between Chicken BRD2 and Newcastle Disease Virus Matrix Protein to Elucidate Host-pathogen Interaction

Paramvir Singh[1], C.S. Mukhopadhyay[1], Shakti Kant Dash[1]

10.18805/ag.D-5508

ABSTRACT

Background: Newcastle disease is one of the most contagious viral diseases caused by the Newcastle disease virus (NDV) which causes a huge loss in the terms of economy as well as mortality level in the poultry sector. The viral matrix protein is one of the proteins which helps in the attachment of virus, budding and assembly of progeny virions in the host cell. Host-pathogen interaction plays a very important role in the changeover of genetic diversity in organisms.

Methods: The amino acid sequences of chicken BRD2 gene and matrix protein of NDV were used for structure prediction and visualization of protein-protein interaction between BRD2 and matrix protein by using Discovery Studio software. The nucleotide sequences of both genes were used for evolutionary analysis using MEGA7 software followed by selection pressure analysis using online Datamonkey server.

Result: The results indicated that the ligand was attached with the matrix protein, which acts as a receptor on the virus surface. The purifying selection has been evident for the viral matrix protein gene. The docking of viral protein with host BRD2 protein has been detected in the study which further warrants experimental validation. In conclusion, the viral matrix protein interacts at amino acid position Tyrosine-148, Valine-339, Alanine-341 with ligand present on the chicken BRD2 protein. The evolutionary analysis revealed that the matrix protein of NDV R2B strain was the closest to that of Peacock, pigeon, duck and quail however, it exhibited maximum distance with the chicken NDV from Ireland, Vietnam and Haryana state of India.

Key words: BRD2 protein, Evolution, Matrix protein, Molecular docking, Newcastle disease virus.

INTRODUCTION

Microbes are found everywhere on this planet and they are going into our body through inhalation and ingestion and we transport thousands of microorganisms on the daily basis. Most of these microbes are not able to produce any adverse effects due to protective mechanisms adapted by our body like sneezing, coughing, urinating and defecation. If a microbe starts to cause damage to the body, it becomes a pathogen and causes the disease.

Newcastle disease is considered as one of the most contagious diseases of the poultry industry. It is caused by Newcastle disease virus (NDV) which falls under the Order *Mononegavirales*, family *Paramyxoviridae*, genus *orthoavulavirus*, species *Avian orthoavulavirus* 1. The velogenic and mesogenic strains of the Newcastle disease virus (NDV) are responsible for causing Newcastle disease in chickens. The Newcastle disease spreads through direct contact of infectious birds with normal birds, feed, water, *etc.* Up to 100% of mortality may be reported in case of disease caused by a velogenic form of NDV in unvaccinated chickens (Anonymous, 2013).

The host-pathogen interaction is defined as how pathogens interact with their host cell and assist within their host organism's body. The host-pathogen interaction can be studied at the molecular, cellular, or population level. The host-pathogenic interactions help us to understand the mechanism of microbial pathogenesis, the susceptibility of a host towards a pathogen, the biology of the host and the host cell as well as the pathogen.

A host is defined as a system that is used by the microbes to survive and reproduce. Besides, the host system can also elicit reactions with the microbes resulting in either damage, or benefit, or indifference in the host body in the states of symbiosis, colonization, commensalism, latency and disease (Casadevall and Pirofski, 2015). The term "pathogen" can be applied to those microbes which are capable of causing host damage at tissue, cellular or organ level. These pathogens disrupt the normal function of the host body and can cause serious troubles (Casadevall and

[1]Department of Bioinformatics, College of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana-141004, Punjab, India.

Corresponding Author: C.S. Mukhopadhyay, Department of Bioinformatics, College of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana-141 004, Punjab, India. Email: csmbioinfo@gmail.com

How to cite this article: Singh, P., Mukhopadhyay, C.S. and Dash, S.K. (2025). *In silico* Prediction of Molecular Docking between Chicken BRD2 and Newcastle Disease Virus Matrix Protein to Elucidate Host-pathogen Interaction. Agricultural Science Digest. 45(2): 337-349. doi: 10.18805/ag.D-5508.

Submitted: 05-10-2021 **Accepted:** 03-01-2022 **Online:** 07-03-2022

Pirofski, 2002). The pathogens consist of various types of microbes like bacteria, viruses, parasites, fungi, *etc.*

The Bromodomain-containing protein 2 (BRD2) gene is present in the nucleus of the host cell and it helps in chromatin and lysine-acetylated histone binding. Matrix protein is one of the important constituents of the enveloped viruses which plays an important role in the assembly of the virus, budding of other progeny virions and several other processes occurred in viral replication.

The protein-protein interaction becomes one of the most convenient fields of system biology. The contacts between proteins are based on different types of interactions and associations formed between them (Rao *et al.*, 2014). However, the experimental methods used for the identification of protein-protein interactions are expensive, time-consuming and challenging. There are many computational approaches had been developed for predicting protein-protein interactions. Although, sometimes these computational methods give false-positive results, however, the development of more efficient and accurate *in silico* methods is pivotal these days (Jiang *et al.*, 2018).

To date, no report is available on host-pathogen interaction using docking between the viral matrix protein of Newcastle disease virus and BRD2 protein of chicken. Moreover, there is no report on the evolutionary analysis of the BRD2 coding sequence (cds) of chicken and other avian species. Keeping in view the existing knowledge gap, the present research proposal was designed to explore the host-pathogen interaction using Bio-computational methods and evolutionary analysis of host and pathogen target genes.

MATERIALS AND METHODS

This experiment was carried out at the Animal Genomics and Functional Genomics Laboratory of the College of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana from the month of November, 2020 to April, 2021. Although no animal was used in this research work, the animal ethics approval was obtained from Institutional Animal Ethics Committee (IAEC) vide number GADVASU/2018/IAEC/47/09 dated 16.11.2018.

Sequence download

The amino acid sequences of BRD2 protein of chicken and Matrix protein-coding protein of Newcastle disease virus were downloaded from the NCBI protein website.

Prediction of 3-D structure

The tertiary structure of both the peptide sequences (BRD2 protein of chicken and Matrix protein of R2B strain of NDV) using the SWISS-MODEL by using *ab initio* method. The best structure was picked based on maximum sequence similarity.

Protein-protein interaction visualization

The PDB format files downloaded from SWISS-MODEL were uploaded on the online PatchDock server (https://bioinfo3d.cs.tau.ac.il/PatchDock/) to predict the interaction between BRD2 protein of chicken and viral Matrix protein of NDV. The result files downloaded from Patch Dock Server were visualized in Discovery Studio software for predicting the protein-protein interaction between host and pathogen genes.

Nucleotide sequence retrieval for evolutionary study

The coding sequences of the BRD2 (of chicken and other species) and viral Matrix protein of various strains of NDV were retrieved from NCBI- GenBank (https://www.ncbi.nlm.nih.gov/genbank/) in FASTA format. A total of 27 and 16 coding sequences were used for evolutionary analyses of host BRD2 and viral matrix proteins, respectively.

Multiple sequence alignment

The sequences of both genes were subjected to multiple sequence alignment using the online tool Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). The jalview format (.jvl) files were opened in Jalview software to find any conserved regions for each of the multiple sequence alignment results.

Pairwise distance

The pairwise distance was calculated in MEGA7 software (Kumar *et al.*, 2016). The reliability bootstrap value inserted for calculating pairwise distance was 500 Bootstraps.

Construction of phylogenetic tree

The best evolutionary model was selected for each of the coding sequences (cds) of chicken BRD2 and viral Matrix protein. The molecular phylogenetic analysis was done by using the Maximum Likelihood method with 500 bootstrap resampling.

Analysis of selection pressure and evolutionary fingerprinting

The text files containing the cds of various strains of the viral Matrix protein gene and the cds of BRD2 from divergent animal species were separately uploaded on Datamonkey Server (Delport *et al.*, 2010). The methods used for analysis were internal branch fixed effects likelihood (IFEL), random effects likelihood (REL), GA-branch, evolutionary fingerprinting and branch site-REL test.

RESULTS AND DISCUSSION

This study of molecular docking was based on the prediction of interaction between BRD2 protein of chicken and the viral matrix protein by using *in silico* methods. The amino acid sequences of chicken BRD2 protein (NCBI GenBank Accession No. NP001025845) and matrix protein gene of Newcastle disease virus (GenBank Accession No. AFX98108) from the NCBI website. The overall length of the host BRD2 protein and the viral matrix protein-coding gene was 729 amino acid (aa) and 364 aa, respectively. The sequence of the BRD2 protein belonged to the white leghorn chicken breed whereas the matrix protein gene amino acid sequence belonged to the NDV R2B mesogenic vaccine strain of India.

Tertiary structure prediction

The structures were obtained based on *ab initio* method *i.e.*, all the parameters were picked up automatically by the server and the best available models were chosen based on maximum percentage identity as shown in (Fig 1 and 2). The molecular models "4yug.1.A" and "4g1g.1.A" were selected as the best structure for the host-BRD2 and the viral matrix proteins. The sequence similarity of the input peptide sequences of host BRD2 and the viral matrix proteins with the best-selected models (4yug.1.A and 4g1g.1.A) was 88.64% and 92.58%, respectively (Fig 1A and 1B). A ligand named as 1 x 4-((2S,4R)-1-Acetyl-4-((4-Chlorophenyl)Amino)-2-Methyl-1,2,3,4-Tetrahydroquinolin-6-yl) Benzoic Acid was present on the BRD2 protein. Hsieh *et al.* (2016) conducted the research on molecular docking between the protein Disulfide isomerase A3 (PDIA3) protein of chicken and Fusion (F) protein of the Newcastle disease virus. They predicted the tertiary structure of NDV F protein and chicken PDIA3 protein by using the SWISS-MODEL tool. The results have shown that the catalytic domain of the chicken PDIA3 gene binds with the F protein of NDV. The outer side of the catalytic domain was covered by a unique binding pocket made up of 11 residues on the F-protein of NDV. They concluded that the unique binding pocket can be further investigated to develop the new vaccines which will decrease the harmful impacts of Newcastle disease.

Tertiary structure validation

The validation results for predicted structures of host-BRD2 and the viral matrix protein indicated that 100% and 97% of the amino acids fell within the highly preferred zone of conformations, respectively (Fig 2A and 2B; Table 1).

Tertiary structure uploading on docking server

The result files from the SWISS-MODEL tool were downloaded to upload on the online PatchDock Server to predict the sites of protein-protein interaction by visualizing them on the suitable software. The results obtained from PatchDock were based on shape complementarity criteria.

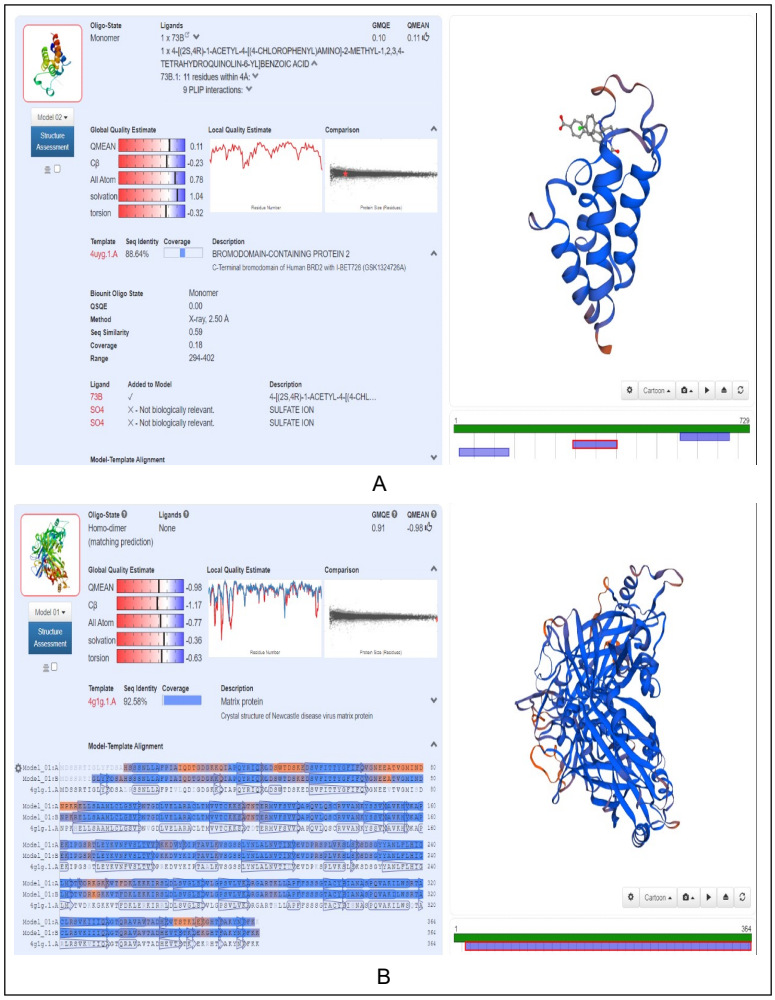

Fig 1: Complete details of the predicted 3-D structure of chicken BRD2 protein (A) and viral matrix protein (B) generated by SWISS-MODEL server.

Protein-Protein interaction visualization

The files were opened in BIOVIA Discovery Studio visualizer software (Dassault Systems). The results shown that the interaction was reported between the host BRD2 protein and the viral matrix protein. The amino acid position Tyrosine-148 of the viral matrix protein has interacted with the Benzoic acid ligand present on the chicken via Pi-Alkyl Bond. The point at which Tyrosine-148 was attached on the benzoic acid ligand, amino acid position Leucine-330 of chicken BRD2 protein was also attached with the help of an alkyl bond (Fig 3A and 3B). Another amino acid of matrix protein, Valine-339 was found to interact with ligand via an alkyl bond. On the same point, the amino acid Leucine-330 of BRD2 protein was also attached via alkyl bonds (Fig 4). There was another amino acid of matrix protein identified as Alanine-341 was interacted with the hydroxyl group of the Benzoic acid ligand via an alkyl bond (Fig 5A). An amino acid from the BRD2 protein, Tryptophan-317 was also attached to the hydroxyl group along with Alanine-341 (Fig 5B). Tryptophan-317 has interacted on the ligand ring with the help of the Pi-Alkyl bond. Apart from that, Alanine-341 was also attached at two other sites at the ligand with other amino acids. Work on the molecular docking between DNA gyrase, Topoisomerase IV and Penicillin Binding Proteins (PBP3) proteins from *Pseudomonas aeruginosa* bacteria and Ester derivative compound of Eugenol (an oil found in cloves) was carried out by Dhurga *et al.* (2016). They visualized the docking results using Bioivia Discovery Studio 4.1 visualizer software. They found that Eugenol ester derivative compound had a significant binding connection with DNA gyrase, Topoisomerase IV and PBP3 proteins. The residue positions Serine-111 and Asparagine-298 of DNA gyrase, Alanine-290, Valine-289, Leucine-286 and Alanine-278 of Topoisomerase IV and Tyrosine-124, Glutamic acid-150 and Histidine-55 of PBP3 proteins bound with Eugenol ester derivative compound. They emphasized the future investigation of Eugenol ester derivative compound to use them as a bactericidal agent.

Evolutionary analysis of host (Chicken BRD2) and pathogen (NDV Viral Matrix protein) gene

Multiple sequence alignment

The multiple sequence alignment of nucleotide sequences of chicken BRD2 indicates that the cds are mostly

Table 1: The percentage of amino acids (of predicted BRD2 and Matrix protein tertiary structures) falling under different regions of the Ramachandran's plot.

Amino acid positions in Ramachandran's plot	BRD2 protein	Matrix protein
Highly preferable conformation	100%	97.79%
Preferable conformation	0%	1.42%
Questionable conformation	0%	0.78%

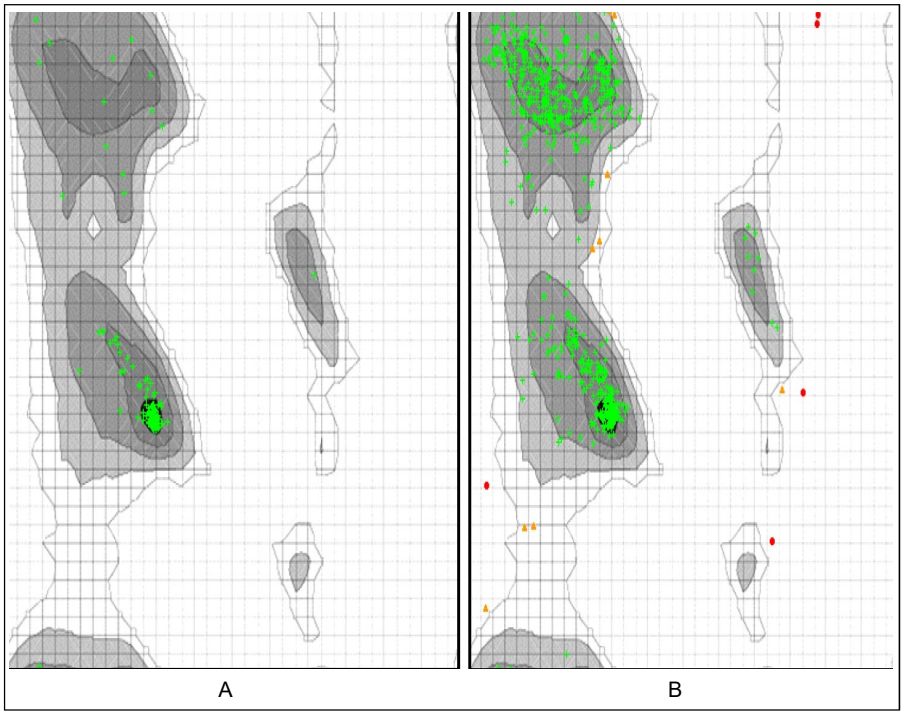

Fig 2: Validation of the predicted structure of BRD2 protein of chicken (Fig 2A) and viral matrix protein (Fig 2B) by using online Ramachandran Plot Server.

conserved among species, however, some divergent species like gecko and alligator depicts clear differences with chicken BRD2 cds. The MSA results for viral matrix protein cds showed that most of the regions of matrix protein are conserved among all the sequences. There were some mismatches seen at some sites in the sequences.

The multiple sequence alignment of 60 nucleotide sequences of the Drosha gene carried out by Kaur *et al.* (2015) revealed that the sequences belonging to different types of organisms have a conserved region of domains between them. Some domains are conserved in different species with minor differences between some species.

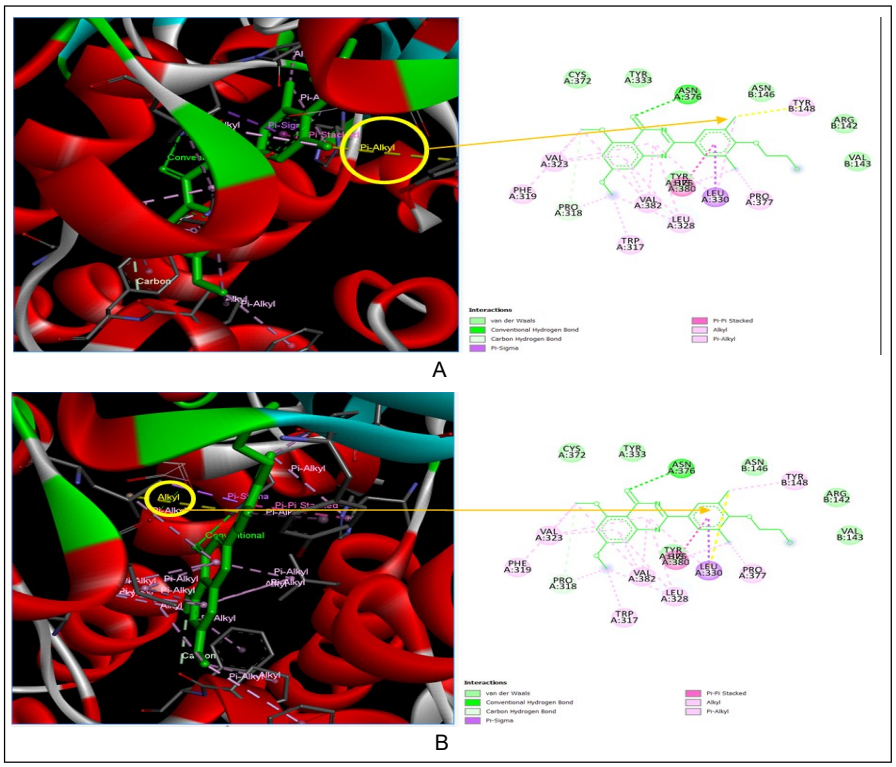

Fig 3: Bonding of Tyrosine-148 of matrix protein at benzoic acid ligand present on chicken BRD2 protein via Pi-Alkyl bond (Fig 3A) and Leucine-330 of BRD2 protein at the benzoic acid ligand *via* Alkyl bond (Fig 3B).

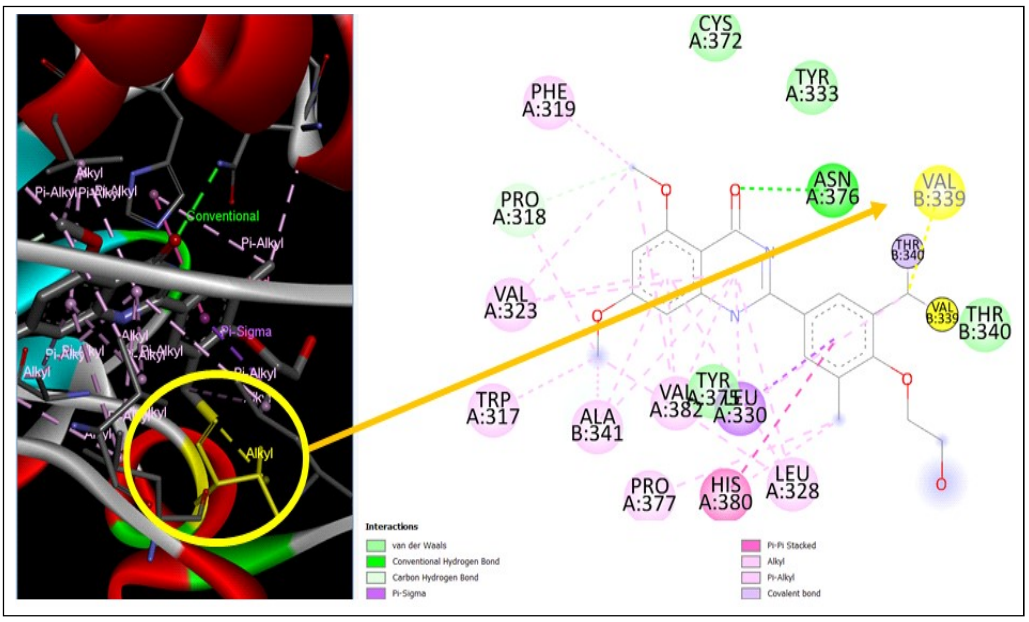

Fig 4: Bonding of Valine-339 of NDV matrix protein on the benzoic acid ligand via Alkyl bond.

Determination of best evolutionary model

The selection of the best model was done based on the lowest BIC (Bayesian information criterion) scores. In each model, AICc (Akaike information criterion, corrected) values, maximum likelihood value (InL) and other parameters were also mentioned. All the position gaps and missing data were removed. The GTR+G+I (General Time Reversible, with Gamma distribution, invariable) model using a discrete Gamma distribution for BRD2 nucleotide sequences was considered as the best evolutionary model and the GTR+G model was considered as the best evolutionary model for the viral matrix protein cds.

Calculation of pair-wise distance

The result showed that the sequences related to chicken and other avian species have the least pairwise distance, whereas the sequences of outgroup species such as -*Gekko japonicus* (Schlegel's Japanese gecko lizard) (more than 0.20), *Alligator sinensis* (Chinese alligator) (more than 0.17) *etc.* have maximum distance from the chicken BRD2 cds (Fig 6A). The pairwise distance calculation of matrix protein of NDV R2B strain resembles close relation with many of the Indian strains and those NDV which was isolated from other avian species like peacock, quail and goose were closely related to matrix protein of chicken-NDV. Although, the R2B strain is closely related with those NDV strains that were isolated from Egypt, USA, China and Ukraine (less than 0.01), it was also evident the maximum distance from NDV strains that were isolated from other geographical locations like Ireland and Vietnam (0.056 to 0.114) (Fig 6B).

Determining the evolutionary selection of sequences

Selection pressure analysis of the chicken BRD2 indicated that the test statistic values of dN-dS were less than 0.05 among the galline BRD2 cds (P>0.05). This indicates that the null hypothesis of "strict" neutral selection (dN=dS) is not rejected. However, for the rest of the combination the null hypothesis is rejected (P<0.05) and the alternative hypothesis of positive selection (dN>dS) has been accepted (Fig 7A). In the case of viral matrix protein, the test of the null hypothesis of "strict neutrality" of selection (dN=dS; P>0.05)

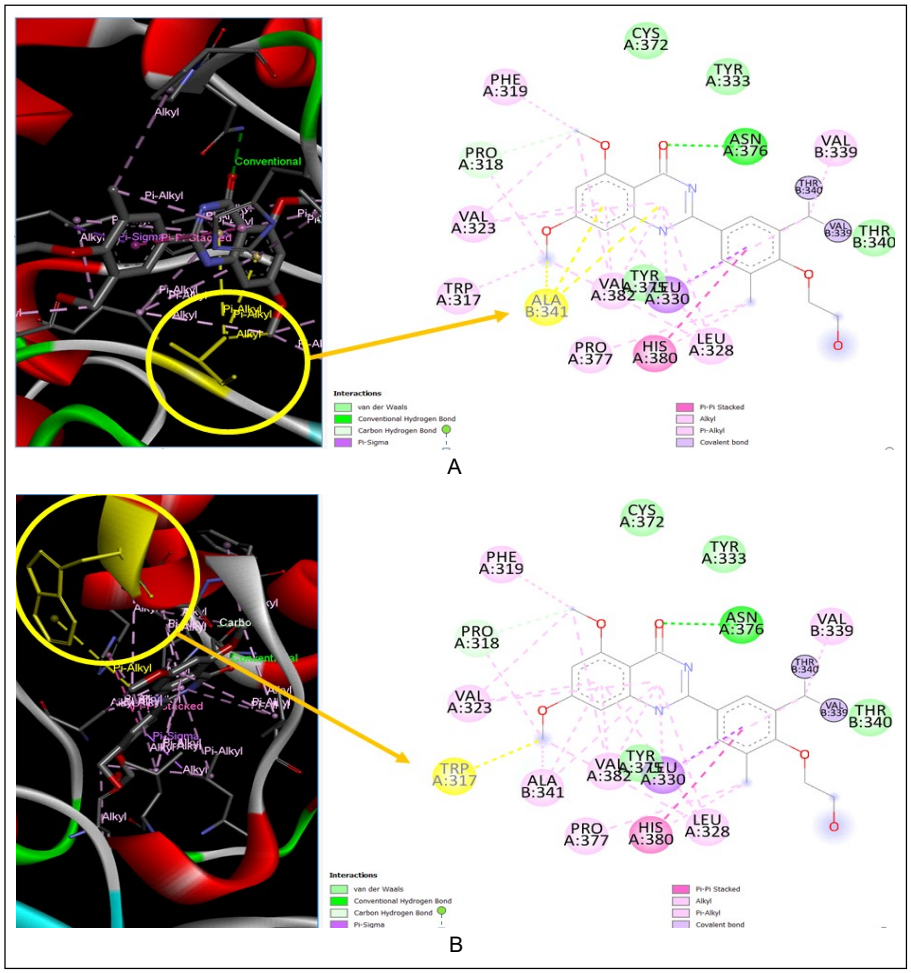

Fig 5: Alkyl bonding of Alanine-341 of matrix protein (Fig 5A) and Pi-Alkyl bonding of Tryptophan-317 (Fig 5B) of BRD2 protein of chicken at hydroxyl group on the benzoic acid ring. Other binding sites of Alanine-341 are also seen.

was rejected for almost all the sequence pairs, except for NDV-R2B vs NDV|Chicken|Egypt and NDV|Quail| Chennai|India; similarly, NDV|Quail| Chennai|India vs ND and NDV|Peacock and NDV|Chicken|Egypt; NDV|B1| Takaaki vs certain strains of chicken and duck of China, *etc.* As the number of synonymous substitutions (dS) is more than non-synonymous substitutions (dN), it depicts that the matrix coding sequences experienced purifying or negative selection (Fig 7B). Jiménez *et al.* (2009) worked on molecular evolutionary analysis of Dormancy associated MADS-box (DAM) genes from peach. MADS-box gene sequences of *Arabidopsis* were retrieved from the *arabidopsis* Information Resource (TAIR), whereas Peach genes were cloned by the research team in their laboratory and Poplar MADS-box genes were downloaded from *Populus trichocarpa* genome dataset v1.1. By using MEGA4 software, they found that purifying selection was observed in MADS-box genes of all three species. It was evident that no positive selection was observed in the MADS-box gene of *Arabidopsis*, Peach and Poplar. The overall research concluded that purifying selection has a strong effect on the rates of molecular evolution.

Construction of phylogenetic tree

The phylogenetic analysis showed that the chicken BRD2 protein resembled 99% similarity with transcript variants of Red Jungle fowl BRD2 protein whereas it was distant from outgroup species like Chinese alligator and Japenese lizard *etc.* (Fig 8A). It was observed that the NDV R2B strain closely relates with NDV isolated from India and Egypt,

NM_001030674.1|Chicken|BRD2|mRNA		0.002	0.002	0.002	0.002	0.002	0.006	0.006	0.006	0.007	0.008	0.009	0.010	0.009	0.010	0.011	0.010	0.010	0.011	0.009	0.012	0.011	0.011	0.014	0.015	0.013	0.014
XM_015294962.2|Chicken|BRD2|TV_X3|mRNA	0.008		0.000	0.000	0.000	0.000	0.006	0.006	0.006	0.007	0.008	0.009	0.010	0.009	0.010	0.011	0.010	0.010	0.011	0.009	0.012	0.011	0.011	0.014	0.014	0.013	0.015
XM_015294960.2|Chicken|BRD2|TV_X1|mRNA	0.008	0.000		0.000	0.000	0.000	0.006	0.006	0.006	0.007	0.008	0.009	0.010	0.009	0.010	0.011	0.010	0.010	0.011	0.009	0.012	0.011	0.011	0.014	0.014	0.013	0.015
XM_015294964.2|Chicken|BRD2|TV_X5|mRNA	0.008	0.000	0.000		0.000	0.000	0.006	0.006	0.006	0.007	0.008	0.009	0.010	0.009	0.010	0.011	0.010	0.010	0.011	0.009	0.012	0.011	0.011	0.014	0.014	0.013	0.015
XM_015294963.2|Chicken|BRD2|TV_X4|mRNA	0.008	0.000	0.000	0.000		0.000	0.006	0.006	0.006	0.007	0.008	0.009	0.010	0.009	0.010	0.011	0.010	0.010	0.011	0.009	0.012	0.011	0.011	0.014	0.014	0.013	0.015
XM_015294961.2|Chicken|BRD2|TV_X2|mRNA	0.008	0.000	0.000	0.000	0.000		0.006	0.006	0.006	0.007	0.008	0.009	0.010	0.009	0.010	0.011	0.010	0.010	0.011	0.009	0.012	0.011	0.011	0.014	0.014	0.013	0.015
XM_031605525.1|Phasianus_colchicus|BRD2|TV_X1|mRNA	0.046	0.041	0.041	0.041	0.041	0.041		0.004	0.000	0.006	0.008	0.009	0.009	0.009	0.010	0.010	0.010	0.010	0.011	0.010	0.011	0.011	0.011	0.013	0.014	0.013	0.014
XM_010728459.3|Meleagris_gallopavo|BRD2|mRNA	0.046	0.043	0.043	0.043	0.043	0.043	0.029		0.004	0.006	0.008	0.009	0.009	0.009	0.009	0.011	0.010	0.009	0.010	0.010	0.011	0.011	0.011	0.012	0.013	0.014	0.014
XM_031605527.1|Phasianus_colchicus|BRD2|TV_X3|mRNA	0.046	0.041	0.041	0.041	0.041	0.041	0.000	0.029		0.006	0.008	0.009	0.009	0.009	0.010	0.010	0.010	0.010	0.011	0.010	0.011	0.011	0.011	0.013	0.014	0.013	0.014
XM_021413460.1|Numida_meleagris|BRD2|TV_X1|mRNA	0.062	0.060	0.060	0.060	0.060	0.060	0.054	0.049	0.054		0.009	0.008	0.007	0.008	0.008	0.009	0.009	0.008	0.010	0.008	0.010	0.010	0.010	0.012	0.013	0.013	0.013
XM_015878533.2|Coturnix_japonica|BRD2|TV_X3|mRNA	0.078	0.076	0.076	0.076	0.076	0.076	0.075	0.078	0.075	0.087		0.012	0.012	0.011	0.012	0.012	0.012	0.012	0.012	0.011	0.013	0.012	0.013	0.015	0.015	0.014	0.016
XM_035572401.1|Cygnus_atratus|BRD2|TV_X1|mRNA	0.102	0.101	0.101	0.101	0.101	0.101	0.096	0.092	0.096	0.070	0.133		0.005	0.005	0.006	0.009	0.007	0.006	0.009	0.008	0.011	0.009	0.009	0.011	0.012	0.011	0.013
XM_035312525.1|Oxyura_jamaicensis|BRD2|TV_X1|mRNA	0.109	0.109	0.109	0.109	0.109	0.109	0.103	0.094	0.103	0.072	0.137	0.037		0.006	0.007	0.009	0.007	0.006	0.010	0.008	0.010	0.009	0.009	0.012	0.012	0.012	0.012
XM_038187752.1|Anas_platyrhynchos|BRD2|mRNA	0.107	0.107	0.107	0.107	0.107	0.107	0.101	0.095	0.101	0.077	0.134	0.043	0.054		0.008	0.010	0.008	0.008	0.010	0.008	0.010	0.009	0.009	0.012	0.013	0.013	0.013
XM_026057225.1|Apteryx_rowi|BRD2|TV_X1|mRNA	0.115	0.113	0.113	0.113	0.113	0.113	0.110	0.101	0.110	0.087	0.143	0.053	0.066	0.080		0.009	0.003	0.004	0.010	0.009	0.011	0.009	0.011	0.011	0.012	0.011	0.014
XM_030006680.1|Aquila_chrysaetos_chrysaetos|BRD2|TV_X1|mRNA	0.125	0.124	0.124	0.124	0.124	0.124	0.123	0.124	0.123	0.106	0.143	0.092	0.102	0.104	0.098		0.010	0.009	0.003	0.007	0.012	0.008	0.010	0.014	0.014	0.015	0.015
XM_013952513.1|Apteryx_australis_mantelli|BRD2|mRNA	0.118	0.117	0.117	0.117	0.117	0.117	0.114	0.106	0.114	0.092	0.145	0.061	0.072	0.087	0.016	0.104		0.005	0.010	0.009	0.011	0.009	0.011	0.012	0.012	0.012	0.014
XM_026119855.1|Dromaius_novaehollandiae|BRD2|mRNA	0.114	0.111	0.111	0.111	0.111	0.111	0.109	0.101	0.109	0.080	0.140	0.046	0.054	0.072	0.026	0.094	0.032		0.010	0.009	0.011	0.008	0.009	0.011	0.011	0.012	0.013
XM_010569907.1|Haliaeetus_leucocephalus|BRD2|TV_X1|mRNA	0.130	0.131	0.131	0.131	0.131	0.131	0.125	0.126	0.125	0.118	0.150	0.103	0.115	0.112	0.108	0.022	0.114	0.104		0.008	0.012	0.008	0.011	0.015	0.015	0.014	0.016
XM_009891112.1|Charadrius_vociferus|BRD2|mRNA	0.118	0.119	0.119	0.119	0.119	0.119	0.118	0.117	0.118	0.099	0.146	0.088	0.101	0.099	0.097	0.081	0.104	0.094	0.090		0.011	0.008	0.010	0.013	0.013	0.014	0.014
XM_034073550.1|Melopsittacus_undulatus|BRD2|TV_X2|mRNA	0.147	0.145	0.145	0.145	0.145	0.145	0.140	0.140	0.140	0.126	0.167	0.127	0.130	0.129	0.138	0.153	0.140	0.134	0.161	0.147		0.012	0.012	0.014	0.015	0.015	0.015
XM_027781969.1|Falco_peregrinus|BRD2|TV_X1|_mRNA	0.134	0.134	0.134	0.134	0.134	0.134	0.130	0.127	0.130	0.108	0.143	0.095	0.104	0.102	0.107	0.086	0.108	0.099	0.097	0.095	0.155		0.010	0.013	0.014	0.014	0.015
XM_032983990.1|Tyto_alba_alba|BRD2|mRNA	0.148	0.146	0.146	0.146	0.146	0.146	0.138	0.135	0.138	0.114	0.164	0.106	0.108	0.102	0.125	0.121	0.130	0.112	0.129	0.114	0.152	0.119		0.012	0.013	0.014	0.013
XM_038370879.1|Dermochelys_coriacea|BRD2|TV_X5|mRNA	0.183	0.182	0.182	0.182	0.182	0.182	0.172	0.163	0.172	0.161	0.205	0.137	0.147	0.150	0.138	0.187	0.143	0.137	0.196	0.177	0.198	0.185	0.154		0.006	0.012	0.013
XM_030547403.1|Gopherus_evgoodei|BRD2|TV_X5|mRNA	0.187	0.186	0.186	0.186	0.186	0.186	0.180	0.174	0.180	0.163	0.210	0.142	0.154	0.155	0.140	0.186	0.146	0.140	0.195	0.180	0.203	0.185	0.156	0.045		0.013	0.013
XM_025216589.1|Alligator_sinensis|BRD2|TV_X6|mRNA	0.175	0.173	0.173	0.173	0.173	0.173	0.168	0.170	0.168	0.163	0.192	0.138	0.156	0.158	0.137	0.196	0.145	0.141	0.191	0.186	0.197	0.192	0.182	0.161	0.162		0.015
XM_015422942.1|Gekko_japonicus|BRD2|TV_X4|mRNA	0.213	0.217	0.217	0.217	0.217	0.217	0.215	0.206	0.215	0.190	0.239	0.177	0.167	0.183	0.177	0.213	0.178	0.168	0.226	0.205	0.219	0.214	0.195	0.179	0.180	0.207	

A

JX316216.1|NDV|R2B|		0.001	0.001	0.001	0.002	0.004	0.004	0.004	0.004	0.004	0.005	0.006	0.018	0.009	0.014	0.004
KJ769262.1|NDV|Peacock|	0.006		0.002	0.002	0.003	0.005	0.005	0.005	0.005	0.005	0.006	0.007	0.018	0.010	0.015	0.005
FJ939313.1|NDV|Chicken|Egypt|	0.003	0.008		0.001	0.002	0.003	0.003	0.003	0.003	0.004	0.005	0.005	0.017	0.008	0.014	0.003
HQ902590.1|NDV||Quail|Chennai|India|	0.007	0.011	0.003		0.002	0.004	0.004	0.004	0.004	0.004	0.005	0.006	0.018	0.009	0.014	0.004
GU978777.1|NDV|Chicken|U.S.|	0.012	0.016	0.009	0.012		0.004	0.004	0.004	0.004	0.004	0.005	0.006	0.018	0.009	0.014	0.004
KJ607167.1|NDV|China|	0.022	0.025	0.019	0.022	0.023		0.000	0.000	0.002	0.002	0.005	0.004	0.018	0.010	0.015	0.002
AF375823.1|NDV|B1|Takaaki|	0.022	0.025	0.019	0.022	0.023	0.001		0.000	0.002	0.002	0.005	0.004	0.018	0.010	0.015	0.002
KJ607168.1|NDV|China|LFJ|	0.022	0.025	0.019	0.022	0.023	0.000	0.001		0.002	0.002	0.005	0.004	0.018	0.010	0.015	0.002
KU133362.2|NDV|Pigeon|Ukraine|	0.021	0.025	0.018	0.022	0.023	0.007	0.007	0.007		0.001	0.005	0.003	0.018	0.010	0.015	0.000
JX193082.1|NDV|duck|China|Guangxi|	0.023	0.027	0.020	0.024	0.025	0.009	0.009	0.009	0.003		0.006	0.003	0.019	0.010	0.015	0.001
MK902652.1|NDV|Toucan|China|	0.027	0.032	0.025	0.028	0.029	0.030	0.030	0.030	0.030	0.032		0.007	0.019	0.008	0.016	0.005
HM357251.1|NDV|Chicken|Tamil_Nadu|India|	0.034	0.038	0.032	0.035	0.036	0.020	0.020	0.020	0.014	0.016	0.044		0.021	0.012	0.017	0.003
EF201805.1|NDV|Mukteswar|	0.111	0.113	0.108	0.112	0.110	0.113	0.113	0.113	0.114	0.116	0.117	0.129		0.018	0.013	0.018
KM056351.1|NDV|Haryana|	0.051	0.055	0.048	0.051	0.053	0.056	0.056	0.056	0.056	0.058	0.046	0.069	0.112		0.016	0.010
AY562991.1|NDV|Chicken|Ireland|Ulster|	0.085	0.089	0.082	0.085	0.085	0.088	0.088	0.088	0.088	0.090	0.092	0.102	0.077	0.096		0.015
MN609929.1|NDV|Chicken|Vietnam|	0.022	0.025	0.019	0.022	0.023	0.007	0.007	0.007	0.001	0.003	0.030	0.014	0.114	0.056	0.088	

B

Fig 6: Pairwise Distance (PWD) indicating the distances between the taxa based on the coding sequences of chicken BRD2 gene (Fig 6A) and indication of distances between the coding sequences of NDV matrix protein (Fig 6B).

whereas it indicated distant relation with NDV isolated from other countries like Ireland, China, Vietnam and Ukraine. NDV isolated from other avian species like Peacock, toco toucan and Quail has shown close relation with R2B strain while NDV isolated from Pigeon, duck and goose has shown distant relation with NDV R2B strain (Fig 8B). There was a study conducted on the molecular phylogeny of the Bubaline Dicer enzyme by Singh *et al.* (2015). They used 115 amino acid sequences of Dicer enzymes of different species using maximum likelihood method with 500 bootstrap resampling value. The phylogenetic tree shows that the Dicer1 transcript variants of some species were fell within the same branch. They found that Bubaline Dicer1 sequence was closely related to yak and cattle Dicer1 sequence with a bootstrap value of 92.

IFEL analysis

The codon-based selection pressure analysis was done on Datamonkey Server (https://www.datamonkey.org). This analysis was performed for chicken BRD2 and viral matrix protein by using different statistical methods, namely, Internal Branch Fixed Effects Likelihood (IFEL), Random Effects Likelihood (REL) and Evolutionary Fingerprinting.

The IFEL analysis for chicken BRD2 shown 22 positive selection sites and 129 negative selection sites with less than 0.05 p-value (Fig 12), whereas the IFEL analysis of the viral matrix protein used with a p-value of less than 0.05 depicted only one positive codon site, whereas 26 negatively selected sites were found. Only 52 codons have shown synonymous substitutions (Fig 9A). Świderská *et al.* (2018) worked on the chicken TLR4 and TLR7 genes.

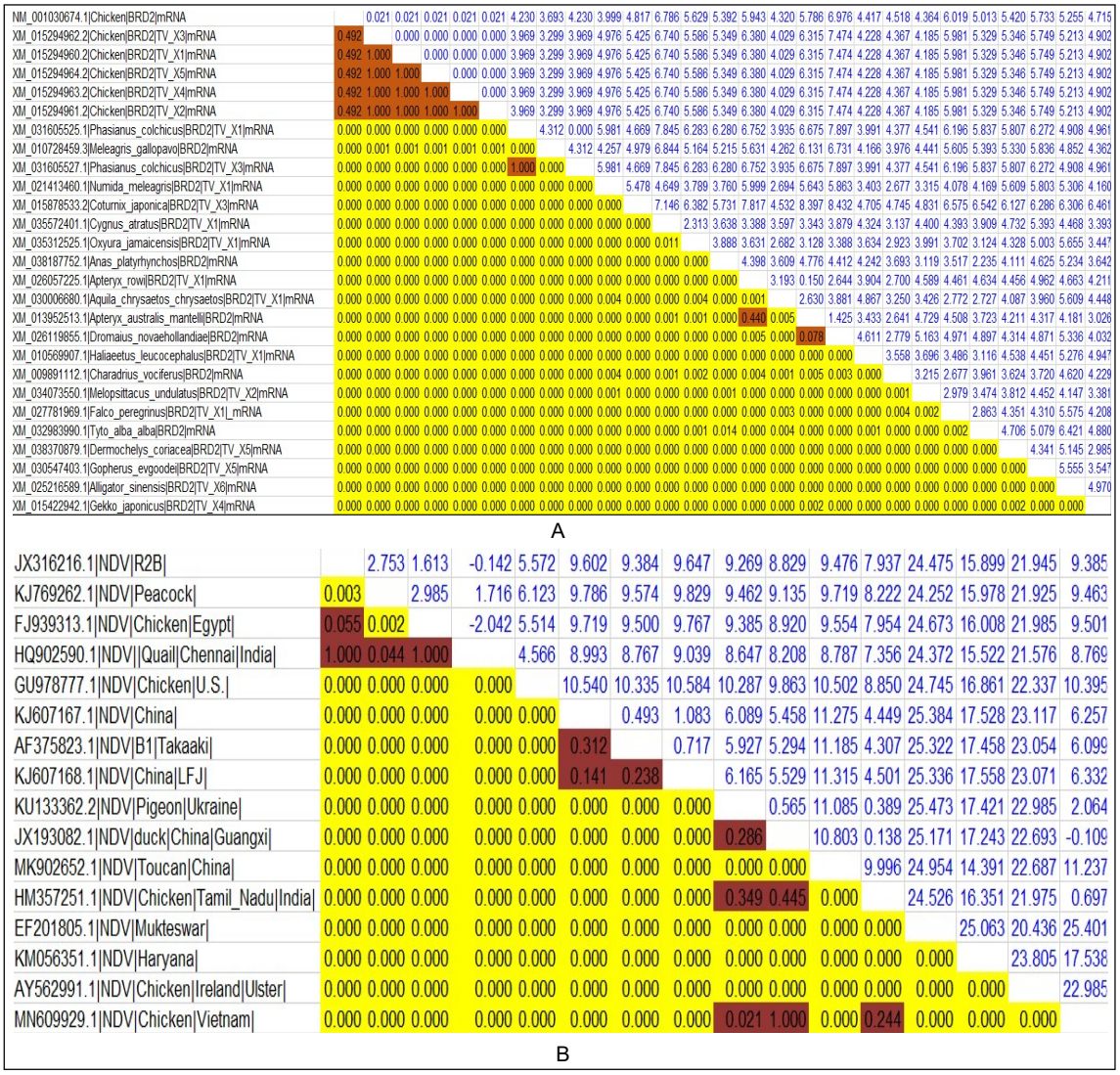

Fig 7: Matrix showing the dN-dS test statistic (above diagonal) and the corresponding probability (below diagonal) for the cds of BRD2 gene of chicken. The non-significant probabilities have been indicated by dark highlight (Fig 7A).

Matrix showing the dN-dS test statistic (above diagonal) and the corresponding probability (below diagonal) for the cds of nucleotide sequences of NDV matrix protein gene. The non-significant probabilities have been indicated by dark highlight (Fig 7B).

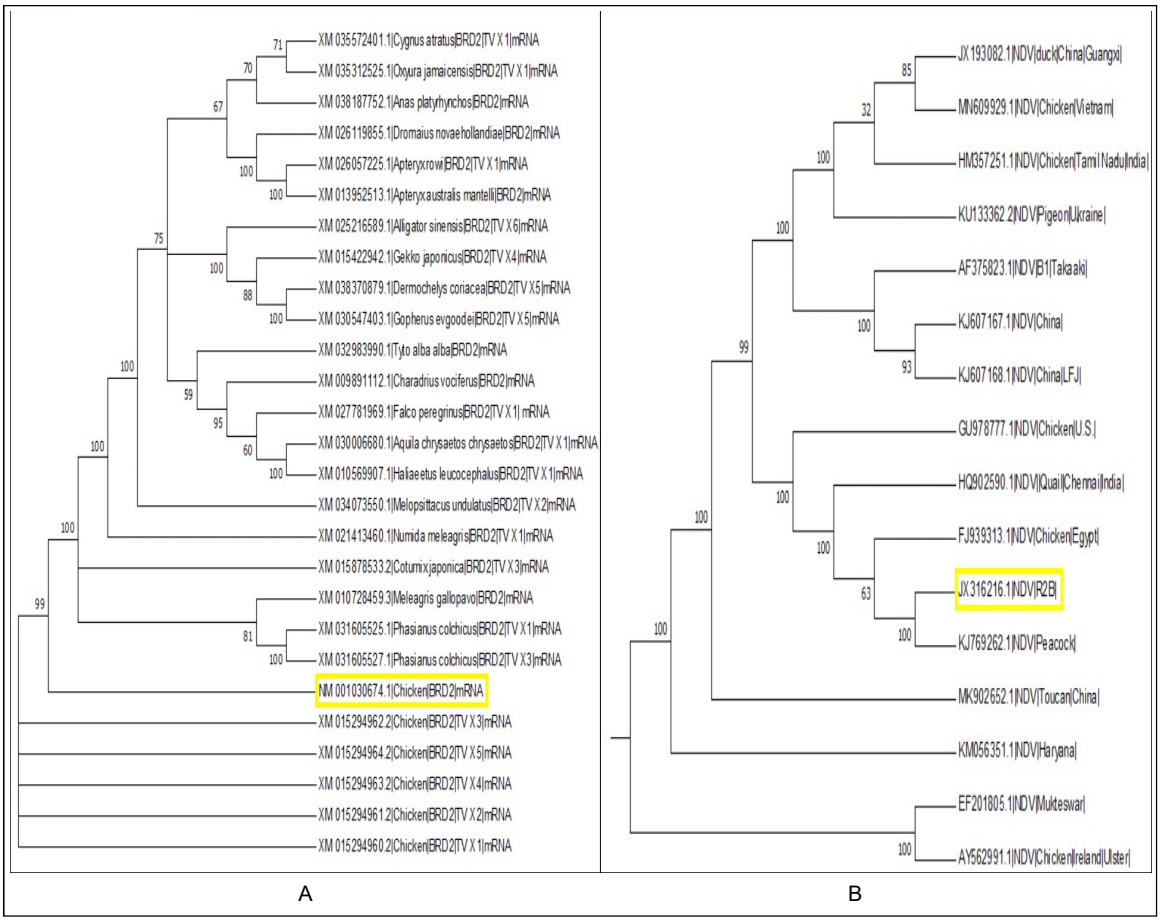

Fig 8: Phylogenetic tree of cds of nucleotide sequences of chicken BRD2 (Fig 8A) and NDV matrix protein (Fig 8B) constructed using maximum likelihood method with 500 bootstrap resampling value.

They performed selection pressure analysis using online Datamonkey server and identified five sites in TLR4 and two sites in TLR7 gene which were undergone positive selection by using IFEL at a p-value of less than 0.05.

REL analysis

REL analysis of nucleotide sequences of BRD2 shown 50 positively selected sites and 371 negatively selected sites were found, respectively (Fig 9B). The REL analysis of the viral matrix protein shown no positive selection site and all sites were negatively selected. This depicted that the virus may hinder the transfer of their deleterious genes for their survival because the REL analysis has shown all the sites as negative selection sites. Xia *et al.* (2020) selected six genes for selection pressure analysis from feline Coronavirus. The name of the genes chosen for the study was non-structural protein (nsp12-nsp14), Spike (S) protein, Nucleocapsid (N) protein and 7b genes. They performed selection pressure analysis through FEL, REL and MEME in the Bayes Factor was greater than 50 for REL method. They found that four positive sites were present in nsp12-nsp14 genes, whereas 12, 4 and 4 positive sites were present in S, N and 7b genes respectively. They identified 106, 168, 25 and 17 negatively selected sites for nsp12-14, S, N and 7b genes, respectively. The overall study concludes that most of the sequences may be conservatively maintained for the survival of the virus.

Evolutionary fingerprinting

The evolutionary fingerprinting results of chicken BRD2 shown that the best fitting model has five rate classes with an AICc value of 84194. 61 in which 37 parameters were considered by the server. The evolutionary fingerprinting analysis done for the NDV matrix protein had two rate classes with an AIC value of 32597. 83 with 28 parameters. Murray *et al.* (2013) worked on the viral suppressors of RNA interference (VSRs) which helps in inhibition of RNA interference (RNAi) to clear the invasive route for entry of the virus into the host body. They selected single-stranded RNA viruses of plants for their study. The evolutionary fingerprinting analysis shown no relationship between VSRs of plant viruses. The coat proteins and polymerases of plant viruses also not shown any type of clusters among them.

GA branch analysis

The GA-branch analysis for the BRD2 protein shown that out of 9 rate classes, the number 8 rate class had been observed with the best dN/dS score with a value of 87227.4

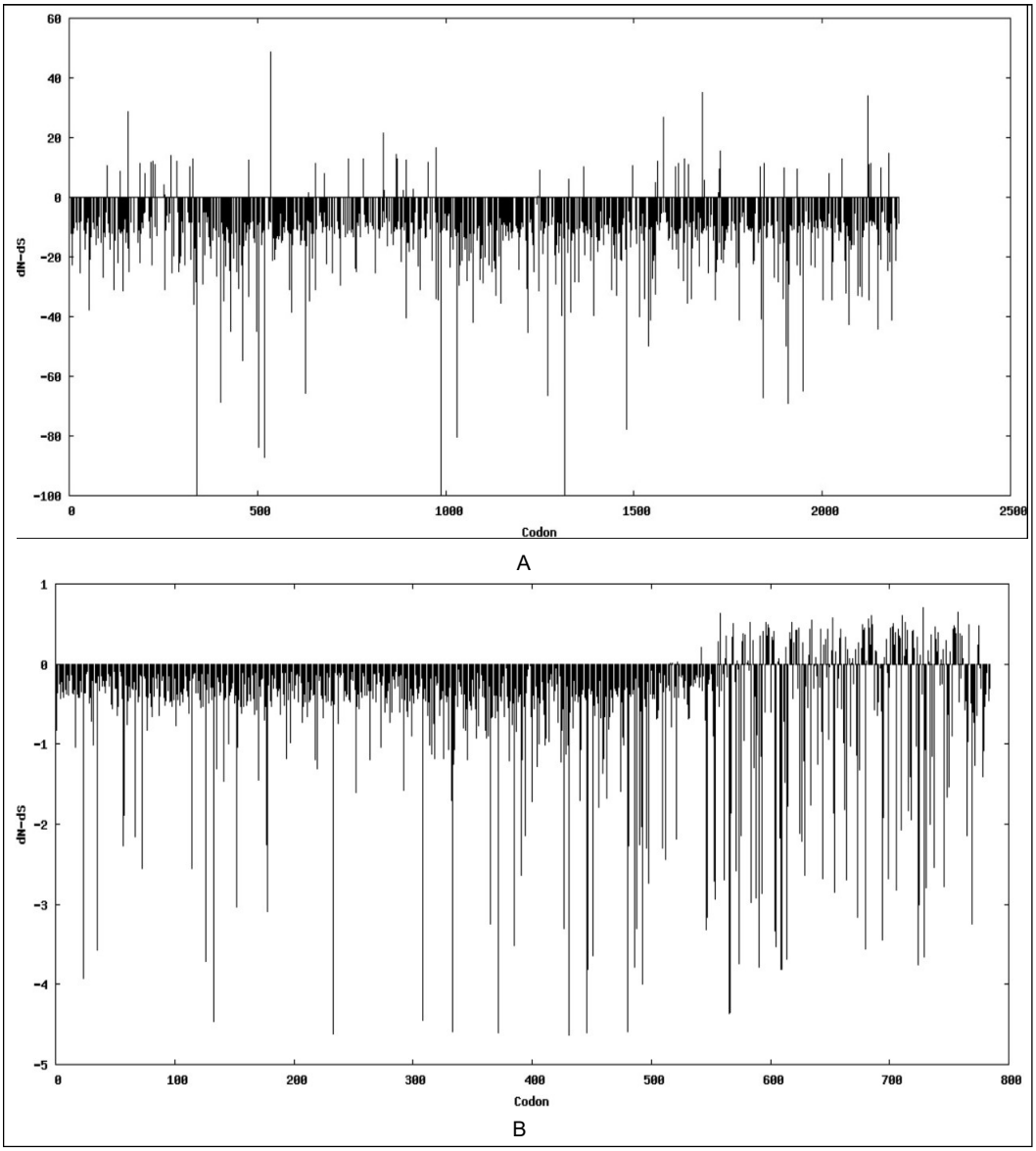

Fig 9: Graphical representation of dN-dS test statistics versus the codon positions obtained from IFEL (Internal Branch Fixed Effects Likelihood) analysis of cds of matrix protein of Newcastle disease virus using Datamonkey server (Fig 9A). Graphical representation of dN-dS test statistics versus the codon positions obtained from REL (Random Effects Likelihood) analysis of cds of chicken BRD2 using Datamonkey server (Fig 9B).

AICc points and 6 branches showed the highest probability value of dN>dS (Fig 10A). For GA-branch analysis of viral matrix protein, it was found that the best c-AIC score was observed with the value of 32508. 7 and 3 branches had shown the highest dN>dS probability value (Fig 10B). Sangula *et al.* (2010) selected eleven isolates of Foot and Mouth (FMD) SAT1 virus isolates from the Embakasi FMD laboratory in Nairobi (Kenya). 42 virus sequences were retrieved from NCBI GenBank which includes 17 sequences from East Africa and 25 from other countries of the African continent. To check the selective environments on phylogenetic branches, they applied the GA branch method to predict the values of dN/dS. They found that FMD is more prevalent in Kenya and Tanzania due to the exchange of transboundary wildlife movements. Uganda also has higher rates of FMD than other countries, but they have a lesser prevalence than in Kenya and Tanzania. These three countries share their borders, so the researchers emphasized approaching to control the transboundary animal disease.

Branch site REL analysis

In BRD2 protein, out of 26 sequences, 19 sequences come under episodic diversifying selection at a probability value of less than 0.05 (Fig 11A). In matrix protein gene, this analysis was done by using 16 nucleotide sequences and

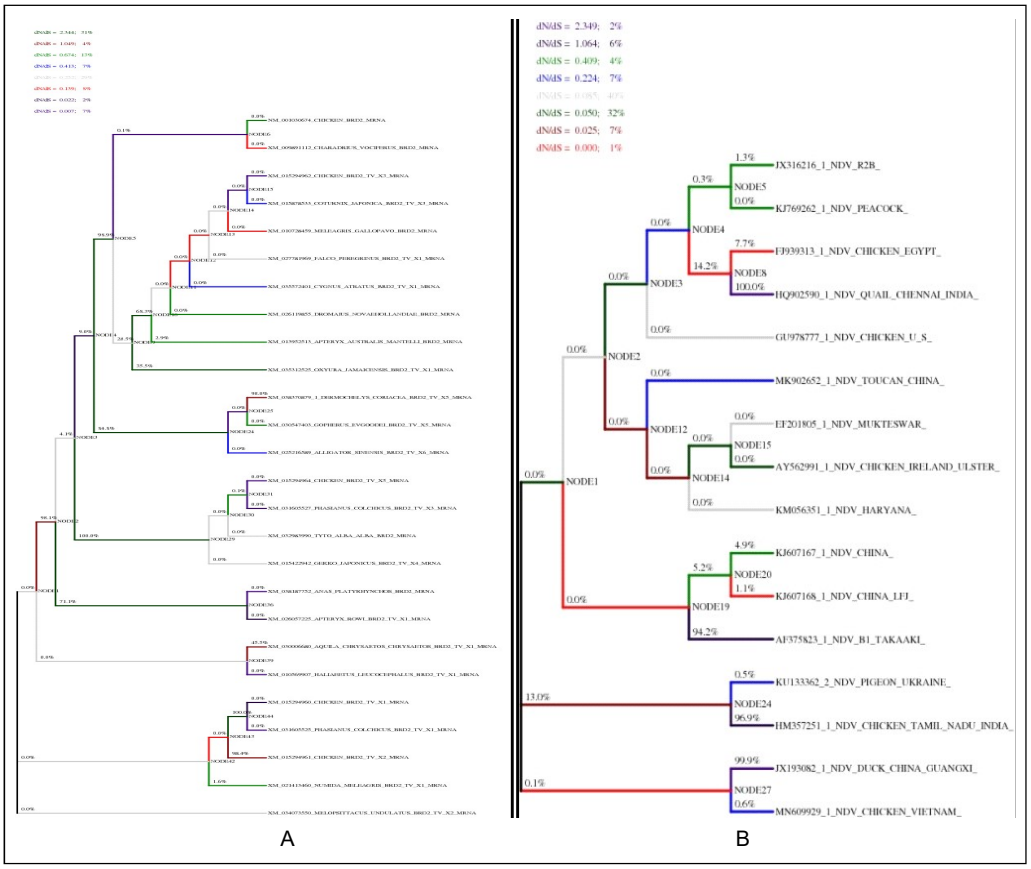

Fig 10: Result of GA-Branch analysis of chicken BRD2 gene (Fig 10A) and NDV matrix protein gene (Fig 10B) shows difference between the branch values of dN-dS in different colors.

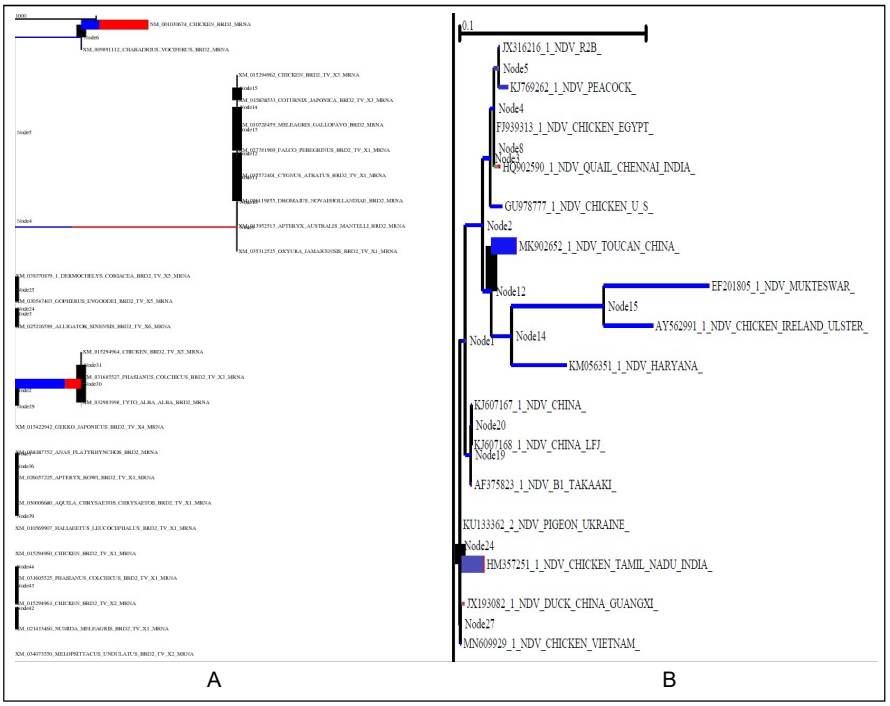

Fig 11: Result of branch site REL (Random effects likelihood) analysis of chicken BRD2 gene (Fig 11A) and NDV matrix protein gene (Fig 11B) by using Datamonkey Server. Thicker branches have undergone episodic diversifying selection.

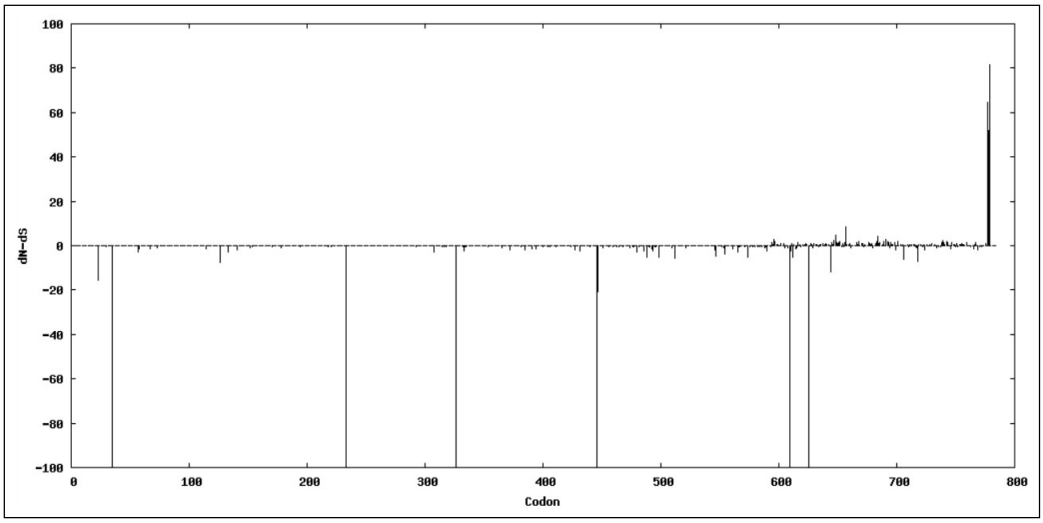

Fig 12: Graphical representation of dN-dS test statistics versus the codon positions obtained from IFEL (Internal Branch Fixed Effects Likelihood) analysis of cds of BRD2 of chicken using Datamonkey server.

the results were shown that out of 16 sequences, 2 sequences were undergone episodic diversifying selection at a p-value lesser than 0.05 (Fig 11B).

CONCLUSION

The viral matrix protein interacts at amino acid position Tyrosine-148, Valine-339, Alanine-341 with ligand present on the chicken BRD2 protein. Amino acid positions Leucine-330 and Tryptophan-317 of BRD2 protein were also found to interact with the ligand on the site where matrix protein amino acids were attached. The matrix protein of the NDV R2B strain was the closest to that of Peacock, pigeon, duck and quail while surprisingly it exhibited maximum distance with the chicken NDV from Ireland, Vietnam and the Haryana state of India. Additionally, positive selection was observed in the chicken BRD2 protein whereas negative selection was observed in the viral matrix protein. The codon-specific positive and negative selection analysis revealed that only 50 out of 784 codons were positively selected in BRD2 cds while 371 codons were negatively selected. On the contrary, all the codons were showing negative selection in the case of viral matrix protein. This indicates that negative selection is conservatively important for the survival of both Newcastle disease virus and chicken.

Conflict of interest

All authors declare that they have no conflict of interest.

REFERENCES

Anonymous, (2013). Report of OIE on “Newcastle Disease”. URL: https://www.oie.int/fileadmin/ Home/eng/AnimalHealth intheWorld/docs/pdf/Diseasecards/NEWCASTLE DISEASE.pdf (Visited on 18th January, 2020).

Casadevall, A. and Pirofski, L.A. (2002). What is a pathogen? Annals of Medicine. 34(1): 2-4.

Casadevall, A. and Pirofski, L.A. (2015). What is a host? Incorporating the microbiota into the damage-response framework. Infections and Immunity. 83(1): 2-7.

Delport, W., Poon, A.F., Frost, S.D. and Kosakovsky Pond, S.L. (2010). Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics. 26(19): 2455-2457.

Dhurga, K., Gunasekaran, G., Senthilraja, P., Manivel, G. and Stalin, A. (2016). Molecular Modeling and Docking Analysis of Pseudomonal Bacterial Proteins With Eugenol And Its Derivatives. Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences. 2(1): 40.

Hsieh, J.C., Jernigan, R.L. and Lamont, S.J. (2016). Host-Pathogen Protein-Protein Interaction Prediction Using an *in silico* Model Iowa State University Animal Industry Report. 13(1).

Jiang, M., Niu, C., Cao, J., Ni, D.A. and Chu, Z. (2018). *In silico*-prediction of protein-protein interactions network about MAPKs and PP2Cs reveals a novel docking site variants in Brachypodium distachyon. Science Reproduction. 8(1): 15083.

Jiménez, S., Lawton-Rauh, A.L., Reighard, G.L., Abbott, A.G. and Bielenberg, D.G. (2009). Phylogenetic analysis and molecular evolution of the dormancy associated MADS-box genes from peach. BMC Plant Biology. 9: 81. doi: https://doi.org/10.1186/1471-2229-9-81.

Kaur, J., Mukhopadhyay, C.S., Dhanoa, J.K., Sethi, R.S. and Arora, J.S. (2015). Expression profiling and in silico characterization of bubaline Drosha in light of evolution. International Journal of Advanced Research. 3(11): 2956-2966.

Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology Evolution. 33(7): 1870-1874.

Murray, G.G., Kosakovsky Pond, S.L. and Obbard, D.J. (2013). Suppressors of RNAi from Plant Viruses are Subject to Episodic Positive Selection. Proceedings. Biological Sciences. 280(1765): 20130965. doi: https://doi.org/10.1098/rspb.2013.0965.

Rao, V.S., Srinivas, K., Sujini, G.N. and Kumar, G.N. (2014). Protein-protein interaction detection: Methods and analysis. International Journal of Proteomics. 147648.

Sangula, A.K., Belsham, G.J., Muwanika, V.B., Heller, R., Balinda, S.N., Masembe, C. and Siegismund, H.R. (2010). Evolutionary analysis of foot-and- mouth disease virus serotype SAT 1 isolates from east Africa suggests two independent introductions from southern Africa. BMC Evolutionary Biology. 10: 371. doi: https://doi.org/10.1186/1471-2148-10-371.

Singh, J., Mukhopadhyay, C.S., Arora, J.S. and Kaur, S. (2015). Biocomputational characterization and evolutionary analysis of bubaline dicer1 enzyme. Asian-Australasian Journal of Animal Sciences. 28(6): 876-887.

Świderská, Z., Šmídová, A., Buchtová, L., Bryjová, A., Fabiánová, A., Munclinger, P. and Vinkler, M. (2018). Avian Toll-like receptor allelic diversity far exceeds human polymorphism: An insight from domestic chicken breeds. Scientific Reports. 8(1): 17878. doi: https://doi.org/10.1038/s41598-018-36226-1.

Xia, H., Li, X., Zhao, W., Jia, S., Zhang, X., Irwin, D.M. and Zhang, S. (2020). Adaptive evolution of feline coronavirus genes based on selection analysis. Bio. Med. Research International. 9089768. doi: https://doi.org/10.1155/2020/9089768.